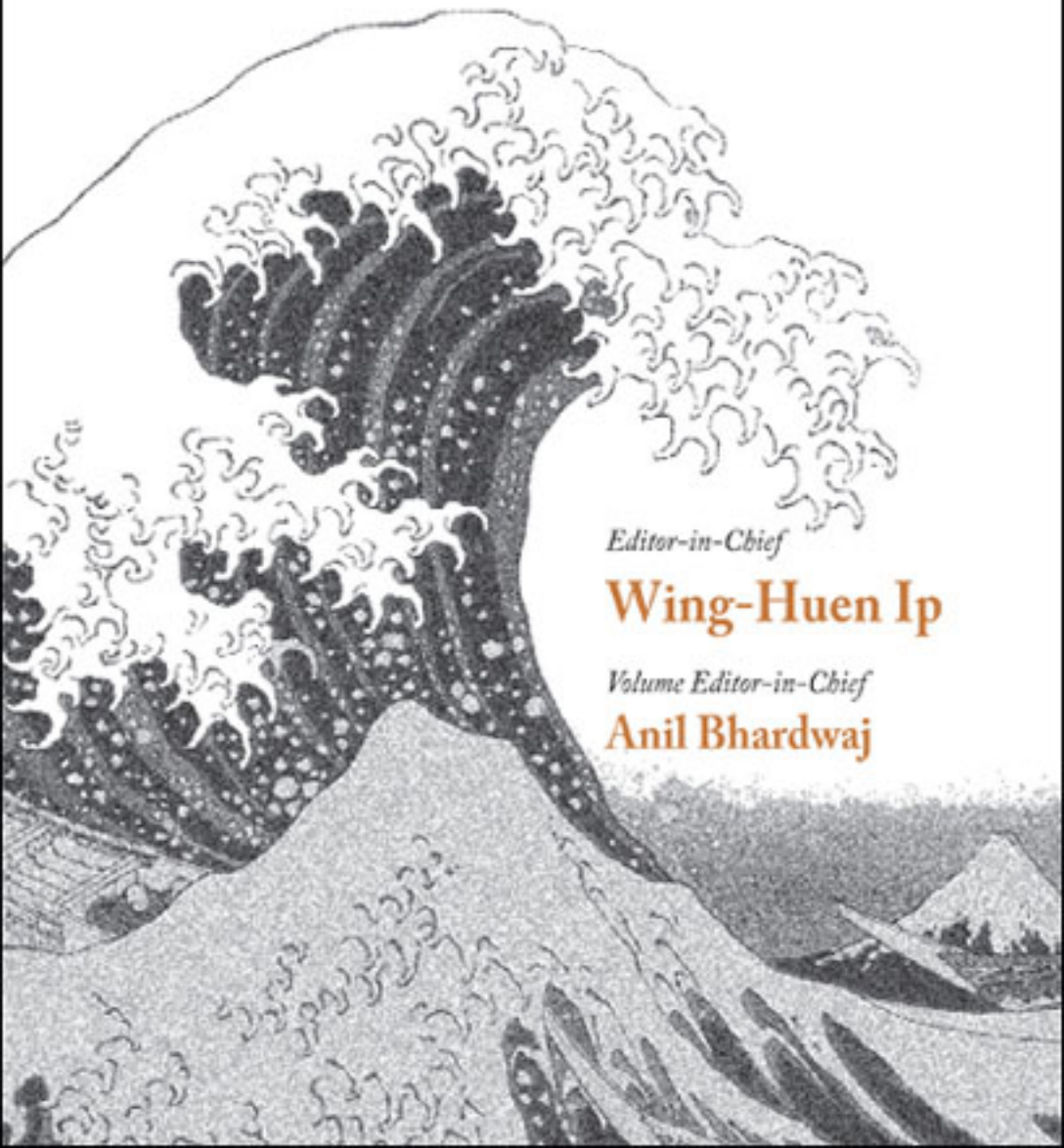
Advances in Geosciences

Volume 3: Planetary Science (PS)

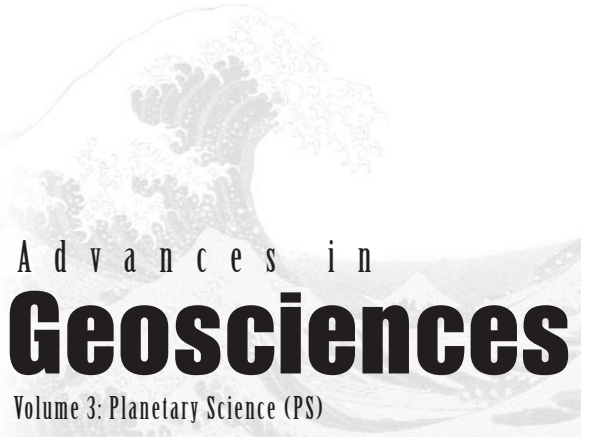


Editor-in-Chief

Wing-Huen Ip

Volume Editor-in-Chief

Anil Bhardwaj


Advances in
Geosciences
Volume 3: Planetary Science (PS)

This page intentionally left blank

Advances in

Geosciences

Volume 3: Planetary Science (PS)

Editor-in-Chief

Wing-Huen Ip

National Central University, Taiwan

Volume Editor-in-Chief

Anil Bhardwaj

Vikram Sarabhai Space Centre, India

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ADVANCES IN GEOSCIENCES
A 5-Volume Set
Volume 3: Planetary Science (PS)

ISBN 981-256-456-X (Set)
ISBN 981-256-983-8 (Vol. 3)

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore.

EDITORS

This page intentionally left blank

REVIEWERS

The Editors of Volume 3 would like to acknowledge the following referees who have helped review the papers published in this volume:

Mario Acuna
Johannes Benkhoff
R. Binzel
Michel Blanc
Hermann Boehnhardt
E. Bois
Graziella Branduardi-Raymont
Emma J. Bunce
Iver Cairns
Maria Teresa Capria
Alberto Cellino
Michael Collier
Gabriele Cremonese
Alexander Dalgarno
D. Davis
Yoshifumi Futaana
Marina Galand
Mark Gurwell
Walt Harris
Martin Hilchenbach
Takeshi Imamura
Wing -H. Ip
Vlad Izmodenov
Esa Kallio
H. Kawakita
Rosemary Killen
A. S. Kirankumar
Hirtsugu Kojima
Andreas Kopp
Norbert Krupp
Michael Kueppers
Harri Laakso
Rosine Lallement
Yves Langevin
Luisa-M. Lara
E. Pilat Lohinger
Maria J. Lopez-Gonzalez
Bjørn Lybekk
B. Marsden
Philippe Masson
Michael Mendillo
Patrick Michel
Rene Michelsen
T. E. Moore
Thomas. G. Mueller
Alessandro Mura
Masato Nakamura
Rumi Nakamura
Juergen Oberst
Stefano Orsini
Leon Phillips
Jolene Pickett
Brain Ramsey
H. Rishbeth
Yoshifumi Saito
Gerd Sonnemann
Ann Sprague
E. Tedesco
Philippe Thebault
Johan Warell
Richard Wayne
Stuart Weidenschilling
Paul Weissman
O. Witasse
Peter Wurz
Pierre Vernazza
Andrew Yau

This page intentionally left blank

Contents

This page intentionally left blank

REVIEW OF MARINER 10 OBSERVATIONS: MERCURY SURFACE IMPACT PROCESSES

CLARK R. CHAPMAN
Southwest Research Institute, 1050 Walnut Street
Suite 400, Boulder, CO 80302, USA
cchapman@boulder.swri.edu

The major evidence concerning Mercury's craters remains the imaging of Mariner 10 from over three decades ago. We are beginning to gain information about Mercury's unimaged side from Earth-based radar. The MESSENGER mission will start making major advances in a few years. Issues that have developed and remain to be resolved include the specific roles of the Late Heavy Bombardment and of hypothetical "vulcanoids" in cratering the planet, which affect calibration of the absolute chronology of Mercury's geological and geophysical evolution, and the role of secondary cratering, by ejecta from both the visible craters and from the numerous large basins.

1. Introduction

Mercury's surface was first revealed by Mariner 10 imaging. Generally, Mercury is heavily cratered, with less cratered regions, superficially similar to the Moon. Craters range from nearly saturated small-sized craters (many show clustering suggestive of secondaries) up to enormous multi-ringed basins, of which Caloris is the most prominent in Mariner 10 images. Dozens of basins have been tentatively identified. Morphometric statistics of Mercury's craters have been compared with those for other terrestrial bodies and interpreted in terms of differences in surface gravity and other factors.[1–4]

Craters are used for both relative and absolute age-dating of geological units on planetary bodies. Early interpretations of Mercury's cratering record drew analogies from the Moon: it was supposed that most craters and basins formed about 3.9 Ga, during the same Late Heavy Bombardment (LHB) that dominated lunar cratering.[5] Through superposition relationships with other geological features (e.g., lobate scarps), a tentative chronology for Mercury's geological history was derived (tied to the old Tolstoj basin and the younger Caloris basin), raising potential incompatibilities

with geophysical inferences about interior cooling rates and processes that generate Mercury's magnetic field.[2,5]

Craters on Mercury (and other bodies) are studied in several ways. Beyond the Mariner 10 imaging, and expected imaging from MESSENGER and BepiColombo missions to Mercury, radar delay-doppler mapping, some at better than 2 km resolution, from Arecibo[6] and other radar telescopes is starting to rival the resolution of the global-scale Mariner 10 imaging. Studies comparing Mercury's crater population with those on the Moon, Mars, and other bodies provides useful understanding. Also relevant is modeling of impactor populations and simulations of their impact history on Mercury. Theoretical studies of cratering mechanics, ejecta distributions, and regolith evolution have sometimes been applied to Mercury.

In detail, Mercury's craters have morphological differences from those on the Moon and Mars, partly due to differences in gravity and impactor environment (e.g., higher velocity impacts on Mercury), but most of the differences are probably due to the different geological processes that erode and degrade craters after they have formed on the various planets. For example, while many craters on Mars extend back to the Late Heavy Bombardment epoch that may be contemporaneous with the formation of many of Mercury's craters, the Martian surface has undergone glaciation, rainfall/runoff, dust storms, sedimentation, exhumation, and many other processes not thought likely to have been relevant on Mercury.

2. Origins of Mercury's Craters

Potential sources for the impactors that formed Mercury's craters are numerous. In principle, the size distributions and the impact rates could have varied with time and in ways not necessarily correlated with the cratering histories of other bodies. Sources include: the near-Earth asteroids and their cousins (of which only three have yet been found) that orbit entirely interior to Earth's orbit (termed Apoheles); short- and long-period comets, including sun-grazers; vulcanoids, an as-yet-undiscovered hypothetical population of remnant planetesimals from accretionary epochs, orbiting mainly inside Mercury's orbit; and secondary cratering by ejecta from basins and large primary craters. Endogenic crater-forming processes (e.g., volcanism) are also possible. Differences in asteroid/comet cratering rates (perhaps including the LHB) are not expected to vary by large factors for Mercury compared with the Earth–Moon system or Mars. But if vulcanoids were/are important, they could have extended the duration of Mercury's

intense cratering into epochs far later than the LHB, and even overprinted the LHB.[7]

It is possible that planetesimals interior to Mercury's orbit never formed or that they were destroyed (e.g., by their mutual, high-velocity collisions) before the epoch of the LHB. But if they did survive such early processes, it appears that their subsequent depletion by Yarkovsky effect drift might have lasted for several billion years, perhaps resulting in appreciable post-LHB cratering of Mercury.[8] (The Yarkovsky effect has been found to have a potent effect in changing the orbits of small solar system bodies; it acts on rotating bodies due to the asymmetry between insolation and re-radiation in the thermal infrared.) The point here is not to assert that current assumptions that tie Mercury's absolute geological chronology to the LHB are wrong, but that such a chronology should not be taken as a strong constraint. There are other geophysical aspects of Mercury that were surprising to the Mariner 10 researchers and originally seemed difficult to reconcile. Mercury appears to have an active dynamo-generated magnetic field, although this has been debated. Mercury's buckled crust (expressed as a global distribution of lobate scarps) suggests global cooling and shrinking in post heavy-cratering epochs, but perhaps the cooling and crustal shortening has not gone to completion. Recent bistatic radar interferometric studies[9] suggest the presence of a molten layer within Mercury. Theoretical modeling combined with hypotheses concerning impurities in the core may have reconciled Mercury's small size and rapid cooling rate with these indications of a still molten portion of the planet's core. Nevertheless, the evaluation of these geophysical issues should not be strongly constrained by any particular cratering chronology. Mercury's heavily cratered regions may, in fact, reflect LHB cratering and its internal geological processes may have shut down soon afterwards (access to molten magma may have been closed off by crustal compression). Alternatively, both the cratering and the faulting could have extended billions of years closer to the present time, if vulcanoid cratering was important.

3. Secondary Cratering

Secondary cratering may be a much more important process than previously thought. Studies of both the sparsely cratered surface of Europa[10] and of Mars[11] have recently suggested that the steep branch of the crater size-frequency relation for craters smaller than a few km (originally identified by Shoemaker[12] as the "secondary" branch but later attributed to the inherent

size-distribution of collisionally evolved asteroids) really is dominated by secondaries from primaries larger than 10 km. One study, for example, finds that a single 10 km diameter crater on Mars may be responsible for as many as a billion secondary craters larger than 10 m diameter.[11] Crater chains, crater rays, and clusters of small craters visible in Mariner 10 images were already attributed to the process of secondary cratering on Mercury. The global high-resolution imaging expected from MESSENGER, combined with the extensive intercrater plains on Mercury as well as the tendency for ejecta on Mercury to be less widely distributed around the globe than is true for the Moon, all should help explicate the role of secondary cratering as a planetary process. Possibly, many supposed primary craters several tens of km in size may instead be secondaries from basin-forming impacts, as Wilhelms[13] believes is true for the Moon.

References

1. F. Vilas, C. R. Chapman and M. S. Matthews, eds., *Mercury* (University of Arizona Press, Tucson, 1988).
2. P. D. Spudis and J. E. Guest, in *Mercury*, eds. F. Vilas, C. R. Chapman and M. S. Matthews (University of Arizona Press, Tucson, 1988), p. 118.
3. R. J. Pike, in *Mercury*, eds. F. Vilas, C. R. Chapman and M. S. Matthews (University of Arizona Press, Tucson, 1988), p. 165.
4. P. H. Schultz, in *Mercury*, eds. F. Vilas, C. R. Chapman and M. S. Matthews (University of Arizona Press, Tucson, 1988) p. 274.
5. R. G. Strom and G. Neukum, in *Mercury*, eds. F. Vilas, C. R. Chapman and M. S. Matthews (University of Arizona Press, Tucson, 1988), p. 336.
6. J. K. Harmon, P. J. Perillat and M. A. Slade, *Icarus* **149** (2001) 1.
7. M. A. Leake, C. R. Chapman, S. J. Weidenschilling, D. R. Davis and R. Greenberg, *Icarus* **71** (1987) 350.
8. D. Vokrouhlichy, P. Farinella and W. F. Bottke, *Icarus* **148** (2000) 147.
9. J. L. Margot, S. J. Peale, R. F. Jurgens, M. A. Slade and I. V. Holin, *Proccedings of the 35th COSPAR* (2004), p. 3693.
10. E. B. Bierhaus, C. R. Chapman and W. J. Merline, *Nature*, **437**, doi:10.1038/nature 04069 (October 2005) 1125–1127.
11. A. S. McEwen *et al.*, *Icarus* **176** (2005) 351.
12. E. M. Shoemaker, in *Ranger VII Pt. II. Experimenters' Analyses and Interpretations*, eds. R. L. Heacock *et al.* (JPL Tech. Rept. No. 32-700, 1965), p. 75.
13. D. E. Wilhelms, *Lunar and Planetary Sci. Conf. 7th* (1976), p. 2883.

EARTH GROUND-BASED OBSERVATIONS OF MERCURY EXOSPHERE — MAGNETOSPHERE — SURFACE RELATIONS

LEBLANC FRANÇOIS
Service d'Aéronomie du CNRS/IPSL, France
francois.leblanc@aerov.jussieu.fr

Mariner-10 flybys of Mercury were the main sources of information on Mercury's exosphere up to the discovery of the exospheric sodium emission from ground based observatories (Potter and Morgan, *Science* **229** (1985) 651–653). These later observations were followed by the discovery of potassium (Potter and Morgan, *Icarus* **67** (1986) 336–340) and calcium emissions in Mercury's exosphere (Bida *et al.*, *Nature* **404** (2000) 159–161). Several ground-based observations have underlined the significant spatial and temporal variations of Mercury's exosphere. Such observations lead to the suggestion of a large number of potential sources of ejection of volatiles from Mercury's surface, but also to the suggestions of strong relations between Mercury's exosphere and its magnetosphere as well as between Mercury's exosphere and upper surface.

1. Introduction

Since the Mariner-10 flybys of Mercury, most of the studies and observations of Mercury's neutral exosphere have concerned the bright sodium exospheric emissions.[1–3] Since its first observation,[4] several observations underlined the significant spatial and temporal variabilities of Mercury's sodium exosphere.[5,6] Such observations lead to the suggestion of a large number of potential sources of ejection of the sodium atoms from Mercury's surface. They also suggest strong relations between Mercury's exosphere and its magnetosphere[7] as well as between Mercury's exosphere and upper surface.[8–10]

This paper is focused on the different sources of variation of the known exospheric abundances and does not intend to be a complete review on Mercury's exosphere already done in several recent papers.[1–3] Most of these variations derived from the strong links between Mercury's exosphere, its magnetosphere, and its upper surface. These exospheric variations can therefore, be also interpreted as indirect observations of the characteristics of the magnetosphere and of the upper surface. This paper emphasizes the

most debated and recent questions and the key information that may be derived from ground-based observations, on Mercury's exosphere — magnetosphere — upper surface system. Such information is particularly needed in preparation of the forthcoming Messenger and BepiColombo missions.

In Sec. 2, we give a brief overview of the known neutral and ion compositions of Mercury's exosphere. Section 3 describes some important sources of variations of the exospheric abundance related to the coupling between Mercury's exosphere, its upper surface, and its magnetosphere. Section 4 concludes this paper.

2. Exospheric Composition

The only few elements that Mariner-10 provided on Mercury's neutral exosphere are:

(i) the measurements of H and He with surface density of $250\,\mathrm{H/cm^3}$ and $6 \times 10^3\,\mathrm{He/cm^3}$,[11]
(ii) an identification of O column density with suggested value for its surface density of $4.4 \times 10^4\,\mathrm{O/cm^3}$,[12,13]
(iii) an upper limit for the total content of the exosphere at the subsolar point: 10^7 neutral particles/cm^3 from the solar occultation UV measurement.[12]

Ground-based observations identified:

(i) Na atoms with zenith sub-solar column density of $10^{11}\,\mathrm{Na/cm^2}$,[4,14]
(ii) K atoms with column density of $10^9\,\mathrm{K/cm^2}$,[4,15,16]
(iii) Ca atoms with column density of $10^8\,\mathrm{Ca/cm^2}$.[17]

Only upper limits are known for other elements from Mariner-10 solar occultation UV experiment: at the surface density $1.4 \times 10^7\,\mathrm{H_2/cm^3}$, $2.5 \times 10^7\,\mathrm{O_2/cm^3}$, $2.3 \times 10^7\,\mathrm{N_2/cm^3}$, $1.5 \times 10^7\,\mathrm{H_2O/cm^3}$, $1.6 \times 10^7\,\mathrm{CO_2/cm^3}$ and $3.1 \times 10^7\,\mathrm{Ar/cm^3}$.[12] Hunten *et al.*[11] suggested that densities of $10^4\,\mathrm{H_2/cm^3}$, $1\,\mathrm{H_2O/cm^3}$, and $1\,\mathrm{CH_4/cm^3}$ are more realistic for these species (for H_2O and CH_4 in particular because of a short-lifetime against photoionization). At the Moon, which is the best analog of Mercury's exosphere, O_2 has never been measured, N_2 has been measured with $8 \times 10^2\,\mathrm{N_2/cm^3}$, CO_2 with $10^3\,\mathrm{CO_2/cm^3}$ (the same density as for CO), and Ar with 4×10^4 to $10^5\,\mathrm{Ar/cm^3}$.[18] Upper limits at the Moon for O, N, C, Fe, and S below 10^3 particles/cm^3, Si and Al below 10^2 particles/cm^3 have also been reported.[18]

There are only two published works providing some estimates of the abundances of other non-identified elements.[19,20]

Only a measurement of Mercury's electron density has been performed by Mariner-10 radio occultation experiment.[21] These authors reported an upper limit for the electron density of 10^3 electrons/cm^3. This is probably why, up to now, very few studies on the ion exosphere have been published.[22–26] However, the ion exosphere has been widely quoted to play.

(i) A crucial role for the neutral exosphere first of all as the product of its ionization[11] but also as agent for sputtering of the surface.[27] Planetary ions are not only at the origin of one mechanism producing the neutral exosphere but also a significant source of enrichment of the upper surface from which the neutral exosphere is formed.[15,28]
(ii) Ions are also seen as a larger source of loss for Mercury's exosphere through solar wind entrainment rather than direct neutral loss.

The ion exosphere is fully unknown but simple calculations based on photo ionization of the neutral component can be done to infer what could be the main ion species in Mercury's exosphere.[24,29]

3. Sources of Variation of Mercury's Exosphere Content

The exospheric column and surface densities given in the previous section should be also considered with regard to all the sources of temporal and spatial variations of Mercury's exosphere that is as follows:

(i) Mercury's orbit around the Sun induces changes in the average solar flux by a factor 2.3 but also an irregular apparent motion of the Sun as seen from Mercury,
(ii) Short and long-term variations of the solar wind should significantly change Mercury's magnetosphere, and as a consequence its exosphere, through changes of solar wind sputtering rate. Coronal mass ejection or flare events could compress Mercury's magnetosphere up to its surface or open large regions to the solar wind. Magnetospheric recycling of planetary ions could also significantly influence Mercury's exospheric content and spatial distribution,
(iii) Mercury's upper surface content in volatiles is most probably non-uniform due to the preferential migration of the volatiles into regions less submitted to sputtering, desorption or vaporization (like nightside

regions or the parts of high-latitude craters permanently in the shadow). Such non-uniformity could lead to a non-uniform spatial exospheric distribution.

3.1. *Variation due to Mercury's orbit*

Mercury's eccentricity (0.2056) is the largest of the inner planets of the solar system. The variation of the heliocentric distance from perihelion (0.306 AU) to aphelion (0.466 AU) implies a variation of the average intensity of the solar wind and photon fluxes by a factor 2.3. Therefore, solar wind sputtering, photon-stimulated desorption or photon ionization will change in efficiency by such a factor in an average (also the brightness of Mercury's exosphere as seen from the Earth). Variation of the solar UV flux and of the solar wind flux on short-time scale is also an important cause of variability in the exosphere (see as an example the increase of the UV flux displayed Plate 6 of Ref. 7 and Sec. 3.2.1).

As a consequence, Mercury's surface temperature also changes significantly from a maximum temperature around 650 K at perihelion down to a maximum temperature around 550 K at aphelion.[30] Thermal desorption and the average energy of the ejected particles by this latter mechanism will also significantly change along Mercury's year.

Mercury rotates around the Sun in a 2–3 resonance between orbit (87.97 Earth days) and sidereal rotation (58.6 Earth days), its diurnal period being therefore of 176 Earth days.[31] This motion leads to an irregular apparent motion of the Sun as seen from Mercury's surface, in particular, a slowing down and subsequent reversal of the Sun's apparent motion during a short period equivalent to few Earth days at perihelion. The Sun's apparent motion at Mercury is faster at aphelion. This 2–3 resonance produces hot and cold longitudes, which correspond to regions of Mercury's surface that receive maxima and minima of solar flux along Mercury's year.[32] It also implies that Mercury's nightside surface enriched in volatiles (see Sec. 3.3.1 for further explanations), moves into the dayside with a much larger speed at aphelion than at perihelion. As a consequence, the quantity of volatiles available for ejection into Mercury's exosphere is larger at aphelion than at perihelion, and therefore, Mercury's total exospheric content may be larger at aphelion than at perihelion, a trend that seems to be confirmed by the now extended set of observation of Mercury's sodium exosphere.[9,10,14] In particular, Killen *et al.*[10] have shown by compiling a large set of observations of Mercury's sodium exosphere that Mercury's sodium exosphere total

content might be up to 2.5 times larger at aphelion than at perihelion. A model which would not consider a variation of the surface content in Mercury's surface would have predicted a global denser exosphere at perihelion than at aphelion (up to a factor 2.3) because all processes ejecting sodium atoms into Mercury's exosphere are in an average 2.3 more efficient at perihelion than at aphelion. The only way that may compensate this effect (and actually to invert it in order to be consistent with Killen *et al.*[10]) is to decrease such an efficiency. A decrease of the available reservoir for ejection from aphelion to perihelion is the only mechanism up to now that has been suggested to explain this loss of efficiency.[9]

3.2. *Magnetosphere–exosphere relations*

3.2.1. *Solar wind penetration*

Observations of enhanced and localized peaks of sodium and potassium emission in Mercury's exosphere (Fig. 1 right panel for the sodium emission) have been reported regularly soon after the first observation of sodium in Mercury's exosphere.[4] Such peaks of emissions occur usually at high latitudes and more frequently than one Mercury diurnal or sidereal cycle. These latter properties strongly suggest a relation with solar wind penetration

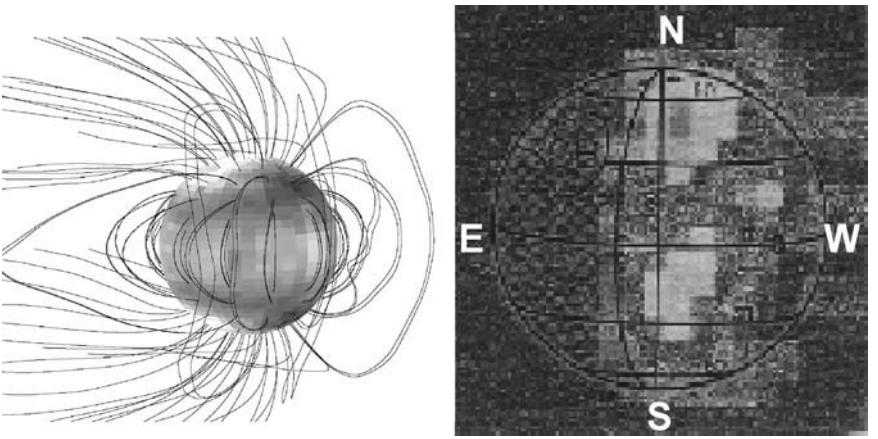


Fig. 1. Left panel: simulated magnetic field lines of Mercury's magnetosphere for nominal solar wind conditions and flux of proton solar wind impacting the surface. Maxima of this flux correspond to gray colors whereas minima correspond to white colors. From Ref. 34. Right panel: observation of high-latitude dayside peaks in Mercury's sodium exosphere (from Ref. 5), E, N, W, and S are for East, North, West, and South of the observers. The Sun is on the right part of the figure.

of Mercury's magnetosphere (Fig. 1 left panel) and subsequent sputtering of the surface by energetic incident solar protons leading to ejection of a significant number of volatiles such that they can be seen from the Earth ground-based observatories.[5] Solar wind sputtering is in particular an energetic process of ejection and is therefore, able to eject refractory species into Mercury's exosphere. Therefore, such refractory species, like calcium, should be found preferentially above the surface submitted to solar wind sputtering.[5,17] As a matter of fact, the first observations of the calcium in Mercury's exosphere reported brighter emissions at high latitudes than at the equator.[17] However, it has been recently suggested after analyzing 4 years of observations, that the origin of the observed exospheric calcium atoms is preferentially formed by micrometeoroid ejection as molecule and subsequent photodissociation.[33]

3.2.2. *Solar event encounter*

The encounter of a solar event with Mercury could lead to a significant enhancement of the incident dynamic pressure, such that the dayside magnetosphere is no longer able to efficiently shield Mercury's surface from the incident solar wind and energetic particles.[34,35] Potter *et al.*[36] suggested a relationship between the observation of a global and rapid increase of the total content of Mercury's sodium exosphere and the occurrence of Mercury-directed Coronal Mass Ejections observed by the SOHO spacecraft coronagraphs. Leblanc *et al.*[37] used a sample energetic particle event observed at Earth to estimate some of the effects of such an event on the surface sputtering contribution to Mercury's sodium exosphere. These authors concluded that fluxes several orders of magnitude larger would be required to produce Potter *et al.*'s[36] inferred sodium enhancements. Such events are fully within the range of variation of known solar energetic particle events.

3.2.3. *Magnetospheric recycling*

Delcourt *et al.*[26] found that planetary Na^+ ions may convect to the nightside, be accelerated and then hit the surface in a non-uniform way.[29,38] This feature is confirmed by hybrid simulations of ion circulation at Mercury.[34] Leblanc *et al.*[28] concluded that less than 15% of the photoion reimpact Mercury's surface. These authors using typical solar wind conditions found that enhancement in the Na emissions could be correlated to these bands of re-implantation. They also suggested that any change in the solar wind

conditions could change the magnetospheric structure[39] and induce significant changes in the recycling of magnetospheric ions inducing correlated variation in the neutral exosphere. Using different neutral exospheric and magnetospheric models, Killen *et al.*[10] found that between 45 and 65% of the photoions reimpact the surface, most on the dayside and underlined the significant role of the electric field of convection in the global dawn to dusk balance of the recycled ion.

3.3. *Upper surface–exosphere relations*

3.3.1. *Migration of the volatiles*

A particle ejected from Mercury's dayside not on an escaping trajectory (that is with less than ∼0.1 eV/amu when ejected) will randomly hop and become temporally absorbed (as example for most of the sodium atoms) or partially energetically accommodated in Mercury's upper surface at each hop (as example for most of the hydrogen and helium atoms).[40] Due to the very high temperature of Mercury's dayside surface (more than 400 K during most of the day, see Fig. 2), the period during which such a particle is absorbed in Mercury's upper surface is shorter than few Earth minutes or few Mercury's seconds (values valid in the case of a sodium atom using recent laboratory measurements[41] and including porosity effect[9]). Therefore, the time needed for such a particle to encounter a cold surface where it can be trapped for a long period (with respect to Mercury's day) will be much shorter than Mercury's hour. These cold regions are essentially late evening, early morning or high-latitude regions. This migration of volatiles into cold regions should lead to a global larger density of trapped volatiles on the nightside than on the dayside.[9] An immediate consequence is a significant morning/evening asymmetries of Mercury's exosphere because of the release of trapped nightside volatiles at morning. Such morning/evening asymmetry in the case of Mercury's sodium exosphere has been discussed for a long-time up to its recent and unambiguous observation.[42]

Another putative consequence of this preferential migration of ejected volatiles into cold surfaces could be the enrichment of Mercury's surface at high-latitude early morning. Indeed as illustrated in Fig. 2, in the morning, equatorial regions get hotter earlier by few degrees in longitude than high-latitude regions. As an example, 60° latitude regions reach equatorial surface temperature few tens of Earth hours later. Such a period is roughly the time for a thermally desorbed particle (with less than 0.06 eV when ejected) to move by a distance of the order of Mercury radius. Such a time

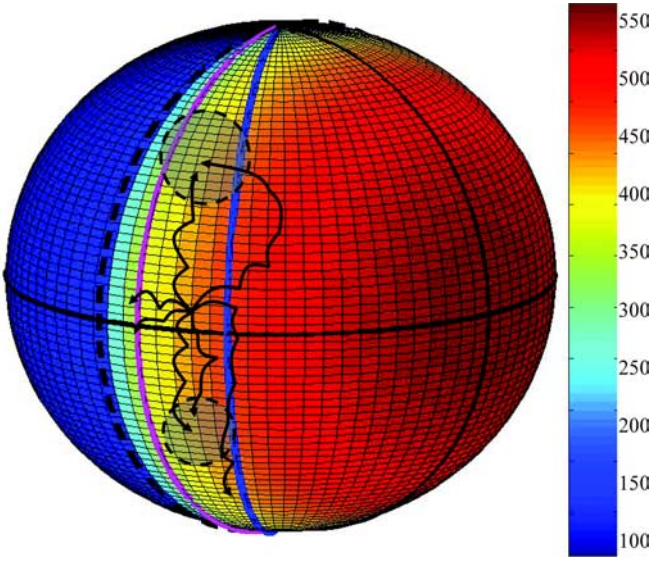


Fig. 2. Mercury's upper surface temperature calculated for average heliocentric Mercury's distance as seen from Mercury's morning side. Thin horizontal dark line is for the equator and vertical thin dark line for the subsolar line. Violet and blue thick lines are longitude 10° and 30° after the dawn terminator (dashed dark line). Also drawn are examples of trajectories of sodium particle in Mercury's exosphere starting from equatorial early morning regions and moving by successive cycles of temporal absorption/ejection/reabsorption up to the moment when the time of residence in the regolith gets of the order to the diurnal cycle.

should be even shorter for particle ejected by a more energetic process (with less than few eVs). Therefore, the depletion of equatorial regions in favor of higher latitude regions may occur during a period sufficient enough to significantly increase the absorbed volatile content of these latter regions. Such regions should then release more absorbed volatiles than equatorial regions and should be associated to high-latitude peaks of exospheric density.[9]

However, Killen *et al.*[10] argued that because of the high temperature of the equatorial regions, the sodium atoms continue to diffuse to the extreme surface of grains at a rate sufficient to maintain all known source rates during few Earth days. Therefore, this diffusion should reduce the signature associated to the enhancement of absorbed volatiles in the upper surface at high latitude when released into the exosphere. Moreover, more energetic processes like photon-stimulated desorption and solar wind sputtering should also deplete these enriched regions.

3.3.2. *Topography effects*

Another relation between upper surface and exosphere is related to the presence of large craters and basins at Mercury's surface. In particular, Sprague[43] suggested relation between the presence in Mercury's morning of Caloris basin, a crater with a diameter equal to half-Mercury's radius, and an apparent increase in Mercury's potassium total content. In the same way, Potter *et al.*[36] observed a three times enhancement of Mercury's total sodium content when Caloris basin was in Mercury's early morning. Such type of large topographic anomalies at Mercury's surface as well as smaller craters[14] could therefore, have an important role in supplying Mercury's exosphere in volatiles.[44]

4. Conclusions

Up to now, only few elements of Mercury's exosphere have been identified, most probably representing only few percents of the total content of the exosphere. Moreover, the influences of Mercury's magnetosphere and of Mercury's surface composition are key points to be solved in order to fully understand the origin, dynamic, and composition of Mercury's exosphere.

Any description of Mercury's exosphere has then to consider the numerous sources of variation which are:

(i) The changes along Mercury's orbit of the different mechanisms of production, destruction, and recycling.
(ii) The role of the solar wind and photon flux and their short and long-times variations.
(iii) The changes of Mercury's magnetosphere with respect to the interplanetary magnetic field and also possible encounters with solar event.
(iv) The role of large and small topographic structures at Mercury's surface.

Very little has been said in this paper about the different mechanisms of ejection from the surface thought to be at work at Mercury.[7,45] Actually, the way each species is ejected from Mercury's upper surface defines their initial energy and therefore, their spatial distribution. Spatial variations also depend on the regions of Mercury's surface where such species are produced and the variations of the ejected flux intensity with respect to solar conditions, to distance to the Sun and to position on Mercury's surface.

At the end, the solar radiation pressure has been suggested to model the shape of Mercury's exosphere for some species.[46] In particular, it has

been suggested that Mercury's exospheric sodium tail recently observed[47] is essentially shaped by the solar pressure acting on the sodium atoms ejected from Mercury's surface.[9]

References

1. R. M. Killen and W.-H. Ip, *Rev. Geophys.* **37** (1999) 361–406.
2. F. Leblanc, E. Chassefière, R. E. Johnson, D. M. Hunten, E. Kallio, D. C. Delcourt, R. M. Killen, J. G. Luhmann, A. E. Potter, A. Jambon, G. Cremonese, M. Mendillo, N. Yan and A. L. Sprague, Submitted to Planetary Space Science (2006).
3. A. Milillo, P. Wurz, S. Orsini, D. Delcourt, E. Kallio, R. M. Killen, H. Lammer, S. Massetti, A. Mura, S. Barabash, G. Cremonese, I. A. Daglis, E. De Angelis, A. M. Di Lellis, S. Livi, V. Mangano and K. Torkar, *Space Sci. Rev.* **117** (2005) 397–443.
4. A. E. Potter and T. H. Morgan, *Science* **229** (1985) 651–653.
5. A. E. Potter and T. H. Morgan, *Science* **248** (1990) 835.
6. A. E. Potter, T. H. Morgan and R. M. Killen, *Plan. Space Sci.* **47** (1999) 1441–1448.
7. R. M. Killen, A. E. Potter, P. Reiff, M. Sarantos, B. V. Jackson, P. Hick and B. Giles, *J. Geophys. Res.* **106**, (2001) 20,509–20,525.
8. D. M. Hunten and A. L. Sprague, *Adv. Space Res.* **19** (1997) 1551–1560.
9. F. Leblanc and R. E. Johnson, *Icarus* **164** (2003) 261–281.
10. R. M. Killen, M. Sarantos, A. E. Potter and P. H. Reiff, *Icarus* **171** (2004) 1–19.
11. D. M. Hunten, T. M. Morgan and D. M. Shemansky, The Mercury atmosphere, in *Mercury*, eds. F. Vilas, C. Chapman and M. Matthews (University of Arizona Press, Tucson 1988), pp. 562–612.
12. A. L. Broadfoot, D. E. Shemansky and S. Kumar, *Geophys. Res. Lett.* **3** (1976) 577–580.
13. D. E. Shemansky, *The Mercury Messenger 2* (Lunar and Planet. Inst., Houston, Tex. 1988), p. 1.
14. A. L. Sprague, R. W. H. Kozlowski, D. M. Hunten, N. M. Schneider, D. L. Domingue, W. K. Wells, W. Schmitt and U. Fink, *Icarus* **129** (1997) 506–527.
15. A. L. Sprague, *J. Geophys. Res.* **97** (1992) 18257.
16. A. L. Sprague, *J. Geophys. Res.* **98** (1992) 1231.
17. T. A. Bida, R. M. Killen and T. H. Morgan, *Nature* **404** (2000) 159–161.
18. S. A. Stern, *Rev. Geophys.* **37** (1999) 453–491.
19. T. H. Morgan and R. M. Killen, *Planet. Space Sci.* **45** (1997) 81.
20. P. Wurz and H. Lammer, *Icarus* **164** (2003) 1–13.
21. G. Fjelbo, A. Kliore, D. Sweetnam, P. Esposito, B. Seidel and T. Howard, *Icarus* **29** (1976) 407–415.
22. A. F. Cheng, R. E. Johnson, S. M. Krimigis and L. Z. Lanzeroti, *Icarus* **71** (1987) 430–440.

23. W. H. Ip, *Geophys. Res. Lett.* **14** (1987) 1191.
24. R. Lundin, S. Barabash, P. Brandt, L. Eliasson, C. M. C. Nairn, O. Norberg and I. Sandahl, *Adv. Space. Res.* **19** (1997) 1593.
25. D. C. Delcourt, T. E. Moore, S. Orsini, A. Milillo and J.-A. Sauvaud, *Geophys. Res. Lett.* **29**, 12 (2002) doi:10.1029/2001GL013829.
26. D. C. Delcourt, S. Grimald, F. Leblanc, J.-J. Bertherlier, A. Millilo and A. Mura, *Ann. Geophysicae* **21** (2003) 1723–1736.
27. W.-H. Ip, *Astrophys. J.* **418** (1993) 451.
28. F. Leblanc, D. Delcourt and R. E. Johnson, *J. Geophys. Res.* **108** (2003) 5136, doi:10.1029/2003JE002151.
29. F. Leblanc, H. Lammer, K. Torkar, J. J. Berthelier, O. Vaisberg and J. Woch, *Notes du Pôle de Planétologie de l'IPSL* http://www.ipsl.jussieu.fr/Documentation, 2004, **5**.
30. A. S. Hale and B. Hapke, *Icarus* **156** (2002) 318–334.
31. G. Colombo, *Nature* **208** (1965) 575.
32. S. Soter and J. Ulrichs, *Nature* **214** (1967) 1315–1316.
33. R. M. Killen, T. A. Bida and T. H. Morgan, *Icarus* **173** (2005) 300–311.
34. E. Kallio and P. Janhunen, *Geophys. Res. Lett.* **30** (2003) 1877, 10.1029/2003GL017842.
35. K. Kabin, T. I. Gombosi, D. L. DeZeeuw and K. G. Powell, *Icarus* **143** (2000) 397–406.
36. A. E. Potter, T. H. Morgan and R. M. Killen, *Planet. Space Sci.* **47** (1999) 1441–1448.
37. F. Leblanc, J. G. Luhmann, R. E. Johnson and M. Liu, *Planet. Space Sci.* **51** (2003) 339–352, doi:10.1016/S0032-0633(02)00207-6.
38. D. C. Delcourt and F. Leblanc, *Notes du Pôle de Planétologie de l'IPSL* http://www.ipsl.jussieu.fr/Documentation, 2005, p. 12.
39. J. G. Luhmann, C. T. Russell and N. A. Tsyganenko, *J. Geophys. Res.* **103** (1998) 9113–9119.
40. W. H. Smyth and M. L. Marconi, *Icarus* **441** (1995) 839–864.
41. B. V. Yakshinskiy, T. E. Madey and V. N. Ageev, *Surf. Rev. Let.* **7** (2000) 75–87.
42. H. Schleicher, G. Wiedemann, H. Wöhl, T. Berkefeld and D. Soltau, *A&A* **425** (2004) 1119–1124.
43. A. L. Sprague, *Icarus* **84** (1990) 93–105.
44. N. Yan, F. Leblanc and E. Chassefière, *Icarus* (2006).
45. R. E. Johnson, Surface boundary layer atmospheres, *Chapter in Atmospheres in the Solar System: Comparative Aeronomy Geophysical Monograph* **130** (2002) 203–219.
46. W. H. Smyth, *Nature* **323** (1986) 696–699.
47. A. E. Potter, R. M. Killen and T. H. Morgan, *Meteoritic and Planetary Science* **37** (2002) 1165–1172.
48. A. E. Potter and T. H. Morgan, *Icarus* **67** (1986) 336–340.

This page intentionally left blank

ON THE DYNAMICS OF CHARGED PARTICLES IN THE MAGNETOSPHERE OF MERCURY

DOMINIQUE C. DELCOURT[*,†] and KANAKO SEKI[‡,§]
[†]*CETP-CNRS, Institut Pierre Simon Laplace*
4 avenue de Neptune, 94107 Saint-Maur des Fossés, France
[‡]*Solar-Terrestrial Environment Laboratory, Nagoya University*
Honohara 3-13, Toyokawa, Aichi 442-8507, Japan
[*]*dominique.delcourt@cetp.ipsl.fr*
[§]*seki@stelab.nagoya-u.ac.jp*

We review some features of ion dynamics in Mercury's magnetosphere using single-particle simulations. Not unexpectedly, the small spatial and temporal scales of this magnetosphere lead to non-adiabatic transport features that have a variety of implications such as extreme sensitivity to initial conditions upon injection into the magnetosphere, formation of beamlets and of a thin current sheet in the magnetotail, or large scale filtering due to the finite width of the magnetosphere with respect to the ion Larmor radius. We show that ions are rapidly transported and energized within Mercury's magnetosphere, with possibly significant recycling of planetary material. The occurrence of reconnection in the magnetotail however, may substantially alter the global convection pattern, and cause enhanced down-stream losses. We demonstrate that large non-adiabatic energization may be achieved for electrons as well, in particular during presumed expansion phases of substorms which lead to short-lived precipitation onto the planet surface and formation of bouncing electron clusters.

1. Introduction

Mariner-10 observations in 1974–1975 revealed an intrinsic magnetic field at Mercury, with a reduced (by about 2 orders of magnitude) dipolar moment as compared to that of Earth. The spatial and temporal scales of the resulting magnetosphere differ widely from those of the terrestrial one (by factors of $\sim$7 and $\sim$30, respectively). The actual structure of Mercury's magnetosphere (in particular, the existence of large-scale plasma cells such as lobes, plasma sheet or boundary layers) remains to be elucidated. Also, it is not known if and to which amount ions originating from the planetary exosphere contribute to the magnetospheric populations as is the case for ionospheric ions at Earth. Although it is expected that the magnetopause is frequently (up to 30% of the time) blown down to the planet surface due to enhanced solar wind pressure, it is neither known whether the solar wind forms a

significant plasma supply to Mercury's magnetosphere. This has motivated several numerical studies using a variety of modeling approaches (MHD, single-particle, hybrid) (e.g., Refs. 1–3). The purpose of this paper is to review some results obtained from single-particle simulations. These simulations were performed using either an analytical model of the hermean magnetosphere (adapted from Ref. 4) or results of MHD simulations. Section 2 is dedicated to consequences of small spatial scales, whereas Sec. 3 focuses on consequences of small temporal scales.

2. Adiabaticity Breaking Due to Small Spatial Scales

Because of the weak intrinsic magnetic field of Mercury and of the enhanced dynamical pressure of the solar wind, the hermean magnetosphere exhibits spatial scales that are much smaller (by about a factor 7) than those of the terrestrial magnetosphere. This raises questions for the non-linear dynamics of charged particles since their Larmor radii must be small compared to the characteristic scale of magnetic field variations for their motion to be adiabatic (equivalently, for the guiding center approximation to be valid). As a matter of fact, a simple calculation of the adiabaticity parameter κ defined as the square root of the minimum curvature radius to maximum Larmor radius ratio[5] reveals that this parameter likely is smaller than 3 throughout most of the magnetotail for keV electrons and ions (e.g., Ref. 2); hence, prominent non-adiabatic features.

In a study intended to examine the entry of solar wind ions in Mercury's magnetosphere, Delcourt and Leblanc[6] showed that access of these ions to the inner magnetosphere is restricted to a limited domain of the phase space. This is illustrated in the right panels of Fig. 1 that show the color-coded altitude, energy, and time of flight of solar wind protons as a function of initial latitude and longitude at the magnetopause. Here, test H^+ were initialized with 100 eV and 90° pitch angle. Test H^+ launched with small pitch angles or large energies ($\sim$1 keV) were found to escape into the magnetotail due to transit times significantly smaller than the convection time scale (see, e.g., Fig. 1 of Ref. 6). In the right panels of Fig. 1, white areas correspond to test ions that reach the model boundary at 6 R_M radial distance without intercepting the magnetopause or the planet surface. It is clearly apparent from these panels that the solar wind H^+ that precipitate onto Mercury's surface originate from a limited longitudinal interval in the dawn sector (note that longitudes of 0° and –90° correspond to local times of 1200 and 0600, respectively).

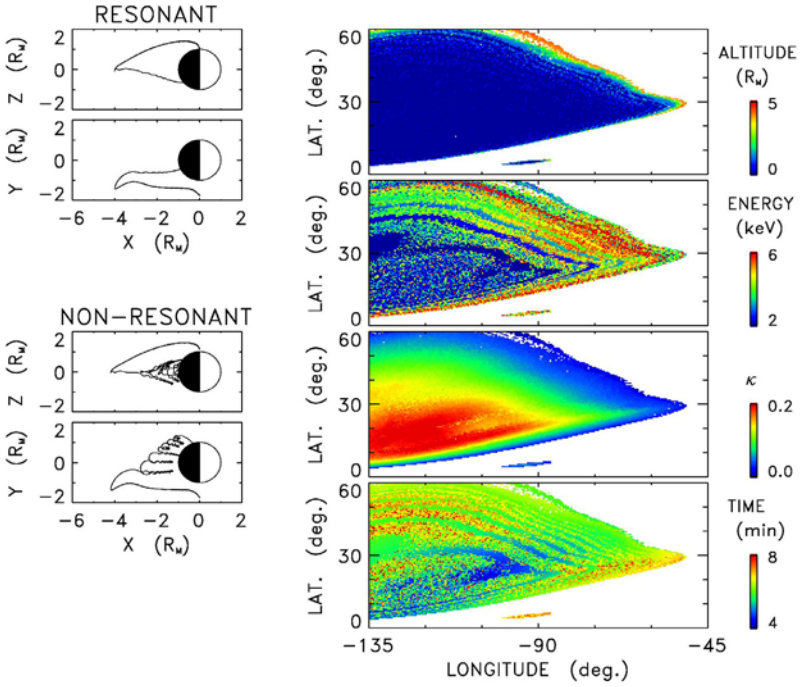


Fig. 1. (*Right; from top to bottom*) Final altitude, energy, minimum adiabaticity parameter κ, and time of flight of solar wind H^+ as a function of initial longitude and latitude at the magnetopause. Test H^+ are traced until they reach the model boundary (i.e., the magnetopause or a radial distance of 6 R_M) or precipitate onto Mercury's surface. (*left*) Examples of (*top*) resonant and (*bottom*) non-resonant particles (adapted from Ref. 6).

In the right panels of Fig. 1, a prominent structuring of the precipitating ion energy and time of flight can also be seen, that directly follows from non-adiabatic transport in the mid-tail. Indeed, in their analysis of non-linear particle dynamics in a field reversal, Chen and Palmadesso[7] put forward that, for some values (smaller than 1) of the κ parameter, particles preferentially execute Speiser-type orbits whereas, for other κ values, most of the particles experience prominent magnetic moment scattering and are temporarily trapped inside the current sheet. This feature was interpreted as the result of resonance between the fast meandering motion about the mid-plane and the slow gyromotion about the normal field component. To characterize these resonances, Chen and Palmadesso[7] derived the following empirical relationship: $\kappa_n \approx 0.8/(n + 0.6)$, where n is an integer ≥ 1. These resonances coincide with the quasi-adiabatic regime put forward by

Büchner and Zelenyi.[5] They contribute to the formation of thin current sheets and are responsible for the beamlets that are frequently observed in the Earth's plasma sheet boundary layer (e.g., Refs. 8 and 9).

In Fig. 1, resonant particles and associated beamlets are characterized by short residence times (upper left panel), whereas quasi-trapped particles exhibit longer residence times as well as larger energies due to repeated interactions with the current sheet and larger drift in the Y-direction (lower left panel). It is thus clearly apparent from Fig. 1 that solar wind protons entering Mercury's magnetosphere exhibit an extreme sensitivity to initial conditions at the magnetopause. That is, a slight change in injection position leads to a large variation in the adiabaticity parameter κ, an effect which is of importance for hybrid simulations and contrasts with the situation prevailing at Earth.

The small width of Mercury's magnetotail has another important implication for the dynamics of charged particles. Indeed, in an attempt to estimate the contribution of planetary ions to the magnetospheric populations, Delcourt *et al.*[2] showed that low-energy Na^+ ions sputtered from the planet surface may gain access to the magnetotail via convection over the polar cap, in a like manner to low-energy ions upflowing from the topside ionosphere at Earth. In the magnetospheric lobes, these ions are subjected to a prominent parallel energization due to an $\mathbf{E} \times \mathbf{B}$ related centrifugal effect that is more pronounced (because of a larger angular speed of the convecting field lines) than at the Earth. Once they interact with the magnetotail current sheet, these Na^+ ions are accelerated in a non-adiabatic manner and may subsequently precipitate onto the planet surface, hence contributing to regolith sputtering. Delcourt *et al.*[2] showed that, on the whole, exospheric Na^+ form a substantial (several tenths of ions/cm^3) source of plasma for the magnetotail. As envisioned in the "geopause" interpretation framework of Moore and Delcourt,[10] this raises the question of the net magnetospheric contribution of planetary material versus that of the solar wind, especially when the planet occupies a large volume of the magnetosphere as is the case at Mercury.

An example of planetary ion behaviors is shown in Fig. 2 that shows the trajectories of test Na^+ launched from two distinct latitudes in the dayside sector (left panels), together with their energy and magnetic moment as a function of time (center panels). As for the right panels of Fig. 2, they show the results of systematic Na^+ trajectory computations, with grey-coded flux and energy of precipitating ions as a function of longitude and latitude. It is apparent from the right panels of Fig. 2 that, after transport into the

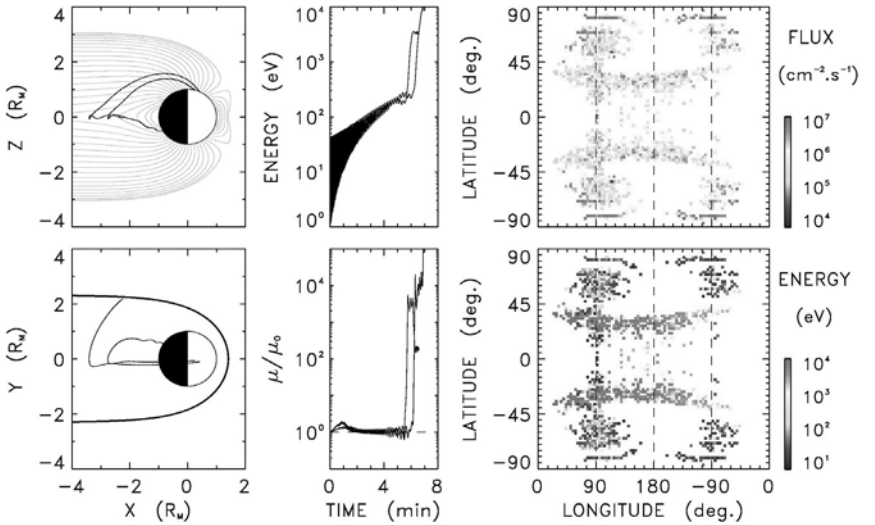


Fig. 2. (*Left*) Model Na^+ trajectories projected in the noon–midnight plane and in the equatorial plane. (*Center*) Corresponding energy and magnetic moment versus time. (*Right*) Grey-coded flux and average energy of precipitating Na^+ ions as a function of longitude and latitude (adapted from Ref. 2).

magnetosphere, Na^+ ions originating from Mercury's exosphere give rise to two narrow bands of precipitation that extend over 10°–20° in latitude. The average ion energy within these bands increases from post-midnight to pre-midnight sectors, as expected from westward drift and associated acceleration by the large-scale convection electric field. These energetic ion bands contrast with those observed at high (above ∼50°) latitudes in Fig. 2, which correspond to downflowing exospheric populations that do not interact with the magnetotail current sheet.

An intriguing feature in the right panels of Fig. 2 is the abrupt cutoff of the Na^+ precipitation at latitudes of 35°–40°. In the Earth's auroral zone, this poleward boundary roughly delineates the transition between polar cap and plasma sheet, that is, the transition between the open flux region and closed magnetic field lines that extend up to the reconnection site in the far tail. In contrast, on the equatorward edge, precipitation is controlled to some extent by stochastic motion in the inner plasma sheet, the Isotropic Boundary (IB) of downward flux delineating the transition between adiabatic and non-adiabatic transport (e.g., Refs. 11 and 12). At Mercury, as far as IB is concerned, it is expected to occur at fairly low

latitudes since particles behave in a non-adiabatic manner throughout most of the magnetotail. As for the poleward boundary, it can be seen in the bottom left panel of Fig. 2 that it does not map into the far tail but at distances of a few planetary radii and that it is due to the fact that ions intercept the dusk magnetopause in the course of their interaction with the current sheet. That is, in contrast to the situation prevailing at Earth, the poleward boundary of precipitation in the right panels of Fig. 2 directly follows from the finite ion Larmor radius. At large energies and/or large mass-to-charge ratios, these Larmor radii become comparable to or larger than the magnetotail width, in which case ions do not execute a full Speiser orbit and are not reflected toward the planet. Here, it may actually be suspected that this finite Larmor radius boundary at the poleward edge of precipitation depends upon ion species (with a cutoff for heavy ions occurring at lower latitudes than for H^+) and provides information of the magnetotail structure, in a like manner to IB.

The results displayed in Figs. 1 and 2 were obtained using a modified version of the Luhmann and Friesen[4] model together with a two-cell pattern of magnetospheric convection. For comparison, Fig. 3 shows the results obtained in the study of Seki *et al.*,[13] that examines the transport of exospheric Na^+ in the electric and magnetic fields derived from MHD simulations of the Mercury–solar wind interaction, assuming a southward orientation of the interplanetary magnetic field (IMF). While the analytical

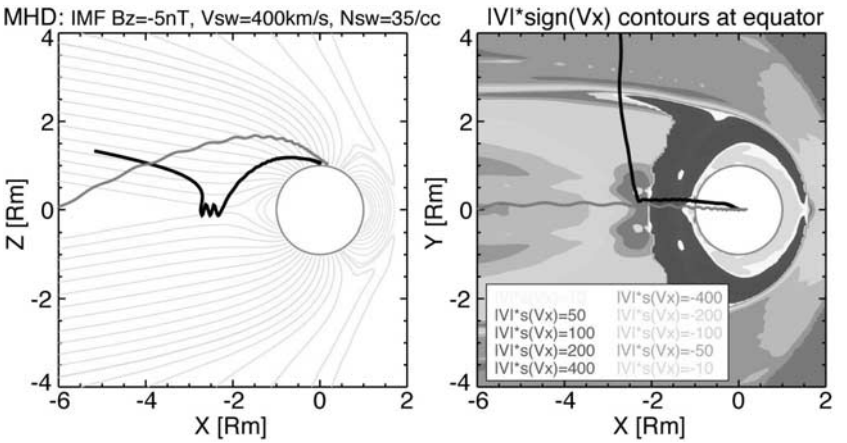


Fig. 3. Identical to the left panels of Fig. 2 but using electric and magnetic fields derived from MHD simulations (adapted from Ref. 13).

model implicitly assumes the existence of a neutral line in the far tail, the MHD results show the formation of a near-Mercury neutral line (NMNL) near 2 R_M radial distance (right panel of Fig. 3). In the terrestrial magnetosphere, the formation of such a neutral line near the planet is commonly associated with substorm activity. It is apparent from Fig. 3 that the existence of such a NMNL can significantly alter the Na^+ trajectories and change the ratio of ions that precipitate onto the planet surface. That is, ions that reach the region of tailward flow in the plasma sheet do not return to the planet regardless of their κ parameter. These results suggest that precipitation and recycling of planetary ions at Mercury depend upon the magnetic activity and the large-scale convection pattern.

3. Adiabaticity Breaking Due to Small Temporal Scales

During the Mariner-10 pass of 1974, the magnetic field first pointed tailward and abruptly turned northward after closest approach of the planet. Interestingly, high-energy (several tens of keVs and above) electron injections were recorded in conjunction with this rapid change of the magnetic field orientation, a feature which is reminiscent of that observed during substorm dipolarization at Earth. This suggests that the hermean magnetotail may be subjected to substorm cycles as well (e.g., Ref. 14), although Luhmann *et al.*[15] put forward that these observations may provide indication of processes directly driven by the solar wind.

If substorms occur at Mercury in a like manner to Earth, significant particle energization is to be expected due to the electric field induced by relaxation of the magnetic field lines. By means of two-dimensional simulations, Ip[16] actually demonstrated that protons may be accelerated up to 10–15 keV during such reconfiguration events. On the other hand, because of the short-time scales that characterize the hermean environment, it may be anticipated that electrons are subjected to significant energization as well. In a recent study, Delcourt *et al.*[17] examined this issue using a rescaled version of a time-dependent particle code previously developed to investigate the dynamics of charged particles during substorms at Earth (see, Ref. 18 for details on the modeling technique).

An example of the results obtained by Delcourt *et al.*[17] is presented in Fig. 4. As mentioned above, characteristic time scales at Mercury are expected to be smaller than those at Earth by about a factor 30 (e.g., Ref. 19), and a short-lived (5 s) dipolarization was accordingly considered here. The left panels of Fig. 4 show selected electron trajectories during

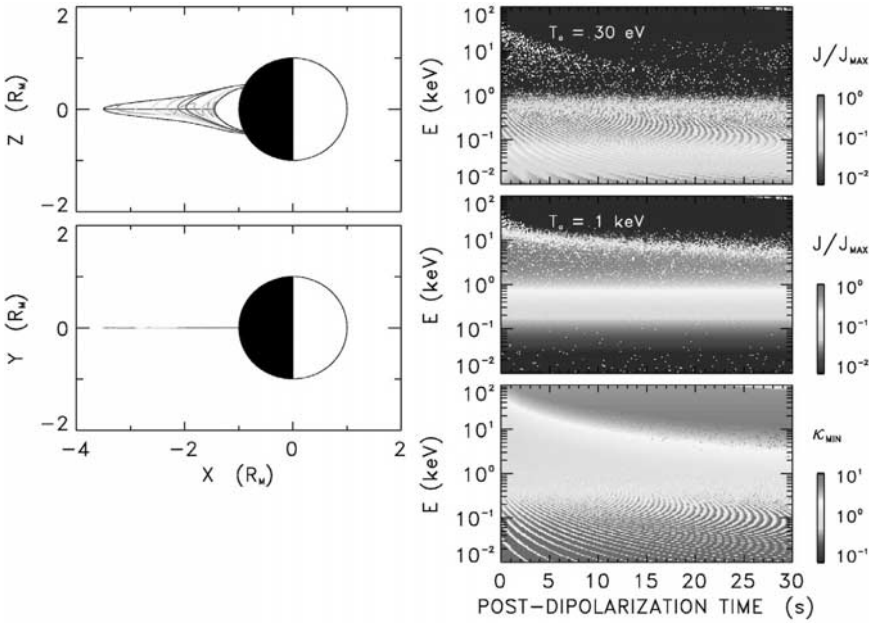


Fig. 4. (*Left*) Model electron trajectories during dipolarization. These trajectories are shown (*top*) in the X–Z plane and (*bottom*) in the X–Y plane, the different colors corresponding to different initial positions along the field line. Black lines in the upper left panel show the magnetic field line before and after dipolarization. (*Right; from top to bottom*) Grey-coded flux (normalized to the maximum value) and adiabaticity parameter κ of electrons as a function of post-dipolarization time and energy. These variations are shown at 1.5 R_{M} radial distance and for two different temperatures (30 eV and 1 keV) of the initial Maxwellian distribution (adapted from Ref. 17).

dipolarization. Here, the test particles were initialized with 10 eV energy and 90° pitch angle at different locations (separated by equidistant steps of 500 km) along the field line. The convection surge due to the large electric field induced by dipolarization of the magnetic field line (e.g., Ref. 20) is clearly noticeable in these panels that display a rapid inward injection of the electrons. Also, because of the low-energy character of the particles (up to a few hundreds of eV after dipolarization), no significant drift in the Y-direction can be seen in the bottom left panel.

The upper right panels of Fig. 4 show the energy–time spectrograms obtained for electrons with 30° pitch angle. Note here that the initial electron distribution is assumed to be Maxwellian with two distinct initial temperatures (30 eV and 1 keV). In addition, the bottom right panel of Fig. 4 displays the minimum adiabaticity parameter κ encountered during

transport as a function of time and energy. A striking feature here is the formation of bands of enhanced flux that gradually evolve with time. This effect is obtained for low-initial temperature (top right panel) and vanishes at higher initial temperatures (center right panel). The reason for this effect may be understood by examining the κ variations in the bottom right panel of Fig. 4. Indeed, it is apparent from this latter panel that electrons with energies up to a few hundreds of eV have κ much large than unity. They accordingly behave in an adiabatic (magnetic moment conserving) manner. These test electrons however experience a different parallel energization depending upon bounce phase at the dipolarization onset (violation of the second adiabatic invariant), and Liouville theorem that states conservation of the particle density in phase space leads to higher flux for those electrons that experience larger energy gains.

On the other hand, it can be seen in the bottom right panel of Fig. 4 that, above a few hundreds of eV, electrons have κ of the order of unity or below. These are accordingly subjected to magnetic moment scattering which hampers the build-up of the above structuring due to differential parallel energization. In other words, for $T_0 = 30\,\mathrm{eV}$ (top right panel), most of the electrons are found to behave adiabatically with respect to the first invariant but non-adiabatically with respect to the second one, whereas for $T_0 = 1\,\mathrm{keV}$ (center right panel) the bulk of the electron population behaves non-adiabatically with respect to the first invariant. Finally, note that at high energies (several tens of keV), electrons are less sensitive to the convection surge induced by the dipolarization and that they drift in the immediate vicinity of the planet; hence their large κ parameter.

In the Earth's magnetosphere, Mauk[20] demonstrated that the bouncing ion clusters frequently observed in the inner plasma sheet (e.g., Refs. 21 and 22) likely follows from such differential parallel energy gains and consequent violation of the second adiabatic invariant during substorm dipolarization. At Earth, reconfiguration of the magnetic field lines typically occurs on time scales of the order of a few minutes and the above flux modulation accordingly affect essentially ions that have comparable bounce periods. In contrast, at Mercury, it is apparent from Fig. 4 that, because of small temporal scales, such a flux modulation may be obtained for electrons and it may be speculated that, in a like manner to bouncing ion clusters at Earth, the build-up of bouncing electron clusters will provide information on the dynamics of the magnetotail.

Finally, it was shown in Ref. 18 that, in the Earth's magnetosphere, the first adiabatic invariant may not be conserved during dipolarization which allows for enhanced perpendicular energization. Given the relaxation time

scale, such a non-adiabatic heating favors ions with large mass-to-charge ratios because of their large gyration periods. This process may be responsible for the species-dependent energization observed in the storm-time ring current (e.g., Ref. 23). At Mercury, one may expect this non-adiabatic heating to affect plasma sheet H^+ populations due to the short-time scale of magnetic transitions.

This is illustrated in Fig. 5 that shows model H^+ trajectories in a 5-second dipolarization identical to that in Fig. 4. Here, test protons were initialized with 100 eV energy and 90° pitch angle at different (grey-coded) positions along the dipolarizing field line. It is clearly apparent from Fig. 5 that most of the test H^+ are subjected to prominent magnetic moment enhancement and energization (up to several tens of keV), with the exception of the ions initialized close to the planet surface that have small Larmor radii and small gyration periods. One expects that a number of these non-adiabatically accelerated ions will be lost at the dusk magnetopause because of their large Larmor radii after dipolarization.

In this regard, note that ions at Mercury generally do not conserve their magnetic moment even in the steady state case (see, e.g., Fig. 1) so

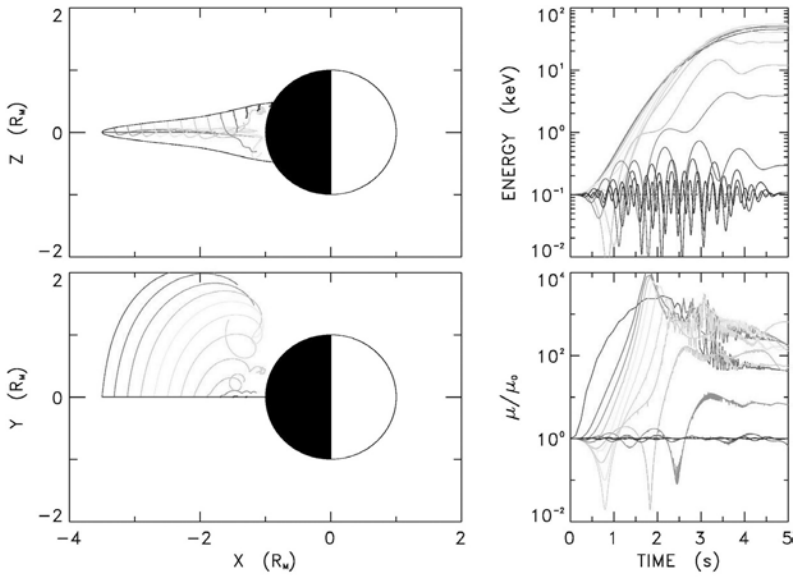


Fig. 5. Model H^+ orbits during a 5-second dipolarization of the magnetic field lines: (*left*) trajectory projections in noon–midnight and equatorial planes, (*right*) energy, and magnetic moment versus time.

that a complex intrication of features due to both spatial and temporal non-adiabaticity is to be expected (e.g., Ref. 24). A future study will be dedicated to comparative analysis of proton and heavy ion transport under such conditions.

4. Summary

Single-particle simulations of ion circulation within Mercury's magnetosphere reveal prominent deviations from adiabaticity, be it due to small spatial scales with respect to the particle Larmor radius or due to small temporal scales with respect to the particle gyration period. Ions that circulate in the inner magnetotail exhibit a prominent sensitivity to injection conditions because of a sharp gradient in the magnetic field magnitude and field line elongation. This leads to rapid changes from quasi-adiabatic to quasi-trapped behaviors and vice versa. The small width of the hermean magnetotail also is of importance in this respect since ions with large Larmor radii may not be reflected toward the planet after interaction with the current sheet and intercept the dusk magnetopause. Ions are found to be rapidly transported and energized within Mercury's magnetosphere, the significance of recycling and/or down-stream loss depending upon the occurrence of reconnection in the magnetotail and upon the global convection pattern. The characteristic scales of Mercury's magnetosphere are such that electrons are found to behave non-adiabatically as well. In particular, violation of the second adiabatic invariant due to the transient electric field induced by dipolarization of the magnetic field lines may lead to short-lived precipitation onto the planet surface as well as formation of bouncing electron clusters in the inner magnetotail.

Acknowledgments

Part of this work was performed while D. C. Delcourt was residing at STEL, Toyokawa, Nagoya University (Japan).

References

1. K. Kabin, T. I. Gombosi, D. L. DeZeeuw and K. G. Powell, *Icarus* **143** (2000) 397.
2. D. C. Delcourt, S. Grimald, F. Leblanc, J.-J. Berthelier, A. Millilo, A. Mura, S. Orsini and T. E. Moore, *Ann. Geophys.* **21** (2003) 1723.
3. E. Kallio and P. Janhunen, *Geophys. Res. Lett.* **30** (2003) DOI 10.1029/2003GL017842.

4. J. G. Luhmann and L. M. Friesen, *J. Geophys. Res.* **84** (1979) 4405.
5. J. Büchner and L. M. Zelenyi, *J. Geophys. Res.* **94** (1989) 11,821.
6. D. C. Delcourt and F. Leblanc, *Notes du Pôle de Planétologie* (Institut Pierre Simon Laplace, 2005), p. 12.
7. J. Chen and P. J. Palmadesso, *J. Geophys. Res.* **91** (1986) 1499.
8. M. Ashour-Abdalla, L. A. Frank, W. R. Paterson, V. Peroomian and L. M. Zelenyi, *J. Geophys. Res.* **101** (1996) 2587.
9. J.-A. Sauvaud and R. A. Kovrazhkin, *J. Geophys. Res.* **109** (2004) DOI 10.1029/2003JA010333.
10. T. E. Moore and D. C. Delcourt, *Rev. Geophys.* **33** (1995) 175.
11. V. A. Sergeev M. Malkov and K. Mursula, *J. Geophys. Res.* **98** (1993) 7609.
12. P. T. Newell, Y. I. Feldstein, Y. I. Galperin and C.-I. Meng, *J. Geophys. Res.* **101** (1996) 10,737.
13. K. Seki, N. Terada, D. C. Delcourt and T. Ogino (abstract), SGEPSS, Kyoto, Japan (2005).
14. G. L. Siscoe, N. F. Ness and C. M. Yeates, *J. Geophys. Res.* **80** (1975) 4359.
15. J. G. Luhmann, C. T. Russell and N. A. Tsyganenko, *J. Geophys. Res.* **103** (1998) 9113.
16. W.-H. Ip, *Adv. Space Res.* **19** (1997) 1615.
17. D. C. Delcourt, K. Seki, N. Terada and Y. Miyoshi, *Annales Geophys.* **24** (2006).
18. D. C. Delcourt, K. Seki, N. Terada and Y. Miyoshi, *Ann. Geophys.* **23** (2005) 3389.
19. G. L. Siscoe, N. F. Ness and C. M. Yeates, Substorms on Mercury?, *J. Geophys. Res.* **80** (1975) 4359.
20. B. H. Mauk, *J. Geophys. Res.* **91** (1986) 15,423.
21. J. M. Quinn and C. E. McIlwain, *J. Geophys. Res.* **84** (1979) 7365.
22. M. Hirahara, T. Mukai, T. Nagai, N. Kaya, H. Hayakawa and H. Fukunishi, *J. Geophys. Res.* **101** (1996) 7749.
23. D. G. Mitchell, P. C. Brandt, E. C. Roelof, D. C. Hamilton, K. C. Retterer and S. Mende, *Space Sci. Rev.* **109** (2003) 63.
24. S. C. Chapman, Properties of single-particle dynamics in a parabolic magnetic reversal with general time dependence, *J. Geophys. Res.* **99** (1994) 5977.

THE DAYSIDE MAGNETOSPHERE OF MERCURY

S. MASSETTI, S. ORSINI, E. DE ANGELIS, V. MANGANO,
A. MILILLO and A. MURA
Istituto di Fisica dello Spazio Interplanetario, IFSI-INAF, Roma, Italy
stefano.massetti@ifsi-roma.inaf.it

Many efforts are currently in progress to depict a realistic model of the Hermean magnetosphere, also in the frame of the researches that will be performed by the next space missions *Messenger* and *BepiColombo*. Although the magnetosphere of this planet can be roughly approximated by scaling the Earth's one (about 1:6.9), several differences are expected due to both the lack of an ionosphere, which rise questions concerning the closure of the magnetospheric currents, and the likely absence of a stable ring current. A remarkable feature, outlined by the models, is represented by the wide area accessible by the solar wind plasma on the Mercury's dayside, due to the broad footprint of the magnetospheric cusp regions. Moreover, the absence of a dense atmosphere/ionosphere implies that a large faction of the impinging plasma can reach the planet's surface. We present the results obtained by comparing two modelization of the Mercury's magnetosphere by means of *ad hoc* modified *Tsyganenko* and *Toffoletto–Hill* models. The analysis is focused on the dayside magnetospheric configurations and plasma entry patterns, derived by taking into account the effect of the magnetic reconnection with the interplanetary magnetic field. The role played by the radial IMF component (B_x) is also discussed.

1. Introduction

We present preliminary results obtained by comparing two modelization of the Mercury's magnetosphere by means of *ad hoc* modified *Tsyganenko* (T96) and *Toffoletto–Hill* (TH93) models, hereafter referred as T96* and TH93*, respectively. The analysis is focused on the dayside magnetospheric configurations and plasma entry patterns, derived by taking into account the effect of the magnetic reconnection with the interplanetary magnetic field (IMF). Figure 1 sketches the reconnection geometry on the dayside magnetopause, and the related key parameters. We applied the stress balance condition to the *de Hoffmann–Teller* (HT) reference frame, moving along the magnetopause at the rotational discontinuity produced by the magnetic reconnection between the planet's magnetic field and the interplanetary magnetic field (e.g., Refs. 1 and 2). Equations (1)–(5) permit us

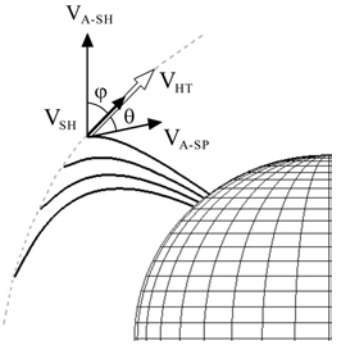


Fig. 1. Sketch of a reconnected field line on the dayside magnetopause, illustrating the key parameters associated with the *de Hoffmann–Teller* (HT) reference frame: $V_{\text{A-SH}}$ and $V_{\text{A-SP}}$ are the Alfvénic speed (tangent to the local magnetic field) on the magnetosheath and the magnetospheric side, respectively; φ and θ are the angles formed by $V_{\text{A-SH}}$ and $V_{\text{A-SP}}$ with the magnetopause (dashed grey line); V_{SH} is the magnetosheath flow velocity (black arrow tangent to magnetopause); and V_{HT} is the HT reference frame velocity (white arrow).

to determine the minimum (V_{min}), peak (V_{p}), and maximum (V_{max}) velocity of the ions crossing the magnetopause along open field lines (please, refer to Ref. 3 for a more detailed discussion):

$$V_{\text{SH}} = -V_{\text{A-SH}} \cos \varphi, \tag{1}$$

$$V_{\text{HT}} = V_{\text{SH}} + V_{\text{A-SH}} \cos \varphi, \tag{2}$$

$$V_{\text{min}} = V_{\text{HT}} \cos \theta, \tag{3}$$

$$V_{\text{p}} = V_{\text{HT}} \cos \theta + V_{\text{A-SP}}, \tag{4}$$

$$V_{\text{max}} = V_{\text{HT}} \cos \theta + V_{\text{A-SP}} + V_{\text{th}}. \tag{5}$$

The acceleration of the injected plasma critically depends on the ratio between the Alfvénic speed on the magnetospheric ($V_{\text{A-SP}}$) and magnetosheath ($V_{\text{A-SH}}$) sides. In this work, we assumed $V_{\text{A-SP}} = 5\ V_{\text{A-SH}}$, while the $V_{\text{A-SH}}$ speed was derived from the model. As in a previous analysis,[3] the magnetosheath plasma velocity V_{SH}, density N_{SH}, and temperature T_{SH}, close to the magnetopause, were derived from the work of Spreiter *et al.*[4] At the orbit of Mercury, the presence of a significant B_x component causes the plasma to precipitate mainly in the dayside hemisphere that is directly connected to the IMF. This is due to the reduced plasma penetration through the open magnetic field lines that are draped in the antisunward direction, forming the polar cap regions. Our results suggest that the magnetic reconnection geometry causes the focusing of the plasma injection within confined

areas of the magnetosphere. However, finite gyro-radius effects, grad-**B** and $\mathbf{E} \times \mathbf{B}$ drifts are expected to spread out the plasma patterns as the injected ions precipitate through the magnetosphere (e.g., Refs. 5 and 6).

2. Open Field Regions

In Fig. 2, we report the fraction of the Mercury's surface that is crossed by open magnetic field lines, mapping on the dayside magnetopause, as a function of the solar wind pressure ($P_{\rm dyn}$) and of the IMF B_z component. On the *left panels*, the results we had obtained by means of the T96* model[3] are shown: *upper panel*, IMF $B = (0, 0, B_z)$ and $P_{\rm dyn} = 20$ nPa; *lower panel*, IMF $B = (0, 0, -10)$ nT. On the *right panels*, we show the results derived with the present model (TH93*) using the IMF geometries reported on the legends: B_z-dominated (black lines) and three configurations with a fixed B_x/B_y ratio along the nominal Parker's spiral direction at Mercury (gray lines).

In the T96* case, the open area is equally distributed in both northern and southern hemispheres. In the TH93* case, the area is completely, or

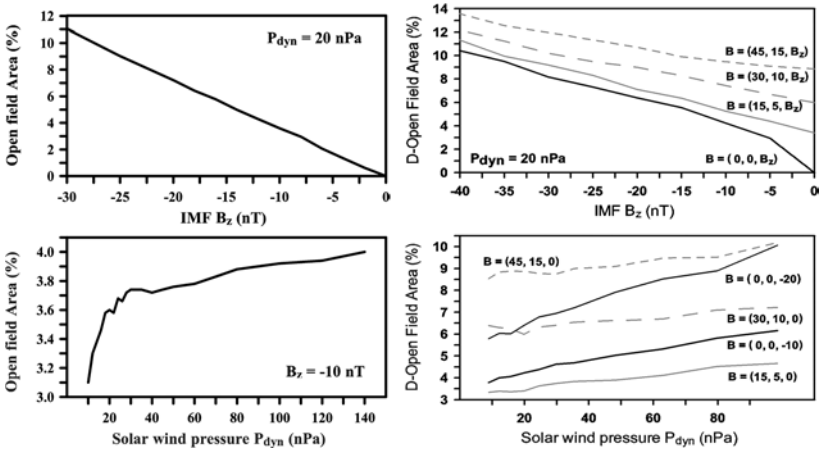


Fig. 2. Fraction of Mercury's surface that is characterized by open magnetic field lines, mapping on the dayside magnetopause, as calculated by means of both the T96* (*left panels*) and the TH93* (*right panels*) models. *Upper panels*: open field area as a function of the IMF B_z component, for fixed solar wind pressure ($P_{\rm dyn} = 20$ nPa): the black lines mark the IMF $B = (0, 0, B_z)$ configurations, while the gray lines show the results obtained with the IMF aligned along the nominal Parker's spiral direction (TH93* only, since T96* is independent from B_x). *Lower panel*: open field area as a function of the solar wind pressure, derived with the IMF configurations indicated in the legends.

mainly, concentrated in the southern (northern) hemisphere for B_x positive (negative). The results obtained with the two models are similar in the case of fixed solar wind pressure ($P_{\rm dyn} = 20\,{\rm nPa}$, upper panels). For B_x/B_y fixed along the Parker's spiral direction (TH93* only), the open field area increases progressively as a function of the total IMF strength. By allowing the interconnection with the IMF B_x component, the TH93* model causes a large fraction of the planetary surface to be exposed to the direct precipitation of magnetosheath plasma, even when $B_z \sim 0$, as previously discussed by other authors (e.g., Ref. 7).

3. Plasma Penetration into Mercury's Dayside Magnetosphere

Figure 3 illustrates the results obtained by applying the stress balance condition to the Mercury's dayside magnetosphere, approximated by the TH93* model. Calculations are made with a reference solar wind speed with $V_{\rm SW} = 430\,{\rm km/s}$ and density $N_{\rm SW} = 76\,{\rm cm}^{-3}$. In the test runs here presented, three different interplanetary magnetic field configurations were used, namely: IMF $B = (30, 10, 0)\,{\rm nT}$ (*top*), IMF $B = (30, 10, 10)\,{\rm nT}$ (*middle*), and IMF $B = (30, 10, -10)\,{\rm nT}$ (*bottom*), with a B_x/B_y ratio close to the one expected at Mercury's orbit (about 20° with respect to the radial direction). The *left panels* show the foot-points of the magnetic field lines on the surface of Mercury: closed (blue dots), open on the dayside (cusp regions, red dots) and open tailward (polar caps, gray dots). Contrary to the T96* model, the TH93* produces patterns in agreement with the Kallio and Janhunen[5] numerical simulations, thanks to the hemispherical asymmetry introduced by the interconnection with the IMF B_x component. The interconnection with the IMF radial component causes the magnetosphere to be "open" also for zero, or positive, B_z component (Fig. 3, *top* and *middle panels*). When B_x is positive, the open hemisphere is the southern one, where we can see the footprint of the local cusp. In this case, the northern cusp becomes visible only when a significant IMF B_z component is applied, but its extension is in any case smaller than the southern one. This is due to the fact that most of the open field lines in the northern hemisphere cross the magnetopause on the nightside side, where the magnetosheath flux is strongly antisunward and few particles can to precipitate along the convected field lines. When the sign of B_x is negative, the situation reverses, and the northern hemisphere is the most exposed to the plasma injection. The sign of the B_y component is responsible of the dawn–dusk asymmetry,

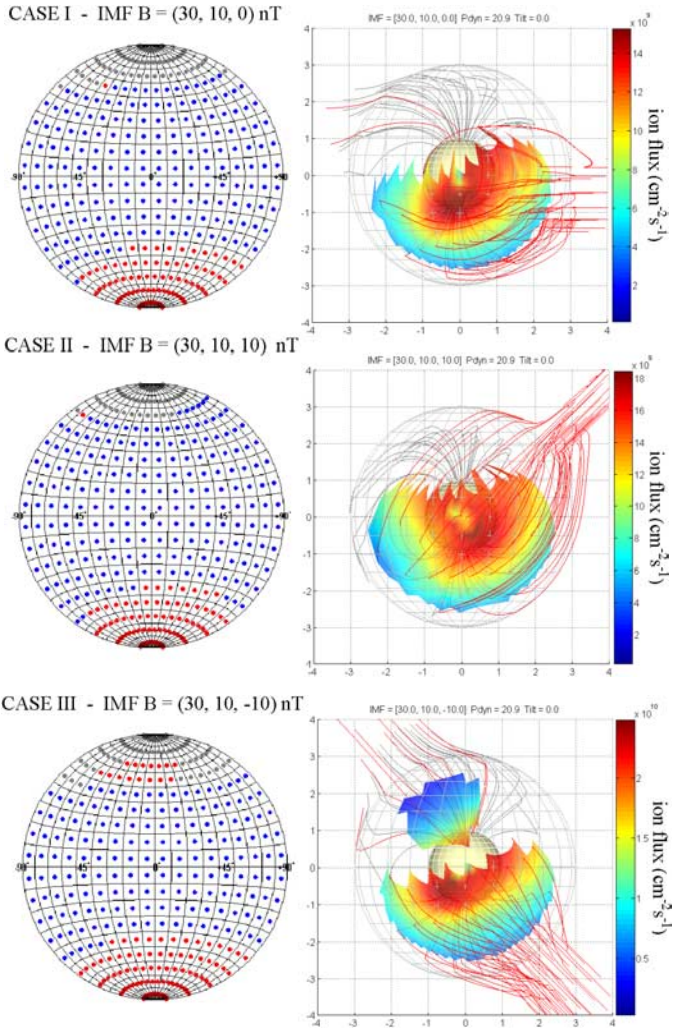


Fig. 3. Preliminary results obtained by means of the TH93* model, with $V_{SW} = 430\,\mathrm{km/s}$, $N_{SW} = 76\,\mathrm{cm}^{-3}$, under three different interplanetary magnetic field configurations: IMF $B = (30, 10, 0)\,\mathrm{nT}$ (*top*), IMF $B = (30, 10, 10)\,\mathrm{nT}$ (*middle*) and IMF $B = (30, 10, -10)\,\mathrm{nT}$ (*bottom*). *Left panels*: foot-point of closed (blue dots), open on the dayside (cusps, red dots) and open tailward (polar cap, gray dots) magnetic field lines. *Right panels*: estimation of the magnetosheath plasma entry through the magnetopause (gray mesh), along the dayside open field lines (red lines). The flux is calculated by taking into account the magnetic reconnection effects (Eqs. (1)–(5)), and by assuming that 50% of the incoming ions are reflected back at the magnetopause boundary. The fluxes shown have to be compared with the unperturbed solar wind flux of about $3.3 \times 10^9\,\mathrm{cm}^{-2}/\mathrm{s}^{-1}$. It can be noted that the highest fluxes are located near the subsolar region, where the magnetosheath plasma is estimated to be about 3.5 higher than the upstream solar wind.

with the cusp region shifted duskward (dawnward) in the northern (southern) hemisphere for positive values. The same effect is observable on the Earth, at high- and low-altitudes, from both particles and auroral data. The *right panels* of Fig. 3 sketch the dayside magnetosphere of the planet, with the magnetic field lines marked as above (red = open on the dayside, gray = open on the tail). The color coded surface represents the estimated magnetosheath particle flux injected along the dayside open field lines, and mapped over the corresponding area of the magnetopause (gray mesh). The fluxes are calculated by taking into account the energy gain produced by the superimposition of the field line dragging caused by the magnetosheath flow, and by the field line tension (as described by Eqs. (1)–(5)). Moreover, it was assumed that 50% of the ions are reflected back by the magnetopause boundary, that is, a transmission factor of 0.5, which is a typical value in the case of the Earth (e.g., Ref. 8). To have an idea of the fraction of ions that can actually reach the surface, the fluxes shown in Fig. 3 must be multiplied by a factor of about 0.2, which is equivalent to a loss-cone of 35°. In addition, finite gyro-radius effects, grad-$\mathbf{B}$ and $\mathbf{E} \times \mathbf{B}$ drifts are expected to spread out, to some extent, the plasma precipitation patterns depicted in Fig. 3, as the ions precipitate throughout the magnetosphere to the planet surface (see e.g., Refs. 5 and 6). A big unknown in this context is the poor knowledge of the intensity and geometry of the electric field, which can strongly affect the displacement of the plasma away from the injection regions. Figure 3 shows that the strongest fluxes are located near the sub-solar region of the magnetopause, where the magnetosheath plasma should be about 3.5 higher than the free solar wind (as expect on the base of Spreiter *et al.*[4] gasdynamic approximation). The energy distribution (not shown) typically peaks on the equatorial side of the open field regions, and then decreases gradually while moving toward the high latitudes, as observed on the Earth's LLBL, cusp and mantle magnetospheric regions.[3]

4. Conclusions

We discussed some results from a modelization of the plasma entry on the dayside magnetosphere of Mercury. This analysis was performed by means of the TH93 model that we adapted *ad hoc* to reproduce the key features of the planet. The results confirm our previous findings based on a modified version of the Tsyganenko (T96) model,[3] but display also other

new aspects. In particular, the interconnection of the Mercury magnetosphere with the IMF B_x component produces a hemispheric asymmetry, as discussed by other authors (e.g., Refs. 5 and 7). The magnetosheath ions crossing the dayside magnetopause are subjected to acceleration at the rotational discontinuity, generated by the merging of the IMF with the planetary magnetic field. This mechanism can generate particles with energies up to several tens of keV, and fluxes of the order of $10^{10}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at the magnetopause boundary, about 10% of which could reach the surface of the planet (depending on the loss-cone).

The exact determination of these values is affected by several poorly known parameters as the $V_{\text{A-SP}}/V_{\text{A-SH}}$ ratio linked to the relative densities on both sides of the rotational discontinuity at the magnetopause, and the efficiency of the magnetic merging at the dayside of Mercury. The latter plays a very important role: in fact, the rate of the magnetic merging, which can be either "bursty" or "steady", can heavily affect the plasma distribution inside the magnetosheath. A high rate of dayside magnetic merging would lead to a fast depletion of the plasma density inside the magnetosheath, causing a sensible deviation from the Spreiter's gasdynamic approximation. Under steady reconnection, the wide cusp region of the B_x-interconnected hemisphere could be almost always "open", and the magnetosheath density would be locally close to the upstream solar wind value. In this case, the fluxes calculated in this work would rapidly decrease, approaching the values of the unperturbed solar wind at Mercury's orbit.

Acknowledgments

Authors wish to thank N. Tsyganenko, F. Toffoletto and T. Hill, together with the NASA/Goddard Space Flight Center (NSSDC) for making the magnetospheric T96, and TH93, models available to the scientific community.

References

1. S. W. H. Cowley and C. J. Owen, *Planet. Space Sci.* **37** (1989) 1461–1475.
2. M. Lockwood and M. F. Smith, *J. Geophys. Res.* **99** (1994) 8531–8553.
3. Massetti *et al.*, *Icarus* **166** (2003) 229–237.
4. J. R. Spreiter, A. L. Summers and A. Y. Alksne, *Planet. Space Sci.* **14** (1966) 223–253.

5. Kallio and Janhunen, *Geophys. Res. Lett.* **3** (2003), doi:10.1029/2003 GL017842.
6. A. Mura, S. Orsini, A. Milillo, D. Delcourt and S. Massetti, *Icarus* **177** (2005) 305–319.
7. M. Sarantos, P. H. Reiff, T. H. Hill, R. M. Killen and A. L. Urquhart, *Planet. Space Sci.* **49** (2001) 1629–1635.
8. M. Lockwood, *Ann. Geophysicae* **15** (1997) 1501–1514.

NEUTRAL ATOM EMISSION FROM MERCURY

A. MURA[†,*], S. ORSINI[†], A. MILILLO[†], D. DELCOURT[§], A. M. DI LELLIS[‡], E. DE ANGELIS[†] and S. MASSETTI[†]
[†]*INAF-IFSI, Roma, Italy*
[‡]*AMDL Srl., Roma, Italy*
[§]*CETP-CNRS, Saint Maurice des Fossés, France*
[*]*mura@ifsi.rm.cnr.it*

Neutral Atom emission from Mercury Magnetosphere is discussed in this study. In particular, we consider those neutrals whose emission is directly related to the circulation of energetic ion of solar wind origin, via both charge-exchange and ion-sputtering. The environment of Mercury, in fact, is characterized by a weak magnetic field; thus, cusp regions are extremely large if compared to the Earth's ones, and intense ion fluxes are expected there. Spatial and energy distributions of ions and neutrals, and energy-integrated simulated ENA images, are obtained by means of a single-particle 3D simulation. The feasibility of neutral atom detection and imaging in the Hermean environment is also discussed here: simulated neutral atom images are investigated in the frame of the Neutral Particle Analyser-Ion Spectrometer (SERENA NPA-IS) experiment, on board the ESA mission BepiColombo/MPO.

1. Introduction

Since Mariner-10 observations of Mercury revealed an intrinsic magnetic field,[1] several studies have been focused on the solar wind/magnetosphere/exosphere/surface interaction at Mercury.[2–4] In particular, numerical simulations[5–7] and theoretical studies[8] suggest that the solar wind may directly have access to the inner regions through magnetospheric cusps and reach the surface of Mercury. This may results both in charge-exchange (CE) between solar wind ions and planetary neutrals, and in surface ion-sputtering process, which is one of the sources of the exospheric populations.

The CE occurs when an energetic ion collides with an exospheric neutral. The ion may grab an electron and become an energetic neutral atom (ENA), which retains both the energy and the direction of the incoming ion. No more affected by electro-magnetic forces, these ENAs travel on straight lines just like photons; ENA imaging is hence a common way to indicate the detection and analysis of ENA fluxes. The ENA detection not only gives information about the effectiveness of CE as a loss process of

magnetospheric plasma, but also is a powerful tool to obtain information about 3-dimensional (3D) exospheric and plasma properties.[9,10] In fact, even if this information needs a deconvolution process to be unfolded from 2-dimensional (2D) ENA images, ENA imaging has the advantage to give a *global* and *instantaneous* information from a remote vantage point. Moreover, since ENAs retain approximately all the energy of the colliding ions, their energy distribution gives information about energy distribution of the magnetospheric plasma.

The result of ion-sputtering is, in most cases, a neutral with a few eV energy. However, a fraction of the sputtered neutrals may have enough energy to travel on straight lines and escape from planetary gravity: in this paper, we will refer to those neutrals as ENAs too. Detection and mass-analysis of this ENA signal will give information about: effectiveness of ion-sputtering as surface loss process and exospheric source; relative surface composition; map of ion-impact on Mercury's surface.[11]

The purpose of the present study is the estimation of the neutral flux generated by CE and ion-sputtering and the discussion of the feasibility of the neutral-particle imaging of the plasma circulation and the surface. Here we use a single-particle model to reconstruct the spatial distribution and the energy spectrum of H^+ ions. We will show that this plasma is able to generate a noticeable amount of neutral particles, both via CE with the neutral exosphere and via ion-sputtering processes on its surface. Moreover, we will show that the neutral flux coming from both those processes is extremely sensible to the external conditions, thus leading to a "neutral imaging" investigation technique for Mercury's magnetosphere and surface, through directional neutral atoms detection.

2. Proton Circulation Model

The magnetic and the electric field models developed in this study originate, respectively, from Tsyganenko T96[12] and Volland[13] models, which refer to the Earth. They have been modified considering the different intrinsic magnetic field of Mercury and the different properties of the solar wind at the orbit of Mercury (0.31–0.47 AU).[11,14] Here we have assumed an unperturbed solar wind density of $52\,\mathrm{cm}^{-3}$, a solar wind velocity of 430 km/s and a dynamic pressure (nmv^2) of 16 nPa.[11]

The magnetic field model accepts the z and the y components of IMF as independent inputs. Even if the magnetic reconnection at the dayside magnetopause is essentially driven by the IMF B_z component,[11]

the x component should play a role in the magnetosphere solar wind coupling,[15,16] because of the different orientation of the Parker's spiral at Mercury with respect to the Earth's case. In principle, one could expect to have more proton penetration with a B_x-interconnected magnetosphere.[17] Unfortunately, the present magnetic field model does not allow us to explore such configurations. In this paper, we consider only the particular configuration $B_{\mathrm{IMF}} = (0, 0, -20)\,\mathrm{nT}$, which allows reconnection in the dayside magnetosphere and thus the entrance of solar wind protons. Different configurations of B_{IMF} have been discussed in a previous study.[7]

The electric potential in any point of the space is obtained assuming that $E \cdot B$ is zero, and hence following the B field line as far as it reaches any region where V is known. As a zero-order hypothesis, here we have assumed that the potential at the surface of Mercury is similar to Volland[13] potential at the ionosphere of the Earth. The absence of an ionosphere at Mercury may make it difficult to comprehend whether and how the field-aligned currents, if present, can close the electric circuit between the planet and the cross-tail potential drop (PD). The argument is still widely debated, even if some evidences of the presence of field-aligned current at Mercury exist.[3] Using the given solar wind parameters, and taking into account the different size of Mercury's magnetospheric cavity with respect to the Earth's one, we can estimate a cross-polar cap PD between 10 and 100 kV.[14] To explore the behavior of the model under different external conditions, here we use PD = 10 kV as "reference/mean" condition,[18,19] and the values of 100 and 1 kV to evaluate the importance of E in the reconstruction of the proton circulation.

The proton distribution in the magnetosphere of Mercury has been obtained by using a Monte–Carlo single-particle model.[7] The trajectories of about 2×10^6 test-particles have been computed using classical equation of motion[14] (including electro-magnetic forces and gravity), and accumulated on a 5-dimensional (5D) grid (space, energy E, and pitch angle PA). Each trajectory begins at the magnetopause of Mercury, with random initial position and velocity; a weight w is associated to the test-particle, taking into account the proton flux where this particle "was born" (Φ_0). The distribution of protons at the MP has been taken from Massetti;[11] in this simulation, unlike a previous paper,[7] we have started test particle trajectories from both north and south magnetopause. The test-particle trajectory ends at the surface of the planet or when it reaches an IMF field line not connected with Mercury's surface. Photoionization is not considered; CE collisions between H^+ and exospheric neutrals are taken into account by

decreasing w:

$$\frac{dw}{dt} = w(t)|\mathbf{v}| \sum_i \sigma_i n_i \,, \qquad w(0) \propto \Phi_0, \tag{1}$$

where v is the proton velocity, n_i the density of the ith neutral population, and σ_i is the relative CE cross-section, i.e., the inverse of ion mean free path for unitary neutral density.

Figure 1 shows a generic H^+ simulated distribution in the north/dayside part of the xz plane (in Mercury Solar Ecliptic coordinates (MSE): $x_{\rm MSE}$ towards the sun, $y_{\rm MSE}$ in the plane of ecliptic, perpendicular to $x_{\rm MSE}$, $z_{\rm MSE}$ perpendicular to $x_{\rm MSE}$ and $y_{\rm MSE}$). Left panel refers to low-energy protons (100 eV–1 keV); right panel refers to high-energy protons (1–10 keV); red circle is the planetary surface and blue lines are magnetic field lines. According to our model, both high and low-energy protons precipitate along magnetic field lines; high-energy protons precipitate at lower latitudes with respect to low-energy ones, since, according to Massetti *et al.*[8] the mean energy of protons over the MP decreases with the latitude. Some of the precipitating protons are reflected by the increasing magnetic field; the others experience CE (1%) or reach the planetary surface (10%). During this motion, protons are drifted northward by the $E \times B$ drift, and westward by the grad-B drift. In this way, a second and a third population are originated. The relative importance of these two drifts depends on fields' strength and proton energy. The second population becomes stronger as the electric field increases; the contrary happens to the third population.

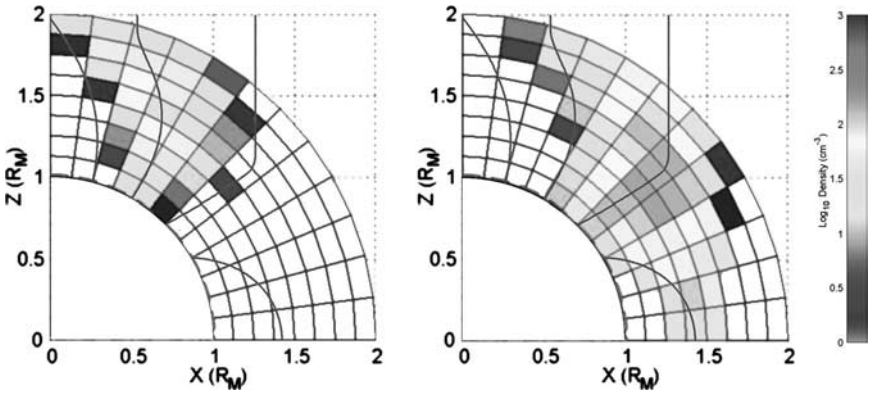


Fig. 1. H^+ density over the x–z plane. Left panel: from 100 eV to 1 keV; right panel: from 1 to 10 keV. Magnetic field lines are in blue.

Eventually, some of the westward-drifting protons precipitate onto the dawn and night surface. Finally, according to this model, the H^+ circulation is negligible on the night side.

3. ENA Simulations

In the present model, we consider only ion-sputtering and CE processes. Ion-sputtering is the result of the impinging of an ion of mass m_1 upon the planetary surface; if the impact energy (E_i) is high enough, a new particle (m_2), mostly neutral, may be released from the surface. Its ejection energy (E_e) distribution function $f_\mathrm{S}(E_\mathrm{e})$ peaks at few eV:[20–22]

$$f_\mathrm{S}(E_\mathrm{e}, E_\mathrm{i}') \propto \frac{E_\mathrm{e}}{(E_\mathrm{e}+E_\mathrm{b})^3}\left[1-\left(\frac{E_\mathrm{e}+E_\mathrm{b}}{E_\mathrm{i}}\frac{(m_1+m_2)^2}{4m_1m_2}\right)^{1/2}\right]. \tag{2}$$

Here, E_b is the surface binding energy of the atomic species extracted. The consequent neutral differential flux is

$$\frac{d\Phi}{dE_\mathrm{e}} = YR\int_{E_\mathrm{min}}^{E_\mathrm{max}} \frac{d\Phi_{\mathrm{H}^+}}{dE_\mathrm{i}} f_\mathrm{S}(E_\mathrm{e}, E_\mathrm{i})\, dE_\mathrm{i}\,, \tag{3}$$

where Y and R are the yield and the surface relative abundance of the atomic species considered, and Φ_{H^+} is the ion flux. In this simulation, E_min and E_max are the same of the H^+ accumulation grid, i.e., 100 eV and 10 keV, respectively. In this study, we have concentrated our attention to Oxygen only, observed in the exosphere of Mercury by Mariner-10.[23] For our speculations, since the surface composition is not known, we assume a uniform O surface abundance of 50% and a rough value for Y of 0.1;[24] the binding energy considered here is 3.5 eV.

The CE process occurs when an energetic ion collides with an exospheric neutral atom; the result is an ENA, which retains both the energy and the direction of the incoming ion.[25] To simulate an ENA image, we start from the ENA production rate for unitary length in any point P of the grid:

$$\frac{d\Phi_\mathrm{ENA}(P)}{ds} = \frac{n_{\mathrm{H}^+}}{w}\frac{dw}{dt}, \tag{4}$$

$$\Phi_\mathrm{ENA}(Q) = \frac{d\Phi_\mathrm{ENA}(P)}{ds}\Delta V f_\alpha(\hat{v})\,\frac{1}{|\overline{QP}|^2}. \tag{5}$$

Equation (5) gives the total flux escaping from a cell located in P and detected in Q, where ΔV is the volume of the grid cell in P; $f_\alpha(\hat{v})$ is the

3D angular distribution function, $\hat{v}$ is the direction from P to Q. It is then sufficient to integrate over all grid cells along the *line-of-sight* ($\hat{v}$) to obtain the ENA image at Q. In this simulation we include only H, H_2, He, and O, and we assume the exosphere composition derived by Wurz and Lammer.[24] They have used a Monte–Carlo model guided by some input parameters: surface density (in m^{-2}, H: 2.3×10^7; H_2: 2.6×10^{13}; He: 6×10^9; O: 4.4×10^{10}), temperature (for thermal desorption, in K, H: 420, H_2: 540, He: 575) and binding energy (for ion-sputtering, O: 3.4 eV). In particular, H_2 gives the most significant contribution to ENA generation. Since H_2 has not been directly observed, its exospheric vertical profile has been intentionally depleted (in accordance with authors' suggestions); the factor adopted here (0.01) is intentionally very low, to avoid any possible overestimation of ENA fluxes; the corrected surface density is hence $2.6 \times 10^{11}\,m^{-2}$.

4. ELENA Sensor

The ELENA sensor, part of the SERENA experiment,[11,26,27] is a time-of-flight (ToF) detector, with energy, mass and azimuth resolution. Particles enter the ToF chamber through the holes of a couple of masks; one of those masks shutters with respect to the other, driven by a piezo-electric engine at frequencies up to 100 kHz.

The holes on the masks let the particles enter only during a defined time-window (*tw*) (approximately one-tenth of the duty-cycle time). During the subsequent flight toward the micro-channel plate (MCP), ions are removed by high-voltage deflectors. The delay between the center of the opening *tw* and the Stop event on the MCP is the measured ToF. The pulse height signal on the MCP gives also information about particle energy (E_p), mass is calculated by knowing E_p and ToF. The position of the event gives the particle direction; the viewing geometry is explained in Fig. 2. Presently, a relative error on energy down to 35% could by hypothesized, and up to six mass channels could be identified. The minimum detectable energy depends on the MCP efficiency, which strongly decreases at low energies, and it is of the order of 20 eV (see, e.g., Fig. 3 in Ref. 28). In principle, there is no limit

Table 1. Performances of ELENA instrument.

Aperture: 1 cm^2	FoV: $-30°$–$46°$ cross track, $\pm 1°$ along tr.
Angular resolution: $2° \times 2°$	ToF resolution $\Delta t/t$: $\sim 8\%$
Energy range: 20 eV–5 keV	Energy resolution $\Delta E/E$: $\sim 35\%$
Mass channels: up to six	Pointing: ∼Nadir,–Z MPO axis+8° tilt
G. Factor (w/o MCP): $2.5 \times 10^{-4}\,cm^2\,sr$	Piezo freq.: up to 100 kHz

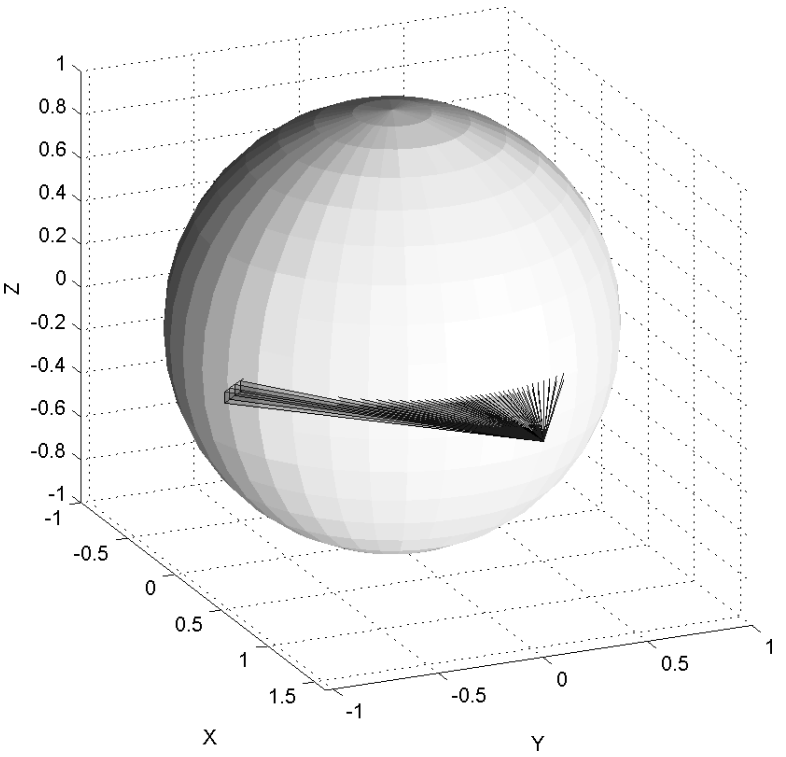


Fig. 2. ELENA viewing geometry at MPO apoherm; the intrinsic field of view of the 38 sectors are indicated. ELENA is at the white circle.

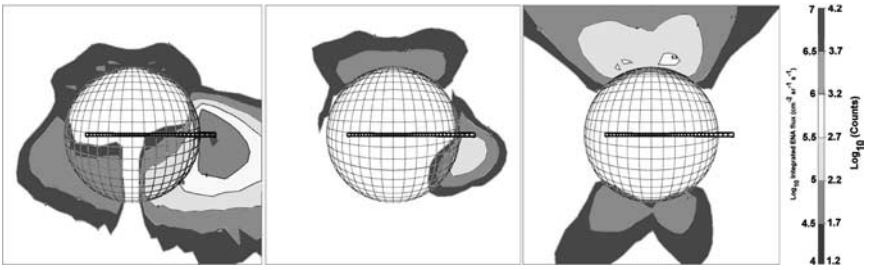


Fig. 3. Fish-eye projection of H-ENA simulated images, as seen from (−1.5, 0, 0.4) R_{M} MSE. The color scale is coded to H-ENA flux, integrated from 100 eV to 10 keV (left scale) and to one-minute count rates (right scale, with geometrical factor of ELENA — but no MCP efficiency — included.). Black grids show ELENA FoV (squares are instrument sectors: $2 \times 2^{\circ}$). Different panels refer to different PD values; from the left: 1, 10, and 100 kV.

on the maximum detectable energy; however, high-energy particles could hit the MCP before the shuttering mask has been closed, and could be lost, being mixed with photons.

The shuttering mask technique does not introduce "disturbing" detector elements, which may affect particle trajectory or energy. This is particularly important in this case, since the detection of ion-sputtered ENAs with energies down to few tens of eV is crucial. Furthermore, this technique has a really good photon rejection, important for measurement in the Hermean environment.

5. ENA Imaging

In this section we present our results in terms of simulated ENA images, and we discuss them in the frame of the MPO mission. The MPO has a polar orbit, with a 400 km periherm and a 1,500 km apoherm; ELENA FoV is always nadir-pointing, and most sectors look toward the planetary surface. We will discuss about CE H-ENAs and ion-sputtering O-ENAs separately, since instrumental simulations show that the related signals can be easily distinguished.

Charge-exchange: According to our proton circulation model, and assuming the exospheric profiles introduced in Sec. 3, CE process occurs mostly in the dayside and dawnside regions close to the planetary surface (up to altitudes of hundreds of km). In a medium PD configuration, the H-ENA production rate for unitary length reaches values up to 10 (cm^2/s/sr/m), close to the dayside planetary surface. This value, if integrated over a long line-of-sight, could lead to high values of H-ENA flux. To facilitate the detection of H-ENAs, the ELENA central axis is tilted, with respect to the s/c nadir axis, by 8°. In this way at least three sectors point outward of the planet limb if the s/c is in a ±15° orbital arc centered at the apoherm, while one sector points outward in a ±90° arc. The H-ENA signal is lower in sectors looking towards the planet, because the integration path is shorter. The s/c apoherm position will move, in longitude, during the MPO mission, thus allowing different, optimal vantage points. Here we examine two possible examples.

Figure 3 shows H-ENA simulated images, as seen from the nightside; the vantage point is in $(-1.5, 0, 0.4)$ R_M MSE. For simplicity, H-ENA fluxes are integrated over all energies, even if ELENA can resolve particle energy. In principle, from this vantage point it is possible to detect the H-ENA signal generated by protons that are precipitating into the cusp regions or

circulating over the North pole, and the H-ENA signal originated by protons that are bouncing and drifting westward. The first produce an intense H-ENA signal in the upper part of the three H-ENA pictures. The second population of protons reaches low altitude at the mirroring points, where exospheric densities are higher: even if, here, H^+ flux is generally lower than H^+ flux in the dayside, they may produce an intense ENA signal that can be seen on the right side of the pictures. This westward circulation is generated by B-grad drift; this drift prevails over the $E \times B$ drift if the PD is low. For this reason, the second ENA peak is more evident for low values of PD.

These considerations apply to a generic ENA sensor; as far as it concerns ELENA, the proper FoV is shown by the black bold grid in the center of Fig. 3, where each square represents an instrument sector ($2^\circ \times 2^\circ$). The color-scale, in this case, gives also simulated counts, integrated over 1 min. Thanks to s/c motion it is possible to investigate this region from different vantage points, but it is not straightforward how to deconvolve these "images" to obtain plasma properties. In principle, the problem is similar to that of obtaining information on a 3D plasma distribution from a single, typical 2D ENA image, since the lack of information is the same; however, this problem will be faced in a future study.

As far as it regards a single ELENA 1D "image", these simulations show that it will be possible to monitor the westward-drifting H^+ population, by looking at the instantaneous H-ENA signal in the last three sectors. An increase or decrease of this signal could, in principle, be correlated to external variations (such as potential drop, in this example); moreover, the $2^\circ \times 2^\circ$ instrument resolution results, at the surface-closest point, in a 200 km altitude resolution, which allows a rough reconstruction of the vertical profile of H^+ flux. Another information that could be obtained from ELENA data in this configuration is how many protons circulate in the nightside, by means of those sectors that look towards the planet.

Figure 4 is similar to Fig. 3, but the vantage point here is in the dawn region ($0.1, -1.6, 0\, R_M$). In this configuration, ELENA looks directly towards the westward drifting population. Hence, from this vantage point it is possible to monitor instantaneously the H^+ population drifting westward (central sectors), as well as the precipitating one (right sectors) since, in most cases (*low* and *medium* PD), all sectors have noticeable count rates. The same considerations made about Fig. 2, concerning PD influence on H-ENA images, apply also here. In particular, if the PD is high (right panel), protons pass over the poles at higher altitudes, where exosphere density is lower, and CE process is less relevant.

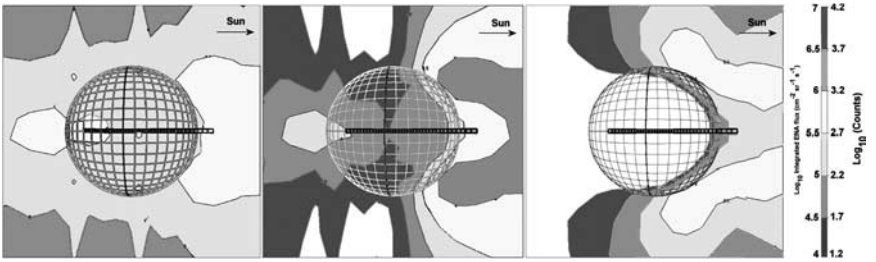


Fig. 4. Same as Fig. 2, but vantage point is at (0.1, −1.6, 0) R_M MSE. Sun is on the right.

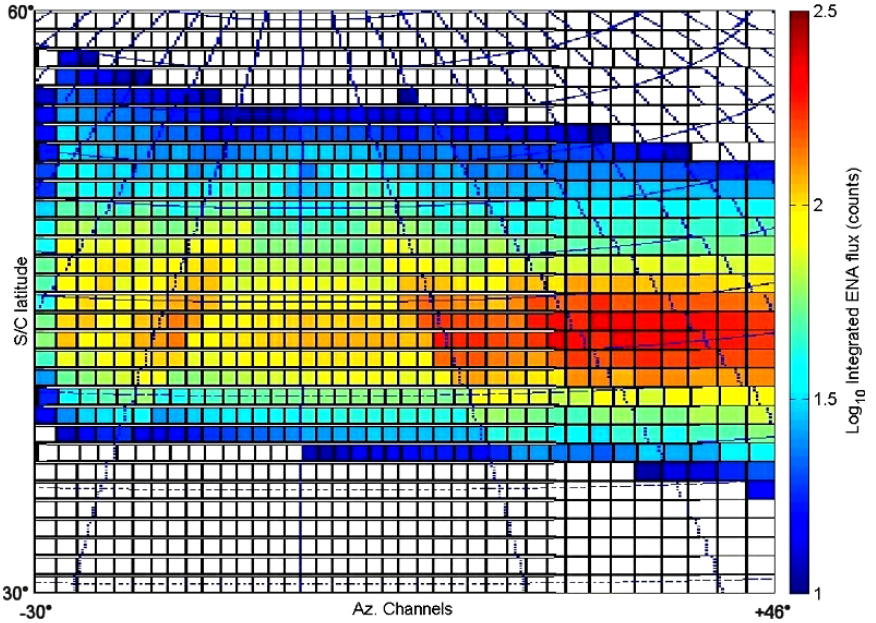


Fig. 5. 2D image of surface oxygen emission, obtained by the superimposition of 31 ELENA simulated measurements. Color is coded according to ELENA 30 s count rates, integrated over all energies above oxygen escape energy. Geometrical factor of ELENA (but no MCP efficiency) is included.

Ion-sputtering: The detection of ion-sputtering ENAs deals with several scientific objects of BC/MPO mission,[11] such as the study of particle loss-rate, surface emissivity, ion-sputtering process, and proton precipitation. Figure 5 is an example of ELENA simulated O-ENA signal. The 2D image of surface Oxygen emission is obtained with the help of s/c motion,

which is orbiting from 30 to 60° N at 12:00 MLT (the virtual periherm is at 0°, 12:00 MLT). The intense proton precipitation in the dayside, northern hemisphere results in an intense O-ENA production in the center of the global 2D image. In general, it is possible to have a global view of proton precipitation by means of ion-sputtered directional ENAs, with surface spatial resolution between 15 and 50 km, depending on s/c altitude. Some studies[7,8] have estimated that, under different IMF conditions, the configuration of the Hermean magnetosphere changes so that the area of high proton precipitation (and subsequent neutral release) moves.

According to our simulations, the ion-sputtering ENA signal from dayside is high enough to be detected by ELENA; the spatial and time resolution is good enough to monitor instantaneous changes of the magnetospheric configuration; the spatial resolution permits to discriminate surface emissivity variations. Those considerations apply to an accumulation time of 30 s, which is a compromise between time resolution, count rates and downlink data-rate.

6. Summary and Conclusions

In this study, we have used a Monte–Carlo model of proton circulation in Mercury's dayside as an input for the simulation of neutral particle emission from both the surface (via ion-sputtering) and the exosphere (via CE). Monitoring the product of such interactions gives important information about Mercury's magnetosphere and exosphere[30] such as effectiveness of CE as a loss process of magnetospheric ions; effectiveness of ion-sputtering as a release process of exospheric neutrals; instantaneous and global monitoring of spatial distribution (via ENA deconvolution) and energy distribution of the magnetospheric plasma. Simulated ENA images, for realistic MPO/SERENA/ELENA vantage points and FoV, have been obtained and discussed; the results show the feasibility of neutral atom remote sensing made by the SERENA/ELENA sensor.

A first result of this study is that ENA fluxes coming from both CE and ion-sputtering, at Mercury, could be high enough to allow using this technique. Moreover, we have show that these neutral fluxes are extremely sensible to the external conditions, thus leading to a "neutral imaging" investigation technique for Mercury's magnetosphere and surface, through directional neutral atoms detection. In this way, it is possible to obtain basic information about plasma circulation, as it happens at the Earth,[9,10,31] and to study surface release processes.[11] According to our model, solar

wind protons are admitted into Mercury's magnetosphere only when B_z is negative. In the "*mean* PD" conditions, the total amount of protons precipitating onto the surface of Mercury is approximately equal to $4 \times 10^{25}\,s^{-1}$. For comparison, the amount of protons that exchange their charge, in 1 s, is $4 \times 10^{24}\,s^{-1}$. The first number depends only on the H^+ model adopted, and is similar to other results.[8] The second number depends also on the exospheric profiles assumed. In summary, approximately 10% of the injected protons precipitate onto the planet, 1% is neutralized due to CE and the remaining part leaves the magnetosphere.

Some more specific issues, related to BC/MPO mission, can be outlined. The MPO periherm is so close to the planet that H-ENAs are generated partially above the s/c. Since ELENA FoV is nadir-pointing, CE H-ENA imaging is better performed when the s/c is close to its apoherm, especially in the dayside, dawn and nightside regions. On the other hand, the MPO orbital configuration permits a very good spatial resolution, for both CE and ion-sputtering ENA imaging. BepiColombo program also foresees an ENA detector on board Mercury Magnetospheric Orbiter (BC/MMO). This spacecraft will be orbiting at higher distances, and will allow a global ENA imaging with lower spatial resolution. For this reason, it will be very useful to have simultaneous ENA images from both instruments.

For some vantage points (in dawn and dayside regions), the instrument is able to monitor, at the same time, proton circulation (via CE H-ENA detection) and precipitation (via the detection of ENAs sputtered from the surface), thus permitting the evaluation of the relative importance of such loss processes. In the case of sputtered ENAs, the s/c motion will help to obtain 2D images from 1D ELENA FoV, in the case of H-ENA, s/c motion could permit to observe the same emission region from different vantage point. The feasibility of such a deconvolution will be studied in a future paper.

The intensity of the directional ENA signal originated from ion-sputtering depends on both proton precipitation flux and surface properties (composition and yield). However, our simulations show that it is possible to discriminate between these two factors. In fact, by visually inspecting Fig. 5, we can note that the proton flux on the surface is roughly constant on a scale of 100 km (about two pixels); since ELENA spatial resolution is lower (between 15 and 50 km), any small-scale spatial change in ion-sputtering ENA signal is probably due to surface property variations. On the contrary, temporal variability in the ENA signal should be ascribed only to modifications in the proton circulation properties. It is worth noting that

the shape and intensity of proton flux on Mercury's surface depends both on the applied magnetic field and on the used parameters (e.g., the IMF). The magnetic field model applied here has been developed from an Earth-based model, so that it does not include a non-zero IMF B_x component. This missing contribution could have some impact in our predictions, since different magnetic models predict a different shape of H^+ surface flux,[16] and will be considered in a further study.

Finally, ELENA data analysis will be supported by the other SERENA experiment units: Strofio will monitor exospheric composition, detecting the thermal, ion-sputtered population; miniaturized ion precipitation analyzer (MIPA) will give instantaneous fluxes of precipitating ions, and help the evaluation of sputtering process efficiency; particle ion CAMera (PICAM) will give spatial and velocity distribution of ionized components of the ion-sputtering process.

References

1. N. F. Ness, K. W. Behannon and R. P. Lepping, *Icarus* **28** (1976) 479–488.
2. B. E. Goldstein, S. T. Suess and R. J. Walker, *J. Geophys. Res.* **86** (1981) 5485–5499.
3. J. A. Slavin, J. C. J. Owen, J. E. P. Connerney and S. P. Christon, *Planet. Space Sci.* **45** (1997) 133–141.
4. R. M. Killen, A. E. Potter, P. Reiff, M. Sarantos, B. V. Jackson, P. Hick and B. Giles, *J. Geophys. Res.* **85** (2001) 145–167.
5. E. Kallio and P. Janhunen, *Ann. Geophysicae* **11** (2003) 2133–2145.
6. Killen, Sarantos and Reiff, *Adv. Sp. Res.* **33** (2004) 1899–1904.
7. A. Mura, S. Orsini, A. Milillo, D. Delcourt and S. Massetti, *Icarus* **175** (2005) 305–319.
8. S. Massetti, S. Orsini, A. Milillo, A. Mura, E. De Angelis, H. Lammer and P. Wurz, *Icarus* **166** (2003) 229–237.
9. E. C. Roelof, D. G. Mitchell and D. J. Williams, *J. Geophys. Res.* **90** (1985) 10991.
10. S. Orsini, I. A. Daglis, M. Candidi, K. C. Hsieh, S. Livi and B. Wilken, *J. Geophys. Res.* **99** (1994) 13489.
11. A. Milillo, S. Orsini, D. Delcourt, E. Kallio, R. M. Killen, H. Lammer, F. Leblanc, S. Massetti, A. Mura, P. Wurz, S. Barabash, G. Cremonese, I. A. Daglis, E. De Angelis, A. M. Di Lellis, S. Livi and V. Mangano, *Space Science Reviews* **117** (2005) 397–443.
12. N. A. Tsyganenko, *Proc. of 3rd International Conference on Substorms (ICS-3)*, ESA SP-389, 1996.
13. H. Volland, *J. Geophys. Res.* **83** (1978) 2695–2699.
14. D. C. Delcourt, S. Grimald, F. Leblanc, J. J. Berthelier, A. Millilo, A. Mura and S. Orsini, *Ann. Geophysicae* **21** (2003) 1723–1736.

15. K. Kabin, T. I. Gombosi, D. L. DeZeeuw and K. G. Powell, *Icarus* **143** (2000) 397–406.
16. M. Sarantos, P. H. Reiff, T. W. Hill, R. Killen and A. L. Urquhart, *Planet. Space Sci.* **49** (2001) 1629–1635.
17. E. Kallio and P. Janhunen, *Adv. Space Res.* **33** (2004) 2176–2181.
18. K. W. Ogilvie, J. D. Scudder, V. M. Vasyliunas, R. E. Hartle and G. L. Siscoe, *J. Geophys. Res.* **83** (1977) 1807–1824.
19. W. H. Ip, *Icarus* **71** (1987) 441.
20. P. Siegmund, *Phys. Rev.* **184** (1969) 383–416.
21. E. M. Sieveka and R. E. Johnson, *Astrophys. J.* **287** (1984) 418–426.
22. G. Betz and K. Wien, *Int. J. Mass Spectrometry and Ion Processes* **140** (1994) 1–110.
23. A. L. Broadfoot, D. E. Shemanky and S. Kumar, *Geophys. Res. Lett.* **3** (1976) 577.
24. P. Wurz and H. Lammer, *Icarus* **164** (2003) 1–13.
25. G. B. Hasted, Butterworths, *Physics of Atomic Collisions* (London, 1964), pp. 416.
26. ESA PL Team, SERENA experiment interface document A, ESA, BC-EST-RS-02522, 2005.
27. BepiColombo Project Team, Experiment interface document B, ESA, BC-EST-RS-1140, 2005.
28. T. M. Stephen and L. B. Peko, *Rev. Sci. Instruments* **71** (2000) 1355–1359.
29. S. Barabash, A. V. Lukyanov, P. C. son Brandt and R. Lundin, *Planet. Space Sci.* **49** (2001) 1685.
30. S. Orsini, A. Milillo, E. De Angelis, A. M. Di Lellis, V. Zanza and S. Livi, *Planet. Space Sci.* **49** (2001) 1659.
31. J. L. Burch, S. B. Mende, D. G. Mitchell, T. E. Moore, C. J. Pollock, B. W. Reinisch, B. R. Sandel, S. A. Fuselier, D. L. Gallagher, J. L. Green, J. D. Perez and P. H. Reiff, *Science* **291**, 5504 (2001) 619–624.

BEPICOLOMBO — MPO SCIENTIFIC ASPECTS AND SYSTEM UPDATE

JOHANNES BENKHOFF* and RITA SCHULZ
ESA Research and Scientific Support Department, ESTEC
Postbox 299, NL-2200 AG Noordwijk zh, The Netherlands
**johannes.benkhoff@esa.int*

The BepiColombo is an interdisciplinary mission to explore the planet Mercury through a partnership between ESA and Japan's Aerospace Exploration Agency (JAXA). From their dedicated orbits two spacecrafts, the Mercury Planetary Orbiter (MPO) and the Mercury Magnetospheric Orbiter (MMO), will be studying the planet and its environment. Both orbiter will be launched together on a single Soyuz–Fregat. The launch is foreseen for August 2013 with arrival in late summer 2019. Solar electric propulsion will be used for the journey to Mercury. The BepiColombo scientific payload has been selected in 2004. The MPO payload comprises eleven instruments/instrument packages; the MMO payload consists of five instruments/instrument packages. Together, both spacecraft and their scientific payload will provide the detailed information necessary to understand Mercury and its magnetospheric environment and to find clues to the origin and evolution of a planet close to its parent star. The MPO will focus on a global characterization of Mercury through the investigation of its interior, surface, exosphere, and magnetosphere. In addition, it will be testing Einstein's theory of general relativity. Major effort was put into optimizing the scientific return by defining the payload complement such that individual measurements can be interrelated and complement each other. A detailed overview of the status of BepiColombo will be given with special emphasis on the MPO and its payload complement.

1. Introduction

Mercury still is the least known planet in the inner solar system and its precise characterization is long overdue. Being the planet closest to the sun it represents the inner end-member of the four terrestrial planets. As such it plays a fundamental role in constraining and testing the competing theories explaining dynamical and compositional aspects of the formation and evolution of the whole group. However, only a space mission can provide the necessary details of its properties. From the Earth, Mercury's maximum

*Corresponding author.

elongation to the Sun is 28°. It is therefore visible for only 2 h at most and at very low altitude either before sunrise or after sunset. This makes detailed observations of the planet for large ground-based telescopes as well as the Hubble Space Telescope very complicated or impossible, because none of them can take measurements this close to the Sun. Nevertheless, only one spacecraft, Mariner 10, ever visited Mercury about 30 years ago, performing three fly-bys in 1974–1975. During these short fly-bys Mariner 10 was able to image about 45% of the planet's surface, to discover its unexpected magnetic field, the existence of gaseous species forming an exosphere and the presence of a unique magnetosphere. However, nothing is known about Mercury's interior structure or its elemental and mineralogical composition.

Drawing conclusions on the formation and evolution of Mercury requires a complete description of the planet and its environment. The BepiColombo mission shall provide the detailed information necessary to understand Mercury and its magnetospheric environment and to find clues to the origin and evolution of a planet close to its parent star. In addition, the mission provides unique possibilities for testing Einstein's theory of general relativity. The scientific background and the objectives of the mission have been outlined in Refs. 1 and 2. The main issues to be addressed are as follows

- Origin and evolution of a planet close to its parent star.
- Mercury's figure, interior structure, and composition.
- Interior dynamics and origin of its magnetic field.
- Exogenic and endogenic surface modifications, cratering, tectonics, volcanism.
- Composition, origin, and dynamics of Mercury's exosphere and polar deposits.
- Structure and dynamics of Mercury's magnetosphere.
- Test of Einstein's theory of general relativity.

The complexity of these fundamental scientific objectives poses high requirements on the payload of the mission. None of the objectives concerning Mercury can be reached through measurements of any single instrument. However, the combination of the measurements of the various instruments aboard BepiColombo will provide answers to many outstanding questions related to the origin and evolution of the innermost member of the group of the terrestrial planets.

2. Updated Mission Overview

BepiColombo is an interdisciplinary mission devoted to the exploration of Mercury and its environment. It aims at understanding the process of planetary formation and evolution in the hottest part of the proto-planetary nebular and at understanding the similarities and differences between the magnetospheres of Mercury and Earth. The mission will be carried out as a joint project between ESA and JAXA. The mission's first priority is to provide information necessary to understand Mercury and its magnetospheric environment. Specific issues of fundamental physics will also be addressed. BepiColombo consist of two elements, the Mercury Planetary Orbiter (MPO) and the Mercury Magnetospheric Orbiter (MMO).

2.1. *Mission scenario*

BepiColombo will be launched from Kourou with an upgraded Soyoz-2B with Fregat M upper stage. The MPO will carry the MMO during launch and cruise. The launch is foreseen for August 2013. Solar electric propulsion will be used for the roughly 6-year journey to Mercury. Upon arrival in 2019, the solar electric propulsion module will be jettisoned and chemical propulsion will provide the required thrust for Mercury capture and injection of both spacecraft into their dedicated polar orbits. The MMO will be released first, after which an additional thrust phase will insert the MPO into its final orbit. Both orbits are elliptical with eccentricity and inclination optimized for the study of Mercury (MPO orbit: $400\,\mathrm{km} \times 1,500\,\mathrm{km}$) and its magnetosphere (MMO orbit: $400\,\mathrm{km} \times 12,000\,\mathrm{km}$). Being as close as $0.3\,\mathrm{AU}$ to the Sun, the spacecraft have to withstand extremely high temperatures and radiation doses, which require inclusion of radiators and extensive shielding in their design. The science operation has been baselined for duration of one Earth year in Mercury orbit with the possibility to extend the mission for another year.

2.2. *The Mercury planetary orbiter*

The MPO is a three-axis-stabilized and nadir-pointing spacecraft focusing on a planet-wide remote sensing. Its low-eccentricity polar orbit ($400\,\mathrm{km} \times 1,500\,\mathrm{km}$) will provide excellent spatial resolution over the entire planet surface. BepiColombo is currently in the so called *definition phase*, during which two competing definition study prime contractors undertake all mission definition activities including the detailed design of the MPO spacecraft. The design of the MPO can therefore not be provided in this paper. Towards the end of this definition phase the two contractors will prepare

and submit industrial proposals on which basis the mission as well as the MPO design, and with this also the prime industrial contractor will be selected. After the completion of this selection process, BepiColombo will move into its implementation phase, starting late 2006.

2.3. *Payload*

In November 2004, the ESA Science Program Committee has provisionally approved the selection of the MPO scientific payload. It comprises 16 instruments, which in several cases have been combined to instrument packages sharing common subsystems and reflects the recommended MPO Reference Payload.

With this payload complement we will be able to perform a global characterization of Mercury. The MPO will concentrate on the investigation of Mercury's interior, surface and exosphere, and support the MMO in the detailed characterization of the magnetosphere and its relation to the exosphere and the planet.

Figure 1 outlines the science topics addressed by the measurements of the MPO payload. The two cameras will provide a global mapping of the surface morphology with a spatial scale of 100 m and in addition image

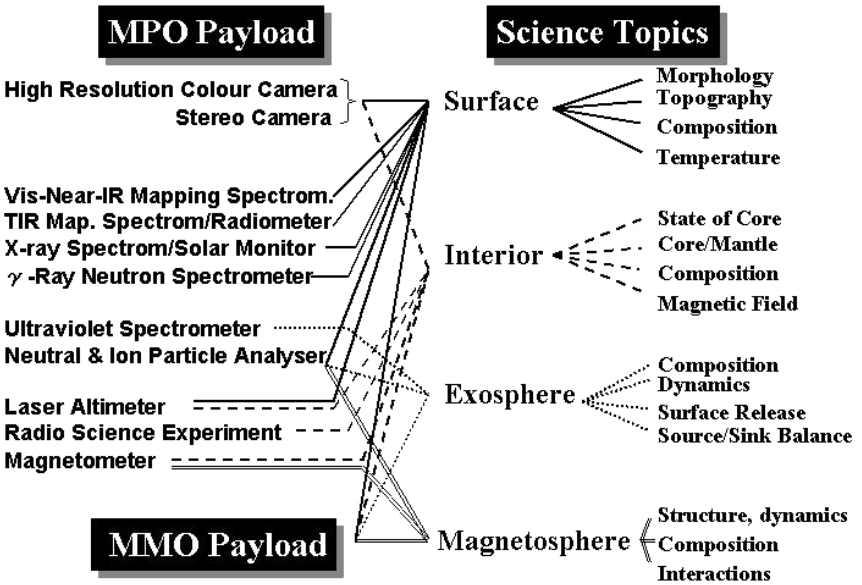


Fig. 1. The topics addressed by the combined measurements of the selected MPO payload complement.

areas of specific interest with spatial resolutions better of up to 5 m. The near- and thermal-infrared spectrometers will obtain the global mineralogical surface composition with a spatial resolution of 500 m, comparable to the regolith mixing length. The X-ray spectrometer will provide a global mapping of the elemental surface composition at a scale of about 50 km, and during higher solar activity measure with high spatial resolution of up to 500 m. The combination of these five instruments allows the morphological and compositional characterization of individual surface features, hence the identification of compositional variations between surface features. This is absolutely necessary in order to conclude on whether specific landmarks have been produced by endogenic processes (e.g., volcanism) or exogenic processes (e.g., impacting objects). Knowledge of Mercury's surface composition will provide a key test of competing models for the formation and evolution of Mercury and with that of all the terrestrial planets. Elements not accessible through X-ray spectroscopy as well as the sub-surface elemental composition will be measured by the γ-ray-neutron spectrometer, albeit at lower spatial resolution. The neutron spectrometer and will additionally take measurements of the radar-bright spots observed from ground in the polar regions[3,4] to identify their composition. If these spots, originally suspected to be due to water ice, are composed of sulphur,[5] this will support the planet's core to be composed of Fe-FeS alloys, which are liquid at lower temperatures than pure iron.[6]

Mercury's interior structure will be investigated by the radio science experiment in combination with the laser altimeter and the high-resolution camera. Together these instruments will provide an accurate determination of the global gravity field and the rotational state of the planet as well as the amplitude of the physical libration. These parameters allow constraining its internal structure. A recent discussion of Mercury's interior and strategies for its exploration using gravity and topography data can be found in Ref. 7. The question whether the solid crust of the planet is decoupled from the inner core by some liquid layer will also be tackled by accurate measurements of the planetary magnetic field.

The exosphere will be studied by remote sensing as well as *in situ* measurements. It is expected that the six elements observed to date (Ca, Na, K, H, He, and O) only constitute a fraction of Mercury's exosphere.[8] Spectrometers in the visible and ultraviolet will provide large-scale measurements of its composition, vertical structure and dynamics, while the neutral and ion particle analyzers will perform an *in situ* analysis of the exosphere and its interactions with the surface and the magnetosphere to

understand the surface release processes and the source/sink balance of the exosphere. Together these measurements will help explaining the cycling of volatile elements between Mercury's interior, surface and exosphere, and the contribution of meteoritic/cometary material and solar wind plasma to Mercury's near-surface volatile budget.

2.4. *Description of the MPO payload instruments*

In this section a description of the scientific goals of each of the payload elements is given.

2.4.1. *Spectrometer and imagers for MPO BepiColombo integrated observatory system*

Spectrometer and images for MPO BepiColombo integrated observatory system (SIMBIO-SYS) is an integrated package for the imaging and spectroscopic investigation of the surface of Mercury. The science goals of SIMBIO-SYS are to examine the surface geology (stratigraphy, geomorphology), volcanism (lava plain emplacement, volcanoes identification), global tectonics (structural geology, mechanical properties of lithosphere), surface age (crater population and morphometry, degradation processes), surface composition (maturity and crustal differentiation, weathering, rock forming minerals abundance determination), and geophysics (libration measurements, internal planet dynamics) of Mercury. It incorporates capabilities to perform medium space resolution global mapping in stereo and color imaging using two pan-chromatic and three broad-band filters, respectively, as well as high space resolution imaging in a pan-chromatic and three broad-band filters and imaging spectroscopy in the spectral range 400–2,000 nm.

It consist of (1) the stereo channel (STC), which will provide the global color coverage of the surface in full stereo at 50 m/pixel resolution with the aim of defining the main geological units, large scale tectonic features, impact crater population and, if present, volcanic edifices, (2) the high spatial resolution imaging channel (HRIC), which objective is the characterization of special surface targets with high-resolution images at ground pixel sizes of about 5 m/pixel from altitude 400 km in four different bands, and (3) the visible infrared hyperspectral imager channel (VIHI), which is a hyper spectral imager in the visible and near-infrared range that will map the planet in order to provide the global mineralogical composition of the surface.

2.4.2. *Mercury radiometer and thermal infrared spectrometer*

It is the goal of Mercury radiometer and thermal infrared spectrometer (MERTIS) to provide detailed information about the mineralogical composition of Mercury's surface layer by globally mapping spectral emittance with a high spectral resolution. The MERTIS will cover a wavelength range from 7 to 14 μm with a spectral resolution up to 90 nm. This allows detecting and identifying the characteristic features of the surface minerals in this spectral region, such as the Christiansen frequencies, reststrahlen bands, and transparency features. In addition MERTIS will be able to measure thermo physical properties of the surface like thermal inertia and surface texture. The MERTIS is an IR-imaging spectrometer, which will make use of the micro-bolometer technology where no cooling is required. The MERTIS will globally map the planet with a spatial resolution of 500 m.

2.4.3. *Mercury imaging X-ray spectrometer*

The Mercury imaging X-ray spectrometer (MIXS) instrument is designed to perform X-ray fluorescence (XRF) analysis of the surface of Mercury. The XRF is a well-known technique used for remote sensing the atomic composition of airless, inner solar system bodies. The primary scientific goal of MIXS is to produce global elemental abundance maps of key-rock forming elements to an accuracy of 10–20%. If solar conditions permit, high-spatial resolution mapping of these elemental abundances will take place. A third scientific objective is to confirm whether the auroral zone, where energetic particles interact with the surface, is an intense source of continuum and line X-rays. In order to achieve its science objectives, MIXS consists of two channels — the MIXS-C, a collimator providing efficient flux collection over a broad range of energies with a wide field of view for planetary mapping and the MIXS-T, an imaging telescope with a narrow field for high-resolution measurements of the surface.

2.4.4. *The solar intensity X-ray and particle spectrometer*

The objective of the solar intensity X-ray and particle spectrometer (SIXS) is to perform measurements of X-rays and particles of solar origin. The instrument is capable of broadband measurements of X-ray, proton, and electron spectra with high-time resolution and a very wide field-of-view. Data provided by SIXS are mandatory for a valid fluorescence analysis of MIXS spectra. The solar X-ray monitor is also the single source of this

necessary data in orbital phases of Mercury while the planet is behind the Sun. Scientific objectives for SIXS are studies of temporal and spectral variability of solar X-ray corona, temporal behavior, and X-ray spectral classification of solar flares, and temporal and spectral variability of proton and electron radiation near Mercury.

2.4.5. *Probing of hermean exosphere by ultraviolet spectroscopy*

Probing of hermean exosphere by ultraviolet spectroscopy (PHEBUS) is a dual ultra-violet spectrometer working in the wavelength range from 55 to 315 nm and devoted to characterize Mercury's exosphere composition and dynamics. The search for surface ice layers in permanently shadowed regions of high-latitude craters is one of the instrument objectives. However, the general scientific objectives of PHEBUS, oriented toward better understanding the coupled surface exosphere-magnetosphere system are to detect new species, including metallic ones, volatile atoms, molecules and radicals, noble gases, and ions, in addition to already detected species; to measure an average exosphere (number densities of constituents, vertical structure), with as much as possible species monitored together, at different positions of Mercury around the Sun; to measure sharp local and temporal variations of the exosphere content (time scale: less than a few hours), at specific times and places of interest and to search for albedo variations of Mercury's night side surface, lighted by the interplanetary H Ly-α glow in order to exhibit possible signatures of surface ice layers (H_2O, SO_2, N_2, CO_2, ...) in high latitude polar craters, and any other signature of interest on the night side.

2.4.6. *Search for exospheric refilling and emitted natural abundances*

The scientific objectives of search for exospheric refilling and emitted natural abundances (SERENA) are to provide information on the whole surface–exosphere magnetosphere system and the processes involved in the system as well as in the interaction with the solar wind and the interstellar medium. The interaction between the Hermean environment with energetic plasma particles, solar radiation, and micrometeorites gives rise to both thermal and directional neutral populations in the near-planet space. The SERENA neutral particle analyser sensors will record such populations, while the SERENA ion spectrometer sensors will record photo-ionized or charged components of the surface release processes as well as the plasma precipitation and circulation in the Hermean magnetosphere.

The SERENA combines four sensors into one scientific package: emitted low-energy neutral atoms (ELENA) is a neutral particle camera that investigates neutral gases escaping from the surface of Mercury, their dynamics and the processes responsible for such a population, start from a rotating field mass spectrometer (STROFIO) is a neutral particle spectrometer which investigates the exospheric gas composition, miniature ion precipitation analyser (MIPA) is an ion monitor investigating the chain of processes by which plasma precipitate towards the surface, and planetary ion camera (PICAM) an ion mass spectrometer operating as an all-sky camera for charged particles to study the chain of processes by which neutrals are ejected from the soil, ionized and transported through the environment of Mercury.

2.4.7. *BepiColombo laser altimeter*

The BepiColombo laser altimeter (BELA) instrument scientific goals are characterizing and measuring the figure, topography, and surface morphology of Mercury. The BELA will provide absolute topographic height and position with respect to a Mercury centered coordinate system. This information will be used to create a digital terrain model that allows to invert the gravity field of Mercury data, acquired by MORE instrument, for internal structure and to quantitatively explore the geology, the tectonics, and the age of the planet's surface. The returned laser signal can be used to measure the local surface roughness and the albedo. In synergy with other instrument results it will also be possible to explore the interior structure and the geologic evolution of the planet.

2.4.8. *Mercury magnetometer*

The primary objective of Mercury magnetometer (MERMAG) is to provide the measurements that will lead to the detailed description of Mercury's planetary magnetic field and its source. The goal is to contribute to the understanding of the origin, evolution, and current state of the planetary interior. The determination of all the terms associated with the internal field up to the octopole with high accuracy as well as higher order terms, depending on the structure of the internal field should also be possible.

The secondary objectives of MERMAG relate to the interaction of the solar wind with the magnetic field of Mercury and the planet itself, the formation and dynamics of the magnetosphere as well as to the processes that control the interaction of the magnetosphere with the planet. In particular,

measurements close to the planet will allow the determination of the conditions for access of the solar wind to the planetary surface and assessing the role and importance of different current systems, including subsurface induction currents and the conductivity of the regolith. These objectives will be greatly assisted by the planned close association with the magnetic field investigation on the MMO.

2.4.9. *Mercury orbiter radio-science experiment*

Mercury orbiter radio-science experiment (MORE) addresses scientific goals in geodesy, geophysics, and fundamental physics. It will help to determine the gravity field of Mercury as well as the size and physical state of its core. It will provide crucial experimental constraints to models of the planet's internal structure and test theories of gravity with unprecedented accuracy. MORE will also measure the gravitational oblateness of the Sun and test and characterize the most advanced interplanetary tracking system ever built. Finally, it will assess the performances of the novel tracking system in precise orbit determination and space navigation. These scientific goals will be achieved by means of several data types, generated by MORE itself at the ground station, other onboard instruments (BELA, ISA, and SIMBIO-SYS) and the onboard attitude determination and control system.

The goal of the BepiColombo rotation experiment is the determination of Mercury's obliquity (i.e., the obliquity of the equator with respect to the orbital plane) and the amplitude of its 88-day physical librations in longitude. These two quantities, together with the coefficients of the second-degree harmonics of the gravity field, will indicate whether or not Mercury has a molten core and provide the radius of this molten core.

2.4.10. *Italian spring accelerometer*

Italian spring accelerometer (ISA) is a three-axis high-sensitivity accelerometer is not only devoted to study the planet but also to test Einstein's theory of general relativity (GR) to an unprecedented level of accuracy. To perform such measurements it is necessary to combine data from four different instruments located on-board the MPO: star tracker, high-resolution camera, accelerometer, and transponder. The scientific objectives of ISA are strongly connected with those of the MORE Experiment. Together the experiments can give information on Mercury's interior structure as well as test Einstein's theory of the GR. The scientific goals are (1) to study the global gravity field of Mercury and its time variations due to solar tides,

(2) to investigate the local gravity anomalies, (3) to determine the rotation state of Mercury, and (4) to measure the motion of Mercury center of mass.

2.4.11. *Mercury gamma-ray and neutron spectrometer*

The scientific goals of Mercury gamma-ray and neutron spectrometer (MGNS) are to measure the elemental surface and sub-surface composition for distinguishable regions over the entire Mercury by the measuring (1) the nuclear lines of major soil-composing elements, (2) the leakage flux of neutrons, and (3) the lines of natural radioactive elements and to determine the regional distribution of volatile depositions on the polar areas of Mercury which are permanently shadowed from the Sun, and to provide a map of column density of this depositions with the accuracy of $0.1\,\mathrm{g\,cm^{-2}}$ and with a surface resolution of about 400 km.

3. Summary

BepiColombo will do a full in-depth exploration of Mercury and its magnetospheric environment. With its two spacecraft interdisciplinary approach, the mission is able to provide the detailed information necessary to resolve the many outstanding questions related to its origin, evolution, structure and environment. The MPO will provide a global mapping of the surface morphology, composition and temperature, study its extremely thin atmosphere (exosphere) and investigate Mercury's interior structure and magnetic field. The MMO investigations will result in a complete characterization of the planet's magnetosphere. Simultaneous measurements of the two spacecraft will resolve spatial and temporal ambiguities in the exosphere and magnetosphere that would arise from single point observations. Given the high scientific potential related to this planet and its environment BepiColombo will open a new frontier in the study of our solar system.

References

1. R. Grard and A. Balogh, Returns to Mercury: Science and mission objectives, *Plan. Space. Sci.* **49** (2001) 1395–1407.
2. R. Grard, *Ad. Sp. Res.* **33** (2003) 2125.
3. J. K. Harmon and M. A. Slade, Radar mapping of Mercury: Full-disk images and polar anomalies, *Science* **258** (1992) 640–642.
4. J. K. Harmon, P. J. Perillat and M. A. Slade, High-resolution radar imaging of Mercury's north pole, *Icarus* **149** (2001) 1–15.

5. A. L. Sprague, D. M. Hunten and K. Lodders, Sulfur at Mercury, elemental at the poles and sulfides in the regolith, *Icarus* **118** 211–215; *Erratum Icarus* **123** (1995) 247.
6. H. Harder and G. Schubert, Sulfur in Mercury's core? *Icarus* **151** (2001) 118–122.
7. T. Spohn, F. Sohl, K. Wieczerkowski and V. Conzelmann, The interior structure of Mercury: What we know, what we expect from BepiColombo, *Plan. Space Sci.* **49** (2001) 1561–1570.
8. D. M. Hunten, T. H. Morgan and D. E. Shemansky, The Mercury atmosphere, in *Mercury*, eds. F. Vilas, C. R. Chapman and M. S. Matthews (Tucson, The University of Arizona Press, 1988), pp. 562–612.

DIAGNOSING THE MERCURY PLASMA ENVIRONMENT USING LOW-FREQUENCY ELECTRIC FIELD MEASUREMENTS

L. G. BLOMBERG* and J. A. CUMNOCK
Alfvén Laboratory, School of Electrical Engineering
Royal Institute of Technology, SE-100 44 Stockholm, Sweden
**lars.blomberg@ec.kth.se*

The magnetosphere of Mercury is most intriguing because of its extreme nature, with scale sizes vastly different from the corresponding terrestrial ones, and a harsh environment at a comparatively small solar distance. The present brief paper summarizes some scientific topics where electric field measurements at low frequency will make a significant contribution to the understanding of the Mercury plasma environment, and discusses possible diagnostics.

1. Introduction

Mercury's magnetosphere is quite different from that of the Earth making it interesting to study in its own right, yet the two magnetospheres are similar enough to allow comparative studies which are highly likely to provide new knowledge and improved understanding of them both. The electric field describes the plasma transport in a magnetosphere. The electric field also plays an important role in the interaction of the solar wind with the planetary magnetosphere and in the interaction of the magnetosphere with the underlying ionosphere, in the case of Earth, or, most likely, the underlying planetary surface or its immediate environment, in the case of Mercury. Finally, it is an important parameter for the energy transport between the different regions.

BepiColombo Mercury Magnetospheric Orbiter (MMO) will make the first observations ever of the low-frequency electric field in the Mercury environment. The spacecraft is equipped with two pairs of sensors for the electric field, Wire Probe anTenna (WPT)[1] and Mercury Electric Field In-Situ TOol (MEFISTO).[1,2] The two pairs are orthogonally mounted in the spacecraft spin plane and will measure the electric field vector in the spin plane over a wide frequency range.

*Corresponding author.

2. Scientific Topics

We here briefly review a number of scientific topics related to the low-frequency electric field in the plasma environment of planet Mercury. For a more thorough overview of the Mercury environment the reader is referred to, e.g., Cumnock and Blomberg.[3]

2.1. *Plasma convection and cross-polar potential*

Plasma convection in a magnetosphere is controlled by the DC electric and magnetic fields. Thus, to determine plasma circulation and plasma transport across boundaries the static component of the electric field needs to be measured. The electric field and plasma flow inside the magnetosphere also give important information on the interaction of the solar wind with the magnetosphere.

In Earth's magnetosphere there is a saturation effect so that the potential drop inside the magnetopause does not increase linearly with an increasing potential drop across a similar distance in the solar wind.[4–7] Hill's explanation of the saturation effect is based on the idea that field-aligned currents are set up that create a magnetic field opposing the field at the magnetopause and thus weakening reconnection, eventually limiting it at some fixed plateau value. The efficiency of this feedback process depends on the ionospheric conductivity. However, at Mercury it is difficult to conceive of a similar process given the low conductivity near the planet and the typically corresponding low ability of its magnetosphere to sustain any appreciable field-aligned currents. Thus, for strong solar wind electric fields, a significantly higher fraction may penetrate into the Hermean magnetosphere than would penetrate into the Terrestrial one, as further discussed elsewhere.[8] *In situ* measurements are needed to clarify the saturation mechanism at Mercury.

2.2. *ULF pulsations*

There are indications from Mariner 10 magnetometer data of the existence of field-line resonances at Mercury.[9] A field-line resonance is a fundamental response of the planetomagnetic field to the solar wind's interaction with the magnetopause. At Mercury they are interesting also because they depend on the (electromagnetic) reflective properties of the surface and thus may be used to estimate the surface conductivity. This requires simultaneous measurements of the electric and the magnetic fields at low frequency. Some

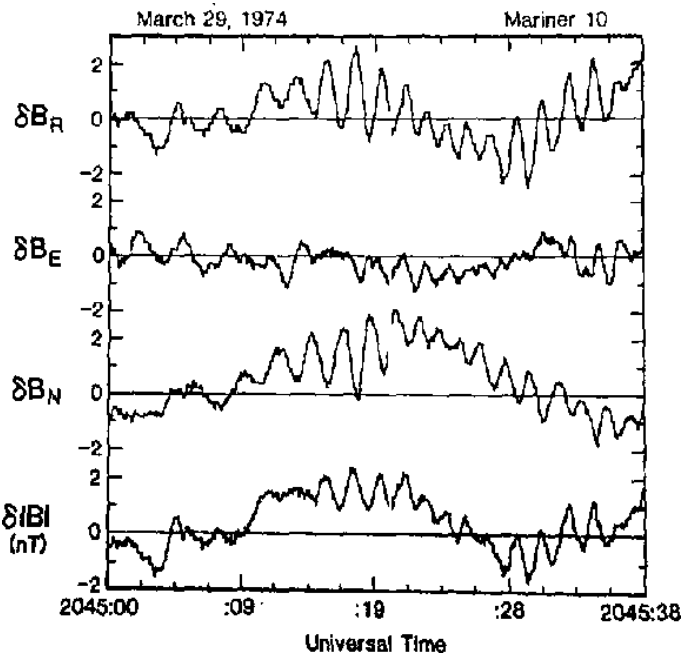


Fig. 1. ULF pulsations recorded at Mercury by the Mariner 10 magnetometer, assumed to be the fourth harmonic of a standing wave (after Russell[9]).

of the "surface" conductivity may reside in a photoelectron cloud just above the surface. By comparing observations in different local time sectors the two components can be separated.

In the terrestrial magnetosphere (damped) standing ULF waves are quite frequently generated by a Kelvin–Helmholtz instability at the flanks of the magnetosphere as the solar wind streams past. A similar effect is conceivable at Mercury, and there have been some indications that it does indeed occur[9] (see Fig. 1). This class of waves is often referred to as field line resonances, since they involve a large-scale fluctuating motion along the entire length of a set of magnetic field lines. Depending on the conducting properties of the planet and its immediate environment, the waves will be reflected either above, at, or below the surface. Also, depending on whether the conductance at the reflection boundary is greater or less than the waveguide conductance either the magnetic field or the electric field of the wave will change phase when it is reflected. Mariner 10 did not measure the electric field and, thus, no firm conclusions regarding the nature

of the waves observed can be drawn. Combined low-frequency electric and magnetic field measurements on BepiColombo will enable us, for the first time, to properly diagnose standing ULF waves at Mercury, as well as tell us more about the conductive properties of the reflective boundary at low altitude. They will also tell us more about the conductive properties across the magnetic field close to the planetary surface.[10]

2.3. *Field-aligned currents*

Mariner 10 observed signatures of field-aligned currents.[11] Whether these were persistent or transient is not known, although the generally poorly conducting low-altitude region may make persistent currents unlikely. Studying these with simultaneous measurements of the electric and the magnetic fields will shed additional light not only on the surface conductivity issue but also on the more general, and largely unsolved, question of current closure.[12,13] Depending on the details of the low-altitude closure mechanism the correlation between the electric and the magnetic "disturbance fields" will occur for different components, thus enabling diagnostics (Sec. 2.5).

2.4. *Exo-ionospheric conductivity*

A conducting layer above the surface of Mercury could, at least in principle, come about in several ways. If the atmosphere were sufficiently dense it would be a plasma created mainly by photo-ionization of the neutral atmosphere. However, another possibility is that of a cloud of electrons, similar to an electron-rich cathode sheath, made up of electrons photoemitted from the surface.

2.5. *Current closure*

A spacecraft overflying a static structure of field-aligned currents closed by Pedersen current at low altitude will see a correlation between orthogonal components of the electric and magnetic vectors transverse to the main magnetic field (Blomberg[10] and references therein). Assuming also homogeneous conductivity, orthogonal components of the transverse **E** and **B** are correlated. Such a correlation is often observed when overflying structures in the terrestrial ionosphere. If photoemission from the surface should indeed be an important source of charge carriers in the near-Hermean environment, there would be a net negative charge present, i.e., there would be a cathode sheath rather than a plasma. In this case, because of the

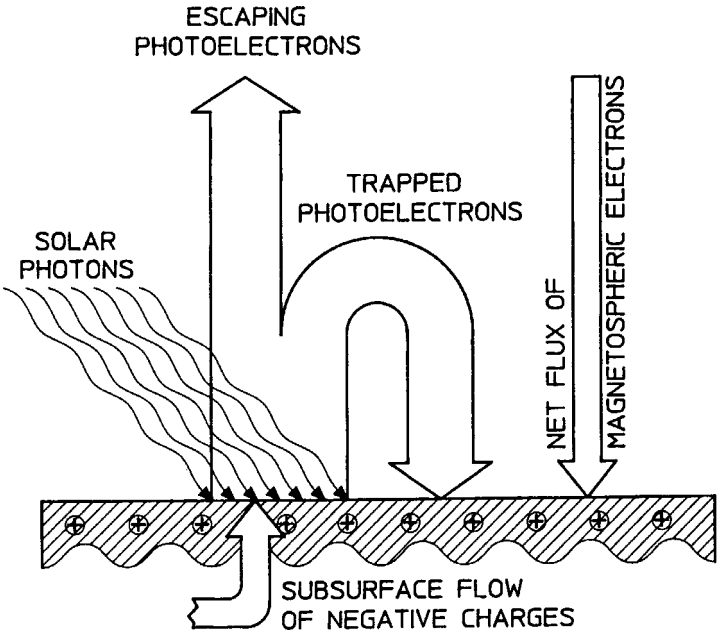


Fig. 2. Possible current closure mechanisms at Mercury (after Grard[12]).

absence of collisions, the current flowing may be a Hall current arising from $\mathbf{E} \times \mathbf{B}$ drifting electrons. With this assumption the correlation is seen for the same components of $\mathbf{E}$ and $\mathbf{B}$ rather than for the orthogonal ones.[10] Figure 2 illustrates some of the possible current closure mechanisms at Mercury.

2.6. *Magnetospheric compressibility*

The compressibility of the magnetosphere under fluctuations is also related to the conductivity of the planet. If the surface and body are non-conducting the field lines are anchored in the dynamo region of the planet, whereas with a perfectly conducting surface they would be anchored at the surface, thus reducing the magnetospheric compressibility. For the intermediate case the anchor depth would be frequency dependent and approximately given by the condition that the magnetic Reynold's number be about unity. This is yet another reason for trying to understand the nature of Mercury's conductivity.

2.7. *Mirror point asymmetry*

Because of the expected dipole offset from the centre of the planet, causing differently wide loss-cones in the two hemispheres, interhemispheric charge flow may arise along Mercury's magnetic field.[13] Such a charge flow will lead to weak field-aligned currents connecting the hemispheres or to a parallel electric potential distributed along the field lines, or both.

2.8. *Particle trapping*

The maximum kinetic energy of the particles that can be trapped in Mercury's magnetic field can be estimated by requiring the gyro radius be less than some fraction of the size of the magnetosphere. The upper limit to the energy of trapped protons is found to be of the order of 100 keV with a fairly large uncertainty depending on the assumptions made both about the size of the magnetosphere, the relevant fraction of its size, and the magnitude of the magnetic field (which may vary significantly over the gyro path). For heavier ion species the upper energy limit is lower, in inverse proportion to the square root of the atomic mass. Electrons can likely be trapped at energies up to several 100 MeV (i.e., beyond the energy range of all magnetospheric electrons). Whether radiation belts exist at Mercury is controversial. If they exist they are likely not present at all times, but rather an occasional phenomenon. Possible radiation belts at Mercury have also been discussed by, for example, Russell *et al.*[14] and Orsini *et al.*[15]

2.9. *Wave activity at higher frequencies*

2.9.1. *LF and MF waves*

Energy transport within the Hermean magnetosphere is partly governed by plasma wave activity, in particular Alfvénic activity. Alfvén waves arise where dynamic processes occur. They can transport energy, in the form of electromagnetic energy, large distances before dissipating it through kinetic, inertial, or wave breaking processes. BepiColombo will be the first mission to Mercury where the detailed physics can be studied, i.e., energy transport, acceleration processes, and dissipation. A number of wave emission processes are expected. Hermean kilometric radiation (HKR) well below 1 MHz may be emitted from the "auroral" acceleration regions that are known to exist at Mercury, and have a similar cause as the Auroral kilometric radiation (AKR) near Earth. It is interesting to note that radio emissions originating near Mercury below about 50–80 kHz will be trapped

inside (reflected back into) the Hermean magnetospheric cavity due to the dense ambient solar wind (30–70 cm^{-3}). Further, the possible existence of radiation belts around Mercury may be inferred from synchrotron radiation emissions with a peak around a few MHz.[8]

2.9.2. *Solar activity*

By monitoring bursts of solar radio emissions that indicate the solar disturbance level, it is possible to investigate the whole chain of events from the Sun to the processes directly responsible for magnetospheric storms and substorms at Mercury. Specific activity on the Sun can then be related to specific activity in the Hermean magnetosphere. The time delay is somewhat less than one day, for example, for a solar coronal mass ejection to reach Mercury. It will be possible to compare space weather effects on Mercury with those on Earth, both directly when Earth and Mercury are close to each other in heliospheric longitude, and indirectly when they are separated in longitude.

3. Summary

Mercury's plasma environment in general and its magnetosphere in particular are exciting targets for future exploration. Since Mariner 10 did not and Messenger will not measure the low-frequency electric field, BepiColombo provides the first opportunity to make direct measurements of the electric fields around Mercury.

References

1. H. Matsumoto, J.-L. Bougeret, L. G. Blomberg, H. Kojima, S. Yagitani, Y. Omura, M. Moncuquet, G. Chanteur, Y. Kasaba, J.-G. Trotignon, Y. Kasahara and BepiColombo MMO PWI Team, Plasma/Radio wave observations at Mercury by the BepiColombo MMO spacecraft, *Advances in Geosciences* 2005, this issue.
2. L. G. Blomberg, H. Matsumoto, J.-L. Bougeret, H. Kojima, S. Yagitani, J. A. Cumnock, A. I. Eriksson, G. T. Marklund, J.-E. Wahlund, L. Bylander, L. Åhlén, J. A. Holtet, E. Kallio, Y. Kasaba, A. Matsuoka, M. Moncuquet, K. Mursula, Y. Omura and J. G. Trotignon, MEFISTO — An electric field instrument for BepiColombo/MMO, *Adv. Space Res.*, in press, 2005.
3. J. A. Cumnock and L. G. Blomberg, The Mercury environment: A literature survey, *Royal Inst. Technol. Report TRITA-ALP-2003-02*, 2003.

4. T. W. Hill, A. J. Dessler and R. A. Wolf, Mercury and Mars: The role of ionospheric conductivity in the acceleration of magnetospheric particles, *Geophys. Res. Lett.* **3** (1976) 429–432.
5. T. W. Hill, Magnetic coupling between solar wind and magnetosphere: Regulated by ionospheric conductance? *Eos Trans. AGU* **65** (1984) 1047.
6. P. H. Reiff and J. G. Luhmann, Solar wind control of the polar-cap voltage, in *Solar Wind-Magnetosphere Coupling*, eds. Y. Kamide and J. A. Slavin (Terra Sci., Tokyo, 1986), pp. 453–476.
7. G. L. Siscoe, G. M. Erickson, B. U. Ö. Sonnerup, N. C. Maynard, J. A. Schoendorf, K. D. Siebert, D. R. Weimer, W. W. White and G. R. Wilson, Hill model of transpolar potential saturation: Comparisons with MHD simulations, *J. Geophys. Res.* **107**, A6, doi:10.1029/2001JA000109 (2002) 1075.
8. L. G. Blomberg, J. A. Cumnock, Y. Kasaba, H. Matsumoto, H. Kojima, Y. Omura, M. Moncuquet and J.-E. Wahlund, Electric fields in the Hermean environment, *Adv. Space Res.*, in press, 2005.
9. C. T. Russell, ULF waves in the Mercury magnetosphere, *Geophys. Res. Lett.* **16** (1989) 1253–1256.
10. L. G. Blomberg, Mercury's magnetosphere, exosphere, and surface: Low-frequency field and wave measurements as a diagnostic tool, *Planet. Space Sci.* **45** (1997) 143.
11. J. A. Slavin, J. C. J. Owen, J. E. P. Connerney and S. P. Christon, Mariner 10 observations of field-aligned currents at Mercury, *Planet. Space Sci.* **45** (1997) 133.
12. R. Grard, Photoemission on the surface of Mercury and related electrical phenomena, *Planet. Space Sci.* **45** (1997) 67–72.
13. L. G. Blomberg and J. A. Cumnock, On electromagnetic phenomena in Mercury's magnetosphere, *Adv. Space Res.* **33** (2004) 2161–2165.
14. C. T. Russell, D. N. Baker and J. A. Slavin, The magnetosphere of Mercury, in *Mercury* (University of Arizona Press, 1988), pp. 514–561.
15. S. Orsini, A. Milillo, E. De Angelis, A. M. Di Lellis, V. Zanza and S. Livi, Remote sensing of Mercury's magnetospheric plasma environment via energetic neutral atoms imaging, *Planet. Space Sci.* **49** (2001) 1659–1668.

PLASMA/RADIO WAVE OBSERVATIONS AT MERCURY BY THE BEPICOLOMBO MMO SPACECRAFT

H. MATSUMOTO*,†, J.-L. BOUGERET‡, L. G. BLOMBERG§, H. KOJIMA†, S. YAGITANI¶, Y. OMURA†, M. MONCUQUET‡, G. CHANTEUR‖, Y. KASABA**, J.-G. TROTIGNON††, Y. KASAHARA¶ and BEPICOLOMBO MMO PWI TEAM‡‡

†*RISH, Kyoto University, Japan*
‡*LESIA-Observatoire de Paris, France*
§*Alfvén Laboratory, KTH, Sweden*
¶*Kanazawa University Japan*
‖*CETP/IPSL, France*
**ISAS/JAXA, Japan*
††*LPCE, CNRS, France*
**matsumot@rish.kyoto-u.ac.jp*

The BepiColombo Mercury Magnetospheric Orbiter (MMO) spacecraft comprises the plasma and radio wave observation system called Plasma Wave Investigation (PWI). The PWI is designed and developed in collaboration between Japanese and European scientists. Since plasma/radio wave receivers were not installed in the former spacecraft, Mariner 10, which observed the planet Mercury, the PWI onboard the MMO spacecraft will provide the first plasma/radio wave data from Mercury orbit. It will give important information for studies of energy exchange processes in the unique magnetosphere of Mercury characterized by the interaction between the relatively large planet without ionosphere and the solar wind with high dynamic pressure. The PWI consists of three sets of receivers (EWO, SORBET, and AM^{2}P), connected to two sets of electric field sensors (MEFISTO and WPT) and two kinds of magnetic field sensors (LF-SC and DB-SC). The PWI will observe both waveforms and frequency spectra in the frequency range from DC to 10 MHz for the electric field

*Corresponding author.
‡‡The BepiColombo MMO PWI team consists of the members participating from the following institutions and universities:
[Japan] RISH/Kyoto University; Kanazawa University; ISAS/JAXA; Ehime University; Kyoto Sangyo University; Toyama Prefectural University; ROIS/National Institute of Polar Research; Graduate scool of Science/Tohoku University
[France] LESIA-Observatoire de Paris; CETP/IPSL; LPCE/CNRS, Orléans
[Sweden] Alfvén Laboratory, Royal Institute of Technology; Swedish Institute of Space Physics
[Finland] University of Oulu; Finnish Meteorological Institute
[Norway] University of Oslo
[Netherlands] ESA/RSSD, Noordwijk
[Hungary] Eötvös University

and from 0.1 Hz to 640 kHz for the magnetic field. In the present paper, we demonstrate the scientific objectives of plasma/radio wave observation around Mercury. Further, we introduce the PWI system, which is designed to meet the scientific objectives in the BepiColombo MMO mission.

1. Introduction

The Mariner 10 encounters with the planet Mercury strongly suggest the existence of an intrinsic magnetic field, which leads to the formation of the magnetosphere through its interaction with the solar wind plasma.[1] Since space plasmas are essentially collisionless, the observations of plasma/radio waves in the Mercury magnetosphere provide important information in studying the energy/momentum exchange processes in the unique plasma environments characterized by the weak intrinsic magnetic fields and the high dynamic pressure of the solar wind.

The Plasma Wave Investigation (PWI) onboard the BepiColombo Mercury Magnetospheric Orbiter (MMO) addresses a wealth of fundamental scientific questions pertaining to the magnetosphere and exosphere of planet Mercury, the solar wind at Mercury's location, and solar radiation from the view point of Mercury (see Fig. 1). The PWI measurements will provide

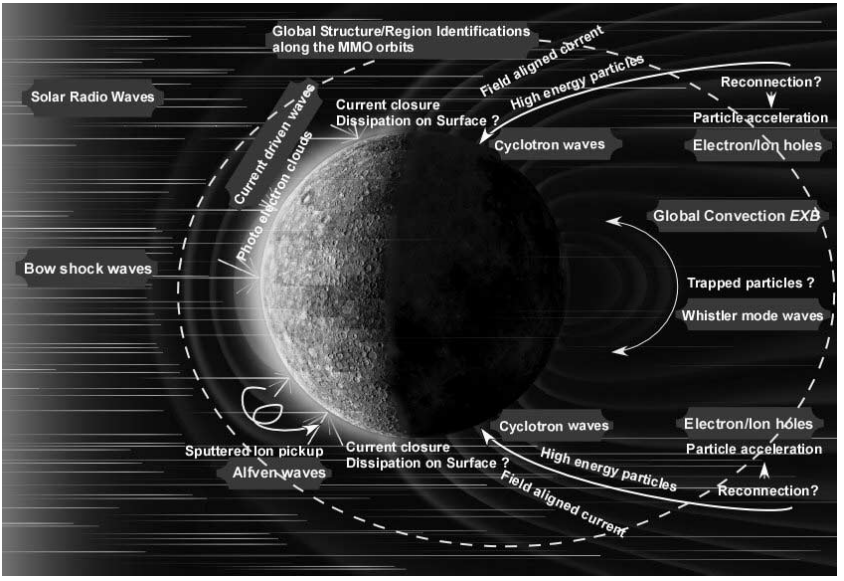


Fig. 1. Schematic views of Mercury's magnetosphere and representative targets of PWI. The dashed circle represents the orbit of the MMO spacecraft.

new information on the structure, dynamics, and physical processes related to energy transfer and scale coupling through wave-particle interactions. Observations of solar radio activity will be used as the solar activity level in the heliocentric sector facing Mercury by stereoscopic studies combined with similar measurements in Earth orbit. In the present paper, we demonstrate the scientific objectives of plasma/radio wave observations at Mercury and introduce the design of the PWI onboard the BepiColombo MMO spacecraft.

2. Scientific Objectives

2.1. *Structure of the magnetosphere*

2.1.1. *Identification of regions and boundaries*

The signatures of plasma waves depending on the local plasma parameters and processes enable us to identify regions and boundaries of the magnetosphere. Boundaries will also be identified by the variations of plasma densities and temperatures obtained by plasma wave measurements as well as plasma drifts obtained by DC electric field measurements.

2.1.2. *Global convection*

Because of the differences in plasma sources and sinks, the global plasma circulation in the magnetosphere may be quite different from Earth's. The DC electric field governs such dynamics. Since the gyro radii of charged particles are large compared to the size of the magnetosphere, kinetic effects associated with non-gyrotropy, stochastic diffusion, etc. may dominate. Plasma waves associated with those processes will suggest the uniqueness of the Mercury magnetosphere.

2.1.3. *Global profile of density and temperature*

Mapping of electron density and temperature in the solar wind, the magnetosphere, and the exosphere will provide fundamental input on the chemistry of ionized species (e.g., Na, K, O) in Mercury's environment and for dynamic modeling of the magnetosphere. Plasma frequencies are expected in the frequency range from a few 10 s of kHz to 300 kHz in solar wind and of a few kHz in the magnetosphere. They are equivalent to the plasma number densities about from 10 to 1,000 cm^{-3} in the solar wind and of about 0.1 cm^{-3} in the magnetosphere.

2.1.4. *Current closure/dissipation at low altitude*

Field-aligned currents are a fundamental mediator of energy and stress in the magnetosphere. There is currently no clear understanding of how field-aligned currents close at low altitude. Combined electric and magnetic field measurements will provide new information on the closure mechanisms.[2]

2.1.5. *Plasma wave propagation*

Propagation of whistler waves along the magnetic field lines guided by plasma density structures will provide information on the magnetic field and density structure of the Mercury magnetosphere. Further, the whistler mode waves resonating with extremely high energy electrons suggest the existence of energetic electrons associated with acceleration phenomena such as substorms.[3]

2.2. *Dynamics of the magnetosphere*

2.2.1. *Solar-wind–magnetosphere coupling*

Mercury does not have any appreciable ionosphere.[4] At Earth, the penetration of the solar wind electric field is weakened by a feedback mechanism from the ionosphere. Such mechanism might be less efficient at Mercury. The role of the ionosphere can be clarified by comparison to the Earth. Further, by measuring the electric field as well as the magnetic field of Eigen-frequency waves (i.e., field-line resonances), their phase difference will yield information on the reflective properties of the ionosphere and/or planetary surface.

2.2.2. *Response to solar wind*

Since the Alfvén Mach number can be less than unity, we may find quite different responses of the magnetosphere, such as formation of a slow shock. Also, distance variations along Mercury's eccentric orbit result in variations of a factor 2 of the solar wind power. In the case the auroral radio emissions exist, their power will provide a diagnostic of the global magnetospheric response.

2.2.3. *Magnetosphere–exosphere coupling*

Due to the lack of a thick atmosphere, plasma supply mechanisms from Mercury can be quite different from those at Earth. Mercury's exosphere could modify the magnetospheric plasma conditions. Plasma waves will

contribute to the understanding of the plasma supply and loss mechanisms associated with ion waves.

2.2.4. *Search for transient radiation belts in the hectometric radio range*

No stable radiation belts are expected around Mercury. But, synchrotron transients could be produced from transient radiation belts with MeV electrons. Synchrotron radiation might peak near a few MHz.

2.3. *Energy transfer and scale coupling*

2.3.1. *Nature of substorms*

Because of the much smaller magnetosphere, the time scale of substorms at Mercury will be shorter. Thus, a better understanding of the loading–unloading process will be gained by comparison with substorms at Earth. Electric fields associated with fast reconfigurations of the magnetosphere will accelerate charged particles.

2.3.2. *Reconnection*

Reconnection of the magnetosphere plays important roles in energy transfer from the solar wind to the magnetosphere. The process of reconnection involves coupling of microscale and mesoscale phenomena such as various types of beam or current driven instabilities resulting in heating and acceleration of electrons and ions, which excite various types of plasma waves through nonlinear wave-particle interaction, dissipating the energy of the accelerated particles.

2.3.3. *Identification of auroral processes*

Current closure may require acceleration of particles along the magnetic field. Thus, a large electric field enhancement similar to the auroral acceleration region may be expected. In the case the auroral radio emissions exist, their source regions are expected to be very close to the planet polar surface. Therefore, we can estimate their frequency range is around 20 kHz. The scaling law predicts an auroral radio power of 10^{6-7} W.

2.4. *Wave-particle interactions*

2.4.1. *Nonlinear kinetic processes*

In the dayside, the solar wind plasma sometimes interacts with the photoelectron cloud above the Mercury surface directly, and leads to current

driven instabilities. Pick-up ions conveyed by the solar wind into the night side magnetosphere can interact with the magnetospheric plasma, which leads to ion–ion beam instabilities. And, high-energy electron beams in the magnetotail generate Langmuir waves and electron phase space holes which travel in the electron beam direction. The spatial distribution of electron hole propagation allows us to estimate the location of reconnection.

2.4.2. *Non-gyrotropic effects*

The sputtered ions from the surface that are picked up are associated with Alfvén or ion cyclotron waves. Those waves may evolve to large amplitude non-linear waves, as at comets. Some complexities arise from the existence of an intrinsic magnetic field at Mercury. Propagation analysis of Alfvén waves allows us to study the connection of the solar wind with the magnetosphere.

2.4.3. *Physics of foreshock*

Langmuir waves are generated by backstreaming electrons reflected at the bow shock. Electromagnetic $2f_{\mathrm{pe}}$ radiation is produced from those, by non-linear coupling or mode conversion, which is the common process with solar radiations. *In situ* measurements will help discriminate the proposed mechanisms.

2.5. *Solar radio emissions and diagnostics*

2.5.1. *Space weather observation*

Monitoring of solar radiation up to 10 MHz (types II and III radio bursts) allows us to correlate them with the Mercury magnetospheric response in order to create a solar activity index from the view point of Mercury, providing context information for the *in situ* measurements.

2.5.2. *Stereoscopic observation*

Combined with other spacecraft, the MMO can perform stereoscopic observations of solar radiation from large-scale plasma structures such as CMEs. Such studies enable us to estimate the global properties of the solar wind in the inner heliosphere.

2.5.3. *Interplanetary shocks*

In-situ observations of interplanetary shocks at 0.3–0.47 AU provide a unique opportunity to identify different characteristics of more energetic radio sources not observed at Earth.

3. Instrument Design

To meet the science objectives, PWI is designed as a sophisticated plasma/radio wave receiver system with high sensitivity electric and magnetic sensors. The PWI has two pairs of electric field sensors, Wire-Probe anTenna (WPT) and Mercury Electric Field *In-situ* TOol (MEFISTO) and two types of magnetic field sensors, Low-frequency search coil (LF-SC) and dual band search coil (DB-SC). The configuration of the PWI sensors is shown in Fig. 2. (Note that MEFISTO and WPT consist of the sensor units and their peripheral electronics such as the deployment system. We address each sensor unit as MEFISTO-S and WPT-S, respectively.) Using these sensors, the PWI covers a very wide frequency range, DC to 10 MHz for electric field and 0.1 Hz to 640 kHz for magnetic field.

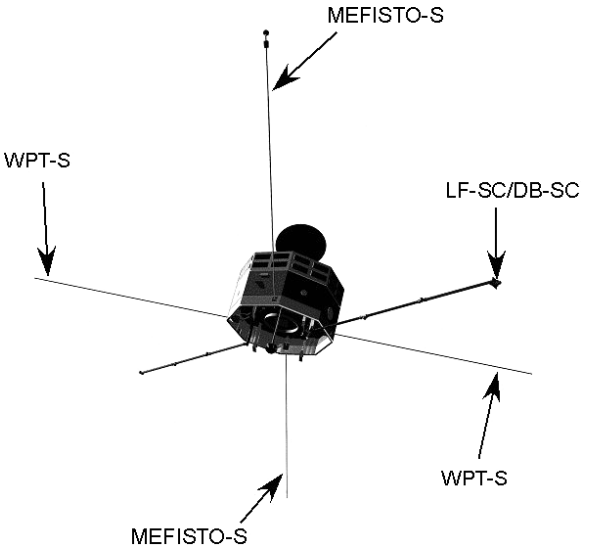


Fig. 2. External view of the MMO spacecraft and PWI sensors.

The PWI system consists of three receiver components as follows:

(1) EWO: Composite of electric field detector (EFD), waveform capture (WFC), and onboard frequency analyser (OFA).
(2) SORBET: "Spectroscopie des Ondes Radio & du Bruit Electrostatique Thermique" (Spectroscopy of radio waves and thermal electrostatic noise).
(3) Active measurement of mercury's plasma (AM^2P).

These receiver components are connected to the corresponding sensors through the sensor preamplifiers. Figure 3 shows a block diagram.

3.1. *Sensors for electric field and magnetic field*

Two pairs of electric field sensors are orthogonally installed in the spacecraft spin plane. They will be deployed after Mercury orbit insertion. The total length after deployment is 32 m tip-to-tip for each of them. One pair of electric field sensor is the WPT based on heritage from the PANT system for EFD and PWI onboard the Geotail spacecraft.[5,6] It consists of a conductive

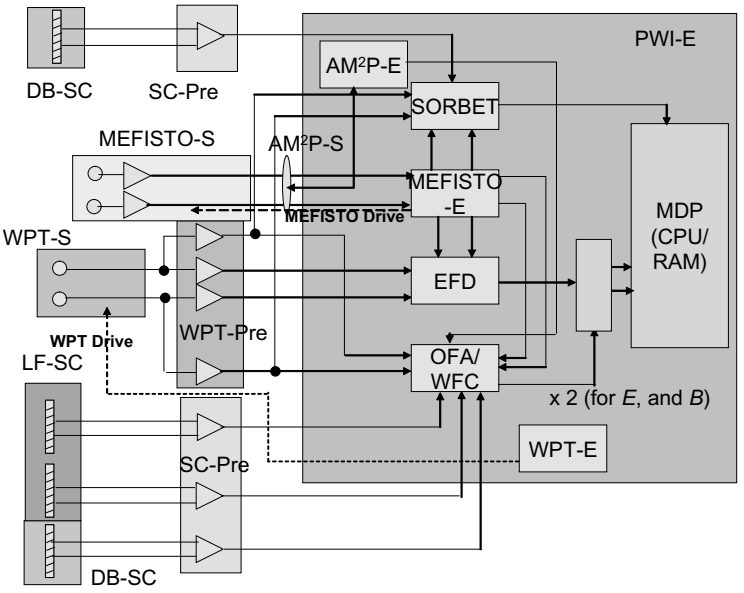


Fig. 3. Block diagram of the PWI system.

wire with a sphere at the end of the wire. It is directly fed to the pre-amplifiers located inside the spacecraft. It responds to frequencies from DC to 10 MHz. The other pair of electric field sensor is MEFISTO, based on heritage from the EFW aboard the Cluster spacecraft.[7] It consists of an extendable boom, a conductive short wire with a sphere at the end, and a pre-amplifier box. The preamplifier box is attached at the end of the extendable boom and the short wire with the sphere is deployed from that preamplifier box. The antenna has good response in the frequency range below 3 MHz.

Two sets of magnetic field sensors are installed at the end of the coilable mast, 5 m in length, in order to avoid electromagnetic contamination radiated from the spacecraft. The LF-SC is two-axial search coils. They are mounted in the spacecraft spin plane, and observe electromagnetic waves in the frequency range below 20 kHz. The dual-band search coil (DB-SC) is a single-axis search coil. It covers two different frequency ranges (one is below 20 kHz, and the other is up to 640 kHz). It is aligned with the spacecraft spin axis. Two different coils with different turn numbers are wound on one axis core. These two different coils provide frequency coverage in two different ranges.

The sensitivities of electric field and magnetic field sensors are shown in Fig. 4.

3.2. *Receivers for electric field and magnetic field*

EWO-EFD is dedicated to observing DC and low-frequency electric fields. It is connected to the WPT and MEFISTO electric field sensors. The frequency coverage of the EWO-EFD is from DC to 30 Hz, and its dynamic range is 110 dB. The EWO-EFD also has a single probe system for measuring spacecraft potentials.

The EWO-WFC/OFA receiver components focus on plasma wave observations in the medium frequency range up to 360 kHz for electric field, and up to 20 kHz for magnetic field. Waveforms picked up by electric/magnetic field sensors are directly sampled (WFC data) and stored in onboard memories. The stored waveform data are used for calculating FFT spectra (OFA data). While the OFA data are sent to Earth in the nominal mode, the WFC waveform data are sent after the data selection and data compression procedures in the high bit rate mode. SORBET has two main functions. One is as thermal noise receiver (TNR), which measures *in situ* plasma temperatures and densities by monitoring continuously the plasma thermal noise.

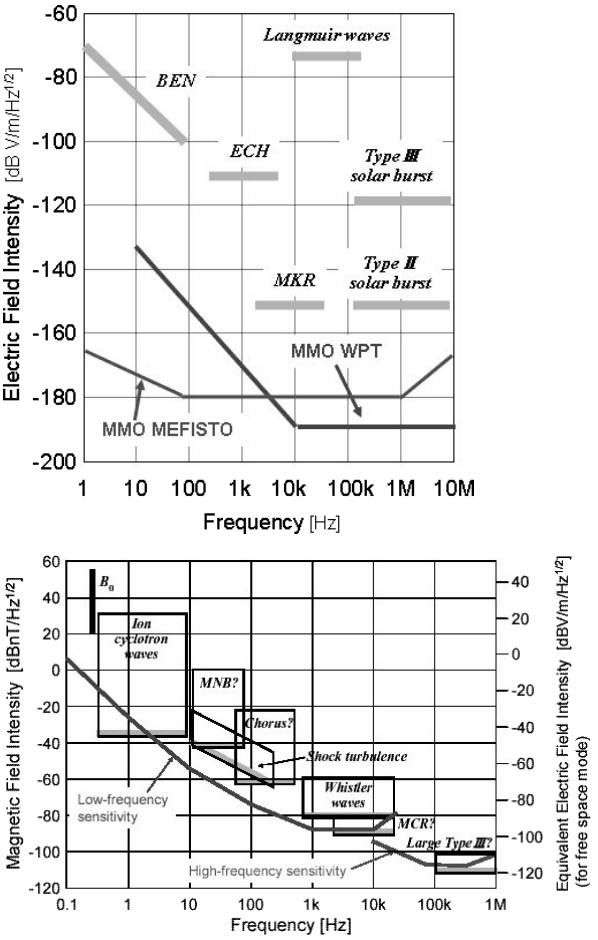


Fig. 4. Sensitivities and expected intensities of natural waves. Top: electric field. *Bottom*: magnetic field: Broadband electrostatic noise (BEN), Electron cyclotron harmonic wave (ECH), mercury kilometric radiation (MKR), magnetic noise burst (MNB) and Mercury continuum radiation (MCR).

Voltage measurements from the electric antennas are processed first into spectral data by a wavelet-like method and then the plasma temperatures and densities are deduced by fitting a quasi-thermal noise spectrum model. SORBET is based on an heritage of the HF receiver of the RPWS experiment onboard Cassini, from which the quasi-thermal noise spectroscopy is now providing electron temperature and density in the inner magnetosphere of Saturn.[8] Another function is to observe very high-frequency radio

waves up to 10 MHz for electric field and up to 640 kHz for magnetic field. SORBET is connected to the WPT, MEFISTO, and DB-SC sensors.

The AM^2P measures the antenna impedance of the MEFISTO electric field sensors. It feeds a series of current pulses onto the electric field sensors and monitors the voltage response. The obtained data sets provide us with information on the antenna impedance, which is used for precise calibration of the observed electric field data. Since the antenna impedance strongly depends on plasma parameters, the data sets of AM^2P are also useful for measuring *in situ* plasma densities and temperatures. The function of measuring the WPT antenna impedance is provided by a part of EWO.

The PWI electronics board will be installed into the MDP-2 (Mission Data Processor) chassis developed by ISAS/JAXA, as the MMO system team. Data processing unit (DPU) in the MDP-2 will manage the data processing and command control for PWI and will employ event detection algorithm to find expected or unknown events and will distribute triggering signal among other experiments.

MDP-2 will also be in charge of controlling the data flows and the operation of the MGF-I (magnetic field-inboard), MSASI (sodium exosphere imager) and MDM (dust monitor).

4. Expected Performance

4.1. *Electric field detection: from DC to 10 MHz*

WPT and MEFISTO each measure orthogonal components of the electric field in the spacecraft spin plane using complementary sensor designs. WPT offers superior sensitivity at high frequencies whereas MEFISTO offers superior sensitivity at low frequencies. For all intermediate frequencies both sensors offer excellent sensitivity and the electric field vector in the spin plane can be reconstructed. Weak radio waves will mainly be registered by WPT whereas weak low-frequency fluctuations will mainly be detected by MEFISTO. For DC electric fields, both sensors provide sufficient sensitivity for all scientific objectives.

For DC electric fields (sampled by EWO), the PWI will cover the dynamic range and resolution of the DC electric field expected in Mercury's magnetosphere and in the solar wind, less than 500 mV/m with resolution of 0.015 mV/m. The sensitivities for AC electric field measurements (by EWO and SORBET) are summarized in Fig. 4.

4.2. *Magnetic field detection: from 0.1 Hz to 640 kHz*

Two search coil sensors, LF-SC and DB-SC, are combined to measure AC magnetic fields: tri-axial vector components for low frequencies (0.1 Hz–20 kHz), and one-axial component along the spacecraft spin axis for high frequencies (10 kHz–640 kHz). They offer excellent sensitivity (Fig. 4) and dynamic range (>80 dB) that will cover all expected electromagnetic waves.

4.3. *Spacecraft potential*

The PWI electric field sensors have sufficient dynamic range for spacecraft potential (±100 V) to operate at all electron densities and temperatures expected in Mercury's magnetosphere and in the solar wind.

4.4. *Electron density and temperature*

Continuous surveying of electron density and temperature will be performed by SORBET, using the technique of QTN spectroscopy[8,9] in the frequency range 2.5–640 kHz. The performance may be roughly summarized as follows: The electron density will be measured from 0.1 to ∼6,000 cm^{-3} (with ∼1% accuracy) and the core temperature from 0.1–100 eV (with ∼10% accuracy). SORBET will also provide a diagnostic of the suprathermal electron population (which is ubiquitous in the solar wind).

4.5. *Calibration of electric field antenna impedance*

The precise antenna impedance is very important for calibration of the observed waveform and spectrum data. Since we have two different types of electric field antennas, we include two instruments for the onboard measurement of antenna impedance. AM^2P and EWO serve to measure the antenna impedance of MEFISTO and WPT, respectively. Both components feed a known source signal to the antennas. By collecting the responses of MEFISTO and WPT, we can calculate their frequency-dependent antenna impedances.

5. Summary

Up to now, the MMO is scheduled to be launched in 2012, and initiate observations at Mercury from 2016. At this stage, we extensively discuss the development of the PWI system, so that it can provide excellent science

data and advance our knowledge of Mercury's plasma environment. The BepiColombo MMO mission provides a unique and important opportunity for the progress of magnetospheric studies by the first detailed comparison of the planetary magnetosphere with the terrestrial one. The PWI realizes detailed diagnostics of macroscopic and microscopic phenomena in the Mercury magnetosphere in the view point of plasma/radio waves, which no one has observed before.

Acknowledgments

This work was supported by JAXA (Japan), CNES (France), SNSB (Sweden), NSC (Norway), and HSO (Hungary).

References

1. N. Ness, K. W. Behannon, R. P. Lepping, Y. C. Whang and K. H. Schatten, Magnetic field observations near Mercury: Preliminary results from Mariner 10, *Science* **185** (1974) 151–160.
2. L. G. Blomberg and J. A. Cumnock, On electromagnetic phenomena in mercury's magnetosphere, *Adv. Space Res.* **33** (2004) 2161–2165.
3. D. Summers, R. M. Thorne and F. Xiao, Relativistic theory of wave-particle resonant diffusion with application to electron acceleration in the magnetosphere, *J. Geophys. Res.* **103** (1998) 20487–20500.
4. C. T. Russell, D. N. Baker and J. A. Slavin, The magnetosphere of Mercury, in *Mercury*, eds. F. Vilas, C. R. Chapman and M. S. Matthews (The University of Arizona Press, Arizona, 1988).
5. K. Tsuruda, H. Hayakawa, M. Nakamura, T. Okada, A. Matsuoka, F. S. Mozer and R. Schmidt, Electric field measurements on the GEOTAIL Satellite, *J. Geomag. Geoelectr.* **46** (1994) 693–711.
6. H. Matsumoto, I. Nagano, R. R. Anderson, H. Kojima, K. Hashimoto, M. Tsutsui, T. Okada, I. Kimura, Y. Omura and M. Okada, Plasma wave observations with GEOTAIL Spacecraft, *J. Geomag. Geoelectr.* **46** (1994) 59–95.
7. G. Gustafsson, M. André, T. Carozzi, A. I. Eriksson, C.-G. Fälthammar, R. Grard, G. Holmgren, J. A. Holtet, N. Ivchenko, T. Karlsson, Y. Khotyaintsev, S. Klimov, H. Laakso, P.-A. Lindqvist, B. Lybekk, G. Marklund, F. Mozer, K. Mursula, A. Pedersen, B. Popielawska, S. Savin, K. Stasiewicz, P. Tanskanen, A. Vaivads and J.-E. Wahlund, First results of electric field and density observations by Cluster EFW based on initial months of operation, *Ann. Geophys.* **19** (2001) 1219.
8. M. Moncuquet, A. Lecacheux, N. Meyer-Vernet, B. Cecconi and W. S. Kurth, Quasi-thermal noise spectroscopy in the inner magnetosphere of Saturn with Cassini/RPWS: Electron temperatures and density, *Geophys. Res. Lett.* **32** (2005) L20S02, doi:10.1029/2005GL022508.

9. N. Meyer-Vernet, S. Hoang, K. Issautier, M. Maksimovic, R. Manning, M. Moncuquet and R. Stone, Measuring plasma parameters with thermal noise spectroscopy, in *Geophysical Monograph* 103: *Measurements Techniques in Space Plasmas*, eds. E. Borovsky and R. Pfaff (1998), pp. 205–210.

LOW ENERGY ION OBSERVATION BY MERCURY MAGNETOSPHERIC ORBITER: MMO

YOSHIFUMI SAITO

Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency
3-1-1 Yoshinodai, Sagamihara, Kanagawa, 229-8510, Japan
saito@stp.isas.jaxa.jp

DOMINIQUE DELCOURT

CETP-IPSL, 4 avenue de Neptune, 94107, Saint-Maur des Fossés, France
dominique.delcourt@cetp.ipsl.fr

ANDREW COATES

MSSL, University College London, Holmbury
St Mary, Dorking RH5 6NT, UK
ajc@mssl.ucl.ac.uk

According to the previous satellite observation (Mariner10 fly-by), Mercury has a magnetosphere with its own strong intrinsic magnetic field. In order to elucidate the detailed plasma structure and dynamics around Mercury, an orbiter BepiColombo Mercury magnetospheric orbiter (MMO) is planned to be launched in the timeframe between 2012 and 2013 as a joint mission between ESA and ISAS/JAXA. For measuring low energy ions, two sensors Mercury ion analyzer (MIA) and Mercury ion mass spectrum analyzer (MSA) will be onboard the MMO. MIA measures energy spectrum (5 eV/q–30 keV/q) of ions around Mercury as well as solar wind ions while MSA measures mass discriminated energy spectrum (5 eV/q–40 keV/q) of ions around Mercury. MIA is a toroidal top-hat type electrostatic analyzer with dynamic range as wide as 10^6, while MSA consists of a similar electrostatic analyzer followed by a time-of-flight section for mass discrimination. In order to realize the wide dynamic range, electrical sensitivity control and attenuation grid are applied at the same time. The thermal environment around Mercury is so severe that the thermal design of the instrument is very important. Each sensor should have its own thermal shield in order to minimize the thermal input and to maintain the sensor temperature within an acceptable range.

1. Introduction

Mercury is one of the least explored planets in our solar system. No spacecraft has visited Mercury since Mariner 10 made three flybys two in 1974 and one in 1975. Although nearly 30 years have passed since the

Mariner-10 Mercury fly-bys, most aspects of Mercury remain unknown. Two missions to explore Mercury are currently in progress: MESSENGER (MErcury Surface, Space ENvironment, GEochemistry, and Ranging) and BepiColombo. MESSENGER is on its way to Mercury after a successful launch on August 3, 2004. BepiColombo is a joint mission between ESA and ISAS/JAXA. Two separate spacecraft, the Mercury planetary orbiter (MPO), and the Mercury magnetospheric orbiter (MMO) are planned to be launched in the timeframe between 2012 and 2013. The MPO will study the surface and internal composition of the planet, and the MMO will study the detailed plasma structure and dynamics around Mercury. The combination between MMO and MPO will provide us with significant scientific discoveries on Mercury.

Although the plasma scientific payload of Mariner 10 was very limited, it made a very important discovery that Mercury possesses an intrinsic magnetic field, whose intensity was in a very intriguing range in terms of comparative planetary magnetospheres.[1] The dominance of the dipole term in the spherical harmonic expansion of Mercury's magnetic field suggests that the interaction between the solar wind and Mercury's magnetosphere should be "Earth-like", in contrast to the cases of Mars and Venus where the planetary magnetic fields have only local effects on the interaction. On the other hand, because of its small size and gravity, Mercury has very different environmental characteristics compared to the Earth. The tenuous atmosphere and the possible dominance of the heavy elements supplied from the surface[2,3] may provide a unique opportunity to assess the relative importance of the solar wind source and the mixing mechanisms of magnetospheric plasma. This is difficult to do at the Earth where the proton component of the ionospheric plasma cannot be uniquely distinguished from the solar wind. The tiny nature of Mercury's magnetosphere requires that totally new ideas be applied to the basic concepts that have been applied to the Earth's magnetosphere. The large gyro-radii of energetic and thermal plasma components (relative to the magnetospheric size) suggests that ideal MHD may no longer apply; we anticipate that the scale-coupling between the microscopic and macroscopic plasma processes, which has been highlighted by recent theoretical work, has a global and dominant importance in Mercury's magnetosphere. As Mariner 10 observations showed, Mercury's magnetosphere is rich in energetic particles of the keV order and higher. Exploration of the particle environment of Mercury is very promising since the spatial distribution and temporal evolution of charged particles within the magnetosphere provide important information on the dynamics

of Mercury's magnetosphere. Mariner 10 observations suggested that acceleration events reminiscent of the Earth's "substorms" occur in Mercury's magnetosphere.[4] Detailed analyses of particle acceleration events by MMO will place significant constraints on some of the acceleration mechanisms discussed so far to explain particle acceleration in Earth's magnetosphere, particularly in terms of magnetic reconnection. In order to fulfill the science objectives, a comprehensive instrument package was proposed for plasma, high energy particle and energetic neutral atom measurements with sufficiently high time resolution, wide energy range, wide dynamic range, wide angle coverage, and high mass resolution under restricted thermal, weight, and power consumption conditions. In this paper, we will briefly introduce the Mercury ion analyzer (MIA), one of the low energy ion sensors.

2. Mercury Ion Analyzer

The scientific objectives of low-energy ion measurement on Mercury orbit are to understand: (1) structure of the Mercury magnetosphere, (2) plasma dynamics of the Mercury magnetosphere, (3) Mercury — solar wind interaction, (4) atmospheric abundances, structure, and generation/loss process, and (5) solar wind between 0.3 and 0.47 AU. In order to enable the required measurements, MIA should measure 3-dimensional (3D) distribution function of solar wind ions around Mercury (0.3–0.47 AU), and Mercury magnetospheric ions simultaneously. According to Mukai *et al.*,[5] the predicted proton fluxes ranges 10^6–10^{12}/cm^2/s/str/keV. Accordingly, MIA has to have very wide dynamic range of $>10^6$. The MIA realizes this wide dynamic range by using electrical sensitivity control and attenuation grid simultaneously. The energy range of MIA is 5 eV/q–30 keV/q. This energy range will cover predicted energy distribution of solar wind ions around Mercury, Mercury cold plasma sheet ions, and Mercury hot plasma sheet ions. The maximum number of energy step is 64 that are necessary for distinguishing solar wind alpha particles from solar wind protons. The angular resolution for measuring solar wind ions is 5.625°, while the angular resolution for measuring Mercury magnetospheric ions is 11.25° or 22.5°, depending on the allocated capacity of the telemetry data. This angular resolution is enough for fulfilling required measurements including non-Maxwellian characteristics of the ion distribution function predicted in the Mercury magnetosphere and solar wind — Mercury magnetosphere interaction.

According to our knowledge of the Earth's magnetosphere, full 3D measurements of low-energy ions with high time resolution are indispensable

for understanding the structure and dynamics of the magnetosphere. Since no low-energy ion data have been obtained around Mercury, low-energy ion data obtained by MIA together with Mercury Ion mass spectrum analyzer (MSA) on MMO will provide us with unique opportunity to understand detailed structure and dynamics of the Mercury magnetosphere.

3. Measurement Principle of MIA

The MIA basically employs a method of a top hat electrostatic analyzer with toroidal deflectors as illustrated in Fig. 1. The center of FOV is designed to be slightly inclined upward from the horizontal plane perpendicular to the axis of rotational symmetry. With the spin motion of spacecraft, 3D ion distribution function is observed. The inner toroidal electrode is supplied with high voltage swept between 0 V and −5 kV. Ions coming through the collimator are attracted down toward the inner electrode by the action of the applied potential. Only the ions with specific energy range can further travel down to the exit of the electrodes. The ions passing through the toroidal deflectors enter to Z-stack Micro-channel plate (MCP) and are intensified to detectable charge pulses. In front of the MCP stack, a grid supplied with slightly lower voltage (−100 V) than the front surface of the MCP is placed in order to increase the detection efficiency of ions. Finally, the charge pulses are received by 41-channel discrete anode. The positions where the charge pulses are detected correspond to the incident azimuthal directions of the ions.

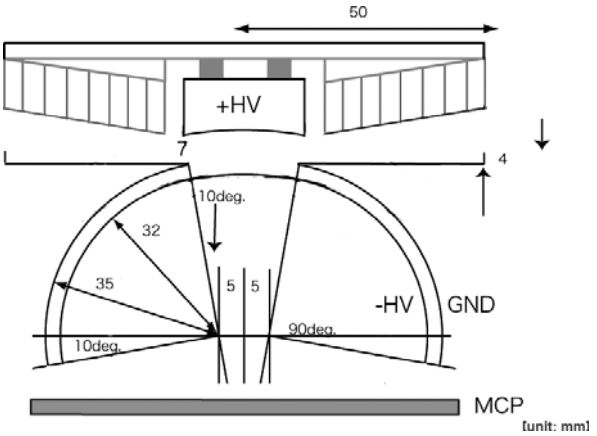


Fig. 1. Measurement principle of MIA.

Table 1. Specifications of MIA.

Energy resolution	$\Delta E/E \sim 15\%$ (FWHM, high g-factor mode) $\Delta E/E \sim 3\%$ (FWHM, low g-factor mode)
Field of view	$10^\circ \times 360^\circ$ (high g-factor mode) $5^\circ \times 360^\circ$ (low g-factor mode)
Angular resolution	$5.625^\circ \times 5.625^\circ$ (min.: solar wind) $22.5^\circ \times 22.5^\circ$ (max. Mercury ion)
Time resolution	2 s/3D distribution function
g-factor	⟨high g-factor mode⟩ 5×10^{-4} cm^2 sr keV/keV at $10^\circ \times 22.5^\circ$ 5×10^{-5} cm^2 sr keV/keV at $10^\circ \times 22.5^\circ$ (low g-factor solar wind sector) ⟨low g-factor mode⟩ 1×10^{-5} cm^2 sr keV/keV (min.) at $5^\circ \times 5.625^\circ$ 5×10^{-7} cm^2 sr keV/keV (min.) at $5^\circ \times 5.625^\circ$ (low g-factor solar wind sector)

The estimated dynamic range of the low energy ion flux around Mercury including both intense solar wind ions and weak magnetospheric ions is as wide as 10^6.[5] In order to measure both solar wind ions without saturation and Mercury magnetospheric ions with enough counting statistics, MIA has a function to change g-factor electrically. Sensitivity of the analyzer is controlled by applying high voltage to the "top hat" part. The center of the "top hat" part is insulated from the surrounding structures. By applying high voltage between 0 V and +5 kV, g-factor can be reduced down to 1/50. In addition to the electrical g-factor control, attenuation grid (10% transmission) is placed at limited sector of the exit part of the analyzer in order to further reduce geometrical factor for solar wind ion measurement. Table 1 summarizes the specifications of MIA.

4. Design of MIA

Figure 2 shows a 3D model of MIA. The MIA is divided into two parts, the analyzer part with toroidal deflectors and an ion detector, and the electronics with pre-amplifier, high-voltage power supply, circuits to control and monitor analyzer, and other circuits including spacecraft I/F. Both parts are integrated into one package. In order to reduce weight, some of the mechanical parts are made of magnesium alloy.

The solar radiation around Mercury is about 10 times more intense than around the Earth. Since solar UV is a significant source of the noise count, the following several steps to reduce solar UV are indispensable. (1) Toroidal

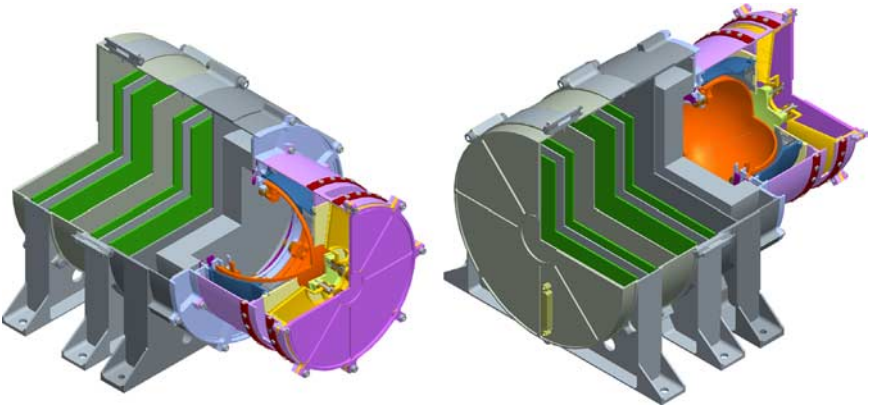


Fig. 2. 3D figure of MIA.

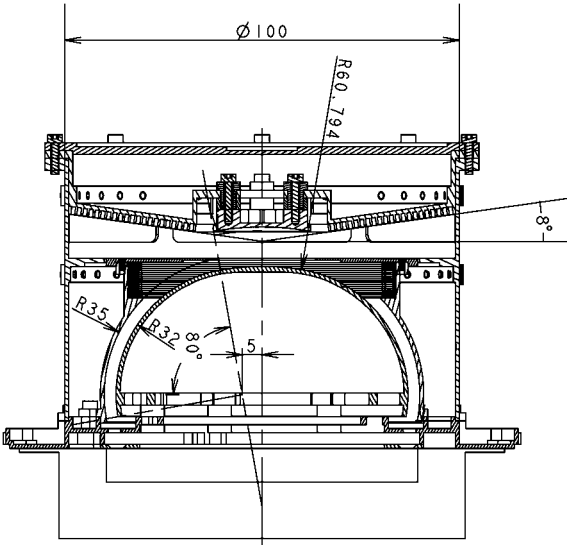


Fig. 3. Cross section of the analyzer part of MIA.

deflector plate is blackened. (2) Striation is made on the toroidal deflectors. (3) Light trap is made at the upper edge of the outer sphere. (4) Baffle structure is made at the entrance part of the analyzer. (5) Light attenuation grid (10% transmission) is placed at limited sector of the entrance. The light trap can be seen in Fig. 3 that shows cross section of the analyzer part of MIA.

Acknowledgments

The authors thank all the members of BepiColombo MMO Mercury plasma particle consortium (MPPC) for valuable discussion on the specifications of Plasma/Particle sensors.

References

1. N. F. Ness, K. W. Behannon, R. P. Lepping, Y. C. Wang and K. H. Schatten, Magnetic field observations near Mercury: Preliminary results from Mariner 10, *Science* **185** (1974) 151.
2. A. E. Potter and T. H. Morgan, Sodium and potassium atmospheres of Mercury, *Planet. Space Sci.* **45** (1997) 95.
3. T. H. Morgan and R. M. Killen, A non-stoichimetric model of the composition of the atmospheres of Mercury and Moon, *Planet. Space Sci.* **45** (1997) 81.
4. G. L. Siscoe, N. F. Ness and C. M. Yeates, Substorms on Mercury?, *J. Geophys. Res.* **80** (1975) 4359.
5. T. Mukai, K. Ogasawara and Y. Saito, An empirical model of the plasma environment around Mercury, *Adv. Space Res.* **33** (2004) 2166.

This page intentionally left blank

ICE ON THE MOON AND MERCURY

DANA H. CRIDER
Department of Physics, Catholic University of America
106 Driftwood Dr., Gibsonville, NC 27249, USA
crider@cua.edu

ROSEMARY M. KILLEN
Department of Astronomy, University of Maryland
College Park, MD 20740, USA
rkillen@astro.umd.edu

RICHARD R. VONDRAK
Solar System Exploration Division, NASA Goddard Space Flight Center
Greenbelt, MD 20771, USA
richard.vondrak@nasa.gov

Observational data suggest that the ice deposits in permanently shaded regions on the Moon are relatively impure, being highly mixed with the regolith. In contrast, the deposits on Mercury appear to be pure and covered by dry regolith. The differences in the deposits on the two bodies may be the result of differing weathering rates, their unique impact histories, or selection effects from the observation methods. We discuss the available data and results from modeling. We find that the most suitable explanation for the Mercury observations is that a comet deposited an ice layer at least 50 cm thick in the northern cold traps less than 50 Myr ago. The lunar model suggests that if a relatively pure ice layer exists, it is old and buried or was originally thin and is now impure.

1. Introduction

Both the Moon and Mercury have regions near their respective poles where the low obliquity of the planet combined with topography provides permanent shadow from direct sunlight. Thermal modeling[1] predicts that temperatures are low enough that water ice would be stable to sublimation in a permanently shadowed region (PSR) for billions of years.[2] The amount of ice near the surface is important as both a historical tracer of processes and conditions in the solar system and as a potential resource for solar system exploration.

Observational evidence supports the existence of volatiles in PSRs at both the Moon and Mercury,[3,4] however, the interpretation of the data as

water ice is neither unique nor conclusive.[5–7] This paper assumes that the observations are indicative of water ice and interpret them in that context. There are differences in the observed deposits at the Moon and Mercury indicating some diversity in the histories of the two bodies. In this paper, we compare the observations at the Moon and Mercury and model simulations of the deposits to understand the distinctions and their implications.

2. Volatile Evolution in Permanent Shadowed Regions

The current depth distribution of volatiles in a PSR is partly a result of the source of volatiles. The amount of time over which emplacement occurred is of particular importance. An episodic source of ice, e.g., a cometary impact, would deposit a relatively pure layer of water ice in a short time. In contrast, a continuous source of water ice, like migration of water vapor through the exosphere after release from the surface by micrometeoroid bombardment or solar wind sputtering, would slowly and steadily add volatiles to the PSR over time. Since there is always new material arriving on the Moon, the migrating water would be mixed into the regolith. The expected depth profile is a thick, impure deposit.

Space weathering also affects the depth distribution of volatiles in the PSR. Space weathering is a term for the collection of processes that modify the properties of a planetary surface from exposure to radiation and impact. After the volatiles condense in the PSR, they are retained neither completely nor without interaction with the surrounding regolith. Some volatiles are lost due to incident starlight, scattered sunlight, and sputtering.[8] However, impacts are the primary modifying factor by moving material both vertically and laterally. Each successive impactor excavates material in one locality and buries material in an adjacent locality. The depth of alteration varies with impactor size and energy. Small impactors are more frequent than large impactors; so most impacts affect only a small area. Thus, every piece of the surface has undergone its own unique impact history and will have a different concentration profile with depth. Material that may be vaporized from an impactor can recondense when it arrives in an adjacent location in the PSR.[9] In this way an existing ice layer gets mixed with the surrounding regolith with time. The crater production function has a very similar shape for Mercury and the Moon; however, Mercury's flux is 5.5 times greater than the Moon's.[10,11]

Our model simulates the evolution of volatiles in a column of regolith in a PSR.[12] It considers various sources of volatiles by slowly adding volatiles over time or by starting with a pure ice layer.[13] It predicts the depth

distribution of volatiles as a function of time by Monte–Carlo calculations of a large number of independent columns of regolith through time as space weathering processes modify them. The profile obtained by averaging over all Monte–Carlo runs of a single initial condition can be viewed as a large-area footpoint average of the depth distribution. The model is described in detail in Refs. 12–14. Although our previous work on the Moon has followed the evolution of H content, all lunar calculations are converted to equivalent H_2O concentration, [H_2O], for presentation in this paper.

3. The Observations

Mercury and the Moon have different signatures of ice in the polar regions. This is due in part to the techniques used to measure the regions. However, real differences in the contents of the PSR probably also exist.

One common observation is radar. Mercury's north polar region observed with the Arecibo radar shows evidence for possible water ice in PSRs[4] extending down to 72° latitude.[15] The polarization signature is consistent with relatively pure water ice of several tens of cm thickness covered by ∼30 cm of dry regolith (Personal communication). Whereas Mercury data indicate pure ice deposits buried by dry regolith; lunar data suggest that no pure ice is there.[16]

In addition, neutron measurements have been conducted in orbit of the Moon and are planned for Mercury. A decrease in epithermal neutron flux is an indicator of the presence of hydrogen, although the chemical form of the hydrogen cannot be determined.[3] Hydrogen concentration in approximately the top meter of the lunar regolith has been determined from Lunar Prospector Neutron Spectrometer (LPNS) measurements.[17] The highest hydrogen concentration, and thus the highest potential for water, is outside of the largest craters in northern hemisphere and inside small craters in the southern hemisphere. The data are consistent with <10% (H_2O) spread throughout top 1 m of regolith.[18,19]

3.1. *Mercury*

The Mercury observations put some strict constraints on the ice deposits there. Figure 1 shows example profiles from the Mercury simulations. Three solid lines show the evolution with time of an ice layer initially 50 cm thick. The thinnest line depicts the profile after 20 Myr. The medium and thick lines are from 50 to 80 Myr, respectively. The broken lines contrast profiles from ice layers of different initial thicknesses at 50 Myr for comparison with

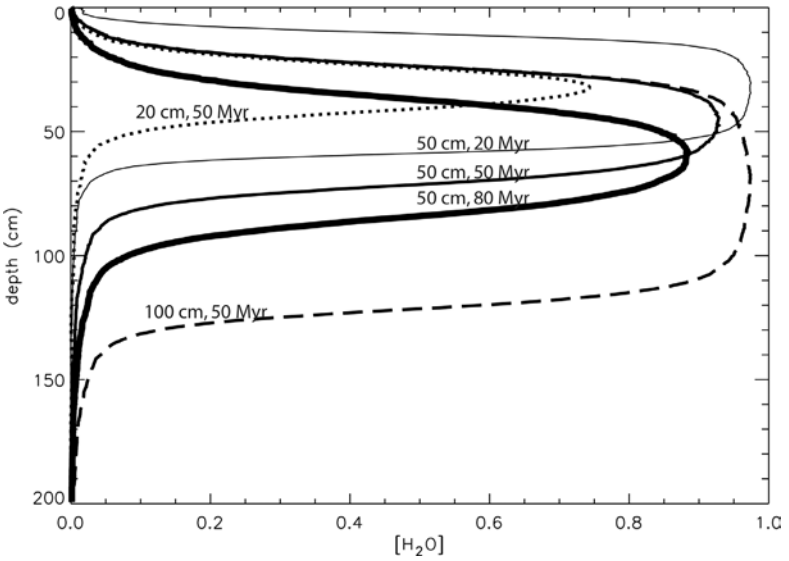


Fig. 1. Several depth profiles are given for a weathered ice layer on Mercury. The three solid lines are for an ice layer initially 50 cm thick after 20, 50, and 80 Myr (with increasing line thickness). The dashed lines show a weathered ice layer after 50 Myr that was initially 20 (dotted) and 100 cm (dashed) thick.

the medium thickness solid line. In order to have a layer of dry regolith cover a few tens of cm thick, the time of emplacement is constrained to about 50 Myr ago. The burial rate at Mercury is 0.43 cm/Myr.[14] Note that the thick solid curve in Fig. 1 shows that at an age of 80 Myr, the deposit would be covered with more dry regolith than indicated by the radar observations. The thin curve suggests that after 20 Myr, there is not enough material covering the deposit to be consistent with the observations.

Similarly, the purity of ice required for the coherent radar backscatter sets a strict condition on the ice in hermean PSR. Unless there is a mechanism that we have not considered that concentrates ice in the PSR, the ice would have to have been deposited in an episodic event, e.g., a comet impact. Also provided in Fig. 1 are the expected profile for each of three ice layers of different initial thickness (20, 50, and 100 cm) 50 Myr after emplacement. An interesting effect is that the peaks are buried to the same depth, offset by the difference in the initial thickness.[14] Also, the profiles look very similar on the top-side until the peak concentration is reached. The dotted profile is for a layer initially 20 cm thick. The dashed profile is for a 100 cm thick layer. Both the 50 and 100 cm layers have sufficient

purity (>90%) to be consistent with the backscatter signature at 50 Myr. The 20 cm layer only has a peak [H_2O] of 75% by weight though. This suggests that the observed ice at Mercury began as a thick deposit.

3.2. *Moon*

The lunar data are consistent with <10% water ice mixed with regolith in the top 1 m of regolith. Because of the nature of the neutron measurements, the inference of the depth distribution is based on a model of neutron leakage through material.[20] The neutron fluxes at the Moon are reproduced with a single layer model. However, that does not preclude an uneven distribution with depth.

The lunar model shows that an ice deposit on the Moon will get buried and mixed with regolith over time.[12,21] The burial rate is, in fact, slower at the Moon than at Mercury. In order for a thick ice deposit on the Moon to be consistent with the neutron data, deposition would have to have occurred at least 500 Myr ago. In Fig. 2, three black curves display the evolution of

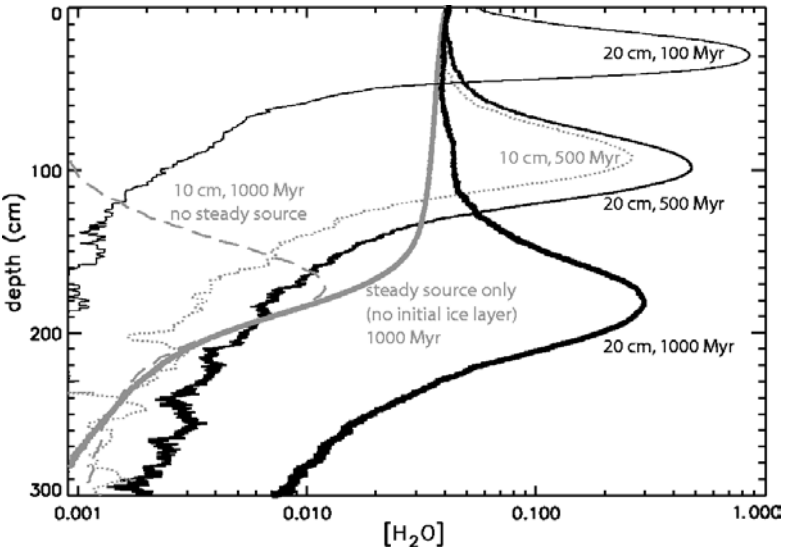


Fig. 2. Depth distribution of water ice in the lunar PSR is given for several model runs. The dashed gray line gives the evolution of a 10 cm thick ice deposit over 1,000 Myr without any additional source of water vapor. The other curves include migrating water vapor. The initial thicknesses and ages of the deposits are labeled next to the curves. The thick gray solid line included no initial ice deposit. The water concentration is on a log scale, in contrast to Fig. 1.

a 20 cm thick initial ice deposit from 100 to 500 to 1,000 Myr. In these simulations, a steady source of water vapor was included at the rate of 1.8×10^{-15} g/cm^2 s.[22]

The effect of the steady source alone is given in the thick, gray line.[12] As time goes on, the enriched layer thickens, but does not become more concentrated. One curve (dashed gray) shows the evolution of a 10 cm thick ice deposit after 1,000 Myr without a steady source of water vapor. Such an ice deposit would be undetectable by LPNS.

The gray dotted curve on Fig. 2 depicts the depth profile of an ice deposit initially 10 cm thick. It is possible that several thin ice layers have been deposited in the PSR of the Moon from cometary impacts. We would expect this to result in several thin, enriched layers of water ice interspersed with drier regolith. As long as the total thickness of the ice layers deposited within the last 500 Myr is <10 cm, the model simulations are consistent with the data.[13]

4. Conclusions

Our model suggests that the putative ice deposits in the PSR of Mercury were emplaced very recently, <50 Myr ago. Space weathering on the Moon would not remove the signature of a recent ice layer on the Moon. If relatively pure water ice layers exist in lunar PSRs, our model suggests they were emplaced quite long ago (>500 Myr) and are buried too deep for detection by neutron spectroscopy. More likely is that any cometary ice on the Moon began as a thin layer (<10 cm).

Gradual accumulation of water vapor (from migration through the atmosphere) cannot explain the observations of pure ice on Mercury, but may contribute to the inventory in addition to cometary ice layers. Gradual accumulation of water vapor is consistent with the observations of <10% water ice on the Moon. Further, if gradual accumulation of water is occurring, it limits the thickness of any ice layer that would be consistent with the lunar observations.

The burial rate of regolith at Mercury (0.43 cm/Myr) exceeds that at the Moon (0.16 cm/Myr) by a factor of about 3. The slow burial rate at the Moon allows the slow source to accumulate before being buried whereas the fast burial rate at Mercury tends to dilute the slow source. Meanwhile, ice stays near the surface longer at the Moon, thus is subjected to loss from exposure more than at Mercury.

Assuming that cometary ice exists at both the Moon and Mercury, it is still puzzling that a thick deposit appears close to the surface on Mercury, where it would be buried very quickly. Are we simply observing a special time, soon after a recent comet impact on Mercury? Recent observations of Sungrazing comets by the Solar Heliospheric Observatory suggest that there is currently a large population of comets (Kreutz family) that cross Mercury's orbit.[23] A similar population could have existed in the past to supply the putative deposits at Mercury. Further, it is true that the impactor flux is greater at Mercury; but not by enough to account for the difference in lunar and hermean near-surface ice.

Perhaps the loss rate from exposure is greater than our model has incorporated. Then the slow burial of ice on the Moon would result in a lower retention fraction. That would account for the observed differences at Mercury and the Moon. Remote sensing data from upcoming lunar and hermean missions will provide data that can help determine the contents of the PSRs. They will give more information and may help better constrain the possible source and retention mechanisms.

Acknowledgments

This work was supported by NASA grants NNG05GB84G and NNG05GA80G.

References

1. A. R. Vasavada, D. A. Paige and S. E. Wood, *Icarus* **141** (1999) 179.
2. K. Watson, B. C. Murray and H. Brown, *J. Geophys. Res.* **66** (1961) 3033.
3. W. C. Feldman *et al.*, *Science* **241** (1998) 1496.
4. J. K. Harmon and M. A. Slade, *Science* **258** (1992) 640.
5. A. L. Sprague, D. M. Hunten and K. Lodders, *Icarus* **118** (1995) 211.
6. L. Starukhina, *J. Geophys. Res.* **106** (2001) 14701.
7. R. R. Hodges, *J. Geophys. Res.* **107** (2002) 5125.
8. J. Arnold, *J. Geophys. Res.* **84** (1979) 5659.
9. B. J. Butler, *J. Geophys. Res.* **102** (1997) 19283.
10. M. J. Cintala, *J. Geophys. Res.* **97** (1992) 947.
11. G. Neukum *et al.*, *Planet. Space Sci.* **49** (2001) 1507.
12. D. H. Crider and R. R. Vondrak, *J. Geophys. Res.* **108** (2003) 5079.
13. D. H. Crider and R. R. Vondrak, *Adv. Space Res.* **31** (2003) 2293.
14. D. H. Crider and R. M. Killen, *Geophys. Res. Lett.* **32** (2005) L12201.
15. J. K. Harmon, P. J. Perillat and M. A. Slade, *Icarus* **149** (2001) 1.

16. B. A. Campbell *et al.*, *Nature* **426** (2003) 137.
17. W. C. Feldman *et al.*, *J. Geophys. Res.* **105** (2000) 4175.
18. W. C. Feldman *et al.*, *J. Geophys. Res.* **106** (2001) 23231.
19. R. C. Elphic *et al.*, *Lun. Plan. Sci. Conf.* **XXXVI** (2005) 2297.
20. W. C. Feldman, R. C. Reedy and D. S. McKay, *Geophys. Res. Lett.* **18** (1991) 2157.
21. J. Arnold, *Proc. Lun. Sci. Conf. 6th* (1975) 5659.
22. D. H. Crider and R. R. Vondrak, *Adv. Space Res.* **30** (2002) 1869.
23. D. A. Biesecker *et al.*, *Icarus* **157** (2002) 323.

GLOBAL HIGH-RESOLUTION STEREO MAPPING OF THE MOON WITH THE SELENE TERRAIN CAMERA

JUN'ICHI HARUYAMA*,†, MAKIKO OHTAKE†,
TSUNEO MATSUNAGA‡ and LISM WORKING GROUP
†*Japan Aerospace Exploration Agency, ISAS*
3-1Yoshinodai, Sagamihara City, Kanagawa 229-8510, Japan
‡*The Social and Environmental Systems Division*
National Institute for Environmental Studies
**haruyama.junichi@jaxa.jp*

The surface features of the Earth and Moon differ substantially. The detailed investigation of morphology is one of the foremost science goals of a Japanese lunar explorer Selenological and Engineering Explorer (SELENE) that will be launched in 2007. The Terrain Camera (TC), one of the mission instruments on SELENE, is a panchromatic stereo camera of 10 m spatial resolution from the SELENE nominal altitude of 100 km. Detailed lunar maps including digital terrain models (DTMs) will be produced from the TC data. These maps will be fundamental information for lunar morphology. In this paper, we describe the TC global high-resolution stereo mapping experiment.

1. Introduction

To provide lunar topographic data is one of the most important subjects of lunar explorations. Contour maps produced using Apollo stereo imaging data have been fundamental data sets for lunar sciences though they cover only 20% or less of the surface in the lunar equator regions. Nearly global (between $\pm 75^\circ$ in longitude) topographic data for the Moon were obtained by Clementine lidar with vertical accuracy of 130 m but the mean distance between its sampling points was considerably large, about 50 km.[1,2] Recently, Cook *et al.*[3–5] have produced excellent lunar global Digital Terrain Models (DTMs) from Clementine UVVIS image data with spatial resolution of 100–150 m/pixel. Local relative height resolution of their DTM tiles is within nearly 100 m. On the other hand, Archinal *et al.*[2] are creating the newest control network, Unified Lunar Control Network (ULCN) 2005 with horizontal accuracy of a few km, and vertical accuracy of 1 km or less.

In 2007, Japan will launch a three axes stabilizing lunar orbiter Selenological and Engineering Explorer (SELENE). One of the most important purposes of SELENE is to obtain source data to provide global lunar morphological maps with higher resolutions than those of past missions. In the 14 mission instruments on SELENE, especially Lunar Altimeter (LALT) and Terrain Camera (TC) are expected to provide data attaining this purpose. The nominal sampling rate of LALT is 1 s that is corresponding to 1.6 km interval on the lunar surface and the ranging error of LALT will be $< \pm 5$ m.[6] The TC will take global stereo pair data with a spatial resolution of 10 m from SELENE nominal altitude of 100 km to provide DTMs with height resolution of about 20 m. In this paper, we address the TC mission with specification, operation plan, and ground data systems of this instrument.

2. Overview of TC Mission Objectives and Specification of TC

On SELENE, three optical cameras are installed: the TC, the Multi-band Imager (MI), and the Spectral Profiler (SP). The TC is a stereo mapping imager,[7–10] MI is a multi-color imager with five-bands in visible range and four-bands in near-infrared range,[11] and SP is a continuous spectral profiler of the range from visible to near infrared.[12] These cameras share some parts of structures and electrical circuits to reduce mission resources of weight, space, and electrical power. We call the three optical cameras of SELENE as the Lunar Imager/Spectrometer (LISM), generally.

The mission objectives of TC are to produce

(1) global high-resolution and high-contrast mosaicked maps of the Moon,
(2) DTMs for the entire surface of the Moon that are expected to be base maps and models of future lunar sciences and lunar explorations.

The TC consists of two slant telescope systems (see Fig. 1). Each system has a linear detector with a panchromatic filter in a visible range of 0.43–0.85 μm. The TC takes images by push-broom method. The slant angle is $\pm 15°$ from the nadir vector, thus the base-height ratio of TC is 0.57. The instant field of view (IFOV) of TC is 0.00553°. The number of each detector pixel is 4096. The TC has three swath modes: (1) full swath mode that uses all detector pixels of 4096, (2) nominal swath mode that uses 3500 pixels, and (3) half swath mode that uses 1750 pixels. The field of view will be 22.4, 19.3, and 9.65° for full, nominal and half swath mode, respectively.

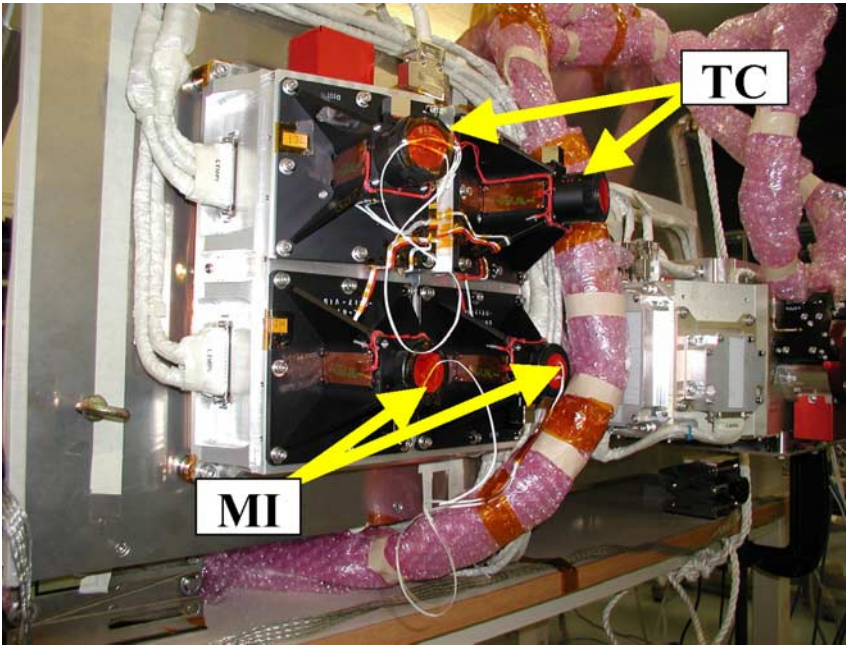


Fig. 1. Terrain Camera (TC) flight model. The TC consists of two slant optical heads, or telescopes. The slant angles are ±15° to the nadir vector. A linear CCD sensor with panchromatic filter in the visible range of 0.43–0.85 μm is attached at the bottom of each head. The spatial resolution of TC is 10 m/pixel at the SELENE nominal altitude of 100 km.

The sampling interval of TC is 6.5 ms. Thus the cross track and along track spatial resolutions of TC will be 10 m at the SELENE nominal altitude of 100 km. Since the TC sampling time is fixed, the along track spatial resolution does not change so much even if the SELENE altitude becomes lower or higher. The TC has three exposure time modes: (1) long-exposure time mode of 6.5 ms exposure, (2) middle exposure time mode of 3.25 ms exposure, and (3) short-exposure time mode of 1.625 ms exposure. The saturation reflectance values are set to be 8, 16, and 32% for long, middle, and short-exposure time mode, respectively. The adopted A/D converter has a rate of 10 bit. The TC data can be transmitted in form of non-compression. However, since the TC data rate will be very large, we have prepared DCT-compression modes that will be selected in 32 Q-tables. We are now investigating how to select an appropriate Q-table in various imaging conditions.

3. Overview of SELENE/TC Operation

The SELENE will be injected into a Moon circular orbit of 100 km altitude, and revolve in 118 min around the Moon. A visible operation time for the Moon explorer at one ground space station is at longest 12 h and possibly about 8 h/day. Considering this restriction, SELENE has onboard data storage of about 100 Gbit memory and performs the transmission data rate of 10 Mbps from the explorer to the Usuda Deep Space Center (UDSC) in Japan. The data amount of TC and MI will be very large since these are high-resolution imaging cameras. For example, the TC nominal data rates onboard that can be calculated from the active pixel number of 3500 for nominal swath mode and A/D converter rate of 10 bit are 10.8 Mbps for TC stereo observation mode and 5.4 Mbps for TC mono-telescoping mode. The total data amount per day will be about 500 Gbit. The mission data recorder on SELENE has no partition by mission instrument and a large amount data will overwrite other mission data. Therefore, an allotment for data volume of 50 Gbit/day has been required to LISM. The operation policies of LISM to satisfy this requirement is that

(1) TC and MI will not operate simultaneously,
(2) TC/MI will take images in a scheduled longitude range, and
(3) TC/MI will compress their data onboard.

In practice, appropriate solar elevation angles for TC and MI are different, thus such an isolated operation of TC and MI is not of disadvantage. During the SELENE nominal mission period of 1 year or 13 moon cycles, we are planning to execute as follows:

(1) TC stereo-mapping in three moon cycles,
(2) TC mono-telescope mapping at illumination conditions of eastward and westward lower solar elevation angles for 60N–60S in two moon cycles, and
(3) MI mapping in 6 months.

The first two months will be used for checkout of instruments.

We are planning to execute two type TC observations regularly in one revolution as follows:

(1) nominal stereo-mapping observation or mono-telescope mapping, and
(2) SP support observation by one telescope system with a high-compression mode to reduce the data to a few percentage.

Besides these regular observations, we will execute the dark current acquisition of about 20–30 s during flights on the night-side at least once a week. Moreover, for the important or interesting areas that are out of the nominal observation, we will plan extra observations with consideration of allowed data amount and prediction of data reduction by compression.

4. Overview of TC Ground Processing Systems

4.1. *Level 2A processing system*

The SELENE data will be transmitted to the SELENE Operation and Analysis Center (SOAC) in the Institute of Space and Astronautical Science (ISAS) at Sagamihara from USDS and stored in the level 0/1 (L0/1) processing system in the SOAC (see Fig. 2). When the data arrive at the L0/1 processing system, a notification message is mailed to the LISM level 2A (L2A) processing system in Channel Access Data Unit (CADU) format. When the L2A processing system receives the mail, it starts running to get the LISM data that is divided into Virtual Channel Data Unit (VCDU) with a mission ID of LISM. The L2A processing system extracts LISM data with the Consultative Committee on Space Data Systems (CCSDS) packet format from data in VCDU. A CCSDS packet is composed with a primary header, secondary headers, user data, and a check word that is optional. The extracted LISM data in CCSDS packet are depacketed and separated into TC, MI, and SP data.

In the L2A processing system, a number will be attached in a sequential TC data taken in a set of same parameters. This number is called as a "strip" number. The strip number is sequentially attached by each revolution. The TC strip numbers are given without distinguishing telescope systems and same strip numbers are given for the stereo observation. The data divided into unit of strip will be thereafter cut into "scenes" by a definite line number with an overlap. The planned line numbers of net and overlap are 4088 and 568 in default, respectively. Since the nominal active pixel number for TC is 3500 (and two dummy pixels for DCT compression), the total volume of a TC scene data processed in L2A processing system will be about 20 MB. The TC data in scene unit are re-compressed by 12 bit JPEG and registered in SELENE level 2 data base (L2DB) system in the SOAC with a thumbnail image and a catalogue information file. Co-investigators (Co-Is) will get the SELENE mission data products via the L2DB system.

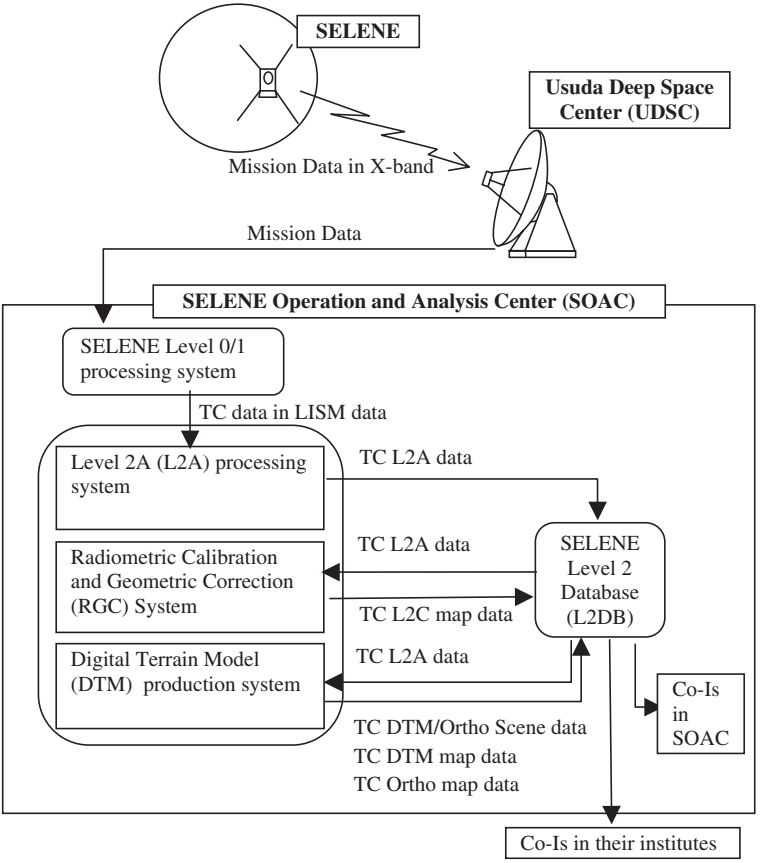


Fig. 2. SELENE Terrain Camera (TC) ground data processing systems: (1) level 2A (L2A) processing system, (2) radiometric calibration and geometric correction (RGC) system, and (3) DTM production system.

4.2. *Radiometric calibration and geometric correction system*

The radiometric calibration and geometric correction will be executed in the LISM RGC processing system in SOAC. The RGC system gets the TC level 2A data from the L2DB system and processes dark level correction, flat field calibration, and non-linearity calibration. The TC data processed in the RGC system are classified into two levels: one is the level 2B (L2B) data that has radiance values and another is the level 2C (L2C) data that has reflectance values. We are planning that the reflectance values of L2C

data are corrected into that of 30, 0, 30° in incident, emission, and phase angles. There are some studies on photometric collection for Clementine bands (e.g. see http://www.actgate.com/act/clemcal/alfred.doc written by A.S. McEwen) but few for panchromatic data like TC data. The phase angle correction for TC data is in consideration. The TC L2C scene data will be mosaicked to be two base-maps of eastward and westward lower solar illumination conditions that will show high-contrast images. These two maps for the entire Moon will be registered in the L2DB system.

4.3. *DTM production system*

We are also developing the TC DTM production system installed in the SOAC. Input data to this system are also L2A scene data stored in the L2DB system. The DTM production system performs radiometric calibration, geometric correction, and destriping prior to stereo-matching. For the DTM production proceeding, we will use the newest SELENE orbital and attitude information provided by the JAXA operation team. The error of orbital information may have an order of km at first. The substantial improvement of accuracy of SELENE orbit information are planned by the operation team and SELENE gravitational science team. The attitude information will be delivered in an interval of 2 s with accuracy of better than 0.0025°. The designed SELENE attitude control stability is $\pm 0.003°/s$ in 3σ corresponding to 0.4% of TC IFOV in the TC sampling time. An analysis of SELENE structure response modes has shown that SELENE will be so stable not to make critical oscillation for TC, less than a few percentage of IFOV in the TC sampling time. Thus, we expect that the DTM errors due to spacecraft oscillation will be very small.

Produced DTM scene data set will be registered in the L2DB system with a quality information data, an ortho-imaging data, a catalog information file, and a thumbnail. A set of adjoining DTM scene data also will be mosaicked to be a map that will be also registered in the L2DB system. One of the most important pre-launch work is to prepare appropriate DTM matching parameters. To build an absolute Digital Elevation Model (DEM) of the Moon with a few 10 m resolution is an ultimate goal of TC global stereo-mapping.

5. Conclusion

The SELENE, scheduled for launch in 2007, will carry the TC onboard. The goal of the camera is high-resolution stereo mapping for the entire lunar

surface. Since the TC spatial resolution is 10 m/pixel and base-height ratio is 0.57, we can expect to generate the lunar DTMs with 20 m or better height resolution. Although there are restrictions for transmitted data rate and the total volume of the mission recorder onboard, we accomplish the global high-resolution mapping by TC with smart operation plans. Although the ground data processing systems that are now in developing, TC data will be carefully processed and distributed to Co-Is and researchers worldwide. The TC global high-resolution data will contribute to study of the origin and evolution of the Moon.

Acknowledgments

We acknowledge the detailed review and helpful comments of the reviewer for this paper.

References

1. D. E. Smith *et al.*, *JGR*, **102** (1997) 1591–1611.
2. B. Archina *et al.*, *LPSC* **XXXVI** (2005) 2106.pdf.
3. A. C. Cook *et al.*, *LPSC* **XXXIII** (2002) 1281.pdf.
4. A. C. Cook *et al.*, *LPSC* **XXXI** (2000) 2978.pdf.
5. A. C. Cook *et al.*, *Planet. Space Sci.* **44** (1996) 1135–1148.
6. H. Araki *et al.*, *ISTS 2006 abst.*, in preparation.
7. J. Haruyama *et al.*, *LPSC* **XXXI** (2000) 1317.pdf.
8. J. Haruyama *et al.*, *ISTS 2000 Proceedings* (2001) 2000-j-25p.
9. J. Haruyama *et al.*, *ISTS 2004 Proceedings* (2003) 2002-k-26p.
10. J. Haruyama *et al.*, *ISTS 2004 Proceedings* (2005) 867.
11. M. Ohtake *et al.*, *ISTS 2002 Proceedings* (2002) 2002-j-27p.
12. T. Matsunaga *et al.*, *SPIE 2000 Proceedings* **4151** (2000) 32–39.

OXYGEN CHEMISTRY IN THE VENUS MIDDLE ATMOSPHERE

F. P. MILLS[*,†,**], M. SUNDARAM[*,‡], T. G. SLANGER[§],
M. ALLEN[¶,‖] and Y. L. YUNG[‖]
[*]*Research School of Physical Sciences and Engineering*
The Australian National University, Canberra, ACT 0200, Australia
[†]*Centre for Resource and Environmental Studies*
The Australian National University, Canberra, ACT 0200, Australia
[‡]*Department of Physics, Faculty of Science*
The Australian National University, Canberra, ACT 0200, Australia
[§]*Molecular Physics Laboratory, SRI International*
Menlo Park, CA 94025, USA
[¶]*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA 91109, USA
[‖]*Division of Geological and Planetary Sciences*
California Institute of Technology, Pasadena, CA 91109, USA
[**]*Frank.Mills@anu.edu.au*

Decades of research have sought to understand the similarities and differences between Venus and Earth. Yet, it is still not clear what chemical processes maintain the long-term stability of Venus' primarily CO_2 atmosphere because, until recently, the observed limit on O_2 was an order of magnitude smaller than predicted by photochemical model calculations. CO_2 dissociates into CO and O after absorbing photons at wavelengths <205 nm. These O atoms should combine to form O_2, and observations of intense airglow, produced as oxygen molecules in the $O_2(a^1\Delta)$ and $O_2(c^1\Sigma)$ states decay radiatively to the ground state, confirm rapid production of O_2 on both day and night sides. Achieving an appropriate balance in numerical models between this rapid production of O_2 and the rapid destruction implied by the observational upper limits on O_2 has been a challenge for the past 25 years. Numerical modeling shows that recent proposals may resolve this gap between theory and observations, depending on the rates of poorly constrained reactions. The laboratory and observational studies needed to help resolve remaining questions regarding oxygen chemistry in the Venus middle atmosphere are outlined.

1. Introduction

Venus is the planet in our solar system that is most similar to Earth but there are significant differences. One of the most striking is the composition of their atmospheres. Understanding how and why they differ can

help elucidate the processes that control the composition and evolution of terrestrial-type atmospheres. An important first step is to understand the present processes that maintain the long-term stability of CO_2 on Venus.

The primary constituent of the Venus atmosphere is CO_2, ~96.5%.[1] CO_2 photodissociates at wavelengths <205 nm to form CO and O. In an initially pure CO_2 atmosphere, O preferentially combines to form O_2 and within a few thousand years, an initially pure CO_2 atmosphere evolves to have ~7% CO and 3.5% O_2.[2,3] The observed abundances for CO and O_2 on Venus and Mars, however, are orders of magnitude smaller, which suggests efficient catalytic cycle(s) aid in re-forming CO_2. For Mars, these catalytic cycles are believed to involve HO_x radicals[2] (HO_x = OH + HO_2 + H), and models have been able to reproduce the observed CO and O_2 using observational constraints for a number of years.[2] Comparable agreement has yet to be achieved for Venus, but recent modeling[4] has shown reasonable agreement with the existing upper limit on O_2.[5]

The leading candidates for gas-phase catalysts that can stabilize CO_2 in the Venus atmosphere are chlorine compounds,[6,7] and recent laboratory work[4] has validated key assumptions made in models since the 1980s. Significant uncertainties remain, however, in the rates for critical reactions in the chlorine catalytic pathways, in the photolysis rates for loss of CO_2, and in the potential efficacy of alternative catalytic pathways.[8,9] In addition, new observations of oxygen airglow suggest current models of Venus' oxygen chemisty may be too simplistic. This manuscript compares the results from gas-phase chlorine catalytic chemistry with two more speculative catalytic mechanisms:

$$CO + O_2(c^1\Sigma) \rightarrow CO_2 + O(^1D), \tag{1}$$

$$CO + O + \text{aerosol} \rightarrow CO_2 + \text{aerosol} \tag{2}$$

and then discusses what laboratory and observational work is needed to improve our understanding of Venus' oxygen chemistry.

The detailed discussion of Reaction (1) from a forthcoming publication[9] is summarized here. The primary net production of O_2 in Venus atmospheric models is via Reaction (3),

$$2O + M \rightarrow O_2^* + M, \tag{3}$$

where M is any third molecule or atom that can collisionally stabilize the intermediate complex and O_2^* is one of the many excited states of O_2. Laboratory and theoretical studies[10,11] suggest a large majority of the O_2 formed in Reaction (3) is initially in a highly excited state $\left(A^3\Sigma_u^+, A'^3\Delta_u, c^1\Sigma_u^-\right.$, or

$^5\Pi_g$) and that there is rapid crossover between these states. CO_2 is particularly efficient at quenching $O_2(c, \nu > 0)$ to $O_2(c, \nu = 0)$,[11–14] so a significant fraction of the O_2^* produced should pass through the $O_2(c, \nu = 0)$ state, and airglow observations[15] support this hypothesis. Collisional quenching by CO_2 of $O_2(c, \nu = 0)$ is slow,[16] so $O_2(c)$ should be available in sufficient quantity to make Reaction (1) a potentially significant source for CO_2 above 90 km altitude. A laboratory study[17] has shown that O_2^* can react with CO to form CO_2, but the identity of the excited O_2 state involved is not known. Thus, it is likely that a reaction like Reaction (1) does occur in the Venus atmosphere but the rate and detailed process is not known yet.

The detailed mechanism underlying Reaction (2) is unknown at present. One proposed mechanism[18–20] is that single- or multi-photon absorption in sulfuric acid aerosols may dissociate H_2SO_4 to form OH or O, which can react with dissolved CO to form CO_2. A second possibility is that photo-excitation of CO may be followed by either Reaction (4)[21] or Reactions (5) and (6),[17,21]

$$CO^* + CO \rightarrow CO_2 + C, \tag{4}$$

$$CO^* + O_2 + M \rightarrow CO_3 + M, \tag{5}$$

$$CO_3 + O_2 \rightarrow CO_2 + O_3. \tag{6}$$

A third proposed mechanism[8,20,22] is that photo-excitation of sulfuric acid with trace amounts of dissolved Fe and/or SO_2 can produce SO_3^-, SO_4^-, and HSO_5^- which can oxidize dissolved CO to CO_2. Given the uncertainties associated with Reaction (2), we have adopted the approach of postulating the existence of Reaction (2), identifying the effective rate at which this reaction becomes important, and then assessing whether that effective rate is plausible.[23]

2. Photochemical Model

Our photochemical model is based on the Caltech/JPL model[24] and is similar to that used to study the photochemistry of Mars.[2] A detailed description of the model inputs is available,[3] so only key features and subsequent modifications are summarized here. The Venus model represents the "global average" photochemistry, so diurnal variations have not been explicitly modeled. Photodissociation was calculated for 45° latitude at local noon and then divided by two. The solar fluxes used represent high solar irradiance conditions. Diffuse radiation was calculated using

the Feautrier method generalized to arbitrary anisotropic scattering with aerosol scattering properties based on *Pioneer Venus* measurements. Photoabsorption and chemical kinetic data were drawn from critically reviewed compilations[25] whenever possible. Temperature-dependent cross-sections were used for CO_2, SO_2, O_3, and OCS. The UV cross-sections for photolysis of $ClCO_3$ from recent laboratory work[4] were included. Changes in kinetic rate recommendations since 1997 are not expected to be significant. We simultaneously solve the one-dimensional continuity equation for 47 species using a finite difference iterative algorithm to converge to a steady-state solution. Vertical transport via Eddy diffusion was set based on observations. The H_2O profile was fixed to match the equilibrium vapor pressure over 75 wt.% sulfuric acid.[26] Our model atmosphere, noon local solar time, extends from 58 to 112 km altitude with 2-km thick layers. At the lower boundary, the mixing ratio for CO_2 was set to 0.965, HCl to 0.4 ppm, OCS to 1 ppb, and SO_2 to 1 ppm. The concentration gradient at the lower boundary for all other species was set to zero. At the upper boundary, the upward flux of molecules, such as CO_2, was set equal to the downward flux of their photodissociation products to simulate photodissociation occurring above the upper boundary.

3. Model Results

Key results from the three alternative catalytic schemes for production of CO_2 are compared with extant global-scale observations and the results from a nominal chemistry model calculation in Table 1. The calculated oxygen profiles are shown in Fig. 1.

The $O_2(c^1\Sigma)$ model examined the effect of Reaction (1). $O_2(c^1\Sigma)$ was produced via $2O + M \rightarrow O_2(c^1\Sigma) + M$ with an assumed branching ratio of

Table 1. Comparison of global average observations and calculations.

Observation	Measured value	Model calculation			
		Aerosol	P2004	$O_2(c^1\Sigma)$	Nominal
SO_2 scale height at 70 km (km)	3 ± 1[a]	3	3	4	4
f(SO_2) at 70 km (ppb)	$20 - 430$[a]	50	40	40	90
SO peak near 65 km (10^{10} cm^{-3})	$1 - 4$[a]	2	1	0.3	0.3
O_2 column (10^{18} cm^{-2})	<1.5[b]	2.6	2.1	16	20
$O_2(^1\Delta)$ airglow (MR)	$1 - 1.5$[c,d]	1.2	1.2	1.5	2.1
H_2SO_4 production (10^{11} cm^2/s)	$\sim 2 - 10$[e,f]	3	4	8	9

Sources for measured values: [a]Ref. 27; [b]Ref. 5; [c]Ref. 28; [d]Ref. 29; [e]Ref. 6; [f]Ref. 30.

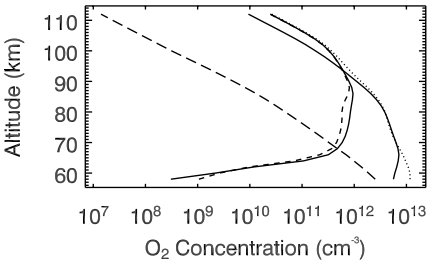


Fig. 1. Comparison of O_2 concentrations calculated from several photochemical models and the O_2 profile used to interpret the spectroscopic upper limit. The dotted line shows the results from the nominal chemistry model. The dashed-triple-dotted line is from the $O_2(c)$ model. The short-dashed line is from the P2004 model. The solid line is from the aerosol chemistry model. The long-dashed line is the O_2 profile used to interpret the spectroscopic upper limit observation.[5]

90%.[11] The rate for loss of $O_2(c^1\Sigma)$ due to collisions with CO_2 was assumed to be 1×10^{-15} cm^3/s,[16] and the rate for Reaction (1) was assumed to be 1×10^{-11} cm^3/s. The rate for Reaction (1) has not been measured. The sensitivity of the model calculations to these parameters will be explored in a forthcoming publication.

The P2004 model[4] included several adjustments to enhance the effectiveness of the gas-phase catalytic chlorine chemistry.[6] The thermal dissociation rate for ClCO was decreased by 1.5 times the recommended uncertainty in the ClCO equilibrium constant,[25] the rate for $ClCO + O_2 + M \rightarrow ClCO_3 + M$ was increased slightly, and temperatures at 84–96 km altitude were decreased by 6 K. The sensitivity of the model calculations to the thermal stability of ClCO is illustrated by the difference between the P2004 and the nominal chemistry models in Fig. 1.

The aerosol model examined the effect of incorporating heterogeneous chemistry, Reaction (2), into a model that otherwise included only gas-phase chemistry. Reaction (2) was added to the P2004 model with a reactive uptake coefficient, γ, of 10^{-5}, which is believed to be a plausible rate (L. Phillips, personal communication, Jan 2005). At this rate, CO_2 production via Reaction (2) was comparable to that via the gas-phase chlorine catalytic pathways. The mixing ratio of OCS at the lower boundary was increased to 10 ppb[31] to better match the cloud level CO observation.[32] If γ is increased to 10^{-4}, then the calculated O_2 column abundance decreases by almost a factor of two. If γ is less than about 3×10^{-6}, then the Cl chemistry dominates production of CO_2, as it does in the P2004 model.

4. Discussion

All three models are generally consistent with the global-scale observations, Table 1, and the aerosol and P2004 models achieve comparably small column abundances for O_2. The calculated O_2 column abundances are larger than the standard interpretation of the observational upper limit.[5] However, the correct way to compare the calculated O_2 with the observational upper limit is via a radiative transfer model simulation of the spectroscopic observation. Prior radiative transfer simulations[33] suggest that if the O_2 profile has the shape predicted by the photochemical models, then an O_2 column abundance which is smaller than about $3 \times 10^{18}\,\mathrm{cm}^{-2}$ is roughly consistent with the observational upper limit. This factor of $\sim$ two uncertainty in interpretation of the observational upper limit arises because multiple scattering within the upper cloud strongly enhances absorption by O_2 within the upper cloud and the O_2 profile assumed for the standard interpretation of the upper limit observation[5] has the vast majority of its column abundance within the upper cloud. Most model calculated profiles, however, are severely depleted in O_2 below 70 km altitude, Fig. 1. In addition, the observational upper limit was derived from observations of a portion of the day side atmosphere. O_2 airglow and CO abundances vary significantly across the night side, so the current upper limit on O_2 may not be representative of the global-average O_2 calculated in the photochemical models. If the O_2 abundance is below the present upper limit then a combination of all three catalytic mechanisms discussed here may be required.

The primary disagreement between current models and observations is in the calculated global-average CO profiles. However, no CO observations have been reported for the day side at 60–90 km for the past two decades and none of the extant observations provide day side profiles. *Venus Express* should help fill this gap.

The three catalytic mechanisms investigated are most effective at different altitudes, Fig. 2: Reaction (1) above 90 km, Reaction (2) below 80 km, and the gas-phase chlorine catalytic mechanism at 65–90 km. As photodissociation of CO_2 occurs primarily at 70–100 km, the latter two mechanisms severely deplete oxygen and CO in the upper cloud. Consequently, a source of CO at the base of the upper cloud, which was photodissociation of OCS in our models, is required to explain the CO abundance observed in the upper cloud.[19]

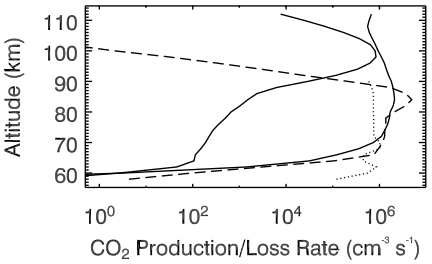


Fig. 2. Comparison of CO_2 production and loss rates calculated from several photochemical models. The solid line is the CO_2 photolysis rate from the nominal chemistry model. The dashed-triple-dotted line shows production of CO_2 via Reaction (1) from the $O_2(c)$ model. The long-dashed line shows production of CO_2 via $ClCO_3$ from the P2004 model. The dotted line shows production of CO_2 via Reaction (2) from the aerosol chemistry model.

5. Conclusions and Recommendations

Either chlorine or aerosol catalytic chemistry may be able to explain the low-O_2 abundance on Venus. The chlorine catalytic mechanism[6] has been validated[4] but the rates for the component reactions are poorly constrained. An aerosol catalyzed mechanism to produce CO_2 has not been demonstrated conclusively yet, so the rate(s) are unknown.

A multi-dimensional chemical transport model is needed and *Venus Express* will provide the information needed to guide development of such a model. Wind fields, temperatures, and the distribution of CO on the day and night sides at 60–100 km from *Venus Express* will be particularly useful for stepping up to a multi-dimensional chemical transport model.

Laboratory work is required to interpret data from *Venus Express* and improve the quality of photochemical modeling. Tighter constraints on the temperature dependent rate for $ClCO + O_2 + M \rightarrow ClCO_3 + M$ and the temperature dependent equilibrum constant for ClCO are needed. The efficacy of aerosol catalytic mechanism(s) for production of CO_2 needs to be quantitatively evaluated, and the rate for Reaction (1) needs to be determined along with the rates of competing loss reactions for $O_2(c^1\Sigma)$.

Several sets of observations are needed and many are possible from *Venus Express*. The abundance of ground-state O_2; coincident observations of the distributions of CO and $O_2(a^1\Delta)$ airglow; simultaneous retrievals of CO distributions, winds, and temperatures; and profiles for O_3 and OCS

in the middle atmosphere are among the most critical for understanding oxygen chemistry in the Venus middle atmosphere.

Acknowledgments

Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. Funding from NASA's Planetary Atmospheres program, the Australian Research Council, the Asia Oceania Geosciences Society, and the International Society of Biometeorology Tromp Travel Fund is gratefully acknowledged.

References

1. U. von Zahn *et al.*, Composition of the Venus Atmosphere, in *Venus*, eds. D. M. Hunten *et al.* (University of Ariz. Press, Tucson, AZ 1983), p. 299.
2. H. Nair *et al.*, *Icarus* **111** (1994) 124.
3. F. P. Mills, Ph.D. Dissertation, California Institute of Technology, 1998.
4. H. Pernice *et al.*, *PNAS* **101** (2004) 14007.
5. J. Trauger and J. Lunine, *Icarus* **55** (1983) 272.
6. Y. Yung and W. DeMore, *Icarus* **51** (1982) 199.
7. V. A. Krasnopol'sky and V. A. Parshev, Photochemistry of the Venus atmosphere, in *Venus*, eds. D. M. Hunten *et al.* (University of Ariz. Press, Tucson, AZ, 1983), p. 431.
8. G. A. Rowland *et al.*, *J. Photochem. Photobiol. A* **153** (2002) 1.
9. T. G. Slanger *et al.*, *Icarus in press* (2006) doi: 10.1016/j.icarus.2005.12.007.
10. R. P. Wayne, *Res. Chem. Intermed.* **20** (1994) 395.
11. T. G. Slanger and R. A. Copeland, *Chem. Rev.* **103** (2003) 4731.
12. T. G. Slanger, *J. Chem. Phys.* **69** (1978) 4779.
13. R. D. Kenner *et al.*, *J. Photochem.* **10** (1979) 199.
14. R. A. Copeland *et al.*, *J. Chem. Phys.* **105** (1996) 10349.
15. V. A. Krasnopol'sky, Venus spectroscopy in the 3000–8000 Å region by Veneras 9 and 10, in *Venus*, eds. D. M. Hunten *et al.* (University of Ariz. Press, Tucson, AZ 1983), p. 459.
16. R. D. Kenner and E. A. Ogryzlo, *Can. J. Chem.* **61** (1983) 921.
17. O. F. Raper and W. B. DeMore, *J. Chem. Phys.* **40** (1964) 1047.
18. C. T. Mills and L. F. Phillips, *J. Photochem. Photobiol. A: Chem.* **74** (1993) 7.
19. C. T. Mills *et al.*, *J. Photochem. Photobiol. A: Chem.* **93** (1996) 83.
20. S. J. Wrenn *et al.*, *J. Photochem. Photobiol. A: Chem.* **129** (1999) 101.
21. P. B. Roussel and R. A. Back, *J. Photochem. Photobiol. A: Chem.* **46** (1989) 159.
22. G. A. Rowland and L. F. Phillips, *Geophys. Res. Lett.* **27** (2000) 3301.
23. M. Sundaram, F. P. Mills, M. Allen and Y. L. Yung, Heterogeneous chemistry in the Venus middle atmosphere, *in preparation* (2006).
24. M. Allen *et al.*, *J. Geophys. Res.* **86** (1981) 3617.

25. W. B. DeMore *et al.*, *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling: Evaluation 12*, Jet Propulsion Laboratory Publication 97-4 (http://jpldataeval.jpl.nasa.gov/) (1997).
26. F. P. Mills, *Adv. Space Res.* **23**, 9 (1999) 1573.
27. C. Y. Na *et al.*, *Icarus* **112** (1994) 389.
28. P. Connes *et al.*, *Ap. J.* **233** (1979) L29.
29. D. Crisp *et al.*, *J. Geophys. Res.* **101** (1996) 4577.
30. J. R. Winick and A. I. Stewart, *J. Geophys. Res.* **85** (1980) 7849.
31. B. Bezard *et al.*, *Nature* **345** (1990) 508.
32. L. D. G. Young, *Icarus* **17** (1972) 632.
33. F. P. Mills, *J. Geophys. Res.* **104** (1999) 30757.

This page intentionally left blank

OBSERVATIONS IN THE SHADOW OF MARS BY THE NEUTRAL PARTICLE IMAGER

M. HOLMSTRÖM*,†,‡, K. BRINKFELDT*, S. BARABASH* and R. LUNDIN*
*Swedish Institute of Space Physics, P.O. Box 812, SE-98128 Kiruna, Sweden
†Currently at NASA Goddard Space Flight Center
Code 612.2, Greenbelt, MD 20771, USA
‡matsh@irf.se

We present observations of energetic neutral atoms (ENAs) in the shadow of Mars by the neutral particle imager (NPI), part of the ASPERA-3 experiment on-board Mars Express. The observations are well into the umbra, where the count rates are low, and contamination by UV light minimal. We present statistics over all available observations, from the 2004 and the early 2005 eclipse seasons. We investigate skymaps of the observations, and their time dependence, and try infer the different origins of the observed fluxes: instrumental effects, ENAs from Mars and its interaction with the solar wind, and ENAs of possibly heliospheric origin. We also study the time evolution of the observed signals. It is found that most of the observed emissions are consistent with UV, but some are not, suggesting ENA fluxes — Mars related and non-Mars related. The NPI measure the integral ENA flux (0.1–60 keV) with no mass or energy resolution, but with high angular resolution ($5 \times 11^\circ$). The sensor is also sensitive to UV light.

1. Introduction

The Neutral Particle Imager (NPI) is part of the ASPERA-3 experiment[1] on-board the European Space Agency's Mars Express mission (MEX) that was launched on June 2, 2003, and entered orbit around the planet early on Christmas morning 2003.

The ASPERA-3 instrument is comprised of four sensors; two ENA sensors, an electron spectrometer and an ion mass spectrometer.

The NPI provides measurements of the integral energetic neutral atom (ENA) flux with no mass or energy resolution but high-angular resolution. The sensor utilizes reflection and ion sputtering from a graphite surface to detect ENAs and suppress the UV background. The NPI is a replica of the ENA sensor used in the ASPERA-C experiment on the Mars-96 mission[2] and was successfully flown on the Swedish microsatellite Astrid,[3] launched in 1995. The UV suppression is not perfect, and we will later discuss how to separate ENAs from UV in the data.

The second ENA sensor, the neutral particle detector (NPD), provides measurements of the ENA flux, resolving velocity and mass (H and O) with a coarser angular resolution. ENAs incident on a surface at a grazing angle of 15° are reflected and cause secondary electrons used for a start signal. The reflected ENAs hit a second surface and again produce secondary electrons utilized for a stop signal. Time-of-flight electronics give the ENA velocity. The pulse-height distribution of the stop signals roughly determine the ENA mass. The NPD sensor is a new development.

ASPERA-3 provides the first ever ENA measurements in the solar wind energy range at another planet. In this work we describe, and discuss, NPI measurements in the shadow of Mars during the 2004 and early 2005 eclipse seasons.

2. Background

The NPI measure the integral ENA flux (in the energy range 0.1–60 keV) with no mass or energy resolution, but with good angular resolution. The incoming ENAs are detected by 32 sectors in a plane for a 360° field of view (FOV), with a 5° FOV perpendicular to the plane. Here we number the sectors from 0 to 31. The direction between sector 0 and 31 is NPI's x-axis, along the Mars Express z-axis that points toward the center of the planet during nadir observations. Sectors 15 and 16 are physically blocked and the count rates should be on the noise level, but we see higher count rates in the data from these sectors. This is an instrumental effect that is not yet understood, and these sectors should be disregarded if they appear in any plot (they are not used in any of the analysis of this paper).

The time resolution is 1 s, so the raw data is 30 integer counts every second. ASPERA-3's main unit is mounted on a scanning platform which means that the NPI can rotate around its x-axis when the scanner is operated. Here we will only discuss observations when the scanner was not operated. Then NPI is fix relative the spacecraft with its FOV plane perpendicular to MEX's y-axis (and the axis of the solar panels). Sector 8 is looking in the general direction of MEX x-axis (and the high-gain antenna). An illustration of the NPI geometry is shown in Fig. 1.

During different parts of Mars Express polar orbit, and during different Martian seasons, the NPI samples different ENA populations. The NPI dayside observations have been reported by Gunell *et al.*,[4] Brinkfeldt *et al.*,[5] and Kallio *et al.* [6] have investigated the decrease and increase in ENA flux at entry and exit of the Martian shadow. Predictions by computer simulations of the ENA generation from solar wind protons[7] and photoionized planetary

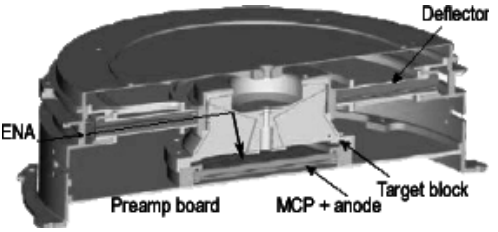


Fig. 1. Illustration of the NPI geometry.

oxygen[8] has previously been done. Here we investigate the ENA fluxes in Mars' umbra. One motivation is to search for interplanetary ENA populations. The interstellar neutral stream of helium, as detected by the ULYSSES/GAS-instrument,[9] is of too low energy (13 eV) to be detected by the NPI, since the NPI only is sensitive to neutrals with an energy of at least 200 eV (see Fig. 11) and the energy of the interstellar neutral helium will be below this energy even if we account for Mars' orbital velocity, and the gravitational acceleration of Mars and the Sun.

However, many different observations hint at the existence of other interplanetary fluxes of neutral atoms (no direct observations have been made), as described by Collier *et al.*[10] Such streams might have energies in the keV range, detectable by the NPI. The most promising part of Mars Express' orbit to search for such streams is in the shadow of the planet since we then do not have any direct solar UV light contamination, and the flux of ENAs from the Mars–solar wind interaction should be low.

In Fig. 2 we see an example of the raw data from an eclipse observation.

Mars Express' orbit goes into the shadow at different time periods (eclipse seasons). The data analyzed here is from two different eclipse seasons, one in 2004 and one in early 2005. This is summarized in Table 1. Figure 3 shows the eclipse part of the orbit, and how it evolves during these two eclipse seasons.

To avoid the ENA fluxes at entry and exit of eclipses (probably ENAs produced by solar wind-exosphere charge exchange that then are scattered by elastic collisions in the exosphere) we restrict the examined data to observations when the spacecraft is at least 0.15 Martian radii inside the umbra (if not otherwise noted). We call these observations deep umbra observations. The value of 0.15 was chosen after examining how the count rates depend on distance inside the umbra for all the observations. It was found that the countrates level off inside this region, as illustrated in Fig. 2(b).

Examples of NPI's FOV in ecliptic coordinates are shown in Fig. 4 with Mars' limb and center also indicated. During Earth communication the

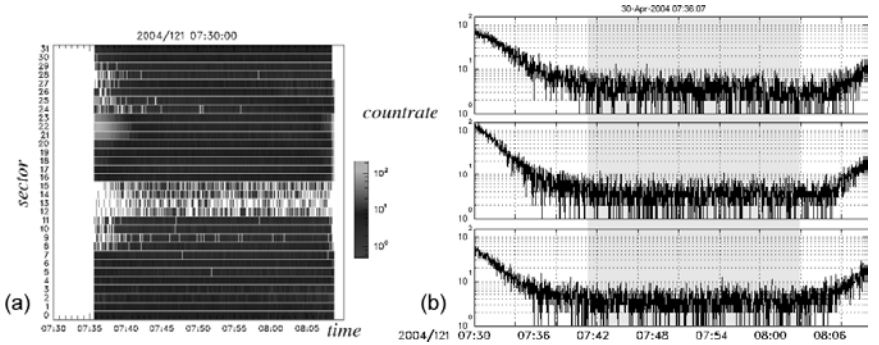


Fig. 2. Typical NPI eclipse observation on April 30, 2004, 07:36–08:08. (a) Countrates for all 32 sectors as a function of time. Sector 31 is the top row and sector 0 at the bottom. We see the sun set in sector 12 early in the observation, and rise again at the end of the observation. (b) Countrates for three of the sectors with the deep eclipse region shaded. The sectors are 21–23 from top to bottom. The y-axis scale goes up to 100 counts per second.

Table 1. Summary of eclipse observations presented in this work.

Dates	Number of observations	Duration of eclipses (min)	Mars' ecliptic longitude (deg)
2004-04-09 to 2004-08-21	259	10–20	102–162
2005-02-14 to 2005-03-31	163	30–60	245–270

NPI FOV plane is in the ecliptic plane. During nadir observations, inertial observations, and slews, the NPI observes other directions.

To examine the NPI observations in ecliptic coordinates, we would like to make sky maps where all the sectors contribute. A complication is that different sectors have different responses to ENAs. Their sensitivity and background noise levels are different. We can use ground calibration values to solve this inter-sector normalization problem, but we think that the in-flight values are slightly different, and that they have changed between 2004 and 2005. Here we use ground calibration values for the sectors' relative sensitivities, and we compute the noise level (dark currents) for each sector from observations, as shown and explained in Table 2. The noise level was estimated by binning the eclipse observations in 60 s bins, and the smallest average count rate was chosen as a noise level estimate for each season. It would have been preferable to also do an in-flight estimate of the the sectors' relative sensitivities, and attempts has been made, e.g., to identify times when different sectors look in the same direction close in time, but no good

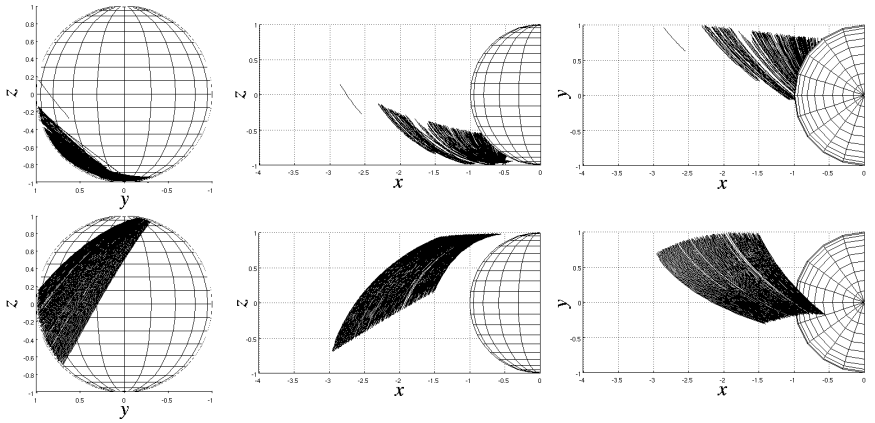


Fig. 3. Eclipse orbit geometry for the NPI observations during 2004 (top) and early 2005 (bottom). The time periods correspond to the dates given in Table 1. From left to right the yz-, xz-, and xy-planes in the Mars Solar Orbital (MSO) coordinate system, where the x-axis is toward the Sun, the z-axis is perpendicular to the planet's velocity, in the northern ecliptic hemisphere, and the y-axis completes the right handed system (and is approximately opposite to Mars' velocity vector).

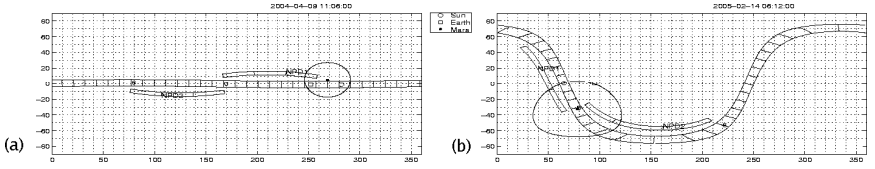


Fig. 4. Examples of NPI FOVs for eclipse observations in (a) 2004 and (b) 2005. The coordinate system is ecliptic with longitudes on x-axes, and latitudes on y-axes. The 32 NPI sectors' FOVs (look directions) is the continuous band with sectors 0 and 8 indicated. For comparison, the FOVs of the six NPD sectors are also shown. Mars limb is the closed line with the planet center indicated. The sun direction is shown by a small disc, and Earth by a square. We can note that MEX is in Earth communication pointing in (a) and nadir pointing in (b).

estimates were achieved. There is also priority effect due to the electronics, where some counts are shifted to higher numbered sectors that we have not attempted to compensate for in the present analysis. This effect can be viewed as a reduction of the in-plane angular resolution of the sensor, since we get a smoothing between adjacent sectors. During calibrations, the effect was responsible for moving approximately 40% of the counts in sector n to sector $n + 1$.

The position of the spacecraft and the look directions of the NPI sectors as a function of time were computed using the CSPICE[11] subroutine

Table 2. Inter-sector normalizations. The normalized counts for sector i, v_i, is computed from the raw counts, u_i, as $v_i = (u_i - \alpha_i)/\beta_i$. The relative sensitivities, β_i, are from ground calibrations, and the off-sets (dark currents), α_i, are estimated from minimum deep umbra count rates averaged in 60 s bins, for respective year.

Sector i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
β_i	1.11	1.07	1.08	1.07	1.00	0.99	1.10	1.04	1.00	1.31	1.30	1.16	1.05	0.85	0.88
α_i 2004	5.8	5.6	7.2	7.2	4.2	2.1	2.6	2.1	2.2	1.2	1.4	1.4	0.43	0.28	0.33
α_i 2005	10.3	9.2	11.5	12.6	5.9	3.4	3.9	3.9	3.1	1.4	1.5	1.4	0.20	0.17	0.15
Sector i	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
β_i	0.65	0.69	0.58	0.64	0.64	0.84	0.97	0.99	0.98	1.15	1.28	1.30	1.44	1.09	0.86
α_i 2004	1.9	2.3	3.4	2.1	1.3	1.6	1.8	1.3	1.2	1.4	1.7	0.62	1.2	1.2	1.8
α_i 2005	1.5	1.3	1.8	0.68	0.45	0.52	0.85	1.2	1.3	1.3	1.6	0.3	0.18	0.23	0.70

package. The raw data was processed using `pkttool-1.103`. The data and geometry information was merged into a text file that subsequently was read into Matlab for further processing and display. The software is available from the authors upon request.

3. Observations

Using the inter-sector normalization described in the previous section, we can compute skymaps for the NPI deep eclipse data that show the average count rate as a function of look direction ecliptic longitude and latitude. For each one-second observation, we first do an inter-sector normalization, according to Table 2, of the counts from each sector, then deposit the count in the cell that corresponds to the sector's boresight direction, and also accumulate the observation time on the grid. This is done for all the data to be included. In Fig. 5, we show the logarithm of the total time and counts accumulated for each cell on the longitude–latitude grid. Finally, by dividing the accumulated counts in each cell by the total observation time we get an average count rate.

Such sky maps for 2004 and 2005 are shown in Fig. 6. In 2004, we note signals from the general Mars direction (see Fig. 4 for Mars' approximate location). These are probably charge exchange ENAs from the Mars–solar wind interaction and UV scattered and refracted in Mars' atmosphere. In

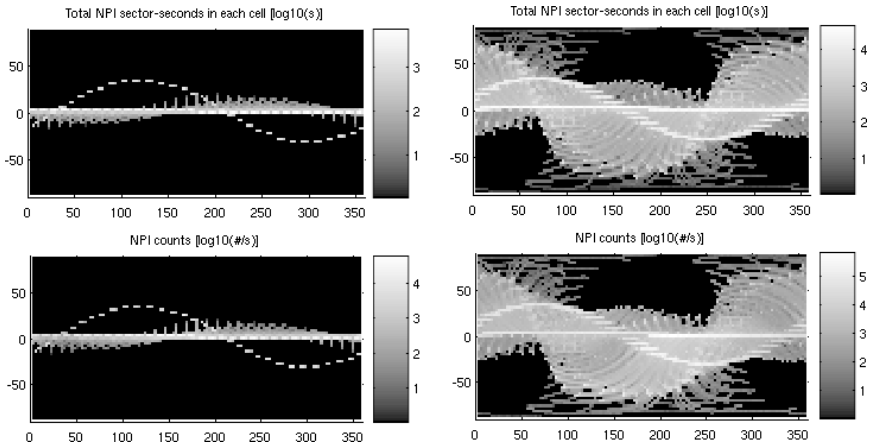


Fig. 5. Logarithmic sky maps of total sector time for each cell (top) and total number of counts (bottom) for NPI eclipse observations in 2004 (left) and 2005 (right). The coordinate system is ecliptic with longitudes on x-axes, and latitudes on y-axes, and the grid is 128×64 cells. Note the different color scales. Black denotes no coverage.

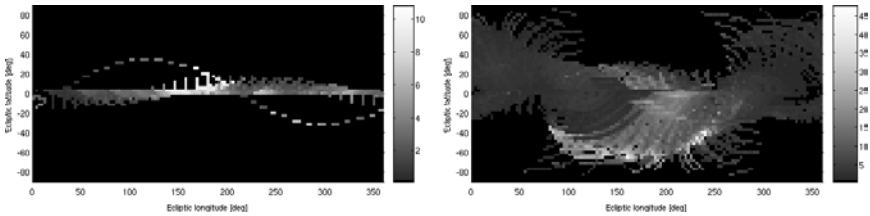


Fig. 6. Sky maps of average count rates (1/s) for NPI eclipse observations in 2004 (left) and 2005 (right), i.e., we see the average count rate that the NPI sectors have registered when pointing in that direction. The coordinate system is ecliptic with longitudes on x-axes, and latitudes on y-axes, and the grid is 128×64 cells. Note the different color scales. The maximum value on the colorbar is 6 in 2004, and 35 in 2005. Black denotes no coverage.

this coordinate system, Mars will not have constant size and position during an orbit and from orbit to orbit, but for the deep umbra data the change is small (Mars can increase in apparent size by 50°) and does not affect the different signals discussed later on. Another signal is seen coming from Mars' ram direction (the direction of the planets velocity vector) around 200° longitude. We will get back to this signal in the 2005 data. Note that there is also a signal from below the ecliptic near 0 longitude. We also see that the coverage out of the ecliptic is sparse during 2004. The cell size of the longitude–latitude grid was chosen smaller than the sector FOV (128×64 cells). Since we have many observations, we can achieve this superresolution. In 2005 we have more observations away from the ecliptic plane and we directly see the very strong emissions from below the ecliptic plane. These correspond well with the galactic plane.

The different characteristics of the individual sectors (sensitivity and noise level as discussed in the previous section) in a way make the NPI composed of 30 different sensors. Therefore, one has to be careful when comparing observations from different sectors. That the strong signals seen in the sky maps are not artifacts of the normalization and averaging of different sectors has been checked by identifying the signals in the time serie of individual sectors, and it is also possible to see the features discussed if one looks at sky maps for individual sectors.

4. Analysis

The big question to try and answer is what signals are from ENAs and what are from UV? What is neutrals and what is photons? Some approaches to separate UV and ENA signals are as follows:

- Compare the observed sky maps with UV sky maps.

- Examine the time evolution of the signals. Stellar UV sources should be stable in position and intensity over time.
- Examine data from other ENA detectors to identify similar signals. Assuming an hydrogen ENA signal in the 1 keV range, the current possibilities are the NPD on MEX and LENA[12] on IMAGE. This is a topic for future studies.

Below we will address the first two items, but first of all: could we be seeing something else than ENAs or UV, e.g., X-rays? Probably not, comparison with published sky maps in other wavelengths show no similarities to the features in Fig. 6.

Let us now compare the NPI skymaps in Fig. 6 with a UV skymap. In Fig. 7, we show SOHO/SWAN UV sky maps.[13] To the left in ecliptic and to the right in galactic coordinates. At the top is the original data, and at the bottom the same data smoothed by an NPI sector angular response function Appendix A. The Milky Way is clearly visible both in the UV and NPI skymaps. In ecliptic coordinates it is the U-shaped structure. We can note that by comparing countrates in Fig. 6 and intensity in Fig. 7 we can estimate the UV response of the NPI: 1 kR (uniformly over a sector) corresponds approximately to 10 counts per second, assuming that the galactic emissions are only UV. We can also compare the lower right plate in Fig. 6 with the NPI sky-map in galactic coordinates shown in Fig. 8. Most of the strong UV signals in the galactic plane correspond well to the NPI measurements. The UV source at longitude −150 is, however, an exception.

Turning now to the time evolution of the observed signals. We see in Fig. 6 that there is a source in the ecliptic around longitude 200° in 2004 and 2005. However, it has moved to a higher longitude by approximately 25° from 2004 to 2005, if it is the same signal. However, this shift is inconsistent with a UV source.

We further investigate the time evolution by examining how the ecliptic plane signals evolve over time in Fig. 9. Both years seem to include signals that are constant in longitude (suggesting UV) and change longitude at the same rate as the Sun (maybe Mars related), but also other changes are visible (possible ENAs), e.g., the emissions around longitude 200 look different in 2004 and 2005, and do also vary over time. In Fig. 10 we show how the galactic plane signals evolve over time. We see a clear change of the signal at around −45° longitude over time. This change is present even if we only look at counts from the hemisphere away from Mars, excluding a Mars-related source.

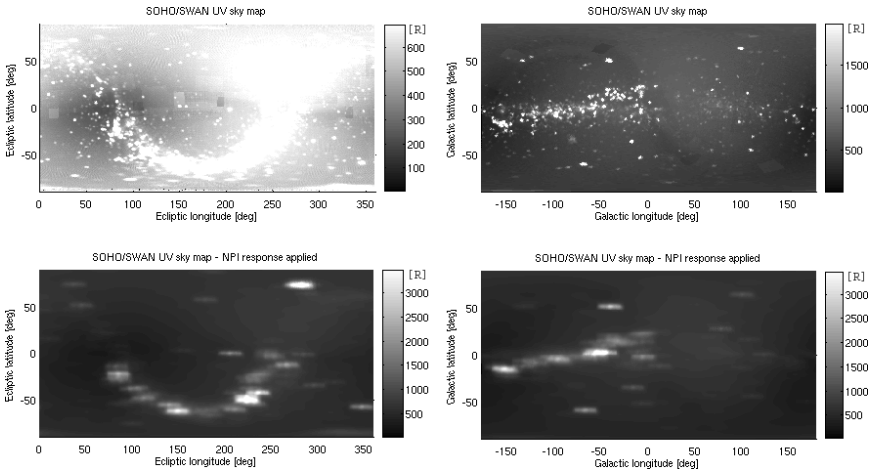


Fig. 7. SOHO/SWAN all-sky UV Maps on one degree resolution longitude–latitude grids in the unit of Rayleighs. The instrument is sensitive to the wavelengths 117–180 nm. The top row shows the raw images and in the bottom row we have convolved the images with the response function of one NPI sector (from calibrations). The left column is in ecliptic coordinates, and the right is in galactic coordinates. Note that the intensity ranges in the plots were chosen to enhance the features. The maximum in the raw images is 65 and 5 kR in the smoothed images. Data provided by Teemu Mäkinen on behalf of the SWAN team.

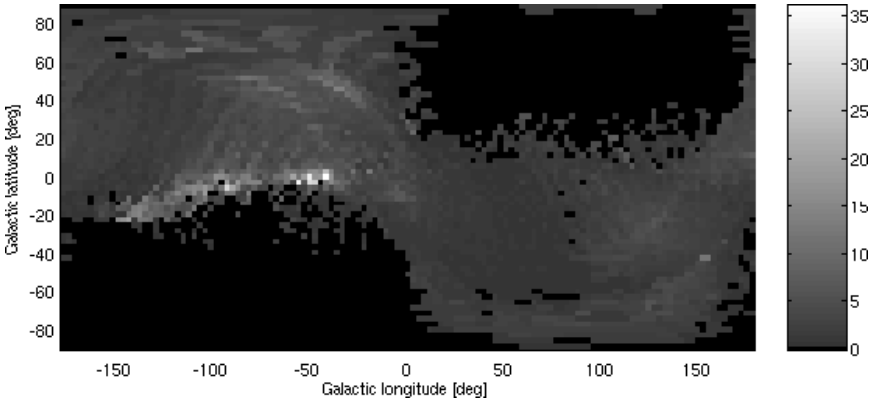


Fig. 8. The NPI skymap for 2005 in galactic coordinates (longitude–latitude). Otherwise the same as Fig. 6.

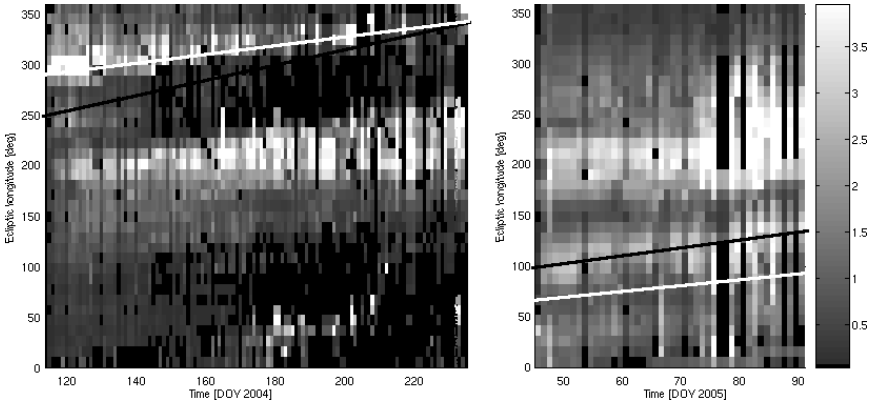


Fig. 9. Time evolution of the signal in the Ecliptic plane during 2004 (left) and 2005 (right). Each column represent average count rates (1/s) during 1 day. The x-axes are day of year, and the y-axes are ecliptic longitudes. Only observations by sectors looking at most 6° off the ecliptic plane were included. For 2004 we included all umbra observations. The dark lines represent Earth's longitude, and the light lines the Sun's longitude.

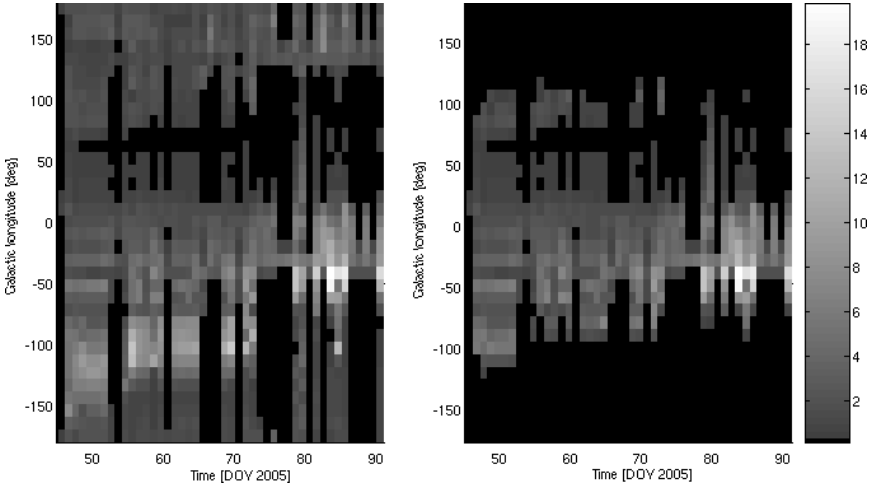


Fig. 10. Time evolution of the signal in the Galactic plane during 2005. Identical format as in Fig. 9 but for the Galactic plane, and with Galactic longitude and latitude. To the right, all observations in the hemisphere of directions centered at Mars were removed.

5. Discussion

The NPI has observed ENA and UV fluxes in the shadow of Mars in 2004 and 2005. The position of most of the strongest emissions in the NPI sky maps correspond fairly well to UV sources, but there is at least one signal that shifts position between the years, indicating that there are ENAs in this signal and that the ENAs seem not to be related to the Mars–solar wind interaction (they are well away from the Mars direction). In addition, there are ENAs seen in the eclipse that emanates from the Mars–solar wind interaction judging from the direction they come from (near Mars' limb).

There does not seem to be any obvious correlation between the NPI observations and the neutral streams discussed by Collier *et al.*,[10] since those fluxes have incoming directions of 260°–290° ecliptic longitude, and we have not identified any possible neutral atom fluxes from these directions in the NPI data. However, all previous observations of interplanetary ENAs are indirect (except the ISN observations[9]). Thus, it is difficult to rule out that there is a connection between the NPI observations and previous observations.

The next step in this investigation is to identify features that change over time and examine the response of individual sectors to these signals, and how they vary over time. Also, one should investigate if any such signal can be identified in MEX NPD or IMAGE LENA data (this would confirm that a signal is interplanetary since IMAGE is in Earth orbit), detectors that are not UV sensitive. Since these detectors also have velocity resolution, we could get more information on the ENA populations seen. A future comparison with measurements by the ASPERA-4 NPI sensor onboard ESA's Venus Express mission could also confirm or rule out a possible interplanetary ENA signal. Venus Express is scheduled to arrive at the planet in April 2006.

Acknowledgment

We thank Teemu Mäkinen at the Finnish Meteorological Institute (FMI), and the SWAN team, for providing UV all-sky data.

Appendix A. NPI Angular and Energy Response

Here, we present the NPI angular response, as used in the smoothing of the UV skymap in Fig. 7. The angular response was measured during on-ground calibrations with an MCP bias voltage of −2500 V and is shown in Table 3.

Table 3. The NPI response azimuth–elevation, 31×14, matrix for one sector. The unit is arbitrary, and normalized to a maximum value of 1. The sum of all entries is 82.0539, so to get a smoothing mask with a unit sum we should divide each entry by that number. We define the boresight direction at the maximum element $(17, 8)$. The step in azimuth and elevation is one degree, so element (i, j) at row i, column j, gives the response in azimuth i–17°, and elevation j–8° relative the boresight direction.

0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.040	0.032	0.032	0.056	0.015	0.079	0.111	0.118	0.094	0.118	0.086	0.062	0.070	0.054
0.007	0.007	0.023	0.023	0.069	0.084	0.131	0.162	0.177	0.177	0.069	0.100	0.046	0.038
0.030	0.030	0.023	0.061	0.076	0.084	0.199	0.245	0.222	0.237	0.130	0.099	0.030	0.030
0.015	0.015	0.015	0.054	0.116	0.193	0.295	0.428	0.453	0.227	0.227	0.039	0.055	0.071
0.015	0.039	0.015	0.063	0.141	0.291	0.668	0.628	0.525	0.383	0.226	0.085	0.070	0.015
0.015	0.023	0.031	0.092	0.208	0.355	0.562	0.646	0.560	0.437	0.237	0.107	0.038	0.022
0.038	0.022	0.030	0.098	0.220	0.304	0.502	0.563	0.494	0.365	0.266	0.083	0.030	0.038
0.022	0.022	0.030	0.083	0.137	0.381	0.519	0.656	0.596	0.367	0.275	0.091	0.038	0.076
0.023	0.015	0.038	0.077	0.216	0.434	0.683	0.583	0.663	0.422	0.203	0.109	0.062	0.039
0.031	0.023	0.047	0.039	0.189	0.490	0.681	0.594	0.475	0.285	0.214	0.079	0.071	0.031
0.015	0.031	0.047	0.086	0.220	0.306	0.571	0.601	0.608	0.373	0.186	0.061	0.054	0.015
0.015	0.015	0.046	0.038	0.232	0.489	0.552	0.601	0.547	0.423	0.164	0.102	0.039	0.047
0.000	0.015	0.039	0.094	0.268	0.418	0.717	0.826	0.534	0.369	0.227	0.094	0.070	0.046

Table 3. (*Continued*)

0.015	0.031	0.047	0.117	0.243	0.417	0.805	0.869	0.641	0.372	0.230	0.095	0.055	0.023
0.015	0.015	0.015	0.151	0.222	0.476	0.808	0.965	0.656	0.568	0.299	0.062	0.055	0.015
0.007	0.015	0.031	0.119	0.270	0.592	0.875	1.000	0.825	0.503	0.277	0.081	0.057	0.016
0.024	0.024	0.073	0.089	0.315	0.353	0.821	0.896	0.701	0.540	0.256	0.030	0.023	0.053
0.022	0.037	0.053	0.068	0.184	0.461	0.610	0.612	0.621	0.319	0.195	0.047	0.023	0.023
0.015	0.047	0.047	0.062	0.217	0.409	0.454	0.819	0.617	0.409	0.248	0.082	0.014	0.007
0.029	0.029	0.052	0.082	0.159	0.320	0.368	0.625	0.233	0.234	0.117	0.086	0.047	0.008
0.023	0.031	0.047	0.127	0.183	0.359	0.537	0.554	0.516	0.307	0.129	0.040	0.016	0.008
0.024	0.024	0.057	0.082	0.215	0.365	0.575	0.710	0.443	0.335	0.159	0.050	0.059	0.025
0.016	0.008	0.041	0.057	0.196	0.356	0.482	0.525	0.457	0.226	0.092	0.030	0.022	0.000
0.000	0.037	0.037	0.045	0.192	0.294	0.625	0.520	0.595	0.312	0.104	0.032	0.049	0.016
0.008	0.008	0.016	0.067	0.202	0.388	0.482	0.688	0.391	0.281	0.068	0.034	0.017	0.034
0.000	0.008	0.026	0.052	0.097	0.212	0.392	0.420	0.333	0.189	0.081	0.009	0.009	0.018
0.018	0.027	0.036	0.045	0.125	0.268	0.461	0.396	0.340	0.208	0.103	0.008	0.016	0.000
0.008	0.008	0.016	0.058	0.084	0.143	0.134	0.253	0.186	0.194	0.059	0.025	0.000	0.034
0.017	0.008	0.051	0.042	0.034	0.051	0.043	0.086	0.094	0.077	0.025	0.008	0.017	0.008
0.017	0.000	0.017	0.026	0.008	0.034	0.008	0.043	0.008	0.008	0.043	0.000	0.008	0.017

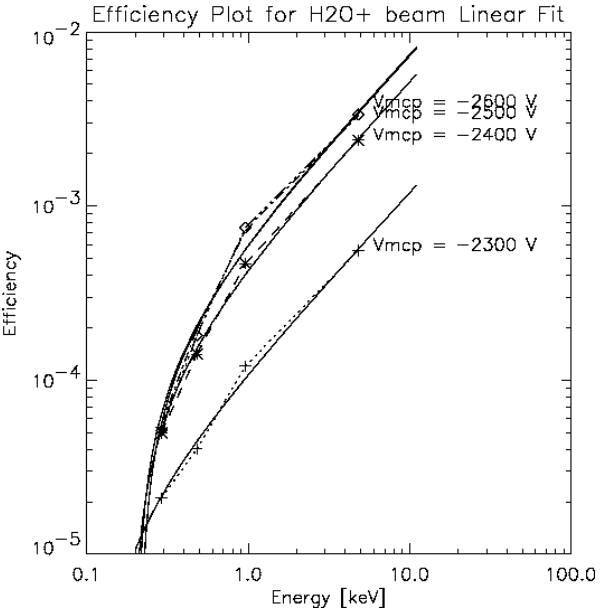


Fig. 11. The NPI efficiency to H_2O^+ ions as a function of energy from on-ground calibrations. The different curves are for different MCP bias voltages. The in flight MCP bias voltage for the time period discussed here was −2400 V.

We also show the NPI energy response from calibrations in Fig. 11. Given an inter-sector normalized count rate, v_i, as defined in Table 2, we can estimate the incoming flux of neutrals as $f \approx v_i/(\epsilon G)$ $(\mathrm{cm}^{-2}/\mathrm{sr}/\mathrm{s})$, where the geometrical factor[14] $G \approx 2.7 \times 10^{-3}$ $(\mathrm{cm}^2\,\mathrm{sr})$, and ϵ is the energy-dependent efficiency from Fig. 11, e.g., for 1 keV neutrals we have $\epsilon \approx 4 \times 10^{-4}$, so a count rate of one per second would correspond to a flux of 2×10^5 $(\mathrm{cm}^{-2}/\mathrm{sr}/\mathrm{s})$.

References

1. S. Barabash, R. Lundin, H. Andersson, J. Gimholt, M. Holmström, O. Norberg, M. Yamauchi, K. Asamura, A. J. Coates, D. R. Linder *et al.*, in *Mars Express: The Scientific Payload*, ed. A. Wilson (ESA Special Publication, SP-1240, 2004), pp. 121–139.
2. S. Barabash, R. Lundin, T. Zarnowiecki and S. Grzedzielski, *Adv. Space Res.* **16** (1995) 81–86.

3. S. Barabash, O. Norberg, R. Lundin, S. Olsen, K. Lundin, P. C:son Brandt, E. C. Roelof, C. J. Chase, B. H. Mauk, H. Koskinen *et al.*, *Measurement Techniques in Space Plasmas*, eds. Robert F. Pfaff *et al.* (AGU, 1998), pp. 257–262.
4. H. Gunell, K. Brinkfeldt, M. Holmström, P. Brandt, S. Barabash, E. Kallio, A. Ekenbäck, Y. Futaana, R. Lundin, H. Andersson *et al.*, submitted to *Icarus* (2006), in press.
5. K. Brinkfeldt, H. Gunell, P. Brandt, S. Barabash, R. A. Frahm, J. D. Winningham, E. Kallio, M. Holmström, Y. Futaana, A. Ekenbäck *et al.*, submitted to *Icarus* (2006), in press.
6. E. Kallio, S. Barabash, K. Brinkfeldt, H. Gunell, M. Holmström, Y. Futaana, W. Schmidt, T. Säles, H. Koskinen, P. Riihelä *et al.*, submitted to *Icarus* (2005).
7. M. Holmström, S. Barabash and E. Kallio, *J. Geophys. Res.* **107** (2002) 10.
8. S. Barabash, M. Holmström, A. Lukyanov and E. Kallio, *J. Geophys. Res.* **107** (2002) 10.
9. M. Witte, M. Banaszkiewicz, H. Rosenbauer and D. McMullin, *Adv. Space Res.* **34** (2004) 61–65.
10. M. R. Collier, T. E. Moore, D. Simpson, A. Roberts, A. Szabo, S. Fuselier, P. Wurz, M. A. Lee and B. T. Tsurutani, *Adv. Space Res.* **34**, 1 (2004) 166–171.
11. C. H. Acton, *Planet. Space Sci.* **44**, 1 (1996) 65–70.
12. T. E. Moore, D. J. Chornay, M. R. Collier, F. A. Herrero, J. Johnson, M. A. Johnson, J. W. Keller, J. F. Laudadio, J. F. Lobell, K. W. Ogilvie *et al.*, *Space Sci. Rev.* **91** (2000) 155–195.
13. J. L. Bertaux, E. Kyrola, E. Quemerais, R. Pellinen, R. Lallement, W. Schmidt, M. Berthe, E. Dimarellis, J. P. Goutail, C. Taulemesse *et al.*, *Sol. Phys.* **162** (1995) 403–439.
14. K. Brinkfeldt, *Instrumentation for Energetic Neutral Atom Measurements at Mars, Venus, and the Earth*, Doctoral Thesis, Umeå University, Sweden, 2005.

OBSERVATIONAL FEATURES OF THE SECONDARY LAYER OF THE MARTIAN IONOSPHERE

HAI-REN LIAO[†,¶], JING-SONG WANG[†,‡,*], HONG ZOU[†]
and XIAO-DONG WANG[†]

[†]*School of Earth and Space Sciences, Peking University, Beijing, China*
[‡]*China Meteorological Administration, Beijing, China*
[*]*wangjs@cma.gov.cn*
[¶]*wjs@pku.edu.cn*

The behaviors of both the density and height of the peak of the primary Martian ionospheric layer have been intensively investigated. On the contrary, the study on the secondary layer is mainly focused qualitatively on the photochemistry processes leading to this layer. The observational behaviors of the secondary peak, i.e., that of the peak height and the peak density, are seldom emphasized. In this paper, the Martian ionospheric profiles obtained by the Mars Global Surveyor (MGS) Radio Science (RS) experiments during two continuous periods are investigated. For 1,678 of the available 3,377 ionospheric electron density profiles, the secondary peak can be easily distinguished. Both the solar zenith angle dependence and the solar activity dependence of the secondary peak are investigated. E10.7, the integrated extreme ultraviolet (EUV) energy flux at the top of the atmosphere reported in units of 10.7 cm radio flux, is used as the proxy of solar radiation in this study. However, the main photoionization source for the secondary layer is not mainly EUV but soft X-ray. This shortcoming should be improved by using X-ray flux measurements; however, this is beyond the scope of this study.

1. Introduction

The observed electron density profiles of the Martian ionosphere always present a primary or upper layer, and a trend of a secondary or lower layer, which could behave as a peak, a ledge or a shoulder, but sometimes even not present.[1,2] A third layer is also noticed by authors (e.g., Refs. 3 and 4).

The behaviors of both the height and density of the primary peak have been intensively studied,[5–11] while the study on the secondary layer has been mainly focused on the photochemistry processes.[3,12,13] It is believed that the secondary layer is produced by the ionization due to the solar soft X-ray radiation and impact of photo- and secondary electrons.

The Radio Science (RS) experiment aboard Mars Global Surveyor (MGS)[18] has obtained many electron density profiles of the Martian

ionosphere (http://nova.stanford.edu/projects/mgs/eds-public.html). Rishbeth and Mendillo[14] got the peak densities of the two Martian ionospheres layers measured by MGS during March 9–27, 1999, and compared their properties with that of the terrestrial E and $F1$ layers. But for the majority of the profiles measured by MGS, the observational features of the secondary layer have not been explored.

In this paper, more than 3,000 ionospheric electron profiles are used to investigate the behaviors of the secondary layer.

2. Data Analysis

Among all of the accessible MGS data sets, two data sets are obtained in periods both lasting for more than half year which is the longest and hence suitable for statistical study. One period is from November 1, 2000 to June 6, 2001 (period I), the other is from November 1, 2002 to July 2, 2003 (period II).

The profiles can be catalogued into four types according to the appearance of the secondary layer, say, the variation of the electron density N with the altitude z at the point of inflexion (as shown in Fig. 1):

$$\begin{aligned}
&\text{Type 1}: \quad \partial N/\partial z = 0 \quad \text{and} \quad \partial^2 N/\partial^2 z < 0,\\
&\text{Type 2}: \quad \partial N/\partial z = 0 \quad \text{and} \quad \partial^2 N/\partial^2 z = 0,\\
&\text{Type 3}: \quad \partial N/\partial z > 0,\\
&\text{Type 4}: \quad \text{no inflexion.}
\end{aligned}$$

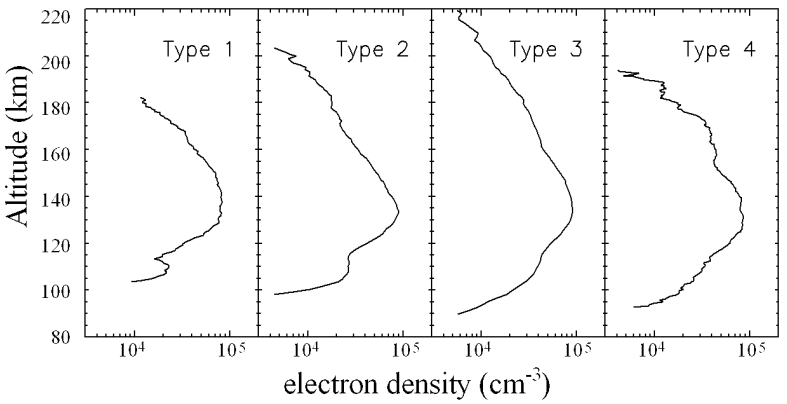


Fig. 1. Catalogues of the ionospheric profiles obtained by MGS/RS according to the appearance of the secondary layer (see text for details).

Without any treatment in advance, among all of the 3,377 profiles, 2,667 can be catalogued to type 1, i.e., with a distinct secondary layer. However, to avoid "pseudo-peak" caused by observation noises, a 5-point running average is applied to all the 2,667 profiles and 1,678 profiles of them can be still treated as type 1. Therefore, these 1,678 profiles have convincing secondary layer and are left for further investigation. The density and height of the peak are easily read from each profile.

The dependence on the solar zenith angle and the solar radiation of both the peak density and peak height of the secondary layer are investigated here. The flux of solar extreme ultraviolet (EUV) radiation near Mars is not available, but it can be derived from the E10.7 at the Earth orbit with distance and direction correction. E10.7 is the integrated EUV energy flux at the top of the atmosphere reported in units of 10.7 cm radio flux (F10.7)[15,16] and is available at http://SpaceWx.com. Actually the 10.7 cm radio flux (presented by F10.7) is has no physical connection with the atmospheric processes, while the EUV radiation (presented by E10.7) is the main ionizing source for ionosphere, so E10.7 is a better proxy than F10.7 for ionizing flux. In this paper, the daily averaged E10.7 index is used as a proxy of the solar EUV flux.[15,19]

3. Results

The occultation location on the Mars surface varies with time during the MGS/RS measurement and about every 12 continuous orbits of MGS around Mars corresponds to a period of the longitude variation of the occultation location.[18] Since planetary scale waves in the neutral atmosphere can appear as longitudinal variations in the Martian ionosphere,[1] a daily running average is applying to the dataset to reduce the effects of the longitudinal variation.

Figure 2 shows the dependence of the daily averaged secondary peak density, N_{m}, and height, z_{m}, upon the solar zenith angle (SZA, denoted as χ). The left panel is for period I, the right panel is for period II. Analogous to the description of the SZA dependence of the Martian ionospheric primary peak, i.e.,

$$N_{\mathrm{m}} = N_{\mathrm{m0}} \cos^{k} \chi, \tag{1}$$

$$z_{\mathrm{m}} = z_{\mathrm{m0}} + H \ln \sec \chi, \tag{2}$$

where N_{m0} and z_{m0} are the sub-solar ($\chi = 0$) peak density and height, respectively, k is a parameter that depends on the photochemical process,[6]

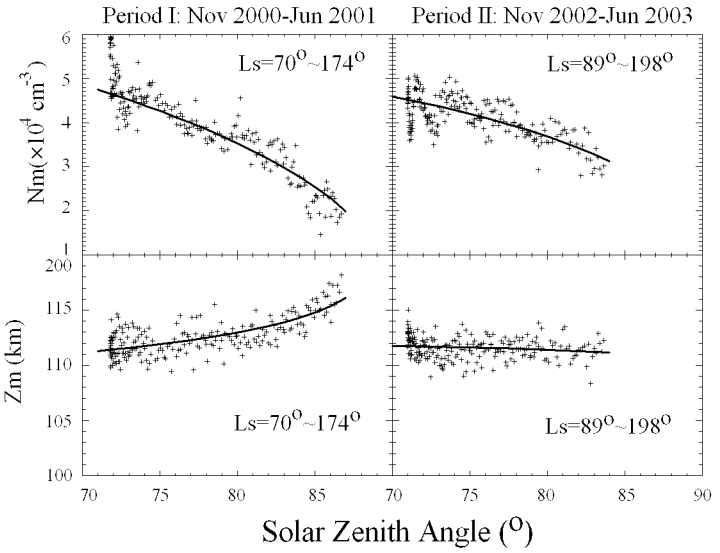


Fig. 2. The solar zenith angle dependence of the peak density and the peak height of the secondary electron density maximum. The solid curves are best fit by the expression indicated in each panel.

and H is the mean neutral scale height in the upper atmosphere, the SZA dependence of the secondary layer can be described as

$$N_{\rm m}(\text{period I}) = 8.1 \times 10^4 \cos^{0.48}\chi, \tag{3}$$

$$N_{\rm m}(\text{period II}) = 6.5 \times 10^4 \cos^{0.32}\chi, \tag{4}$$

$$z_{\rm m}(\text{period I}) = 108 + 2.6 \ln \sec\chi, \tag{5}$$

$$z_{\rm m}(\text{period II}) = 112 - 0.5 \ln \sec\chi. \tag{6}$$

It is found that in both periods the apparent behaviors of the secondary layer depart from an ideal Chapman-α layer for which $k = 0.5$, while $k = 0.57$ was obtained for the primary layer.[6] And most noticeable is that the peak height during period II is nearly constant with the increase of SZA. It is quite different from the general behaviors of the primary layer as well as that of the secondary layer during period I.

Figure 3 shows the variations of the secondary peak density and height, the corresponding E10.7* and SZA with the solar longitude ($L_{\rm s}$, which can be used as the proxy of time herein), where E10.7* is the solar E10.7 index at the Mars orbit derived from the Earth orbit value after two corrections. The time shift caused by the angle of Earth–Sun–Mars can be corrected by

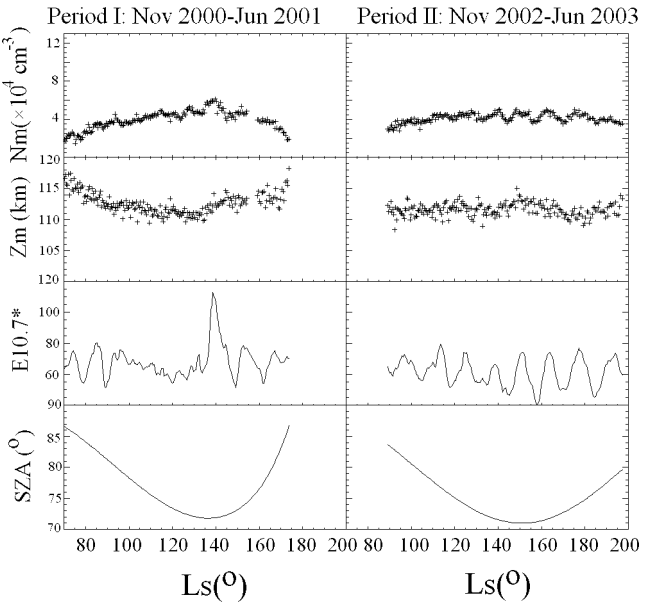


Fig. 3. The peak density and peak height of the secondary layer, the corresponding E10.7* and solar zenith angle (SZA) versus the solar longitude (L_s, used herein as the proxy of time).

considering the locations of both planets and the rotation speed of the Sun. For example, at the time of observation Mars was located about x degrees away from the Earth in the direction of their orbital motion, then the date of E10.7* is equal to the date of E10.7 plus $(x/360) \times 27$ days, where 27 is the average period of solar rotation. The maximum correction for the two used datasets is ∼13 days (when the Earth and Mars nearly opposite), i.e., half of a solar rotation period. Possible individual variation of the solar radiation with time scale less than one solar rotation period will not influence the statistic analysis for nearly half year. The attenuation of the solar radiation from the Earth's orbit (∼1 AU) to the Mars' orbit (∼1.6 AU) can be corrected by the law of $1/r^2$ with the measured Sun–Planet distances (r). The left column in Fig. 3 is for period I, and the right column is for period II.

It can be seen that other than the SZA dependence shown in Fig. 2, there are still obvious variations of shorter time scale, especially the positive correlation between the peak density and E10.7*, while no prominent correlation exists between the peak height and E10.7*.

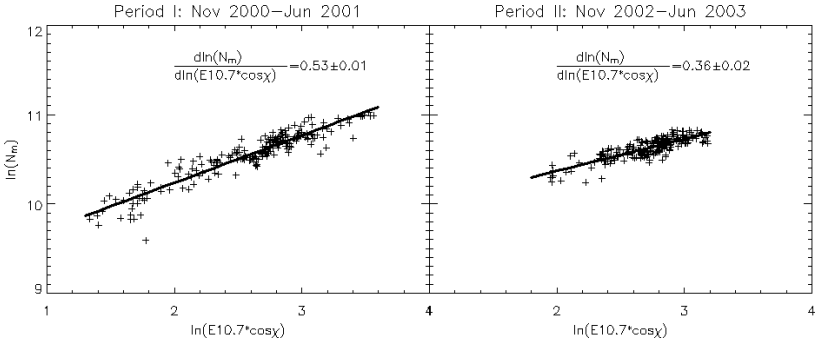


Fig. 4. The solar radiation dependence of the secondary peak density for both periods.

Table 1. Summary on the data sets of the two periods.

Period (Number of profiles)	Date	L_s(°)	SZA(°)	LAT(°N)
Period I (1572)	November 1, 2000–June 6, 2001	70–174	71–87	63–86
Period II (1805)	November 1, 2002–June 4, 2003	89–198	71–84	60–85

The dependence of the ionospheric peak density on the solar radiation as well as SZA under a photochemical equilibrium is normally described as (e.g., Ref. 10)

$$\frac{d \ln N_{\mathrm{m}}}{d \ln(E10.7 \times \cos\chi)} = k. \tag{7}$$

For an ideal Chapman-α layer k is the same as in Eqs. (1) and (2) and expected to be 0.5.

Figure 4 shows the dependence of N_{m} on E10.7* and SZA during both considered periods. The correlation between the solar radiation and the ionospheric peak density is convincing. For period I, $k = 0.53$ and period II, $k = 0.36$, both close to the value in Eqs. (3)–(6) but still deviate from 0.5.

However, the discrepancy between the two periods is noteworthy. As shown in Table 1 and Fig. 3, L_{s} and SZA variations as well as the occultation locations are quite similar for both periods, the systematical disagreement between these two periods needs further discussion.

4. Discussion and Summary

Total 3,377 Martian ionospheric electron density profiles measured by MGS/RS during period I (from November 1, 2000 to June 6, 2001) and

period II (from November 1, 2002 to June 4, 2003) are investigated, and about 50% (1,678) profiles present a detectable secondary layer.

The peak density of the secondary layer for both periods decrease with the increase of the solar zenith angle, and can also described by $N_{\mathrm{m}} = N_{\mathrm{m0}} \cos^{k} \chi$ just like that for the primary layer. However, the SZA dependence of the peak height is different for each period: for period I, the peak height increase with SZA, similar to that for the primary layer; on the contrary, for period II the peak height is nearly constant when SZA is increasing (Fig. 2).

The peak density of the secondary layer is obviously modulated by the solar radiation (Fig. 3) just like the case for the primary layer. The peak density is positively correlated to the solar radiation (Fig. 4). Again, there are apparent differences between the solar radiation dependence for the two periods.

Since the variations of SZA, L_{s} and the latitude coverage for the two datasets are quite similar, the disagreement of the behaviors of the secondary layer for both periods needs to be attributed to other reason(s).

The first candidate is the solar irradiance. Table 2 summarizes the solar irradiance conditions during these two periods. The average E10.7 is 172.7 during period I, and 148.4 during period II. These correspond to E10.7* value of 67.8 and 61.5, respectively. Obviously, the solar activities are different for these two periods.

In this paper, the E10.7 index is used as a measure of the solar irradiance input to the Martian secondary layer. In fact, however, one of the dominant ionization source for the Martian ionospheric secondary layer is soft X-rays.[12,14] The increase in the fluxes of soft X-rays is not linear with the increase of average E10.7 and the secondary peak behaves differently at different solar radiation conditions.[12,13] These could directly contribute to the disagreement of the secondary peak's behaviors at different time. A reasonable way to improve is to use direct X-ray flux measurements for further investigation; however, it is beyond the goal of this paper.

Table 2. Summary of solar fluxes conditions of two periods.

Period (days)	Minimum E10.7*	Maximum E10.7*	Average E10.7*	Deviation of E10.7*
Period I (209)	51.3	113.1	67.8	10.4
Period II (216)	40.1	79.4	61.5	8.6

Second, solar soft X-ray radiation is not the only ionization source for the Martian secondary layer. The other important source is impact of photo- and secondary electrons. Modeling by Fox *et al.*[12] shows that at different solar activities, the relative importance of the photoionization and electron impact ionization is different. In the analysis above, either SZA or solar radiation dependence of the peak behaviors is derived under the assumption of photoequilibrium, and of course the effects of the photo- and secondary electron impact are not included. This consideration could be the most important factor to interpret the peak height's "abnormal" SZA dependence: nearly constant with the increase of SZA.

Moreover, if the scale height of the upper atmosphere, H, has a gradient, k is not 0.5 but becomes $0.5(1 + dH/dz)$.[17] Models show that such a gradient does exist in the Martian upper atmosphere (e.g., Ref. 11). If the scale height gradients are different in the two considered periods, it will contribute to the behavior differences discussed here. However, it should be also kept in mind that this value was obtained after only solar zenith angle correction. There are many other factors that can modulate the morphology of the ionosphere (e.g., Refs. 7–11). A comprehensive consideration is needed for the determination of an "average" or "background" k value, which is beyond the scope of this paper.

Anyway, the quantitative comprehension to the behaviors of the secondary layer of the Martian ionosphere still needs further studies, especially on the disagreement of the behaviors between different observation periods.

Acknowledgments

We thank the MGS Radio Science Team for making their ionospheric profile data available on a public website. This research was funded by the National Natural Science Foundation of China (No. 40374058).

References

1. S. W. Bougher, S. Engel, D. Hinson and J. Forbes, Mars Global Surveyor radio science electron density profiles: neutral atmosphere implications, *Geophys. Res. Lett.* **28** (2001) 3091–3094.
2. C. R. Martinis, J. K. Wilson and M. J. Mendillo, Modeling day-to-day ionospheric varirability on Mars, *J. Geophys. Res.* **108** (A10) (2003) 1383, doi:10.1029/2003JA009973.
3. J. L. Fox, Response of the Martian thermosphere/ionopshere to enhanced fluxes of solar soft X-rays, *J. Geophys. Res.* **109** (2004a) A11310, doi:10.1029/2004JA010380.

4. M. Pätzold, S. Tellmann, B. Häusler, D. Hinson, R. Schaa and G. L. Tyler, A sporadic third layer in the ionosphere of Mars, *Science* (2005) 837–839.
5. J. S. Bauer and M. H. Hantsch, Solar cycle variation of the upper thermosphere of Mars, *Geophys. Res. Lett.* **16** (1989) 373–376.
6. M. H. Hantsch and S. J. Bauer, Solar control of the Mars ionosphere, *Planet. Space Sci.* **38** (1990) 539–542.
7. J.-S. Wang and E. Nielsen, Behavior of the Martian dayside electron density peak during global dust storms, *Planet. Space Sci.* **51** (2003) 329–338.
8. J.-S. Wang and E. Nielsen, Evidence for topographic effects on the Martian ionosphere, *Planet. Space Sci.* **52** (2004a) 881–886.
9. J.-S. Wang and E. Nielsen, Solar wind modulation of the Martian ionosphere observed by Mars Global Surveyor, *Ann. Geophys.* **22** (2004b) 2277–2281.
10. T. K. Breus, A. M. Krymskii, D. H. Crider, N. F. Ness, D. Hinson and K. K. Barashyan, Effect of the solar radiation in the topside atmosphere/ionosphere of Mars: Mars Global Surveyor observations, *J. Geophys. Res.* **109** (2004) A09310, doi:10.1029/2004JA010431.
11. S. W. Bougher, S. Engel, D. P. Hinson and J. R. Murphy, MGS Radio Science electron density profiles: interannual variability and implications for the Martian neutral atmosphere, *J. Geophys. Res.* **109** (2004) E03010, doi:10.1029/2003JE002154.
12. J. L. Fox, P. Zhou and S. W. Bougher, The Martian thermosphere/ionosphere at high and low solar activities, *Adv. Space Res.* **17** (1995) 203–218.
13. J. L. Fox, Advances in the aeronomy of Venus and Mars, *Advances in Space Research* **33** (2004) 132–139.
14. H. Rishbeth and M. Mendillo, Ionospheric layers of Mars and Earth, *Planet. Space Sci.* **52** (2004) 849–852.
15. W. K. Tobiska, T. Woods, F. Eparvier, R. Viereck, L. Flyod, D. Bouwer, G. Rottman and O. R. White, The SOLAR2000 empirical solar irradiance model and forecast tool, *J. Atmos. Sol. Terr. Phys.* **62** (2000) 1233–1250.
16. W. K. Tobiska, SOLAR2000 irradiances for climate change, aeronomy, and space systems engineering, *Adv. Space Res.* **34** (2004) 736–1746.
17. M. Nicolet, *J. Atmo. Terr. Phys.* **1** (1951) 141.
18. D. P. Hinson, R. A. Simpson, J. D. Twicken, G. L. Tyler, Initial results from radio occultation measurements with Mars Global Surveyor, *J. Geophys. Res.* **104** (1999) 26997–27012.
19. W. K. Tobiska, Validating the solar EUV proxy E10.7, *J. Geophys. Res.* **106**, A12 (2001) 29969–29978.

This page intentionally left blank

MARTIAN ATMOSPHERE DURING THE 2001 GLOBAL DUST STORM: OBSERVATIONS WITH SWAS AND SIMULATIONS WITH A GENERAL CIRCULATION MODEL

TAKESHI KURODA*, ALEXANDER S. MEDVEDEV and PAUL HARTOGH
Max-Planck-Institute for Solar System Research
Max-Planck.Str.2, D-37191 Katlenburg-Lindau, Germany
**kuroda@mps.mpg.de*

The Earth-orbiting submillimeter wave astronomy satellite (SWAS) observed the global mean surface and atmosphere temperature on Mars as a function of the altitude. Unlike for the infrared spectrometers, the temperature retrievals from submillimeter instruments can be performed in the presence of the atmospheric dust. During the 2001 global dust storm on Mars, SWAS measured the atmosphere and surface temperature for aerocentric longitudes from $L_s = 166^\circ$ to 233°, and observed the temperature inversion in the lower atmosphere. We use a recently developed general circulation model of the Martian atmosphere to simulate the temperature and other atmospheric fields under the conditions corresponding to those during the SWAS measurements. The model takes into account the radiative effects of the atmospheric dust. Simulations show an overall agreement with the SWAS measurements. In particular, the model reproduces the inversion of the global mean temperature and the surface cooling detected by SWAS. Time series of the globally averaged surface and atmospheric temperature at 0.05 hPa (∼49 km) are also in a good agreement with the measurements. Without the dust, the model cannot reproduce these features. A brief discussion of the differences between the model and observations as well as their possible reasons is provided.

1. Introduction

The June 21, 2001 Mars opposition was a good opportunity for observations from Earth and Earth-orbiting satellites because the apparent size of the planet was largest in more than 10 years. In addition, soon after the opposition, the rise and growth of the global dust storm were observed. It started from smaller regional storms in and around the Hellas Basin on June 26, 2001 (aerocentric longitude, $L_s = 185^\circ$), and grew to cover the entire Mars in 2 weeks, as was recorded in the snapshots from Hubble telescope.[1]

*Corresponding author.

Global dust storms are the characteristic phenomena in the Martian meteorology. They occur only around the Martian perihelion, in southern springs or summers, although not on every Martian perihelion. Observations by the Mars-orbiting thermal emission spectrometer onboard Mars global surveyor (MGS-TES) revealed that the globally averaged daytime surface temperature decreased by ∼20 K during the 2001 global dust storm, whereas the globally averaged daytime atmospheric temperature at 0.5 hPa (∼25 km) increased by ∼40 K.[2]

The Earth-orbiting submillimeter wave astronomy satellite (SWAS) also observed the globally averaged surface temperature and vertical atmospheric temperature profiles of Mars during the onset of the 2001 global dust storm between $L_s = 166$ and 233°.[3] The peculiar features of Martian temperature observations in the submillimeter range are that the source function is nearly linear with temperature, and that the scattering and emission by the dust do not affect temperature retrievals due to the small particle sizes relative to the observed wavelength (the typical mean diameter of dust particles is ∼2 μm).[4] The latter is important because the temperature measurements are not contaminated by the large amount of aerosol during storms in the Martian atmosphere. The observations from SWAS revealed ∼20 K decrease in the surface temperature and ∼40 K increase in the atmospheric temperature at 0.5 hPa during the global dust storm. This is consistent with the MGS-TES measurements, and indicates that the temperature retrievals of SWAS are credible. In addition, SWAS can retrieve the atmospheric temperature at altitudes higher (up to 75 km) than MGS-TES (up to ∼35 km for nadir-geometry observations).[2] The SWAS measurements showed no global thermal effects of the storm above the 0.01 hPa (∼60 km) level.[3]

Temperature inversions are the phenomena which occur on Mars when the atmospheric temperature increases with height. The averaged over the visible disk temperature inversion was apparently observed by SWAS during the global dust storm (Gurwell, personal communication). It is thought that the inversions are caused by the increased absorption of the solar radiation by the dust lifted into the air.

In this paper, we use a recently developed general circulation model (GCM) of the Martian atmosphere[5] to simulate the circulation and temperature under the scenario corresponding to the conditions of the 2001 dust storm, and compare the model results to the SWAS measurements. Section 2 presents a brief description of the model and the dust scenario employed in the calculations. The results of the simulations are shown in Sec. 3. The discussion of the accuracy of the model is given in Sec. 4.

2. Model Description

The Martian GCM used for simulations has been derived from the Cologne model of the middle atmosphere (COMMA),[6] a terrestrial GCM which was extensively used for studies of the dynamics and photochemistry of the middle atmosphere. A detailed description of the Martian model is given in Hartogh *et al.*[5] The model has 32 grid points in the longitude and 36 grid points in the latitude; 118 vertical levels span the atmosphere from the deepest point on the Martian surface to approximately 130 km with a vertical grid step of about 1.14 km. The model employs the realistic topography, surface albedo and thermal inertia. The comprehensive set of physical parametrizations includes the surface energy balance scheme, CO_2 and dust radiation codes, gravity wave drag and turbulent diffusion parameterizations, and the CO_2 condensation/sublimation scheme with only thermal effects included. Cooling in 15 μm CO_2 band is calculated with the speeded-up version of the exact non-LTE radiative transfer code, the accelerated lambda iteration for atmospheric radiation and molecular spectra (ALI-ARMS).[7,8] Heating in the CO_2 near-infrared bands is computed after Forget *et al.*[9] The accuracies of these schemes are tested in Ref. 5.

The absorption, emission and scattering by the atmospheric dust are calculated for wavelengths between 0.2 and 200 μm using the two-stream discrete-ordinate/adding method.[10] The dust radiation scheme utilizes 19 representative wavelength bands: nine in the visible and 10 in the infrared spectral range. The corresponding radiative fluxes depend on the dust parameters, the refractive indices and the particle size distribution. The refractive indices of the aerosol for the visible wavelengths (shorter than 5 μm) are taken from Ref. 11, the montmorillonite 219b data of Toon *et al.*[12] are employed for the wavelengths between 5 and 17 μm, and constant values of Forget[13] are utilized for the wavelengths longer than 17 μm.[11–13] The dust particle size distribution is borrowed from Ref. 4. The latter uses the modified gamma function with the effective radius of 1.6 μm and the effective variance of 0.2 μm.[4] The dust parameters employed here are essentially the same as in CCSR/NIES Martian AGCM.[14] Hence, the further information on the dust parameters can be found therein. The spatial dust distribution in our simulations is prescribed by the functions which depend on the aerocentric longitude, latitude and pressure (or height). These functions were introduced to fit the observations by MGS-TES in 2001.[2,15,16] The vertical dust distribution is assumed theoretically from the vertical mixing and gravitational settling.[17] This scenario corresponds to the TES3

dust scenario in Ref. 14, and can be represented by the following analytic formulae:

$$\begin{aligned}\tau_*(L_\mathrm{s}) &= 0.05 + \frac{\tau_\mathrm{max} - 0.05}{2}[1 + \tanh\{(L_\mathrm{s} - L_\mathrm{s0})\alpha\}] \quad \text{for } L_\mathrm{s} < L_\mathrm{s1}, \\ \tau_*(L_\mathrm{s}) &= 0.05 + (\tau_\mathrm{max} - 0.05)\{\cos((L_\mathrm{s} - L_\mathrm{s1})/2)\}^\beta \quad \text{for } L_\mathrm{s} \geq L_\mathrm{s1},\end{aligned} \tag{1}$$

where $\tau_*(L_\mathrm{s})$ denotes the infrared (9–10 μm) dust opacity at 7 hPa calculated at the equator, south, and north poles as τ_eq, τ_S, and τ_N, respectively. The factor τ_max is the maximum opacity in each region, L_s0 is the aerocentric longitude at the time of the most intense dust storm, L_s1 is the aerocentric longitude for the time when the dust opacity begins to decrease, and α and β are the rates of the opacity increase and decrease, correspondingly. These factors are presented in Table 1. The optical depth at all latitudes $\phi, \tau(\phi, L_\mathrm{s})$, is then prescribed as follows:

$$\begin{aligned}\tau(\phi, L_\mathrm{s}) &= \tau_\mathrm{N} + \frac{\tau_\mathrm{eq} - \tau_\mathrm{N}}{2}\{1 + \tanh[5(45^\circ - \phi)]\}, \quad \phi > 0, \\ \tau(\phi, L_\mathrm{s}) &= \tau_\mathrm{S} + \frac{\tau_\mathrm{eq} - \tau_\mathrm{S}}{2}\{1 + \sin[2(45^\circ + \phi)]\}, \qquad \phi < 0.\end{aligned} \tag{2}$$

The vertical distribution of the dust mixing ratio Q is given by the following formulae, in accordance with those adopted by LMD-AOPP Martian AGCMs.[9,18]

$$\begin{aligned}Q &= Q_0 \exp\left\{0.007\left[1 - \left(\frac{p_0}{p}\right)^{70/z_\mathrm{max}}\right]\right\}, \\ z_\mathrm{max}(\phi, L_\mathrm{s}) &= 60 + 18\,\sin(L_\mathrm{s} - 160^\circ) - 22\,\sin^2\phi,\end{aligned} \tag{3}$$

where z_max is the cut-off altitude of the dust [km], p the pressure, p_0 the standard pressure (7 hPa), and Q and Q_0 are the dust mixing ratios at the levels p and p_0, respectively.

Table 1. Values of the factors used in Eq. (1) for prescribing the time-latitude profiles (from Ref. 14). N, EQ, and S correspond to the factors for calculating τ_N, τ_eq, and τ_S, respectively.

	N	EQ	S
τ_max	0.1	0.7	1.0
L_s0	210°	190°	192°
L_s1	230°	205°	205°
α	10	18	22
β	20	20	15

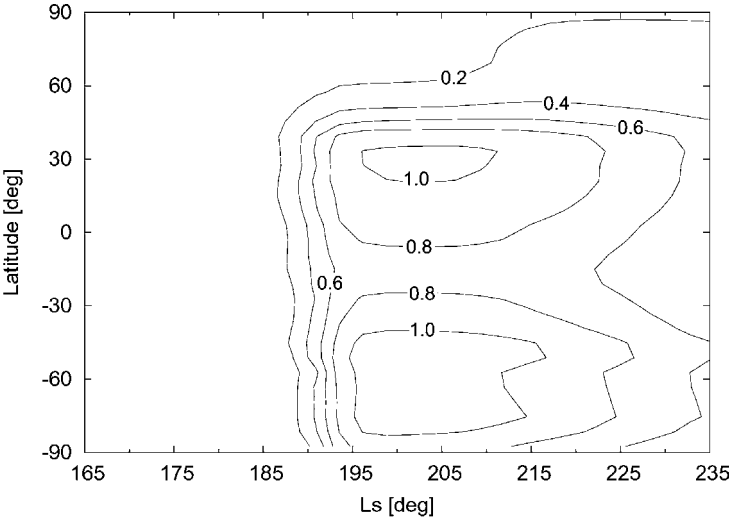


Fig. 1. The time series of the zonal mean dust opacity at 9 μm employed in the simulations.

The zonal-mean dust opacity distribution at the surface used in the simulations is shown in Fig. 1. Distributions of this type are based on observations,[16] and are commonly accepted in GCM studies. Note that the dust opacity values in Fig. 1 are affected by the topography with the global north-to-south elevation. As the dust is not transported in this simulation, no dynamical feedback is taken into account.

3. Model Results

At first, the model was spun up for 10 sols from the state of the rest and a globally uniform temperature $T = 160\,\mathrm{K}$ at the fixed L_s =160°. After reaching a practically equilibrium state, the model was run with L_s continuously varying from 160° to 240° to cover the entire period of measurements by SWAS. During the SWAS observations, the local time at the center of the visible Martian disk varied significantly from early afternoons at the start of the measurement sessions to mid-mornings at the end of these sessions.[3] The sub-Earth latitude of the observation also varied around the Martian equator,[3] although only slightly. Using the accurate ephemeris to obtain the sub-Earth local time and latitude for any particular date, we reconstructed the appearance of the planet's disk out of the model output. Then, "disk average" temperatures were calculated from the GCM results

employing the weighting function for the "visible" grid points based on the corresponding viewing geometry. Thus constructed temperatures allow for a direct comparison with the SWAS measurements. For simplicity, they are referred to as "global means" in this paper.

Figure 2 compares the simulated global "visible" surface temperature between $L_s = 165$ and 235° with the SWAS and MGS-TES measurements. Since the data from MGS-TES exist for local times 2:00 and 14:00 only, they cannot be directly compared with the SWAS. Instead we made an assumption that diurnal variations of the Martian surface temperature follow the time dependence reproduced by the model. Thus, the corresponding plot in this figure is actually composed of the MGS-TES data at the two local times interpolated according to the assumed diurnal time dependence, and followed by averaging over the visible disk in the manner similar to the one described for the model output. Figure 2 shows an overall consistency in the time series. The temperature drops with the onset of the dust storm in both the measurements and simulations, although the drop is more gradual in the observations than in the model. To emphasize the role of the dust in the radiative energy transfer, we included the surface temperature simulated with the dust radiation scheme turned off. As seen in the figure, the

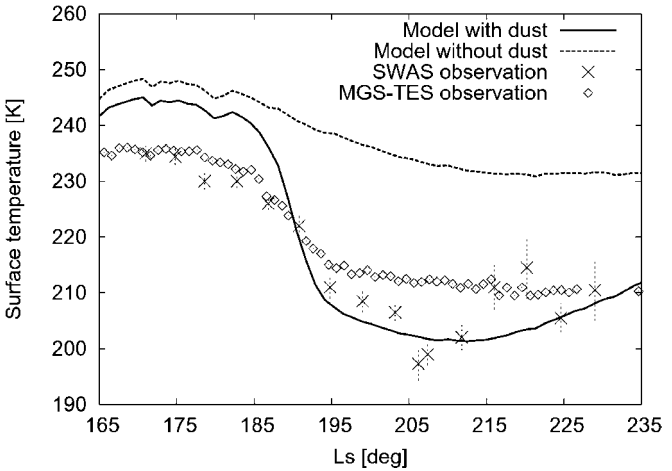


Fig. 2. Global mean surface temperature simulated with the Martian GCM and measured from SWAS and MGS-TES. The solid line represents the temperature for the run with dust radiation included (the dust distribution is as in Fig. 1); the dashed line is for the run without the dust. The crosses and dotted error bars show the observational results by SWAS.[3] The diamonds present the MGS-TES observation.

global-mean surface temperature is up to 30–40 K colder in the run when the dust radiative effects are included. Pure seasonal change in the surface temperature accounts for about 10 K during this time of the Martian year, as the run without the dust shows.

Figure 3 presents the vertical distributions of the simulated global-mean atmospheric temperature between $L_s = 165$ and 235°, and the comparison

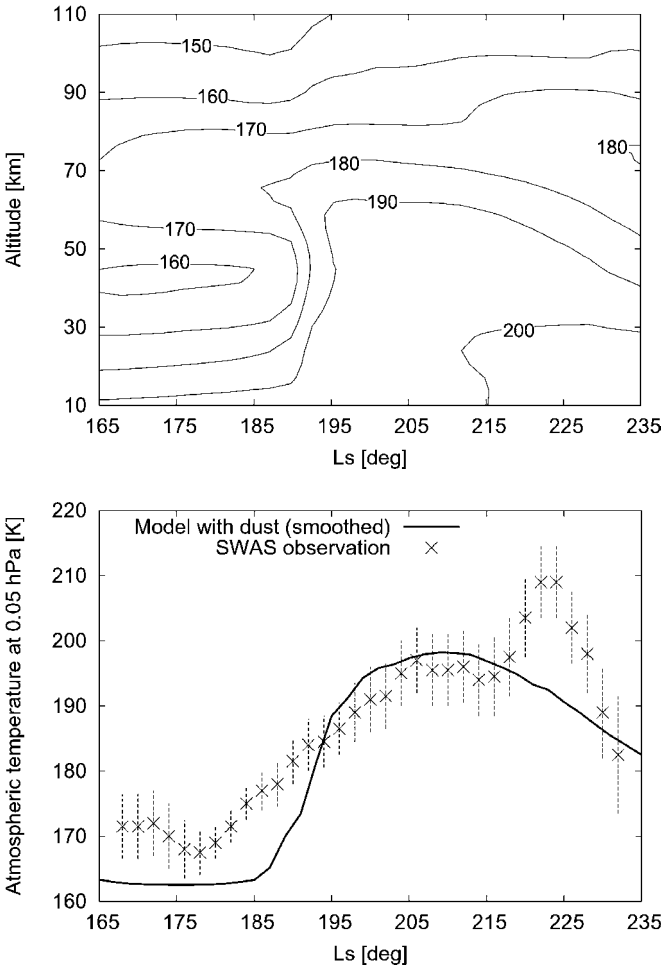


Fig. 3. (a) Vertical distribution of the global-mean atmospheric temperature simulated in the Martian GCM with the dust radiation effects included. (b) The time series of the global mean atmospheric temperature at 0.05 hPa from the model (the solid line), and from the SWAS observations (crosses and dotted error bars).[3]

with the SWAS measurements at 0.05 hPa (∼49 km). The temperature at 0.5 hPa (∼25 km) is roughly consistent with the SWAS measurements[3] before the onset of the dust storm, and rises in both the observations and simulations. However, in the model, the temperature increase owing to the global dust storm is 20–30 K lower than in the measurements. The model result in Fig. 3(b) was smoothed over 8.5° of L_s to match the SWAS data which were reportedly averaged over the same time interval. At 0.05 hPa the temperature time series is almost consistent with the observations during the global storm, although is ∼10 K colder before the global dust storm (see Fig. 3(b)). At 0.005 hPa (∼73 km), the model does not produce the temperature increase with the onset of the storm. This is consistent with the SWAS measurements, although the simulated temperature is generally ∼10 K higher.

4. Discussion

We compare the global mean atmospheric and surface temperatures simulated with the recently developed Martian GCM[5] with those observed by the Earth-orbiting submillimeter instrument (SWAS). The model includes the parametrization of the absorption, emission, and scattering of the solar and infrared radiation by the atmospheric dust.

The model reproduces a moderate atmospheric temperature inversion below 30 km around $L_s = 215°$, as was measured by SWAS (Gurwell, personal communication) and observed with MGS-TES during the dust storm of 2001. The global-mean daytime (local time 14:00) atmospheric temperature from MGS-TES is shown in Fig. 4. The temperature inversion is clearly seen there around $L_s = 215°$. The simulated time series of the surface

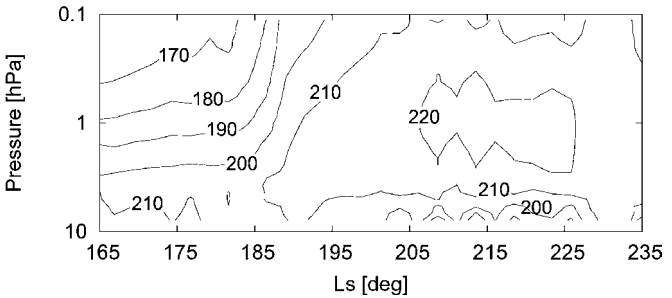


Fig. 4. Vertical distribution of the daytime (local time 14:00) global-mean atmospheric temperature measured by MGS-TES for the same period as in Fig. 3.

and 0.05 hPa atmospheric temperature are consistent with the observations. These observed features cannot be reproduced in the control run with the dust radiation scheme turned off. At 0.005 hPa, the temperature changes little during the global dust storm. This agrees well with observations. At 0.5 hPa, the simulated temperatures rises with the onset of the storm, but remain lower than in the observations.

There are several possible reasons which may explain these inconsistencies. First, the lack of measurements of the dust parameters results in the uncertainty in prescribing the dust in the model. There is a wide range of the observed aerosol particle distributions.[4] The size has a profound effect on the radiative transfer, and the corresponding heating/cooling rates. Besides, there are no measurements concerning the vertical dust distribution. Therefore, the dust parameters prescribed in Martian GCMs still represent a very crude approximation, and should be treated cautiously. Further observations are required.

Second, possible improvements in model results can be expected if the radiative transfer in the 15 μm CO_2 band due to the CO_2 and the dust is considered interactively, unlike in all Martian GCMs we are aware of. Third, at lower altitudes, the determination of height in observations becomes difficult, and the corresponding error increases. More care should be applied in treating the results from the layers close to the surface.

Acknowledgment

This work was supported by Deutsche Forschungsgemeinschaft (DFG), project HA 3261/1-2. The authors are grateful to Dr. Teruyuki Nakajima for providing the parts of the dust radiation code.

References

1. NASA, J. Bell, M. Wolff and the Hubble Heritage Team, Scientists track "perfect storm" on Mars, *HubbleSite*, STScI-2001-31, 2001.
2. M. D. Smith, Interannual variability in TES atmospheric observations of Mars during 1999–2003, *Icarus* **167** (2004) 148.
3. M. A. Gurwell, E. A. Bergin, G. J. Melnick and V. Tolls, Mars surface and atmospheric temperature during the 2001 global dust storm, *Icarus* **175** (2005) 23.
4. M. G. Tomasko, L. R. Doose, M. Lemmon, P. H. Smith and E. Wegryn, Properties of dust in the Martian atmosphere from the imager on Mars Pathfinder, *J. Geophys. Res.* **104** (1999) 8987.

5. *J. Geophys. Res.* **110**, E11 (2005) doi:10.1029/2005JE002498.
6. U. Berger and U. von Zahn, The two-level structure of the mesopause: A model study, *J. Geophys. Res.* **104** (1999) 22083.
7. A. A. Kutepov, O. A. Gusev and V. P. Ogibalov, Solution of the non-LTE problem for molecular gas in planetary atmospheres: Superiority of accelerated lambda iteration, *J. Quant. Spectrosc. Radiat. Transf.* **60** (1998) 199.
8. O. A. Gusev and A. A. Kutepov, Non-LTE gas in planetary atmospheres, in *Stellar Atmosphere Modeling*, eds. I. Hubeny, D. Mihalas and K. Werner, *ASP Conference Series* **288** (2003) 318.
9. F. Forget, F. Hourdin, R. Fournier, C. Hourdin, O. Talagrand, M. Collins, S. R. Lewis, P. L. Read and J.-P. Huot, Improved general circulation models of the Martian atmosphere from the surface to above 80 km, *J. Geophys. Res.* **104** (1999) 24155.
10. T. Nakajima and M. Tanaka, Matrix formulations for the transfer of solar radiation in a plane-parallel scattering atmosphere, *J. Quant. Spectrosc. Radiat. Transfer* **35** (1986) 13.
11. M. E. Ockert-Bell, J. F. Bell III, J. B. Pollack, C. P. McKey and F. Forget, Absorption and scattering properties of the Martian dust in the solar wavelengths, *J. Geophys. Res.* **102** (1997) 9039.
12. O. B. Toon, J. B. Pollack and C. Sagan, Physical properties of the particles composing the Martian Dust Storm of 1971–1972, *Icarus* **30** (1977) 663.
13. F. Forget, Improved optical properties of the Martian atmospheric dust for radiative transfer calculations in the infrared, *Geophys. Res. Lett.* **25** (1998) 1105.
14. T. Kuroda, N. Hashimoto, D. Sakai and M. Takahashi, Simulation of the Martian atmosphere using a CCSR/NIES AGCM, *J. Meteorol. Soc. Japan* **83** (2005) 1.
15. M. D. Smith, B. J. Conrath, J. C. Pearl and P. R. Christensen, Thermal Emission Spectrometer observations of Martian planet-encircling dust storm 2001A, *Icarus* **157** (2002) 259.
16. J. Liu, M. I. Richardson and R. J. Wilson, An assessment of the global, seasonal, and interannual spacecraft record of Martian climate in the thermal infrared, *J. Geophys. Res.* **108**, E8 (2003) 10.1029/2002JE001921.
17. B. J. Conrath, Thermal structure of the Martian atmosphere during the dissipation of the dust storm of 1971, *Icarus* **24** (1975) 36.
18. S. R. Lewis, M. Collons, P. L. Read, F. Forget, F. Hourdin, R. Fournier, C. Hourdin, O. Talagrand and J.-P. Huot, A climate database for Mars, *J. Geophys. Res.* **104** (1999) 24177.

THE BULK DENSITY OF COMETARY NUCLEI

BJÖRN J. R. DAVIDSSON
Department of Astronomy and Space Physics, Uppsala University
Box 515, SE-751 20 Uppsala, Sweden
bjorn.davidsson@astro.uu.se

This paper reviews the current knowledge about bulk densities and porosities of cometary nuclei. Grain agglomeration and formation of small planetesimals in the Solar Nebula, as well as the subsequent collisional and thermal evolution of cometary nuclei are discussed, in order to provide a theoretical expectation on the range of bulk densities for the present comet population. Various methods for estimating cometary bulk densities, and corresponding results are reviewed, showing a reasonable agreement with the previously mentioned theoretical range. Current empirical estimates indicate that comets are bodies with rather low density and high porosity (typically $\rho_{\rm bulk} = 600 \pm 400\,{\rm kg/m^3}$ and $\psi = 0.6 \pm 0.3$, respectively).

1. Introduction

Comets are the most primitive and least evolved bodies in the Solar System. As such, studies of their chemical composition and physical properties will yield invaluable information about the conditions prevailing in the Solar Nebula more than 4 billion years ago. For example, their bulk densities and porosities may place important constraints on planetesimal formation scenarios. However, it is not clear to what extent the present population of comets differ from the small planetesimals which built the Jovian planets and their moons — a substantial collisional and thermal evolution has possibly taken place, and it is necessary to investigate how this may have affected the structure of these bodies. Furthermore, bulk densities are difficult to measure and increased efforts are needed to improve current estimates.

It is first necessary to define a plausible upper limit on the nucleus density. Greenberg and Hage[1] and later Greenberg[2] tried to define the composition of a "typical" comet nucleus by considering the elemental abundances in the interstellar medium, the Solar System, the dust of Comet 1P/Halley, and the gas comae of several comets. They obtained mass fractions of silicates, organic refractories, volatiles

(dominated by water), and carbonaceous particles as 0.20:0.19:0.55: 0.06 and 0.26:0.23:0.42:0.09, respectively. By assigning the densities $\{\rho_{\rm sil}, \rho_{\rm or}, \rho_{\rm vol}, \rho_{\rm car}\} = \{3{,}500, 1{,}800, 1{,}200, 2{,}000\}\,{\rm kg/m^3}$ for these components, a total compact density of $\rho_{\rm comp} = 1{,}540$–$1{,}650\,{\rm kg/m^3}$ was obtained. However, the assumed densities may be on the high side (e.g., the pyroxene end member enstatite, which dominates the silicate composition of Comet C/1995 O1 Hale-Bopp,[3,4] only has $\rho_{\rm sil} = 2{,}700\,{\rm kg/m^3}$, and since water ice by far is the most common volatile, $\rho_{\rm vol} = 930\,{\rm kg/m^3}$ may be a more reasonable value). If the lower values for $\rho_{\rm sil}$ and $\rho_{\rm vol}$ are used, the density of compact material drops by $\sim 300\,{\rm kg/m^3}$ with respect to the value suggested by Greenberg.[2] Therefore, the range $1{,}400 \lesssim \rho_{\rm comp} \lesssim 1{,}700\,{\rm kg/m^3}$ is applied here. *The basic question in the current paper is, if the bulk densities of comets are in this range, or substantially lower due to porosity.*

The density expected for small planetesimals in the Solar Nebula is discussed in Sec. 2. The subsequent evolution of these planetesimals is discussed in Sec. 3, leading to some theoretical expectations on the bulk density of comets. Empirical estimates of cometary bulk densities are summarized in Sec. 4, and are compared with the theoretical range in Sec. 5.

2. Grain Agglomeration and Planetesimal Formation

The Solar Nebula consisted of ∼98% hydrogen and helium (by mass) while heavier species predominantly were trapped in solid dust grains with a plausible average radius[5] of $\sim 0.1\,\mu$m. These grains may have had a layered structure,[6] with a silicate core, a mantle of organic substances, and a crust of ices (mostly H_2O, CO, and CO_2) mixed with very small carbonaceous particles (e.g., polycyclic aromatic hydrocarbons).

The gas disk was supported by radial pressure gradients and therefore orbited the Sun at sub-Keplerian velocities. The grains, slowed down further by gas drag, had generally very small relative velocities (on the order of $V_{\rm rel} = 10^{-4}\,{\rm m/s}$), governed by Brownian motion. Under such conditions, laboratory experiments on ground and in microgravity[7,8] show that grains readily stick to each other upon collisions and build highly porous fractal clusters, having a fractal dimension $D_{\rm f} = 1.3$–1.9 (see Fig. 1). This means that the cluster mass $m \propto r^{D_{\rm f}}$ grows much slower with cluster size r, than for solid particles (for which $D_{\rm f} = 3$). As the clusters grow, their cross section to mass ratio σ/m slowly decreases, leading to reduced gas drag and higher relative cluster velocities.

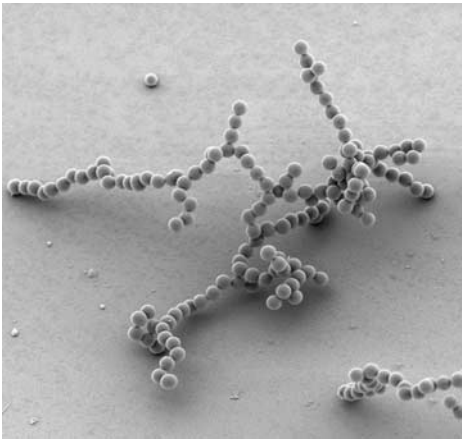


Fig. 1. Dust agglomerate of $r = 0.95\,\mu$m SiO_2 spheres formed by differential sedimentation. Image reproduced from Ref. 10 with permission from the author.

Increasing impact velocities leads to a certain degree of compaction, and at some point the clusters no longer grow as fractals (although the porosity remains high). Wurm and Blum[7] showed experimentally that this restructuring limit is reached at $V_{\rm rel} \approx 0.2\,$m/s. At even higher velocities, the sticking efficiency decreases and collisions no longer lead to growth. Blum and Wurm[9] showed that this fragmentation limit occurs abruptly at $V_{\rm rel} \approx 2\,$m/s for spherical SiO_2 grains. At that point,[10] the clusters have only grown to $\sim$0.1 m at $r_{\rm h} = 1\,$AU and $10^{-4}\,$m at $r_{\rm h} = 100\,$AU (where $r_{\rm h}$ is the heliocentric distance).

The fragmentation limit may therefore be a serious problem, since it is difficult to explain how larger bodies could have formed, especially since the largest relative velocities expected[10,11] at $r_{\rm h} \gtrsim 30\,$AU during planetesimal formation is $V_{\rm rel} \approx 50\,$m/s. However, Poppe *et al.*[12] showed that grains with *irregular* surfaces have a strikingly different behavior compared to spherical grains — even at $V_{\rm rel} \approx 50\,$m/s, 30–80% of the grains stick upon impact. Furthermore, Bridges *et al.*[13] have demonstrated experimentally that the presence of frost on the grain surfaces may facilitate sticking, especially if the water ice is mixed with methanol.

What is the expected porosity for bodies built by particle-by-particle agglomeration at relative velocities $V_{\rm rel} \lesssim 50\,$m/s? Blum and Schräpler[10,14] have investigated experimentally how the porosity of so-called "dust cakes" (i.e., large grain aggregates with an initial porosity of $\psi \approx 0.85$) is affected

Table 1. "Dust cake" porosity ψ measured by Blum and Schräpler[10,14] as function of static pressure, or equivalently, impact velocity $V_{\rm rel}$.

Pressure (Pa)	$V_{\rm rel}$ (m/s)	ψ
10^2	1	0.85
10^4	10	0.73
10^5	50	0.64
10^6	100	0.64

by static compression, simulating an impact. Table 1 summarizes some of their results. It is seen that mass growth due to agglomeration by individual grains (or very small clusters) only leads to a small compaction even at relatively high impact velocities. For macroscopic bodies ($r \sim 1\,$m) built by this mechanism, it is therefore expected that ∼65% or more of the body volume is empty space.

As the bodies grow, the dominant source of mass growth changes from agglomeration of small clusters to collisions between macroscopic bodies. The main contributor to the relative velocity between bodies is now a size-dependent radial drift relative to the gas, caused by the head-wind acting on the small planetesimals (the bodies spiral toward the Sun). Table 2 shows results obtained by Weidenschilling,[11] who performed numerical large-scale planetesimal growth simulations. After a peak in $V_{\rm rel}$ of ∼50 m/s for bodies of ∼1 m size, the relative velocity decreases again due to the lesser effect of gas drag on larger bodies. For example, a ∼500 m boulder, formed after $\sim 1.5 \times 10^5$ years of agglomeration, therefore mostly accretes bodies of size ∼200 m, and the collision velocity is only $V_{\rm rel} \approx 0.2\,$m/s.

The consequences of collisions between macroscopic porous targets have been investigated, e.g., by Donn,[15] Sirono and Greenberg.[16] Typically,

Table 2. Results from numerical simulations of planetesimal growth at $r_{\rm h} = 30\,$AU according to Weidenschilling.[11] For various time steps t after onset of calculations, the size of the largest bodies formed are shown (Targets), as well as the size of the bodies they predominantly accrete (Impactors), and the relative impact velocity $V_{\rm rel}$.

t (year)	Target size (m)	Impactor size (m)	$V_{\rm rel}$(m/s)
6×10^4	0.1	$(3-5) \times 10^{-5}$	10
7×10^4	10	0.2–0.3	2
8×10^4	70	20	0.8
1.5×10^5	500	200	0.2
2×10^5	6,000	1,000	2

two equal-sized boulders of radius r and porosity $\psi = 0.6$, colliding at $V_{\rm rel} \leq 10\,{\rm m/s}$ penetrate each other to a distance of $\sim 0.03r$, i.e., the bodies barely attach. The porosity in the compaction zone may be reduced substantially, but due to its small size, the overall effect is rather small. Furthermore, one might expect that macroscopic void spaces are formed between blocks which may compensate (or even overcompensate) the compression during collisions. For this reason, it is expected that planetesimals maintain a porosity of $\psi \geq 0.65$ until they reach the typical size of a comet, $r = 1$–$10\,{\rm km}$. Since a collision even at $V_{\rm rel} = 50\,{\rm m/s}$ only leads to a local heating of $\sim 20\,{\rm K}$, it is expected that planetesimal growth proceeds without substantial loss of volatiles.[15] Applying $1{,}400 \lesssim \rho_{\rm comp} \lesssim 1{,}700\,{\rm kg/m^3}$, and a porosity $\psi \geq 0.65$ makes it plausible that *newly formed planetesimals of cometary size in the outer Solar System had bulk densities* $\rho_{\rm bulk} \lesssim 600\,{\rm kg/m^3}$.

3. Collisional and Thermal Evolution of Cometary Nuclei

Dynamical studies[17–21] have shown that the Jupiter Family Comets (JFCs), i.e., comets with a Tisserand parameter $2 \leq T_{\rm J} \leq 3$ with respect to Jupiter, and with orbital periods $P \leq 20\,{\rm years}$, most likely have their origin in the Edgeworth-Kuiper Belt (EKB). Edgeworth-Kuiper Objects (EKOs) larger than $\sim 100\,{\rm km}$ have a size distribution $N \propto r^{-\alpha}$, where the cumulative power law index is $\alpha = 3.4 \pm 0.3$ according to Gladman *et al.*,[22] and $\alpha = 3.2^{+0.4}_{-0.3}$ according to Trujillo *et al.*[23] However, for smaller objects ($40 \lesssim r \lesssim 70\,{\rm km}$), the power law index seems to decrease abruptly[24] to $\alpha \approx 1.9 \pm 0.6$ (although this estimate is based on a very small sample of objects).

For a collisionally relaxed population, $\alpha = 1.9 - 2.5$ is expected,[25–27] which seems to indicate that the smaller EKOs are severely affected by collisions. It has been suggested[25,28–30] that 50–90% of the $r \lesssim 10\,{\rm km}$ EKOs (and thereby, the JFCs) in fact are collisional fragments.

Additional support for a collisional history may be found in the size distribution for JFCs themselves, although the estimates of α for this population vary substantially. Fernández *et al.*[31] found $\alpha = 2.65 \pm 0.25$, Lowry and Weissman[32] estimated $\alpha = 1.59 \pm 0.03$, while Meech *et al.*[33] report $\alpha = 1.45 \pm 0.05$. However, Jewitt *et al.*[34] pointed out that JFCs generally are more elongated (average photometric range $\langle \Delta m_{\rm R} \rangle = 0.54 \pm 0.07$) than small main belt asteroids ($\langle \Delta {\rm m_R} \rangle = 0.32 \pm 0.05$), which may indicate that JFCs are not a population of collision fragments after all

(although the discrepancy could be explained by sublimation-induced shape modification).

The presence of active and inactive areas on comets (the former typically covering less than 20% of the surface of JFCs, according to Fernández *et al.*[35]), could in principle be yet another sign of cometary collisional evolution. One can imagine that water ice was a globally distributed surface species for newly formed comets, given the previously mentioned small degree of heating during the formation processes. However, impact velocities in the present-day EKB are substantially higher[29] ($\langle V_{\rm rel} \rangle \approx 500\,{\rm m/s}$ at 42 AU) than in the Solar Nebula, due to the absence of gas. The energy delivered by an impact could then deprive the surface material of ice in the immediate vicinity of the crater, thereby forming a "dead spot" on the surface (i.e., little or no gas would subsequently be produced there, even when solar illumination is strong). Durda and Stern[30] estimated that a 1 km body in the EKB would experience 90–300 impacts by bodies larger than 4 m during 3.5 Gyr, and that 20–224% of its surface would be cratered (i.e., the same place might be hit more than once) — numbers that are comparable to the cometary surface fraction of inactive areas. The reason that substantial fractions of JFC surfaces are inactive could therefore be that surface volatiles have been removed long ago in the EKB due to collision-induced melting. Consequently, "active areas" on cometary surfaces may therefore represent old pristine material which has not experienced impacts (i.e., surface ice is still present). It is noted that the Oort cloud comets may also have been collisionally evolved[36] before leaving the planetary region.

Due to the potential importance of cometary collisional evolution, it is not obvious that the density $\rho_{\rm bulk} \lesssim 600\,{\rm kg/m^3}$, characterizing the earliest planetesimals, also is representative for the current population of comets. Whether the collisions predominantly have led to compaction (increase of $\rho_{\rm bulk}$) or crack formation (decrease in $\rho_{\rm bulk}$), is hard to say. However, important lessons can be learned by considering the rapidly growing body of information available for asteroids (see, e.g., the review by Britt *et al.*[37]). With the exception of a few of the largest asteroids, which appear to be compact and coherent objects, most investigated asteroids seem to have porosities in the range $0.3 \leq \psi \leq 0.7$, with a tendency for clustering at the lowest value. The dominant mechanism for producing a porosity in excess of $\sim$30% appears to be impact-induced shattering, followed by dispersion of the fragments and, later, gravitational reaccumulation

(i.e., rubble-pile formation). Crack formation, on the other hand, is only considered to yield very modest increases in porosity. If this also is true for comets, substantial increases of their porosity during collisions should not be expected.

The perhaps most likely outcome of an impact on a porous target like a comet, is therefore compaction. However, the degree of porosity decrease may not be severe, as illustrated by the oddly shaped asteroid 253 Mathilde, which presently has a porosity of about $\psi \approx 0.4$. It has endured at least five giant collisions which, according to Housen *et al.*,[38] only has increased its bulk density by ∼20%. If similar numbers apply for the collisional evolution that JFCs may have experienced in the EKB, the expected cometary bulk density limit would just be adjusted to $\rho_{\rm bulk} \lesssim 700\,{\rm kg/m^3}$. But even if collisions had a more severe effect on the porosity, it is not likely to decrease below $\psi \approx 0.36$, the random close packing limit (as long as the constituent $\sim 0.1\,\mu$m grains are not forced to merge, or are ground down to even smaller pieces). The reason is that it becomes increasingly difficult to crush particles with decreasing size l (for silicate material, compressive strength roughly scales as $l^{-1/2}$), and to compact silicate soils significantly beyond the close packing limit requires much higher pressures than can be delivered in collisions at a few kilometers per second.[37] This may in fact be the reason why most non-monolithic asteroids do not have a porosity below $\psi \approx 0.3$. If the close packing limit is applied to comet material, this would suggest a bulk density limit of $\rho_{\rm bulk} \lesssim 1{,}000\,{\rm kg/m^3}$.

Another mechanism that could modify the density, is radiogenic core melting. Prialnik *et al.*[39] showed that if the planetesimal abundance of ^{26}Al was as high as in the Allende meteorite, icy bodies with $r \gtrsim 6\,{\rm km}$ would melt. Liquid water could then fill microcavities before freezing, potentially creating a core region in larger bodies characterized by $\rho_{\rm comp}$. If such a body was disrupted in a collision in the EKB, the fragments (of which some may be seen as JFCs today) could potentially be relatively solid bodies. However, the authors pointed out that presence of amorphous water ice (which may be responsible for the frequently observed cometary outbursts), rules out such radiogenic core melting.

To conclude, it is expected that truly pristine comets have bulk densities $\rho_{\rm bulk} \lesssim 600\,{\rm kg/m^3}$. In case collisions have compacted the bodies, it is still not likely that the bulk density exceeds $\sim 1{,}000\,{\rm kg/m^3}$, since collisions hardly could have compressed the bodies beyond the close packing limit. However, if radiogenic core melting indeed took place, at least a fraction of

the comets could have a bulk density approaching 1,700 kg/m^3, if they are collision fragments from the core of a once molten parent body.

4. Methods for Estimating Cometary Bulk Densities

4.1. *Non-gravitational force modeling*

As a comet sublimates, a so-called non-gravitational force is exerted on the nucleus due to conservation of linear momentum. Since this force modifies the cometary orbit with respect to the purely gravitational orbit, measured non-gravitational changes of the orbit may be used in combination with thermophysical model calculations in order to estimate the mass and bulk density of the comet. In a pioneering study, Rickman[40–42] estimated the bulk density of Comet 1P/Halley, and found $500 \lesssim \rho_{\rm bulk} \lesssim 1{,}200\,{\rm kg/m^3}$, with a preference for the lower value. Sagdeev *et al.*,[43] estimated $\rho_{\rm bulk} = 600^{+900}_{-600}\,{\rm kg/m^3}$, while Peale[44] obtained a preferred value $\rho_{\rm bulk} = 700\,{\rm kg/m^3}$ (although he could not exclude any value in the range $30 \leq \rho_{\rm bulk} \leq 4{,}900\,{\rm kg/m^3}$). In spite of the large error bars associated with each individual estimate, the combined results seem to indicate that the bulk density of 1P/Halley is below $\sim$1,000 kg/m^3.

Rickman *et al.*[45] estimated the bulk densities of 17 "old" short-period comets (objects having made at least 30 revolutions since capture), and although the results were uncertain for individual objects, the sample as a whole indicated a typical density of $\rho_{\rm bulk} \lesssim 500\,{\rm kg/m^3}$.

Comet 19P/Borrelly was studied by Farnham and Cochran,[46] who found a density $290 \lesssim \rho_{\rm bulk} \lesssim 830\,{\rm kg/m^3}$, with a preferred value of $\rho_{\rm bulk} \approx 490\,{\rm kg/m^3}$. Davidsson and Gutiérrez[47] studied the same object, obtaining $180 \lesssim \rho_{\rm bulk} \lesssim 300\,{\rm kg/m^3}$. The reason for the discrepancy may be that Farnham and Cochran assumed a momentum transfer coefficient that is substantially larger than obtained through direct simulation Monte Carlo modeling of the gas in the near-nucleus coma, as performed by Davidsson and Gutiérrez.

Davidsson and Gutiérrez[48] also studied Comet 67P/Churyumov-Gerasimenko, obtaining $100 \lesssim \rho_{\rm bulk} \lesssim 370\,{\rm kg/m^3}$ for the best possible reproduction of observational data (including the rotational lightcurve amplitude, water production rate versus time, and non-gravitational changes in the orbital period, longitude of perihelion and longitude of the ascending node). An upper limit $\rho_{\rm bulk} \lesssim 600\,{\rm kg/m^3}$ was suggested if uncertainties in the empirical data were considered. The study of Comet

81P/Wild 2 by the same authors[49] yielded $\rho_{\rm bulk} \lesssim 600$–$800\,{\rm kg/m^3}$ (depending on the applied nucleus volume).

4.2. *Tidal disruption*

A handful of comets have been observed to split due to tidal forces in the vicinity of Jupiter or the Sun.[50] The most well studied of these objects is Comet D/1993 F1 Shoemaker-Levy 9, which disrupted in 1992 and later impacted Jupiter in 1994. By modeling the morphology of the post-disruption nucleus chain, Asphaug and Benz[51] found $\rho_{\rm bulk} = 600 \pm 100\,{\rm kg/m^3}$ if the parent body did not rotate, and $\rho_{\rm bulk} \lesssim 1{,}000\,{\rm kg/m^3}$ if a prograde 9 h period was assumed. A similar study by Solem[52] yielded $\rho_{\rm bulk} = 500$–$600\,{\rm kg/m^3}$. Furthermore, a study of the impact features on Jupiter made by Crawford[53] suggested that the impactors had densities around $\rho_{\rm bulk} \approx 250\,{\rm kg/m^3}$.

4.3. *Rotational stability*

In order for a comet nucleus to remain intact, self-gravity and material forces must provide a sufficient centripetal force to avoid breakup due to nucleus rotation. Formulae derived by Davidsson[54,55] show how the critical breakup rotational period depends on body properties such as size, shape, bulk density, tensile, and shear strength. When the size, shape, and spin period of a comet are estimated empirically, these formulae may also be used to place a *lower limit* on the bulk density, although an assumption must be made regarding the material strength. None of the 14 objects studied by Davidsson[55] therefore *required* a bulk density higher than $\sim 600\,{\rm kg/m^3}$ to stay intact, when assuming zero strength (see Table 3), although it cannot be excluded that the true densities are higher. The same conclusion was reached by Lowry and Weissman,[32] when studying a sample of 13 objects.

Although this technique has limited importance for individual comets or small samples of objects, a larger sample may reveal a true physical cut-off in the parameter space, i.e., non-existence of rotational periods below a certain value for comets with a particular size, which in turn places a constraint on the bulk density and material strength of the objects. The investigation by Toth and Lisse,[56] in which 29 objects were studied, is an important step in this direction — the cut-off appear consistent with a bulk density of $\sim 300\,{\rm kg/m^3}$ and tensile strength of $\sim 250\,{\rm Pa}$. A tendency of finding shorter rotational periods for smaller objects may indicate that this in fact represents the *physical* cut-off limit.

Table 3. Lower limits on $\rho_{\rm bulk}$ obtained by Davidsson,[55] using empirical data on the nucleus size, shape, and rotational period. Note that the material strength was assumed negligible and that only 7 out of 14 objects required $\rho_{\rm bulk}$ in excess of $200\,{\rm kg/m^3}$ (due to long rotational periods in comparison with the critical breakup periods).

Comet	Lower limit on $\rho_{\rm bulk}$ (kg/m^3)
31P/Schwassmann-Wachmann 2	530
6P/d'Arrest	440
95P/Chiron	370
46P/Wirtanen	350
107P/Wilson-Harrington	340
10P/Tempel 2	250
C/1991 L3 Levy	200

It is noted that the unusual fast-spinning Comet 133P/Elst-Pizarro, with a rotational period $P = 3.471\,{\rm h}$ and nucleus semi-axes $a = 3.0\,{\rm km}$, $b = c = 2.1\,{\rm km}$ (according to observations by Hsieh *et al.*[57]), requires a bulk density of at least $\rho_{\rm bulk} = 1{,}300\,{\rm kg/m^3}$ to remain intact, in case it is strengthless. However, if the comet material possesses some tensile strength (which does not even have to be as large as the strength predicted by Greenberg *et al.*,[58] for porous core-mantle grain material), the bulk density could easily be as low as $100\,{\rm kg/m^3}$, without risking rotational breakup (all according to a formula valid for prolate ellipsoids, derived by Davidsson[55]). In particular, the largest tensile strength required in the interval $100 \leq \rho_{\rm bulk} \leq 1{,}300\,{\rm kg/m^3}$ amounts to $\sim$200 Pa (needed if the bulk density is $\rho_{\rm bulk} \approx 600\,{\rm kg/m^3}$).

4.4. *Radar observations*

Harmon *et al.*[59,60] have detected radar echoes from a number of comet nuclei. The ratio between the radar and geometric cross sections (i.e., the radar albedo) can be used to estimate the dielectric constant in the surface material. This in turn yields estimates of the bulk density in the uppermost layer of the nucleus, for various assumptions regarding the composition. Table 4 shows such estimates for a number of comets, assuming either pure water snow or a silicate powder to cover the nucleus surface. Note that these values may not be representative for the entire nucleus (e.g., presence of macroscopic void spaces in the deep comet interior may reduce $\rho_{\rm bulk}$ further).

Table 4. Near-surface bulk densities (kg/m^3) derived from radar cross sections observed by Harmon and colleagues,[59,60] assuming two different compositions — water snow and silicate powder. The average bulk density may represent ice/dust mixtures.

Comet	Density (snow)	Density (Si powder)	Density (average)
C/1983 H1 IRAS-Araki-Alcock	500	900	700
2P/Encke	500	900	700
C/1983 J1 Sugano-Saigusa-Fujikawa	700	1,300	1,000
26P/Grigg-Skjellerup	700	1,300	1,000
C/1996 B2 Hyakutake	200	400	300

4.5. *Nucleus fragmentation and disruption*

About 30 comet nuclei have been observed to split or partially fragment (see the excellent review by Boehnhardt[61]), often for no apparent reason (i.e., non-tidally). Desvoivres *et al.*[62] have developed a model of the dynamics of fragments, based on the nucleus gravity and sublimation-induced non-gravitational forces acting on them, which can be used to estimate the mass of such fragments. The authors applied their model to Comet C/1996 B2 Hyakutake, which expelled a number of fragments in March 1996. The estimated bulk density of the fragments depend on their assumed sizes. Taking the sizes as 20, 10 and 7 m, yields corresponding bulk densities of 300, 600, and 800 kg/m^3, respectively, showing that their densities hardly could have been as high as $\rho_{\rm comp}$ unless they were extremely small.

4.6. *Binary Edgeworth-Kuiper objects*

Given the probable relation between small EKOs and JFCs, it is of interest to study the bulk densities of the former objects, at least for bodies small enough ($r \lesssim 100$ km) to be unaffected by gravitational compression.[10] A number of binary EKOs have already been discovered,[63] which allows system mass estimates to be made if the orbital periods and separation distances are determined. In case optical and thermal infrared observations are performed simultaneously, estimates of the object albedos and sizes are obtained, which would yield the sought-for bulk density of the pair. Although the primaries discovered so far typically have sizes above the gravitational compression limit, this technique may provide important information on bulk densities of small EKOs in the near future.

4.7. *Direct measurements by spacecraft*

To date, no less than six comets have been the targets of spacecraft missions — 21P/Giacobini-Zinner (*ICE*), 1P/Halley (*Giotto, Vega 1 and 2, Suisei, Sakigake, ICE*), 26P/Grigg-Skjellerup (*Giotto*), 19P/Borrelly (*Deep Space 1*), 81P/Wild 2 (*Stardust*), and 9P/Tempel 1 (*Deep Impact*). Of these mission, only *Deep Impact* has resulted in mass and bulk density estimates, obtained by studying the dynamics of ejecta created when a 370 kg impactor collided with 9P/Tempel 1, at a velocity of 10.3 km/s. According to A'Hearn *et al.*,[64] the bulk density is $\rho_{\rm bulk} = 620^{+470}_{-330}\,{\rm kg/m^3}$. The other missions have not resulted in mass or bulk density estimates since flyby distances and comet-spacecraft relative speeds have been too large to yield measurable gravitational perturbations on the spacecraft trajectories.

The ESA mission *Rosetta*, which includes the lander probe *Philae*, will hopefully obtain accurate mass and bulk density measurements of Comet 67P/Churyumov-Gerasimenko, upon arrival in 2014.

5. Conclusions

In Sec. 2, it was argued that freshly formed comet-sized planetesimals were likely to have porosities $\psi \gtrsim 0.65$, which translates to $\rho_{\rm bulk} \lesssim 600\,{\rm kg/m^3}$ for material with $\rho_{\rm comp} \leq 1{,}700\,{\rm kg/m^3}$. The subsequent evolution of these bodies into the comets we see today may have included collisional alteration, thermal processing, and possibly other mechanisms not considered here. Collisions could hardly increase the bulk density above $\sim$1,000 kg/m^3, unless the constituent grains were deformed and merged in the process, which is rather unlikely. Radiogenic core melting and gravitational compaction of larger parent bodies (i.e., objects larger[10] than $\sim$100 km) may partially have removed microcavities and increased the bulk density to $1{,}400 \lesssim \rho_{\rm bulk} \lesssim 1{,}700\,{\rm kg/m^3}$, at least for a fraction of the comet population. However, if amorphous water ice indeed is an important constituent in comets, as suggested by frequent outbursts, such dense objects may be very rare, if existing at all.

The methods for bulk density estimates of comets available at present, are all indirect and not very accurate. Estimates based on non-gravitational force modeling may be biased toward fairly porous objects, since comets with very small and uncertain non-gravitational orbit changes (possibly caused by a comparably large mass and bulk density) are not selected for study. Other methods only yield a lower limit (rotational stability), cannot be applied for a sufficiently large sample (tidal breakup), or only target the

surface regions of the bodies (radar observations). But taken together, existing estimates consistently seem to indicate that comet bulk densities generally are lower than $\sim$1,000 kg/m^3, with typical values around $\sim$600 kg/m^3. This implies that $\sim$60% of the body volumes typically is empty space, and that comets indeed seem to be rather "fluffy" bodies. For further information and discussion, the review by Weissman *et al.*[65] is highly recommended.

Acknowledgments

I wish to thank Drs. J. Watanabe, H. Kawakita, and L. M. Lara, organizers of the *Physical Properties of Comets and Related Bodies* session, for inviting me to present this paper during the Asia Oceania Geosciences Society's 2nd Annual Meeting, held in Singapore, 20–24 June 2005. Furthermore, I am indebted to the referees of this paper, Drs. P. R. Weissman and S. J. Weidenschilling, for having provided me with thoughtful, constructive, and highly appreciated comments and suggestions.

References

1. J. M. Greenberg and J. I. Hage, *Astrophys. J.* **361** (1990) 260.
2. J. M. Greenberg, *Astron. Astrophys.* **330** (1998) 375.
3. D. H. Wooden, D. E. Harker, C. E. Woodward, H. M. Butner, C. Koike, F. C. Witteborn and C. W. McMurtry, *Astrophys. J.* **517** (1999) 1034.
4. T. L. Hayward, M. S. Hanner and Z. Sekanina, *Astrophys. J.* **538** (2000) 428.
5. A. Li and J. M. Greenberg, *Astron. Astrophys.* **323** (1997) 566.
6. J. M. Greenberg, in *Comets*, ed. L. L. Wilkening (The University of Arizona Press, Tucson, 1982), p. 131.
7. G. Wurm and J. Blum, *Icarus* **132** (1998) 125.
8. J. Blum, G. Wurm, S. Kempf, T. Poppe, H. Klahr, T. Kozasa, M. Rott, T. Henning, J. Dorschner, R. Schräpler, H. U. Keller, W. J. Markiewicz, I. Mann, B. A. S. Gustafson, F. Giovane, D. Neuhaus, H. Fechtig, E. Grün, B. Feuerbacher, H. Kochan, L. Ratke, A. El Goresy, G. Morfill, S. J. Weidenschilling, G. Schwehm, K. Metzler, W.-H. Ip, *Phys. Rev. Lett.* **8** (2000) 2426.
9. J. Blum and G. Wurm, *Icarus* **142** (2000) 138.
10. J. Blum, in *Astrophysics of Dust*, ASP Conference Series, Vol. 309, eds. A. Witt, G. Clayton and B. Draine (The Astronomical Society of the Pacific, San Francisco, 2004), p. 369.
11. S. J. Weidenschilling, *Icarus* **127** (1997) 290.
12. T. Poppe, J. Blum and T. Henning, *Astrophys. J.* **533** (2000) 454.
13. F. G. Bridges, K. D. Supulver, D. N. C. Lin, R. Knight and M. Zafra, *Icarus* **123** (1996) 422.

14. J. Blum and R. Schräpler, *Phys. Rev. Lett.* **93** (2004) 115503-1.
15. B. D. Donn, *Astron. Astrophys.* **235** (1990) 441.
16. S. Sirono and J. M. Greenberg, *Icarus* **145** (2000) 230.
17. J. A. Fernández, *Mon. Not. R. Astron. Soc.* **192** (1980) 481.
18. M. Duncan, T. Quinn and S. Tremain, *Astrophys. J. Lett.* **328** (1988) L69.
19. T. R. Quinn, S. Tremain and M. J. Duncan, *Astrophys. J.* **355** (1990) 667.
20. H. F. Levison and M. J. Duncan, *Icarus* **127** (1997) 13.
21. V. V. Emel'yanenko, D. J. Asher and M. E. Bailey, *Mon. Not. R. Astron. Soc.* **350** (2004) 161.
22. B. Gladman, J. J. Kavelaars, J.-M. Petit, A. Morbidelli, M. J. Holman and T. Loredo, *Astron. J.* **122** (2001) 1051.
23. C. A. Trujillo, J. X. Luu, A. S. Bosh and J. L. Elliot, *Astron. J.* **122** (2001) 2740.
24. G. M. Bernstein, D. E. Trilling, R. L. Allen, M. E. Brown, M. Holman and R. Malhotra, *Astron. J.* **128** (2004) 1364.
25. P. Farinella and D. R. Davis, *Science* **273** (1996) 938.
26. D. P. O'Brien and R. Greenberg, *Icarus* **164** (2003) 334.
27. M. Pan and R. Sari, *Icarus* **173** (2005) 342.
28. S. A. Stern, *Astron. J.* **110** (1995) 856.
29. D. R. Davis and P. Farinella, *Icarus* **125** (1997) 50.
30. D. D. Durda and S. A. Stern, *Icarus* **145** (2000) 220.
31. J. A. Fernández, G. Tancredi, H. Rickman and J. Licandro, *Astron. Astrophys.* **352** (1999) 327.
32. S. C. Lowry and P. R. Weissman, *Icarus* **164** (2003) 492.
33. K. J. Meech, O. R. Hainaut and B. G. Marsden, *Icarus* **170** (2004) 462.
34. D. Jewitt, S. Sheppard and Y. Fernández, *Astron. J.* **125** (2003) 3366.
35. J. A. Fernández, G. Tancredi, H. Rickman and J. Licandro, *Revista Mexicana de Astronomia y Astrofisica de Serie de Conferencias* **11** (2001) 15.
36. S. A. Stern and P. R. Weissman, *Nature* **409** (2001) 589.
37. D. T. Britt, D. Yeomans, K. Housen and G. Consolmagno, in *Asteroids III*, eds. W. F. Bottke Jr., A. Cellino, P. Paolicchi and R. P Binzel (The University of Arizona Press, Tucson, The Lunar and Planetary Institute, Houston, 2002), p. 485.
38. K. R. Housen, K. A. Holsapple and M. E. Voss, *Nature* **402** (1999) 155.
39. D. Prialnik, A. Bar-Nun and M. Podolak, *Astrophys. J.* **319** (1987) 993.
40. H. Rickman, in *The Comet Nucleus Sample Return Mission*, ESA SP-249, ed. O. Melita (ESA Publications Division, Noordwijk, 1986), p. 195.
41. H. Rickman, in *Cometary Environments*, *Advances in Space Research*, Vol. 9(3), eds. T. I. Gombosi, S. K. Atreya, E. Grün and M. S. Hanner (Maxwell Pergamon Macmillan plc, Oxford, 1989), p. 59.
42. Y. V. Skorov and H. Rickman, *Planet. Space Sci.* **47** (1999) 935.
43. R. Z. Sagdeev, P. E. Elyasberg and V. I. Moroz, *Nature* **331** (1988) 240.
44. S. J. Peale, *Icarus* **82** (1989) 36.
45. H. Rickman, L. Kamél, M. C. Festou and C. Froeschlé, in *Symposium on the Diversity and Similarity of Comets*, eds. E. J. Rolfe and B. Battrick (ESA Publications Division, Noordwijk, 1987), p. 471.

46. T. L. Farnham and A. L. Cochran, *Icarus* **160** (2002) 398.
47. B. J. R. Davidsson and P. J. Gutiérrez, *Icarus* **168** (2004) 392.
48. B. J. R. Davidsson and P. J. Gutiérrez, *Icarus* **176** (2005) 453.
49. B. J. R. Davidsson and P. J. Gutiérrez, *Icarus* **180** (2006) 224.
50. Z. Sekanina, *Astron. Astrophys.* **318** (1997) L5.
51. E. Asphaug and W. Benz, *Icarus* **121** (1996) 225.
52. J. C. Solem, *Astron. Astrophys.* **302** (1995) 596.
53. D. Crawford, *Lunar Planet. Sci.* **XXVIII** (1997) 267.
54. B. J. R. Davidsson, *Icarus* **142** (1999) 525.
55. B. J. R. Davidsson, *Icarus* **149** (2001) 375.
56. I. Toth and C. M. Lisse, *Icarus*, in press (2005).
57. H. H. Hsieh, D. C. Jewitt and Y. R. Fernández, *Astron. J.* **127** (2004) 2997.
58. J. M. Greenberg, H. Mizutani and T. Yamamoto, *Astron. Astrophys.* **295** (1995) L35.
59. J. K. Harmon, D. B. Campbell, A. A. Hine, I. I. Shapiro and B. G. Marsden, *Astrophys. J.* **338** (1989) 1071.
60. J. K. Harmon, D. B. Campbell, S. J. Ostro and M. C. Nolan, *Planet. Space Sci.* **47** (1999) 1409.
61. H. Boehnhardt, *Earth, Moon and Planets* **89** (2002) 91.
62. E. Desvoivres, J. Klinger, A. C. Levasseur-Regourd, J. Lecacheux, L. Jorda, A. Enzian, F. Colas, E. Frappa and P. Laques, *Mon. Not. R. Astron. Soc.* **303** (1999) 826.
63. W. M. Grundy, K. S. Noll and D. C. Stephens, *Icarus* **176** (2005) 184.
64. M. F. A'Hearn, M. J. S. Belton, W. A. Delamere, J. Kissel, K. P. Klaasen, L. A. McFadden, K. J. Meech, H. J. Melosh, P. H. Schultz, J. M. Sunshine, P. C. Thomas, J. Veverka, D. K. Yeomans, M. W. Baca, I. Busko, C. J. Crockett, S. M. Collins, M. Desnoyer, C. A. Eberhardy, C. M. Ernst, T. L. Farnham, L. Feaga, O. Groussin, D. Hampton, S. I. Ipatov, J.-Y. Li, D. Lindler, C. M. Lisse, N. Mastrodemos, W. M. Owen Jr., J. E. Richardson, D. D. Wellnitz and R. L. White, *Science* **310** (2005) 258.
65. P. R. Weissman, E. Asphaug and S. C. Lowry, in *Comets II*, eds. M. C. Festou, U. Keller and H. A. Weaver (The University of Arizona Press, Tucson, 2004), p. 337.

This page intentionally left blank

LYMAN-α OBSERVATIONS OF SUNGRAZING COMETS WITH THE SOHO/UVCS INSTRUMENT

A. BEMPORAD*, G. POLETTO†, J. RAYMOND‡ and S. GIORDANO§
*Astronomy Department Florence University
L.go E. Fermi 2, 50125 Florence, Italy
†INAF-Arcetri Astroph. Obs., L.go E. Fermi 5, 50125 Firenze, Italy
‡Harvard-Smithsonian CFA, 60 Garden St., Cambridge, MA 02138, USA
§INAF-Torino Astronom. Obs., Strada Osservatorio 20
10025 P. Torinese, Italy
*bempy76@arcetri.astro.it

The Large angle and spectrometric coronagraphs aboard the Solar and heliospheric observatory (SOHO) spacecraft observed a large (more than 1000) number of sungrazers. This led to many studies which tried to explain their origin and the peculiar shape of the observed cometary lightcurves. However, in the last years a few sungrazers have been observed also in the hydrogen Lyman-α spectral line by the UV coronagraph spectrometer on SOHO. This instrument allowed to perform UV spectroscopical observations of sungrazing comets on their final stage of life at projected heliocentric distances between 1.4 and 10 solar radii. Ultra violet coronagraph spectrometer (UVCS) detected in the sungrazer UV spectra mainly the Lyman-α spectral line. Typically, emission in this line originates in the hydrogen cloud produced by the water photodissociation, but, at these low heights, it is necessary to take into account also strong interaction processes (e.g., mass-loading and charge exchange) between the solar wind and the outgassed materials. From these observations, it has been possible to estimate cometary parameters such as the outgassing rates and the nucleus sizes, as well as parameters of the coronal plasma encountered by the comet. In this work we review the main results derived from the UVCS observations of sungrazing comets: the detection of a "hidden" mass below $\sim 6\,R_\odot$, the indirect and direct evidences for the occurrence of fragmentation processes and a tentative estimate for the pyroxene dust grain number density. Moreover, we present here preliminary results on the UVCS data interpretation of a sungrazer observed in 2002.

1. Introduction: State of Art

Sungrazer comets have a perihelion distance of a few solar radii and in some cases collide with the Sun. The first reliable discovery of a sungrazer occurred in 1680 (the Great Comet); then, after the Great Comet of 1843, four individual sungrazing comets have been observed in the 1880s, while the next one did not appear until 1945. Two further sungrazers were discovered in the 1960s, Comet Pereyra in 1963 and Comet Ikeya-Seki in 1965.

Hence, before 1979 only about 10 sungrazers had been identified. Then, in the decade 1979–1989, 16 sungrazers were discovered: Six by the the SOLWIND instrument on the P78-1 satellite and 10 by the coronagraph aboard the *Solar Maximum Mission* (SMM) spacecraft. After the beginning of the *SOlar and Heliospheric Observatory* (SOHO) mission the number of discovered sungrazers increased very quickly: since 1996, the two *Large Angle and Spectrometric COronagraph* (LASCO) cameras (see below) aboard the SOHO spacecraft detected more than 1000 sungrazing comets,[a] revealing the sungrazers to be much more numerous than expected from previous observations. Over 700 of these belong to the Kreutz sungrazer group (i.e. the group of sungrazing comets which, following the theory first formulated by Daniel Kirkwood in 1880, are fragments of a single sungrazer progenitor). This sudden increase in the number of discovered sungrazers is due to many advantages introduced by the LASCO coronagraphs, first of all its higher photometric accuracy ($\sim\pm 0.1$ mag) and the very large field of view (FOV) which allows the detection of sungrazers up to a distance of about 0.15 AU ($32\,R_\odot$).

The LASCO[1] instrument consists of three optical systems (C1, C2, and C3), each equipped with its own 1024×1024 pixel CCD camera. The C2 is a white-light coronagraph which observes the solar corona from 2 to $6\,R_\odot$ with a projected pixel size of 11.4 arcsecs $\sim$8270 km, while the C3 FOV ranges from 3.7 to $32\,R_\odot$ with a projected pixel size of 56.0 arcsecs $\sim$40,600 km. Typical LASCO data consist of a sequence of images taken, respectively, at a rate of about three and two images *per* hour, viewed as "movies".

The *Ultra Violet Coronagraph Spectrometer* (UVCS)[2] instrument aboard SOHO allowed for the first time the spectroscopic observations in the EUV range of sungrazing comets in the final stage of their life. This instrument consists of two UV channels and a white light polarimeter; the instantaneous FOV in the UV channels, given by the $42'$ long spectrometer entrance slit (normal to the solar radius), may be moved along the radial direction in order to observe the solar corona from 1.4 to 10 solar radii. The slit may be rotated around an axis pointing toward the center of the Sun in order to observe the whole corona from polar to equatorial latitudes.

Thanks to the large extent of the C3 FOV it is possible to compute the orbital parameters of a newly discovered sungrazer and set the position of the UVCS slit to intercept its trajectory. This allows EUV spectroscopic observations of the same comet at different (typically 4–5) heliocentric

[a] See http://ares.nrl.navy.mil/sungrazer/.

distances. From these observations we have an unique opportunity to obtain information on the constituents of the cometary dust. However, despite the very large amount of sungrazing comets discovered by LASCO, as of today UVCS observed only 13 comets; 10 of these belong to the Kreutz group. So far, detailed study of sungrazers C/1996 Y1,[3] C/2000 C6,[4] and C/2001 C2[5] have appeared: in the following sections we describe the main results from UVCS observations of these comets.

2. The UVCS Sungrazer Emission: Main Properties

Once the cometary orbital parameters have been computed from LASCO observations, the UVCS slit is set along the trajectory to observe the comet transit at 4–5 different heliocentric distances. At each position, the instrument starts to observe at a fixed altitude ∼15–20 min before the comet enters the UVCS FOV and stops about 30–40 min later, depending on the cometary orbital parameters. When the comet, traveling along its orbit, intersects the slit, cometary emission is seen typically over a section ∼100–200″ long. The spectral intervals for sungrazer observations typically include, beside neutral Hydrogen Lyα λ1215.7 Å Lyβ λ1025.7 Å and the OVI $\lambda\lambda$1031.9–1037.6 Å lines, spectral lines from species that might be present in cometary spectra such as HeI, NI, AlIII, SiI, SiII, PII, OI, CII, and ArI. However, in the selected spectral ranges, UVCS usually detected a cometary emission only in the H Lyα line[b]; hence, in the following we discuss the origin and main characteristics of this line in sungrazer spectra. The H Lyα line is observed also in coronal spectra: hence, in order to identify the cometary signal, we need to calculate its average coronal intensity in spectra taken before the comet entered the UVCS FOV and to subtract this background signal from the following exposures. This procedure removes also the possible Lyα interplanetary emission and the detector dark counts. From the cometary Lyα intensity observed along the spectrograph slit in different exposures it is possible to reconstruct the cometary Lyα image and then the shape of the Hydrogen cloud. Figure 1 shows the reconstructed Lyα images for the sungrazers C/2000 C6 (left panel) and C/2001 C2 (right panel). In both cases it is clear that the Hydrogen clouds appear to be strongly "stretched" along the comet–Sun direction leading to the formation of a "Lyα tail": as we discuss later, this is due mainly to

[b] OI, CII, and CIII have been detected in UVCS spectra of non-sungrazing comets,[6] partly because these comets are much brighter.

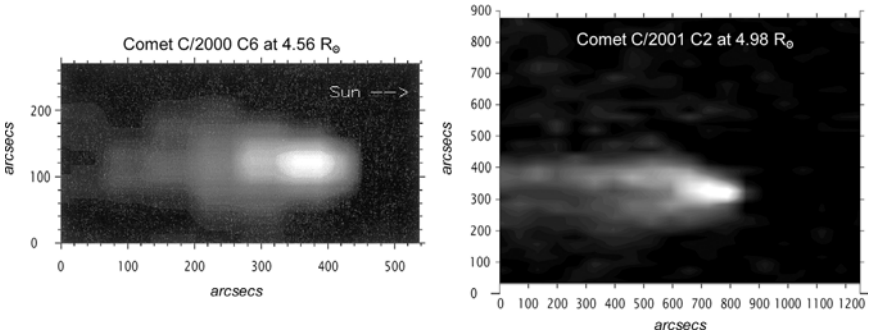


Fig. 1. Left: composite image of the C/2000 C6 Lyα tail at 4.56 $R_\odot$ after the subtraction of the coronal emission (adapted from Ref. 4). Right: the same for the C/2001 C2 sungrazing comet (adapted from Ref. 5). In both images the pixel size along the y-axis correponds to the spatial resolution along the UVCS slit, while the pixel size along the x-axis corresponds to the comet's radial velocity multiplied by the exposure time.

the interaction of the cometary ejecta with the solar wind. We notice here that the correction for the comet motion along the slit revealed the Lyα tails shown in Fig. 1 to be both oriented with the cometary orbital path. The angle (projected on the plane of the sky) between the radial from the Sun and the orbital path is on the order of 14 and 13°, respectively, for the C/2000 C6 (Fig. 1, left) and the C/2001 C2 (Fig. 1, right) Lyα images. We discuss this properties of the sungrazers Lyα tails in Sec. 3. As revealed by a comparison between the two panels of Fig. 1, while the C/2000 C6 sungrazer shows a single Lyα tail, the C/2001 C2 image shows two tails.[5] We discuss this issue in Sec. 4: here we anticipate only that the presence of two tails has been interpreted as a signature of a subfragment traveling with the main nucleus.

In Sec. 3 we briefly discuss the origin for the enhanced (with respect to the coronal plasma density) number of H atoms responsible for the observed cometary emission.

3. The Origin of the Observed Lyα Sungrazer Emission

In typical coronal conditions of high-electron temperature ($T_e \sim 10^6$ K) and low-electron density ($n_e \sim 10^6$–10^7 cm^{-3}) the observed spectral line emission is due to the collisional excitation with thermal electrons and the resonant scattering of the chromospheric radiation (only for spectral lines which connect to the lowest levels). Hence, the first question we have to solve is whether the observed cometary Lyα emission arises from collisional and/or

radiative excitation. In order to distinguish between these two processes, we may resort to the ratio Lyβ/Lyα between the observed intensities of these spectral lines. At typical coronal temperatures, the ratio between the collisional components of these line is expected to be ∼0.13–0.14,[3] while the ratio between the radiative components has a much lower value on the order of ∼0.001–0.002.[3] Because in the UVCS observations of sungrazers C/1996 Y1, C/2000 C6, and C/2001 C2 the Lyβ line was very faint or absent, it can be concluded that the observed cometary Lyα emission arises almost entirely from radiative excitation (as is usually the case in the corona[7]).

Another question concerning the observed Lyα tails is the origin of the neutral Hydrogen atoms responsible for this emission. In principle these neutrals could be a "first generation" of H atoms formed from the photodissociation of the H_2O molecules outgassed from the nucleus, and/or a "second generation" of H atoms formed *via* the charge exchange process between these atoms and the coronal protons p^+ of the solar wind. First, we have to take into account that the velocity imparted to the H atoms from outgassing[8] and photodissociation processes[9] is less than about 40 km/s, much smaller than the typical sungrazer speed close to the Sun[3–5] (∼250–300 km/s). This implies that the H atoms from the photodissociation of H_2O molecules move, in first approximation, with the comet, hence (opposite to what Fig. 1 shows) the Lyα emission from these atoms would drop as soon as the comet travels beyond the spectrograph slit. Second, because of the high radial component of the cometary speed, the atomic absorption profile of these H atoms is Doppler shifted with respect to the chromospheric emission profile, and the radiatively excited Lyα emission is strongly reduced (by a factor 0.2–0.1 for radial speeds between 250–300 km/s[10]) by the Swings effect. The emission from outgassed H atoms would result in a narrow line profile, while the FWHM of the cometary and coronal profiles turns out to be equal (Fig. 2). Moreover, because of the comet velocity component along the line of sight, we should expect a Doppler shift of the Lyα line profile from outgassed H atoms with respect to the coronal profile, while no significant shifts are observed[c] (see Fig. 2). We conclude that the H atoms responsible for the Lyα sungrazer emission belong to the "second generation" of neutrals, i.e. to the H atoms from the photodissociation of H_2O which, before being collisionally or radiatively ionized, undergo a charge

[c]Partial filling of the UVCS aperture by the sungrazer at the time the comet entered the UVCS FOV results in a small shift in the Lyα line profile observed only in that exposure.

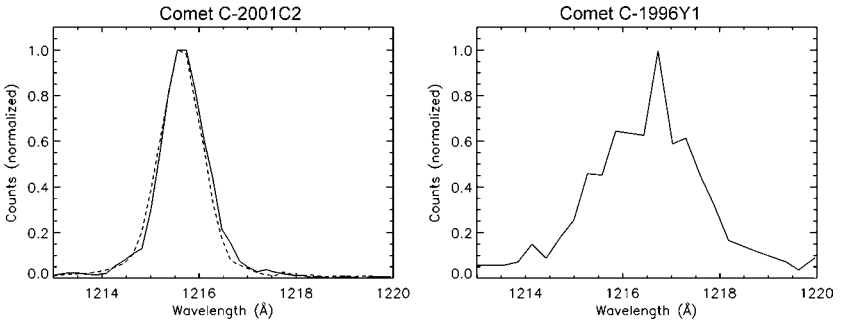


Fig. 2. Normalized cometary Lyα profiles for the C/2001 C2 (left panel, solid line) and the C/1996 Y1 (right panel, adapted from Ref. 3) sungrazers, respectively, at 3.6 and 5.3 $R_\odot$. In the left panel the cometary profile is compared to the average coronal profile (dashed line).

transfer with coronal protons. This is in contrast to the normal comets, in which the first generation neutrals typically dominate due to the smaller velocity of the comet relative to the Sun.[6] Because of the small momentum transfer in the charge exchange process,[11] the secondary H atoms have the same velocity distribution of the coronal atoms. Following the latter conclusion, we should expect that the orientation of the Lyα tails match the anti-Sun direction, because of the nearly radial direction of the outflowing solar wind at 4.56 and 4.98 $R_\odot$. On the contrary, as mentioned in Sec. 2, the observed Lyα tails are oriented along the cometary orbital path. This can be explained as follows: as we better discuss in the next section, from the variation with time of the Lyα cometary intensity it has been possible to evaluate the characteristic times for the charge exchange and ionization processes. It turns out that cometary H atoms formed by charge exchange have a short lifetime (on the order of 500 and 1,000 s, respectively, for the C/2000 C6 and C/2001 C2 sungrazers observed at 4.56 and 4.98 $R_\odot$). This may explain the orbital orientation of these tails: because of the short H atoms lifetime, before being ionized they covered path too short to align the observed tail with the local outflow velocity.

In Sec. 4 we discuss sungrazers parameters derived from the observed Lyα emission.

4. Estimate of Sungrazers Outgassing Rate and Nucleus Size

The number of neutrals moving with the comet $N_{\rm coma}$ is proportional to the unknown outgassing rate $\dot{N}$ (where $\dot{N}$ is the number of neutrals per

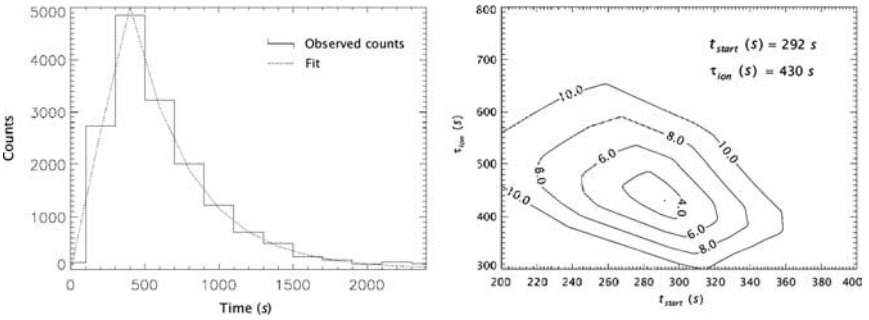


Fig. 3. Left: the C/2001 C2 Lyα counts (summed over the slit length) versus time curve as observed at 3.6 $R_\odot$ (solid line) and the curve fit (dotted line). Right: the iso-χ^2 curves for the determination of the ($\tau_{\rm ion}$, $t_{\rm st}$) values at which the χ^2 value is minimum (see text).

second produced by outgassing) and to the unknown ionization and charge exchange rates ($\tau_{\rm ion}^{-1}$ and $\tau_{\rm cx}^{-1}$, respectively). These are proportional[11,12] to the local electron density $n_{\rm e}$ encountered by the comet. The $N_{\rm coma}$ atoms undergo a charge transfer with the ambient protons creating a number $N_{\rm tail}$ of neutrals (responsible for the observed Lyα emission) at a rate $\tau_{\rm cx}^{-1}$: this process locally increases the number density of H atoms with respect to the coronal density. As a consequence, the number of H atoms $N_{\rm tail}$ left by the comet along its path exponentially decays as $\exp(-t/\tau_{\rm ion})$ until the cometary Lyα signal disappears (Fig. 3, left panel). By fitting with an exponential curve (depending on $\dot{N}$, $\tau_{\rm ion}$, the time $t_{\rm st}$ at which the comet first entered the slit and other known parameters) the observed Lyα intensity decay after the comet transit, it is possible to derive an estimate for the cometary outgassing rate $\dot{N}$ and the local electron density $n_{\rm e}$ (Fig. 3, right panel). Assuming a balance between the energy supplied by the solar radiation impinging on a spherical nucleus and the energy required to sublimate the quantity of ice derived from the $\dot{N}$ value, we may estimate (for isotropic outgassing) the nucleus equivalent radius.

Table 1 gives the results obtained at different heliocentric distances h ($R_\odot$) of the comet for the sungrazers C/1996 Y1, C/2000 C6, and C/2001 C2, in particular the ice mass outgassed per second $Q_{\rm H_2O}$ (kg/s) and the estimated nucleus radius r (m). An interesting result from Table 1 is the UVCS detection of a significant surviving mass even after the comets disappeared from the LASCO/C2 FOV (hence below 6 $R_\odot$), implying the presence at very low heliocentric distances of a hidden mass "inconsistent with the nominal models based on the coronagraphic observations alone".[13] This

Table 1. Derived sungrazer parameters.

Comet name	Q_{H_2O} (kg/s)	r (m)	h ($R_\odot$)
C/1996 Y1	20.0	3.4	6.80
C/2000 C6	71.8	3.0	3.26
	140	5.8	4.56
	34.6	3.4	5.71
	10.5	2.5	6.36
C/2001 C2	820[a]	20.3	3.60
	58.9[a]	7.8	4.98
	28.5[b]	5.4	4.98

[a]Main nucleus.
[b]Subfragment.

apparent inconsistency may be removed by introducing in the sungrazer erosion model the presence of one or more companions traveling with the main nucleus: if these subfragments have a lower erosion rate, they can survive longer than the main nucleus becoming visible very close to the Sun.[13] Another interesting result shown in Table 1 is the observation of sudden increases in the Q_{H_2O} values as the comets approach the Sun: these may be ascribed to a fragmentation of the comet nucleus which increases the surface exposed to the solar flux (even if there are other alternative explanations for this phenomenon[4]).

Table 1 has different entries for the two fragments of the C/2001 C2 sungrazer. As we anticipated in Sec. 2, this comet shows two Lyα tails at $4.98\,R_\odot$ and a single tail at $3.60\,R_\odot$: a possible interpretation[5] is that at $4.98\,R_\odot$ we are observing the outgassing from two fragments. By applying the technique described above to the two Lyα intensity versus time profiles, we derived an estimate for the radii of both fragments. At $3.60\,R_\odot$ the subfragment probably disappears because it completely sublimates between these two heights (as we verified starting from the estimated radius of 5.4 m and integrating over the cometary orbit the rate of change in radius[14] caused by the erosion process[5]). Again, the sudden increase in the estimated radius at $3.60\,R_\odot$ indicates that at this heliocentric distance the main nucleus is going through a complete disruption due to either or both thermal or gravitational stresses. This explains also why we did not observe the comet later in UVCS data at $2.20\,R_\odot$: the comet never reached this heliocentric distance.

Fragmentation processes are believed to play a key role in the origin and evolution of sungrazing comets: first, the occasional secondary brightenings observed in some sungrazer lightcurves (i.e., the variation of the sungrazer

magnitude with heliocentric distance) have been explained by introducing one or more nearby companions traveling with the main nucleus.[14] Second, from a comparison between the orbital parameters of different comets, it has been demonstrated[15] that many sungrazers arriving in pairs or triplets originated via fragmentation events from a single sungrazer far from the Sun. More recently, an intriguing scenario has been published involving splitting both close and far to the Sun[16]: this could explain the generation, via runway fragmentation, of the observed sungrazers from a single progenitor. The UVCS instrument has a spatial resolution higher than the LASCO resolution, hence is able to resolve objects which are unresolved in the LASCO data. The UVCS observations of sungrazers can strongly improve our knowledge on this subject, as confirmed by the presence of subfragments inferred for the C/2000 C6 and directly observed for the first time for the C/2001 C2 sungrazers.

Another important result from the observations of the C/2001 C2 sungrazer arises from the detection, in the Lyα intensity versus time profile of its main tail at $4.98\,R_\odot$, of a very slow exponential decay. By fitting the observed curve, the derived local electron density n_e was lower than expected. This led the authors[5] to hypotesize the presence of a background Lyα emission superposed onto an exponential decay: after the subtraction of this background, the fit to the exponential Lyα decay gave a more reliable n_e value. This additional emission has been ascribed to neutral H atoms formed by the interaction between the products from the cometary dust photodissociation and the coronal protons. The idea arises from the association that has been made between the heliocentric distance of 4–5 $R_\odot$ at which the cometary Lyα emission peaks and the sublimation rate of pyroxene dust grains which peaks at about the same height.[17] In principle, it is possible that pyroxene dust grains act as agent to neutralize coronal protons creating an additional number of neutrals and the C/2001 C2 sungrazer has been observed at $4.98\,R_\odot$, hence around this height. In this scenario, we expect a mixing between the Lyα emission formed by the charge transfer between cometary neutrals and coronal protons and the Lyα emission from the charge transfer between coronal protons and pyroxene grains. Given the cross sections σ for the charge trasfer processes between protons and O or Si atoms from the photodissociation of SiO_2 molecules and assuming a typical bulk density and radius for the pyroxene grains,[17] the authors[5] derived an expression for the expected number of H neutrals produced by the charge transfer as a function of the unknown dust number density. Equating this formula to the value of the additional H atoms estimated from

the observed constant background of the Lyα counts, they finally find that the pyroxene dust grain number density needed to reproduce the observations is $\simeq 6.2 \times 10^{-10}\,\mathrm{cm}^{-3}$. This has been the firt order-of-magnitude determination of the pyroxene dust grain number density in sungrazing comets.

Before concluding this section, we briefly discuss the formation of the cometary bow shock for sungrazers. Close to the Sun the magnetic field, the solar wind velocity v_w and the sungrazer velocity v_com are nearly radial; a shock should form with a shock velocity $v_\mathrm{sh} = v_\mathrm{w} + v_\mathrm{com}$ and a fraction of the energy dissipated in the shock should go into plasma heating (detectable as an increase in the FWHM of the Lyα line profile). However, sungrazers C/2000 C6 and C/2001 C2 were immersed in a slow wind region ($v_\mathrm{w} \sim 150$–$200\,\mathrm{km/s}$) and no significant plasma heating was observed: as we already mentioned, the Lyα cometary and coronal profiles had the same FWHM. On the contrary, sungrazer C/1996 Y1 crossed a fast wind region: its Lyα profile was much broader (kinetic temperature $T_\mathrm{k} \sim 9 \times 10^6\,\mathrm{K}$) than the background coronal profile and Doppler shifted because of the comet motion along the line of sight (see Fig. 2). From the observed proton kinetic temperature, the authors[3] inferred a wind speed of $v_\mathrm{w} \simeq 620\,\mathrm{km/s}$. Obviously the formation of the bow shock for sungrazers is strongly dependent on the physical parameters of the coronal region crossed by the comet.

5. Ultraviolet Coronagraph Spectrometer Observation of C/2002 S2: First Results

On September 18–19, 2002 UVCS observed the sungrazer C/2002 S2[18] at four different heliocentric distances. In this section we briefly describe the first results obtained from an analysis of these data. Figure 4 (top left panel) shows the reconstructed Lyα appearance of this comet at the heliocentric distances of 4.64 and $6.84\,\mathrm{R}_\odot$. These panels clearly show the presence of 2 Lyα tails at $6.84\,\mathrm{R}_\odot$ and possibly a similar multiplicity also at lower heliocentric distance. The $6.84\,\mathrm{R}_\odot$ image seems to be similar to the $4.98\,\mathrm{R}_\odot$ image of the C/2001 C2 sungrazer (see Fig. 1), but there is a strong difference between the two. As revealed by a Gaussian fit of the C/2002 S2 Lyα line profiles, while the emission from the main tail is red-shifted by more than $\sim 60\,\mathrm{km/s}$, the Lyα profiles of the secondary tail are blue-shifted by more than $\sim 120\,\mathrm{km/s}$. The resulting Doppler shift image of the comet is puzzling and the two observed Lyα tails cannot be easily interpreted as the signature of two fragments.

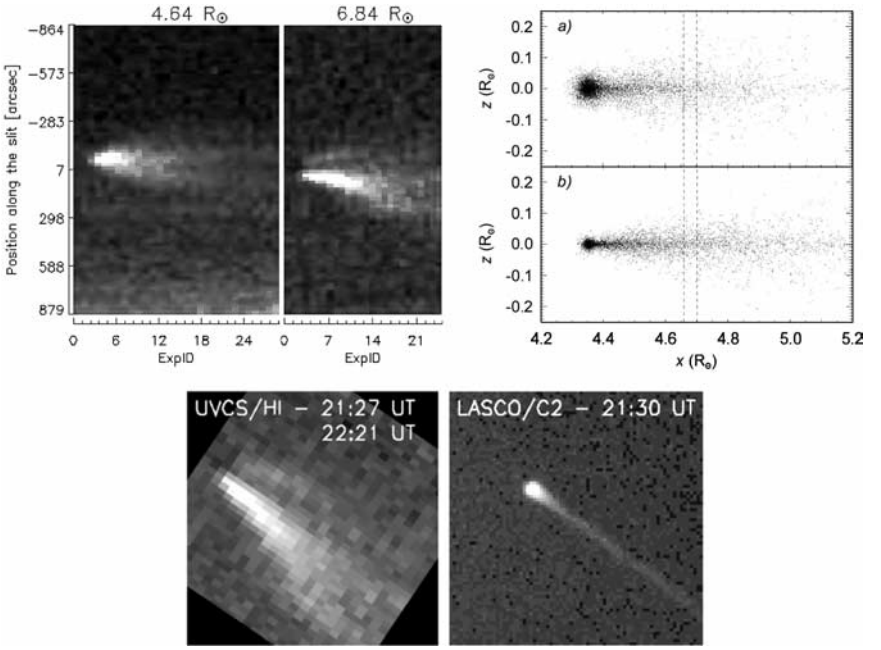


Fig. 4. Top left: composite images of the C/2002 S2 Lyα tails at 4.64 and 6.84 $R_\odot$ (coronal background emission not subtracted). Top right: the simulated distribution of the H atoms with two different assumption on the wind speed for the comet crossing at 4.64 $R_\odot$ (dashed line represents the UVCS slit FOV). Bottom: a comparison between the C/2002 S2 Lyα image (left) and the white light image (right) from LASCO data.

A first alternative explanation could be that the observed tails both originate from the interaction of cometary dusts with the coronal protons, as described in Sec. 4. However, we notice that, opposite to what happens for the C/2001 C2 observation, during the C/2002 S2 observations the orbital plane was seen edge-on. In general, the ion and dust tails or tails of dusts ejected at different times have a different inclination, but lie on the orbital plane. Hence, in this interpretation, we should expect to see two tails superposed along the line of sight, and this is not the case. Moreover, a comparison between the Lyα image reconstructed from UVCS data and the white light image at the same time from LASCO (Fig. 4, bottom panel), shows a single tail lying in between the two Lyα tails. Because, as it is well known, white light emission is due to the scattering from the dust grains, this may lead us to consider a separate origin for the C/2002 S2 Lyα tails and the dust tail.

A second alternative explanation for the two tails could be that the comet is crossing an inhomogeneous coronal region. The secondary H atoms travel with the solar wind, hence the sungrazers Lyα image strongly depends on the local magnitude and direction of the wind velocity vector. If the comet encounters, for instance, a transition region between slow and fast wind regime, we could expect to observe an "inhomogeneous Lyα image" of the comet: in this sense, UVCS observations of sungrazers could be used as "tracers" of the solar wind inhomogeneities. In order to solve these problems it is necessary to better understand the relationships between the shape of the observed Lyα tails and the physical parameters of the coronal plasma crossed by the comet. To this end, Giordano *et al.*[18] are working on a comet simulation code based on the Monte Carlo technique. The aim of this code is to reproduce the observed sungrazer Lyα image as a function of both coronal (e.g. kinetic temperature, electron density, electron temperature, wind speed, etc. ...) and cometary (e.g. outgassing rate, velocity distribution and kinetic temperature of the outgassed, etc. ...) free parameters. Figure 4 (right panels) shows the simulated dependence of the H atoms distribution from the wind speed $v_{\rm w}$ for $v_{\rm w} = 0\,{\rm km/s}$ (panel a) and $v_{\rm w} = 200\,{\rm km/s}$ (panel b). These results are preliminary: a full paper is in preparation.

6. Conclusions

Sungrazing comets have been detected by UVCS in the Lyα spectral line. Data analysis provided many new results about these comets, first of all the presence of a hidden mass, unobserved by LASCO coronagraphs and ascribed to the presence of slowly eroding subfragments traveling with the main nucleus. The fragmentation processes play a fundamental role in the formation and evolution of sungrazers; from UVCS data the presence of subfragments has been inferred from the observed sudden increases in the Lyα intensity. A subfragment has been directly observed for the first time (thanks to the high UVCS spatial resolution), as derived from the observation of a secondary Lyα tail. Recent UVCS observation of a sungrazer revealed again a secondary Lyα tail: however, because of the puzzling Lyα Doppler shift image, this cannot explained as a signature of two fragments. In order to find an alternative explanation, the creation of a Lyα tail simulation code is in progress: this code will help us to better understand the dependence of the observed Lyα image on the physical parameters of the coronal plasma encountered by the comet.

References

1. G. Brueckner, R. Howard, M. Koomen *et al.*, *Sol. Phys.* **162** (1995) 357.
2. J. L. Kohl, R. Esser, L. D. Gardner *et al.*, *Sol. Phys.* **162** (1995) 313.
3. J. C. Raymond, S. Fineschi, P. L. Smith *et al.*, *ApJ* **508** (1998) 410.
4. M. Uzzo, J. C. Raymond, D. A. Biesecker *et al.*, *ApJ* **558** (2001) 403.
5. A. Bemporad, G. Poletto, J. C. Raymond *et al.*, *ApJ* **620** (2005) 523.
6. M. S. Povich, J. C. Raymond, G. H. Jones *et al.*, *Science* **302** (2003) 1949.
7. J. C. Raymond, J. L. Kohl, G. Noci *et al.*, *Sol. Phys.* **175** (1997) 654.
8. A. H. Delsemme, in *Comets*, ed. L. L. Wilkening (Tucson, 1982), p. 85.
9. W. F. Huebner, J. J. Keady and S. P. Lyon, *Ap&SS* **195** (1992) 291.
10. J. L. Kohl, G. Noci, E. Antonucci *et al.*, *Sol. Phys.* **175** (1997) 613.
11. G. W. McClure, *Phys. Rev.* **148** (1966) 47.
12. T. T. Scholz and H. R. J. Walters, *ApJ* **380** (1991) 302.
13. Z. Sekanina, *ApJ* **597** (2003) 1237.
14. M. Iseli, M. Küppers, W. Benz and P. Bochsler, *Icar.* **155** (2002) 350.
15. Z. Sekanina, *ApJ* **576** (2002) 1085.
16. Z. Sekanina and P. W. Chodas, *ApJ* **607** (2004) 620.
17. H. Kimura, I. Mann, D. A. Biesecker *et al.*, *Icar.* **159** (2002) 529.
18. S. Giordano, J. C. Raymond *et al.*, *ApJ* (in preparation).

This page intentionally left blank

AN INVESTIGATION OF THE LIGHT CURVE OF DEEP IMPACT TARGET COMET

VITALY FILONENKO* and KLIM CHURYUMOV
Astronomical Institute of V.N.Karazin Kharkov National University
Sumskaja str., 35, Kharkov-22, 61022, Ukraine
**filonenko@astron.kharkov.ua*

The light curves of integrated visual brightness of the periodic comet 9P/Tempel during 1972, 1983, 1994, and 2005 appearances were constructed and studied. The values of photometrical parameters H_0, n, and H_{10} were computed. A secular fading of this comet was studied. The photometrical peculiarities of comet and long-term evolution of comet's activity probably will change after impact. Therefore, the results, which are presented in this manuscript, are important.

1. Introduction

The mission Deep Impact was successfully finished. Probably the photometrical behavior and a photometrical evolution of target comet will change as a consequence of the artificial impact to the comet nucleus. Therefore, the study of photometrical peculiarities and brightness behavior of target comet in all comet's observed returns is of interest. The results of these investigations are presented here.

2. Observations and Magnitude Analysis

To study the photometrical behavior of comet 9P/Tempel 1, we used more than 1,000 visual estimations of its integrated brightness collected from archives of the International Comet Quarterly, of the Comet Section of British Astronomical Association and from some Internet sources.

All estimations have been corrected to the standard telescope's aperture.[1]

The brightness behavior of comets with respect to heliocentric distance (light curves) is normally represented by equation

$$m_\Delta = H_0 + 2.5n \log r, \tag{1}$$

*Corresponding author.

where $m_\Delta = m_1 - 5 \log \Delta$ is heliocentric magnitude of comet, m_1 is comet's observed integrated magnitude, H_0 is absolute magnitude of comet, n is photometrical exponent, Δ and r are the geocentric and the heliocentric distances of comet, respectively.

An absolute magnitude H_{10}, which had been introduced by S. Vsekhsvyatsky, is important for the study of the secular evolution of cometary activity. This parameter is represented by the equation

$$H_{10} = m_1 - 5 \log \Delta - 10 \log r. \tag{2}$$

3. Visual Brightness Behavior

The light curve of visual integrated brightness of comets characterizes the change of their activity with time and with heliocentric distance. The comparison of light curves of short-period comets in their different returns gives the information about the evolution of cometary activity and on the relation between cometary brightness behavior, solar activity, and condition in the interplanetary space.

3.1. *Light curves*

The brightness behavior of comet 9P/Tempel 1 before perihelion was similar during all four returns (see Fig. 1). As it can be seen, the peaks on brightness (associated to comet outbursts) occur at equal about similar heliocentric distances in all returns. Especially it can be clearly seen that the outbursts occurred on distances 1.8, 1.76, 1.62, and 1.51 AU from Sun [$\log(r) = 0.255$, 0.246, 0.210, and 0.179 accordingly]. Authors have found

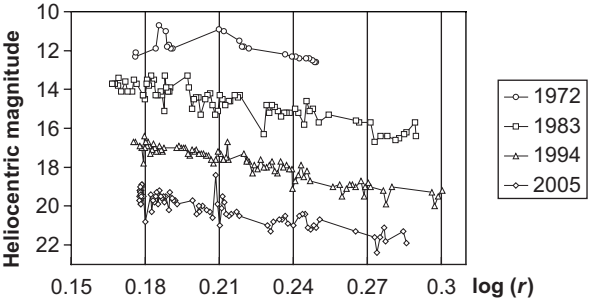


Fig. 1. The light curves of comet 9P/Tempel 1 in four returns. For a better display m_Δ has been displaced by 4^m for 1983, 7^m for 1994, and 9^m for 2005.

for the first time the analogous phenomenon for comet 10P/Tempel 2.[2,3] Later on the authors did show that this phenomenon of cometary outburst activity is typical for all comets.[4]

Unfortunately, post-perihelion light curve can be constructed only for comet's return on 1983. For that only 24 visual brightness estimations have been used. Due to bad conditions of observations for other comet's returns sufficient observations are absent. Therefore, post-perihelion brightness behavior of comet 9P can be not studied in detail.

3.2. *Photometrical parameters*

The values of the photometrical parameters H_0, n, and H_{10} were collected in Table 1. Authors have determined the values of all photometrical parameters for comet's returns on 1972, 1983, 1994, and 2005. For other comet's returns these parameters were taken from Refs. 5–8.

3.3. *Secular brightness variations*

The secular fading of comet 9P is important for possible influence of artificial impact to the nucleus of comet and for cometary photometrical evolution. The change of absolute magnitude H_{10} is presented in Fig. 2, which is constructed on the basis of data, presented in Table 1. From these data the mean fading of comet was calculated to be 0.0060^m for year or about

Table 1. Photometrical parameters of comet 9P/Tempel 1 at all apparitions.

Apparition	H_0	n	H_{10}	Notes	Reference
1867 II	—	—	8.4^m	—	5
1873 I	—	—	9.2	—	5
1879 III	—	—	10.4	—	5
1972 V	—	—	12.0 ± 0.1	—	6
	$10.57^m \pm 0.35^m$	3.00 ± 0.64	9.75 ± 0.09	Pre Per	This paper
1978 II	—	—	13.94	—	7
1983 XI	5.36 ± 0.15	—	8.61 ± 0.03	All	8
	5.82 ± 0.26	9.04 ± 0.50	8.47 ± 0.06	All	This paper
	5.30 ± 0.16	10.02 ± 0.28	8.64 ± 0.03	Pre Per	8
	5.96 ± 0.28	8.82 ± 0.51	8.55 ± 0.07	Pre Per	This paper
	7.86 ± 3.06	10.16 ± 0.29	8.25 ± 0.15	Post Per	8
	5.99 ± 1.41	8.36 ± 3.02	8.01 ± 0.14	Post Per	This paper
1994	5.23 ± 0.19	10.41 ± 0.35	8.62 ± 0.06	Pre Per	This paper
2005	7.19 ± 0.42	8.05 ± 0.71	9.37 ± 0.08	Pre Per	This paper

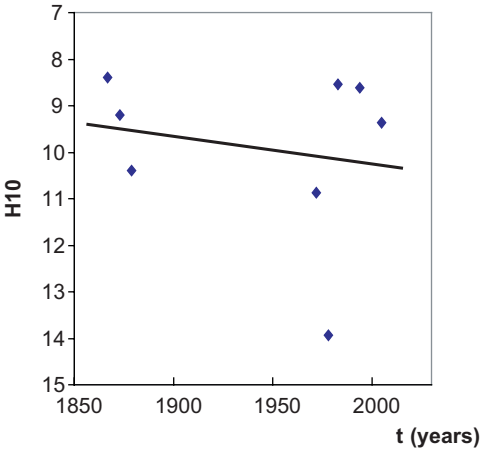


Fig. 2. The secular fading of comet 9P/Tempel 1.

0.04^m for period. This value well corresponds to the mean value of secular fading for short period comets belonging to the Jupiter family (0.05^m for period).[9]

4. Conclusions

(a) The light curve of comet 9P/Tempel 1 was constructed and studied for all observed comet's returns. The values of photometrical parameters H_0, n, and H_{10} were determined.
(b) The value of secular fading of comet was calculated.
(c) The comparison of our results with results of studying of comet's photometrical peculiarities after impact will be able to give information on the change of cometary activity as a result of artificial impact to comet nucleus.

Acknowledgments

Authors are very grateful to Dr. Junichi Watanabe for presenting their work at the AOGS 2005 meeting. Also authors are grateful to Dr. Anil Bhardwaj, editor-in-chief of Advances in Geosciences 2005 (PS volume) and to two anonymous referees for some remarks. Thanks to their considerations the text of our manuscript was greatly improved.

References

1. C. S. Morris, *Astron. Soc. Pacific* (1973), p. 470.
2. K. I. Churyumov and V. S. Filonenko, in *Observations and Physical Properties of Small Solar System Bodies* (Liege, 1992), p. 295.
3. K. I. Churyumov and V. S. Filonenko, *Pisma in Astron. Journal* (1992), p. 922 (in Russian).
4. K. I. Churyumov and V. S. Filonenko, *Astron. Vestnik* (1997), p. 43 (in Russian).
5. S. K. Vsekhsvyatsky, *Physical Characteristics of Comets* (Moscow, 1958) (in Russian).
6. S. K. Vsekhsvyatsky, *Physical Characteristics of Comets 1971–1975* (Kiev, 1979) (in Russian).
7. D. A. Andrienko and A. V. Karpenko, *Physical Characteristics of Comets 1976–1980* (Moscow, 1987) (in Russian).
8. D. A. Andrienko and A. V. Karpenko, *Physical Characteristics of Comets 1981–1985* (Kiev, 1993) (in Russian).
9. J. Svoren, *Contributions of the Astronomical Observatory Skalnate Pleso* (1991), p. 15.

This page intentionally left blank

THREE-DIMENSIONAL MHD SIMULATION OF THE SOLAR WIND INTERACTION WITH COMETS

MEHDI BENNA* and PAUL R. MAHAFFY
NASA — Goddard Space Flight Center, code 699
8800 Greenbelt Road, Greenbelt, MD 20771, USA
**mehdi.benna@gsfc.nasa.gov*

A three-dimensional multi-fluid magnetohydronamic model of the interaction of a cometary coma with the solar wind has been applied to an active Halley-class gas source. The model includes both photo and electron impact ionization, ion recombination, elastic collisions, Coulomb ion/electron interactions, cooling of electrons by water vapor, and the coupling with the magnetic field. Heavy and light ions are treated separately, as necessary. An adaptive mesh refinement is employed to model both near-nucleus kilometer scale phenomena and the much larger structures of the bow shock and the coma tail. With 5.8 million computational cells and 25 levels of refinement, the structures in both the inner coma (within 20,000 km from the nucleus) and in the outer coma were satisfactorily resolved. The model output includes ion density, velocity, temperature, and pressure, electron and neutral temperatures, and magnetic fields. This model is compared to previous simulations and to the location of major plasma boundaries from the 1986 Halley flyby missions. There is satisfactory agreement from this model with the available mission data in features of the plasma structure such as the location of the bow shock, the inner shock, and the mass loading.

1. Introduction

Mass loading is a very common phenomenon in space plasmas. In the case of comets, mass loading controls the structure and the dynamic of the interaction between the solar wind and the cometary atmosphere. During the past decade some of the most sophisticated magnetohydrodynamic (MHD) numerical simulations have been carried out to study the dynamics of the solar wind comet atmosphere interaction and its mass-loading phenomena. The article of Szegö *et al.*[1] provides an excellent review of mass-loading phenomenon and their underling physics. We refer the reader to the papers of Schmidt *et al.*[2] and Gombosi *et al.*[3] for cometary applications of mass-loaded plasmas.

*Corresponding author.

The numerical simulation of the interaction between the cometary atmosphere and the solar wind is a complex problem due to the difference between the two plasmas involved in shaping this coma. The solar wind is a hypersonic hot plasma composed of light ions (mainly protons) and electrons embedded in an interplanetary magnetic field. The plasma in the region of the inner coma formed from a cometary gas released from the comet nucleus is much denser than the solar wind and is composed of ion and neutral species with a range of masses. On release from the nucleus a complex network of chemical reactions produces new chemical species. The cometary gas expands and cools down rapidly once it is released from the nucleus and is often considered to be completely cold ($T_n \sim 0\,\mathrm{K}$ near the nucleus surface) relative to the warmer unperturbed solar wind plasma ($T_{\mathrm{sw}} \sim 10^5\,\mathrm{K}$). Our model includes the major charged particule species for a water-rich nucleus.

In this paper we present early results of the three-dimensional (3-D) version of the multi-fluid MHD simulator (CASIM3D: Comet Atmosphere Simulator 3-D). This code was progressively assembled during the last 2 years to advance cometary modeling techniques by combining adaptive mesh techniques and multi-fluid solvers. The basic multi-dimensional equations of this code and along with its two-dimensional (2-D) application have been presented by Benna *et al.*[4] Section 2.1 describes the physics and chemistry, included in this model, while Sec. 2.2 details the adaptive mesh refinement technique that we implement on a 64-node parallel processing high performance computer. Section 3 presents the output of the MHD model for a comet of the activity level of Halley, using the solar wind conditions relevant for the dates of the Giotto spacecraft flyby. The plasma boundaries we compute are compared to the Giotto derived values.

2. The Cometary Atmosphere Simulator

2.1. *The model*

In the CASIM3D model, we assumed that water vapor molecules dominate the cometary neutrals. When the neutral gas is released from the nucleus by sublimation, it quickly expands, cools down and interacts with the solar UV-EUV radiation and the solar wind plasma. The rich cometary atmosphere is subject to a complex network of chemical reactions and physical interactions that involve neutrals, ions, and electrons. In CASIM3D, we

approximate these interactions to the first order by considering:

- Photoionization and electron impact ionization of heavy neutrals (water molecules).
- Photoionization of light neutrals (atomic hydrogen).
- Recombination of heavy ions (water ions).
- Recombination of light ions (protons).
- Dual elastic collisions between neutrals, ion, and electrons.
- Coulomb interaction between ion and electrons.
- Electronic, vibrational, and rotational cooling of electrons by water vapor.
- Coupling with the magnetic field.

Table 1 summarizes these interactions and shows how they are implanted into the CASIM3D code. In these interactions, we assume that the newly created neutrals, ions, or electrons will instantaneously accommodate, respectively, in the neutral, ion, or electron fluid. The balance between ionization and recombination leads to a "mass-loading" phenomenon, where the cometary ions are picked-up and loaded into the solar wind plasma. The mutual friction between the fluids does not act as a net "mass loading" term. Nonetheless, it acts in shaping the cometary coma through a strong inter-fluid momentum and energy exchange.

Ionization and recombination rates and collisional processes are modeled using the literature models and values. We based our work on the

Table 1. Chemical and physical interactions included in CASIM3D.

	Neutral fluid		Ion fluid		Electron fluid	
Interaction	H_2O	H	H_2O^+ H_3O^+	H^+	Thermal, and suprathermal electrons	References
Photoionization	▼	▼	▲	▲	▲	5 and 6
Impact ionization	▼		▲		▲	7
Recombination	▲	▲	▼	▼	▼	8
Elastic and inelastic collisions	♦	♦	♦	♦	♦	9, 10 and 11
Electronic, vibrational, rotational cooling, or heating					♦	12

Note: (▼) Designates a mass, energy or momentum loss term, (▲) a gain term, and (♦) a loss or gain term.

results of Häberli *et al.*[5,8] to describe the photoionization and recombination reactions. The ionization frequency profiles of Cravens *et al.*[7] take into account the electron impact ionization. Schunk and Nagy[9] provide an excellent review of the collision processes in planetary atmospheres. The computation of the elastic collision effect between neutrals and ions and between ions and electrons is based on their work. We used the theoretical and experimental profiles of Itikawa[10,11] to incorporate the neutral-electron collision effect. The electronic, vibrational, and rotation cooling of electrons by water vapor is based on the study of Gan and Cravens.[12]

2.2. *The adaptive mesh refinement technique*

The size of the comet nucleus is of the order of 10 km, but the distance to bow shock from the nucleus can be five orders of magnitude larger, and the coma/tail extends downstream to a distance up to 10^8 km. To resolve the finest of these scales while including the largest scale requires the use of a mesh for the calculation that is adaptive and can adjust dynamically as the structure in the solution evolves. To achieve this, the CASIM3D code uses the PARAMESH tool kit of MacNeice *et al.*[13] Figure 1 shows an example of a 2-D adaptive mesh.

PARAMESH implements a subset of the Berger–Oliger type of block-structured or patch-based adaptive mesh refinement.[14,15] In these schemes, space is divided into a set of contiguous blocks. Each block is a Cartesian, uniformly spaced sub-mesh. Each block can itself be refined into a set of "child" blocks with half the grid spacing of their parent block. In PARAMESH, space is recursively divided by bisecting the computational volume in each coordinate direction. In a 3-D application, each block is a structured sub-mesh with $n_x \times n_y \times n_z$ cells, and each block has an identical number of cells. For CASIM3D, $n_x = n_y = n_z = 4$. PARAMESH enforces the condition that jumps in resolution by more than a factor of 2 are not allowed. Each sub-mesh block is surrounded by guard cells so that the mesh cells in the sub-mesh that are near to the sub-mesh boundary have enough information about the solution in the neighboring sub-mesh blocks. For CASIM3D, there are two layers of guard cells at each sub-mesh boundary. PARAMESH fills these guard cells by copying data directly from the neighboring blocks if the neighbor is at the same refinement level. In the case of a refinement discontinuity at the block boundary the guard cells are filled either by restriction if the neighboring sub-mesh is more refined or by interpolation if it is less refined. PARAMESH allows the user to specify the

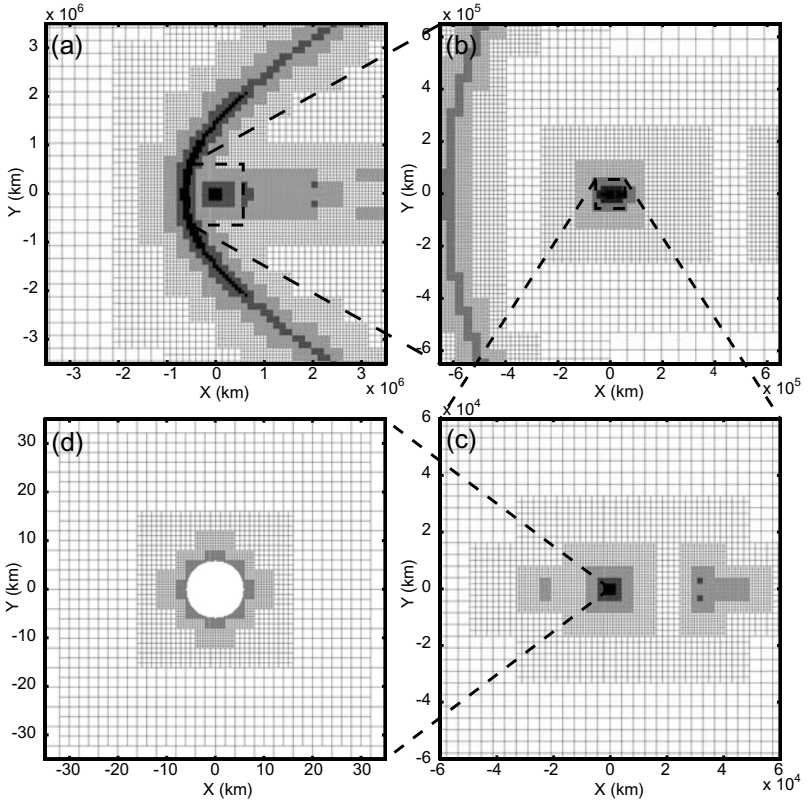


Fig. 1. Mesh obtained at the end of the 2-D CASIM simulation for a Halley-type comet. (a)–(d) showing the progression of the grid refinement. The nucleus can be seen in the center of (d). Note that (a) shows only a part of the more extended computational domain.

restriction and interpolation functions to be used in these situations. For CASIM3D, restriction is done using simple averaging of data at cell center. Interpolation is performed using bilinear interpolation with a MUSCL-type limiter (see Ref. 16).

3. Modeling of a Halley-Type Comet

In this simulation we consider a set of physical parameters globally consistent with comet Halley's parameters as they were observed during the 1986 flybys. We only report in this paper the main variables that define

Table 2. Parameters used for the modeling of a Halley-type comet.

Parameter	Value	Parameter	Value
Nucleus radius	6 km	Solar wind speed	$375\,\mathrm{km\,s^{-1}}$
Gas production rate	$7 \times 10^{29}\,\mathrm{mol\,s^{-1}}$	Solar wind mach number	10
Neutral gas surface temperature	200 K	Solar wind mean molecular mass	1 amu
Mean neutrals molecular mass	18 amu	Solar wind number density	$7\,\mathrm{ions\,cm^{-3}}$
Initial neutral gas speed	$0.6\,\mathrm{km\,s^{-1}}$	Solar wind ion temperature	10^5 K
Ionization life time	10^6 s	Solar wind electron temperature	2.5×10^5 K
IMF magnitude	5.6 nT	IMF angle	45°

the boundary conditions and describe the major physical interactions. The values of these variables are summarized in Table 2.

In CASIM3D the MHD equations are integrated in 3-D using the total variation diminishing Lax-Friedrichs (TVDLF) solver. The TVDLF solver is one of the high-resolution schemes designed to solve hyperbolic conservation laws and capture complex hydrodynamic shock structures without oscillations. Yee[17] provides a good review of this class of schemes. We have described in detail in Ref. 4 these MHD equations that govern the behavior of a cometary atmosphere composed of three fluids: ions, electrons, and neutrals.

Figures 2–4 show the final steady state solution obtained through this process. The solution was computed over 5,836,800 physical cells spread over 25 levels of refinements. The largest computational cell is about 4,200,000 km and the smallest cell is about 250 m. These figures exhibit, respectively, the ion density, temperature, and velocity profiles in the inner coma ($R_c \leq 20{,}000$ km) and in the outer part of the coma ($R_c > 20{,}000$ km). This validates the capabilities of the CASIM3D code to model the cometary atmosphere from the surface to large cometocentric distances.

We present in Fig. 5 a cross section along the the Sun–Comet line of the ion Mach number, pressure and density along with the magnetic field magnitude. These profiles show clearly the formation of the bow shock at $R_b = 5.5 \times 10^5$ km and the formation of the inner shock at $R_i = 2{,}700$ km along the Sun–Comet line. The bow shock is formed as a result of the fast deceleration of the hypersonic solar wind stream to subsonic speeds while ions of cometary origin are imbedded in it. The inner shock is a result of the deceleration of the hypersonic radially expanding cometary ions and its

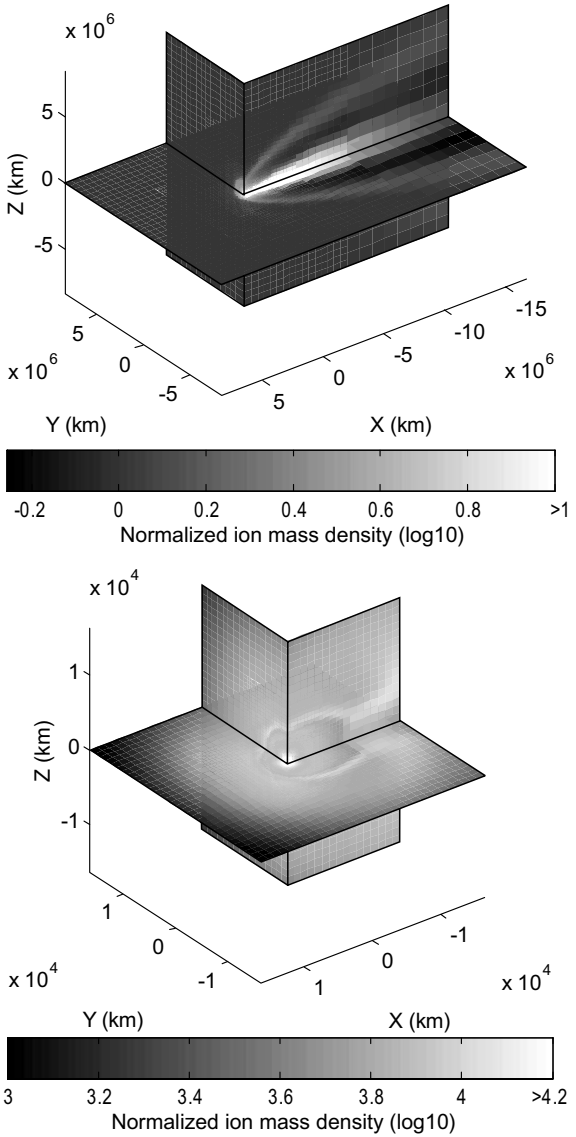


Fig. 2. Normalized ion mass density of a Halley-type comet obtained with CASIM3D. Upper figure: Large structures in the outer coma. Lower figure: Structures in the inner coma. The mass density is given in units of mass density of the unperturbed solar wind. The Sun is along the $X > 0$ axis.

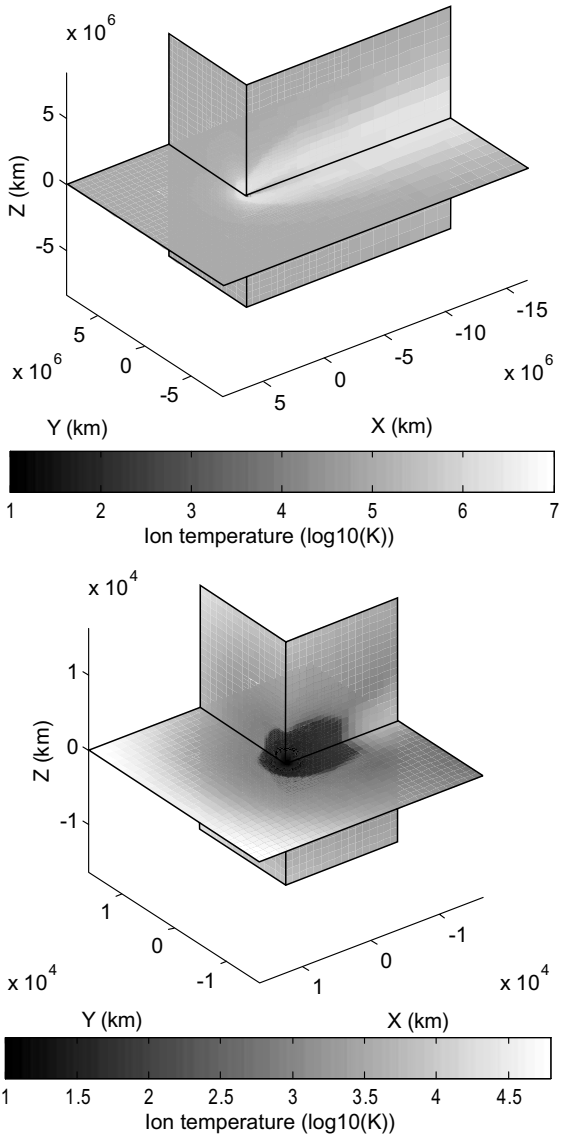


Fig. 3. Ion tempertaure of a Halley-type comet obtained with CASIM3D. Upper figure: Large temperature structures in the outer coma. Lower figure: Temperature structures in the inner coma. The Sun is along the $X > 0$ axis.

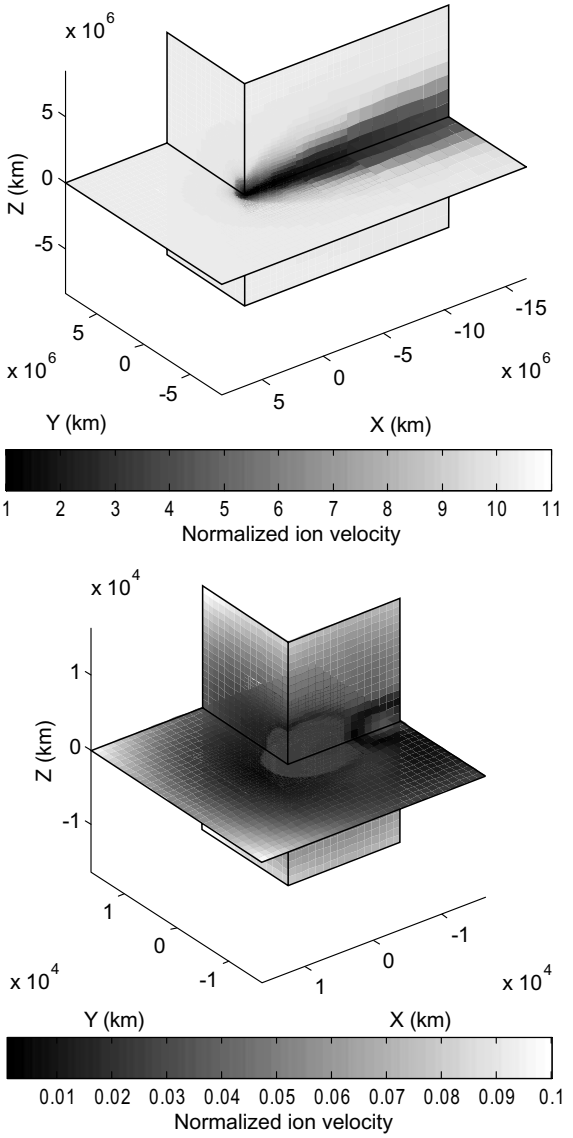


Fig. 4. Normalized Ion velocity of a Halley-type comet obtained with CASIM3D. Upper figure: Large velocity structures in the outer coma. Lower figure: Velocity structures in the inner coma. The ion velocity is given in units of velocity of the unperturbed solar wind. The Sun is along the $X > 0$ axis.

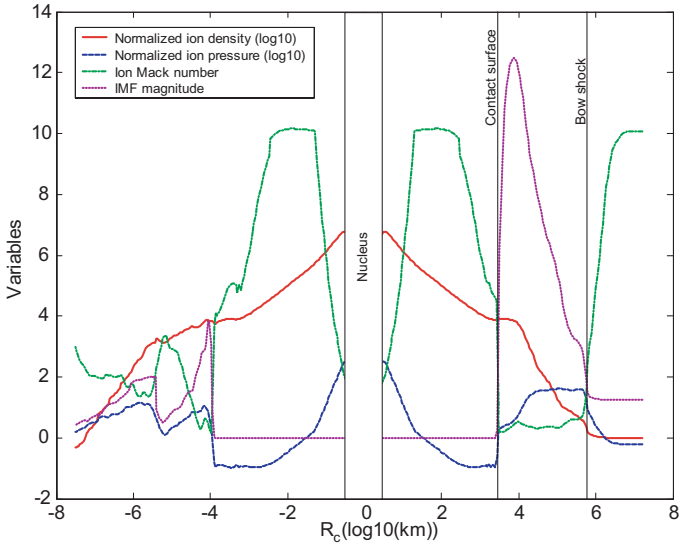


Fig. 5. Plasma parameters along the Sun–Comet line. Both variables and distances are logarithmic. The ion mass density and pressure and IMF magnitude are normalized compared to the values in the unperturbed solar wind. The Sun is along the $X > 0$ axis.

interaction with the contaminated solar wind stream at the contact surface. The position of the bow shock and the inner shock are in a good agreement with Gombosi *et al.* and Schmidt-Voigt numerical models.[18,19]

We also estimated the position of the bow shock and the contact surface along the inbound Giotto trajectory to be, respectively, $\tilde{R}_{\mathrm{b}} = 1.3 \times 10^6$ km and $\tilde{R}_{\mathrm{i}} = 4{,}300$ km. These values are very close to the values derived from the data of the RPA-COPERNIC and the GIOTTO-MAG instruments ($\tilde{R}_{\mathrm{b}} = 1.15 \times 10^6$ km and $\tilde{R}_{\mathrm{i}} = 4{,}660$ km) reported by Rème and Neubauer in their publications.[20,21]

Figure 6 presents a cross section along the the Sun–Comet line of the ion temperature, the thermal electron temperature and the neutral temperature. This profile gives a clear idea of the cooling effect affecting the electrons in the inner coma du to rotational, vibrational, and electronic excitation of water vapor.

Currently, we are adjusting the CASIM3D free parameters to better fit Giotto data. Deeper analysis of the resulting model and comparison of all the fluid parameters with the available Halley's flybys data are in preparation. We will account of these results in a future publication.

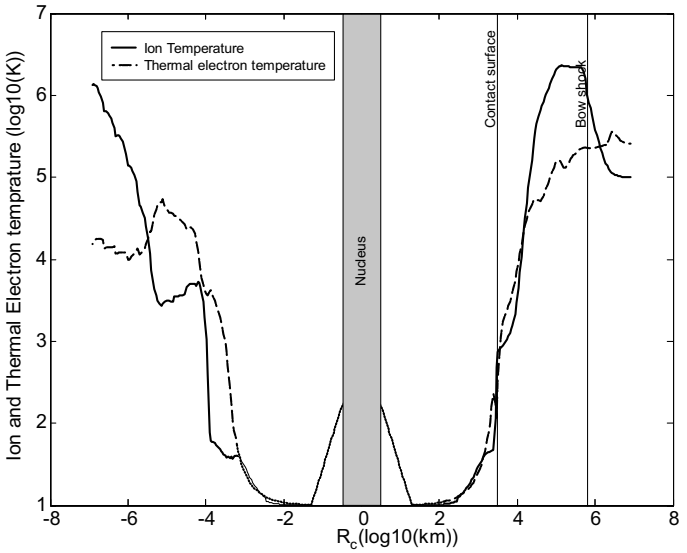


Fig. 6. Ion and thermal electron temperature along the Sun-Comet line. Both variables and distances are logarithmic. Note the electron cooling effect in the inner coma due to the Electronic, variational, and rotation excitation of water vapor. The Sun is along the $X > 0$ axis.

4. Conclusion

In this work we present few results of the our 3-D multi-fluid model of a Halley-type comet. The results show a good agreement with the expected fluid boundaries found by other teams, and early their analysis agrees with the Giotto data gathered during the Halley's flyby. We have demonstrated the capability of CASIM3D as a multi-fluid MHD model that not only derives the ions parameters, but also computes the thermal and supra-thermal electron distributions and temperatures and the neutral parameters. Deeper analysis and comparison with theoretical models and experimental data are in process and will be released in future publications.

Acknowledgments

The work presented herein has been funded by the National Academies and by NASA Goddard Space Flight Center. Calculations were carried out on the high-performance computing resources of the NASA Center for Computational Sciences.

References

1. K. Szegö, K.-H. Glassmeier, R. Bingham, A. Bogdanov *et al.*, *Space Sci. Rev.* **94** (2000) 429.
2. H. U. Schmidt, R. Wegmann, W. F. Huebner and D. C. Boice, *Comput. Phys. Commun.* **49** (1988) 17.
3. T. I. Gombosi, D. L. DeZeeuw and R. M. Häberli, *J. Geophys. Res.* **A7**, 15 (1996) 233.
4. M. Benna, P. Mahaffy, P. MacNiece and K. Olson, *Astrophys. J.* **617** (2004) 666.
5. R. M. Häberli, K. Altwegg, H. Balsiger and J. Geiss, *Astron. Astrophys.* **297** (1995) 881.
6. J. Vön Oertzen, in PhD Thesis, University of K¨ln, Germany, 2003.
7. T. E. Cravens, J. U. Kozyra, A. F. Nagy, T. I. Gombosi and M. Kurtz, *J. Geophys. Res.* **92** (1987) 7341.
8. R. M. Häberli, K. Altwegg and H. Balsiger, *J. Geophys. Res.* **101**, 15 (1996) 579.
9. R. W. Schunk and AD. F. Nagy, *Rev. Geophys. and Space Phys.* **18** (1980) 813.
10. Y. Itikawa, *Planet. Space Sci.* **19** (1971) 993.
11. Y. Itikawa, *Phys. Fluids* **16** (1973) 831.
12. L. Gan and T. E. Cravens, *J. Geophys. Res.* **95** (1990) 6285.
13. P. MacNeice, K. M. Olson, C. Mobarry, R. De Fainchtein and C. Packer, *Comp. Phys. Comm.* **126** (2000) 330.
14. M. J. Berger and J. Oliger, *Comp. Phys. Comm.* **53** (1984) 484.
15. M. J. Berger and P. Colella, *Comp. Phys. Comm.* **82** (1989) 64.
16. D. L. DeZeeuw, in PhD Thesis, University of Michigan, USA, 1993.
17. H. C. Yee, in *NASA TM-101088*, NASA, Washington, 1989.
18. T. I. Gombosi, D. L. DeZeeuw and R. M. Häberli, *J. Geophys. Res.* **101** (1996) 15233.
19. M. Schmidt-Voigt, *Astronom. Astrophys.* **210** (1987) 433.
20. H. Rème, in *Comet Halley: Investigations, Results, Interpretations*, Vol. 1, ed. J. Masson (Ellis Horwood, London, 1990), p. 87.
21. F. M. Neubauer, in *Comet Halley: Investigations, Results, Interpretations*, Vol. 1, ed. J. Masson (Ellis Horwood, London, 1990), p. 79.

XMM-NEWTON OBSERVATIONS OF X-RAY EMISSION FROM JUPITER

G. BRANDUARDI-RAYMONT*, A. BHARDWAJ†,**,
R. ELSNER†, R. GLADSTONE‡, G. RAMSAY*, P. RODRIGUEZ§,
R. SORIA*, H. WAITE¶ and T. CRAVENS‖
*Mullard Space Science Laboratory, University College London
Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK
†NASA Marshall Space Flight Center, NSSTC/XD12
320 Sparkman Drive, Huntsville, AL 35805, USA
‡Southwest Research Institute, P.O. Drawer 28510
San Antonio, TX 78228, USA
§XMM-Newton SOC, Apartado 50727, Villafranca, 28080 Madrid, Spain
¶University of Michigan, Space Research Building
2455 Hayward, Ann Arbor, MI 48109, USA
‖Department of Physics and Astronomy
University of Kansas, Lawrence, KS 66045, USA
*gbr@mssl.ucl.ac.uk

XMM-Newton observations of Jupiter show prominent soft X-ray emission from the auroral spots and, albeit at a lower intensity, from the equatorial regions. While the spectra of the auroral X-rays can be modeled with a superposition of unresolved emission lines, including most prominently those of highly ionized oxygen, Jupiter's equatorial 'disk' emission has a spectrum consistent with that of solar X-rays scattered in the planet's upper atmosphere. Remarkably, a large solar X-ray flare, which took place on the Sun's Jupiter-facing side in November 2003, is found to be associated with a corresponding feature in the Jovian disk X-ray lightcurve. This suggests that the non-auroral X-ray emission from Jupiter is directly controlled by the Sun. However, the *XMM-Newton* results support the view that Jupiter's auroral emissions originate from the capture and acceleration of heavy ions, followed by X-ray production by charge exchange. They presently favor a solar wind scenario, although a magnetospheric origin for some of the ions cannot be excluded.

1. Introduction

XMM-Newton,[1] the state-of-the-art X-ray astronomical observatory launched by the European Space Agency in 1999, provides us with an unparalleled combination of high spectral sensitivity and excellent energy

**Present address: Space Physics Laboratory, Vikram Sarabhai Space Center, Trivandrum 695022, India.

resolution; as such, it is a powerful diagnostic tool in the study of cosmic X-ray sources, from distant active galaxies to planets in our own solar system; a tool which has opened new perspectives on the Universe. Here we report on two recent *XMM-Newton* observations of Jupiter.

Jupiter was first detected at X-ray energies with the *Einstein Observatory*,[2] and was later studied with *ROSAT*.[3–5] *Chandra* HRC-I observations (Fig. 1) provide the clearest view yet of the planet's X-ray emission: they reveal two bright, high-latitude sources associated with the aurorae, as well as low-latitude emission from the planet's disk.[6,7] Moreover, strong 45 min quasi-periodic X-ray oscillations were discovered by *Chandra* in the North auroral spot in December 2000. No correlated periodicity was seen at the time in *Cassini* solar wind data, nor in *Galileo* and *Cassini* energetic particle and plasma wave measurements.

Two possible origins (both involving charge exchange) have been proposed for Jupiter's X-ray auroral emissions: energetic (>1 MeV/nucleon) S and O ions from the outer magnetosphere, or highly ionized, but low energy, solar wind ions, entering the polar/auroral regions along field lines; in both cases acceleration by field-aligned potentials is required.[7,8] Oscillations in brightness may arise from processes related to the magnetosphere (such as

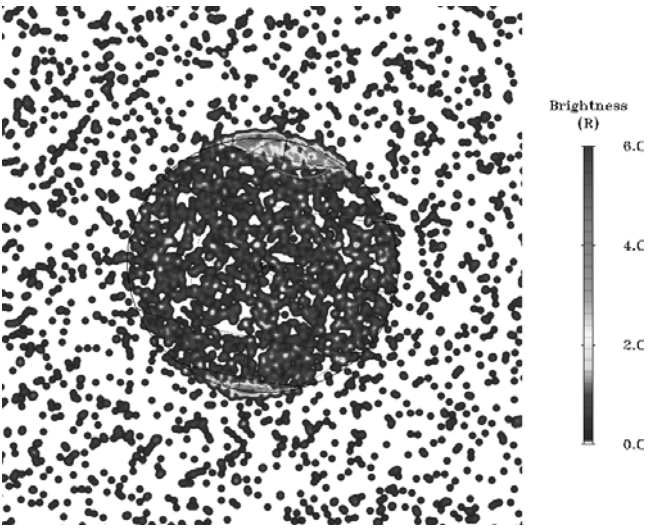


Fig. 1. *Chandra* HRC-I image of Jupiter (December 18, 2000), clearly separating the polar, auroral X-ray sources, and the low-latitude disk X-ray emission. The scale bar is in units of Rayleighs (R).

the pulsed reconnection at the day-side magnetopause suggested by Bunce *et al.*[9]) or in the solar wind.

2. *XMM-Newton* Observations: Lightcurves

XMM-Newton observed Jupiter twice in 2003: between April 28, 16:00 and April 29, 22:00 UT (for a total observing time of 110 ks[10]), and between November 25, 23:00 and November 29, 12:00 UT (245 ks,[11,12] see Fig. 2). Lightcurves from the November observation, shown in Fig. 3, resemble very closely those obtained the previous April.[10] In the following we focus on the *XMM-Newton* November observation and make comparison, where appropriate, with the one in April, which is already published.[10]

Exclusion of data affected by high background at the end of both spacecraft orbits (top panel in Fig. 3, >10 keV energy) leaves 210 ks of good quality data for analysis. The planet's 10 h rotation period is clearly seen in the

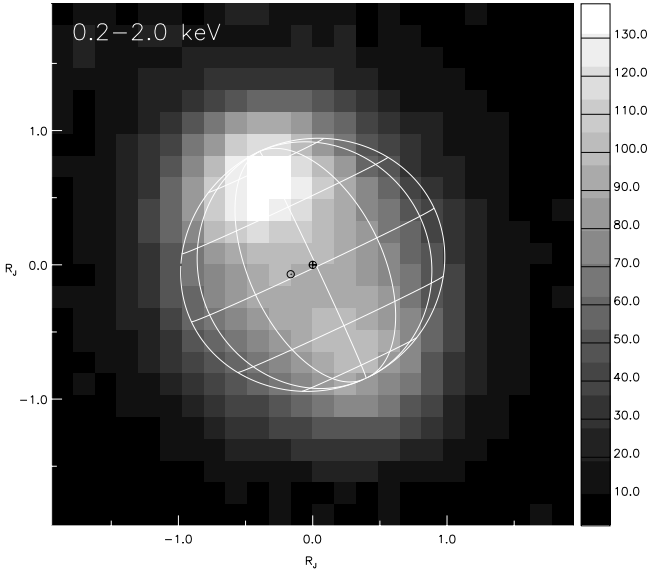


Fig. 2. Smoothed *XMM-Newton* EPIC image of Jupiter (0.2–2.0 keV; 2.9″ pixels), obtained by integrating the X-ray emission detected over the whole November 2003 observation. The bright auroral emissions at the planet's poles, as well as the low-latitude disk emission, are clearly visible. A graticule showing Jupiter orientation with 30° intervals in latitude and longitude is overlaid. The circular mark indicates the sub-solar point; the sub-Earth point is at the center of the graticule. The scale bar is in units of EPIC counts.

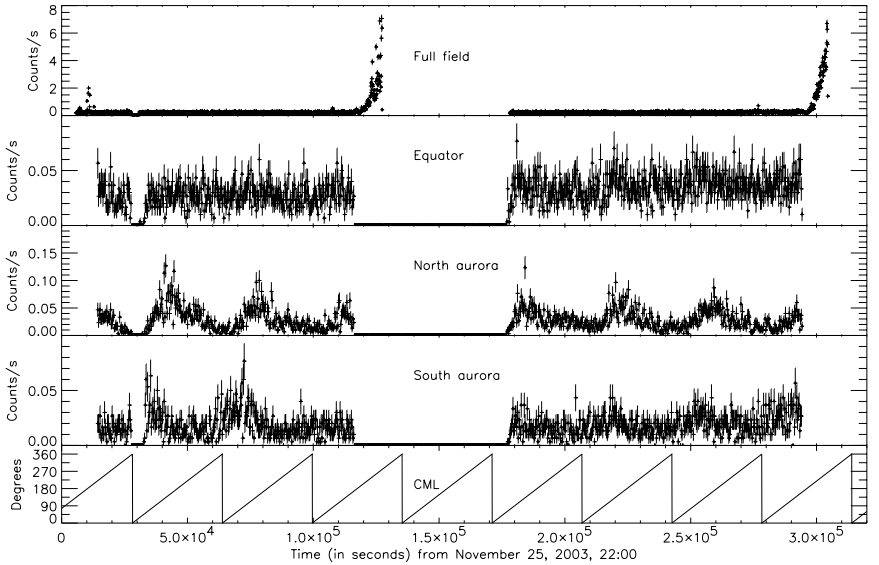


Fig. 3. Jupiter lightcurves (0.2–2.0 keV, 300 s bins) from the November 2003 *XMM-Newton* observations. CML: System III Central Meridian Longitude.

lightcurve of the North and South auroral spots, but not in the equatorial region. However, there is an ∼40% increase in the equatorial flux between beginning and end of the November observation, which is found to correlate with a similar increase in solar X-ray flux. The bottom panel in Fig. 3 shows the System III Central Meridian Longitude (CML). The North spot is brightest around CML = 180°, similar to the December 2000 *Chandra* and the April 2003 *XMM-Newton* results.

Remarkably, a large solar X-ray flare which took place on the Sun's Jupiter-facing side during the *XMM-Newton* observations has been found to be associated with a corresponding feature in the Jovian equatorial lightcurve (Fig. 4).[11] This, and the overall increase in the equatorial X-ray flux during the November observation, suggest that the non-auroral X-ray output from Jupiter is directly controlled by the Sun.

A search for periodic behavior on short timescales in the auroral soft X-rays (i.e. the *Chandra* 45 min oscillations) has been carried out by generating amplitude spectra from the lightcurves; Jupiter's 10 h rotation period is clearly detected in the North and South aurorae (see Fig. 5), but no evidence is found for short timescale periodicities, as none was found in the

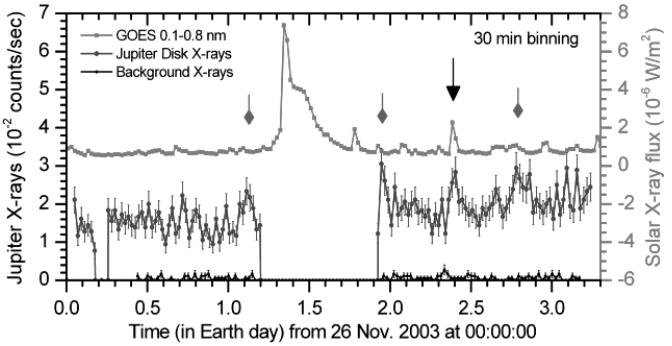


Fig. 4. Comparison of 30-min binned Jupiter disk X-rays (middle curve) with *GOES* 10 1.5–12 keV solar X-ray data (top curve). The X-ray background lightcurve is at the bottom, just emerging from zero level. The Jovian X-ray time is shifted by −4948 s to account for light travel time delays. The gap between 1.2 and 1.9 days (when the largest solar flare of this period took place!) is due to *XMM-Newton* perigee passage. The big, triangular-head arrow at 2.4 days refers to the time of the largest solar flare visible from both Earth and Jupiter, which has a clear matching peak in the Jovian lightcurve. Smaller, diamond-head arrows mark peaks in the Jupiter lightcurve which could correspond to solar flares occurring on the western (Earth-hidden) side of the Sun.

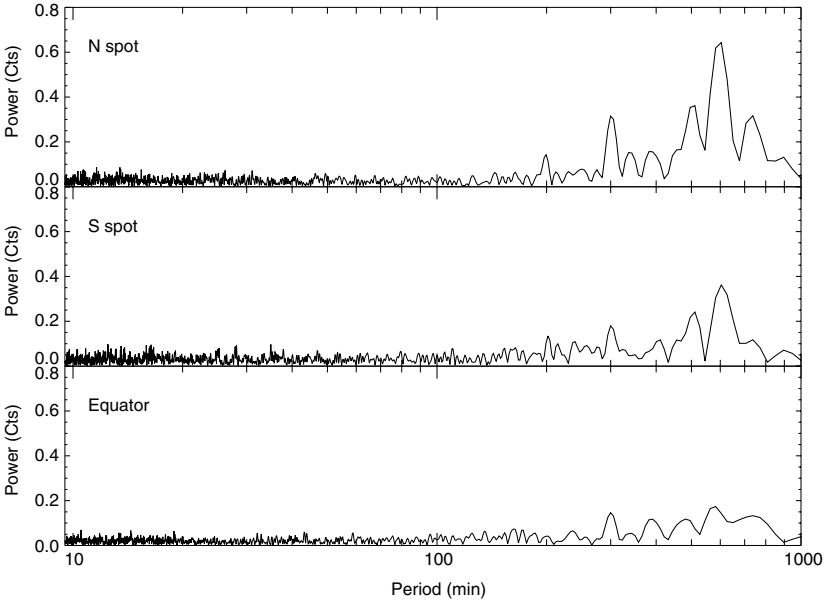


Fig. 5. Amplitude spectra generated from the lightcurves in Fig. 3: Jupiter's 10 h rotation period is clearly detected in the North and South aurorae, but not in the low-latitude emission; there is no evidence for periodic or quasi-periodic variability at short timescales.

April 2003 *XMM-Newton* data. In February 2003 *Chandra* detected quasi-periodic variability on timescales in the range 20–70 min. This suggests that over time the character of the variability in the auroral X-ray emissions can change from well organised to chaotic.

3. *XMM-Newton* Spectral Images

The first observation of Jupiter by *XMM-Newton* in April 2003[10] indicated that the auroral soft X-ray spectra can be modelled with a superposition of emission lines, including most prominently those of highly ionized oxygen (OVII and OVIII). Instead, Jupiter's low-latitude X-ray emission displays a spectrum consistent with that of solar X-rays scattered in the planet's upper atmosphere. These results are strengthened by the November 2003 observations.

Figures 6–9 show EPIC[13,14] CCD images in narrow spectral bands corresponding to the OVII, OVIII, FeXVII, and MgXI lines detected in Jupiter's spectra: OVII emission is concentrated mostly in the North and (more weakly) the South auroral spots, OVIII extends to lower latitudes,

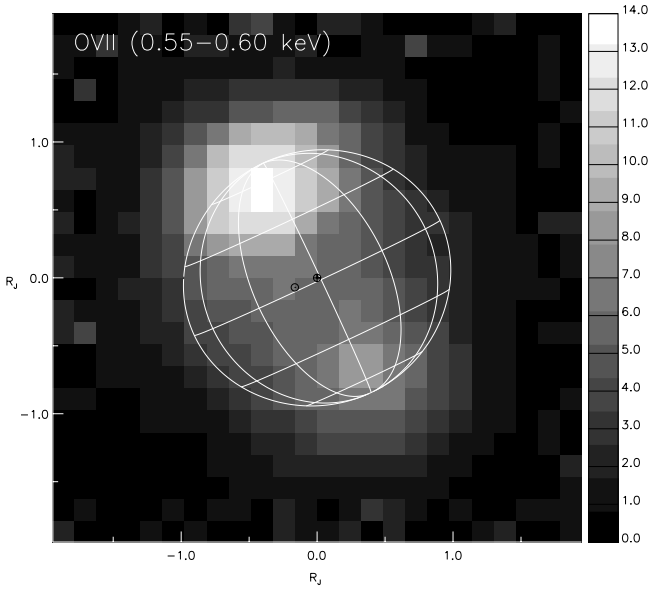


Fig. 6. Smoothed *XMM-Newton* EPIC image of Jupiter in a narrow spectral band centered on OVII. The scale bar is in units of EPIC counts.

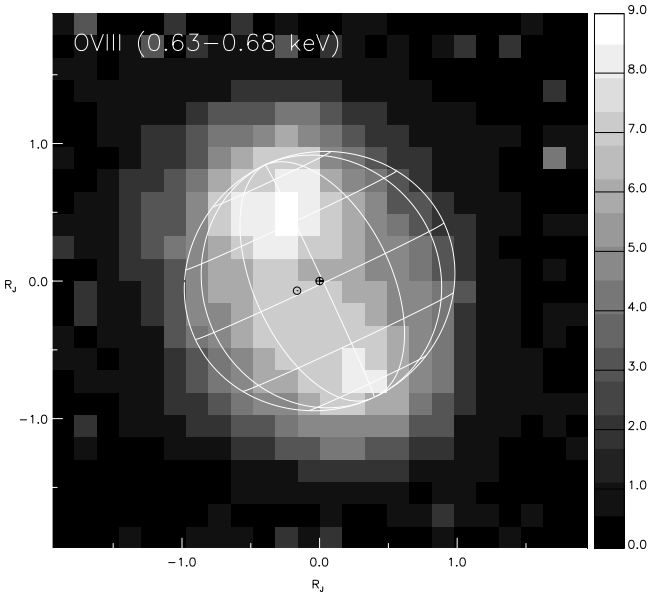


Fig. 7. As Fig. 5, for OVIII.

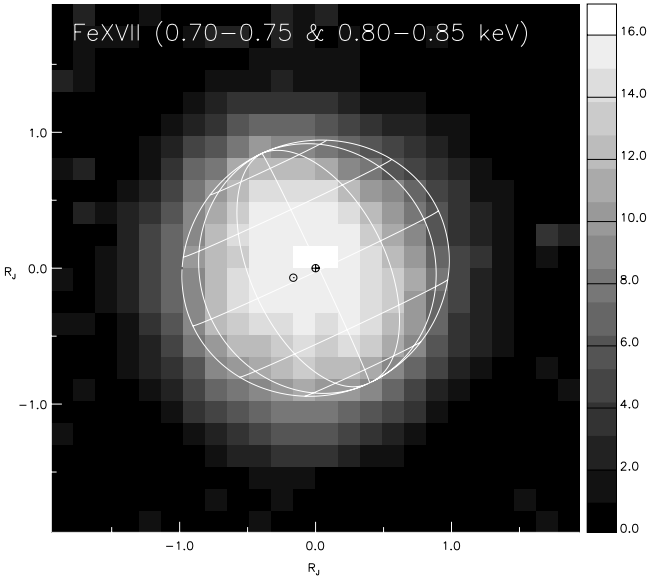


Fig. 8. As Fig. 5, for FeXVII.

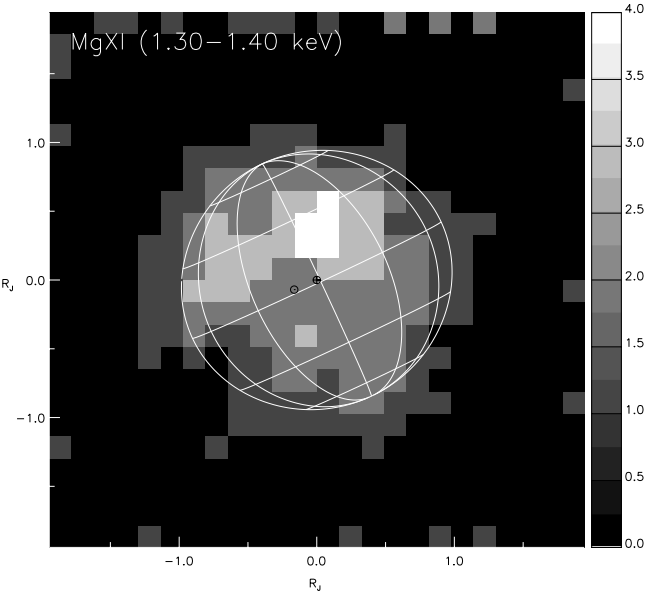


Fig. 9. As Fig. 5, for MgXI.

while FeXVII and MgXI display a rather uniform distribution over the planet's disk, consistent with an origin from scattered solar X-rays.

4. *XMM-Newton* EPIC and RGS Spectra

EPIC CCD spectra of Jupiter's auroral zones and of the low-latitude disk emission were extracted using the spatial selection regions shown in Fig. 10; their spectral 'mixing' (due to the *XMM-Newton* Point Spread Function) was corrected for by subtracting appropriate fractions of disk and auroral spectra, respectively.[12]

A combination of two collisional plasma models (`mekal` in XSPEC) fits the EPIC soft X-ray spectra of Jupiter's aurorae well (see Fig. 11 for the North aurora); the plasma temperatures are $kT = 0.17 \pm 0.01$ and 0.45 ± 0.10 keV: the cooler component is required to explain the prominent OVII emission, and the higher temperature one to describe the higher energy part of the spectrum. In practise, we use a thin plasma code to mimic the recombination process inherent to the charge exchange mechanism. A single `mekal` plasma model ($kT = 0.46 \pm 0.03$ keV, with a normalization six times larger than that required for the auroral spectral component with the same

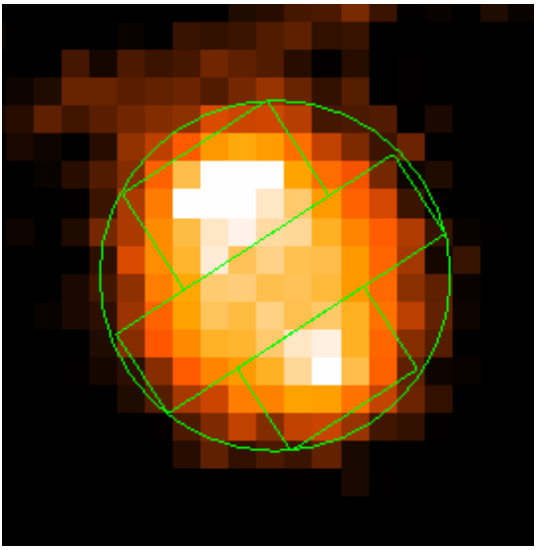


Fig. 10. *XMM-Newton* EPIC image of Jupiter from the November 2003 observation, with extraction regions for the auroral and equatorial spectra superimposed.

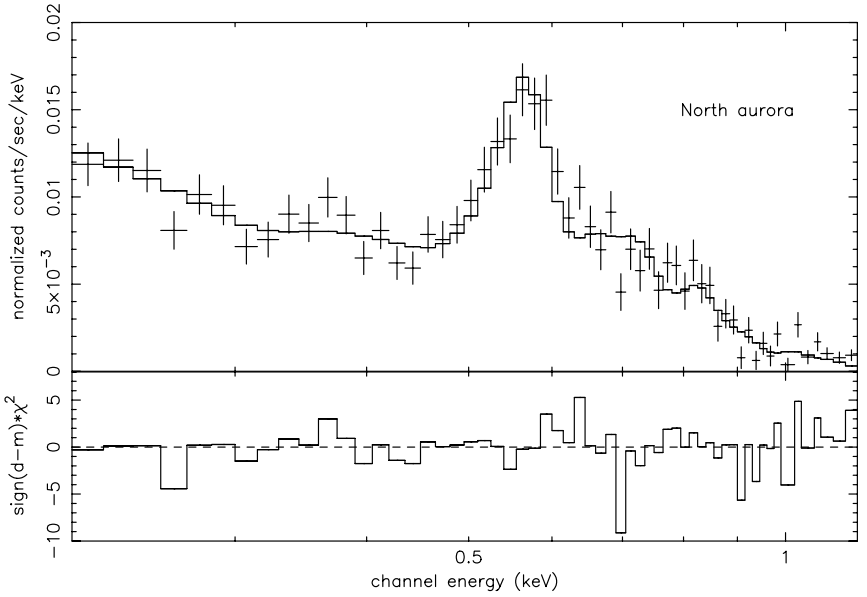


Fig. 11. *XMM-Newton* EPIC spectrum of Jupiter's North aurora and 2-`mekal` model best fit, with solar wind abundances (see text for details).

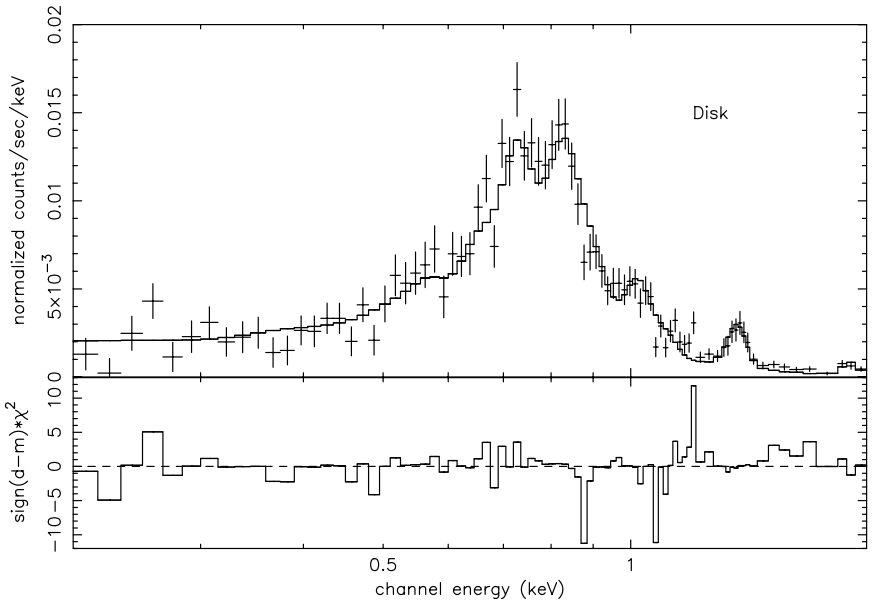


Fig. 12. *XMM-Newton* EPIC spectrum of Jupiter's disk and `mekal` best fit, with solar abundances (see text for details).

temperature) is also a good representation of the low-latitude disk spectrum (Fig. 12), after including additional MgXI and SiXIII emission (at 1.35 and 1.86 keV, respectively, likely consequences of enhanced solar activity) and a small contribution of OVII (0.57 keV) and OVIII (0.65 keV, both residual auroral contamination). However, the auroral spectra are better fitted when using elemental abundances appropriate to the solar wind,[15] while solar values[16] apply to the equatorial spectrum.

The higher resolution RGS[17] spectrum, which includes X-ray light from the whole planet, agrees well, in flux and profile, with the EPIC one integrated over the full disk of Jupiter (Fig. 13). The RGS clearly resolves the strongest emission lines. A simultaneous fit (combination of `mekal` models as for the auroral spectra) shows some discrepancies from the data: this highlights the difficulty of fitting disk and auroral spectra together (because they are clearly different and the RGS cannot easily separate the two spatially) as well as the fact that a thin plasma model is not an adequate approximation of the charge exchange mechanism (especially when fitting high resolution spectra). More appropriate model descriptions are being investigated.[12]

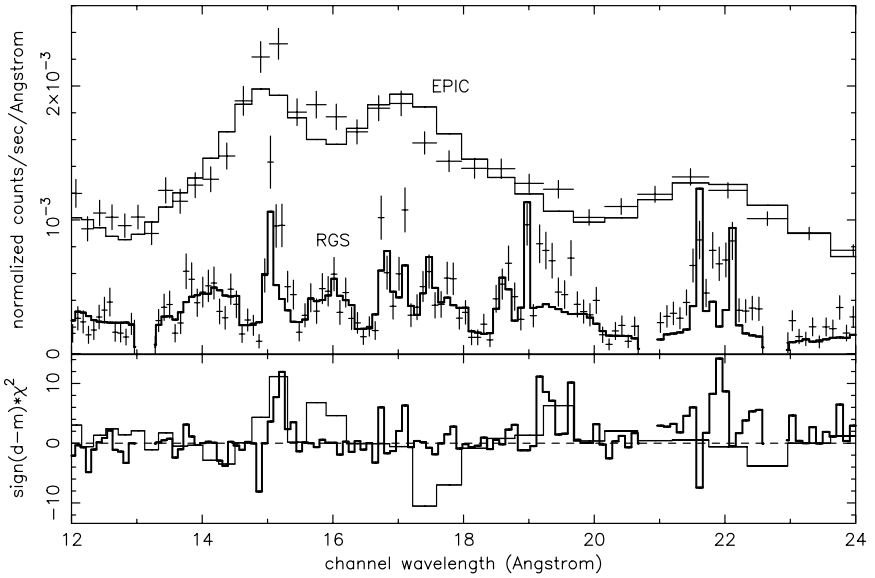


Fig. 13. Top panel: *XMM-Newton* EPIC (upper curve) and RGS (lower curve) soft X-ray spectra of Jupiter (full planet) and best fit with a combination of `mekal` models. Bottom panel: Individual spectral channel contributions to the χ^2 value (thin line: EPIC; thick line: RGS).

5. Discussion and Conclusions

Our *XMM-Newton* results support the view that Jupiter's auroral emissions originate from precipitating energetic heavy ions, producing soft X-rays by charge exchange. They presently favor a solar wind scenario[10] while *Chandra* data[7] are suggestive of a magnetospheric origin. A pure solar wind scenario, though, is probably excluded because it would lead to Lyα production in excess of what is observed.[8] More detailed modeling is being carried out to establish the true origin of the X-ray lines, and studies are underway to resolve the differences.[12] It is conceivable that both scenarios play a rôle in what is certainly a very complex planetary structure.

Acknowledgments

This work is based on observations obtained with *XMM-Newton*, an ESA science mission with instruments and contributions directly funded by ESA

Member States and the USA (NASA). The MSSL authors acknowledge financial support from PPARC.

References

1. F. Jansen, D. Lumb, B. Altieri, J. Clavel, M. Ehle, C. Erd, C. Gabriel, M. Guainazzi, P. Gondoin, R. Much, R. Munoz, M. Santos, N. Schartel, D. Texier and G. Vacanti, *Astron. Astrophys.* **365** (2001) L1.
2. A. E. Metzger, J. L. Luthey, D. A. Gilman, K. C. Hurley, H. W. Schnopper, F. D. Seward and J. D. Sullivan, *J. Geophys. Res.* **88** (1983) 7731.
3. J. H. Waite Jr., F. Bagenal, F. Seward, C. Na, G. R. Gladstone, T. E. Cravens, K. C. Hurley, J. T. Clarke. R. Elsner and S. A. Stern, *J. Geophys. Res.* **99** (1994) 14799.
4. J. H. Waite Jr., G. R. Gladstone, W. S. Lewis, P. Drossart, T. E. Cravens, A. N. Maurellis, B. H. Mauk and S. Miller, *Science* **276** (1997) 104.
5. G. R. Gladstone, J. H. Waite Jr. and W. S. Lewis, *J. Geophys. Res.* **103** (1998) 20083.
6. G. R. Gladstone, J. H. Waite Jr., D. Grodent, W. S. Lewis, F. J. Crary, R. F. Elsner, M. C. Weisskopf, T. Majeed, J.-M. Jahn, A. Bhardwaj, J. T. Clarke, D. T. Young, M. K. Dougherty, S. A. Espinosa and T. E. Cravens, *Nature* **415** (2002) 1000.
7. R. F. Elsner, N. Lugaz, J. H. Waite Jr., T. E. Cravens, G. R. Gladstone, P. Ford, D. Grodent, A. Bhardwaj, R. J. MacDowall, M. D. Desch and T. Majeed, *J. Geophys. Res.* **110** (2005) A01207.
8. T. E. Cravens, J. H. Waite, T. I. Gombosi, N. Lugaz, G. R. Gladstone, B. H. Mauk and R. J. MacDowall *J. Geophys. Res.* **108** (2003) 1465.
9. E. Bunce, S. Cowley and T. Yeoman, *J. Geophys. Res.* **109** (2004) A09S13.
10. G. Branduardi-Raymont, R. F. Elsner, G. R. Gladstone, G. Ramsay, P. Rodriguez, R. Soria and J. H. Waite Jr., *Astron. Astrophys.* **424** (2004) 331.
11. A. Bhardwaj, G. Branduardi-Raymont, R. F. Elsner, G. R. Gladstone, G. Ramsay, P. Rodriguez, R. Soria, J. H. Waite Jr. and T. E. Cravens, *Geophys. Res. Lett.* **32** (2005) L03S08.
12. Branduardi-Raymont *et al.*, *Astron. Astrophys.*, in preparation.
13. L. Strüder, U. Briel, K. Dennerl and 55 co-authors, *Astron. Astrophys.* **365** (2001) L18.
14. M. J. L. Turner, A. Abbey, M. Arnaud and 60 co-authors, *Astron. Astrophys.* **365** (2001) L27.
15. P. Boschler, *Physica Scripta* **T18** (1987) 55.
16. K. Lodders *Astrophys. J.* **591** (2003) 1220.
17. J. W. den Herder, A. C. Brinkman, S. M. Kahn and 35 co-authors, *Astron. Astrophys.* **365** (2001) L7.

X-RAY EMISSION FROM JUPITER, SATURN, AND EARTH: A SHORT REVIEW

ANIL BHARDWAJ
Space Physics Laboratory, Vikram Sarabhai Space Center
Trivandrum 695022, India
anil_bhardwaj@vssc.org

Jupiter, Saturn, and Earth — the three planets having dense atmosphere and a well developed magnetosphere — are known to emit X-rays. Recently, Chandra X-ray observatory has observed X-rays from these planets, and XMM-Newton has observed them from Jupiter and Saturn. These observations have provided improved morphological, temporal, and spectral characteristics of X-rays from these planets. Both auroral and non-auroral (low-latitude) "disk" X-ray emissions have been observed on Earth and Jupiter. X-rays have been detected from Saturn's disk, but no convincing evidence for X-ray aurora on Saturn has been observed. The non-auroral disk X-ray emissions from Jupiter, Saturn, and Earth, are mostly produced due to scattering of solar X-rays. X-ray aurora on Earth is mainly generated via bremsstrahlung from precipitating electrons and on Jupiter via charge exchange of highly ionized energetic heavy ions precipitating into the polar atmosphere. Recent unpublished work suggests that at higher (>2 keV) energies electron bremsstrahlung also plays a role in Jupiter's X-ray aurora. This paper summarizes the recent results of X-ray observations on Jupiter, Saturn, and Earth mainly in the soft energy (∼0.1–2.0 keV) band and provides a comparative overview.

1. Introduction

Terrestrial X-rays were discovered in the 1950s. Launch of the first X-ray satellite UHURU in 1970 marked the beginning of satellite-based X-ray astronomy. After about two decades of search with balloon-, rocket-, and satellite-based experiments,[1] X-ray emission from Jupiter was discovered with the Einstein observatory.[2] During 1990s, Rontgensatellite (ROSAT) made important contributions to planetary X-rays by discovering emissions from Moon and comet and providing better observations on X-rays from Jupiter. With the advent of sophisticated X-ray observatories, viz., Chandra and XMM-Newton, the field of planetary X-ray astronomy is advancing at a faster pace. Several new solar system objects are now known to shine in X-rays at energies generally below 2 keV.[3–5] These include Venus, Mars, Saturn, Galilean moons Io and Europa, Io plasma torus, rings of Saturn,

Table 1. Characteristics of Jupiter, Saturn, and Earth.

Parameter	Earth	Jupiter	Saturn
Distance from Sun (AU)	1	5.2	9.5
Equatorial radius (Earth = 1)	1	11.2	9.5
Rotation period (Earth = 1)	1	0.415	0.445
Inclination (°)	23.5	3.1	26.7
Main atmospheric species	N_2, O_2, O	H_2, H, He	H_2, H, He
Magnetic field (G)	0.31	4.28	0.22
Magnetic moment (Earth = 1)	1	20,000	600
Dipole tilt wrt rotation axis	+11.3°	−9.6°	−0.0°
Magnetosphere size ($R_{\rm planet}$)	6–12 $R_{\rm E}$	50–100 $R_{\rm J}$	16–22 $R_{\rm S}$
Auroral input power (Earth = 1)	1	10^3–10^4	10–100
Energy source(s)	Solar wind	Rotation	Solar wind + rotation
Magnetospheric plasma source(s)	Ionosphere, solar wind	Io, Galilean satellites	Satellites, rings, ionosphere

Notes: 1 AU (astronomical distance) = 1.496×10^{13} cm, Earth equatorial radius ($R_{\rm E}$) = 6,378 km. Average auroral input power at Earth ∼1–100 × 10^9 W.

and Earth and Martian exospheres. Higher spatial and spectral resolution information on planetary X-rays is improving our understanding on the physics of the X-ray production on the planetary bodies, which are much colder than traditional million-degree K or higher temperature plasmas in the solar corona and astrophysical objects.[6]

In this paper, we summarize the recent results of soft (∼0.1–2.0 keV) X-ray observations on Jupiter, Saturn, and Earth: all the three planets having dense atmospheres and intrinsic magnetospheres, and known to emit X-rays. Table 1 provides some of the characteristic parameters of these planets. Reader are referred to other reviews for more details.[1,3–5,7]

2. Earth: Auroral Emissions

It is well known that the X-ray aurora on Earth is generated by energetic electron bremsstrahlung,[8–10] and the X-ray spectrum of the aurora has been very useful for studying the characteristics of energetic electron precipitation.[9,11–15] The PIXIE X-ray imager on the Polar spacecraft measured X-rays in the range 2–60 keV.[16] The high apogee of the Polar satellite (∼9 $R_{\rm E}$) enabled PIXIE to image the entire auroral oval (Fig. 1) with

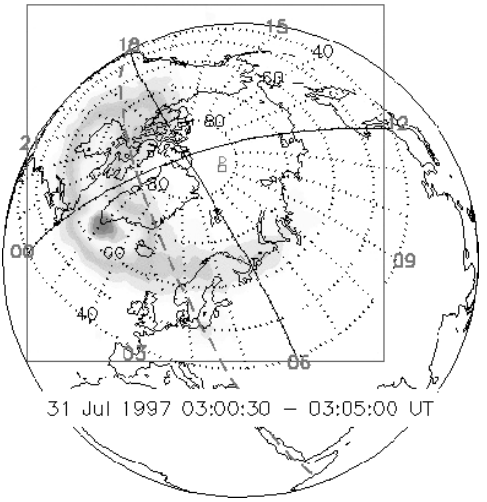


Fig. 1. Auroral X-ray image of the Earth from the Polar PIXIE instrument (energy range 2.81–9.85 keV) obtained on July 31, 1997. The box denotes the PIXIE field-of-view. Solid black lines represent the noon-midnight and dawn-dusk magnetic local time meridians, while red dashed line represents the day/night boundary at the surface. The grid in the picture is in geomagnetic coordinates and the numbers shown are magnetic local time (from PIXIE; courtesy N. Østgaard).

a spatial resolution of ∼700 km. PIXIE data showed that the substorm X-rays brighten up in the midnight sector and have a prolonged and delayed maximum in the morning sector due to the scattering of eastward-drifting electrons.[13] Statistically, the X-ray bremsstrahlung intensity is largest in the midnight substorm onset, is significant in the morning sector, and has a minimum in the early dusk sector.[10] During the onset/expansion phase of a typical substorm the electron energy deposition power is about 60–90 GW, which produces around 10–30 MW of bremsstrahlung X-rays.[17]

While harder X-ray emissions from electron bremsstrahlung are well known in the terrestrial aurora,[9–15] surprisingly, there were no searches for emissions at auroral latitudes at energies <2 keV until recently. Northern auroral regions of Earth were imaged using the high-resolution camera (HRC-I) of the Chandra X-ray observatory at ten epochs (each ∼20 min duration) between mid-December 2003 and mid-April 2004,[18] to search for Earth's soft (<2 keV) X-ray aurora. The first Chandra soft X-ray observations of Earth's aurora showed that it is highly variable — sometimes intense arcs (Fig. 2), other times multiple arcs, or diffuse patches, and at times absent.[18] In at least one of the observations an isolated blob of

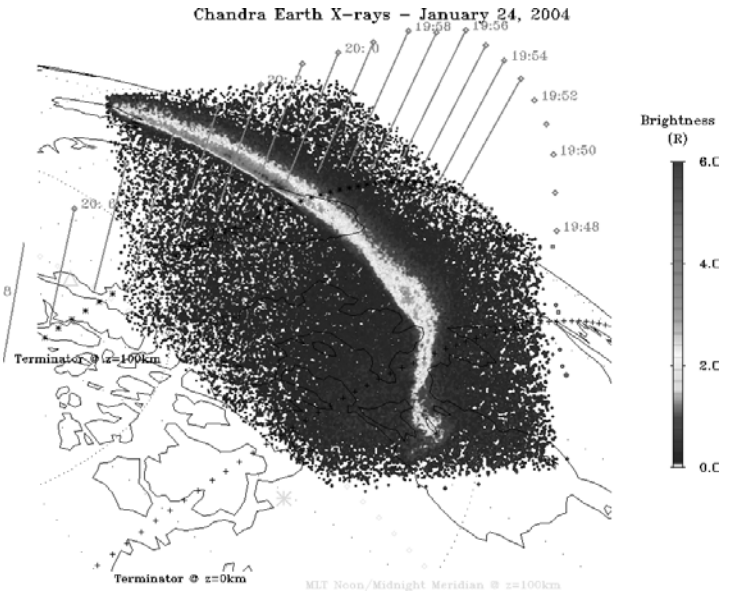


Fig. 2. Chandra HRC-I X-ray image of auroral region on January 24, 2004 showing a bright arc. The orbital location of satellite DMSP F13 is shown by diamonds, with 2-minute time ticks and vertical lines extending down to an altitude of 100 km [see Ref. 18 for details].

emission is observed near the expected cusp location. A fortuitous overflight of DMSP satellite F13 provided SSJ/4 energetic particle measurements above a bright arc seen by Chandra on January 24, 2004, 20:01–20:22 UT. A model of the emissions expected strongly suggests that the observed soft X-ray signal is produced by electron bremsstrahlung.[18]

3. Earth: Non-Auroral Disk Emissions

The non-auroral X-ray background above 2 keV from the Earth is almost completely negligible except for brief periods during major solar flares.[10] However, at energies below 2 keV soft X-rays from the sunlit Earth's atmosphere have been observed even during quite (non-flaring) Sun conditions.[19,20]

The two primary mechanisms for the production of X-rays from the sunlit atmosphere are: (1) the Thomson (coherent) scattering of solar X-rays from the electrons in the atomic and molecular constituents of the atmosphere and (2) the absorption of incident solar X-rays followed

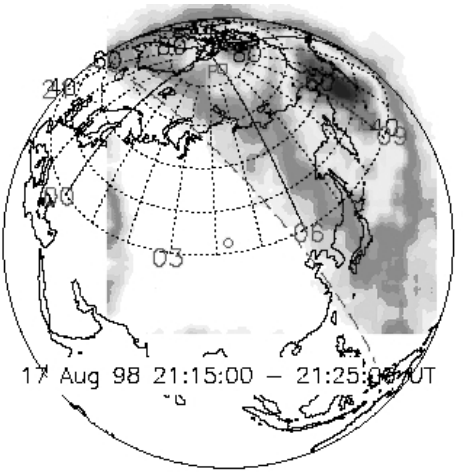


Fig. 3. X-ray image of Earth on August 17, 1998 from Polar PIXIE instrument (energy 2.9–10.1 keV), showing the dayside X-rays during a solar X-ray flare. The grid in the picture is in geomagnetic coordinates, and the numbers shown are magnetic local time. The day–night terminator at the surface of the Earth is shown as a dashed line.

by the emission of characteristic K-shell lines of Nitrogen, Oxygen, and Argon. Figure 3 shows the PIXIE image of Earth demonstrating the X-rays (2.9–10 keV) production in the sunlit atmosphere during a solar flare of August 17, 1998. The X-ray brightness can be comparable to that of a moderate aurora. For two solar flare events during 1998 examined using the data from PIXIE, the shape of the measured X-ray spectra was in fairly good agreement with modeled spectra of solar X-rays scattered and fluoresced in the Earth's atmosphere.[10]

4. Jupiter: Auroral Emissions

Auroral X-rays from Jupiter were first detected by Einstein observatory in 1979,[2] and later studied by ROSAT observations.[21,22] The pre-Chandra understanding of Jovian auroral X-rays was that these emissions are mostly line emissions resulting from recombination and charge exchange transitions in high-charged states of S and O ions precipitating from inner ($\sim$8–12 R_J) region of the magnetosphere.[1,3,21–24]

The Chandra observations of Jupiter in December 2000[25] and February 2003[26] have revealed that: (1) most of Jupiter's northern auroral X-rays come from a "hot spot" that is fixed in latitude and longitude and located

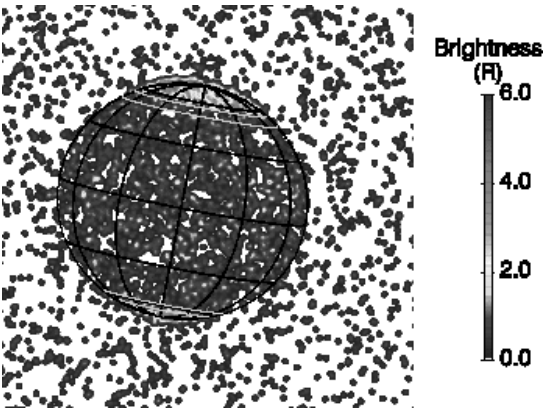


Fig. 4. Chandra X-ray image of Jupiter on December 18, 2000 generated from 10 h of continuous observations. A jovicentric graticule with 30° intervals is overplotted, along with the $L = 5.9$ and $L = 30$ footprints of the magnetic field model. The image shows strong auroral X-ray emissions from high latitudes and rather uniform emissions from the disk [see Ref. 25 for details].

significantly poleward ($>30\,R_J$) of the latitudes connected to the inner magnetosphere (Figs. 4 and 8) and (2) the auroral hot spot X-rays pulsate with periodicity that is regular ($\sim$45 min) at time[25] and chaotic at other times[26] (vary over the 20–70 min range, cf. Fig. 5). Chandra observations also found (Fig. 5) that X-rays from the north and south auroral regions are neither in phase nor in anti-phase, but that the peaks in the south are shifted from those in the north by about 120° (i.e., one-third of a planetary rotation).[26] Periodic oscillations on time scale of 20–70 min are not observed in the XMM-Newton data,[27,28] perhaps due to lower spatial resolution of XMM-Newton relative to the Chandra.

The Chandra-ACIS[26] and XMM-Newton[27–29] observations have provided soft X-ray spectra from the Jovian aurora, which consist of line emissions that are consistent with high-charge states of precipitating heavy (C, O, S) ions, and not a continuum as might be expected from electron bremsstrahlung (see Figs. 6 and 9(a,b)).

XMM-Newton has provided spectral information on the X-rays from Jupiter, which is somewhat better than Chandra. The RGS on XMM-Newton clearly resolves the strongest lines in the spectra, while the EPIC camera has provided images of the planets in the strong OVII and OVIII lines present in the Jovian auroral emissions.[28,29]

The spectral interpretation of Chandra and XMM-Newton observations is consistent with a source due to energetic ion precipitation that undergoes

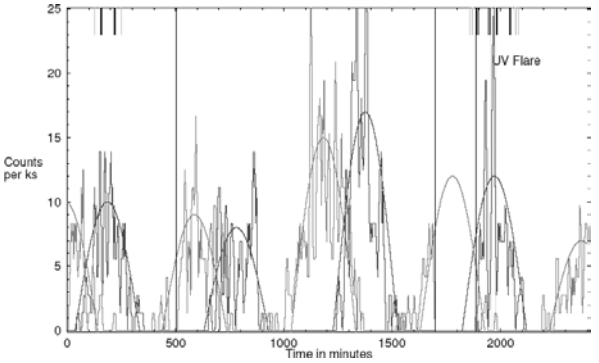


Fig. 5. Chandra observed X-ray count rate for the north (dark) and south (light) auroral zones of Jupiter. The time origin corresponds to 1558:06 UT on February 24, 2003. The black vertical lines mark the transitions from ACIS-S to HRC-I exposures and back to ACIS-S. The bars at the top mark the simultaneous HST observations in different mode (see Ref. 26 for details). Note that the set of exposures containing the UV flare coincides with the tallest peak in the ACIS-S light curve for the northern auroral zone. Smooth sections of sine waves provide crude representations of projected area effects arising from the planet's rotation.

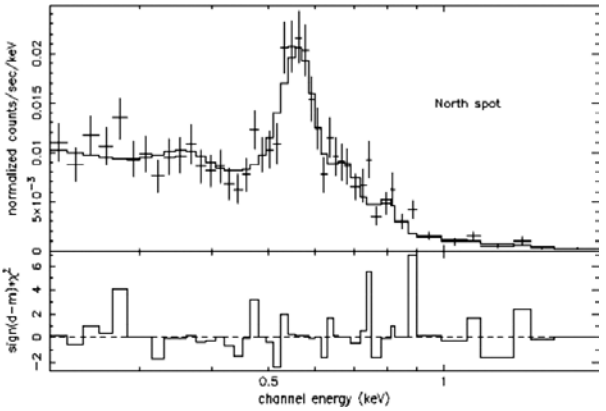


Fig. 6. XMM-Newton EPIC spectrum of Jupiter's northern auroral spot.[27] Best-fit (solid line) for the data (symbol) is obtained with five narrow Gaussian lines and a power law continuum. Bottom panel shows the chi-square of the residual (see Ref. 27 for details).

acceleration to attain energies of >1 MeV/nucleon before impacting the Jovian upper atmosphere.[26–30] However, the source of precipitating ions — whether it is outer magnetospheric or solar wind origin, or a mixture of both, is currently not clear and arguments in favor of either of them have been presented.[26–30]

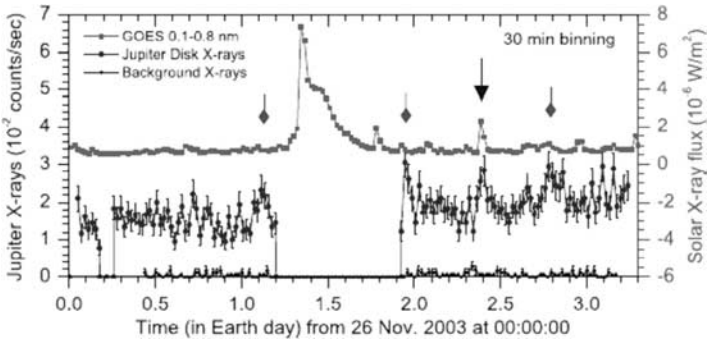


Fig. 7. Comparison of 30-minute binned Jupiter disk X-rays with GOES 10 0.1–0.8 nm solar X-ray data. The light curve of background X-rays is shown in black. The Jovian X-ray time is shifted by −4948 s to account for light travel time delay between Sun–Jupiter–Earth and Sun–Earth. The small gap at ∼0.2 days is due to a loss of telemetry from XMM-Newton, and the gap between 1.2 and 1.9 days is caused by the satellite perigee passage. The black arrow (at 2.4 days) refers to the time of the largest solar flare visible from both, Earth, and Jupiter, during the XMM-Newton observation, which has a clear matching peak in the Jovian light curve. The green arrows represent times when the Jupiter light curve shows peaks, which we suggest correspond to solar flares that occurred on the western (Earth-hidden) side of the Sun. The phase angle (Sun–Jupiter–Earth angle) of the observations was 10.3°, the solar elongation (Sun–Earth–Jupiter angle) was between 76.7 and 79.8° during the observation [see Ref. 35 for details].

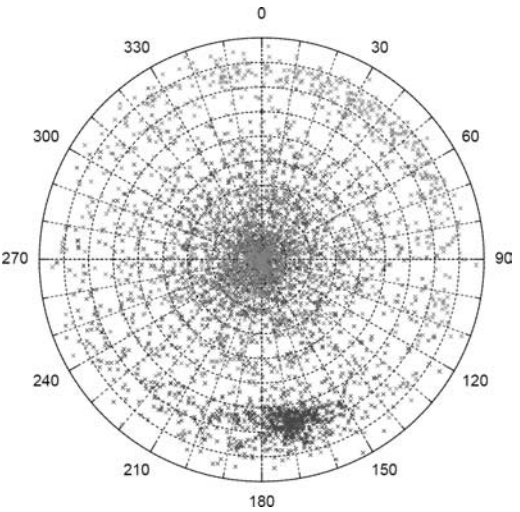


Fig. 8. Equatorial projection of X-ray photons (crosses) as seen by the Chandra ACIS-S and HRC-I instruments in the northern (dark) and southern (light) hemispheres. The concentric circles denote the latitude at intervals of 10° starting with 0° (equator) at the centre. The Jovian S_{III} longitude coordinates are labeled.

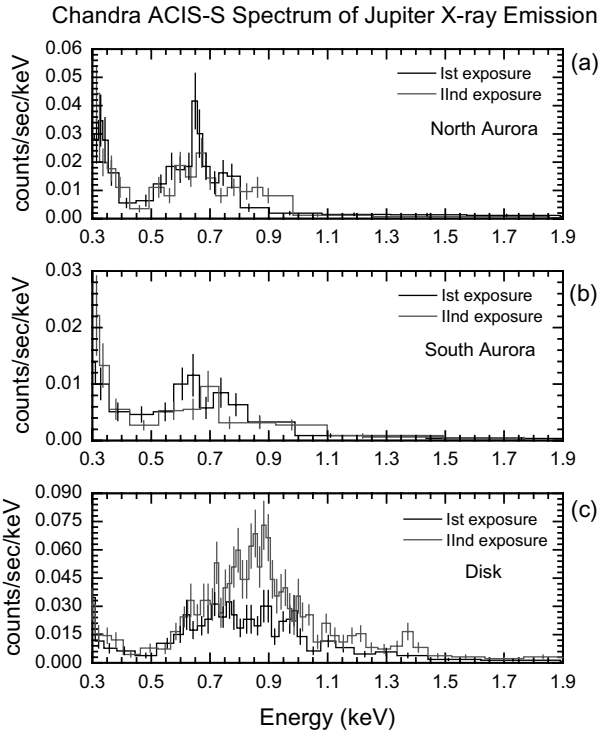


Fig. 9. Jupiter disk X-ray spectra (bottom panel, c) compared with auroral spectra from the north (top panel, a) and south (middle panel, b) observed during the two Chandra ACIS-S exposures taken on February 24 and February 25–26, 2003. Both ACIS-S exposures were of almost same duration and taken about a day apart. Each spectral point represents $\geq$10 events. The differences between the disk and auroral spectra are evident.

Very recently, XMM-Newton and Chandra data[26] have suggested that there is a higher (>2 keV) energy component present in the spectrum of Jupiter's aurora. The observed spectrum and flux, at times, tentatively appears consistent with that produced by electron bremsstrahlung[1,32] at energies greater than 2 keV.

5. Jupiter: Low-Latitude Disk Emissions

X-ray emission from Jupiter's low latitudes was first reported using the ROSAT-HRI.[33] It was proposed that disk X-rays may be largely due to the precipitation of energetic sulfur or oxygen ions into the atmosphere

from Jovian inner radiation belts. Later it was suggested[34] that elastic scattering of solar X-rays by atmospheric neutrals (H_2) and fluorescent scattering of carbon K-shell X-rays from CH_4 molecules located below the Jovian homopause are also potential sources of disk X-rays.

XMM-Newton's 69 h of Jupiter observation, in November 2003, demonstrated that day-to-day variation in disk X-rays of Jupiter are synchronized with variation in the solar X-ray flux (Fig. 7), including a solar flare that has a matching feature in the Jovian disk X-ray light curve.[35]

The X-rays from the disk are quite uniformly distributed across the low-latitudes (Fig. 8) — in contrast to the auroral X-rays. Auroral X-rays from the north (60–75°N latitude) are dominantly confined to ∼150–190° longitude and those from the south (70–80°S latitude) spread almost halfway across the planet (∼300–360° and 0–120° longitude), while the disk X-rays are quite uniformly distributed and are largely confined to <50° latitude in both hemispheres.[36] The spectrum of X-rays from the disk is also harder and extends to higher energies than the auroral spectrum (Fig. 9). No periodicity has been observed in disk X-ray lightcurve.[26,36,37]

Recent studies suggested that the X-ray emission from the Jovian disk is largely due to scattered solar X-rays and that processes occurring on the Sun control the X-rays from Jupiter's disk.[35,36–38]

6. Saturn

The X-ray emission from Saturn was unambiguously detected by XMM-Newton in October 2002[39] and by Chandra in April 2003.[40] X-rays were detected mainly from the low-latitude disk and no clear indication of auroral X-rays was observed.

Recent observation of Saturn (Fig. 10) by Chandra in January 2004 showed that X-rays from Saturn are highly variable — a factor of 2–4 variability in brightness in a week's time.[41] In these observations an X-ray flare has been detected from the non-auroral disk of Saturn, which is seen in direct response to an M6-class flare emanating from a sunspot that was clearly visible from both Saturn and Earth (Fig. 11). This is the first direct evidence suggesting that Saturn's disk X-ray emission is principally controlled by processes happening on the Sun.[41] Also a good correlation has been observed between Saturn X-rays and F10.7 solar activity index. The spectrum of X-rays from Saturn disk is very similar to that from the disk of Jupiter (Fig. 12).

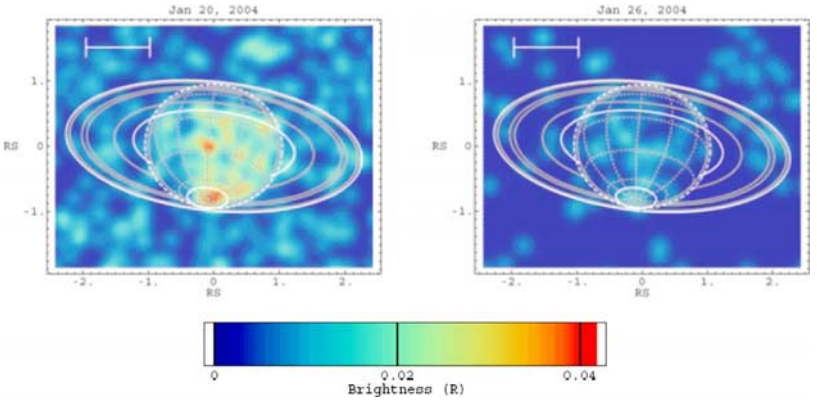


Fig. 10. Chandra ACIS X-ray 0.24–2.0 keV images of Saturn on January 20 and 26, 2004. Each continuous observation lasted for one full Saturn rotation. The white scale bar in the upper left of each panel represents $10''$. The superposed graticule shows latitude and longitude lines at intervals of $30''$. The solid gray lines are the outlines of the planet and rings, with the outer and inner edges of the ring system shown in white. The dotted white line defines the region within which events were accepted as part of Saturn's disk unless obscured by the rings. The white oval around the south pole defines the polar cap region [see Ref. 41 for details].

The Chandra observations in January 2004 also revealed X-rays from Saturn's south polar cap on January 20 (see Fig. 10, left panel). However, the analysis suggest[41] that X-ray emissions from the south polar cap region on Saturn are unlikely to be auroral in nature; they might instead be an extension of its disk X-ray emission.

7. Discussion

Table 2 presents a summary of the main characteristics of X-rays from the three planets. X-rays from the low-latitude (non-auroral) disk of all the three planets are mostly produced by scattering of solar X-rays by atmospheric species. On Jupiter and Saturn the scattering is dominantly resonant scattering with minor ($\sim$<10%) contribution from fluorescent scattering.[34,38] However, not all the incident solar X-rays in the $\sim$0.2–2.0 keV are scattered back. The energy-average geometric X-ray albedo of Jupiter and Saturn over this energy range is $\sim 5 \times 10^{-4}$ (Refs. 35, 41). At Jupiter precipitation of radiation belt ions can also make some contribution to the disk X-rays.[33]

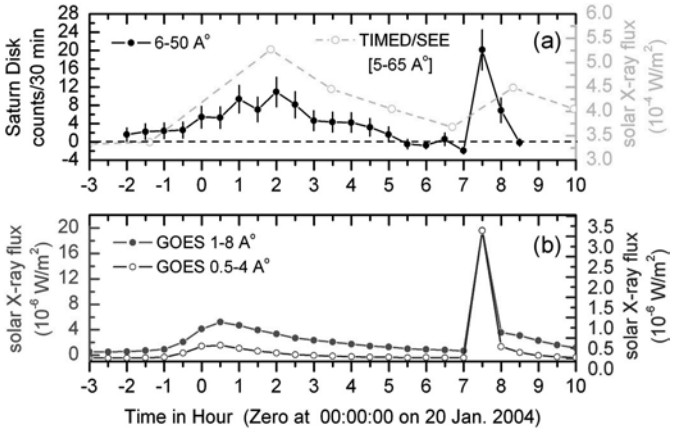


Fig. 11. Light curve of X-rays from Saturn and the Sun on January 20, 2004. All data are binned in 30 min increments, except for the TIMED/SEE data, which are 3 min observation-averaged fluxes obtained every orbit (∼12 measurements per day). (a) Background-subtracted low-latitude (non-auroral) Saturn disk X-rays (0.24–2.0 keV) observed by Chandra ACIS, plotted in black (after shifting by −2.236 h to account for the light-travel time difference between Sun–Saturn–Earth and Sun–Earth). The solar 0.2–2.5 keV fluxes measured by TIMED/SEE are denoted by open circles and are joined by the dashed line for visualization purpose. (b) Solar X-ray flux in the 1.6–12.4 and 3.1–24.8 keV bands measured by the Earth-orbiting GOES-12 satellite. A sharp peak in the light curve of Saturn's disk X-ray flux — an X-ray flare — is observed at about 7.5 h, which corresponds in time and magnitude with an X-ray solar flare. In addition, the temporal variation in Saturn's disk X-ray flux during the time period prior to the flare is similar to that seen in the solar X-ray flux [see Ref. 41 for details].

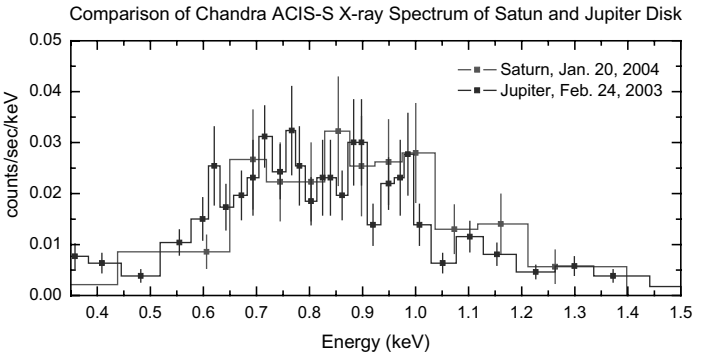


Fig. 12. Disk X-ray spectrum of Jupiter and Saturn. Values for Saturn spectrum are plotted after multiplying by a factor of 5.

Table 2. Characteristics of X-ray emissions of Jupiter, Saturn, and Earth.

Planet	Emitting region	Emitted power[a]	Special characteristics	Major production mechanism
Earth	Auroral atmosphere	10–30 MW	Correlated with magnetic storm and substorm activity	Bremsstrahlung from precipitating electrons
	Non-auroral atmosphere	40 MW	Correlated with solar X-ray flux	Scattering of solar X-rays
Jupiter	Auroral atmosphere	0.3–1 GW	Pulsating (∼30–60 min) X-ray hot spot in north; in south emitted from a band ∼180° wide in longitude	Ion precipitation (outer magnetosphere and/or solar wind) + electron bremsstrahlung
	Non-auroral atmosphere	0.3–2 GW	Relatively uniform over the disk, correlated with solar X-rays	Scattering of solar X-rays + ring current ion precipitation (?)
Saturn	Auroral and non-auroral atmosphere	0.1–0.4 GW	Correlated with solar X-ray flux	Scattering of solar X-rays + electron bremsstrahlung (?)

[a]The values quoted are "typical" values at the time of observation. X-rays from all bodies are expected to vary with time. For comparison the total X-ray luminosity from the Sun is 10^{20} W.

It has been suggested that the upper atmospheres of the giant planets Saturn and Jupiter act as "diffuse mirrors" that backscatter solar X-rays. Thus, these planets might be used as potential remote-sensing tools to monitor X-ray flaring on portions of the hemisphere of the Sun facing away from near-Earth space weather satellites.[35,38,41]

The X-ray aurora on Earth is generated by energetic electron bremsstrahlung.[8–10] The auroral X-rays from Jupiter are produced by charge-exchange of highly-ionized energetic heavy ions precipitating from the outer magnetosphere and/or solar wind.[1,25–30] At higher energies (>2.0 keV) the auroral X-rays at Jupiter[31] could be produced by electron bremsstrahlung process. However, at lower (∼<2.0 keV) energies electron bremsstrahlung falls short by orders of magnitude in explaining the Jupiter auroral X-ray flux. Also the spectrum shape at lower energies is inconsistent with the bremsstrahlung shape (see Figs. 6 and 9).[26,37] At Saturn there is no clear indication of an X-ray aurora.[40,41] X-ray aurora produced by electron bremsstrahlung is expected at Saturn, but it will probably be weak and could escape detection by present-day instruments, because Saturn aurora is relatively weaker than that on Jupiter (see Table 1), and Saturn does not

have copious heavy ion source, like Io on Jupiter. Recently, XMM-Newton has observed Saturn, for two planet rotations, in April and November 2005; the data are being analyzed.

In addition to X-rays from the planet itself, in the Jupiter system the X-ray emission has been observed from the Io plasma torus and from Galilean satellites Io and Europa,[42] while in the Saturn system X-rays have been detected from the rings of Saturn.[43]

Acknowledgments

I acknowledge the fruitful collaboration with Randy Gladstone, Ron Elsner, Graziella Branduardi-Raymont, Hunter Waite, Tom Cravens, and Nikolai Østgaard on several studies reported in this paper. I also acknowledge productive collaboration with several other colleagues in the field of planetary X-rays some of whom are listed as authors or co-authors in the references mentioned in this paper. A part of this work was supported by the NRC Senior Research Associateship at the NASA Marshall Space Flight Center.

References

1. A. Bhardwaj and G. R. Gladstone, *Rev. Geophys.* **38** (2000) 295.
2. A. E. Metzger, *J. Geophys. Res.* **88** (1983) 7731.
3. A. Bhardwaj, G. R. Gladstone, R. F. Elsner, J. H. Waite Jr., D. Grodent, T. E. Cravens, R. R. Howell, A. E. Metzger, N. Ostgaard, A. N. Maurellis *et al.*, *Eur. Space Agency Spec. Publ.* **ESA-SP-514** (2002) 215.
4. A. Bhardwaj *et al.*, *Planet. Space Sci.*, submitted (2005).
5. T. E. Cravens, *Adv. Space Res.* **26** (2000) 1443.
6. M. Gudel, *Astron. Astrophys. Rev.* **12** (2004) 71.
7. A. Bhardwaj, *Bull. Astron. Soc. India* **31** (2003) 159.
8. M. J. Berger and S. M. Seltzer, *J. Atmos. Terr. Phys.* **34** (1972) 85.
9. J. Stadsnes, K. Aarsnes and J. Bjordal, *Adv. Space Res.* **20**, 4/5 (1997) 1043.
10. S. M. Petrinec, D. L. McKenzie, W. L. Imhof, J. Mobilia and D. Chenette, *J. Atmos. Solar Terr. Phys.* **62** (2000) 875.
11. K. K. Vij *et al.*, *J. Geophys. Res.* **80** (1975) 2869.
12. J. R. Sharber *et al.*, *Geophys. Res. Lett.* **20** (1993) 1319.
13. N. Østgaard, J. Stadsnes, J. Bjordal, R. R. Vondrak, S. A. Cummer, D. L. Chenette, G. K. Parks, M. J. Brittnacher and D. L. McKenzie, *J. Geophys. Res.* **104** (1999) 10191.
14. P. C. Anderson, D. L. McKenzie, M. J. Brittnacher, M. W. Chen, M. Hairston and M. F. Thomsen, *J. Geophys. Res.* **105** (2000) 15757.

15. N. Østgaard, J. Stadsnes, J. Bjordal, G. A. Germany, R. R. Vondrak, G. K. Parks, S. A. Cummer, D. L. Chenette and J. G. Pronko, *J. Geophys. Res.* **106** (2001) 26081.
16. W. L. Imhof *et al.*, *Space Sci. Rev.* **71** (1995) 385.
17. N. Østgaard, G. A. Germany, J. Stadsnes and R. R. Vondrak, *J. Geophys. Res.* **107** (2002) 1233.
18. A. Bhardwaj, G. R. Gladstone, R. F. Elsner, N. Østgaard, J. H. Waite Jr., T. E. Cravens, S.-W. Chang, T. Majeed and A. E. Metzger, *J. Atmos. Solar-Terr. Phys.*, in press (2005).
19. D. L. McKenzie, H. R. Rugge and P. A. Charles, *J. Atmos. Terr. Phys.* **44** (1982) 499.
20. S. L. Snowden and M. J. Freyberg, *Astrophys. J.* **404** (1993) 403.
21. J. H. Waite Jr., F. Bagenal, F. Seward, C. Na, G. R. Gladstone, T. E. Cravens, K. C. Hurley, J. T. Clarke, R. Elsner and S. A. Stern, *J. Geophys. Res.* **99**, 14 (1994) 799.
22. G. R. Gladstone, J. H. Waite Jr. and W. S. Lewis, *J. Geophys. Res.* **103** (1998) 20083.
23. T. E. Cravens, E. Howell, J. H. Waite Jr. and G. R. Gladstone, *J. Geophys. Res.* **100**, 17 (1995) 153.
24. V. Kharchenko, W. Liu and A. Dalgarno, *J. Geophys. Res.* **103**, 26 (1998) 687.
25. G. R. Gladstone, J. H. Waite Jr., D. Grodent, W. S. Lewis, F. J. Crary, R. F. Elsner, M. C. Weisskopf, T. Majeed, J.-M. Jahn, A. Bhardwaj, J. T. Clarke, D. T. Young, M. K. Dougherty, S. A. Espinosa and T. E. Cravens, *Nature* **415** (2002) 1000.
26. R. F. Elsner, N. Lugaz, J. H. Waite Jr., T. E. Cravens, G. R. Gladstone, P. Ford, D. Grodent, A. Bhardwaj, R. J. MacDowall, M. D. Desch and T. Majeed, *J. Geophys. Res.* **110** (2005) A01206.
27. G. Branduardi-Raymont, R. F. Elsner, G. R. Gladstone, G. Ramsay, P. Rodriguez, R. Soria and J. H. Waite Jr., *Astron. Astrophys.* **424** (2004) 331.
28. G. Branduardi-Raymont, A. Bhardwaj, R. Elsner, G. R. Gladstone, G. Ramsay, P. Rodriguez, R. Soria, J. H. Waite Jr. and T. E. Cravens, *ESA-SP*, in press (2005).
29. G. Branduardi-Raymont, A. Bhardwaj, R. Elsner, G. R. Gladstone, G. Ramsay, P. Rodriguez, R. Soria, J. H. Waite Jr. and T. E. Cravens, *Advances in Geosciences 2006*, in press (2005).
30. T. E. Cravens, J. H. Waite, T. I. Gombosi, N. Lugaz, G. R. Gladstone, B. H. Mauk and R. J. MacDowall, *J. Geophys. Res.* **108** (2003) 1465.
31. G. Branduardi-Raymont and R. F. Elsner, private communication (2005).
32. R. P. Singhal, S. C. Chakravarty, A. Bhardwaj and B. Prasad, *J. Geophys. Res.* **97** (1992) 18245.
33. J. H. Waite Jr., G. R. Gladstone, W. S. Lewis, P. Drossart, T. E. Cravens, A. N. Maurellis, B. H. Mauk and S. Miller, *Science* **276** (1997) 104.

34. A. N. Maurellis, T. E. Cravens, G. R. Gladstone, J. H. Waite Jr. and L. Acton, *Geophys. Res. Lett.* **27** (2000) 1339.
35. A. Bhardwaj, G. Branduardi-Raymont, R. F. Elsner, G. R. Gladstone, G. Ramsay, P. Rodriguez, R. Soria, J. H. Waite Jr. and T. E. Cravens, *Geophys. Res. Lett.* **32** (2005) L03S08.
36. A. Bhardwaj, R. F. Elsner, G. R. Gladstone, T. E. Cravens, J. H. Waite Jr., G. Branduardi-Raymont and P. Ford, *Bull. Am. Astron. Soc.* **36**, 4 (2004) 18.04.
37. A. Bhardwaj *et al.*, in preparation (2005).
38. T. E. Cravens, J. Clark, A. Bhardwaj, R. F. Elsner, J. H. Waite Jr., A. N. Maurellis, G. R. Gladstone and G. Branduardi-Raymont, *J. Geophys. Res.*, submitted (2005).
39. J.-U. Ness, J. H. M. M. Schmitt and J. Robrade, *Astron. Astrophys.* **414** (2004) L49.
40. J.-U. Ness, J. H. M. M. Schmitt, S. J. Wolk, K. Dennerl and V. Burwitz, *Astron. Astrophys.* **418** (2004) 337.
41. A. Bhardwaj, R. F. Elsner, J. H. Waite Jr., G. R. Gladstone, T. E. Cravens and P. G. Ford, *Astrophys. J. Lett.* **624** (2005) L121.
42. R. F. Elsner, G. R. Gladstone, J. H. Waite Jr., F. J. Crary, R. R. Howell, R. E. Johnson, P. G. Ford, A. E. Metzger, K. C. Hurley, E. D. Feigelson, G. P. Garmire, A. Bhardwaj *et al.*, *Astrophys. J.* **572** (2002) 1077.
43. A. Bhardwaj, R. F. Elsner, J. H. Waite Jr., G. R. Gladstone, T. E. Cravens and P. G. Ford, *Astrophys. J. Lett.* **627** (2005) L73.

INSTRUMENTATION AND OBSERVATIONS OF THE X-RAY SPECTROMETER ONBOARD HAYABUSA

TATSUAKI OKADA*,†,‡, KEI SHIRAI†, YUKIO YAMAMOTO†,
TAKEHIKO ARAI†,§, KAZUNORI OGAWA†,¶,
KOZUE HOSONO†,‡ and MANABU KATO†,‡,¶

†*Institute of Space and Astronautical Science*
Japan Aerospace Exploration Agency
3-1-1 Yoshinodai, Sagamihara, Kanagawa 229-8510, Japan

‡*Department of Earth and Planetary Science, University of Tokyo*
7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan

§*Department of Space and Astronautical Science*
University of Advanced Studies
3-1-1 Yoshinodai, Sagamihara, Kanagawa 229-8510, Japan

¶*Department of Earth and Planetary Sciences*
Tokyo Institute of Technology
2-12-11 Ookayama, Meguro, Tokyo 152-8550, Japan
**okada@planeta.sci.isas.jaxa.jp*

The X-ray spectrometer (XRS) onboard Hayabusa, a Japanese asteroid explorer, has been developed to determine major elemental composition of the surface of 25143 Itokawa through remote X-ray fluorescence spectrometry, which is planned during the asteroid rendezvous phase in 2005. During the cruising phase, the XRS observed cosmic X-rays. In addition, it also observed X-rays excited by the Sun off the surface of the Earth and the Moon just before the Earth swing-by conducted in May, 2004. We present the instrumentation of the XRS and discuss the recent results of observations from the viewpoint of design concept.

1. Introduction

The X-ray spectrometer (XRS) is a remote X-ray fluorescence (XRF) spectrometer onboard Hayabusa (former known as MUSES-C) with the aim to determine major elemental composition of the surface of asteroids.[1] The instrument is based on the charge-coupled device that shows much higher energy resolution in comparison with the conventional proportional counters. A standard sample plate is mounted aboard, which allows *in situ* calibration of XRF excited by the Sun.

*Corresponding author.

The XRF is a well-established technology for both qualitative and quantitative elemental analysis in the laboratory. The difference of the space XRF is that the excitation source is not a generator but the Sun. As was proven during the Apollo 15 and 16 missions,[2] the major elemental composition of an atmosphere-free planetary surface can be determined through this method. The Near-Earth Asteroid Rendezvous mission to Asteroid 433 Eros and the SMART-1 lunar orbiter mission also have the XRS.[3,4] The instrument has been used for many previous missions, and will be also used for the coming lunar orbiter missions such as the Japanese SELENE,[5] the Indian Chandrayaan-1, and the Chinese Chang'E as well as the Mercury missions of the NASA's MESSENGER[6] and the ESA-Japan's Bepi Colombo.[7]

In this paper, we present the objectives and specifications of the XRS instrument onboard Hayabusa. The method of onboard analysis and some results of pre-flight and in-flight experiments are also described.

2. Overviews of the Hayabusa Mission

Hayabusa[8] has been launched on May 9, 2003, by the fifth M–V launch vehicle. After one-year cruise by using the ion engine system, the spacecraft has swung by the Earth for acceleration with terrestrial gravity assist. In September of 2005, it will arrive at 25143 Itokawa, which is a near-Earth asteroid of elongated shape (about $540\,\mathrm{m} \times 290\,\mathrm{m} \times 210\,\mathrm{m}$) and classified as S(IV) type. Hayabusa will characterize the asteroid through remote sensing, followed by sample collections of the uppermost surface material. These samples will be recovered by direct re-entry from the interplanetary orbit in June, 2007.

The mission aims at the technology demonstration of sample return from a near-Earth object. But it also has high scientific significance, since several remote sensing instruments including the XRS and the robotic lander for characterization of the asteroid are mounted on it, as well as instrumentation for the very first rocky sample return from a solar system body other than the Moon.

3. Science and Instrument of the XRS

3.1. *Scientific objectives*

The XRS is to determine major elemental composition of Itokawa such as Mg, Al, Si, and S. Heavier elements such as Ca, Ti, and Fe will be also

measured when the Sun is active. Since Hayabusa hovers at the Home Position (HP) several kilometers sunward (or earthward) the asteroid, apparent size of Itokawa becomes just in the fields of view of the XRS. With long-term observation, the XRS will map the surface in the longitudinal direction by asteroid rotation. Composition at the rotation phase especially when the enhanced solar activity occurs will be accurately determined for larger numbers of elements. That of local areas can be measured from the lower altitude during the descent for sample collection. For its small solar phase angle less than 10°, the HP is favored for spectrometry but not suitable for geomorphology or detailed shape modeling. For this reason, Hayabusa takes a few weeks' trip to the zones at larger solar phase angle (XO). Dependency of the XRF intensities on the solar phase angle can be compared, which informs the surface microscopic roughness by particle size effect[9].

3.2. *Design concept of the XRS*

The configuration [or schematic diagram] of the XRS is shown in Fig. 1. The specifications of the XRS are summarized and tabulated in Table 1. Much improvement is achieved for the precision of the XRS observation, relative to the conventional proportional counters. As the X-ray detector we adopt Hamamatu's Deep-1 type two-dimensional CCD[10] of 1-inch square because energy resolution (better than 150 eV at 5.9 keV) is required high enough to discriminate each XRF peak: moreover, a relatively large detection area of 25 cm^2 is available with four CCDs arrayed. A mesh-supported ultra-thin beryllium window of 5 μm thick has been developed to achieve complete light shield and high transparency of soft X-rays. With its large detection area, the XRS can obtain statistically significant counts in several minutes, especially at the energy of 1–3 keV. A thin latticed collimator made of phosphor blonde is developed to limit the fields of view of $3.5° \times 3.5°$, just-in size of the asteroid when observed from the HP, to achieve high signal to background ratio. In addition, a standard sample plate is mounted so that elemental analysis becomes more accurate with conducting *in situ* calibration of XRF with which to compare the observation data.

There remains severe constraints in mass and power supply. The suitable operation temperature of CCD is less than minus 50° C. Since the use of coolant or mechanical cooler is not favored for this small mission, the thermal design of the instrument is such that can keep the detector cool enough with radiation cooling through a small radiator. The compact electronics is realized by using the onboard computer SH-OBC that has three

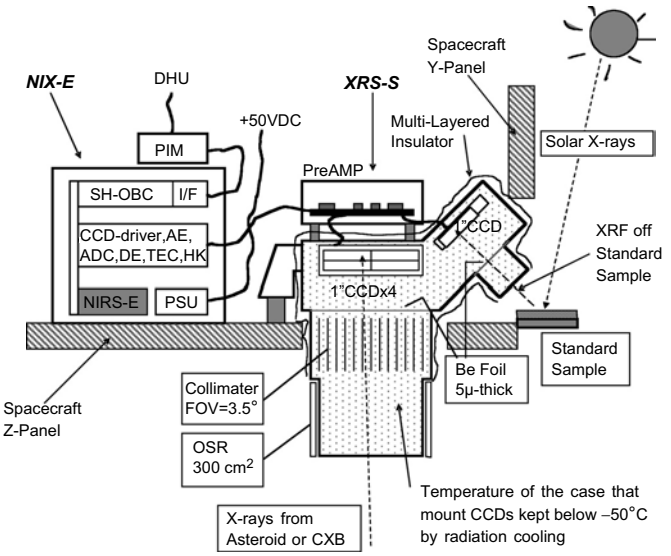


Fig. 1. A schematic diagram of the XRS is shown. The XRS is composed of the XRS-S (sensor unit) and the NIX-E (electronics). Charge-coupled devices (CCD) are used as X-ray detectors. Four chips are arrayed to observe X-rays off the asteroid or from space, and another single chip is used to calibrate XRF off the standard sample excited by the Sun. Thin beryllium windows are mounted in from of the CCD for light shield and high transparency of soft X-rays. Latticed collimator limits the FOV of $3.5 \times 3.5°$, just-in size of the asteroid when observed from the "HP". The CCD is kept cool enough by radiation cooling. Effective onboard data analysis and reduction for telemetry is conducted using the SH-OBC (onboard computer).

Table 1. Specifications of the XRS is shown.

Energy detection	0.7–10 (keV)
Energy resolution	160 (eV @Mn–K = 5.9 keV) or 120 (eV @Al–K = 1.5 keV) at −40°C
Detection area	25 cm^2
Detector	Hamamatsu Deep-1: front illuminated, 2D, 1-inch-square, FFT
Standard sample	Glassy plate of 36 mm × 41 mm.
ADC	DATEL ADCDS1403: 14 bit (12 bit effectively used)
Total mass	1.7 kg for XRS-S, 2.8 kg for NIX-E [including NIRS-E(*)]
Total power	21.2 W (max 30 W)
Time resolution	30/60/300/600 s
Operation mode	Nominal (Histogram): 0.1–0.2 kbps in average, diagnostic: image dump, event extraction raw data

*: NIRS-E is the electronics of NIRS instrument that use the CPU and power supply in common.

Hitachi SH7708 CPUs with majority voting technique to enhance radiation tolerance, and also by using flexible programmable gate arrays. A single unit of driver and readout system operates in turn each of five CCDs, four as main detector and another one as solar calibrator.

As a deep space mission, the downlink bit rate of telecommunication severely constrains the total amount of data production. Data reduction by onboard analysis as well as the programmable sequenced observation is required to operate effectively and achieve scientific goals.

4. Method of Onboard Data Analysis

4.1. *Event extraction*

The XRS requires high-data reduction since its CCD readout rate is 2 Mbytes in 8 s (125 kHz), which is almost equal to the total data allocation for the XRS in a day. However, the X-ray flux is not so high even during enhanced solar activity, and most of pixels have no X-ray derived electrons. Then only X-ray events should be effectively extracted. For that purpose, we apply line-binning of 16 pixels for simplifying the hardware logics, in which case the data of the both side pixels have only to be compared rather than those of the surrounding eight pixels. Those pixels whose value exceeds the average by commanded threshold level are regarded as X-ray events and data of 5 pixels are stored into FIFO memory for software analysis. The schematic diagram is shown in Fig. 2.

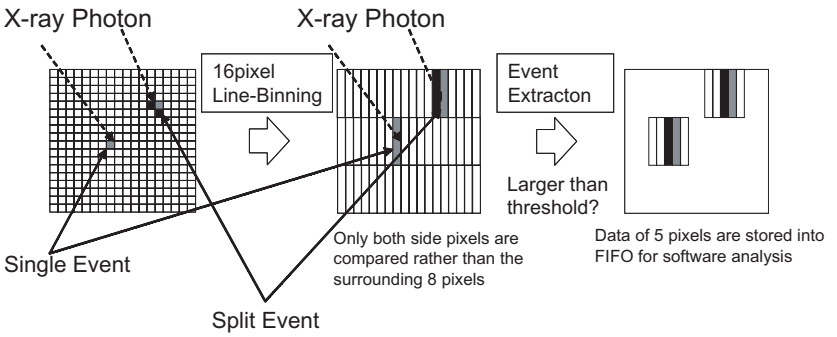


Fig. 2. A schematic diagram of the event extraction is shown. Single and split events are classified as proper grades with the onboard software analysis.

4.2. *Grade analysis*

When X-ray photons or high-energy charged particles enter into the CCD, the derived electrons do not always stay in the single pixel (as single event) but often split into neighboring pixels (as split event). The value of each pixel is not proportional to the irradiated energy so that the value should be summed over a split event. For each extracted X-ray event, data of 5 pixels stored in the FIFO memory are dealt with by the software analysis according to whether the event is single or split. Grade-0 is attributed to the single event, and grades 1–3 is to left-side split, right-side split, and both-side split, respectively. Split threshold levels to judge whether the event is single or split can be set by commands individually for each CCD.

After the classification of grades, energy histograms of 4096 energy channels with 16 bits depth are once produced in a given integration time for each CCD chip and registered in the buffer area on the SH-OBC. Furthermore, those histograms are compressed typically into less than 1/10 size and compiled in the CCSDS telemetry packet before delivery.

5. Laboratory Experiments

We have examined the XRS performance in the laboratory. Figure 3 shows an example of results for powder specimen with simple composition such as MgO, Al_2O_3, SiO_2, $CaCO_3$, TiO_2, and Fe_2O_3. A Cr-target X-ray generator is used as the primary X-ray source. The temperature of CCD is cooled at $-50°C$. In this analysis, only the grade-0 data is used. The XRF line spectra of major elements are clearly detected and discriminated each other. The scattered Cr peaks are also found for all the spectra. The energy resolution ranging 1.25–7.1 keV is almost achieved to that we have expected. Performance of the XRS[11] has been also proven for complex powder compounds as shown in Fig. 4, whose composition is simulated to a typical A12 (Apollo 12 site) basaltic soil.[12] Each XRF peak of component major elements such as Mg, Al, Si, K, Ca, Ti, and Fe and scattered Cr peak are clearly detected and discriminated each other. In addition to K-α peaks, K-β peaks are detected for Si, Ca, Ti, Cr, and Fe, which indicates sufficiently good energy resolution for quantitative elemental analysis as expected.

6. In-Flight Experiments

After Hayabusa was launched, the XRS have been operated for observation and in-flight calibration. Figure 5 shows an example of *in situ* XRF

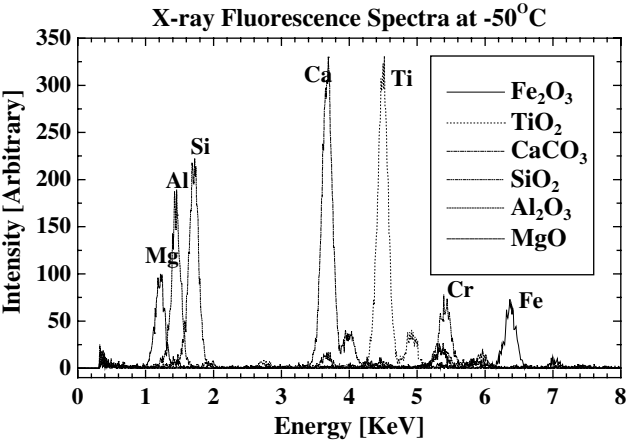


Fig. 3. The results of laboratory experiments are summarized for each specimen in simple composition. The XRF line spectra of each element are detected and clearly discriminated each other. A Cr-target X-ray generator is used as excitation source. The operation temperature of CCD is minus 50°C. In addition to each XRF peak, scattered Cr peak is observed. For this analysis, only the grade-0 data are used.

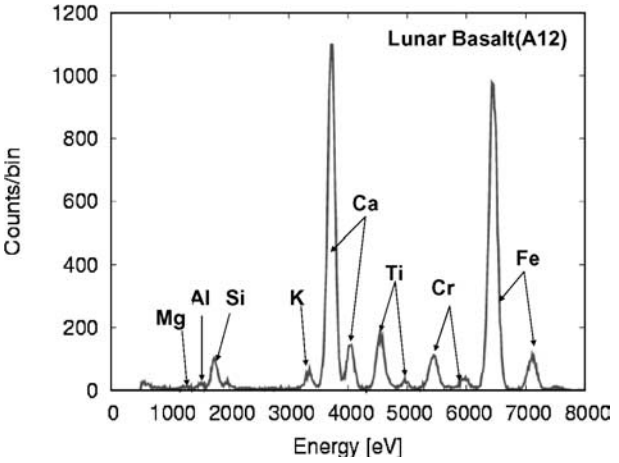


Fig. 4. A result of laboratory experiment is shown for powder stimulant in lunar basaltic composition. A Cr-target X-ray generator is used as excitation source. Grade-0 data is used for this analysis. Each XRF peak of Mg, Al, Si, K, Ca, Ti, and Fe and scattered Cr peak are clearly detected and discriminated each other. Not only K-α but K-β peaks are detected for Si, Ca, Ti, Cr, and Fe, indicating its good energy resolution of the XRS.

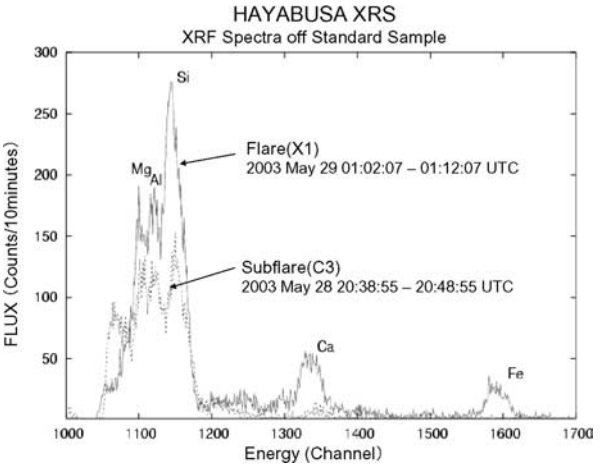


Fig. 5. In-flight observation of XRF off the standard sample during the active (X1) and moderate (C3) solar conditions are shown with clear detection of major elements.

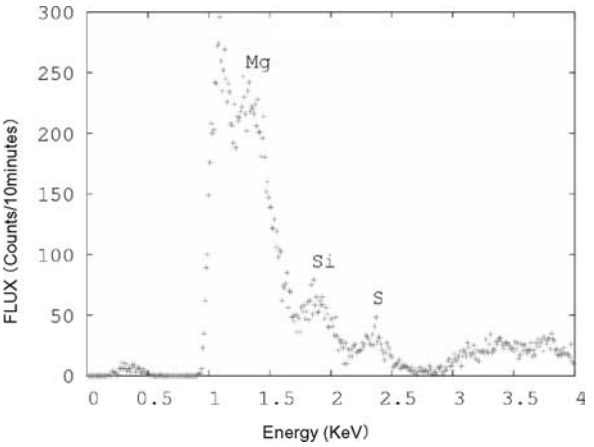


Fig. 6. Kepler's Supernova Remnant was observed on June 15, 2003. Emission lines of Mg, Si, and S are clearly detected as expected.

off the standard sample under the active (X1) and moderate (C3) solar conditions.[13] Intensities of each XRF are enhanced, especially for heavier elements. In the active case shows higher enhancement of Ca and Fe as well as larger Si/Mg that is consistent for higher temperature flare events. Figure 6 shows cosmic X-rays from Kepler's Supernova Remnant.[14] Emission line of Mg, Si, and S are clearly detected as expected.

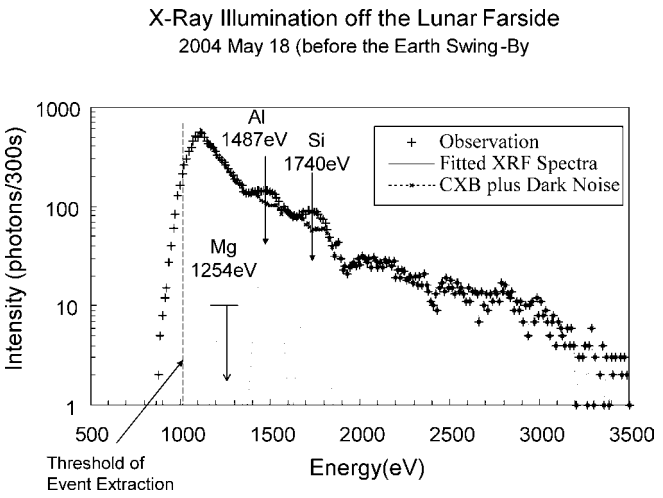


Fig. 7. First X-ray observation of the lunar farside is shown. The XRF at Al and Si are detected along with the cosmic X-ray backgrounds and dark noise. The XRF at Mg is not detected so that Al/Si and the upper limit of Mg/Si are derived.

Just before the Earth swing-by on May 17, 2004, the XRS has observed X-rays off the farside of the Moon for the very first time.[15] The apparent diameter of the Moon was 0.5° so that the X-ray spectrum was a mixture of the X-rays from the Moon with the cosmic X-ray backgrounds. The line spectra of Al and Si that are considered from the Moon can be separated from the continuous CXB but not for that of Mg, indicating high Al/Si and low Mg/Si. This result shows that the global average composition of the lunar far-side is Al-enriched and Mg-poor, most suitable to anorthosetic rather than basaltic (shown in Fig. 7).

7. Conclusions

The XRS has been developed for Hayabusa to determine major elemental composition of Asteroid 25143 Itokawa under such severe conditions as small mass and low bit-rate telemetry. In order to achieve scientific goals, new technologies have been adopted such as charge-coupled devices as X-ray detectors, an ultra-thin beryllium windows for light-shield, a latticed collimator to limit the fields of view, onboard standard sample for *in situ* calibration, the thermal design to keep CCDs cool enough through radiation cooling, and the compact electronics using SH-OBC and FPGAs. Onboard analysis including event extraction and grade classification highly

contributes to data reduction. The functions and performances have been examined in the laboratory. The XRS has proven to work well in space even outside the Earth's magnetosphere and successfully observed some cosmic X-ray bodies and X-rays off the farside of the Moon for the first time. The XRS is expected for accurate elemental analysis of Itokawa during the rendezvous phase in 2005.

Acknowledgments

The authors greatly appreciate the reviewers for their technical advices and grammatical corrections to improve our manuscript.

References

1. T. Okada, M. Kato, A. Fujimura, H. Tsunemi and S. Kitamoto, *Adv. Space Res.* **25** (2000) 1833–1837.
2. I. Adler, J. I. Trombka and P. Gorenstein, *Anal. Chem.* **44** (1972) 28a–35a.
3. L. R. Nittler, R. D. Starr, L. Lim, T. J. McCoy, T. H. Burbine *et al.*, *Meteorit. Planet. Sci.* **36** (2001) 1673–1695.
4. M. Grande, R. Browning, N. Waltham *et al.*, *Planet. Space Sci.* **51** (2003) 427–433.
5. T. Okada, M. Kato, A. Fujimura, H. Tsunemi and S. Kitamoto, *Adv. Space Res.* **23** (1999) 345–348.
6. S. Solomon, R. McNutt, Jr., R. Gold *et al.*, *Planet. Space Sci.* **49** (2001) 1445–1465.
7. R. Grard and A. Balogh, *Planet. Space Sci.* **49** (2001) 1395–1407.
8. J. Kawaguchi, H. Kuninaka, A. Fujiwara and K. T. Uesugi, in *Proc. 5th IAA Intl Conf. Low-Cost Planetary Missions* **ESA SP-542** (2003) 25–32.
9. Y. Kuwada, T. Okada and H. Mizutani, in *Proc. ISAS Lunar Planet. Symp.* eds. H. Mizutani and M. Kato, **30** (ISAS, 1997) 212–215.
10. K. Miyaguchi, H. Suzuki, J. Dezaki and K. Yamamoto, *Nucl. Instr. Meth.* **A436** (1999) 24–31.
11. Y. Yamamoto, Doctor's Thesis, University of Tokyo, 2002.
12. S. R. Taylor, *Planetary Science: A Lunar Perspective* (LPI, Houston, 1982), p. 481.
13. Y. Yamamoto, T. Arai, T. Okada, K. Shirai and M. Kato, in *Proc. ISAS Lunar Planet. Symp.* eds. H. Mizutani and M. Kato, **36** (ISAS, 2003), pp. 286–289.
14. T. Arai, Y. Yamamoto, K. Shirai, T. Okada and M. Kato, in *Proc. ISAS Lunar Planet. Symp.* eds. H. Mizutani and M. Kato, **36** (ISAS, 2003), pp. 282–285.
15. T. Okada, T. Arai, K. Hosono, K. Shirai, Y. Yamamoto, K. Ogawa and M. Kato, *Lunar Planet. Sci.* **36** (2005) 1175.

A MISSION CALLED SAPPORO

W.-H. IP[*,†], I.-G. JIANG[*], D. KINOSHITA[*], L. N. HAU[†], A. FUJIWARA[‡], Y. SAITO[‡], F. YOSHIDA[§], K. W. MIN[¶], ANIL BHARDWAJ[‖], H. BOEHNHARDT[**], P. HARTOGH[**], T. M. CAPRIA[††], G. CREMONESE[‡‡], A. MILILLO[§§], S. ORISINI[§§], D. GAUTIER[¶¶], D. JEWITT[‖‖] and T. OWEN[‖‖]

[*]*Institute of Astronomy, National Central University Chung-Li, Taiwan 32054*

[†]*Institute of Space Science, National Central University Chung-Li, Taiwan 32054*

[‡]*Institute of Space and Astronautical Science Japan Aerospace Exploration Agency, 3-1-1 Yoshinodai, Sagamihara-shi Kanagawa-ken 229-8510, Japan*

[§]*National Astronomical Observatory, Osawa, Mitaka Tokyo 181-8588, Japan*

[¶]*Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea*

[‖]*Space Physics Research Laboratory, Vikram Sarabhai Space Center ISRO, Trivandrum 695022, India*

[**]*Max-Planck-Institute for Solar System Research 37191 Katlenburg-Lindau, Germany*

[††]*Istituto Nazionale di Astrofisica — Istituto di Astrofisica Spaziale e Fisica Cosmica, Via del Fosso del Cavaliere 100, I-00133 Roma, Italy*

[‡‡]*Istituto Nazionale di Astrofisica — Osservatorio Astrronomico di Padova, Vicolo dell'Osservatorio 5, 35122, Padova, Italy*

[§§]*Istituto di Fisica dello Spazio Interplanetario, Via del Fosso del cavaliere 100 I-00133 Roma, Italy*

[¶¶]*LESIA, Observatoire de Paris, F-92195 Meudon, France*

[‖‖]*Institute for Astronomy, University of Hawaii 2680 Woodlawn Drive, Honolulu, HI 96822, USA*

A possible scenario of future collaborative efforts among the Asian, European, and American space science communities is described with scientific objectives focused on the Saturnian system. The project tentatively named Saturn Atmosphere Probe and Phoebe Rover (SAPPORO) mission could probably be realized in around 2025 if organizational steps are taken now to facilitate dialogs and scientific exchanges among the scientists on a global scale.

The best way to know God is to love many things.

- Vincent Van Gogh

1. Introduction

The brilliant scientific and engineering successes of the NASA Cassini Orbiter and of the near picture-perfect atmospheric descent of the ESA Huygens Probe to Titan have now given us new optimism on the next step toward studying of the Saturnian system in near future (Fig. 1). In doing so, we must make several assumptions, based partly on past experience and partly on reasonable projection.

- A post-Cassini-Huygens mission to Saturn — realistically speaking — probably would not take place until 20 years from now (i.e., 2025). During this time interval, the economy and technology of Asia as a whole will likely move from stride to stride with the possibility of matching those of Europe and USA. We must therefore, anticipate the participation of the space organizations in Asia in ambitious space projects with long-lead times.
- In addition to names of eminent scientists like Galileo, Cassini, and Bepi-Colombo, flagship missions at that time would probably carry the names of great cities of the world like Xian, Seoul, Miami, Delhi, Taipei, Rome, Macao, and Sapporo.

We therefore venture to call the present mission concept Saturn Atmospheric Probe and Phoebe Rover (SAPPORO) project.

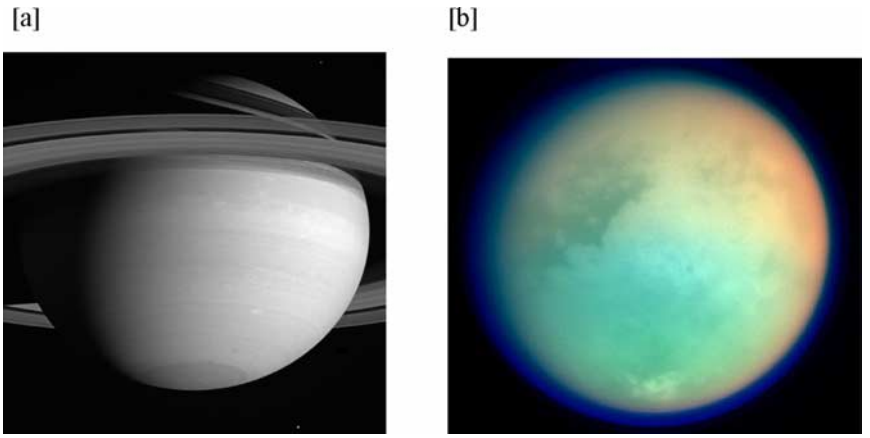


Fig. 1. The very successful Cassini–Huygens mission of NASA and ESA will path the way for the advanced investigations of (a) the Saturnian atmosphere and ring systems and (b) Titan and the icy satellites, by the next generation of planetary scientists in the world-wide community. *Courtesy*: Jet Propulsion, NASA, and Space Science Institute.

In this mission an atmospheric probe to Saturn with scientific payload similar to those on the Galileo and Huygens Probe is a must. A microwave sounder on the SAPPORO Orbiter is essential for the study of the deep interior of the Saturnian atmosphere. Depending on the power system, a ring hopper is not out of the question. Another novel scientific platform that will be the delight of the Kuiper belt community has to do with the lander on Phoebe which is the largest irregular satellite in retrograde orbit at the outskirt of the Saturnian system. The close-up observations by the remote sensing instruments on the Cassini spacecraft have shown Phoebe to be a primitive body captured from the planetesimal population feeding the accretion zone of the outer planets — which in turn might be akin to the KBOs. Chemical analysis of the sample materials on Phoebe's surface has therefore, the potential of serving as a time machine leading us back to the epoch when the solar system was originally formed.

This ensemble of mission ingredients might seem extravagant at first sight. But we must remember that most of them can be ready-made even at present days on the basis of previous and current ESA, JAXA, and NASA missions. Following the trend like a weathercock, we have very good reason to believe that the combined momentum of space technological development and the cooperative efforts of several major advancing nations in Asia will help pave the road to return to Saturn by 2025 as envisaged here. With this spirit, we hope that a preliminary assessment study could be given to this mission concept so that some preparatory inter-agency dialogs might be initiated. In Sec. 2, we will describe the scientific objectives of the SAPPORO mission. In Sec. 3, we will compare the SAPPORO mission concept to projects in planning for the exploration of outer planets. In Sec. 4, we will discuss how the concept of SAPPORO could be extended to the rendezvous missions to Uranus and Neptune of truly international in scope.

2. Scientific Objectives

From the discussions among a team of scientists, a preliminary list of scientific objectives covering a wide spectrum of important research fields in planetary science have been identified. Four of those that would answer fundamental questions are outlined here. As described below, these investigations would also require different instrumentations and platforms.

2.1. *Saturn atmospheric probe*

Figure 2 compares the scenario of the NASA Galileo probe mission to Jupiter to that of the ESA Huygens probe mission to Titan. Both successful

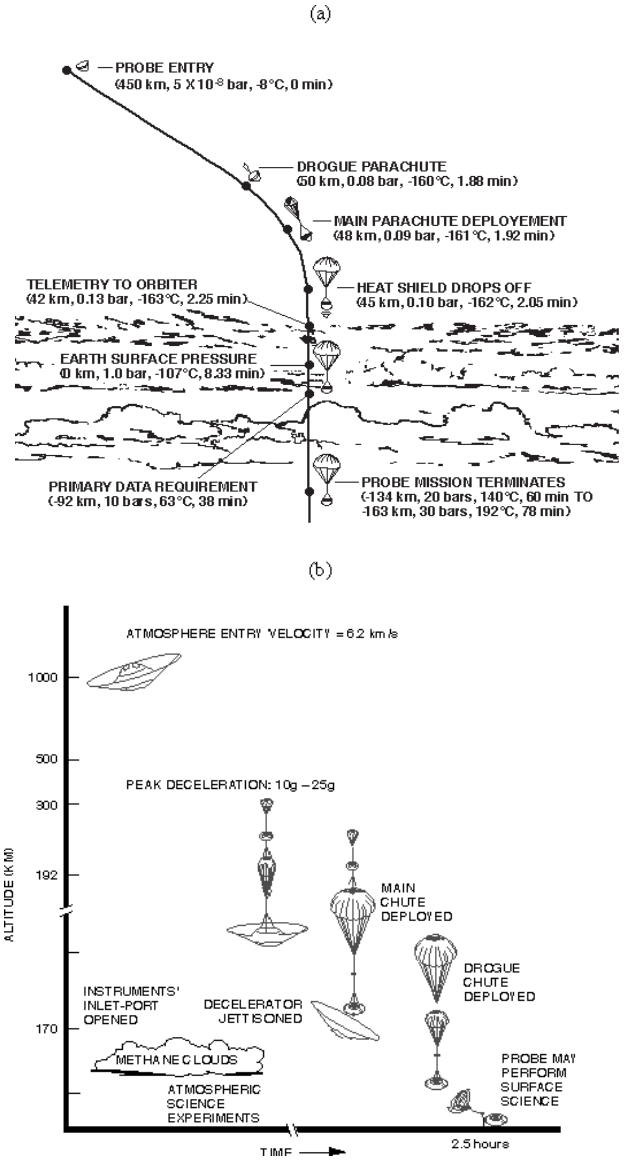


Fig. 2. The descent profiles of (a) the Galileo atmospheric probe and (b) the Huygens atmospheric probe to Titan. *Courtesy*: Jet Propulsion Laboratory, NASA (a) and European Space Agency (b).

projects are considered to be the masterpieces in space and planetary science. While the Huygens probe provided basic information on the composition and origin of Titan's atmosphere, the scientific measurements by the Galileo probe has shown that the noble gases and carbon, nitrogen, and sulfur are enriched in the Jovian atmosphere by a factor of about two to four in comparison to the solar abundances. Such near-uniform chemical enrichment has been interpreted in terms of the accretion of cold icy planetesimals by Jupiter.[1–3] On the other hand, Gautier *et al.*[4] advocated the role of the selective gas trapping effect of the clathrate hydrate icy grains in the solar nebula. Hersant *et al.*[5] have further suggested that the clathrate hydrate model could explain the non-uniform chemical enrichments in S (∼12.5), N (∼2), and C (∼5–9) in the Saturnian atmosphere. Other key measurements include the isotopic ratios of $^{14}N/^{15}N$ and $^{16}O/^{18}O$. If the Saturnian $^{14}N/^{15}N$ ratio is the same as the Jovian value, it would mean that N_2 was the dominant nitrogen reservoir in the solar nebula. This shows that *in situ* measurements of the noble gas and isotopic ratios in Saturn's upper atmosphere will be crucial in answering how the outer planets were formed. From an engineering point of view, these scientific goals can probably be achieved by sending the atmospheric probe to regions with pressure no more than 20 bars thus making the SAPPORO probe design relatively simpler than that of the Galileo Probe.

2.2. *The microwave sounder*

The Cassini mission is entering its second year of scientific operations. It is very likely that the mission will be extended beyond 2008 after the end of the prime mission. What are the important questions still left to be answered after such long-term monitoring study is difficult to assess at this point. However, if we follow the pattern of the Jovian exploration that began with the flyby observations of the Pioneers 10 and 11 spacecraft, then the Voyagers 1 and 2 spacecraft followed by the Galileo Orbiter, we would probably consider the new JUNO Project for a polar orbiter mission a most logical next step. The JUNO Orbiter will carry out particles-and-fields measurements and microwave sounding observations on polar orbits as illustrated in Fig. 3. Such orbital configurations will provide precious information on the polar Jovian magnetosphere and ionosphere so far missing from the Galileo mission. Of course, the polar trajectories of the Cassini Orbiter at the end of the primary mission were cleverly designed to study the polar regions of the Saturnian magnetosphere. But what remains to be

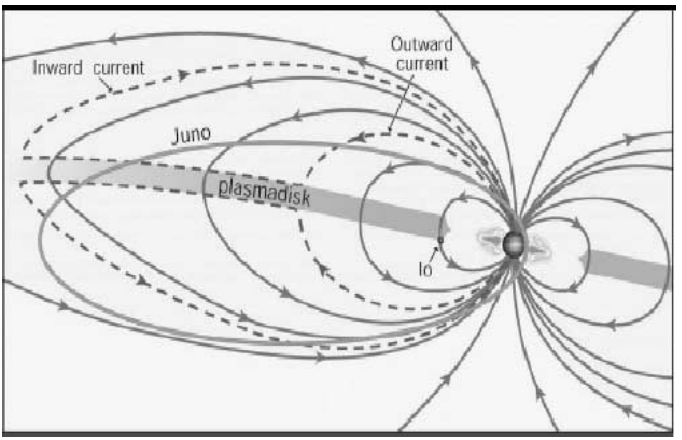


Fig. 3. A sketch of the mission profile of the JUNO Jupiter polar orbiter. *Courtesy*: Jet Propulsion Laboratory, NASA, and Southwest Research Institute.

done is probably the sounding of the deep atmosphere by microwave spectrometer as in the case of the JUNO project. This instrument will probe the temporal and spatial variations of the water and ammonia distributions and hence is the key to our understanding of the convection process and wind profiles in different regions of Saturn's atmosphere. This knowledge will in turn lead to improved understanding of the internal structure of this second largest gas planet. A comparison of the SAPPORO and JUNO measurements will be an important step for a comprehensive study of the interiors of the many extra-solar planets of Jovian masses.

The Jovian dust ring is very tenuous, especially at the top of the upper atmosphere. This is why the JUNO Orbiter can avoid dust impact by following trajectories as shown in Fig. 3. It is doubtful though whether such a free-fly zone exists at Saturn because of the existence of the D ring. Ring gaps such as the Cassini Division might not be completely empty of dust impact hazard either. A JUNO-like orbital configuration might therefore, not be viable for the SAPPORO mission. This could be a challenging factor in terms of gaining microwave remote-sensing data of sufficient spatial resolutions for a wide latitudinal range. It would be very useful to produce a survey of the low-dust zones (if any) in the Saturnian ring system from the Cassini mission with a view of designing a future polar grazing orbiter mission. If this could not be done, we would need to consider the sounding of the low-latitudinal region by having the SAPPORO Orbiter skimming

above the upper atmosphere near he equator and the two (ascending and descending) nodal points to be located outside the main rings.

An advantage of the equatorial orbits is that Titan and the icy satellites could be targeted. Microwave sounding of Titan's atmosphere must be very interesting. A microwave spectrometer experiment was in fact proposed for the Cassini mission but was not accepted. With such an instrument, the global wind field and meteorology of Titan would be studied in detail. Furthermore, the emission and expansion of water and other vapor gas from the hot south pole of Enceladus can also be probed. It would probably not come as a surprise if one day the cometary-like Enceladus is to become the center-piece of the next mission to the Saturnian system. In passing, we note that the Rosetta spacecraft to comet 67P/Churyumov-Gerasimenko carries a microwave instrument (MIRO) onboard just for the very purpose of detailed investigation of the gas and dust emission from active regions of the cometary surface. From this point of view, a microwave spectrometer will be an important instrument for future investigations of Enceladus and other Saturnian icy satellites.

2.3. *The Phoebe rover or penetrators*

On the way in towards Saturn before Saturn Orbit Insertion (SOI), the camera system on the Cassini spacecraft obtained spectacular images of the outer icy satellite, Phoebe (see Fig. 4) which is the largest member of the outer Saturnian satellites in irregular orbits. Because of its retrograde motion — with orbital inclination $i \sim 175^\circ$ — Phoebe is generally believed to be of capture origin, i.e., not a part of the family of regular icy satellites like Titan, Rhea, Dione, Tethys, and Enceladus. The fundamental question is where has it come from. One possibility is that Phoebe might have been a Kuiper belt object in transneptunian orbit before being trapped into the Saturnian system. From a dynamical point of view, capture from the residual population of planetoids in the vicinity of the Saturnian accretion zone would have been more probable. Either way, Phoebe represents the primordial building blocks of the outer planets holding the key to the solar system formation. Detailed analysis of the surface samples might provide very interesting information on the origin and evolution of this unique object. Such measurements should be done in as wide an area as possible by a mobile vehicle. Could the airbag hard-landing techniques for Martian rovers be adapted for a Phoebe Rover? If possible, we should consider the possibility of ejecting a daughter vehicle carrying the Rover towards Phoebe

Fig. 4. A close-up view of the outer irregular Saturnian satellite, Phoebe, by the Imaging science subsystem on the Cassini spacecraft. *Courtesy*: Jet Propulsion Laboratory, NASA, and Space Science Institute.

as the SAPPORO spacecraft goes by. Considering the low-gravity environment and the rugged terrains of Phoebe, the deployment of a rover could in fact be very challenging. An alternative approach is to use hoppers as proposed previously for the study of the Martian moon, Phobos, by the Russian Space Research Institute (IKI). A penetrator system similar to the one designed by ISAS/JAXA for the SELENE mission to the Moon could potentially be very valuable.

2.4. *The ring hopper/smasher*

Many unsolved mysteries would probably be identified by the Cassini mission. A most probable one could be related to the internal structure of the ring particles. How do the ring particles collide and evolve under the influence of the planetary tidal force and mutual gravitational attraction? How do the small and large bodies follow repeated cycles of coagulation and breakup? And then how can we find the ground truth? An idealized approach would be to equip a long-wavelength radar system similar

to those carried on the Mars Express and Rosetta spacecraft to a Saturn Orbiter moving in the ring system (if it could be done). The corresponding subsurface radar sounding would probe the interiors of different ring particles. Could such parking orbits be achieved by the operation of an electric propulsion system is an interesting engineering issue to be investigated.

The Galileo mission was ended by plunging the Galileo Orbiter towards Jupiter. Along this line of thinking, we should consider the option of keeping the Cassini spacecraft in orbit until the arrival of the SAPPORO Orbiter with enough fuel to crash into a chosen ring particle. Even though a little cruel, such a Kamikaze mission now made famous by the Deep Impact on comet 9P/Tempel 1 on July 4, 2005 could yield a lot of important information on the meteoroid bombardment effect and the subsurface structure of the ring particles.

2.5. *The cruise and magnetosphere science*

In 1980's and 1990's when very few opportunities existed for the space study of asteroids, space projects to the outer solar system like Galileo had always been required to introduce into the mission plans one or two flybys of asteroids in Earth-crossing orbits (i.e., the Near-Earth Asteroids) or in the major asteroidal belt. In future, our attention will likely be shifted to the remote observations of asteroids in the outer asteroidal belt like the Hilda asteroids and the Trojans in 1:1 orbital resonance with Jupiter. In addition to their being the primitive bodies left over from the formation of the solar system, the Hilda and Trojan asteroids have the additional important property that they were formed in the vicinity of the so-called snowline dividing the terrestrial rocky planets and jovian gaseous planets. It might be fair to say that only when we have the first-hand information on the chemical compositions of these outer-flying quasi-snowball objects would we be able to understand the structure of the solar nebula.

As for study of the Saturnian magnetosphere, we could envisage multi-spacecraft observations combining the SAPPORO and Cassini Orbiters — before the latter meets its demise by crashing into the rings. With the advance in nanotechnology and miniaturization, it is expected that a new generation of remote-sensing and plasma experiments will be onboard the SAPPORO spacecraft. It can be imagined that an array of multi-wavelength instruments ranging from infrared to X-ray wavelengths will be used to

study Saturn's auroral zones and equatorial X-ray emission among other scientific objectives.

3. Related Mission Plans

The above wish list contains a number of mission components which could be integrated into one single project or divided into several different projects based on the considerations of cost-effectiveness, technological, and programmatic constraints. For example, the Saturn atmospheric probe could be one element of the multi-probe mission to outer planets under study by NASA. The Phoebe Rover and Ring Hopper/Smasher could be included into one single mission. Even though the outer solar system beyond the orbit of Mars is still to be penetrated by Asian space agencies, we have learned from reports in recent scientific meetings that both Japan and China have initiated planning activity in this direction. For example, ISAS/JAXA is investigating the application of the solar sail technique to a Jupiter Orbiter mission, and CSSAR is exploring the feasibility of a solar polar observatory via a Jupiter flyby. It is therefore, hopeful that more and more scientists and engineers in these countries will soon focus their attention to the Saturnian system — because of the many exciting results from the Cassini mission. Indeed, there are likely to be a number of new mission plans to Titan in the next few years making the Saturnian system in future a place just as busy as the Wang-Fu-Jeng district in Beijing.

4. Discussion

It has been said that the initial difficulty with the Cassini proposal — when it was first conceived in 1982 — was that the proposers had to convince senior colleagues (20 years older in age, say) about the scientific importance of the mission. It is now our turn to convince junior colleagues who are 20 years younger about the scientific significance of the SAPPORO mission in 2025. The situation is in fact far better than it sounds. This is partly because several major Asian countries are beginning to enter the field of planetary exploration following the footsteps of NASA, ESA, and JAXA. The Chinese Chang-e mission and the Indian Chandrayaan-1 mission to the Moon are only the forerunners. We expect great leaps forward in planetary science in Asia which also happens to be the new powerhouse of the world economy. It is hard to escape the impression that there will soon be emerging interests in planetary atmospheres, planetary aeronomy and

magnetospheres, cosmochemistry, and small bodies in the AOGS scientific community. And the Saturnian system will be the place to go after the young generation of Asian space scientists has moved from the Moon to Mars, and to Jupiter in cooperation with their colleagues in Europe and America.

Galileo said that the Galilean satellite system appeared to be another solar system for him. The same thing can be said of the Saturnian satellite system.

To study these wonderful gifts and creations from Nature — as in the case of our Earth — would take time and patience. It would take many different missions and possibly many generations before human could finally understand the origin, evolution, and structures of the Solar system and the planetary satellite systems in it. What we have sketched here is mainly to size up the different options which are important to address different major scientific issues. Some of these (i.e., the ring hopper) are clearly very demanding in technical terms and would require advance technology such as nuclear propulsion systems still to be developed. And some emerging hot topics like Enceladus and Iapetus are so new that, very little can be said about it — except for amazement — at the present moment. It is our belief that future discussions will clarify this situation. In this spirit, we use Table 1 to summarize the mission elements which would play essential roles in the next generation of space exploration of the Saturnian system.

In Table 1, we have divided six mission elements into three sets. The idea is to show that, realistically speaking, a multi-mission approach would be necessary if we want to carry out all desired scientific investigations. Depending on the technological readiness, financial constraints, and focused scientific goals, we could envisage the combinations of the mission elements into two to three integrated projects. For example, we could have the combinations of (A,B = Saturn Atmosphere Probe plus Titan Orbiter), (A',B' = Polar Orbiter and Microwave Sounder plus Enceladus Lander), and (C,C' = Phoebe Rover plus Iapetus Rover) or the combinations of (A,B'), (A',B), and (C,C') with the ring hopper (D = Ring Hopper) as a wild card. The main point is to allow for the coverage of as broad a spectrum of scientific interests as possible. We might be criticized as being too idealistic and optimistic in proposing such an endeavor which will require pooling of all scientific resources of Asia, Europe, and Americas in the next 20 some years. To answer this, we must return to the statements made in Sec. 1 projecting the Asian economy and science to be comparable to those

Table 1. Listing of mission elements.

	Mission element	Scientific objectives	Technical study requirements
A	Saturn atmosphere probe	Atmospheric elemental and isotopic abundances	Entry probe and telemetric relay
A'	Polar orbiter and microwave sounder	Atmospheric composition and structure	Trajectory design and ring dust hazard mitigation
B	Titan orbiter	Titan atmosphere, surface, and internal structure	Possible balloon or roving platforms
B'	Enceladus lander	Surface properties and search for biomarkers	Rover system
C	Phoebe rover	Surface properties and subsurface structure	Rover or penetrator system
C'	Iapetus rover	Surface properties and subsurface structure	Rover or penetrator system
D	Ring hopper	Ring particle structure and dynamics	Hopper (or multi-hopper) system

Note that we have not specifically included a post-Cassini–Huygens mission dedicated to Titan into our general consideration here for the simple reason that the scientific community involved in this ongoing NASA–ESA project must have been tempted to produce a game plan leading to this direction and such an initiative should be best pursued in parallel to our effort. Furthermore, in the continuing observations of the Saturnian system and Titan in the next few years, there might still be some completely unexpected results (such as the active outgassing phenomenon of Enceladus) which would shift our attention elsewhere other than Titan.

of Europe and USA combined together in 2025. This would only be possible if peaceful cooperation and co-existence can be established among the Asian countries and the rest of the world. If this condition can be reached, much more grandiose plans would be within our reach.

Acknowledgments

The work was partly supported by grants NSC 94-2752-M-008-001-PAE, NSC 94-2112-M-008-002, and NSC 94-2112-M-008-019.

References

1. T. Owen, P. R. Mahaffy, H. B. Niemann, S. K. Atreya, T. Donahue, A. Bar-Nun and I. De Pater, *Nature* **402** (1999) 269.

2. S. K. Atreya, M. H. Wong, T. C. Owen, P. R. Mahaffy, H. B. Niemann, I. De Pater, P. Drossart and Th Encrenaz, *Planet. Spaced Sci.* **47** (1999) 1243.
3. S. K. Atreya, P. R. Mahaffy, H. B. Niemann, M. H. Wong and T. C. Owen, *Planet. Space Sci.* **51** (2003) 105.
4. D. Gautier, F. Hersant, O. Mousis and J. I. Lunine, *Astrophys. J.* **550** (2001) L227.
5. F. Hersant, D. Gautier and J. I. Lunine, *Planet. Space Sci.* **52** (2004) 623.

This page intentionally left blank

OBSERVATION OF LUMINOUS TRANSIENT PHENOMENA ON PLANETARY BODIES

MARIO DI MARTINO and ALBINO CARBOGNANI
INAF — Osservatorio Astronomico di Torino, 10025 Pino Torinese, Italy
dimartino@to.astro.it

In this paper, we review the possibility of space observations for the most important luminous transient phenomena occurring on planetary bodies. The construction of a breadboard of a CCD camera having a field of view of 120° and a sensitivity able to detect events of visual magnitude +6, together with the realization of the relative operation software, has been funded by the European Space Agency (ESA). This project has been developed by the firm Galileo Avionica S.p.A. (Campi Bisenzio, Italy). The estimated cost for such a space qualified camera is estimated between 3 and 4 million US$. It could fly as a piggy-back payload on a space platform and/or a constellation of satellites conceived for different space missions.

1. Introduction

Transient luminous phenomena on planetary bodies are defined as events of different intensities, which occur in planetary atmospheres and surfaces, their duration spans from about 0.1 s to several days. They consist of meteors, bolides, lightning, impact flashes on solid surfaces, auroras, etc. If well monitored, they represent a very useful tool to study the smallest component of meteoroids and the electric phenomena in planetary atmosphere.

At present only terrestrial transient phenomena are studied. However, ongoing programs devoted to systematic meteor observation worldwide are a few. In the USA only the Florida Fireball Patrol is still active, while in Europe, since 1959, the European Fireball Network (EN) is working. No information exist about extraterrestrial meteor occurring on the solar system planets.

The earth-lightning detection from space it is not a novelty, but a systematic and global monitoring of extraterrestrial lightning has not been carried out. So far, the study of these phenomena has been very limited, due to the lack of an *ad hoc* instrumentation, and their detection has been performed mainly on a serendipitous basis.

Recently, European Space Agency (ESA) has issued an announcement of opportunity for the development of systems devoted to the systematic

detection from space of luminous transient events in the Earth atmosphere and/or on the dark side of other planetary objects. One of such a detector as been designed and a prototype (smart panoramic optical sensor head, SPOSH) has been constructed at Galileo Avionica S.p.A (Campi Bisenzio, Italy). Efforts have to be done now to place this instrument on space platforms conceived for other missions in order to low the costs.

In what follows we review the possibility of space observations for the most important luminous transient phenomena. For sake of clarity we classify the transient phenomena in "Earth phenomena" and "Planetary phenomena", even though some of them originate in a similar physical context.

2. Earth Phenomena

Transient luminous phenomena on Earth occur mainly in its atmosphere at different heights. Their origin is due essentially to the interaction of cosmic debris with the high layers of atmosphere and the electric activity that is present in it at different levels.

Interest for meteors is due to the information they can provide about the history of the solar system and the properties of the original planetesimals, as well as to the large influence of meteors on the upper atmosphere. The Hunten particles (suspended atmospheric particles at altitudes between 80 and 100 km of meteoroid origin) act as condensation nuclei for sporadic phenomena, as noctilucent clouds and meteor trails provide valuable information in turbolence and wind diagnostics.

Concern is rising also about the space debris issue, given the increasing risk for aerospace vehicles. From the event identification standpoint, it can be difficult in many cases to discriminate meteors from space debris events.

Lightning is a phenomenon of high interest in different scientific fields, including studies of the water cycle, and also taking into account less known events like jets, sprites, etc. Among the different electric phenomena in the atmosphere, satellite observations of lightning and auroras are relatively easy. More difficult are the observations of red sprites, blue jets, and elves, due to their superposition to stormy cells where lightning are numerous.

2.1. *Meteors*

Asteroids and comets are the sources of small interplanetary bodies. The collisions between asteroids in the main belt, besides producing km-sized asteroids, generate huge numbers of fragments having sizes spanning from a hundredth of millimeter to some tens of meters, intermediate between

classical asteroids and interplanetary dust. These bodies are called *meteoroids*. The International Astronomical Union (IAU) in 1961 established that meteoroids are defined as the small bodies having mass in the range 10^{-6}–10^{10} g. Assuming a density of 3.5 g/cm^3, the radius of a meteoroid spans from 40 μm to 10 m. Even comets produce meteoroids, although their density is likely lower than those of asteroidal origin. Comet nuclei produce "meteoroid streams" which follow the orbital trajectory of the parent body and originate *meteor showers* visible from Earth.

Approximately 4×10^4 metric tons (40×10^6 kg) of extraterrestrial matter enter the Earth atmosphere every year and eventually settle on the ground. This matter originates from two main components, cosmic dust and meteoroids. Whereas, there is a large uncertainty in the absolute fluxes, there seems to be agreement that mass influx of particles having masses in the range 10^{-5}–10^{-6} g accounts for about 20% of the total flux per year, with the remaining 80% arising from objects in the mass range 10^{-4}–10^{15} g.[4]

Whatever the origin of a meteoroid belonging to the solar system is, its geocentric velocity spans from 11.2 km/s (the escape velocity from Earth) to 72.8 km/s (42.5 km/s, the escape velocity at the perihelion of the Earth orbit, plus 30.3 km/s, the Earth orbital velocity at perihelion). During the atmospheric flight the atoms of the meteoroid disperse in the atmosphere forming a trail, the *meteor*, which looks like a long and narrow cylindrical column. The trail initial dimensions are equal to the mean free path at that height, i.e., about 1 m at 120 km and about 10 cm at 90 km. The trail length can reach several kilometers. A typical value spans from 10 to 20 km, while the heights where the phenomenon begins are between about 120 and 75 km. Time duration of meteor phenomenon spans from 0.5 to 3 s.

A meteor is composed by two elements: the *head* and the *trail*. The meteor head contains the meteoroid, which is ablating, and the ionized gases; the trail is the region where the ionized gases recombine.

Like for other celestial bodies, we can define an *absolute magnitude* for meteors: the apparent magnitude of a meteor located at the zenith at a height of 100 km. Taking into account that the average height where meteors occur is around this value, the apparent magnitude of a meteor at the zenith is roughly comparable with its absolute magnitude.

The visual magnitude, M, is expressed in a logarithmic scale related to the luminous intensity, I, via the relationship:

$$M = 6.8 - 2.5 \log I, \tag{1}$$

where I is in Watt.

The activity of sporadic meteors shows diurnal and seasonal variations.

2.2. *Bolides and superbolides*

Due to orbital perturbations with Jupiter, Saturn, and Mars, and also under the influence of radiative mechanisms like the Yarkowsky effect, meteoroids originating in the asteroid main belt can be inserted into orbits crossing those of the terrestrial planets: Mercury, Venus, Earth, and Mars. Therefore, there are a lot of small bodies, having dimensions of some meters, which can interact with our planet. If the meteoroid has a size larger than about 20 cm, during the passage in the atmosphere the meteor head can reach a high luminosity. When the zenithal apparent magnitude is lower than −8 magnitudes, the meteor is called *bolide*. If the bolide magnitude is lower than −17, this is called a *superbolide*. Bodies having mass larger than 1,000 kg generate superbolides. For meteoroids having diameter of some tens of meters the bolide can be brighter than the Sun (apparent magnitude −27). Superbolides are rare events, which would need a global observing network, in order to be studied in a systematic way.[5]

Often the meteoroid undergoes multiple fragmentations, generating a multiple bolide. If the meteoroid is big enough, it can survive the ablation process. When the velocity decrease below 3 km/s, the mass loss and the radiation emission end, and the meteoroid enters the so-called *dark flight* phase. From this moment a cooling process begins, while at the same time the body trajectory becomes more and more vertical.

2.3. *Meteor observations from space*

Still to be explored is the systematic monitoring of meteors from satellites in orbit around Earth. The USA military surveillance satellites reveal 30–50 superbolide explosions in the atmosphere per year, but frequently the data on these events, especially the less bright, are discarded. A satellite network equipped with dedicated cameras is thus necessary to observe in a global and systematic way these phenomena.

If q is the height in km of the satellite from the Earth surface, and M the meteor absolute magnitude, the apparent magnitude of the meteor observed from orbit is given by:

$$m = M - 10 + 5 \log(q - 100). \tag{2}$$

Considering the sensor limit magnitude on board the satellite +6.0 and the height 400 km, from Eq. (2) it follows that from orbit only meteors having absolute magnitude lower than $M = +3.6$ will be detectable. The trends of Eq. (2) in function of the height and for different M values are plotted in

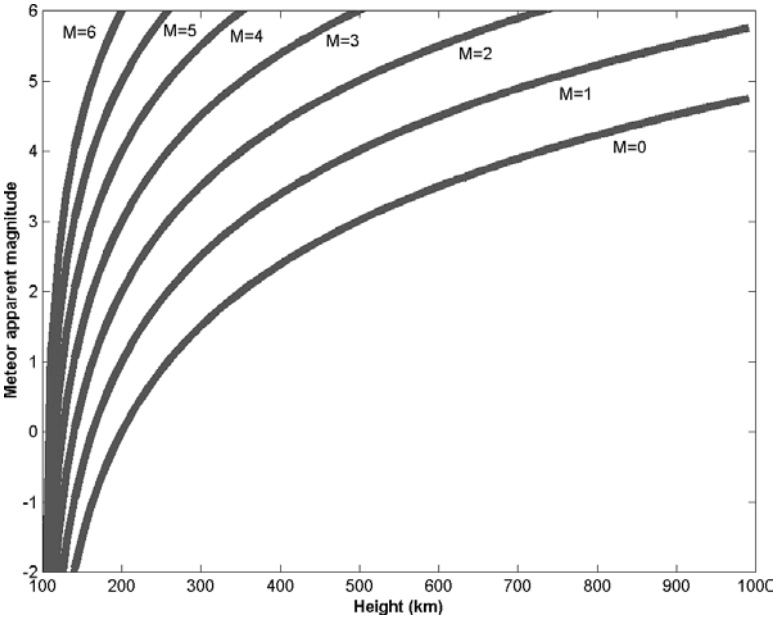


Fig. 1. Meteor apparent magnitude versus orbit height and absolute magnitude M.

Fig. 1. At a height of 1,000 km only meteors having $M = +1$ are still visible. Only in the case of bolides problems do not exist. From (2), a typical bolide having absolute magnitude $M = -12$ observed from an height of 400 km becomes a -9.6 meteor, whereas from 36,000 km its magnitude is $+0.8$, well visible by a sensor designed to detect meteors having magnitude $+6.0$.

Due to the largest distances, from orbit a meteor trail, on average, shows an apparent length shorter if compared with that observed from the ground.

The scenario is schematized in Fig. 2, where the meteor trail is parallel to the radiant direction.

If θ is the angle between the satellite radial direction and the meteor radiant, l the trail length, and h_{m} the average meteor height from the Earth's surface, the angle β subtended by the trail which is at the center of the field of view is given by:

$$\tan\beta = \frac{l\,\sin\theta}{q - h_{\mathrm{m}}}. \tag{3}$$

In the most common cases the angle will be $\theta \neq 0°$. Taking into account an average meteor height of 100 km, a typical angle $\theta = 45°$, a trail length

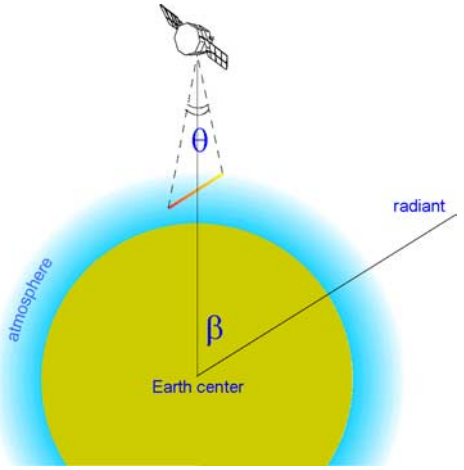


Fig. 2. Geometry for the computation of the length of a trail as viewed from space.

Table 1. Meteor trail apparent length at the FOV center versus orbit height.

Height (km)	Apparent length β (degrees)	Angular velocity (degrees/s)
400	2.00	2.70
1,000	0.70	0.90
10,000	0.07	0.09
36,000	0.02	0.02

of 15 km, and assuming a value of 20 km/s as geocentric velocity we obtain the data listed in Table 1.

As can be seen, even from high orbits, the trail length is long enough and the angular velocity are not prohibitive for modern CCD sensors.

2.4. *Lightning*

Lightning play a key role in the dynamic interplay of forces occurring in the Earth's atmosphere. Lightning distribution on Earth is not uniform. Most recent data have been obtained by the NASA satellite optical transient detector (OTD). The OTD has observed about 1.2 billion lightning per year. On the oceans lightning are less frequent than on dry land and thicken between equator and tropics. The electric activity is stronger in the north hemisphere than in the southern one. The most "stormy" areas of our planet

are: Zaire, Sudan, Cuba, Mexico (Sierra Madre), India (Himalaya), and South–East Asian countries.

A typical lightning is composed by a series of electron discharges, in average four. The duration of each discharge is 35 ms, while the duration of the series spans from 0.1 to 0.25 s, but can reach 2–3 s. In the electric discharge the temperature of the atmospheric gas can reach 20,000 K, while the peak optical power is in the range 7×10^7–3×10^9 W.[9] The spectrum is black body like with a peak emission around 144 nm. On the Earth every time about 2,000 stormy cells exist, producing about 100 discharges per second.

From the Pogson relation between flux density and magnitude, and using the above-mentioned typical values, we can derive the magnitude difference between two bodies having different fluxes. Considering as reference flux the optical solar constant F_1 (1.83×10^2 W/m^2), the magnitude m_1 is the apparent Sun visual magnitude (-26.74). The apparent visual magnitude of a lightning, having an optical power of 3×10^9 W and observed from a height q (km) on the Earth's surface, is:

$$m_{\text{lightning}} = 5 \log(q) - 27.0. \tag{4}$$

Assuming a 400 km height, the apparent peak magnitude of a lightning is -14, larger than the Moon apparent magnitude (-12.5). Then, a lightning can be revealed without difficulty through the clouds and on the background of anthropic lights.

The observation of lightning from space differs in a substantial way from meteor observation not only for the peak magnitude, but because the electric discharges occur in short time in the same position.

2.5. *Sprites*

The terrestrial electric activity involves not only the troposphere but also the highest atmospheric levels up to the ionosphere. As already seen, a cloud-to-ground lightning transports negative charges from the cloud to the ground. Sometimes, the positive cloud top can lose its charge which discharges on the ground. When this occurs, the isolated negative region can create an intense electric field within the ionosphere. This field accelerates the electrons towards the ionosphere and, hitting the atmospheric molecules, they excite their energetic levels. When these molecules return to the fundamental state, they emit visible radiation. This is the origin of the so-called *red sprites*.

Red sprites appear as luminous globes red in color with a quite low-surface intensity. Their frequency is about 1% with respect to the common lightning. Red sprites occur over the top of stormy cells and can be single or multiple. They can reach a height of 95 km with a maximum brightness around 65–75 km. The duration of a red sprite is of few milliseconds. The optical peak power is about 5–25 MW, while the dimensions are of the order of 10–20 km. If observed from a height q (in km), the apparent magnitude, for a power of 10^7 W in the area of maximum luminosity is

$$m_{\mathrm{RSprites}} = 5\,\log(q - 70) - 20.8. \tag{5}$$

A red sprite viewed from a height of 400 km has a total apparent magnitude of −8. The angular dimension will be about 4°. A hindrance to the observation of red sprites from space can be due to the fact that they tend to be projected on the stormy cell and then to be hidden by the lightning dazzle.

2.6. *Blue jets*

Much more enigmatic are the *blue jets*, high height luminous phenomena whose origin is still unknown. Blue jets appear as luminous blue-colored beams coming from the top of a stormy cell.

They have a conical form with an aperture of about 15°. Their velocity is about 100 km/s and reach 40–50 km in height, then their luminosity decreases and they disappear. The phenomenon duration is of the order of 0.1 s, while the peak power is 10^4 W. If observed from orbit, the total apparent magnitude in function of the height is

$$m_{\mathrm{BJets}} = 5\,\log(q - 20) - 13.3. \tag{6}$$

At 400 km a blue jet has an apparent magnitude of about −0.4. The problems arising when these phenomena are observed from space are analogous to those related to the red sprites: they are geometrically projected against the stormy cell, and then the emitted radiation is dominated by the light coming from the lightning. In any case, if the sensor limit magnitude is +6, the observing height cannot be larger than 1,000 km.

2.7. *Elves*

Elves are produced by the interaction between the mesosphere (between 75 and 100 km) and the electromagnetic fields originated by very powerful lightning. They have a flat and circular shape with diameters of about

200 km. They can be associated also with red sprites but form before. Their duration is less than 1 ms. The elves power is far less than the red sprites and they can be observed only if they are viewed in profile. Viewed from orbit, elves are projected against the stormy cell, so their observation can be very difficult.

2.8. *Noctilucent cloud*

Noctilucent clouds (NLC) are the highest clouds observable in the atmosphere, their average height is 83 km. The NLC are formed by water ice and are similar to cirrus, but are thinner and blue or silver colored. Due to the absence of winds, their evolution is very slow, if compared to the tropospheric clouds. Due to their low-optical thickness, they can be observed at sunset and dawn, when the Sun height is 6–12° below the horizon. In these conditions their contrast with the sky is maximum. The NLC are usually observed during the summer from regions having latitude included between 50 and 60° in both hemispheres. During one year, 10–20 NLC are visible, thus it is a quite rare phenomenon. From space they can be only visible from the terminator region, they are very thin, so, if observed from space, their localization is very difficult. From the ground, in fact, they can be observed mainly when they are low on the horizon, a condition in which their optical thickness is higher.

From space and using a large field of view, the observation of NLC is hampered from their faintness, which requires high-exposure times near the terminator light.

2.9. *Auroras*

Auroras are caused by the interaction of charged particles of the solar wind with the Earth's magnetic field and ionosphere atoms and molecules (Table 2). The aurora heights span from 70 to a maximum of 1,100 km with the higher occurrence being between 90 and 110 km. The duration of auroras can span from some seconds to some days and their intensity is enough high to pose no difficulties to satellite observations.

3. Planetary Phenomena

They essentially consist in the same events we can observe on Earth, except impact flashes on the surface of atmosphereless bodies. Of course, these phenomena show different characteristics due to the diverse physical environments.

Table 2. Characteristic emission line of an aurora spectrum.

Wavelength (nm)	Molecule/atom	Height (km)
391.4	N^+	1,000
427.8	N^+	1,000
557.7	O	90–150
630.0	O	>150
636.4	O	>150
661.1	N_2	65–90
669.6	N_2	65–90
676.8	N_2	65–90
686.1	N_2	65–90

3.1. *Meteors on Mars*

As we have seen, on Earth a meteor occurs when a meteoroid interacts with the atmosphere between 120 and 70 km. A part Mercury and Pluto, in principle, all the solar system planets (even the Saturn satellite, Titan, and Neptune satellite, Triton) have sufficiently dense atmospheres to generate meteors. Mars is a planet where meteors could be easily observed.

In spite of the difference in the chemical composition of the Earth and Mars atmospheres, at the same conditions (atmosphere density, meteoroid mass, and velocity), the meteors show the same luminosity because less than 3% of the trail emitted radiation is produced by atmospheric atoms. Considering the same meteoroid mass and velocity, the meteors on Mars occur about 20 km lower.[1]

Meteors on Mars are very similar to the terrestrial ones, thus the instruments designed for observation from orbit of terrestrial meteors can also be used for the observation of Martian meteors.

3.2. *Meteors on Venus*

Meteors on Venus could be observable from orbit and, due to the higher density of the Venusian atmosphere, the interval 120–70 km for the meteors on Earth corresponds to 300–200 km on Venus. This means that the apparent brightness from orbit will be higher with respect to a similar event on Earth. Moreover, the height where the meteor phenomenon occurs on Venus is larger than the haze and cloud upper limit (80 km), allowing the observation from space. Assuming a height of 600 km and a distance of 250 km from the Venusian meteor, we estimate that on Venus meteors are apparently 0.2 magnitudes brighter than on Earth. So, as in the case of Mars,

the instruments designed to observe meteors from space on the Earth are also valid for Venus.

3.3. *Meteors on giant planets*

Considering the same conditions (mass and velocity of the meteoroids and atmospheric density), even on giant planets (Jupiter, Saturn, Uranus, and Neptune) meteors have to be as luminous as on the other planets.

In the case of Jupiter, by using the law of the atmospheres:

$$\rho(q) = \rho_0 \exp(-q/H), \tag{7}$$

where q is the height, ρ_0 the density at the atmospheric level with 100 mbar pressure ($q = 0$ km), and H the atmosphere scale height, we obtain that the interval 120–70 km for the meteors on Earth corresponds to the interval 290–115 km on the giant of the solar system. These values have been computed considering as level zero the atmospheric layer at the pressure of 100 mbar.

A meteor trail in the Jupiter atmosphere has been observed during the Voyager 1 fly-by.[6] From the observations, it results that the bolide reached the atmospheric level at 3.5 mbar, at about 100 km over the zero level at 100 mbar, in good agreement with the estimated height range for the Jovian meteors. The instruments used for the observation of the terrestrial meteors from orbit are also valid for Jupiter, where the meteor phenomenon is substantially analogous.

4. Impact Flashes

The Moon, as all the solar system bodies, has undergone a continuous bombardment by asteroids, comets, and meteoroids since its formation. Due to the absence of an atmosphere, during the last phases of the fall the meteor phenomenon does not occur and all the meteoroid kinetic energy is released on the ground. During the impact a fraction (spanning from 10^{-3} to 10^{-2}) of the impact energy is converted in a luminous flash.

Meteoroid impacts on the Moon have been detected in the past visually and by the Apollo Lunar Seismic Network. Several observers have claimed to have detected optical flashes on the Moon, but any of such events has never been confirmed in an independent way, until the impacts on 18 and 19 November 1999 due to Leonid shower. Unfortunately, different phenomena can cause the appearance of a flash, such as reflection in the instrument optics, cosmic rays hitting the retina, reflections due to artificial satellite

crossing the lunar disk, etc. This is why we need two independent observations of the same flash.

The detection of impact flashes by using a camera on board a lunar satellite should be much more efficient, if compared with the ground-based observations. In fact, being the distance much lower, even the impact of meteoroid having small mass should be detectable. Moreover, we could avoid the constraints due to the geometry of the Earth–Moon system, having the possibility to monitor the lunar dark hemisphere even when it is invisible from Earth.

Considering q_{L} the spacecraft height (in km) from the Moon surface, the impact flash magnitude is given by

$$m = m_{\mathrm{T}} + 5\,\log(q_{\mathrm{L}}) - 27.9, \tag{8}$$

where m_{T} is the event magnitude as seen from the Earth. Assuming that the limiting magnitude of the sensor is +6, from low orbits (within 5,000 km from the lunar surface), flashes that from the Earth should appear as magnitude +15 could be visible. For impact flashes having magnitude between +3 and +7, as seen from the Earth, for distances from the lunar surface lower than 50,000 km, the apparent magnitudes are negatives. This fact should allow to study very low-mass meteoroids and to take spectra of the impact phenomenon, at present completely unknown.

To estimate the impact flash magnitudes on Mercury we have to obtain the impact scaling law. The emitted radiation during an impact is proportional to the fourth power of the impact velocity.[7] Thus, the radiation flux density, F, is proportional to:

$$F \propto \frac{V^4}{q^2}, \tag{9}$$

where V is the impact velocity and q is the observer distance from the impact point. Using the Pogson relation, the difference in magnitude between the flash apparent magnitudes on the Moon and Mercury are given by:

$$m_{\mathrm{M}} - m_{\mathrm{L}} = -2.5\,\log\lfloor (V_{\mathrm{M}}/V_{\mathrm{L}})^4 (q_{\mathrm{L}}/q_{\mathrm{M}})^2 \rfloor, \tag{10}$$

where V_{M} and V_{L} are the impact velocities on Mercury and on the Moon, respectively.

The Galilean satellites are bodies of planetary dimensions having a solid surface where is possible to observe the flash produced by a meteoroid impact. As already done in the case of Mercury, to estimate the flash magnitude we have to use the scaling law for the impacts.

5. Electric Discharges

Lightning have been detected on every planet with an atmosphere, except Mars. On Saturn, Uranus, and Neptune, atmospheric electric activity has been recorded in the radio VLF band.[9] Due to the high opacity of the Venus, Saturn, Uranus, and Neptune atmospheres, which hamper the observation of lightning, candidate planets to image atmospheric electrical discharges are Mars and Jupiter.

5.1. *Observation of Mars lightning*

Due to the prevalence of Martian dust devils and dust storms, an understanding of the underlying physics of electrical discharges in Martian dust clouds is critical for future Mars exploratory missions. Mars low-atmospheric pressure and arid, windy environment suggest that the dust near the surface of Mars is even more susceptible to triboelectric charging than terrestrial dust. Electrical discharges on Mars should occur more frequently but at lower intensities than those seen on Earth.

On the Earth, lightning discharges occur in association with desert sandstorms and volcanic ash plumes. In these events, electric charge separation begins when dust particles collide or brush past each other in the turbulent air. In a Martian dust cloud the electric potential should remain near zero. Variable winds and/or gravitational settling, however, may separate particles by size, building substantial electric fields within the cloud. When these fields reach a critical value, a lightning discharge occurs. Dust storms have been observed on Mars to develop and spread over the entire planet and may last for months. Smaller storms are also known to occur and should provide an ideal environment to search for Martian lightning. On Mars lightning may be very different from terrestrial lightning due to the low-atmospheric pressure. Some authors have suggested diffuse glows or flashes, filamentous discharges, or small arcs.

The most likely candidate for the creation of electrostatic charges and fields is triboelectric charging of dust, i.e., the friction between blown dust particles and the ground (or between) dust particles with each other. Terrestrial experience demonstrates that electric fields of 5–15 kV/m are not uncommon during dust storms and dust devils.[10] Olhoeft[8] suggests that Martian lightning will be a diffuse flash (similar to summer heat lightning). Because of Mars low-atmospheric density, electrical discharges occur at lower electric potential than on Earth, and therefore, should be more frequent. The breakdown electric field on Mars is

expected to be between ∼5 and 20 kV/m, compared to ∼3,000 kV/m on Earth.

To have a rough idea of the brightness of a Mars lightning, we consider an electric field, E, of 5 kV/m, the energy density is

$$u = (1/2)\varepsilon_0 E^2 = 1.11 \times 10^{-4}\,\mathrm{J/m^3}, \tag{11}$$

where ε_0 is the vacuum dielectric permittivity.

If the dust cloud is assumed to be 1,000 m high, and to have a projected surface area of 1 $\mathrm{km^2}$, then the total cloud volume is 10^9 cubic meters, and the total energy due to charge separation within the cloud (assuming that E is uniform throughout) must be 1.11×10^5 J. If the camera recording the flash were 200 km away from the cloud, and if 1/10 of the total energy in the discharge produces light, then 1.11×10^4 J of light would be spread over a sphere of area $5 \times 10^{11}\,\mathrm{m^2}$. If the discharge takes 1 ms to occur, the intensity at the camera is approximately $2.2 \times 10^{-5}\,\mathrm{W/m^2}$. Considering that the Mars dust clouds can cover the whole planet surface, our estimate intensity has to be considered a lower limit.

5.2. *Observation of Jupiter lightning*

The possible occurrence of lightning in the Jupiter atmosphere was firstly predicted by Bar-Nun.[2] The Voyagers 1 and 2 spacecrafts performed the first observations of lightning on Jupiter in 1979. Later, even the Galileo spacecraft has monitored the electric activity in the Jupiter atmosphere. A typical Jovian storm is about 1,500 km in diameter and produces about 20 flashes per minute. The heights of the flashes are between 2 and 5 bar atmospheric pressure layer, in the region where the H_2O clouds are located. This suggests that the lightning generation mechanism is analogous to the terrestrial one (convective electrification of the clouds). In the visible band, the flash intensity ranges from 4.3×10^8 (for those having mean energy) to 6.6×10^9 J for the more energetic ones.[9] The total power is larger with respect to the optical one for a factor between 10^2 and 10^3.[3] In average, in the above-mentioned optical range 0.01 flashes per $\mathrm{km^2}$ per year occur. Assuming a typical duration of 35 ms, as for the terrestrial lightning,[11] the optical power span from 1.2×10^{10} to 2×10^{11} W and, if observed from orbit, the corresponding apparent magnitudes are negative till a distance of 10^5 km from the top of the Jupiter clouds.

6. Conclusions

So far, the study of luminous transient phenomena on planetary bodies has been very limited, due to the lack of an *ad hoc* instrumentation, and their detection has been performed mainly on a serendipitous basis. The times seem now mature to plan the development of a new generation of dedicated, space-based observing platforms. Visible and infrared cameras should be the primary sensors for these facilities, which do not require the launch of independent satellites or space probes, since they might be carried aboard of orbiters conceived to carry out also other activities, to reduce the overall costs of the deployment of an observing network. As a purely speculative example, for instance, one could think about the possibility of adding, as piggy-back, dedicated cameras to some satellite of the fleet which will be launched in the near future for the imminent European Global Navigation Satellite System (Galileo) or on satellite constellations devoted to Earth observation.

High-satellite orbits are better for sky covering issues, but there is also a corresponding intrinsic reduction in sensitivity performances, and in precision in the determination of the entry trajectory of the bodies. For this reason, an ideal space-based network would include also a number of satellites on lower altitude orbits. In addition, also the International Space Station could include some of the sensors of a more general network.

In practical terms, recently, ESA issued an announcement of opportunity for the development of systems devoted to the detection of transient events in the Earth atmosphere and on the dark side of other planetary objects. One of such a detector has been designed and a prototype, together with the relative operation software, has been constructed at Galileo Avionica S.p.A. (Campi Bisenzio, Italy). The estimated cost for a space qualified instrument is between 3 and 4 million US$.

Its characteristics are the following:

Field of view	120°.
Limit magnitude	$+6\,m_V$.
Size	$10\,cm \times 10\,cm \times 10\,cm$.
Mass	1,560 g.
Power consumption	4.9 W.

References

1. L. G. Adolfsson, S. Gustafson and C. D. Murray, *Icarus* **119** (1996) 144–152.
2. A. Bar-Nun, *Icarus* **24** (1975) 86–94.

3. W. J. Borucki and C. P. McKay, *Nature* **328** (1987) 509–510.
4. P. Brown *et al.*, *Nature* **420** (2002) 295–296.
5. Z. Ceplecha *et al.*, in *Meteoroids 1998, Astron. Inst., Slovak Acad. Sci.*, 1999, pp. 37–54.
6. A. F. Cook and T. C. Duxbury, *Journal of Geophysical Research* **86** (1981) 8815–8817.
7. G. Eichorn, *Planet. Space Sci.* **24** (1974) 771.
8. G. R. Olhoeft, in *Sand and Dust on Mars*, NASA CP-10074, 1991, p. 44.
9. C. T. Russel, *Ann. Rev. Earth Planet. Sci.* **21** (1993) 43–87.
10. D. D. Sentman, in *Sand and Dust on Mars*, NASA CP-10074, 1991, p. 53.
11. P. Zarka, *Astronomy & Astrophysics* **146** (1985) L15–L18.

SHORT ELECTRIC-FIELD ANTENNAE AS DIAGNOSTIC TOOLS FOR SPACE PLASMAS AND GROUND PERMITTIVITY

JEAN-GABRIEL TROTIGNON
Laboratoire de Physique et Chimie de l'Environnement, Centre National de la Recherche Scientifique, Orléans University
3 A, Avenue de la Recherche Scientifique, F-45071 Orléans cedex 02, France
Jean-Gabriel.Trotignon@cnrs-orleans.fr

Certain techniques used in geophysical prospection to measure ground permittivity have been transposed to space plasmas. The principle is to measure the self impedance of a single electric antenna or the mutual impedance between two sets of Hertz dipoles. Since probe impedances depend on dielectric properties of the immersed medium, space plasma parameters such as the density and temperature of thermal electrons may be reliably and accurately deduced. Natural waves may also be investigated in a large frequency range including characteristic plasma frequencies. As any electrode immersed in a plasma acquires a charge, it perturbs the plasma in its immediate neighborhood. The way to get around this difficulty is to use four electrodes, two for transmitting and two for receiving. Transmitting electrodes are excited from a signal generator, while receiving electrodes are connected to a high input impedance voltmeter. The transmitted current I and the received voltage V being known, plasma properties may be deduced from both the imaginary and real parts of the mutual impedance $Z = V/I$. Such quadripole probes have been used on sounding rockets and spacecraft.

1. Introduction

The impedance of an electric antenna depends on the dielectric properties of the medium in which the antenna is immersed. Electric sensors have therefore to be calibrated in flight in order to get usable E-field measurements. It was, for example, carefully done onboard the EIDI 1 ionospheric rocket,[1] and onboard the EXOS-D (Akebono)[2] and GEOTAIL[3] magnetospheric spacecraft by measuring the antenna impedances all along the spacecraft trajectories. This is not the purpose of this paper to recall how these calibrations were achieved, but we can mention than the main objective is to determine what deformations actually affect the amplitude and phase of the detected electric fields when the antenna is immersed in a plasma.[3] As no plasma chamber is large enough to contain such long electric-field

antennae, the only way is therefore to implement onboard calibrations. The other consequence is that main characteristics of the encountered plasma, such as the density and temperature of the thermal component of electrons may be determined from the antenna impedance measurements. If a suitable probe can be designed and if a theory exists that describes how the antenna impedance depends on these parameters, impedance measurement indeed becomes a powerful technique for space plasma diagnostic. As a bonus, natural waves may also be investigated in a large frequency range including characteristic plasma frequencies such as the lower- and upper-oblique resonances, the electron gyrofrequency and its harmonics, the plasma frequency, the upper hybrid frequency, and so on.

The capability of such a technique has been proven onboard rockets and spacecraft since several decades. The objective of the current paper is to briefly recall how the impedance probes are working and, in particular, to show how the major difficulty that their electrodes perturb the plasma is got around. Then, not necessarily very recent but well representative quadripole impedance probes that flew onboard a rocket (CISASPE), two spacecraft (GEOS-1 and AUREOL-3/ARCAD-3), and finally CASSINI/HUYGENS are presented.

2. How the Impedance Probe is Working

The impedance probe consists of the probe itself and the electronics that measure the impedance. The major difficulty is that electrodes actually perturb the plasma, this is due to the presence of ion sheaths around the electrodes that insulate them partially from the unperturbed plasma. Another difficulty comes from the antenna sensitivity to quasi-static disturbances originating in the spacecraft body and structures, that is why monopoles are rarely used and symmetrical dipole preferred because differential measurements nominally reduce these disturbances. An elegant way to get around these difficulties[4] is to use four electrodes, two for transmitting and two for receiving, and to measure the mutual impedance $Z = V/I$. Figure 1 illustrates the process. This technique has been used for many years in geophysical prospection to measure the ground permittivity. Transmitting electrodes are excited from a signal generator, while the receiving electrodes are connected to a voltmeter with a very high input impedance. By exciting the transmitting electrodes from a constant current source (source internal impedance π transmitting electrodes self-impedance), a single measurement (V) is actually necessary to obtain the mutual impedance $Z = V/I$. As long as transmitting electrodes are placed at "large" distances from the receiving

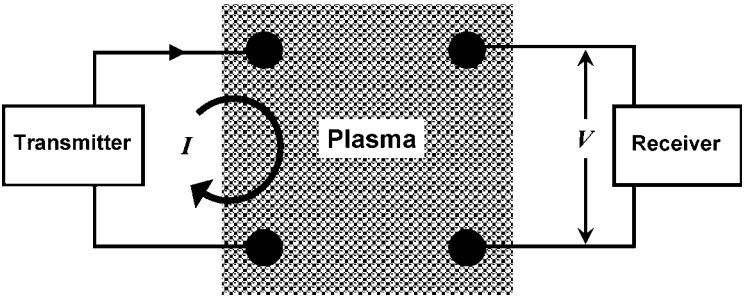


Fig. 1. A schematic illustration showing how the mutual impedance $Z = V/I$ of a quadripole probe is determined.

ones, potential difference V, on open circuit, is independent of perturbations (ion sheaths). The transmitter–receiver distance has thus to be larger than twice the plasma Debye length. The antennae are nevertheless short compared with the wavelength of EM waves so that the quasi-static approximation applies. Both imaginary and real parts of Z may then be interpreted to deduce plasma properties. Typical responses of the mutual impedance probe as a function of frequency for two Debye lengths (1.6 m above and 6 m below) and a plasma density of 11.5 cm^{-3} are shown in Fig. 2. Solid lines refer to the modulus of Z normalized to its value in vacuum. Dotted

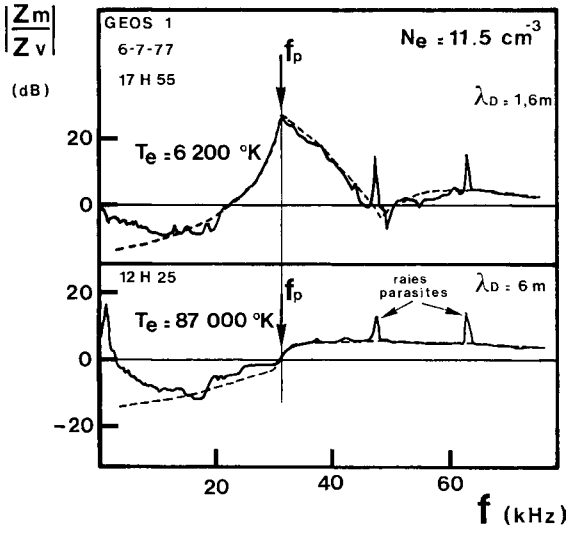


Fig. 2. Typical mutual impedance frequency responses measured onboard GEOS-1 for a plasma density of 11.5 cm^{-3} and two Debye lengths, 1.6 m (*top*) and 6 m (*bottom*).[5]

lines are theoretical responses. The resonance peak occurs at the plasma frequency f_p from which the electron density N_e is deduced. The shape gives the Debye length λ_D and therefore the electron temperature T_e.

3. Representative Quadripole Impedance Probes

A h.f. quadripole probe was launched from Kourou aboard the CISASPE rocket, on 16 December 1971, one hour before sunset. The apogee, 227 km altitude, was reached in 8 min, and the ionosphere was explored for 6 min 30 s. The sensor consists in 4 booms, each of them bears a sphere, 3 cm in diameter, at its end. The spheres are connected to the matching impedance device located inside the central mast (Fig. 3). Once deployed the spheres are at the corners of a plane square, 17 cm in length.[6] The warm magnetoplasma theory predicts: a resonance at the upper-hybrid frequency, f_T, an anti-resonance (at least) at f_D that depends on the Debye length (the Landau wave beats with a cold plasma term and leads to a minimum in the impedance modulus when the receiver–transmitter distance is close to a whole number of electrostatic wavelengths). Anti-resonances may also be observed at harmonics of the electron cyclotron frequency, nf_b. Observations have actually confirmed these predictions (Fig. 3). From the nf_b, f_T, and f_D the magnetic field strength, the plasma density N_e, and the electron temperature T_e were determined.[6] The Debye length was 5 mm at 215 km altitude and reached 5.3 mm at apogee. These values have to be

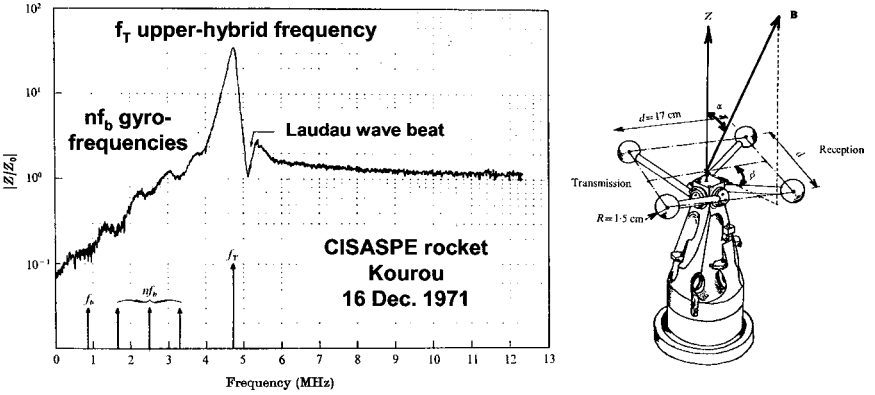


Fig. 3. Quadripole impedance probe on the ionospheric rocket CISASPE and modulus of the mutual impedance normalized with respect to its value in vacuum, at 227 km altitude.[6]

compared with the sphere size, 3 cm in diameter, and the sensor size, a square of 17 cm. The antenna length was therefore much larger than the Debye length, which guarantees reliable measurements. The sphere size is larger that the Debye length, i.e., measurements cannot be considered as point measurements, it has no severe consequence provided the sphere size is much smaller than the antenna size.

The GEOS-1 spacecraft carried the mutual impedance instrument S304 to determine characteristics of the cold plasma component in the Earth's magnetosphere. The two receiving spheres were separated by a distance of 42 m (Fig. 4). The two transmitting spheres were along a line parallel to the receiving antenna, 3.3 m away from the spacecraft, and they made a 1.4 m long dipole.[7] Most of the GEOS-1 data were recorded in a quasi-equatorial plane from 4 to 7 R_E. It allowed the Earth's magnetosphere to be extensively studied, and in particular the plasmasphere. Due to the size of the

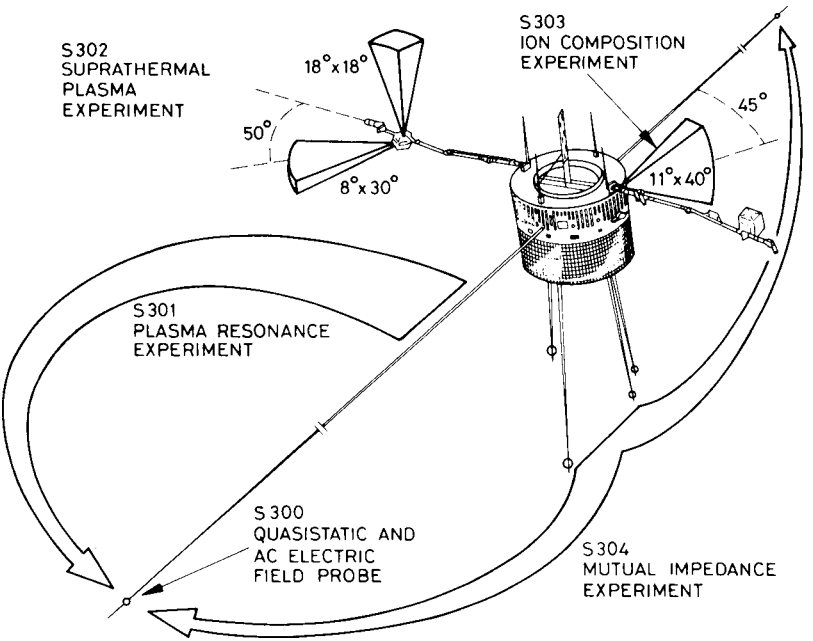


Fig. 4. Schematic view of the GEOS-1 satellite showing the locations of some of the onboard plasma and wave instruments. The S304 mutual impedance is composed of two transmitting 10 cm diameter wire mesh spheres, located at the tips of 3.3 m long axial booms and separated by a distance of 1.4 m, and two receiving 8 cm diameter spherical spheres, mounted on radial booms of length 42 m tip to tip.[7]

mutual impedance sensors, the N_e and T_e might be determined so long as the Debye length was in the 0.5–6 m range. The mutual impedance probe onboard GEOS-1 has performed the first reliable radial profile of the electron temperature in the outer plasmasphere, with L varying from 3 to 6. Here, the thermal plasma is rather cold, $T_e \leq 1$ eV (3,000–10,000 K).[8] The sphere size, 8–10 cm in diameter, was much lower than the Debye length, 0.5–5 m. In most cases, the transmitter–receiver distance, about 20 m, was much larger than two times the Debye length λ_D, which is required for a reliable thermal plasma diagnostic. Let us remark that the term short antenna is still valid for such a quite long antenna, because antennae are here considered as short whenever their lengths are much smaller than the wavelengths of electromagnetic waves so that the quasi-static approximation remains valid. They are also not too long compared with the characteristic lengths of the plasma, and in particular the Debye length.

ISOPROBE onboard AUREOL-3 (ARCAD-3 project) was designed to investigate thermal plasmas in the Earth's ionosphere and magnetosphere. The spacecraft was placed into a quasi-polar orbit in September 1981 (410–2,010 km altitude). As T_e measurements appear to be the most accurate whenever $10\,\lambda_D < D < 40\,\lambda_D$, where D is the distance between transmitting E and receiving R electrodes, 3 E and 2 R electrodes were implemented to cover all possible Debye lengths.[9] The sensor and an example of mutual impedance modulus recorded during the flight are shown in Fig. 5.

A quadrupolar array of ring electrodes has been implemented aboard the CASSINI/HUYGENS probe to measure the Titan atmosphere permittivity as well as the ground surface material permittivity.[10,11] Two of these electrodes could also be used as an electric antenna to detect atmospheric electromagnetic waves, such as those that might be produced by lightning. The electrodes were mounted on a pair of deployable booms in order to minimize the shielding effects of the HUYGENS probe body (Fig. 6). As usual, the mutual impedance was measured by applying a sine wave pulse with amplitude 0.1–10 V at fixed frequencies, from 45 Hz to 5.7 kHz, on the two transmitting ring electrodes that are mounted on the two booms, on both sides of the Huygens probe. The amplitude and phase of impedance were then computed by fast fourier transform using the analysis of the differential signal measured between the two receiving ring electrodes also mounted on the two booms. As the electrodes were quite close (40 cm) to the meter sized conductive body of the HUYGENS probe, it is worth noting that the influence of the latter on the mutual impedance, and therefore the permittivity determination, had to be taken into account.

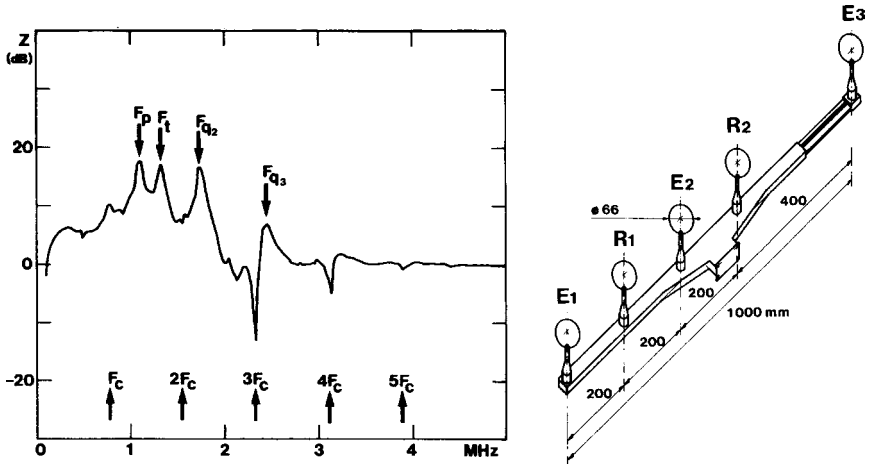


Fig. 5. Impedance modulus (*left*) measured at 1475 km altitude and 44° invariant latitude by the ISOPROBE quadripole probe (*right*).[9] Anti-resonances occur at gyroharmonics nF_c, while peaks are observed at the plasma frequency F_p, the upper-hybrid frequency F_t, and maximum frequencies of the Bernstein's modes F_{qn}.

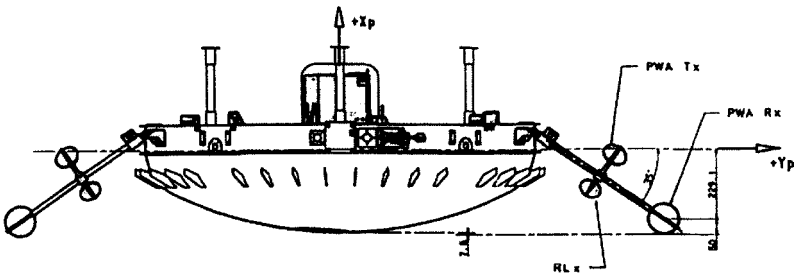


Fig. 6. Mutual impedance probe accommodation on the CASSINI/HUYGENS probe. The two transmitting (Tx) and two receiving (Rx) ring electrodes form a trapezoid in a plane that contains the HUYGENS probe vertical axis. The mutual impedance probe is one of the permittivity, wave and altimetry, PWA, sensors.[10]

4. Conclusion

Impedances of short electric field antennae versus frequency are directly related to plasma characteristics, it therefore allows electron plasma density and temperature to be reliably and accurately determined. Let us consider for example a one (or two) meter long antenna whose closest transmitting/receiving electrodes distance is 40 cm (80 cm) driven by a 2 kHz–2 MHz frequency synthesizer, the plasma density, Debye length,

and temperature ranges that can be determined for such an antenna are respectively: 1–4 $\times$ $10^4\,cm^{-3}$, 0.3–20 cm, 10–2 $\times$ 10^5 K (1–4 $10^4\,cm^{-3}$, 0.5–40 cm, 40–8 $\times$ 10^5 K). Perturbation effects are eliminated by measuring the transfer impedance between two pairs of small electrodes well separated from each other (quadripole probe). Recent theoretical works have been done.[12–16] This technique has been successfully used onboard rockets and many spacecraft (GEOS-1,[7,8] Swedish satellite VIKING,[17] ARCAD/AUREOL-3,[9] MARS-96,[18] CASSINI/HUYGENS lander.[10,11] It was proposed[19] for the Mars Express mission. It is actually a part of the ROSETTA plasma consortium[20] and a similar technique will be implemented inside the active measurement of the Mercury's plasma (AM^2P) instrument which is a part of the plasma wave investigation (PWI) wave consortium of the BepiColombo/MMO mission to Mercury.[21]

References

1. N. Vernet, R. Manning and J. L. Steinberg, *Radio Science* **10** (1975) 517.
2. K. Hashimoto, I. Nagano, T. Okada, M. Yamamoto and I. Kimura, *Geophys. Rev. Lett.* **18** (1991) 313.
3. M. Tsutsui, I. Nagano, H. Kojima, K. Hashimoto, H. Matsumoto, S. Yagitani and T. Okada, *Radio Sci.* **32** (1997) 1101.
4. L. R. O. Storey, P. Meyer and P. Aubry, in *Plasma Waves in Space and in the Laboratory*, eds. J. O. Thomas and B. J. Landmark, vol. 1 (Edinburgh University Press, Edinburgh, Scotland, 1969), p. 303.
5. G. Belmont, P. Canu, J. Etcheto, H. de Féraudy, B. Higel, R. Pottelette, C. Béghin, R. Debrie, P. M. E. Décréau, M. Hamelin and J. G. Trotignon, in *Proc. Conf. Achievements of the IMS, 26–28 June 1984, Graz, Austria,* ESA SP-217 (September 1984), p. 695.
6. J. M. Chassériaux, R. Debrie and C. Renard, *J. Plasma Phys.* **8** (1972) 231.
7. P. M. E. Décréau, J. Etcheto, K. Knott, A. Pedersen, G. L. Wrenn and D. T. Young, *Space Sci. Rev.* **22** (1978) 633.
8. P. M. E. Décréau, C. Béghin and M. Parrot, *J. Geophys. Res.* **87** (1982) 695.
9. C. Béghin, J. F. Karczewski, B. Poirier, R. Debrie and N. Massevitch, *Ann. Geophys.* **38** (1982) 615.
10. R. Grard, H. Svedhem, V. Brown, P. Falkner and M. Hamelin, *J. Atmos. Terr. Phys.* **57** (1995) 575.
11. M. Hamelin, O. Aydogar, G. Bianchini, V. Brown, P. Falkner, M. Fulchignoni, I. Jernej, J. M. Jeronimo, R. Grard, J. J. Lopez-Moreno, G. Molina-Cuberos, R. Rodrigo, K. Schwingenschuh and H. Svedhem, *Adv. Space Res.* **26**, 10 (2000) 1697.
12. C. Béghin, *Radio Sci.* **30** (1995) 307.
13. C. Béghin and E. Kolesnikova, *Radio Sci.* **33** (1998) 503.
14. E. Kolesnikova and C. Béghin, *Radio Sci.* **33** (1998) 491.

15. J. Geiswiller, J. G. Trotignon, C. Béghin and E. Kolesnikova, *Astrophys. Space Sci.* **277** (2001) 317.
16. J. Geiswiller, C. Béghin, E. Kolesnikova, D. Lagoutte, J. L. Michau and J. G. Trotignon, *Planet. Space Sci.* **49** (2001) 633.
17. P. M. E. Décréau, M. Hamelin, R. Massif, H. de Féraudy, E. Pawela, S. Perraut, R. Pottelette and A. Bahnsen, *Ann. Geophys.* **5A** (1987) 181.
18. ELISMA experimenters, *Planet. Space Sci.* **46** (1998) 701.
19. J. G. Trotignon, H.-C. Séran, C. Béghin, N. Meyer-Vernet, R. Manning, R. Grard and H. Laakso, *Planet. Space Sci.* **49** (2001) 155.
20. J. G. Trotignon, R. Boström, J. L. Burch, K.-H. Glassmeier, R. Lundin, O. Norberg, A. Balogh, K. Szegö, G. Musmann, A. Coates, L. Ahlén, C. Carr, A. Eriksson, W. Gibson, F. Kuhnke, K. Lundin, J. L. Michau and S. Szalai, *Adv. Space Res.* **24**, 9 (1999) 1149.
21. J. G. Trotignon, C. Béghin, D. Lagoutte, J. L. Michau, H. Matsumoto, H. Kojima, K. Hashimoto, Y. Kasaba, L. G. Blomberg, J. P. Lebreton, A. Masson, M. Hamelin and R. Pottelette, Active measurements of the thermal electron density and temperature from the Mercury Magnetospheric Orbiter of the BepiColombo mission, in press in *Adv. Space Res.* (2004).

This page intentionally left blank

INFRARED HIGH-RESOLUTION SPECTROSCOPY OF PLUTO BY SUBARU TELESCOPE

TAKANORI SASAKI*, MASATERU ISHIGURO, DAISUKE KINOSHITA
and RYOSUKE NAKAMURA
Department of Earth and Planetary Science, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
** takanori@eps.s.u-tokyo.ac.jp*

Here we report a infrared high-resolution spectroscopy of Pluto in the L band. The spectroscopic observation was performed by the Subaru telescope with the adaptive optics system. The spectrum is dominated by the strong and broad absorption features of methane, but includes some additional features. Comparing the spectrum with model calculations, we suggest that absorption features could be an indication of nonmethane hydrocarbons on Pluto's surface.

1. Introduction

Spectroscopic observations of Pluto at visible to 2.5 μm have showed absorption features of solid methane and carbon monoxide diluted by nitrogen ice, and recently intimated the existence of ethane ice on surface of Pluto.[1,2] Only a few groups obtained photometric observations of Pluto at 2.8–4.1 μm range, where they found CH_4 ice absorption bands as well as other features which were attributed to CO_2 and/or SO_2 ices.[1] However, the spectral resolution of their data was insufficient confirm the existence of hydrocarbons' ice, because of difficulties deriving precise spectra at wavelengths longer than 2.5 μm with increasing telluric sky brightness. In this study, we report the additional composition of Pluto's surface from infrared spectroscopy at 2.8–4.0 μm conducted at Subaru telescope with infrared camera and spectrograph (IRCS)[3] and adaptive optics (AO).[4] We present a spectroscopic analysis of this observation as well as numerical modelling of the spectra.

*Corresponding author.

2. Observations and Data Reduction

A near-infrared spectroscopic observation of Pluto was performed by the 8-m Subaru telescope with IRCS in conjunction with its AO system on May 28, 2002 (UT) near the opposition (RA = 17 h 05 m 33 s, dec = 12°39′25″). The sub-Earth longitude was 40–50° and the separation from Charon was 0.9″. Spectral resolving power was approximately 400 at the L band in our observation, meanwhile that was about 60 in Ref. 1. The typical psf width was 0.3–0.4″ during the observation and the total on-source integration time was 2600 s. For the cancellation of telluric absorption features, a reference star (G3V star SAO141540) was observed just after Pluto observation. The air mass differences between Pluto and the reference star were less than 0.035 throughout the observation.

We used NOAO IRAF astronomical software package to reduce near-infrared spectra obtained by IRCS. Argon lamp frames were used for the wavelength calibration. Since Pluto was clearly separated from Charon with the FWHM as sharp as 0.4″, the resulting spectrum includes no contamination from Charon. For details refer to Ref. 5.

3. Results

The spectrum of Pluto is shown in Fig. 1 along with the previous low resolution data[1] and synthetic spectra of simple ternary intimate mixture of N_2-CH_4-CO. As Nakamura *et al.*,[2] we employed Hapke's bidirectional model[6] to calculate the spectrum of a uniform half-infinie layer covering the whole surface of Pluto. It should be noted that the detailed model calculations are beyond the scope of this paper because we cannot account for the realistic solid solutions, vertically layered structure and spatially segregated patches unlike some previous approaches. Apparently, the shape of the synthetic specta is dominated by CH_4 because N_2 and CO have no significant absorptions in the wavelength range.

Grundy *et al.*[1] found a decrease of the reflectance lower than 3.2 μm and beyond 3.95 μm. They interpreted the result as the contribution from non-volatile ices, such as H_2O for the former, SO_2 and CO_2 for the latter. While we cannot confirm the putative band of SO_2 and CO_2 due to the limited wavelength coverage, our spectra shortward of 3.2 μm are consistent with their results.

Our observations show lower reflectance around 3.45 μm and additional absorptions around 3.1, 3.2, and 3.35 μm. In order to reproduce the features,

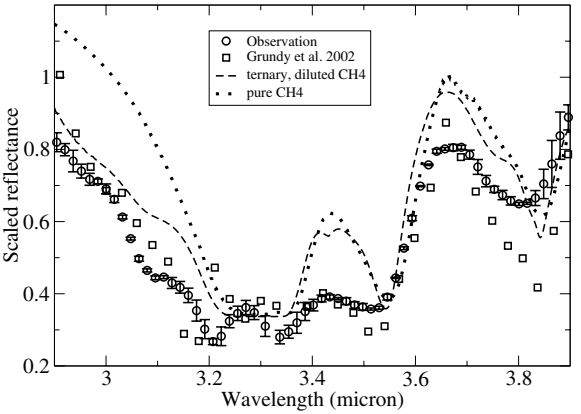


Fig. 1. Reflectance spectrum of Pluto with the previous low-resolution data and synthetic spectra of intimate ternary mixture of N_2-CH_4-CO with the mass ratio 1:0.01:0.002. The circles represents our observation smoothed by taking a running average of 31 pixels (0.08 μm) with their error bars of standard deviation. The dotted and dashed curves indicate synthetic spectra for pure CH_4 and for CH_4 diluted in the solid molecular nitrogen, respectively. We normalized all the data at 3.58 μm and the synthetic spectra were convolved to the same wavelength resolution as the observation.

we incorporated some non-methane hydrocarbons, namely C_2H_2, C_2H_4, C_2H_6, and C_3H_8 into the model calculations. Assuming their mass ratio to CH_4 as 10%, we computed the model spectra in Fig. 2. The optical constants were measured by Quirico *et al.*[7] Although HCN, CH_3OH, and H_2CO are also found in both interstellar and cometary ices, the optical constants for them are not available. In Fig. 2, CH_4 is assumed to be diluted in the solid molecular nitrogen. It is found that C_2H_2, C_2H_4, and C_2H_6 would produce the 3.1 μm absorption if their relative concentration to CH_4 is approximately 10 wt%. This value should be regarded as the upper limit for C_3H_8 because the absorption at 3.0 μm is not seen in the observed spectrum. Figure 2 indicates that C_2H_4 and C_2H_6 could be associated with the absorptions around 3.2 and 3.35 μm.

Adding C_2H_4 as the fourth component, we obtained a better agreement with the observations than the simple ternary mixture model, but there still remains discrepancy between 3.4 and 3.55 μm. On the other hand, C_2H_6 improves the fit in this range, although we see no clear absorption around 3.65 μm in the observed spectrum. So we plotted two synthetic spectra (ternary + C_2H_6 and ternary + C_2H_6 + C_2H_2) in Fig. 3 for the detailed

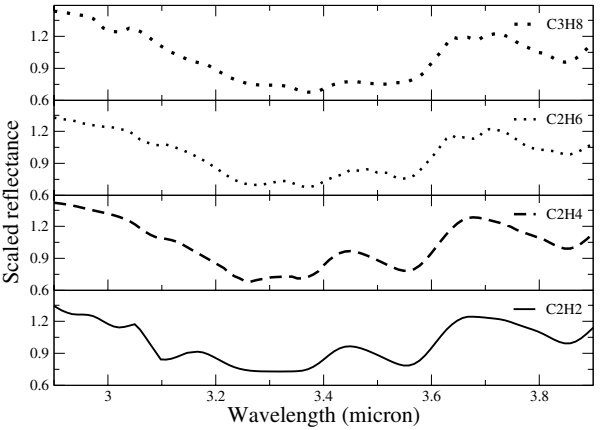


Fig. 2. Modeled spectra including nonmethane hydrocarbons (NMHCs) as the fourth component. The fourth components are C_2H_2, C_2H_4, C_2H_6, and C_3H_8. The model parameters of the basic ternary mixture are the same as those in Fig. 1, and the mass fraction to CH_4 of the fourth component is 0.1.

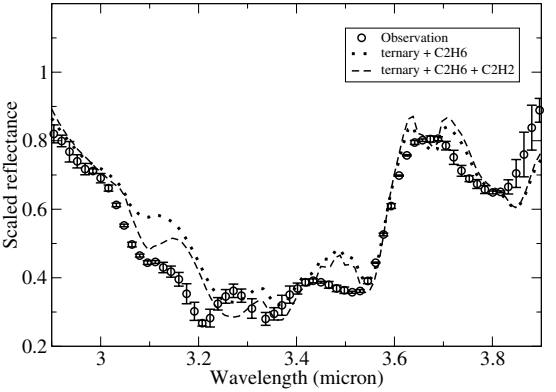


Fig. 3. Reflectance spectrum of Pluto with the modeled spectra. The circles show our observation as shown in Fig. 1. The dashed and dotted curves represent the synthetic spectrum with C_2H_6:CH_4 = 1:10 and C_2H_2:C_2H_6:CH_4 with mass ratios 1:1:10, respectively.

comparison with the observed spectrum. In both cases, the observed absorptions around 3.2 and 3.35 μm shifts toward shorter wavelengths compared with the model spectra. Quirico and Schmitt[8] have shown that the dilution in solid N_2 ice cause the shift of hydrocarbons' absorptions. More laboratory experiments are needed to explore the dilution effect more quantitatively.

Adding C_2H_2 as the fifth component in the model calculation makes the spectrum match the observed spectrum quite well, as shown in Fig. 3.

4. Discussion

In the the protoplanetay disk and the interstellar cloud that proceeded it, the gas-phase ion-molecule reactions could yield C_2H_2. The reaction of H with C_2H_2 on cold dust grains could produce C_2H_6 more effectively than C_2H_4.[9] In fact, C_2H_2 and C_2H_6 have been found in the comas of Oort-cloud comets,[10,11] but the detection of C_2H_4 has not been reported so far. The C_2H_6/CH_4 ratio estimated from Fig. 3 is consistent not only with the Oort-cloud comets but also with a value of a short-period comet 21P/Giacobini-Zinner[12] and the upper limit for interstellar materials.[13] Moreover, the C_2H_2/CH_4 ratio approximately agrees with the values for Oort-cloud comets.[10,14]

Non-methane hydrocarbons could be the secondary products generated from methane. Krasnopolsky and Cruikshank[15] simulated photochemical reactions in Pluto's tenuous atmosphere and the subsequent precipitation to the surface. C_2H_2 and C_4H_2 have the highest precipitation rates among numerous photochemical products, but C_2H_2 cannot produce the observed 3.2 and 3.35 μm features and we do not have the optical constants for solid C_4H_2. The precipitation rates of C_2H_4 and C_2H_6 are smaller than those of C_2H_2 and C_4H_2 by an order of magnitude. If the 3.2 and 3.35 μm features are indication of C_2H_4 and/or C_2H_6 produced by the photochemical reactions, we must increase their relative concentrations to C_2H_2 and C_4H_2 somehow. Moore and Hudson[16] conducted the systematic infrared study of proton-irradiated N_2-rich ices containing CH_4 and CO. They found C_2H_2, C_2H_4, C_2H_6, and C_3H_8 along with HCN and HNC in the irradiated analog materials of Pluto's surface. It should be emphasized that they identified C_2H_6 with the absorption at 3.35 μm but the expected absorption around 3.65 μm was not detected as our observation. In addition, HCN is expected to decrease the gap between the models and observation around 3.1 μm. Therefore, we do not rule out possibilities that observed features were associated with non-methane hydrocarbons produced by the *in situ* surface reactions induced by cosmic-ray irradiation to the original ternary mixture of N_2, CH_4, and CO. In either case, thin deposits are vulnerable to micrometeorite impact. The relative mass ratio to the parent methane, derived from our observations, could be a key to understanding the gardening process on Pluto, such as poorly known dust flux[17] and vertical mixing timescale.[15]

Acknowledgments

We thank Dr. E. Quirico for providing us with the optical constant data of various ices in electronic form. We also thank the anonymous reviewer for a constructive review.

References

1. W. M. Grundy, M. W. Buie and J. R. Spencer, *AJ* **124** (2002) 2273.
2. R. Nakamura *et al.*, *PASJ* **52** (2000) 551.
3. N. Kobayashi *et al.*, *Proc. SPIE* **4008** (2000) 1056.
4. H. Takami *et al.*, *SPIE* **3353** (1998) 512.
5. T. Sasaki *et al.*, *ApJ* **618** (2005) L57.
6. B. Hapke, *Theory of Reflectance and Emittance Spectroscopy* (Cambridge University Press, Cambridge, 1993).
7. E. Quirico *et al.*, *Icarus* **139** (1999) 159.
8. E. Quirico and B. Schmitt, *Icarus* **127** (1997) 354.
9. K. Hiraoka *et al.*, *ApJ* **532** (2000) 1029.
10. M. J. Mumma *et al.*, *Science* **272** (1996) 1310.
11. H. Kawakita *et al.*, *ApJ* **590** (2003) 573.
12. M. J. Mumma *et al.*, *ApJ* **531** (2000) L155.
13. N. Boudin *et al.*, *A & A* **331** (1998) 749.
14. T. Y. Brooke *et al.*, *Icarus* **166** (2003) 167.
15. V. A. Krasnopolsky and D. P. Cruikshank, *J. Geophys. Res.* **104** (1999) 21979.
16. M. H. Moore and R. L. Hudson, *Icarus* **161** (2003) 486.
17. A. Krivov *et al.*, *Planet. Space Sci.* **51** (2003) 251.

UNDERSTANDING THE ORIGIN OF THE ASTEROIDS THROUGH THE STUDY OF VESTA AND CERES: THE ROLE OF DAWN

ALBERTO CELLINO[*,**], FABRIZIO CAPACCIONI[†,††], MARIA TERESA CAPRIA[†,‡‡], ANGIOLETTA CORADINI[†,§§], MARIA CRISTINA DE SANCTIS[†,¶¶], HORST U. KELLER[‡,‖‖], THOMAS H. PRETTYMAN[§,***], CAROL A. RAYMOND[¶,†††] and CHRISTOPHER T. RUSSELL[‖,‡‡‡]

[*]*INAF — Osservatorio Astronomico di Torino*
Strada Osservatorio 20, 10025 Pino Torinese, Italy
[†]*INAF — IASF, via del Fosso del Cavaliere 100, 00133 Roma, Italy*
[‡]*Max Planck Institute for Aeronomie, Germany*
[§]*Los Alamos National Laboratory, Los Alamos, New Mexico, USA*
[¶]*Jet Propulsion Laboratory, Pasadena, California, USA*
[‖]*University of California, Los Angeles, USA*

[**]*cellino@to.astro.it*
[††]*Fabrizio.Capaccioni@rm.iasf.cnr.it*
[‡‡]*MariaTeresa.Capria@rm.iasf.cnr.it*
[§§]*Angioletta.Coradini@rm.iasf.cnr.it*
[¶¶]*MariaCristina.Desanctis@rm.iasf.cnr.it*
[‖‖]*keller@linmpi.mpg.de*
[***]*tprettyman@lanl.gov*
[†††]*craymond@mail1.jpl.nasa.gov*
[‡‡‡]*ctrussell@igpp.ucla.edu*

The Dawn mission is aimed at providing a wealth of fundamental information concerning the two giant asteroids (1) Ceres and (4) Vesta. According to current ground-based data, these two objects are thought to be radically different, and offer a nice example of the huge variety of physical properties characterizing the main belt asteroid population. In particular, (1) Ceres, the largest main belt asteroid, is thought to be a mostly primitive object, whereas (4) Vesta is the best example of a fully differentiated asteroid. A major goal of modern planetary science is to understand how and why these two objects, which are orbiting at fairly similar heliocentric distances, can be so different. In turn, this question is related to our understanding of the early phases in Solar Systems history, when planetary growth took place in the protoplanetary disk. *In situ* exploration of Ceres and Vesta by Dawn is expected to provide new fundamental data, and to possibly open new paths to theoretical investigations.

1. Introduction

In spite of the tremendous improvement in our understanding of the physical properties and overall histories of the asteroids achieved in recent years through major theoretical advancements and detailed observational data, from both remote and *in situ* investigations, several problems are still open. The asteroids are bodies of fundamental importance for a number of reasons. First, they are the outcomes of a process of partly aborted planetary growth. As such, they are bodies in some sense intermediate between fully grown major planets, and the original planetesimals from which planets were originated. The asteroids are a strongly heterogeneous population, and can be classified into a number of distinct taxonomic classes. At least some of these classes are thought to include very primitive objects, likely identical in mineralogy to the most primitive meteorites, the carbonaceous chondrites, which are samples of the oldest material present today in the Solar System. Other taxonomic classes are thought to correspond to bodies that, during their histories, experienced different degrees of thermal evolution, ranging from short episodes of mild heating, to complete melting and differentiation. Given this large diversity, asteroids are very important remnants of the early epochs of the Solar System history, and represent a strong constraint for the theoretical models of formation of our planetary system. In addition to purely theoretical considerations, asteroids are also important from more pragmatic points of view. In particular, it is now widely recognized that episodes of catastrophic impacts with asteroidal bodies have likely determined big changes in the evolution of the Earth's biosphere in the past. Understanding the inventory and size distribution of the potentially hazardous objects is a task to whom large resources have been allocated in recent years in several countries. It is clear, however, that also the basic physical properties of the asteroids, including their overall composition and their internal structure, must be understood in order to be able to develop any credible strategy of mitigation of the impact hazard. Finally, it should also be mentioned that asteroids may be in the future an important source of useful materials for the development of space activities and the survival of human settlements in the inner solar system.

The above considerations are sufficient to explain why asteroids have been and are extensively studied. In addition to theoretical investigations and remote-sensing observations, in recent years a major breakthrough has been obtained by means of direct *in situ* investigations by means of space probes. The first data came from missions that had other bodies as their

main targets, as in the case of the first asteroid flybys performed by the Galileo space probe. More recently, a number of space missions have been and are designed specifically for investigations of asteroidal targets. The most recent fully successful mission of this kind has been accomplished by the NEAR-Shoemaker probe,[1] which has performed an extensive exploration of the near-Earth asteroid 433 Eros. Another mission which is currently active and has produced spectacular results is the Japanese Hayabusa mission, which touched down on the near-Earth asteroid 25143 Itokawa, but is not thought to have succeeded in its final goal of collecting a surface sample, to be brought back to Earth. In this paper, we briefly describe another space mission devoted to asteroids, that is currently under development, and is supposed to be launched in 2006. This mission has been named Dawn, and will have as specific targets, for the first time, two of the largest main-belt asteroids, (1) Ceres and (4) Vesta. The main goal of this mission will be to solve the "Ceres–Vesta paradox", but more in general it will be of primary importance to obtain a better understanding of the early phases of planetary accretion in our Solar System. The rest of this chapter is devoted to give a brief explanation of the "Ceres–Vesta paradox", to explain how Dawn will be important to solve it, and to give a general description of the scientific payload of the probe, and its planned trajectory from the Earth to its targets, as well as the current status of the mission preparation activities.

2. The Ceres–Vesta Paradox

What we call the "Ceres–Vesta paradox" can be easily understood if we first make a few preliminary assumptions concerning the mechanisms of planetary growth and early heating of planetary bodies. In the first part of this section we will describe the theoretical scenario as it has been until a few months ago, when some new important developments have been proposed by some authors. These new ideas will be discussed at the end of this section. Let us then assume that (1) in the process of planetary growth the biggest planetary embryos grow faster; (2) the fundamental sources of heating in early planetary bodies were impact heating and the decay of short-lived radiogenic nuclei like Al^{26}; (3) The largest bodies produced by planetary growth, specially in the inner Solar System, likely preserved their relative heliocentric distances over the age of the solar system.

All the above assumptions appear to be very reasonable according to currently accepted models of Solar System formation, based on a large

body of theoretical investigations and numerical simulations published in the literature. Now, look at the real bodies that are today present in the asteroid main belt, and in particular focus on the two "giant" asteroid (1) Ceres and (4) Vesta. Having a diameter of about 1,000 km (Ceres) and 500 km (Vesta), these two asteroids are indeed giant with respect to the rest of the asteroid population, apart from a few bodies like (2) Pallas or (10) Hygiea. In fact, asteroids like Ceres and Vesta are fundamentally different with respect to normally-sized asteroids in several respects, and are in some sense real "planets". A fundamental property which is shared by both Ceres and Vesta is that they have sizes (and consequently masses) that make them practically indestructible by the process of collisional evolution which steadily grinds the rest of the asteroid population into smaller pieces. Not surprisingly, the overall shapes of Ceres and Vesta are mostly oblate spheroids (with some more pronounced triaxiality in the case of Vesta), compatible with the fact that self-gravitation starts to be dominant for objects of this size.

If we consider the current observational evidence about Ceres and Vesta, however, we easily realize that we could hardly imagine two asteroids more different with respect to each other. Ceres belongs to the G taxonomic class, which is related to the wider C complex, that dominates the asteroid population, specially in the outer regions of the belt. Its reflectance spectrum is mostly flat and featureless, and is diagnostic of a primitive composition, likely similar to that of carbonaceous chondrites. Some observations have suggested the presence of hydrated minerals on its surface, compatible with an overall primitive composition.[2] Bodies like Ceres are generally interpreted in terms of a mineralogical composition that has never been affected by substantial heating episodes. Ceres should be therefore strongly reminiscent of the overall composition of the planetesimals located in the region where it was grown.

In contrast, Vesta is a classical example of an asteroid whose reflectance properties are clearly diagnostic of an overall basaltic crust. As a consequence, Vesta is thought to be a fully differentiated body, including a metallic core, a mantle and a lighter crust. Bodies like this must have been produced by early planetary growth at these heliocentric distances, to justify the existence of iron meteorites and asteroids having a mostly metallic composition (at least a fraction of today's M-type asteroids), interpreted as exposed cores of differentiated parent bodies disrupted by collisions and consistent with meteoritic evidence. The taxonomic classification of Vesta, the so-called V-type, has long been unique among asteroids. The reflectance

spectrum of Vesta turns out to be very similar to achondrite meteorites having an overall basaltic composition (HED meteorites). Only after the reliable identification of a dynamical family associated to Vesta,[3] it was shown that the other members of the family share the same reflectance properties of Vesta.[4] More recently, a number of V-type objects have been discovered also among the population of near-Earth asteroids. As a consequence, it is now widely believed that the Vesta family was produced by a violent impact on Vesta's surface, which produced a very large crater and excavated a substantial amount of material, that escaped the gravitational well of Vesta and produced a swarm of fragments that we see today as family members. The presence of an hemispheric-sized crater on Vesta's surface had been first suggested by polarimetric observations,[5] and had been also confirmed by an extensive analysis of available lightcurves of the asteroid.[6] The crater has been more recently directly seen by means of HST images.[7] It is also believed that Vesta is the origin of a number of other small objects which have been found to belong to the V taxonomic class, and are located in the space of orbital elements in the region between Vesta and the locations of the 3/1 mean-motion resonance with Jupiter, and of the so-called ν_6 secular resonance. These resonances are known to be efficient transfer pathways to the region of the inner planets. Thus, it is believed that fragments of Vesta can reach these resonances under the effect of dynamical evolution likely dominated by the Yarkovsky effect.[8] It is now commonly accepted that V-type near-Earth asteroids and HED meteorites originate from Vesta, although some HED meteorites exhibit a variety of spectral features that are not immediately understood in terms of a common origin from a unique parent body.[9]

What we call the "Ceres–Vesta" paradox, is the fact that, if we accept the (1)–(3) assumptions quoted above, it is very difficult to explain why and how the properties of Ceres and Vesta can be as they appear to be from observations. In particular, what is intriguing is that, among these two bodies, the bigger one seems to be primitive, whereas the smaller one is fully differentiated. If we accept assumption (2), Ceres, being twice as large as Vesta and having a similar composition, should have contained in its interior an amount of radiogenic nuclei more than sufficient to completely melt it. But in this case we should expect also for Ceres a surface of basaltic composition. One could explain this apparent paradox by assuming that the growth of Ceres took place later with respect to Vesta, and in particular when a substantial fraction of the short-lived radiogenic nuclei had already decayed. This possible explanation, however, is in conflict with

our assumption (1), because it is very difficult to imagine how the growth of Ceres, which is much larger than Vesta, might have been slower, taking also into account that the two objects accreted at approximately the same heliocentric distance, and we do not expect that Ceres underwent a significant migration since the time of its accretion (assumption 3). Moreover, as we will see below, there are good reasons to believe that the melting of Ceres would be in any case unavoidable, due to the presence of a substantial amount of long-lived radionuclides in its original composition.

There are not many possible explanations of this paradox, that can be reconciled with both current understanding of planetesimal growth and the available observational evidence. A possibility compatible with our assumptions (1)–(3) is that in the region of Ceres the original planetesimals were significantly different in composition with respect to those located in the region of Vesta. In particular, if the gradient in composition of the solid material in the protoplanetary dust disk was sufficiently sharp, the planetesimals located in the region of growth of Ceres, at about 2.8 AU, could have included a substantially higher content of volatiles with respect to the planetesimals that were located in the region of growth of Vesta, around 2.35 AU from the Sun. If this hypothesis is correct, the amount of volatiles in the interior of Ceres might have been sufficient to efficiently dissipate the heat generated by short-lived radiogenic nuclei. This might be a possible solution of the paradox. In this scenario, what is needed is mainly a sufficiently strong gradient in composition of the protoplanetary disk, which may be plausible. Of course, in the absence of such kind of mechanism of heat dissipation by volatile elements, it is impossible to efficiently cool down a body like Ceres, which, being twice as large as Vesta, in principle requires longer times to dissipate its internal heat. Even accepting the presence of volatiles, however, to be able to produce at the end of the phase of heating dissipation a body having the properties of Ceres, is in any case not so straightforward.

Of course, we cannot exclude other possibilities to solve the "Ceres–Vesta" paradox. A first possibility might be that our assumption (2) is not completely correct. In particular, it has been suggested in the past that electric induction heating could have been very important during the early phases of the Solar System history, particularly when the Sun experienced a T-Tau phase.[10] If this is true, bodies located at smaller heliocentric distances might have been heated up more efficiently, and this might explain why Vesta, and not Ceres, was completely melt. This possible explanation, however, must face two major objections. First, according to current knowledge of this phenomenon, the strong solar wind developed

during a T-Tau phase, should have been strong mainly along a direction perpendicular to the plane of the protoplanetary disk.[11] Second, the strong dependence of the electric induction heating upon heliocentric distance should also be justified.

Finally, we cannot exclude that something in the current observational evidence concerning the properties of Vesta and Ceres might be wrong. This may be particularly true in the case of Ceres, because, in spite of its large size, we do not know really so much about it. In particular, currently accepted interpretations of its spectrophotometric properties might turn out to be wrong, and evidence of some kind of thermal evolution might conceivably be found in the future. This is of course a field in which direct *in situ* exploration could play a decisive role.

In particular, as mentioned at the beginning of this section, some new developments in our general understanding of Ceres have been produced by some very recent investigations. In particular, recent data on the shape of Ceres suggest the presence of a central mass concentration indicative of differentiation.[12] Moreover, the results of a new theoretical investigation[13] indicate that Ceres must necessarily have undergone extensive melting, an inevitable outcome given the cosmic abundances of long-lived radioactives. In this respect, neither Al^{26} nor electromagnetic induction heating would be needed, though they would enhance the overall outcome. The difference in thermal history between Ceres and Vesta would be entirely due to the variation in initial water content: Ceres would have had from the beginning a significant water (17–27%) while Vesta was essentially dry. Subsequent thermal evolution led to the present very different bodies. According to these ideas Ceres, due to its high water content, is expected to have a crust, a mantle and a core structure, the details of which depend upon initial composition and structure. The crucial question in this scenario is, again, why was there such a large water difference in the starting conditions for bodies which presumably formed quite close to each other.

In the case of Vesta, the basaltic composition of its crust seems to be based on a very solid ground of observational evidence. This does not mean, however, that there are not uncertainties in the interpretation of the data. As quoted before, in fact, the simple association Vesta — V-type asteroids — HED meteorites seems to be less straightforward than commonly thought, when a detailed analysis of the reflectance spectra of these different objects are taken into account.[9]

In the presence of such kind of theoretical problems and observational uncertainties it is clear that the possibility of carrying out detailed *in situ*

investigations must be considered as decisive to solve the apparent paradoxes and shed some light on processes active during the early stages of planetary growth. As explained above, a detailed assessment of the overall composition of Ceres and Vesta has also deep implications for our understanding of the compositional gradient of the early protoplanetary disk. For these reasons, Dawn has been conceived and is expected to be able to provide answers to the most important open questions mentioned in this Section. In particular, a crucial test of current ideas will be to find evidence or not of the presence of large amounts of water in Ceres' composition. The presence of water, if not directly detected, will be evidenced by the presence of hydrated minerals on the surface of Ceres. Moreover, an accurate determination of the gravitational field of Ceres will allow us to find evidence of the presence of a metallic core produced by thermal differentiation. For what concerns Vesta, a detailed mineralogical map of the surface, including layers at different depth exposed by the mentioned hemispheric crater that is believed to characterize the surface of this asteroid, will be of paramount importance to really understand the composition gradient of Vesta and the connections with achondrite meteorites. In what follows, we give a brief outline of the instrumental configuration of the probe, that has been chosen with the aim of maximising the scientific output of the mission.

3. Dawn: A General Outline

The Dawn mission is the result of the collaboration of the scientists and agencies of these countries: the US, Germany, and Italy. In the US the mission is led by the University of California, Los Angeles (UCLA), funded by NASA, managed by the Jet Propulsion Laboratory (JPL) with an instrument provided by Los Alamos National Laboratory. In Germany, the framing camera is being provided by the Max Planck Institut fur Sonnensystemforschung with assistance from the Deutsches Zentrum fur Luft-und Raumfahrt (DLR). In Italy the mapping spectrometer is provided by the Agenzia Spaziale Italiana (ASI), and oversight is provided by INAF. The Principal and deputy Principal Investigators are C. T. Russell (UCLA) and Carol Raymond (JPL). The UCLA is responsible for the overall direction of the project, as well as management of science operations, data, products, archiving, analysis and education and public outreach.

The space probe will be ion-propelled. The launch is now expected to occur in 2007 using a Delta 2925H launcher. Independent of the exact launch date Dawn uses a Mars gravity assist in 2009, and will reach Vesta

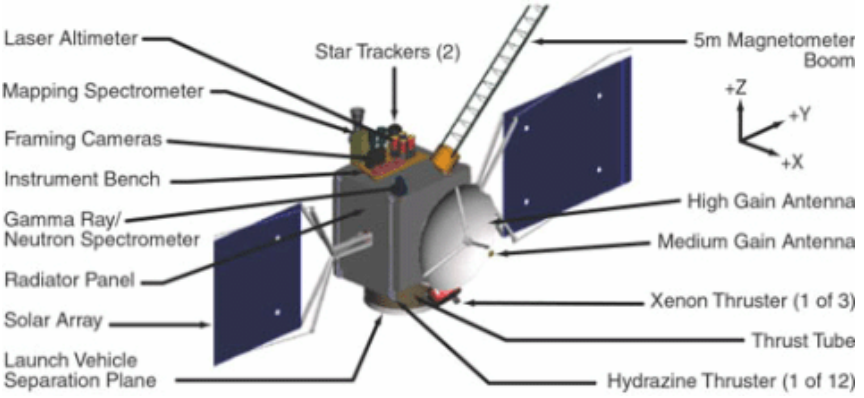


Fig. 1. A schematic illustration of the general design of the spacecraft.

in October, 2011. The departure from Vesta should be in May, 2012, and the arrival to Ceres is planned in August, 2015. The mission end will be in January, 2016. According to this schedule, Dawn will spend several months orbiting around Vesta and Ceres, and will have time to produce extensive investigations of these two asteroids.

A preliminary design of the Dawn spacecraft is shown in Fig. 1. The on-board scientific instruments will be two redundant framing cameras (FCs), produced by DLR/MPAe, a visible and infrared mapping spectrometer (VIR), produced by ASI, and a Gamma ray and neutron detector (GRaND), produced by the Los Alamos National Laboratory. In addition, the JPL will take charge of the gravity science data, from which the masses of the two target asteroids will be accurately determined.

The on-board instruments will permit to derive an impressive amount of data, including (1) a full surface imagery of Vesta and Ceres in at least three colors; (2) full spectrometric maps of the surfaces in three bands, namely 0.35–0.9, 0.8–2.5, and 2.4–5.0 μm; (3) abundances estimations for the Fe, Ti, O, Si, Ca, U, Th, K, H, Al, and Mg elements; (4) determination of the spherical harmonics development of the gravity field up to ninth degree for Vesta and fifth degree for Ceres.

Accordingly, we can expect that Dawn will be able to produce an impressive list of scientific results, including:

- A very accurate determination of the shape, size, composition and mass of both Ceres and Vesta.
- The determination of the internal structure, density and homogeneity of these two asteroids.

- From direct and comparative imaging of the surface morphologies, the determination of the cratering of the surfaces and an assessment of their overall geologic histories.
- The determination of the likely thermal history of the two bodies, including a reliable estimate of the size of the core.
- A new understanding of the role of water in controlling asteroid evolution.
- A reliable test of the current paradigm of Vesta as the HED parent body and the determination of which, if any, meteorites come from Ceres.
- The determination of a quantitative geologic context for HED origin.

4. Current Status of the Mission and Conclusions

From the above mentioned list of scientific results that we expect from Dawn, including but not limited to a definitive solution of the Ceres–Vesta paradox, it is evident that, if successful, this mission will mark an important milestone in the history of planetary science. In particular, we will obtain information on two bodies whose properties are directly reminiscent of the very early phases of planetary growth in our Solar System. Moreover, these two large asteroids are located in a region which marks a transition between the rocky terrestrial planets and the icy bodies of the outer Solar System. In this way, we have excellent reasons to hope that Dawn will provide a new wealth of quantitative information, to build our overall understanding of the origin of our planetary system on a much more solid background of observational evidence.

With the caveat that the launch and subsequent operation of a system as complex as a spacecraft entails much risk, it appears that the Dawn operations are going smoothly. Dawn is currently in the Assembly, Test, Launch, and Orientation (ATLO) phase and no real show-stoppers have been identified so far. The instruments have been delivered, and nearly all the hardware has been integrated to the spacecraft. However, the activities are currently (December 2005) in a state of stand down. This action was taken in response to concerns about the availability of funding in FY2006 to cover any problems that might arise during environmental and performance testing, particularly with regard to several pieces of subsystem hardware perceived to have experienced significant problems. As a consequence, the Dawn launch date, originally planned in mid-2006, moves now toward 2007. As for ourselves, we will obviously keep our fingers crossed not just until then but for many years to come!

Acknowledgments

Very useful reviews by D. R. Davis and E. F. Tedesco are kindly acknowledged. This work was partly supported by the Italian MIUR — COFIN (PRIN 2004).

References

1. A. F. Cheng, in *Asteroids III*, eds. W. F. Bottke, A. Cellino, P. Paolicchi and R. P. Binzel (University of Arizona Press, Tucson, 2002), p. 351.
2. L. A. Lebofsky, M. A. Feierberg, A. T. Tokunaga, H. P. Larson and J. R. Johnson, *Icarus* **48** (1981) 453.
3. V. Zappalà, A. Cellino, P. Farinella and Z. Knežević, *Astron. J.* **100** (1990) 2030.
4. R. P. Binzel and S. Xu, *Science* **260** (1993) 186.
5. J. E. Degewij, E. F. Tedesco and B. Zellner, *Icarus* **40** (1979) 364.
6. A. Cellino, V. Zappalà, M. Di Martino, P. Farinella and P. Paolicchi, *Icarus* **70** (1987) 546.
7. R. P. Binzel, M. J. Gaffey, P. C. Thomas, B. H. Zellner, A. D. Storrs and E. N. Wells, *Icarus* **128** (1997) 95.
8. W. F. Bottke, D. Vokrouhlický, D. P. Rubincam and M. Brož, in *Asteroids III*, eds. W. F. Bottke, A. Cellino, P. Paolicchi and R. P. Binzel (University of Arizona Press, Tucson, 2002), p. 395.
9. T. Hiroi, C. Pieters, F. Vilas, S. Sasaki, Y. Hamabe and E. Kurahashi, *Earth, Planets & Space* **53** (2001) 1071.
10. F. Herbert and C. P. Sonett, *Astrophys. & Space Sci.* **55** (1978) 227.
11. R. Mundt, in *Nearby Molecular Clouds* (Springer-Verlag, Berlin and New York, 1985), p. 160.
12. P. C. Thomas, J. W. Parker, L. A. McFadden, C. T. Russell, S. A. Stern, M. V. Sykes and E. F. Young, *Nature* **437** (2005) 224.
13. T. McCord and C. Sotin, *J. Geophys. Res.* (2005) 110.

This page intentionally left blank

THE EXPECTED ROLE OF GAIA FOR ASTEROID SCIENCE

ALBERTO CELLINO[*,‡], ALDO DELL'ORO[*,§]
and PAOLO TANGA[†,¶]
[*]*INAF — Osservatorio Astronomico di Torino*
I-10025 Pino Torinese (TO), Italy
[†]*Observatoire de la Côte D'Azur*
B.P. 4229 06304 Nice Cedex 4, France
[‡]*cellino@to.astro.it*
[§]*delloro@to.astro.it*
[¶]*tanga@obs-nice.fr*

According to current plans of ESA, Gaia will be launched in 2011. With a systematic survey of the whole sky down to magnitude $V = 20$, Gaia will provide a fundamental contribution in practically all fields of modern Astrophysics. In particular, Gaia will be also a major milestone in the history of asteroid science. Based on its unprecedented astrometric performances, complemented by spectroscopic and photometric capabilities, Gaia will be able to measure the masses of about 100 asteroids. It will directly measure sizes of about 1,000 objects, will derive spin properties and overall shapes of about 10,000 objects, and will derive much improved orbits and taxonomic classification of hundreds of thousands asteroids. The post-Gaia era in asteroid science will be one in which we will know average densities of about 100 objects belonging to all the major taxonomic classes, we will have a much more precise knowledge of the inventory and size and spin distributions of the population, of the distribution of taxonomic classes as a function of heliocentric distance, and of the dynamical and physical properties of dynamical families.

1. Introduction

In some respect, the exploration of the asteroids by means of *in situ* exploration by space probes seems to be an endless adventure. The reason is that in spite of impressive results that have been obtained in recent years after some successful space missions (Galileo, Stardust, NEAR-Shoemaker, and currently Hayabusa) it is clear that there is no hope to have in the near future a number of missions sufficient to explore all the large variety of physical properties (in terms of sizes, compositions, structures, thermal, collisional, and dynamical histories) which characterize this vastly heterogeneous population of minor bodies. At the same time, it is also true that, apart from a handful of objects that are fairly well known due to

in situ investigations, we know still so little about the most fundamental physical properties of these objects, including masses, sizes, and average densities.

The chances to improve significantly and in reasonably short times our overall knowledge of the physical properties of asteroids by means of remote observations are not so high, in spite of the imminent development of new instruments which are expected to produce large amounts of spectrophotometric data (Pan-STARRS, the Large-Aperture Synoptic Survey Telescope, etc.). These new observing facilities will be very important to achieve an improved knowledge of the overall inventory of Main Belt asteroids and near-Earth objects (NEOs). However, the capabilities of these surveys will be forcedly limited in their scopes (for instance, asteroid sizes cannot be measured by means of spectrophotometry) and will be affected by intrinsic limitations due to the simple fact of being ground-based. To make an example, several years of observations are usually needed to derive detailed information upon properties like the spin axis orientation by means of disk-integrated photometry, a task which is particularly suited to dedicated asteroid surveys. The reason is that spin axis orientation can be derived from photometric observations only when these cover a sufficiently large interval of observing geometries (aspect angles), which are usually obtained from observations corresponding to different oppositions of the same object. Moreover, as mentioned above, none of new observing facilities will be able to make direct measurements of fundamental physical parameters like size and albedo (surface reflectivity), not to mention asteroid masses, which will remain also essentially unknown in the future, but in cases of flybys by space missions, and/or of discovered binary systems.

For the above reasons, a primary objective of asteroid science today is to develop new tools to derive, possibly in relatively short times, reliable measurements of masses, sizes (and hence, average densities), in addition to reflectance and rotational properties, for a statistically significant sample of the whole population. The development of such facility(ies) would mark a real milestone in asteroid science, and would produce a real revolution in this field. In this paper, we show that Gaia, one of the cornerstone missions of the European Space Agency, can be this tool.

Gaia will be primarily an astrometric mission, reaching a level of unprecedented accuracy in the measurement of positions and proper motions of celestial bodies. In addition, Gaia will also have powerful photometric and spectroscopic capabilities. A large number of major advances in practically all fields of modern Astrophysics are expected to come from

this mission. In particular, Gaia will also produce a major breakthrough in our knowledge of the asteroids.

In what follows we will present a forcedly brief summary of the results of a number of investigations carried out since a couple of years by ourselves and by a number of other scientists from different European countries which have been active in the Gaia Solar System Working Group. Interested readers may find copies of interesting presentations made during different meetings of the Gaia SSWG at the URL address http:www.obs-nice.fr/tanga/SSWG/ (and links within it).

According to a large body of simulations carried out these years, Gaia astrometric data will be so accurate that asteroid orbital elements computed using Gaia observations alone, spanning over only 5 years, will be more precise than those resulting from all ground-based observations obtained over 200 years of ground-based astrometry, including also new observations that will be presumably obtained from the ground in the lapse of time between now and the end of the operational lifetime of the mission, expected to be launched in 2011. This apparently extravagant conclusion is fully justified by the unprecedented accuracy of Gaia astrometry, which is a major property of the overall Gaia concept. Simulations indicate that Gaia astrometric data will allow us to derive the masses of about 100 Main Belt asteroids, by means of measurements of tiny mutual perturbations occurring during mutual asteroid–asteroid close approaches. This particular topic is extensively discussed elsewhere,[1] and will not be further developed in the present paper. In addition to masses, the unprecedented astrometric accuracy of Gaia observations will also make it possible in many cases to measure the tiny deviations of the observed motion with respect to the predictions of newtonian dynamics, due to the first correction terms predicted by General Relativity. In particular, for the most eccentric orbits with small semi-major axes, like in the case of several near-Earth asteroids, the secular perihelion precession will be measurable and compared to the relativistic predictions. With respect to the current situation, in which relativistic effects have been measured only in the case of the planet Mercury, Gaia will extend the available measurements to several tens of objects. Finally, Gaia data are expected to provide also direct determinations of a Yarkovsky-driven drift in orbital semi-major axis,[2] another major advancement with further implications for studies of the thermal properties of asteroid surfaces and of the dynamical evolution of asteroid families.

Another major application of Gaia will be the direct measurement of sizes of objects having diameters down to a few tens of kilometers. Coupled

with the above-mentioned mass derivations, Gaia will produce a data-set of about 100 average densities of asteroids belonging to most taxonomic classes. This will be certainly a major milestone in the history of asteroid research. Major Gaia contributions will not come, however, uniquely from astrometric and high-resolution data. Very important applications will come also from disk-integrated spectrophotometric observations. An obvious application of the multi-band data collected by Gaia will be asteroid taxonomy. Another major application will be the derivation of spin properties and overall shapes for thousands of asteroids as a nice exploitation of Gaia disk-integrated photometric data. The present paper is mainly devoted to summarize the results of preliminary analyzes of these topics.

2. The Measurement of Sizes

The determination of asteroid sizes with Gaia will be obtained by means of a careful analysis of their generated signals in the Gaia focal plane. We give in what follows a general summary of the basic ideas and the algorithms that have been developed to attack the problem of asteroid size measurements, and we discuss the general results of a large body of numerical simulations.

An assessment of the capability of Gaia in determining asteroid sizes is based on a detailed simulation of the signals that will be produced by these objects. In particular, for the purposes of size determinations, we deal with the Gaia astrometric focal plane. We remind here that every object will be observed many times (typically several tens, according to current simulations) in different sky locations, during the operational lifetime of the mission. Asteroidal sources detected by Gaia will have two major properties: first, they will be (at least those above the resolution limit of the optics) extended sources; second, they will have an apparent and measurable motion across the Gaia focal plane during each single detection. We note that the measurement of this motion will be of primary importance for the purposes of asteroid orbit reconstruction, but this is not the topic of the present discussion.

The measurement of asteroids' apparent angular sizes will be possible, in principle, down to a limit that must be determined. The fact of dealing with moving objects, in this respect, is very important because this will influence the actual signals recorded by the Gaia detectors.

A detailed model of the expected asteroid signals is an obvious prerequisite to carry out any analysis of the predicted performances of Gaia in measuring asteroid sizes, and also to derive an estimate of the shift

between the photocenter of a recorded signal and the projection of the object's barycenter on the sky plane. In turn, this shift will play a major role in determining the final accuracy of asteroid astrometric measurements, with obvious implications for the derivation of the orbital parameters of the objects. From these simple considerations it is easy to understand how the reduction of the dynamical and physical properties of the asteroids observed by Gaia will be as a whole complex and iterative procedure.

The big advantage of Gaia, with respect to ground-based telescopes of larger aperture, will be obviously the fact of being diffraction-limited. The image of an object on the Gaia focal plane will be the result of the convolution of the incoming wave front with the optical system of Gaia. In addition, several other subtle effects will play a role in the generation of the final signal recorded by the detectors: these effects include the motion of the asteroid across the focal plane, the discrete (not ideally uniform) implementation of the time-delayed integration (TDI) read-out technique, the fact that the detector consists of arrays of discrete pixels, so that the exact location of the incoming light on the grid of pixels has some influence on the recorded signal. Other important elements to be taken into account are the quantum efficiency of the detectors at different wavelengths, and, more important, the fact that the signals will consist of the numbers of photoelectrons collected along a limited number of pixel columns. The physical size and rectangular shape of the pixels has also, obviously, a major influence and must be taken into account. In addition to the above properties of the interaction of the incoming wave front with the Gaia optics and detectors, it is also necessary to model the intrinsic properties of the signals. Asteroid emission at visible wavelengths consists of sunlight scattered by the surface. The actual flux of photons incident on the Gaia focal plane will then be the final result of a complex interaction of solar photons with the asteroid surface. The object's size, albedo, shape, macroscopic and microscopic roughness, and light-scattering properties will all play a role in the final properties of the incoming signal. We should stress that this complicated physical process is currently not fully tractable by means of purely analytic means, and no definitive theory of light scattering is presently available.

The formation of an asteroid signal on the Gaia astrometric focal plane is modeled by means of a numerical algorithm, in which all the effects mentioned above are taken appropriately into account. The general algorithm is based on a ray-tracing approach, and a Monte–Carlo implementation. An example showing how an incoming signal is finally converted into a recorded Gaia signal is shown in Fig. 1.

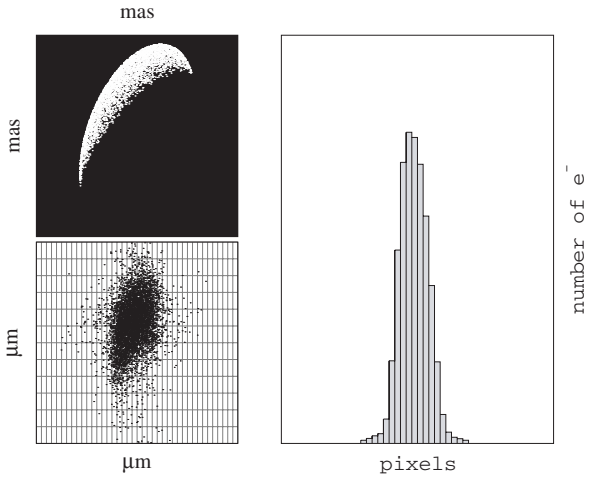


Fig. 1. Example showing how the incoming signal from an ellipsoidal object is converted into a recorded Gaia signal. The ideal optical image is randomly sampled (top left) and then combined with a sampling map of the instrument Point Spread Function, obtaining the photoelectron distribution on the CCD plate (bottom left, in which the grid represents the pixels). Finally the number of the photoelectons in each pixels column is counted obtaining the real 1D signal (right).

According to current ideas, the signals from asteroid sources actually recorded by Gaia will be sets of six or 12 numbers (depending on apparent magnitude), corresponding to the numbers of photoelectrons received in six (12) pixel columns centered around the object's photocenter. The basic goal is then to be able to develop suitable techniques to discriminate the properties of these signals with respect to signals coming from point-like sources of a similar apparent magnitude, as well as to distinguish among signals from extended objects of different angular sizes. The basic property of the signals that can be used to perform the above mentioned tasks is simply the resulting signal width (the σ of the photon distribution over the six, or 12, recorded channels). In other words, if one can derive an expected angular size versus signal width relation, a simple measurement of the signal width would be in principle sufficient to derive the angular size of the asteroidal source. The most important effect that has to be taken into account, in this respect, is photon noise.

Due to photon noise, the angular size versus signal width relation is not a rigorous one-to-one relation, but it is affected by a random variation due

to photon statistics. As a consequence when one measures a given signal width, due to the intrinsic uncertainty of the measurement due to photon noise, there is a corresponding range of angular sizes which are compatible with the measured signal width. This uncertainty is obviously increasing for decreasing signal intensity, or in other words for decreasing apparent brightness.

The complex interplay of magnitude (signal intensity), apparent angular size and resulting signal width has been explored by means of extensive numerical simulations. According to preliminary results, one can separate the apparent magnitude — apparent angular size plane into two distinct domains, according to the possibility to derive an angular size measurement having a relative accuracy better or worse than 10%, as shown in Fig. 2.

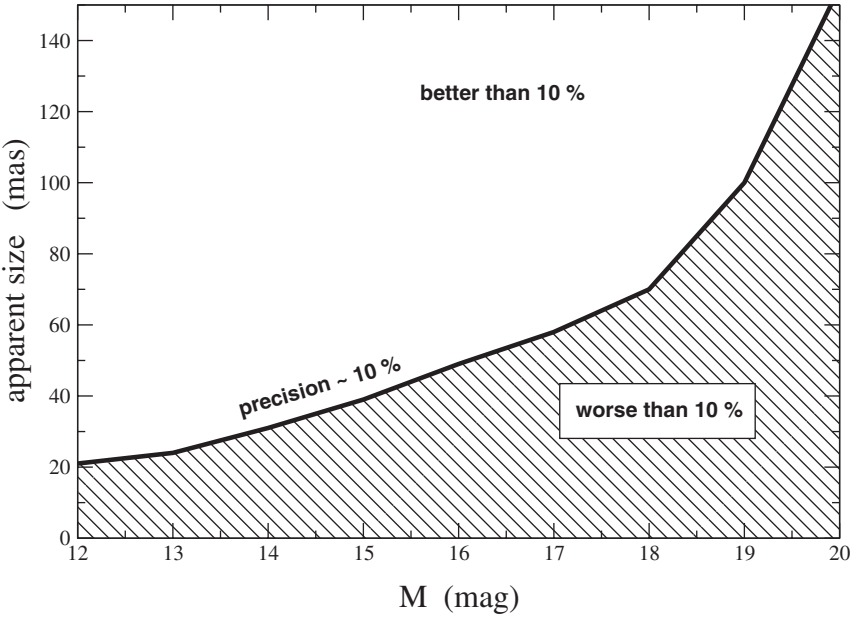


Fig. 2. Precision in measuring the sizes of asteroids with Gaia depends mainly on photons noise statistics. Precision depends on magnitude and angular size. The apparent magnitude–angular size plane can be divided in two regions by the solid line representing the cases of measurements with a precision of 10%. The upper region corresponds to cases of measurements with a precision better than 10%, while the lower one corresponds to the measurements with a precision worse than 10%.

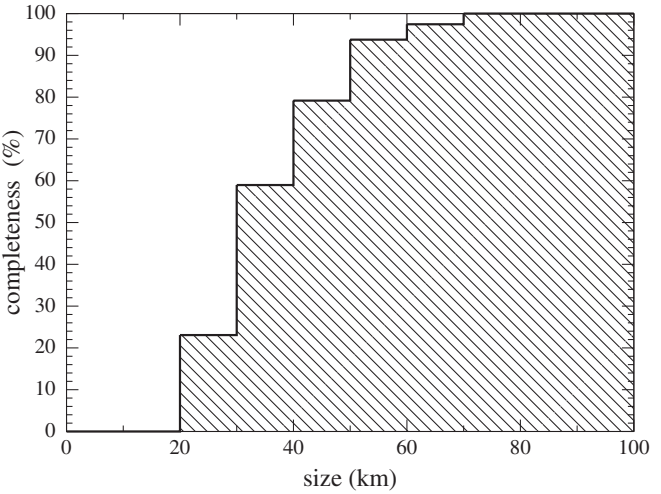


Fig. 3. The fraction (in %) of existing Main Belt asteroids whose sizes will be measured at least once with an accuracy better than 10% as a function of size, during the 5 years operational lifetime of Gaia.

An alternative way to present the results of these simulations, is given in Fig. 3, in which we take also into account the results of detailed simulations (performed by F. Mignard) of the actual detections of known main belt asteroids in a simulated Gaia survey lasting 5 years (starting since a given simulated starting date, which is not relevant for the present purposes). In particular, Fig. 3 shows, as a function of the linear sizes of the objects (in km) the fraction (in %) of the real asteroids of that size that will be measured with an accuracy better than 10% at least once during the operational lifetime of the mission.

The results of the simulations show that above 30 km in diameter, more than one half of the known main belt asteroids will have their size measured at least once during the Gaia operational lifetime. The number of useful measurements rapidly increases for increasing size. Between 20 and 30 km, a fraction larger than 20% of the objects will also be measured at least once, or even a few times. The number of objects that will be measured is thus of the order of 1,000. These results will have a tremendous impact on our knowledge of the asteroid population, as will be discussed in Sec. 5. Here, we only stress the fact that such a result will not be possible using any other observing facility existing today or planned to be operative in the next 10 years.

3. Determination of Spin Properties

Photometry at visible wavelengths has been extensively used since many years to derive information on the rotational state of asteroids. Lightcurves are obtained to determine the spin periods, and lightcurve morphology is also analyzed in order to derive some estimates of the overall shapes of the rotating bodies. Moreover, having at disposal lightcurves taken at different oppositions of the same object, corresponding to a variety of aspect angles (the aspect angle being defined as the angle between the direction of the spin axis and the direction of the observer as seen from the object's barycenter), makes it possible to derive the direction of orientation of the spin axis ("asteroid pole"). Different techniques have been developed for this purpose.[3] Predictions concerning asteroid shapes and spin axis directions based on ground-based photometry have been found to be fairly accurate, according to the results of *in situ* investigations carried out by space probes.[4]

One major advantage of observing from an orbiting platform with respect to conventional ground-based observations is that in principle from space it is easier to observe the objects in a wide range of observing geometries, not limited to those corresponding to a limited time "window" around the epoch of opposition. In particular, from space the asteroids can be seen at small solar elongation angles, which are hardly achievable from the ground. In this respect, Gaia will not be an exception. Each Main Belt asteroid will be typically observed tens of times during 5 years of planned operational lifetime. The simulations indicate that each object will be detected at a wide variety of ecliptic longitudes. In particular, Gaia will observe each object over a large fraction of the interval of its possible aspect angles. The same variety of aspect angles, which is strictly needed to derive the orientation of the spin axis, can be covered from the ground only over much longer times. This *a priori* opens exciting perspectives concerning the possibility to use Gaia disk-integrated photometry as a very efficient tool to derive the poles of the asteroids, as well as the sidereal periods and the overall shapes.

The main difference with respect to the situation usually occurring in traditional asteroid photometry, is that in the case of Gaia we will not have at disposal full lightcurves, but only a number of sparse photometric measurements lasting a few seconds (corresponding to a transit across the Gaia focal plane), obtained according to the law that determines the scanning rate of the sky by the satellite. This would seem in principle a crucial limitation, but it is more than compensated by the high number

of single photometric measurements for each object, by the fact to have data belonging to one single, homogeneous photometric system, and by the good accuracy of Gaia photometry. The latter depends, in turn, on the brightness of the target and varies for different detectors. It is expected to be better than 0.01 mag for single detections of objects as faint as at least $V = 18.5$, using the Gaia astrometric field detectors. The photometric accuracy is worse for the multi-band photometer (MBP), therefore MBP data will not be used for the faintest asteroids, whereas they will be very useful for brighter objects.

The magnitudes of the objects detected by Gaia at different epochs will depend on several parameters: the most important ones are the sidereal period, the shape and the orientation of the spin axis, and the illumination circumstances, described by the phase angle. Additional variations can come in principle also from possible albedo variegation of the surfaces, but this is not expected to be very relevant for the majority of the objects. The possible existence of a non-negligible fraction of binary systems must also be taken into account, but for the moment we will neglect this particular complication of the problem. Gaia will obviously measure apparent magnitudes that will be immediately converted to absolute magnitudes (i.e., after reducing apparent brightness to unit distance from Sun and Earth). More precisely, it will be convenient to work in terms of *differences* of absolute magnitude with respect to a reference observation of each object.

Assuming for sake of simplicity an object having a triaxial ellipsoid shape, orbiting around the Sun along a typical main belt asteroid orbit, one can plot how the absolute magnitudes are expected to vary as a function of time, depending on the coordinates of its pole. An example is given in Fig. 4. In this figure, we consider a simulated object having the same orbit as the one of the main belt asteroid 39 Laetitia (which has a typical Main Belt orbit). We assume a given triaxial shape ($b/a = 0.7$ and $c/a = 0.5$) and two slightly different choices for the spin axis orientation (ecliptic longitude of the pole at 30°, and ecliptic latitude of the pole at +60°, left, and +30°, right). The plot shows the region in the time versus magnitude-difference plane in which the object can be measured during a time span of 5 years, in the two cases. We remind that we plot here the absolute magnitude difference with respect to an arbitrary first observation obtained at a given epoch. As a consequence, the shape of the domain in which the asteroid can be observed during 5 years of mission depends not only on the assumed asteroid pole, axial ratios and sidereal period, but also on the rotational phase of the object at the epoch of the assumed first detection.

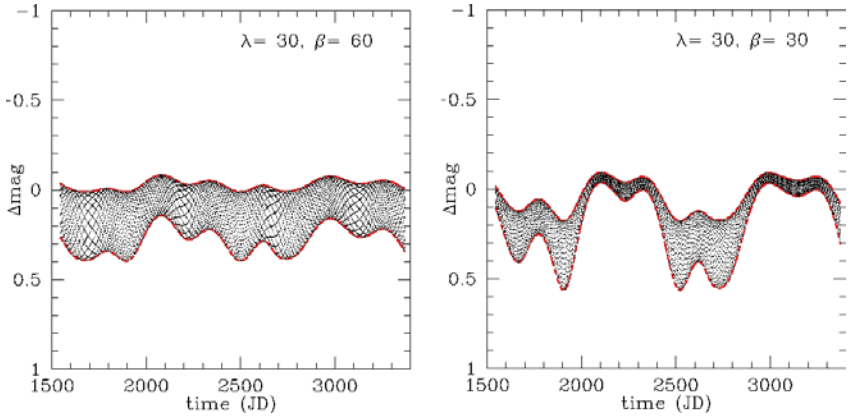


Fig. 4. Absolute magnitude difference with respect to first detection as a function of time, for a simulated object having a given sidereal period, axial ratios, and longitude of the pole λ, but with different values of the latitude β of the pole (see text).

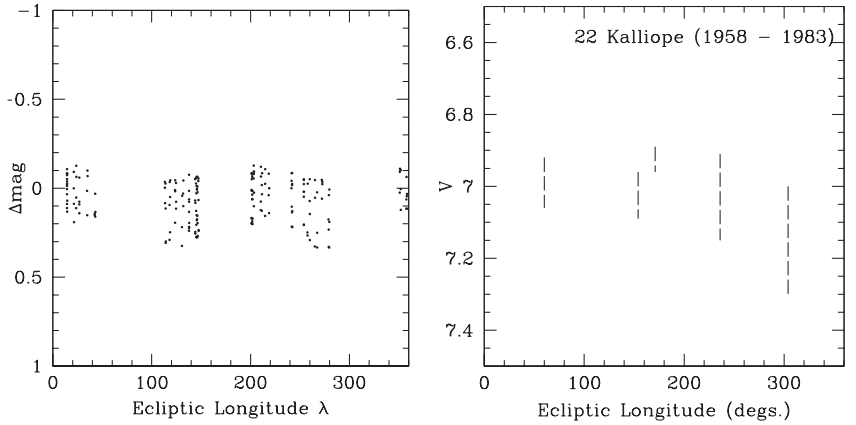


Fig. 5. Absolute magnitude difference with respect to first detection as a function of ecliptic longitude for the same object simulated in Fig. 4 (left panel). Gaia photometric measurements will sample the interval of possible observing circumstances more efficiently than typical ground-based observations. The right panel shows as an example the photometric data-set of asteroid 22 Kalliope, obtained since 1958–1983, used by Zappalà and Knežević[5] to derive the pole of this object.

It is more instructive to plot the magnitude difference with respect to first detection as a function of ecliptic longitude, instead of time. Figure 5 (left) shows this kind of plot. Instead of the full theoretical domain of the plane where the object can be observed, in this figure we show the

location of the actual detections of the object by Gaia. This computation is based on a general simulation of asteroid detections by Gaia performed by F. Mignard. In this plot, we assume the same set of parameters (axial ratio, ecliptic longitude of the pole, sidereal period and rotational phase at first detection) as used in producing Fig. 4 (left). As one can see, Gaia will observe an object having the same orbit of 39 Laetitia many times (around 200 in this example, corresponding to the computed number of detections in the field of view of the Gaia medium band photometer), and over a wide interval of ecliptic longitudes. In order to compare this with a typical set of ground-based data that have been sufficient to derive a reliable estimate of the asteroid pole, we show in the right panel of Fig. 5 an analogous plot, showing the typical set of data (full lightcurves) used in the past to derive the pole of asteroid 22 Kalliope.[5] It is easy to see that Gaia sampling of the longitude–magnitude plane in only 5 years of mission will be much denser with respect to ground-based data obtained, in this particular example, by means of observations spanning over an interval of 25 years. For reasons like this, we expect that Gaia photometry will be a great resource for asteroid science in the future.

Since the number of unknowns (the value of the spin period, two coordinates for the pole axis orientation, a few parameters describing the shape, and an initial rotation phase at $t = 0$ which has also to be determined) is much smaller than the number of observations, it is in principle possible to develop techniques of inversion of photometric data, to determine from them the spin period, the spin axis orientation, and the overall shape. Different approaches are possible in principle, and have been independently considered by different teams. In this paper, we explain in greater details the algorithm developed by the Torino team.

In our approach, as a first approximation, we describe the shapes of the objects by means of triaxial ellipsoids. This is certainly a major simplification, but it has been found in the past to be reasonable for the purposes of asteroid pole computation in a wide variety of cases.[3] The shape parameters to be derived by photometry inversion are then the two axial ratios b/a and c/a. In addition to the rotation period, the pole coordinates, and the unknown rotational phase at the zero-epoch, other parameters to be derived are the linear slope of the magnitude-phase angle variation, which is known to characterize the photometric behavior of real objects, and, at a higher degree of complexity, the coefficient of the linear increase of the lightcurve amplitude as a function of phase angle. This effect, observed for real asteroids,[6] means that one must take into account that the linear

variation of the magnitude as a function of phase angle is not constant at different rotational phases of the object.

In order to solve this inversion problem, we have explored a fully numerical option, based on techniques of "genetic" computation. This approach seems to be particularly promising. This computational technique is based on concepts of "survival of the fittest" which are inspired to classical studies of the evolution of living species. In particular, the possible solutions of the problem are characterized by a set of parameters (spin period, pole coordinates, axial ratios, etc.) which can be seen as the "genes", or the "DNA" of a solution. The goodness of any given solution is assessed on the basis of its corresponding residuals with respect to a set of real or simulated observations. Better solutions give a better fit of the observational data, and the idea of the genetic approach is to simulate an evolution of the solution parameters, by selecting at each step only those giving the best fit. This idea is implemented by means of a numerical algorithm, which initially generates a large number of completely random solutions, saving in memory only a limited subset, corresponding to those producing the smallest residuals. In general, these preliminary solutions are very bad, as one would expect *a priori* from a set of completely random attempts. At this time, the "genetic" mechanism is switched on. This consists in random coupling of the parameters of the saved solutions, and in random variations ("genetic mutation") of some of the parameters constituting the "DNA" of a single solution. If the newly born "baby" solution is better than some of those saved until that step, it enters the "top list", whereas the previously worst solution is removed from the same set. In this way, after a number of the order of one or two millions of "genetic experiments", a very good solution is usually found, which produces small residuals and basically solves the inversion problem. Due to the intrinsically random nature of this "genetic" approach, the right solution is not forcedly found always, but if the genetic algorithm is repeatedly applied to the same set of observed data (typically 30 and 40 times) the right solution (giving the minimum residuals in different attempts) is usually found several times.

So far we have been able to satisfactorily invert many sets of simulated photometric observations. The simulations indicate that the algorithm succeeds in determining with a very high accuracy the spin period, the parameter that would seem *a priori* most difficult to determine, because it should be taken into account that accuracies of the order of 10^{-5} h are needed for this parameter. In spite of this difficulty, we have always been able so far to solve the inversion problem in practically all the simulated cases, in

which we have considered wide variations of all the relevant parameters (sidereal period, axial ratios, pole coordinates, initial rotational phase of the object). The tests have been performed so far using different simulators of Gaia photometric data developed independently by different teams in the Gaia Solar System Working Group. Most tests have been carried out so far by simulating Gaia photometric observations of asteroids, both in the main belt and belonging to the near-Earth population, having ideal triaxial ellipsoid shapes. In these cases, the inversion algorithm is able to find the "true" solution (spin period, pole coordinates, and axial ratios) with an excellent accuracy. Preliminary tests are encouraging, however, even in cases of simulated objects having more complicated shapes. Moreover, a few preliminary tests concerning real asteroids previously observed by the Hipparcos satellite are also encouraging. For instance, the correct spin period, and a reasonable general shape have been obtained in the case of asteroid 216 Kleopatra, in spite of a small number of available Hipparcos measurements (of the order of 20). In this respect, the simulations carried out so far allow us to derive an estimate of the performances of the inversion algorithm as a function of the number of available observations for a given object, and as a function of the photometric accuracy of the observations. In particular, it turns out that excellent inversions of triaxial objects can be obtained for data-sets of about 30 measurements, characterized by Gaussian observational errors having a σ of 0.02 mag. This is significantly larger than the expected Gaia photometric error bars for objects fainter than $V = 18.5$, to be pessimistic. Moreover, we have also checked that the inversion algorithm is able to determine the right sense of rotation of the objects, as theoretically expected.

We note that another fully independent approach, adopted by other colleagues from Helsinki is also being applied for the treatment of simulated Gaia photometric data. The method is based on a different approach.[4] According to simulations, also this technique seems to be extremely effective.

It should be stressed that the data-set of sparse Gaia photometric detections will be qualitatively similar to what we can expect to come in the future from ground-based asteroid surveys like Pan-STARRS or the Large-Aperture Synoptic Survey Telescope. This means that the same techniques of photometric inversion that are being developed for Gaia will be also useful for the above-mentioned surveys. The main difference, however, is that Gaia will be able to sample the possible range of ecliptic longitudes (hence, aspect angles) of the objects at a much faster pace than ground-based

surveys, due to the above-mentioned capability to obtain accurate photometric measurements also far from opposition, when the objects are located at small solar elongation angles and are seen at different aspect angles.

The results of the simulations performed so far indicate that, whatever choice will be made of the most suitable algorithm of inversion of photometric data, the Gaia data-set of disk-integrated photometry will be a major resource to derive spin properties and overall shapes for a number of objects that will not be smaller than 10,000, to be quite conservative. Again, we stress that this will be possible by means of only 5 years of Gaia observations.

4. Taxonomic Classification

The derivation of an asteroid taxonomy will be a direct by-product of Gaia's spectrophotometric capability. According to current knowledge, there are more than 200,000 main belt asteroids that will exhibit apparent magnitudes brighter than 20 in V light when detected by Gaia. A new taxonomic classification will then be produced, as a fairly trivial exploitation the spectrophotometric data coming from the Gaia MBP, which will observe all celestial objects in 11 bands across the electromagnetic spectrum between about 3,000 and 9,500 Å. In this respect, Gaia will have a couple of major qualities: first, this huge spectrophotometric data-base will be obtained using a unique, homogeneous photometric system, and not merging together data coming from different instruments. Second, and equally important, the spectral coverage of the MBP will certainly include also some bands in the the blue region of the reflectance spectrum. This is quite important, because the U and B regions of asteroid spectra, which were well covered by classical UBV spectrophotometry based on photoelectric photometers several years ago, tend now to be missed by the most recent spectroscopic surveys, like SMASS and SMASS2.[7,8] This is a problem, because the blue region is very useful to distinguish between several different sub-classes of primitive objects. Among the hundreds of thousands of asteroids that are expected to be taxonomically classified using the Gaia spectrophotometric data-base, a large fraction will consist of primitive, dark objects belonging to the so-called C-class complex, which dominate the asteroid inventory in the outer region of the asteroid belt. In contrast to spectroscopic surveys like SMASS and SMASS2 that were mostly limited to an interval of wavelengths between 5,000 and 9,500 Å,[8] Gaia spectrophotometric data are expected to be better to discriminate among different subclasses of the big C complex,

and to determine the relative abundance of these different subclasses, also as a function of heliocentric distances. This will be important for studies of the compositional gradient of the solid matter in the Solar System.[9] Other important applications of taxonomy will be discussed in Sec. 5. Here, we stress that a new taxonomy is something that can be conceivably obtained in the near future also by means of dedicated ground-based surveys like the planned Pan-STARRS and/or the Large-Aperture Synoptic Survey Telescope. *Per se*, the new taxonomic classification that will be produced by Gaia is not a result that, alone, would be sufficient to justify the costs of a space mission. The mentioned coverage of the B region of the spectrum, however, and the general link with other, unique results that will be produced by Gaia asteroid observations, make Gaia taxonomy a very welcome and extremely useful addition to the already very rich scientific output of this mission in the field of asteroid science.

5. Conclusions

Based on the considerations developed in this paper, we can expect that in the future we will speak of two distinct eras in the history of asteroid science: a pre-Gaia and a post-Gaia era. The post-Gaia era will be much more advanced in many fundamental respects. Gaia astrometry will provide us reliable measurements of mass for about 100 objects. For these same objects we will know also the size and the overall shape, again obtained from Gaia data, then we will have at disposal reliable estimates of average densities. These objects will belong to different taxonomic classes, then it will be possible to assess the relation between density and taxonomy, to be possibly interpreted in terms of overall composition and structural properties. This kind of knowledge cannot be realistically expected to be achieved by other means in the next two decades.

The direct measurement of sizes will allow us to have an improved knowledge of the asteroid size distribution down to 20 km in diameter. Today, the vast majority of size data at our disposal come from indirect measurements (thermal radiometry, polarimetry) and are subject to considerable uncertainties. Direct size measurements, on the other hand, will allow us to trivially derive the albedos of the same objects, being known their distances at the epochs of the observations. In this way, we will check whether there are dependencies of albedo upon size, as old results of IRAS radiometric observations seem to indicate.[9] An albedo variation due to space weathering processes has been convincingly shown to exist for S-type objects by the Galileo close-range images of asteroid 243 Ida,[10] but it is not completely

clear if the same processes can alter also the reflectance properties of objects belonging to other taxonomic classes (i.e., having different surface compositions). Gaia will give the correct answers to these major open questions.

The determination of the rotational state for a set of more than 10,000 main belt asteroids will make it possible to use the spin properties of main belt objects as a new powerful constraint to the models of the collisional evolution of this population. So far, the models have been mainly constrained by the size distribution, but since collisions strongly affect the spins, Gaia data will add a new full dimension to the problem.

The multi-band observations carried out by Gaia for several tens of thousands of objects will be useful not only to derive a new taxonomy, but also to add spectral reflectance data as a further constraint to identify asteroid families. These are groupings of objects having similar orbital properties, which constitute the outcome of catastrophic events of collisional disruptions of single parent bodies. Families have been so far identified mostly on the basis of similarities in the orbital proper elements of their members,[11] but it has become evident that spectral properties can be added as an additional element to identify objects having a common collisional origin.[11] At the same time, the spin properties derived for family objects will allow us to test the possible existence of preferential alignments in the spin properties, which have been recently found to likely exist in the case of the Koronis family[12] and might be interpreted as a direct evidence of the so-called YORP effect, a kind of thermal torque that is believed to be important in the evolution of asteroidal spins.[13]

If current plans will be respected, Gaia will be launched in 2011. The present generation of planetary scientists will have a chance to participate in this predicted revolution in our knowledge of minor planets.

Acknowledgments

We thank F. Mignard for putting at disposal of the Gaia Solar System working group the results of his simulations of Gaia asteroid observations. We would like to thank him, as well as K. Muinonen, D. Hestroffer, J. Virtanen, M. Kaasalainen, C.-I. Lagerkvist and V. Zappalà for very useful discussions.

References

1. P. Tanga, in *Proceedings of the Symposium "The Three Dimensional Universe with Gaia"*, eds. C. Turon, K. S. O'Flaherty and M. A. C. Perryman (ESA Publication Division SP-576, ESTEC, Netherland, 2005), p. 243.

2. W. F. Bottke, D. Vokrouhlický, P. Rubincam and M. Brož, in *Asteroids III*, eds. W. F. Bottke, A. Cellino, P. Paolicchi and R. P. Binzel (University of Arizona Press, Tucson, 2002), p. 395.
3. P. Magnusson, M. A. Barucci, J. D. Drummond, K. Lumme, S. J. Ostro, J. Surdej, R. C. Taylor and V. Zappalà, in *Asteroids II*, eds. R. P. Binzel, T. Gehrels and M. S. Matthews (University of Arizona Press, Tucson, 1989), p. 66.
4. K. Kaasalainen, S. Mottola and M. Fulchignoni, in *Asteroids III*, eds. W. F. Bottke, A. Cellino, P. Paolicchi and R. P. Binzel (University of Arizona Press, Tucson, 2002), p. 139.
5. V. Zappalà and Z. Knežević, *Icarus* **59** (1984) 436.
6. V. Zappalà, A. Cellino, M. A. Barucci, M. Fulchignoni and D. F. Lupishko, *Astron. Astrophys.* **231** (1990) 548.
7. S. J. Bus, PhD Thesis, Massachusetts Institute of Technology, Boston, 1999.
8. S. J. Bus and P. R. Binzel, *Icarus* **158** (2002) 146.
9. A. Cellino, *Space Science Rev.* **92** (2000) 397.
10. C. R. Chapman, *Meteoritics & Planet. Sci.* **31** (1996) 699.
11. Ph. Bendjoya and V. Zappalà, in *Asteroids III*, eds. W. F. Bottke, A. Cellino, P. Paolicchi and R. P. Binzel (University of Arizona Press, Tucson, 2002), p. 613.
12. S. M. Slivan, R. P. Binzel, L. D. Crespo da Silva, M. Kaasalainen, M. M. Lyndaker and M. Krco, *Icarus* **162** (2003) 285.
13. D. Vokrouhlický, D. Nesvorný and W. F. Bottke, *Nature* **425** (2003) 147.

LIGHTCURVES OF THE KARIN FAMILY ASTEROIDS

TAKASHI ITO and FUMI YOSHIDA*
National Astronomical Observatory of Japan
Osawa 2–21–1, Mitaka, Tokyo 181–8588, Japan
**yoshdafm@cc.nao.ac.jp*

We present the first results of a long-term campaign of photometric observations of the Karin family asteroids. This family is very compact, and is supposed to be extremely young, with an estimated age of about 5.8 Myrs. The purpose of our observations is to determine the rotational properties, the colors and hopefully the overall shapes of the largest possible number of family members, since this might provide important information about the physics of the original break-up event that quite recently produced this family. The lightcurves that we have already obtained for 12 objects are of a generally good quality. We have also obtained some indication that the largest member of the family, (832) Karin, might exhibit some color variation across its surface. This might be an interesting result, but it has to be confirmed by future observations.

1. Introduction

The Karin family was recognized quite recently, with the estimated age of only about 5.8 Myrs.[1] This family consists of about 70 asteroids with sizes ranging from about 1.5 to 20 km in diameter.[2] Most asteroid families are very old, and they have undergone significant collisional and dynamical evolution since their formation, which likely masks the properties of the original collisions. But the remarkably young Karin family asteroids possibly preserve some signatures of the original collisional event that formed the family. This extraordinary feature of the Karin family provides us with several significant opportunities for the research of young asteroids such as potentially detecting tumbling motion, obtaining distribution of rotation period, and estimating the shapes of newly created asteroid fragments.

Driven by these motivations, we have begun a program since November 2002 to observe the lightcurves of all the Karin family members. The potential result derived from our observation could be a strong constraint on laboratory and numerical experiments of collisional fragmentation.[3] In the

*Corresponding author.

first part of this article, we report our preliminary results concerning the lightcurves of the twelve Karin family members, though detailed statistical discussions will be presented in a separate paper.[4] Next, we move on to the result of our multicolor observation of the largest member of this family, (832) Karin. Since (832) Karin is the largest fragment of a recent asteroid disruption, it is possible that this asteroid has both young and old surfaces together: a young surface that was exposed from the interior of the parent body by the family-forming disruption, and an old surface that used to be the parent body surface exposed to space radiation over a long time. If the mixture of these two surfaces is detected by our multicolor observation, it could have significant implication for research on the evolution of asteroid surface spectra.

We briefly report the lightcurves of several Karin family members in Sec. 2. In Sec. 3 we describe our multi-color observations of (832) Karin in 2003 as well as in 2004. The method and results of our multicolor observations are summarized in this section. Section 4 goes to some discussions and interpretation of the results.

2. Lightcurves of the Karin Family Asteroids

For our lightcurve observations of the Karin family asteroids, we have used eight telescopes: The 90-inch Bok reflector at the Steward Observatory (AZ, USA), the 1.8-m Vatican Advanced Technology Telescope (AZ, USA), the 1.5-m telescope at Maidanak Observatory (Uzbekistan), the 1-m telescope at Lulin Observatory (Taiwan), the 1-m Schmidt telescope at Kiso Observatory (Japan), the 0.5-m telescope at the National Astronomical Observatory (Japan), the 0.4-m telescope at Fukuoka University of Education (Japan), and the 0.25-m telescope at Miyasaka Observatory (Japan). We used R-band filter all through the observations because asteroids are generally brightest in the R-band wavelength. Exposure time was 2–8 min so that asteroids had the appearance of point sources. We also observed several Landolt photometric standard stars[5] to determine extinction coefficients. Photometric reduction and aperture photometry were performed using the APPHOT/IRAF package. Magnitudes of the asteroid at different air masses were corrected by the extinction coefficient at each band. Asteroid brightness was measured with respect to that of the field stars in the USNO-A2 catalogue in the same frame.

Lightcurves from the photometric data are constructed following the procedure proposed by Harris and Lupishko.[6] Principally it is an iterative

repetitions of frequency analysis and fitting to Fourier series. We use Lomb's spectral analysis[7] or the WindowCLEAN analysis[8] for the frequency analysis of lightcurves, and fit the data with an eighth order Fourier series.[9] We have to be particularly careful when we combine the lightcurves of several observing runs because they generally have different zero-level magnitudes. We combine the lightcurves of multiple observing runs based on these zero-levels to obtain our final result.

(832) Karin was observed in 1984 by Binzel.[10] Here we report new observations of (832) Karin at 2003 opposition and additional observations at 2004 opposition (see also Sec. 3). We report lightcurves for the first time for the rest of the selected objects. Moreover, 1999 CK16 was observed at two oppositions (2002 and 2004). All these informations together with additional information obtained from future lightcurve data will be used to increase our knowledge of these objects, not only of their rotational periods, but also of other features such as their pole axis.[11]

All the lightcurves that we have determined are displayed in Figs. 1 and 2. The resulting rotation periods, peak-to-peak variations of the lightcurves, and the solar phase angles are listed in Table 1. Looking at Figs. 1 and 2, it is obvious that some of the asteroids, such as Einer, 1997 GT36, or Svojsik, need more and better observations to obtain more accurate lightcurves. Our preliminary data analysis, though the result is not apparent from the figures, indicates that a few members might perform the so-called tumbling motions (i.e., non-principle axis rotation).

3. Multi-Color Observations of (832) Karin

We twice performed multi-color observations of the largest member of the Karin family, (832) Karin, in September 2003, and in September 2004 after an interval of one year.

3.1. *Observation procedure*

For our multi-color observation of (832) Karin, we used the $2\,\mathrm{K}\times 2\,\mathrm{K}$ CCD of the 1.8-m Vatican advanced technology telescope (VATT) on Mt. Graham, Arizona, USA. Our first observation was performed in September 2003, and the second one was done in September 2004. Some of the major parameters used during this observation are listed in Table 2.

The procedures of these two observations are entirely the same. We use B, V, R, and I-filters whose wavelengths are centered at 4359.32, 5394.84,

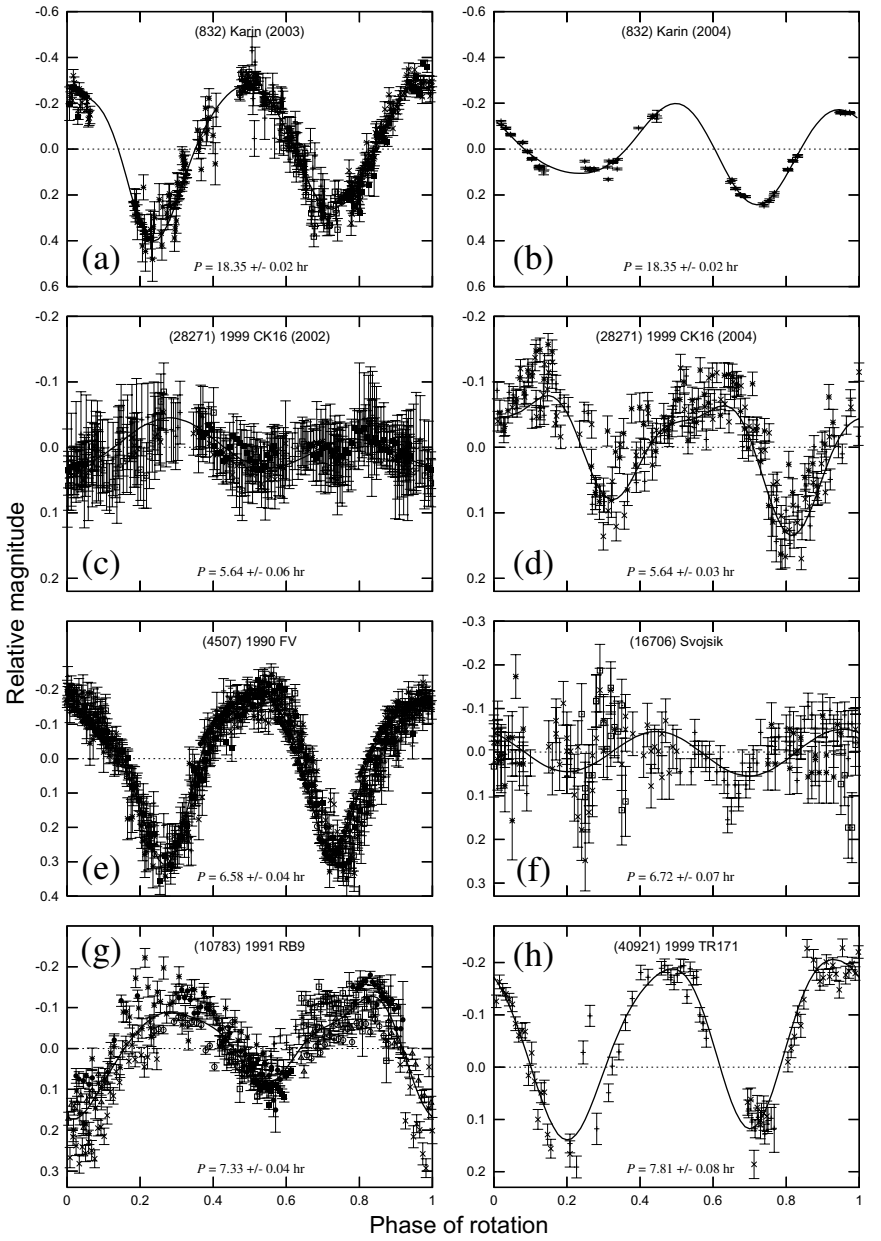


Fig. 1. Lightcurves of six Karin family members.

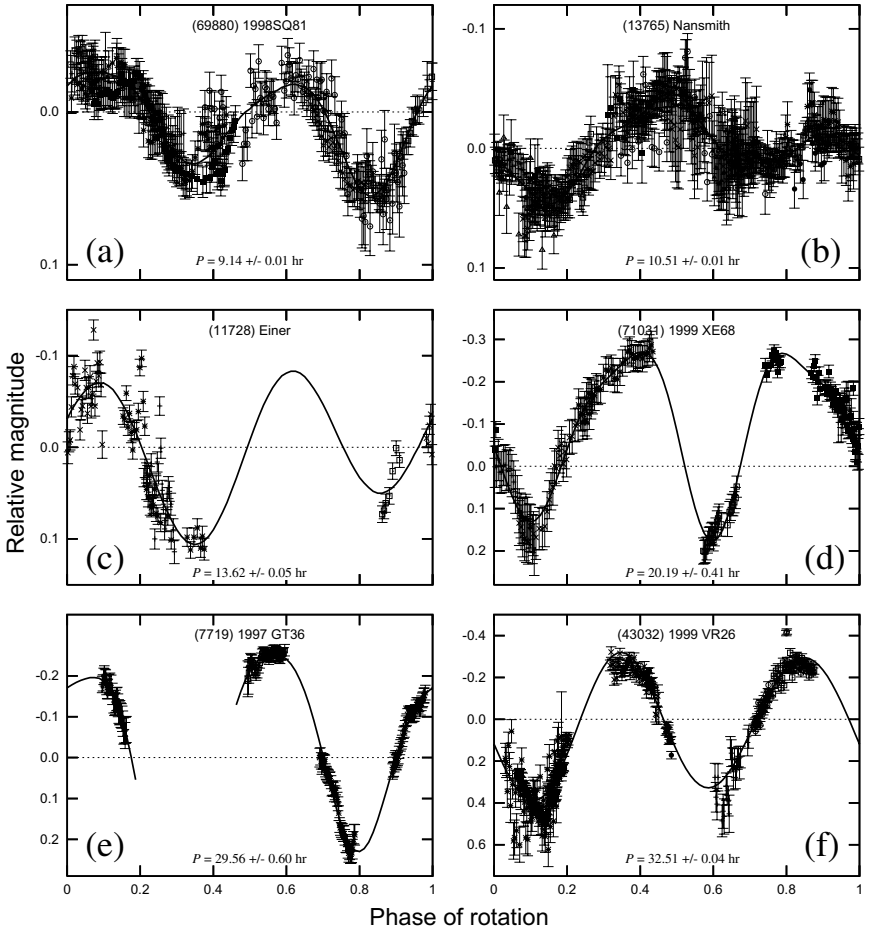


Fig. 2. Lightcurves of another six Karin family members.

6338.14, and 8104.87 Å. In order to remove the effect of magnitude variation due to an asteroid's rotation that could affect the asteroid's color, we always take a pair of R-band images before and after we use other filters. Hence we define one observation sequence as RR–BB–RR–II–RR–VV–RR. Each of the R magnitudes is interpolated (or extrapolated) to the value at the same UT when we use other filters for comparison.

Since the exposure time for each image is 2–3 min, each of these sequences takes about 40 min. While taking R-band images consecutively

Table 1. Some properties of the Karin family asteroids. P is rotation period (hour), δM is peak-to-peak variation magnitude, and α is the solar phase angle (degree). Lightcurves of all asteroids are shown in Figs. 1 and 2. For (832) Karin and (28271) 1999CK16, we had two opportunities to observe their lightcurves: 2003 and 2004 for (832) Karin, 2002 and 2004 for (28271) 1999CK16.

Asteroid	P	δM	α	Figure #
(832) Karin (2003)	18.35 ± 0.02	0.61 ± 0.02	1–14	1(a)
(832) Karin (2004)	18.35 ± 0.02	0.61 ± 0.01	21.7	1(b)
(28271) 1999CK16 (2002)	5.64 ± 0.06	0.08 ± 0.04	3.1	1(c)
(28271) 1999CK16 (2004)	5.64 ± 0.03	0.21 ± 0.02	8.4	1(d)
(4507) 1990FV	6.58 ± 0.04	0.49 ± 0.03	7.9	1(e)
(16706) Svojsik	6.72 ± 0.07	$\sim$0.3	12.7	1(f)
(10783) 1991RB9	7.33 ± 0.04	0.50 ± 0.02	5.6	1(g)
(40912) 1999TR171	7.81 ± 0.08	0.35 ± 0.02	1.5	1(h)
(69880) 1998SQ81	9.14 ± 0.01	0.08 ± 0.01	5.3	2(a)
(13765) Nansmith	10.51 ± 0.01	0.09 ± 0.02	11.3	2(b)
(11728) Einer	13.62 ± 0.05	0.19 ± 0.01	10.3	2(c)
(71031) 1999XE68	20.19 ± 0.41	0.45 ± 0.04	5.0	2(d)
(7719) 1997GT36	29.56 ± 0.60	0.50 ± 0.02	18.6	2(e)
(43032) 1999VR26	32.51 ± 0.04	1.00 ± 0.06	10.6	2(f)

Table 2. Major parameters during our multi-color observations of (832) Karin. From the left, UT referring to the mid-time of each night, distances (AU) between the asteroid and the Sun (r) and the Earth (Δ), the ecliptic longitude (λ) and latitude (β), and the solar phase angle (α) of this asteroid. The unit of angles is degree.

Date (UT)	r	Δ	λ	β	α
20030926.19	2.666	1.803	324.8	1.5	13.36
20030927.19	2.666	1.811	324.7	1.5	13.68
20030928.17	2.665	1.819	324.6	1.5	13.99
20030929.17	2.665	1.827	334.5	1.5	14.30
20040922.44	2.706	2.442	63.8	0.2	21.71
20040923.44	2.707	2.429	64.1	0.2	21.68
20040924.45	2.707	2.417	64.3	0.2	21.64

for the lightcurve observation that we described in the previous section, we performed the multi-color observing sequence several times with intervals of a few hours. Since we were able to observe this asteroid for 4–5 h every night, we repeated this procedure seven times in our 2003 observation and ten times in our 2004 observation. As a result, we obtained color differences such as V–I or B–V. We calculated the errors of these values from the photometry error of each of the B, V, R, and I images: For example, the error of V–I is $\sqrt{\delta V^2 + \delta R^2}$, where δV and δR are the photometry errors of the V and R images.

3.2. *Observation results*

The resulting time variation of the surface color of (832) Karin in our 2003 observation[12] is summarized in Fig. 3(b). For reference, we show the lightcurve of this asteroid during the summer to autumn of 2003 obtained from a couple of telescopes including VATT (Fig. 3(a), equivalent to Fig. 1(a)). As seen in Fig. 3(b), we obtained the color data of this asteroid for over more than 80% of its rotational period at this observation.

The results of our 2004 multi-color observation are summarized in Fig. 3(e), as well as this asteroid's lightcurve obtained at this observation (Fig. 3(d), equivalent to Fig. 1(b)). This time we obtained the color data of this asteroid over almost the entire period of its rotation.

Looking at Fig. 3(b), which shows major results of the 2003 observation, the V–R value is almost constant throughout the rotation. The change in B–V is slight in the early phase of rotation, then gradually becomes larger during the period of this observation. What most draws our attention in this data is an obvious anomaly in V–I value at phase $\sim$0.2. To inspect this anomaly in more detail, we calculated the wavelength dependence of the relative reflectance of this asteroid by subtracting the solar colors of B–$V = 0.665$, V–$R = 0.367$, and V–$I = 0.705$[13] from our original color data. The relative reflectance is normalized at a wavelength of the V filter, 5394.84 Å. Then, as shown in Fig. 3(c), we found that the relative reflectance of this asteroid at long wavelengths (i.e. in the I-band) is much larger at the rotation phase $\sim$0.2 than at other phases. The steep slope of the relative reflectance in Fig. 3(c) at phase $\sim$0.2 should be called "red", as is often seen in regular S-type asteroids.[14,15]

Note that the magnitude errors in Fig. 3(b) look smaller than the magnitude errors in Fig. 3(a), which might seem strange. This is because we have used lightcurve data from many other smaller telescopes in Fig. 3(a), not only that from the 1.8-m VATT, while we drew Fig. 3(b) with only the data from the 1.8-m VATT. If you compare Figs. 3(d) and (e) for both of which we used the only data from VATT, you can see that the magnitude errors in Fig. 3(e) are as large as, or larger than, those in Fig. 3(d), which seems reasonable.

In our 2004 observation results, lightcurve of (832) Karin (Fig. 3(d)) looks different from what we saw a year before (Fig. 3(a)). This is reasonable because the relative orbital configuration of (832) Karin and the Earth is different from our 2003 observation. A remarkable fact of this observation is that we no longer saw a particularly "red" surface on this asteroid.

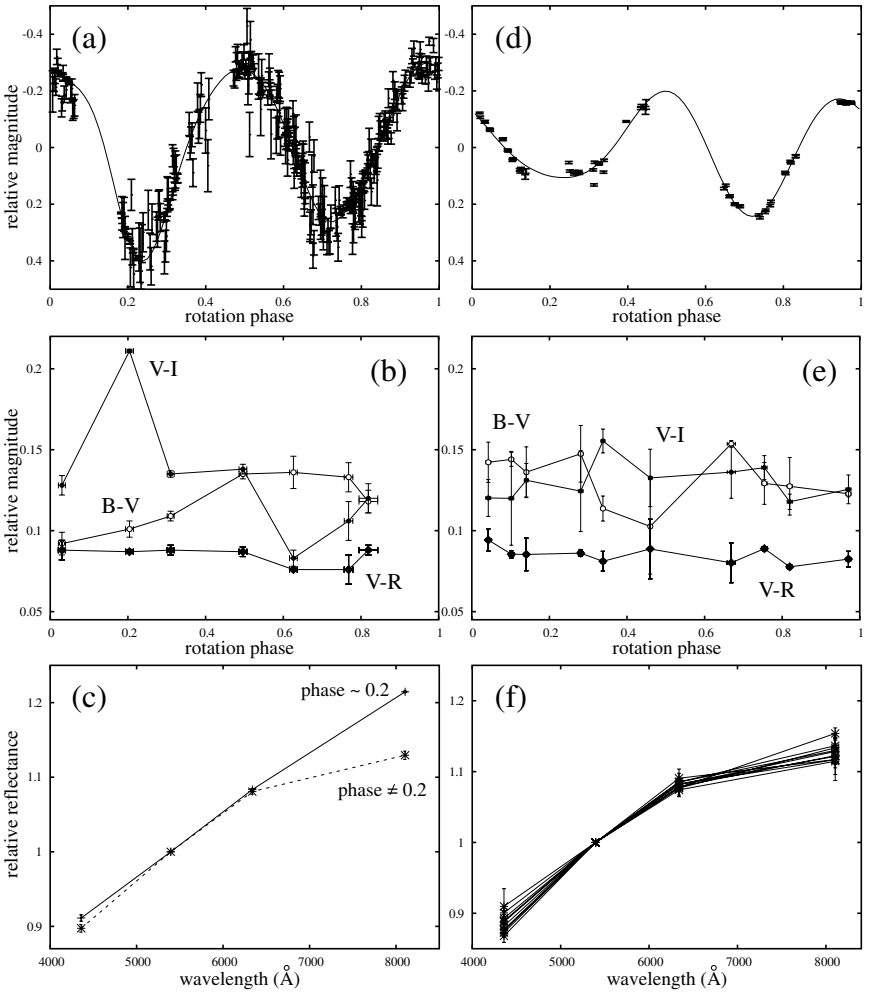


Fig. 3. Lightcurve, relative magnitude, and wavelength dependence of relative reflectance of (832) Karin in our two observations at VATT in September 2003 and September 2004. The left three panels (a)–(c) are for the 2003 observation, and the right three panels (d)–(f) are for the 2004 observation. (a) and (d): Lightcurve. Note that in (a) we have included the data not only from VATT in September 2003, but the data from other smaller telescopes with larger errorbars. (b) and (e): Relative magnitude of B–V, V–I, and V–R. (c) and (f): Wavelength dependence of relative reflectance in B-, V-, R-, and I-band normalized at the V-band wavelength, 5394.84 Å.

Time variation of relative magnitude of V–I in Fig. 3(e) does not show any definite anomaly, unlike what was seen in Fig. 3(b) in September 2003. The wavelength dependence of the relative reflectance of this asteroid in Fig. 3(f) is more like that of phase $\neq 0.2$ in Fig. 3(b) in September 2003 than that of phase $\sim$0.2. In a word, (832) Karin did not show a mature (red) surface in September 2004, exhibiting only a fresh surface with low relative reflectance at longer wavelengths.

4. Discussion

4.1. *Interpretation of the observing results*

So far we do not have a very good explanation for the unexpected color mismatch between the 2003 and 2004 multi-color observations. The key to solving this problem might lie in the difference in the amplitude of two lightcurves in Figs. 3(a) and (d): The lightcurve of September 2003 has a larger amplitude than that of September 2004. In general, when we look at an asteroid from its pole direction, especially at around opposition, the brightness of the asteroid can be nearly constant. Considering the relative orbital configuration between (832) Karin and the Earth, we have drawn a rough and possible schematic figure for deducing why we did not see a red surface on this asteroid in our 2004 observation (Fig. 4). Following Sasaki *et al.*[16] considerations, (832) Karin might be a cone-shaped asteroid fragment with a small portion of mature surface that used to be part of the parent body's surface. If the rotation axis of this fragment is highly inclined, nearly parallel to its orbital plane as in Fig. 4, it might account for the fact that we see its red surface occasionally as it rotates at the position of September 2003. If the orbital configuration, the spin axis orientation, and the location of the red surface are as in Fig. 4, it might also be that we could not see any red surface on this asteroid in September 2004 when we were supposed to look at this asteroid from nearly the pole direction. This geometric configuration could explain why the lightcurve amplitude is smaller in our 2004 observation than in the 2003 observation, depending on the shape of this asteroid.

The surface color variation of (832) Karin suggests that this asteroid possesses an inhomogeneous surface. Judging from the recent breakup history of the Karin family, a part of it could be fresh and newly exposed by the family-forming disruption. Meanwhile there could be a mature surface, once the parent body surface, and had been exposed to space radiation or particle bombardment over a long time.

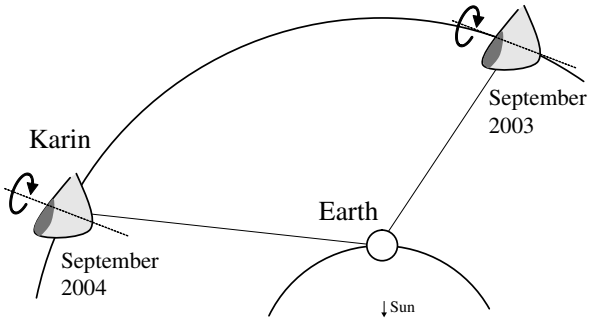


Fig. 4. A rough schematic illustration of the orbital configuration of (832) Karin and the Earth in September 2003 and September 2004. Relative location of the two bodies is determined by the solar phase angle in Table 2. The Earth was roughly at the same position at our 2003 and 2004 observations. We assume that rotation axis of this (maybe) cone-shaped asteroid is almost parallel to its orbital plane, and it has a small portion of red surface (dark gray area).

The existence of the color variations found in our 2003 observation is supported by a near-infrared spectroscopic observation of this asteroid that was performed at nearly the same time as our observation. Sasaki *et al.*[16] deployed the Cooled Infrared Spectrograph and Camera for OH-airglow Suppressor (CISCO) at the 8.2-m Subaru telescope on MaunaKea, Hawaii, and observed (832) Karin in near-infrared wavelength on September 14, 2003, close to the date of our 2003 observation. As a result, Sasaki *et al.*[16] obtained the near-infrared spectra of this asteroid at three different rotational phases; 0.30–0.33, 0.34–0.38 and 0.45–0.51 in our Fig. 3(a). They found a significant difference in the slope between the spectrum obtained at phase = 0.30–0.33 and the others. The former is similar to the spectra of ordinary S-type asteroids (i.e., "red" spectrum), while the latter two match well with the spectra of ordinary chondrites. Sasaki *et al.*[16] interpreted this asteroid's spectrum difference as being due to the mixed distribution of matured and fresh surfaces. This trend of color variation is quite similar to what we obtained in our 2003 observation (Figs. 3(b) and (c)).

A small inconsistency between our and Sasaki's observations is the difference in the rotation phase where the "red" spectrum was observed: In our 2003 result, the surface of (832) Karin seemed mature when the rotation phase was $\sim$0.2, while Sasaki's result claims that the mature surface appeared when the phase was around 0.3. We think this mismatch was caused by an uncertainty in rotational period determination, and does not have a significant influence on our assertion that this asteroid

had a red surface at the rotation phase around 0.2–0.3 when seen in September 2003.

Here we also have to add details as to what we observed in our 2003 observation. Although the existence of old and mature surface on (832) Karin is surely interesting, we need to be aware that the detection of the mature surface could be caused by an artificial effect. In our 2003 observation, the major color change occurred only through the I band color (Fig. 3(b)) at the rotation phase corresponding to the minimum brightness of this asteroid (Fig. 3(a)). Hence, another explanation might be possible: "The apparent magnitude of this asteroid was close to the instrumental limit in the I band color sensitivity, and the derived I magnitudes are not correct". This hypothesis will be denied or confirmed by our future observations.

4.2. *Future observation*

Our observation of the Karin family asteroids has just begun, and will continue getting better and more accurate lightcurves of more asteroids until we cover all the members ($\sim$70) of this family. We also need to return to the same asteroids more than once in order to determine their spin axis orientation and shape.

From the photometric information of (832) Karin, if its surface color variation is real, this could be a firm explanation of the relationship between the spectrum of the asteroid surface and its dynamical history. We will keep observing this asteroid, which sometimes shows us a red surface and sometime does not, to determine its rotational and shape properties. We anticipate that the observation of this asteroid at the opposition in March 2006, when this asteroid will be observed at a different aspect angle from the Earth, will add to our knowledge of this intriguing asteroid.

Recent study has revealed that there are many more asteroid families that are as young as the Karin family: For example, an S-type cluster called the Iannini family is about 5 Myr-old, and a C-type cluster called the Veritas family is about 8.3 Myr-old.[17] We have also started photometric observation research on some of these young asteroid families to compare their characters with that of the Karin family as well as of well-known old families. In the near future, an impending deluge of large-scale sky surveys will yield a far larger amount of information with much higher accuracy about younger (and probably smaller) asteroid families, which will

be critical keys to understanding of the collisional and dynamical evolution of the main belt asteroids.

Acknowledgments

A large part of the result presented in this paper is based on our observations with the VATT (the Alice P. Lennon Telescope and the Thomas J. Bannan Astrophysics Facility). We thank Elizabeth Green, Paula White, Ed Olszewski and Andy Odell for our observing use at the Steward 90-inch telescope. The authors owe a lot to many people who helped us observe and analyze our asteroid data: Shigeru Takahashi, Mansur A. Ibrahimov, Budi Dermawan, Tsuko Nakamura, Toshifumi Yanagisawa, Masanori Hirai, Seidai Miyasaka, Hideo Fukushima, Hideo Sato, and Yusuke Sato, as well as Béatrice Mueller who benefited us with stimulating discussions. Two referees, Alberto Cellino and Maria J. Lopez-Gonzalez, suggested directions which bettered the quality of this paper a great deal. Detailed and constructive review by Yolande McLean has considerably improved the English presentation of this paper. This study is supported by the Grant-in-Aid of the Ministry of Education of Japan (16740259/2004–2005). F.Y. acknowledges financial support from the Sumitomo Foundation (030755/2003–2004).

References

1. D. Nesvorný, W. F. Bottke, L. Dones and H. F. Levison, *Nature* **417** (2002) 720.
2. D. Nesvorný and W. F. Bottke, *Icarus* **170** (2004) 324.
3. P. Michel, W. Benz and D. C. Richardson, *Nature* **421** (2003) 608.
4. F. Yoshida, in preparation.
5. A. U. Landolt, *Astron. J.* **340** (1992) 436.
6. A. W. Harris and D. F. Lupishko, in *Asteroids II*, eds. R. P. Binzel, T. Gehrels and M. S. Matthews (The University of Arizona Press, Tucson, 2002), p. 39.
7. N. R. Lomb, *Astrophys. Space Sci.* **39** (1976) 447.
8. D. H. Roberts, J. Lehar and J. W. Dreher, *Astron. J.* **93** (1987) 968.
9. B. Dermawan, T. Nakamura, H. Fukushima, H. Sato, F. Yoshida and Y. Sato, *Publ. Astron. Soc. Japan* **54** (2002) 635.
10. R. P. Binzel, *Icarus* **72** (1987) 135.
11. P. Magnussion, in *Asteroids II*, eds. R. P. Binzel, T. Gehrels and M. S. Matthews (The University of Arizona Press, Tucson, 2002), p. 1180.
12. F. Yoshida, B. Dermawan, T. Ito, Y. Sawabe, M. Haji, R. Saito, M. Hirai, T. Nakamura, Y. Sato, T. Yanagisawa and R. Malhotra, *Publ. Astron. Soc. Japan* **56** (2004) 1106.
13. D. L. Rabinowitz, *Icarus* **134** (1998) 342.

14. S. Sasaki, K. Nakamura, Y. Hamabe, E. Kurahashi and T. Hiroi, *Nature* **410** (2001) 555.
15. B. E. Clark, B. Hapke, C. Pieters and D. Britt, in *Asteroids III*, eds. W. F. Bottke, A. Cellino, P. Paolicchi and R. P. Binzel (The University of Arizona Press, Tucson, 2002), p. 585.
16. T. Sasaki, S. Sasaki, J. Watanabe, T. Sekiguchi, F. Yoshida, H. Kawakita, T. Fuse, N. Takato, B. Dermawan and T. Ito, *Astrophys. J.* **615** (2004) L161.
17. D. Nesvorný, W. F. Bottke, H. Levison and L. Dones, *Astrophys. J.* **591** (2003) 486.

This page intentionally left blank

DIFFERENCE IN DEGREE OF SPACE WEATHERING ON NEWBORN ASTEROID KARIN

TAKANORI SASAKI*, SHO SASAKI, JUN-ICHI WATANABE,
TOMOHIKO SEKIGUCHI, FUMI YOSHIDA, TAKASHI ITO,
HIDEYO KAWAKITA, TETSUHARU FUSE,
NARUHISA TAKATO and BUDI DERMAWAN
Department of Earth and Planetary Science, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
**takanori@eps.s.u-tokyo.ac.jp*

We have carried out a near-infrared (J, H, and K bands) spectroscopy of the asteroid Karin with cooled infrared spectrograph and camera for OHS on the Subaru telescope. This asteroid is the brightest asteroid among the Karin cluster group, which is thought to be remnants of a collisional breakup only 5.8 million years ago. For different rotational phases of Karin, we derived different spectra such as a reddened spectrum like that of S-type asteroid and an un-reddened spectrum like that of ordinary chondrite. Our result supports the idea that S-type asteroids are parent bodies of ordinary chondrites.

1. Introduction

Although S-type asteroids are the most common among the inner-part main belt asteroids as well as near-Earth asteroids, reddened reflectance spectra and derived mineralogy of S-type asteroids are different from those of ordinary chondrites, the most common meteorites. Space weathering is thought to be able to explain these spectral mismatches between asteroid types and meteorite classes.[1–3] Recent asteroid surveys suggest strong link between S-type asteroids and ordinary chondrites.[4,5] Multispectral observation of Ida by Galileo spacecraft showed that relatively fresh surface such as crater interiors and ejecta have reflectance like ordinary chondrites.[1] Moreover NEAR-Shoemaker mission revealed, thanks to X-ray and near-infrared spectrometer measurements, an ordinary chondrite composition of 433 Eros despite a reddened S-type spectrum, suggesting once again that the space weathering has altered the surface optical properties.[6,7]

The detailed mechanism of space weathering has been remained to be unsolved until recently, when the laboratory experiments using high-energy

*Corresponding author.

pulse laser irradiation[8] confirmed Hapke's hypothesis[9] that the reflectance change forming S-type spectra is caused by formation of nanophase iron particles within vapor-deposited rim around regolith particles. It was suggested that the degree of space weathering can be used to discuss the age of asteroids.[8] However, there had been no observational confirmation on the relation between asteroid age and the degree of the surface alteration. Here, we have an excellent target. Recently, using numerical integration of asteroid orbits, Nesvorný *et al.* revealed a new-born group of asteroids named Karin cluster group,[10] which is thought to be remnants of a recent breakup of only 5.8 million years ago.[10] In the present study, we observed the brightest asteroids 832 Karin in this group to consider the relation between the asteroid age and the effect of space weathering.

In this study, we observed the brightest asteroid 832 Karin (diameter being about 19 km, and absolute magnitude being 11.18[10]) in this group in order to investigate whether the new asteroid really has fresh and unreddened surface.

2. Observations and Data Reductions

A spectroscopic observation of Karin was carried out on September 14, 2003 UT using a cooled infrared spectrograph and camera for OHS (CISCO)[11] with the 8.2-meter SUBARU telescope. The slit size was $108'' \times 0.8''$ in our observation, which led no loss of light of the asteroid. The typical seeing size was 0.3–0.4″ during the observation. A nearby G2V star HIP3990 was observed as a spectroscopic standard. In order to obtain a wide range near-infrared spectrum, grisms named zJ (0.88–1.40 μm), JH (1.06–1.82 μm), and wK (1.85–2.51 μm) were used.

We obtained one-dimensional spectra of Karin by using the NOAO Image reduction and analysis facility (IRAF) aperture extraction package. We obtained OH sky emission lines frames at the end of the night to determine the linear dispersion for wavelength calibration. Relative magnitudes were computed using aperture photometry. Then, we scaled relative spectra to be consistent with each other by this photometry. For details refer to Ref. 12.

3. Results

We observed Karin at 7:57–8:40 UT, 8:46–9:29 UT, and 10:45–11:50 UT. The rotational period of Karin is 18.348 h, which was derived from a

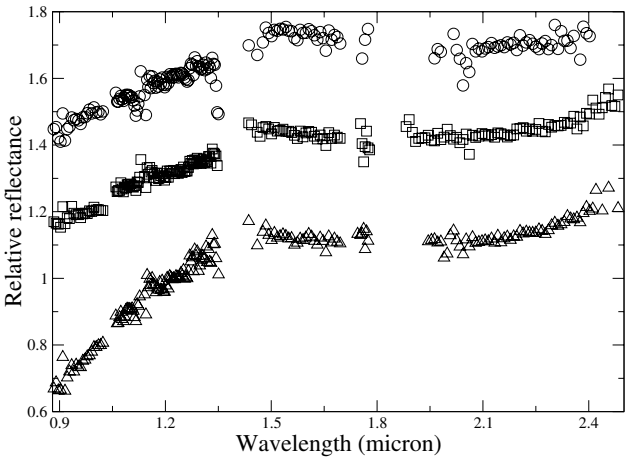


Fig. 1. Relative spectra of 832 Karin: bottom one is the spectrum relative to the first set, middle one is that relative to the second set, and top one is that relative to the last set. Spectra are smoothed by running average of five data points. The top and the bottom spectra are vertically shifted by +0.2 and by −0.2, respectively. We removed data in the wavelength range where the telluric absorptions were strong and the error bars are large.

light-curve obtained by supporting observations.[13] In comparison with the light curve, rotational phases of Karin in our observation are 0.30–0.34 (the first set), 0.35–0.38 (the second set), and 0.45–0.50 (the last set).

Figure 1 shows relative reflectance spectra of Karin. Bottom, middle, and top spectra in Fig. 1 are those of the first, the second, and the last observational sets, respectively. There is obvious difference between the top two and bottom spectra. Near-infrared (0.9–1.4 μm) reflectance slope of the bottom spectra is twice as steep as that of the top spectra. In general, major color changes with rotation are very rarely observed on asteroids, e.g., only a little difference in gradients of spectra at different rotational phase was observed in Vesta.[14] The present color change of Karin would be the biggest color change ever observed with rotational phase.

The shape of 0.88–2.5 μm in the first set's spectrum with a steep slope at shorter wavelength is consistent with an S-type object. We identified which S subclass of the classification scheme[15,16] can best describe the Karin spectrum. Since the first set's spectrum of Karin has the peak position of 1 μm band shorter than 1.0 μm and has an apparent 2 μm band, it is placed among the range of S(III), S(IV), and S(V) classes.

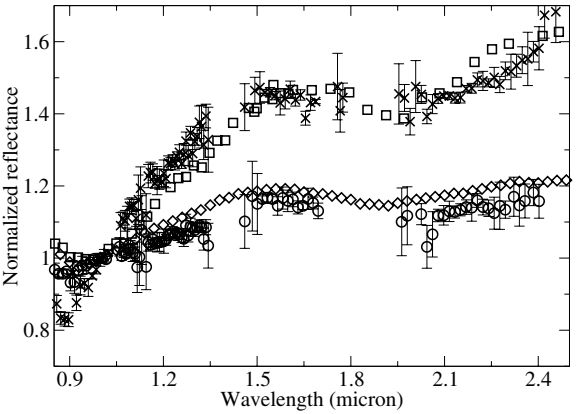


Fig. 2. Reflectance spectra of Karin (the first set: ×, and the last set: ○) along with the previous observations of S(IV)-type asteroid 584 Semiramis (denoted by squares) and L6 ordinary chondrite Paranaiba (denoted by diamonds) normalized to the unity at 1.0 μm. The mean values for every 10 data points are plotted with their error bars of standard deviation.

Figure 2 shows normalized reflectance spectra of Karin (the first set and the last set) along with those of S(IV)-type asteroid 584 Semiramis and L6 ordinary chondrite Paranaiba. The first set's spectrum of Karin matches well the spectrum of S(IV) class asteroid. In contrast, the last set's spectrum (blue) of Karin, which has a relatively gentle slope at the wavelength shorter than 1.6 μm, matches (within noise of our spectrum) a typical normalized spectrum of L6 ordinary chondrite. The first set seems to be the reddened spectrum of the last set by space weathering. Laboratory simulations of space weathering also showed relatively intense reddening at wavelength shorter than 1.6 μm.[8,17]

4. Discussions

Our result indicates that Karin has two different surfaces, reddened and un-reddened surfaces, and the difference among spectra would reflect the degree of space weathering. The mature and fresh surfaces' spectra strongly stand up for the idea that space weathering is responsible for the mismatches between asteroid types and meteorite classes. Our result supports the idea that S-type asteroids are parent bodies of ordinary chondrites.[1] At rotational phase of 0.35, which is the boundary between the first set and the second set, abrupt shift in the degree of space weathering is observed.

As cross-section area of Karin increases (Karin being brighter), matured redder surface is replaced by immature fresh surface. The light-curve of Karin has double peaks of magnitude (where phase are around 0.2 and 0.7, respectively), and the magnitude at 0.2 is larger than that at 0.7, which could imply that the surface around 0.2 is mature and darkened by space weathering while the surface around 0.7 is fresh, although shape difference might be more responsible for this peak height difference.

Let us reflect the spectral change of Karin according to rotational phase. Around rotational phase of 0.35, which is the boundary between the first set and the second set, rapid shift in the degree of space weathering is observed. As cross-section area of Karin increases (Karin being brighter), matured redder surface is replaced by immature fresh surface. More reddened spectra could be observed at phase earlier than 0.3, and the spectrum of our first set of Karin could be an averaged data of un-reddened spectrum and much more reddened spectrum. Actually, visible observations of Karin also suggested that the surface at earlier rotational phase (<0.5) is weathered compared with that at latter phases (>0.5).[13] They also showed the surface at rotational pahse of 0.2 can be most matured (dark and red). Their results are consistent with our observation.

5. Planning for Re-observation

What is very puzzling in our results is that the major spectral change occurs between sets of averaging only 50 min apart, which represents only about 15° of rotation of the asteroid. We verified that the standard star's spectra were not changed before the first set and after the second set of Karin observation, i.e., we would remove the possibility that the spectral change was caused by an instrumental or atmospheric effect through the first and second sets of Karin observation.[12] One hypothesis to explain the observed rapid spectral change is that Karin is one of cone-shaped fragment at low-velocity impact forming the Karin family,[12] however, the problem is still unsolved. It is necessary to perform spectral observation seamlessly in full rotational phase.

Actually, we have conducted a multispectral Karin observation using the 1 m Schmidt telescope at the Kiso observatory on January 14–18, 2005 to get the spectral data in full rotational phase. But, unfortunately, the weather was so bad through the observation that we cannot get a meaningful data. Now, we are planning for re-observation of the asteroid on the next year. This will be the proof that the variation is intrinsic to the rotation of

the asteroid and not due to any other instrument, observing, atmospheric, airmass, hour angle, or dispersion angle effect.

Acknowledgments

We thank J. F. Bell for asteroid Semiramis data and C. M. Pieters for meteorite Paranaiba data. A review by R. P. Binzel has improved the paper.

References

1. C. R. Chapman, *Meteor. Planet. Sci.* **31** (1996) 699.
2. C. M. Pieters *et al.*, *Meteor. Planet. Sci.* **35** (2000) 1101.
3. B. E. Clark *et al.*, in *Asteroid III*, eds. W. F. Bottke *et al.* (Arizona Press, 2003), p. 585.
4. R. P. Binzel *et al.*, *Science* **273** (1996) 946.
5. T. H. Burbine and R. P. Binzel, *Icarus* **159** (2002) 468.
6. L. A. McFadden *et al.*, *Meteor. Planet. Sci.* **36** (2001) 1711.
7. J. F. Bell III *et al.*, *Icarus* **155** (2002) 119.
8. S. Sasaki *et al.*, *Nature* **410** (2001) 555.
9. B. Hapke, W. Cassidy and E. Wells, *Moon* **13** (1975) 339.
10. D. Nesvorný *et al.*, *Nature* **417** (2002) 720.
11. K. Motohara *et al.*, *PASJ* **54** (2002) 315.
12. T. Sasaki *et al.*, *ApJ* **615** (2004) L161.
13. F. Yoshida *et al.*, *PASJ* **51** (2004) 1255.
14. M. J. Gaffey, *Icarus* **127** (1997) 130.
15. M. J. Gaffey *et al.*, *Icarus* **106** (1993) 573.
16. E. A. Cloutis *et al.*, *J. Geophys. Res.* **91** (1986) 11641.
17. M. Yamada *et al.*, *Earth Planets Space* **51** (1999) 1255.

SIZE DISTRIBUTION OF ASTEROIDS AND OLD TERRESTRIAL CRATERS: IMPLICATIONS FOR ASTEROIDAL DYNAMICS DURING LHB

TAKASHI ITO*,†, ROBERT G. STROM‡, RENU MALHOTRA‡,
FUMI YOSHIDA† and DAVID A. KRING‡
†*National Astronomical Observatory of Japan*
Osawa 2–21–1, Mitaka, Tokyo 181–8588, Japan
‡*Lunar and Planetary Laboratory, The University of Arizona*
Tucson, AZ 85712–0092, USA
**tito@cc.nao.ac.jp*

Recent progress in asteroid surveys has revealed the fine structures down to sub-km in diameter of the size-frequency distributions (SFD) of main belt asteroids (MBAs), as well as near-Earth asteroids (NEAs). These SFDs can be compared with the SFD of lunar and planetary crater projectiles. The SFD of the projectiles that created the oldest craters on the lunar highlands, which are considered a fossil of the Late Heavy Bombardment (LHB) impactors of ∼4 Ga ago, shows a very good agreement with that of the current MBAs. This fact indicates that the LHB craters were created by the bombardment of ancient asteroids ejected from the main belt by a short-term, size-independent event, such as the radial movement of strong resonances due to the migration of giant jovian planets. On the other hand, the SFD of the projectiles that have created younger craters such as those on Mars is very different from that of the MBAs; instead, it is quite similar to the SFD of NEAs. This newer population of projectiles might be created by a long-term, size-dependent transportation mechanism of asteroids such as the Yarkovsky effect, which preferentially pushes smaller objects into strong resonances.

1. Introduction

At the dawn of Earth's history, there were intense and cataclysmic impact events, collectively called the Late Heavy Bombardment (LHB).[1,2] The most intense period of the LHB appears to have occurred ∼3.9 Gyr ago, i.e., 500–600 million years after the formation of the Earth–Moon system.[3,4] Evidence of this event began to accumulate when Ar–Ar isotopic analyses of Apollo and Luna samples suggested that several impact basins on the nearside of the Moon had been produced 3.88 and 4.05 Ga. Additional analysis of Apollo samples indicated the U–Pb and Rb–Sr systems had been disturbed nearly uniformly at ∼3.9 Ga, which was attributed to metamorphism of

some portions of the lunar crust by a large number of collisions in a short time, less than 200 million years.

To better characterize the cause, mechanism, and extent of LHB, we can resort to the recent progress in extensive asteroid surveys that have revealed fine structures of the size-frequency distribution (SFD) of main belt asteroids (MBAs) and near-Earth asteroids (NEAs). These data can be compared with the SFD of the crater projectiles on the Moon and on other planets. In this paper, we provide compelling new evidence that the source of the LHB impactors was the main asteroid belt, and that the dynamical mechanism that caused the LHB was unique in the history of the solar system and distinct from the processes producing the flux of objects that currently hit planetary surfaces.[5]

2. Crater SFDs

Throughout this manuscript, SFDs of craters and asteroids are expressed by the so-called R-plot, which expresses differential SFD of crater/asteroid populations relative to D^{-3}, where D denotes diameter.[6] Since many populations of inner solar system small bodies and craters have differential SFDs (dN/dD) more or less proportional to D^{-3}, it is reasonable to normalize them by D^{-3} so that we can see their differences in detail. Also, R values of craters are generally divided by the surface area A, where we count the number of craters in order to estimate the number density of craters; R is defined as $R \equiv (D^3/A) \times dN/dD$.

Expressing the SFD of lunar and planetary craters by R-plot, clearly we see two distinctive SFD populations (Fig. 1). The first crater population is what is typically observed on the oldest lunar highlands; LHB craters as old as $\sim$4 Ga. This crater population is characterized by a wavy R curve as in Fig. 1(a), which is also seen on the oldest highlands on Mercury (Fig. 1(b)) as well as on Mars (Fig. 1(c)).

There is another crater population characterized by rather flat R patterns. These craters are younger than the LHB craters, and their number density is lower. In general, these craters have a wide variety of ages, indicating they have been formed over a timespan as long as Gyr. A typical example of this population is seen on young and smooth northern martian hemisphere where there are a lot of relatively new craters. The R curve for these craters is almost flat, showing $R \propto D^{-3}$, as in the lower part of Fig. 1(c). Another example of this crater population is observed on the Moon as Class-1 craters with quite fresh morphology. These craters also

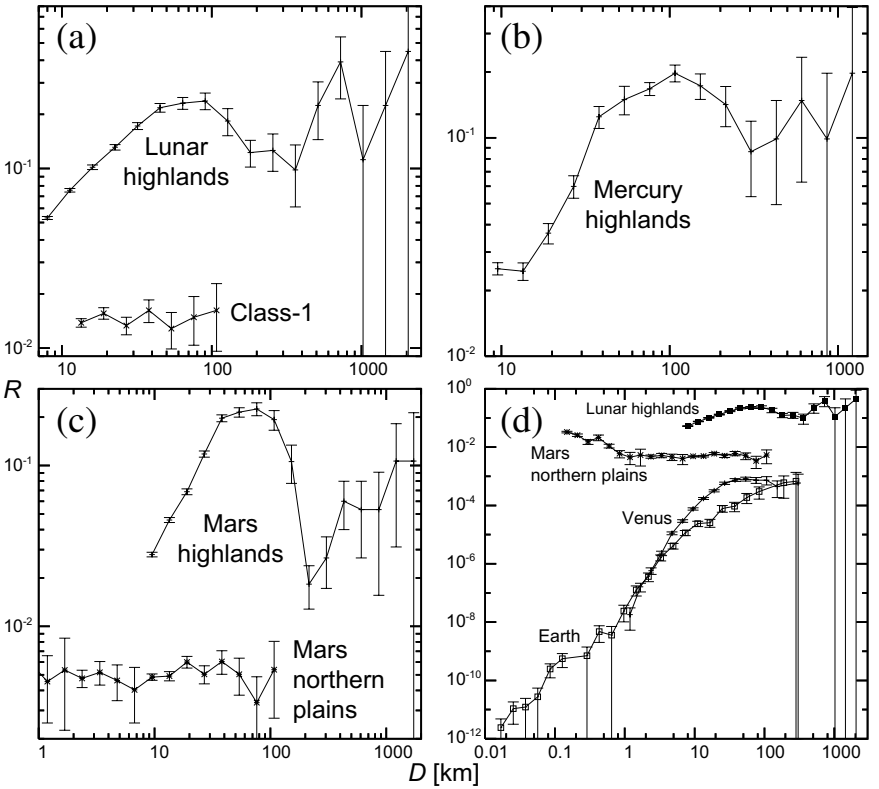


Fig. 1. *R*-plots of some crater SFDs. (a) Lunar highlands craters and lunar Class-1 craters. (b) Craters on the oldest mercurian highlands. (c) Craters on the oldest martian highlands and on young northern plains. (d) Craters on Venus and on the Earth, compared with the lunar highlands craters and young martian craters.

have flat R curves with lower density than the old highland craters (lower left part of Fig. 1(a)). The crater records on Venus and the Earth have been severely obliterated and lost, or many of their projectiles have been screened out by the atmosphere, and we cannot use them to estimate their projectile sources (Fig. 1(d)).

The similarities amongst the wavy R curves of the oldest crater populations on the Moon, Mercury, and Mars indicate that they were created by a single projectile population at the same age, probably during LHB with a short timescale. On the other hand, the younger crater populations characterized by the flat R curves having a variety of ages have presumably been created by a different projectile population with a different mechanism.

3. Asteroid SFDs

Chemical analyses of Apollo samples of impact melts point to a dominantly asteroid reservoir for LHB impactors, rather than to cometary objects.[4] Also, recent asteroid surveys with high resolution and large sky coverage such as Spacewatch, SDSS, or LINEAR have given us a significant degree of understanding of the SFD of MBAs as well as of NEAs. However, we cannot directly compare the SFD of asteroids with that of craters; it is necessary to convert crater SFD into that of projectiles (or the other way around). Procedures of this sort have already been established with the help of the scaling relationship between crater size and projectile size. In this paper, we assume the typical asteroidal impact velocity and asteroidal density as those of crater projectiles, and convert crater SFD into projectile SFD so that we can compare the SFD of crater projectiles with asteroid SFD. As for the SFD of MBAs, we use the results of three survey programs: Spacewatch,[7] SDSS,[8] and Subaru.[9] We rely on the results of the LINEAR program[10] for the SFD of NEAs. We used the so-called Pi scaling laws[11–14] to derive the projectile size from the crater size.

When we draw the asteroid SFDs on R-plot graphs, we immediately see their remarkable similarities to the SFD of crater projectiles. At first, the SFD of MBAs fits that of lunar LHB crater projectiles very well over wide diameter ranges (Figs. 2(a)–(c)). Second, the SFD of NEAs seems quite close to the SFD of the younger projectile population that has created craters on the younger martian plains (Fig. 2(d)). Thus, we obviously have two distinct SFD populations among current asteroids, not only among craters. These similarities should be more than just a coincidence, having firm physical/dynamical reasons.

4. Discussion

The fact that the ancient LHB projectiles had an SFD almost identical to that of the current MBAs could imply several things: First and foremost, there was a size-independent transport process for asteroids during the LHB period. If the LHB duration was as short as 50–200 Myr as previous research suggests,[1] the timescale of the transport process must have been as short. Not so many dynamical mechanisms can make this drastic asteroid transport happen. A plausible candidate is the radial movement of strong resonances in the main belt caused by the migration of giant jovian planets.

Currently some ideas along this line are being proposed in terms of the late formation of Uranus and Neptune and their interaction with

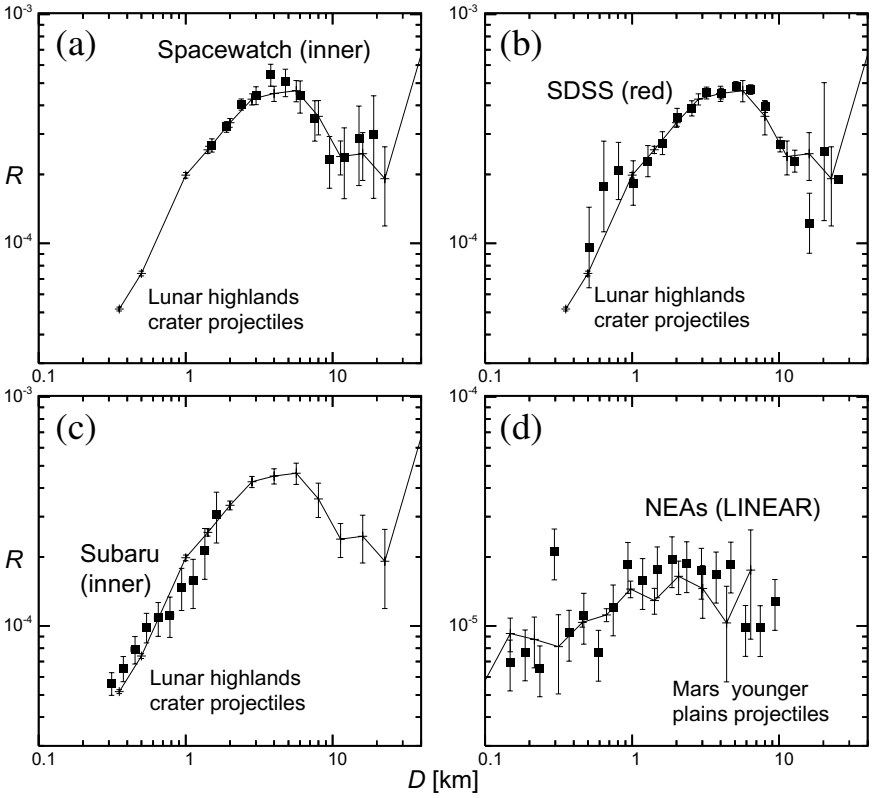


Fig. 2. Comparison between the SFDs of current asteroids (the symbol ■ with errorbars) and crater projectiles (solid lines with errorbars). (a)–(c) are for the oldest lunar highlands crater projectiles and the MBAs mainly in the inner belt surveyed by (a) Spacewatch,[7] (b) SDSS,[8] and (c) Subaru.[9] (d) is the comparison between the young martian crater projectiles and NEAs surveyed by the LINEAR project.[10] Note that the R range is different in (d) from other panels.

planetesimals.[15,16] LHB occurred too late to invoke a nebula gas dissipation as the cause of resonance sweeping, so the only alternative to provoke the resonance sweeping is that the giant planets migrated at that time due to interaction with a swarm of planetesimals. The planetesimal disk must be massive enough to make giant planets radially migrate, hence it should be a distant, massive planetesimal disk beyond the large planets. In addition, a mechanism is needed to produce a late start of giant planet migration around 4 Gyr ago. One possible theory invokes the change

in the eccentricities of Jupiter and Saturn when they pass through a 1:2 mean motion resonance during their orbital migration[16] under the gravitational influence of a swarm of planetesimals. Such a resonance passage would have destabilized the planetesimal disk beyond the orbits of the large planets, causing a sudden massive delivery of cometary planetesimals to the inner solar system. In this scenario, the asteroid belt is also destabilized because of sweeping gravitational resonances; together, these cause a major spike in the intensity of cometary as well as asteroid impacts on the inner planets. The relative intensity of comets versus asteroids in the projectile population of the LHB is not well determined in currently published dynamical simulations. Because the impact signature of the crater record in the inner solar system is asteroidal, we conclude that either comets played a minor role or their impact record was erased by later-impacting asteroids.

Another important implication of our results comes from the fact that we have compared the SFD of 4 Gyr-old projectiles with that of the current MBAs, and found them strikingly similar. This could mean that there has been almost no collisional evolution in the main asteroid belt over the last $\sim$4 Gyr, ever since the LHB. This is seemingly weird, but recent numerical models of the collisional evolution of MBAs support this fact,[17,18] revealing the rather stationary SFD of MBAs. Therefore, it is probably safe to regard the SFD of the current MBAs as a fossil of the LHB projectiles.

From the comparison between the SFDs of younger crater projectiles and the current NEAs, it seems that NEAs have been the impacting source of the newer craters since LHB ceased. Although most NEAs are considered to have originated in MBAs from a dynamical point of view,[19] we have a greater number of smaller objects among the NEA population than the MBA population, judging from the slope difference between Figs. 2(a)–(d). This evidence, in addition to the wide variety of ages of the young crater population, leads us to the conclusion that there has been a size-dependent, long-term transport process that conveys MBAs (preferentially smaller ones) to the inner solar system. A plausible candidate for this kind of mechanism is the Yarkovsky effect, a (generally slow) dynamical effect caused by the thermal time lag of asteroids when they absorb and re-emit solar radiation. Since the Yarkovsky effect works much more effectively on smaller asteroids, it is perfectly eligible to selectively transport small asteroids from the main belt to the terrestrial planet region over a long timespan.

Acknowledgments

The authors have greatly benefited from the electronic version of asteroid SFD data provided by Željko Ivezić and Joseph Scott Stuart. Two anonymous referees suggested directions which bettered the quality of this paper. Detailed and constructive review by Yolande McLean has considerably improved the English presentation of this paper. This study is supported by the Grant-in-Aid of the Ministry of Education of Japan (16740259/2004–2005). F.Y. acknowledges financial support from the Sumitomo Foundation for the research funding (030755/2003–2004).

References

1. G. Ryder, *EOS, Trans. Am. Geophys. Union* **71**, 10 (1990) 313.
2. W. K. Hartmann, G. Ryder, L. Dones and D. Grinspoon, in *Origin of the Earth and Moon*, eds. R. M. Canup and K. Righter (University of Arizona Press, Tucson, 2000), p. 493.
3. B. A. Cohen, T. D. Swindle and D. A. Kring, *Science* **290** (2000) 1754.
4. D. A. Kring and B. A. Cohen, *J. Geophys. Res.* **107**, E2 (2002) 4–1.
5. R. G. Strom, R. Malhotra, T. Ito, F. Yoshida and D. A. Kring, *Science* **309** (2005) 1847.
6. B. A. Ivanov, G. Neukum and R. Wagner, in *Collisional Processes in the Solar System,* eds. M. Ya. Marov and H. Rickman (Kluwer Academic Publishers, Dordrecht, 2001), p. 1.
7. R. Jedicke and T. S. Metcalfe, *Icarus* **131** (1998) 245.
8. Ž. Ivezić and the SDSS collaboration, *Astron. J.* **122** (2001) 2749.
9. F. Yoshida, T. Nakamura, J. Watanabe, D. Kinoshita, N. Yamamoto and T. Fuse, *Publ. Astron. Soc. Japan* **55** (2003) 701.
10. J. S. Stuart and R. P. Binzel, *Icarus* **170** (2004) 295.
11. K. A. Holsapple and R. M. Schmidt, *J. Geophys. Res.* **87**, B3 (1982) 1849.
12. S. K. Croft, *J. Geophys. Res.* **90** (1985) 828.
13. R. M. Schmidt and K. R. Housen, *Int. J. Impact Eng.* **5** (1987) 543.
14. H. J. Melosh, *Impact Cratering: A Geologic Process* (Oxford University Press, New York, 1989).
15. H. F. Levison, L. Dones, C. R. Chapman, S. A. Stern, M. J. Duncan and K. Zahnle, *Icarus* **151** (2001) 286.
16. R. Gomes, H. F. Levison, K. Tsiganis and A. Morbidelli, *Nature* **435** (2005) 466.
17. A. F. Cheng, *Icarus* **169** (2004) 357.
18. W. F. Bottke, D. D. Durda, D. Nesvorný, R. Jedicke, A. Morbidelli, D. Vokrouhlický and H. F. Levison, *Icarus* **175** (2005) 111.
19. A. Morbidelli and D. Vokroulický, *Icarus* **163** (2003) 120.

This page intentionally left blank

STATUS OF THE TAOS PROJECT AND A SIMULATOR FOR TNO OCCULTATION

SUN-KUN KING[†,*], CHARLES ALCOCK[‡], TIM AXELROD[§], FEDERICA B. BIANCO[¶,‡], YONG-IK BYUN[‖], WEN-PING CHEN[‖‖], KEM H. COOK[††], YUNG-HSIN CHANG[‖‖], RAHUL DAVE[¶], JOSEPH GIAMMARCO[¶], TYPHOON LEE[†], MATTHEW LEHNER[‡,¶], JACK LISSAUER[‡‡], STUART MARSHALL[§§], SOUMEN MONDAL[‖‖], IMKE DE PATER[¶¶], RODIN PORRATA[¶¶], JOHN RICE[¶¶], MEGAN E. SCHWAMB[¶], ANDREW WANG[†], SHIANG-YU WANG[†], CHIH-YI WEN[†] and ZHI-WEI ZHANG[‖‖]

[†]*Institute of Astronomy & Astrophysics, Academia Sinica, Taiwan, ROC*
[‡]*Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA*
[§]*Steward Observatory, The University of Arizona, Tucson, AZ, USA*
[¶]*Department of Physics and Astronomy, University of Pennsylvania, PA, USA*
[‖]*Department of Astronomy, Yonsei University, Seoul, South Korea*
[‖‖]*Institute of Astronomy, National Central University, Chung-Li, Taiwan, ROC*
[††]*Lawrence Livermore National Laboratory, Livermore, CA, USA*
[‡‡]*NASA, Ames Research Center, Mountain View, CA, USA*
[§§]*Stanford Linear Accelerator Center, Stanford, CA, USA*
[¶¶]*University of California, Berkeley, CA, USA*
[*]*skking@asiaa.sinica.edu.tw*

The majority of trans-Neptunian objects (TNOs) are probably small comets beyond the orbit of Neptune. A study of TNOs may enable a better understanding of the origin of short-period comets and of the process of planet formation and the early history of the solar system. An occultation survey is currently the only way to detect these objects down to a size of a few kilometers at such a distance. The status of the Taiwan–America Occultation Survey (TAOS) project is reported. In order to monitor thousands of stars on the order of a fraction of a second using CCD cameras, a novel CCD readout technique, the "shutterless zipper mode," is applied. Two predicted asteroid occultation events were successfully observed. Instead of a simple number count of occultation events, an interpretation of a TNO occultation survey result can be obtained by using the simulator described here. Through comparison of the results from an observation and from our simulator, a specific astronomical or astrophysical model can be constrained.

[*]Corresponding author.

1. Introduction

It has long been suspected that short-period comets might have a different origin than those of long-periods.[1,2] The existence of this Edgeworth–Kuiper belt was not realized until the discovery of the first (other than Pluto and Charon) trans-Neptunian object (TNO) 1992 QB_1 by Jewitt and Luu.[3] Currently, more than a thousand of them have been detected. Some of them are even compatible in size with Pluto.[4] The discovery of Sedna[5] with a perihelion around 76 AU might indicate the existence of a whole new family of TNOs in the outer solar system, where the gravitation of Neptune plays a less significant role. However, the distribution of comet-size (a few kilometers) objects is still poorly known. Preliminary results from recent observation[6] and simulation[7] show the possibility of a broken power law in the TNO size distribution. Some physical properties can be derived for the largest TNOs, but certain properties like albedo would be pure speculation and extrapolation for such small objects. One cannot study those comets in detail until their orbits bring them closer to us, such as some of the Centaurs. On the other hand, to explain the current population of the Jupiter-family comets, which is generally believed to originate from the trans-Neptunian region, 10^9–10^{10} comets of a size around 1–10 km in 30–50 AU might be required.[6,9]

Instead of trying to detect the reflected light from those distant small objects, an occultation survey[8] could probe a comet-size TNO at a few hundreds AU or even farther away. The Taiwan–America Occultation Survey (TAOS) project[10–12] has a design based on this indirect strategy. Some other, smaller in scope, surveys have been conducted in recent years,[14,15,26] though convincing statistics describing small comet occultations have yet to be obtained. However, future improvement in both ground-based and space-based experiments are expected to provide deeper insight. We briefly report the status of TAOS in Sec. 2. Related information and details are available at the TAOS website.[a] Diffraction and the angular size of a target star[11,13,14] are two major factors which determine the "detectability" of a presumed occultation event or the "visibility" of an occultation dip in a light curve. Occultation is a technique which is capable of detecting a small comet as far as the inner Oort cloud. A poor distance resolution may be obtained from the angular size of a target star and the size of the foreground TNO. Recent analysis[16] suggest that it might be possible to

[a] http://taos.asiaa.sinica.edu.tw

break the degeneracy between size and distance under certain conditions when a cylindrical symmetric solution and a center-crossing event were assumed. It is not clear whether this will still be the case with a more complete parameter space when certain photometric error or noise is considered. Starting from the theory of diffraction, an ensemble of light curves can be derived with a specific astronomical distribution and/or astrophysical model assumed. These light curves will be fed into a data pipeline for photometric analysis and efficiency test. The output of this simulator can be compared with an observational result, which will be a test against the astronomical/astrophysical model assumed earlier. Details and progress of our simulator and some additional comments on a simple geometric model are presented in Sec. 3.

2. Taiwan–America Occultation Survey Status and Zipper Mode Test

The objective of TAOS is to simultaneously monitor thousands of stars in a field with three or four robotic telescopes running in synchronous mode at a rate of 5 Hz (the rate can be varied). This rate corresponds to a shadow of a few kilometers in size passing through a specific site at a relative speed around 25 km/s, which is dominated by the orbital velocity of the earth itself. These four dedicated small (50 cm) telescopes have a wide field of view ($1.7^\circ \times 1.7^\circ$). The use of at least three telescopes is to reduce false positives below a statistically significant threshold.[17,18] While having the telescopes separated by several kilometers would have provided better resolution and the possibility to determine size and shape directly, practical considerations have lead to have all the telescopes operating at one site within a hundred-meter range. Follow-up observation with a larger telescope could be helpful in telling a foreground asteroid event apart.

Instead of using a multi-object photometer[15] or fiber optics, a high-quantum efficiency, back-illuminated CCD camera (SI-800) with 2048×2048 pixels[b] was installed on each telescope. Following the experience from the ROTSE project,[19] control software has been developed which can perform a sky patrol automatically. A special function which can respond to a gamma-ray-burst coordinates network (GCN) alert was also integrated with the software. Given the high data rate of TAOS, each telescope might

[b]There are actually four more rows unmasked. Some of them have different sensitivities.

collect 20–30 GB of data in a clear night, a nearly real-time data pipeline is needed for processing.

The typical readout time of our 2-channel CCD camera is around 2 s. To use a CCD for a 5 Hz rate survey mission, a special technique is required. The strategy we use is called shutterless "zipper" mode operation.[17,18] This technique takes advantage of the standard CCD readout process, where photoelectrons have to be read out sequentially in each channel. The readout time for each row (2048 pixels) is approximately 1 ms. With some modification in both its firmware and our control software, the readout process can be "held" for, say, 200 ms (the "hold-time") after a few rows were read out. These few rows (for instance, 64 rows) of data constitutes a fundamental unit of TAOS raw data, namely, a "rowblock". All stars and sky patches in a 2048×2048 field will be squeezed into each rowblock (say, 2048×64) at the final readout. However, a star image in each rowblock might be an exposure in one of the, say, 32 cycles ($32 = 2048/64$). A different star in the same rowblock may come from an exposure in a different cycle. Figure 1 is a simplified version which illustrates some special features of this operation. The rowblock readout reaches a steady state after a few cycles. (It takes four cycles in Fig. 1.) In reality, stars spread across the whole field randomly as shown in Fig. 1(e). Without an original stared image, there would be no way to recover the coordinates of a star from a rowblock. This hold-readout cycle can be continued indefinitely with the shutter left open and the telescope tracking a target field. Usually, tens of rowblocks will be stored as a single file for the convenience of data processing.

Through remote control, two predicted asteroid occultation events were recorded successfully in our zipper mode images as shown in Figs. 2 and 3. In the first event, a bright star HIP 079407 ($m_V = 8.80$) was occulted by an asteroid (51) Nemausa ($m_V = 11.9$, diameter = 150 km) at around 18:55 21 February 2004 (UTC). The hold time of each rowblock is about 0.5 s in Fig. 2. A magnitude drop of about 3.5 magnitude for approximately 10 s was detected by the TAOS B telescope. In the second event, a bright star HIP 050535 ($m_V = 8.6$) was occulted by (1723) Klemola ($m_V = 15.7$, diameter = 31 km) at around 12:10 5 June 2004 (UTC). The hold time is 0.25 s in Fig. 3. An occultation lasted for approximately 1 s was recorded by two telescopes simultaneously.

There are a few disadvantages to shutterless zipper mode which must be addressed. For fast data acquisition the cameras are run in shutterless zipper mode. The shutter remains open, small rowblocks (e.g., 64 rows) of the CCD are read out (readout time less than 200 ms). The CCD counts

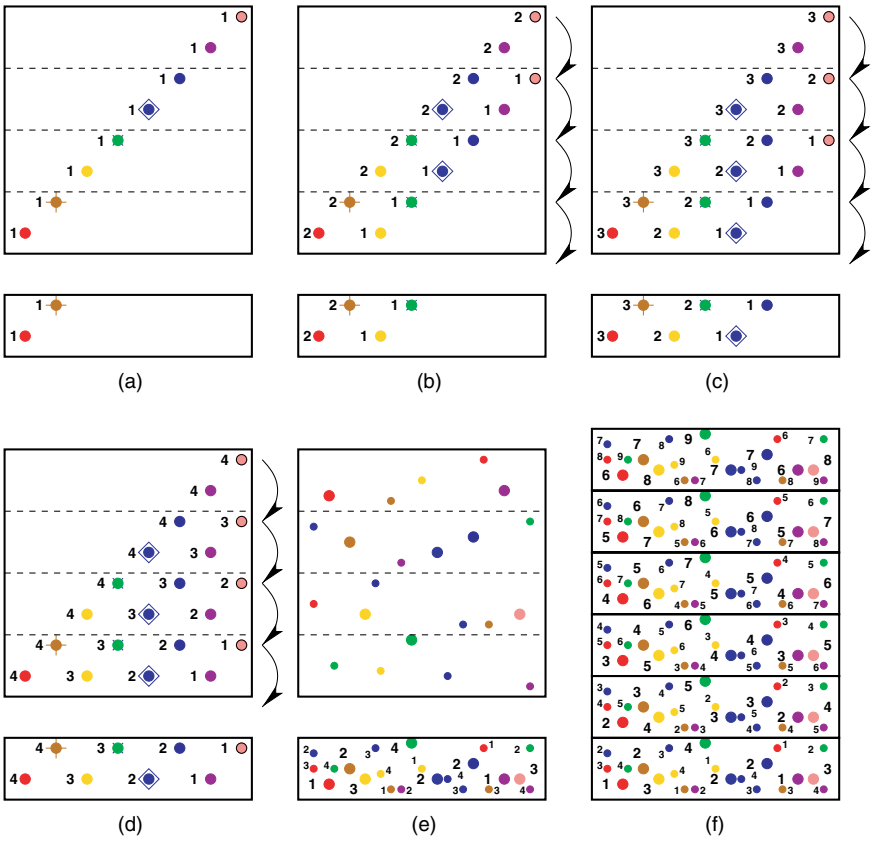


Fig. 1. Shutterless zipper mode. (a)–(d) illustrate the process of CCD readout under shutterless zipper mode. A square CCD chip covers a field where eight different stars happen to align across that field as shown in (a). One rowblock, which is shown right below each "CCD snapshot", is read out in each cycle. The readout direction is downward. It reaches a steady state at (d) after four cycles here while all eight different stars in this field are "squeezed" into one rowblock. Moreover, the star images in each rowblock may come from exposures in different cycles (as shown by their numerical labels). In reality, target stars spread over a field randomly as illustrated in (e). It is impossible to reconstruct a field based on its zipper image. The TAOS raw data is collected as a stack of tens of rowblocks usually. Figure 1(f) shows what it may look like if the target field is the same one as shown in (e). But, keep in mind that as many as 32 cycles might be involved in a TAOS rowblock, which is too complicated to be shown here clearly. To move a whole rowblock downward takes time as well. Therefore, a bright star usually leaves a clear streak behind which is really exposure during the movement (not shown in the plot).

Fig. 2. Occultation of HIP 079407 by (51) Nemausa on February 21, 2004 (UTC). Part of our raw zipper data is shown here. There are 32 rowblocks (90° rotated) in this figure. Each rowblock includes 2048×64 pixels, where 151×64 pixels are shown. Each pixel corresponds to 3 arcsec in the sky. Starting from the left, each cycle lasted for about 0.5 s. The two circles, centered on the target star, indicate the start and the end of this occultation roughly. The asteroid itself ($m_{\mathrm{V}} = 11.9$) is barely visible during occultation where the target star is missing for 10 rowblocks.

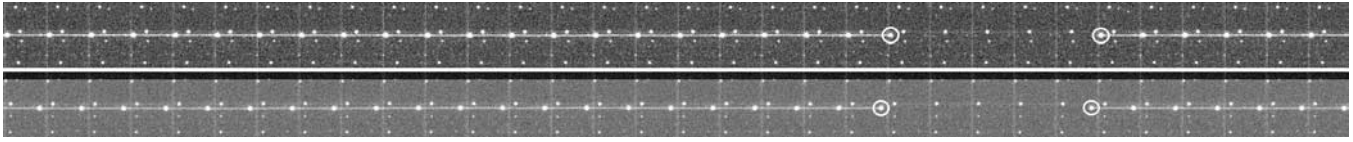


Fig. 3. Occultation of HIP 050535 by (1723) Klemola on June 5, 2004. It is similar to Fig. 2 except that data from two telescopes A (top) and B (bottom) are shown here. Each cycle lasted for about 0.25 s, only. The target star HIP 050535 ($m_{\rm V} = 8.6$) is clearly missing in four rowblocks between the two circles in both cases. This is the signal for an occultation lasted for about just 1 s. The asteroid itself is too faint ($m_{\rm V} = 15.7$) to be identified during this event in our zipper image. Thirty-two rowblocks with 100×64 pixels per rowblock are shown. The two telescopes pointed at slightly different directions. Hence, stars were "zipped" into slightly different patterns.

are rapidly shifted with respect to the star by the size of the rowblock. By the time a rowblock has traversed from the top of the CCD to the bottom where it is read out, it will accumulate all the stars that were in the original 2048 row image into a rowblock. Depending on sky and seeing conditions, a different size of the rowblock can be used. One downside of this technique is that sky background is always collected especially during each hold-time. The sky background could be 5 or 6 times ($\sqrt{2048/64}$) brighter than that of a stared image. Because it takes a finite amount of time to read in each row of the rowblock, a small percent of each star's flux is imprinted in each row creating streaks in the image for stars brighter than 10th magnitude. This will further decrease the number of usable stars for detections. Closing the shutter during readout can remove the streak if a fast and robust shutter is available. To reduce the sky background, one will need something different. For instance, the use of a mask, a fiber system or a novel electronic design such as a multi-channel frame-transfer CCD or CMOS should have an improvement in the sky background level by one or two magnitudes.

Nevertheless, the system capability was demonstrated successfully. TAOS is supposed to conduct a "blind" occultation survey. That does not prevent us from observing a rare predicted event whichever might come close to the TAOS site. A local network is organized for such an event as well. Direct measurements of known Centaurs (e.g., Pholus[20]) or big TNOs[21] are possible. The light curves derived can also be used to study a variable star which varies on a larger time scale.[22] Similar techniques have been developed elsewhere independently to detect a lunar occultation event[23] and to study speckle imaging of binary stars.[24]

3. A Simulator for TNO Occultation Survey

Synchronous mode operation with three TAOS telescopes began in the winter of 2004. Reliable light curves, free from bad weather, poor seeing, lost tracking and other non-ideal conditions, are expected to be produced in large amounts. An immediate question would be, "What do the occultation survey results mean?" First of all, it is straightforward to compare the event number counts with a model prediction. The simplest model should be a geometric model where a certain solid angle is occupied by the foreground TNOs (with some kind of size and orbital distribution assumed) for a given time duration. A certain probability that a background star will be occulted can be calculated.[11,14,25,26] However, it has also been pointed

out that a geometric picture may be too simple to be true because diffraction effect may increase the event rate drastically.[14] The number of TNO occultation events is dominated by the faint-end objects where, near the diffraction limit, Fresnel diffraction applies. The criterion of an occultation event is really an artificial one. An arbitrary cut-off may introduce an error which could be as big as the number count itself. Moreover, the so-called "occultation event rate" depends on the system and the algorithm used. We'd say that the geometric picture is good for an estimate, but, something else is needed for an interpretation. A simulator is just what we need to understand the underlying astronomy and astrophysics.

The following steps are taken to develop the simulator. First, an ensemble of light curves are derived from a specific astronomical or astrophysical model. Next, all these light curves will be run through a specific data pipeline with a photometric algorithm applied. An efficiency test is also required here. By adding a fake signal, the efficiency of an algorithm can be obtained. The final result can now be compared with an observation to justify the model assumed. To derive an ensemble of light curves from the first principles, the physical process of TNO occultation can be simulated with the following four levels of consideration.

(1) *Star level*: The spectral type, luminosity class, apparent magnitude (or angular size) and, if available, a limb darkening model of a target star determine the physical properties of the diffracted light source.
(2) *TNO level*: The size, distance and shape of a TNO together with the astronomical or astrophysical assumption in the total population, size distribution and orbital distribution of TNOs specify the obstacle in a diffraction calculation. A diffractive shadow can be derived with the above two levels of parameters.
(3) *Shadow level*: Impact parameter and shadow velocity should be involved here. An occultation event does not have to be a center-crossing event. Nevertheless, the impact parameter can only be a random number that leads to a certain distribution. The shadow ground velocity will be coupled to the TNO distance by assuming a circular orbit. This might be a good approximation for the classical TNOs. Scattered objects have a larger dispersion in their shadow velocities.
(4) *Light curve level*: Filter is introduced in this level. It is proper to associate a specific term "configuration" with everything mentioned above. A configuration corresponds to the old geometric picture which relates to an ensemble of light curves with a certain probability. In addition,

sampling rate and sampling phase, namely, the (random) start point of each sampling, are the last two things to be considered before an ensemble of (normalized) light curves can be derived.

The ensemble of normalized light curves will be integrated with some simulated light curves at a given flux where various kinds of noise can be added for further analysis. Photometry is the major part of the data pipeline. A preliminary photometric algorithm can be used for testing.

A few concrete examples should be helpful in understanding some of the steps above. The importance of the star level can be understood through Table 1. The limiting magnitude is around 14 in a typical TAOS zipper mode image. It is obvious that the angular size of a TAOS target star could be compatible with a comet-size TNO itself. Another interesting feature is that Fresnel diffraction could dominate a TAOS event. Suppose a TNO is at a distance (d) of 50 AU. The related occultation is observed at a wavelength (λ) of 500 nm. The size of a TNO (s) will satisfy $s^2/\lambda d \sim 1$ if s is around 1 km. The code should be able to handle something like Fresnel integrals rather than just a Fourier transformation. A diffraction code was developed. It is now part of this simulator. Given all proper input parameters, a diffraction pattern and a CCD response at a certain sampling rate with various sampling phase can be derived as illustrated in Fig. 4. Only spherical TNOs of comet-size at trans-Neptunian distance are considered

Table 1. Colors, magnitude, temperatures, and radius of normal stars are shown. The projection of a stellar disk at 50 AU is estimated for various apparent magnitudes (m_V) around 10–14. These are typical magnitudes of target stars in a TAOS zipper mode image. Some of these stellar disks are similar to a comet-size TNO in their angular sizes.

				R (km) at 50 AU				
S_p	M_V	T_{eff}	$R/R_\odot$	$m_V = 10$	11	12	13	14
O5V	−5.7	42,000	12	0.15	0.09	0.06	0.04	0.02
A0V	+0.65	9,790	2.4	0.55	0.35	0.22	0.14	0.09
F5V	+3.5	6,650	1.3	1.10	0.69	0.44	0.28	0.17
K0V	+5.9	5,150	0.85	2.17	1.37	0.86	0.55	0.34
K5V	+7.35	4,410	0.72	3.58	2.26	1.43	0.90	0.57
M0V	+8.8	3,840	0.60	5.83	3.68	2.32	1.46	0.92
M5V	+12.3	3,170	0.27	13.14	8.29	5.23	3.30	2.08
K0III	+0.7	4,660	15	3.49	2.20	1.39	0.88	0.55
K5III	−0.2	4,050	25	3.85	2.43	1.53	0.97	0.61
M0III	−0.4	3,690	40	5.61	3.54	2.23	1.41	0.89

Source: Allen's Astrophysical Quantities, 4th Edition, ed. Cox, A. N. (Springer-Verlag, New York, 2000).

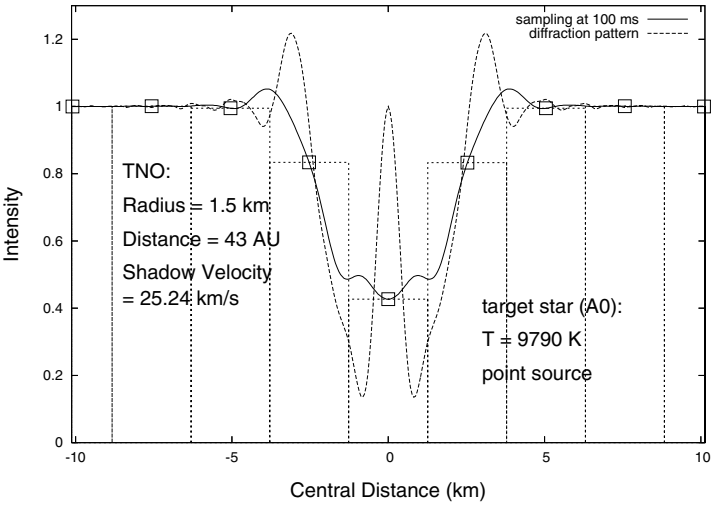


Fig. 4. Diffraction pattern and detector's response. Target star: A0 (black body at 9,790 K), point source. TNO: radius 1.5 km, 43 AU away, shadow velocity 25.24 km/s. Filter: 400–800 nm band-pass. The diffraction pattern through the center is shown as the dash curve. Response at a sampling rate of 100 ms is shown as nine small squares. A set of light curves will fill the solid line if all possible sampling phases are considered.

here. Shadow velocity, however, is a free input parameter here. It can be related to the angular separation from the opposition with a general (circular or non-circular) Keplerian orbit assumed. That velocity is a function of TNO orbital elements and the two-dimensional TNO phase angle. Its dispersion can thus be related to a distribution in the orbital elements. There are some interesting relations between a random impact parameter and the distribution of geometric occultation duration. Given a spherical object with its circular shadow, this distribution can be calculated analytically. A simulation is shown in Fig. 5 (the rightmost solid line with eccentricity $e = 0$) where a target star of point source is assumed. A few more two-dimensional cases were studied. An ellipse with certain orientation is similar to a circle basically. An ellipse of certain size with random orientation has an eccentricity-dependent distribution. Its geometric occultation duration is dominated by the size of its minor axis. Figure 5 shows the cumulated probability with respect to the normalized occultation duration. (For an ellipse, it is normalized with respect to its major axis where it has the longest occultation duration.) Geometric rectangular shadows were analyzed as well. Two folding points relate to its length and its width respectively as can be

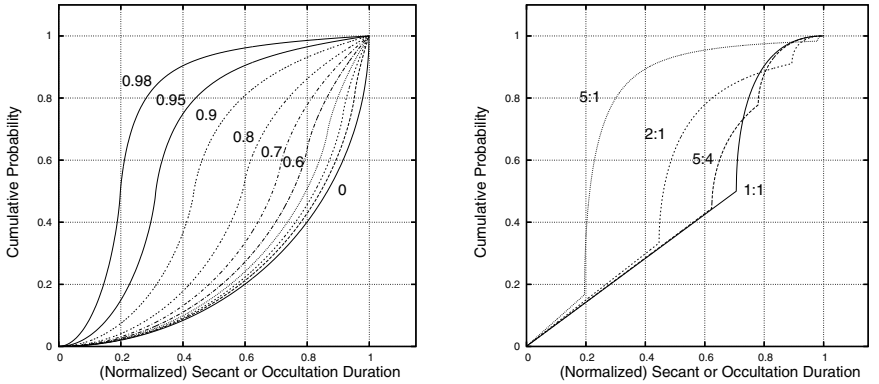


Fig. 5. Cumulative probability of the normalized geometric occultation duration for a given two-dimensional shape at a certain distance. The orientation is randomly distributed with a specific direction of shadow velocity. Target star is assumed as a point source. These are simulated results with one million events involved in each curve. The binning factor in these (cumulative) histograms is 1000. The maximum differential probabilities happen at the maximum slopes of these curves. (Left) Cumulative probability for ellipses of different eccentricities (Starting from the leftmost curve, the eccentricities are 0.98, 0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, and 0, respectively). The occultation duration is normalized with respect to its major axis. The maximum slope in each curve is associated with its minor axis. The slopes at the maximum occultation duration are related to the complete elliptic integral of the second kind. (Right) Cumulative probability for rectangles with different length to width ratios. The occultation duration is normalized with respect to its diagonal. Two maximum slopes in each curve are associated with its length and its width. The slopes of the straight lines range between $1/\sqrt{2}$ (for square) and 1 (if length $\gg$ width).

seen in Fig. 5. These two points degenerate for a square shadow. However, it is not clear whether these results can be applied while diffraction is considered. By assuming theoretical distributions with enough computing power, ensembles of light curves can be generated systematically. To get larger or non-spherical objects involved will require extending the existing code.

4. Perspectives

Photometry of zipper mode data is something new that the TAOS project has developed. Synchronization of the data coming from all telescopes is one of the major features in this experiment. We are planning to apply the "rank statistics"[17,18] method to eliminate a statistical false positive. This way, a simultaneous dip in the data from all telescopes can be given a probability of detection and compared against the chance occurrence of

a statistical fluctuation. We will need to compare the performance of this rank statistic algorithm against that of other photometric algorithms. With color information from other surveys available, we will be able to estimate the distribution of target stars in a given field as a function of angular size. Currently, 500–720 nm band-pass filters are used for all telescopes. It is possible that we could obtain color information by using different filters. By considering the real spectrum of a distant star, it is also likely that we can have some calibration method for our system. Small comets are believed to be non-spherical in general. How a result derived from a spherical model can be extrapolated for a more general case is not clear at this time. Space-based experiments were proposed[14,26] for COROT and Kepler missions. The velocity of a satellite (typically, 7–8 km/s in some direction along its orbit) will contribute to the shadow speed in the shadow level. This alone might change the occultation event rate, which is proportional to the shadow speed, by 20–30%.

Acknowledgments

We thank Dr. Isao Sato (Nakano Star Gazers Club) and Dr. Kinoshita Daisuke for providing detail information on all the asteroid occultation predictions. KHC's was performed under the auspices of the U.S. Department of Energy, National Nuclear Security Administration by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

References

1. K. E. Edgeworth, *Mon. Not. R. Astron. Soc.* **109** (1949) 600.
2. G. P. Kuiper, in *Proceedings of a Topical Symposium, Commemorating the 50th Anniversary of the Yerkes Observatory and Half a Century of Progress in Astrophysics*, ed. J. A. Hynek (McGraw-Hill, New York, 1951), p. 357.
3. D. Jewitt and J. Luu, *Nature* **362** (1993) 730.
4. M. E. Brown, C. A. Trujillo and D. Rabinowitz, *IAUC* **8577** (2005).
5. M. E. Brown, C. A. Trujillo and D. Rabinowitz, *Astrophys. J.* **617** (2004) 645.
6. G. M. Bernstein, D. E. Trilling, R. L. Allen, M. E. Brown, M. Holman and R. Malhotra, *Astron. J.* **128** (2004) 1364.
7. S. J. Kenyon and B. C. Bromley, *Astron. J.* **128** (2004) 1916.
8. M. E. Bailey, *Nature* **259** (1976) 290.
9. M. J. Duncan and H. F. Levison, *Science* **276** (1997) 1670.
10. K. H. Cook, C. Alcock, T. Axelrod and J. Lissauer, *Bull. Am. Astron. Soc.* **27** (1995) 1124.

11. S.-K. King, in *ASP Conf. Ser. 246: IAU Colloq. 183: Small Telescope Astronomy on Global Scales*, eds. W.-P. Chen, C. Lemme and B. Paczyński (The Astronomical Society of the Pacific, San Francisco, 2001), p. 253.
12. C. Alcock, R. Dave, J. Giammarco *et al.*, *Earth Moon and Planets* **92** (2003) 459.
13. F. Roques, M. Moncuquet and B. Sicardy, *Astron. J.* **93** (1987) 1549.
14. F. Roques and M. Moncuquet, *Icarus* **147** (2000) 530.
15. F. Roques, M. Moncuquet, N. Lavillonière, M. Auvergne, M. Chevreton, F. Colas and J. Lecacheux, *Astrophys. J.* **594** (2003) L63.
16. A. Cooray, *Astrophys. J.* **589** (2003) L97.
17. C. Liang, *The Detection of Stellar Occultation by Kuiper Belt Objects* (PhD thesis, University of California, Berkeley, 2001).
18. C. Liang, J. A. Rice, I. de Pater, C. Alcock, T. Axelrod, A. Wang and S. Marshall, *Statistical Science* **19** (2004) 265.
19. S. Marshall, C. Akerlof, R. Kehoe *et al.*, *Bull. Am. Astron. Soc.* **29** (1997) 1290.
20. R. C. Stone, S. W. McDonald and J. L. Elliot, *Astron. J.* **118** (1999) 591.
21. D. V. Denissenko, *Astronomy Letters* **30** (2004) 630.
22. W. P. Chen, Z. W. Zhang, S. K. King *et al.*, *Baltic Astronomy* **12** (2003) 568.
23. O. Fors, J. Núñez and A. Richichi, *Astron. & Astrophys.* **378** (2001) 1100.
24. O. Fors, E. P. Horch and J. Núñez, *Astron. & Astrophys.* **420** (2004) 397.
25. M. J. I. Brown and R. L. Webster, *Mon. Not. R. Astron. Soc.* **289** (1997) 783.
26. B. S. Gaudi, *Astrophys. J.* **610** (2004) 1199.

NEO-SURVEY AND HAZARD EVALUATION

YUEHUA MA* and GUANGYU LI†
Purple Mountain Observatory, CAS, Nanjing 210008, China
National Astronomical Observatories, CAS, Nanjing 210008, China
**yhma@pmo.ac.cn*
†*gyl@pmo.ac.cn*

Chinese scientists have done many research works in this field of asteroid survey and related aspects. They found Asteroid 1125 — China at Yerkes Observatory in 1928 and found Asteroid Purple 1 at Purple Mountain Observatory (PMO) in 1955. Up to the middle period of 1980s, over 130 new numbered asteroids were found at PMO and among them two are Mars-crossing asteroids. The Schmidt CCD Asteroid Program (SCAP) of Beijing Astronomical Observatory (BAO) was put in practice in 1995. Two NEOs and a Mars-crossing asteroid were found in 1997. In 1990s, 575 asteroids were found by SCAP, which was the first place in the world at that time. The construction of NEO Search Telescope was started in 1999 and has been finished in principle. The telescope is a 1.0/1.2 m Schmidt telescope with $4\,\mathrm{K} \times 4\,\mathrm{K}$ shift scanning CCD detector. It is to be expatiated in this paper. The telescope can also inspect space debris. In addition to ground base NEO Survey and investigation, we have tabled a proposal "auto-navigation for NEO exploration" and hope to have international cooperation.

1. The Historical Asteroid Survey in China

The enormous power of celestial bodies impacts is matchless. The danger of NEOs impacts with earth threatens our living space all the time. The annihilation of dinosaur 65 Myrs ago, the great blast of Tunguska in 1908, the impacts of comet Shoemaker-Levy 9 on Jupiter in 1994 and NEOs passing Near Earth orbit with short distances these years warn us that we cannot treat them lightly.

Chinese scientists have done many research works in the field of asteroid survey and related aspects. In 1928, Asteroid 1125-China was found at Yerkes Observatory by C. Y. Chang, who was the former director of Purple Mountain Observatory (PMO). In 1955, Asteroid Purple 1 was found at PMO by professors C. Y. Chang and J. X. Zhang, which was the first asteroid found by Chinese at home. Up to mid-period of 1980s, the asteroid research group lead by Professor J. X. Zhang found over 130 new numbered asteroids at PMO with 40 cm two-tube refracting telescope and two of them

are Mars-crossing asteroids. That was the fifth place in the world at that time. Among them Asteroids 2077 (Jiangsu) and 2078 (Nanjing) are Mars-crossing asteroids. In 1990, together with a French astronomer Alain Maury, Professor Q. Wang found a Near Earth Asteroid 1990 AD in France.

The Schmidt CCD Asteroid Program (SCAP) of Beijing Astronomical Observatory (BAO) was put in practice in 1995. The telescope used is a 60/90 cm Schmidt telescope with 2 K × 2 K CCD detector. In 1997, near-Earth Asteroids 1997 BR and 1997 TT25 and a Mars-crossing asteroid 1997 LY4 were found by this scheme. In 1990s, 575 asteroids were found by SCAP, and the observational quantity and the number of asteroids found headed the list in the world in successive years in that period.

To commend the Chinese astronomer contributing to asteroid research, Harvard College Observatory named Asteroids 2051 and 4760 as Chang (C. Y. Chang) and Jia-Xiang (J. X. Zhang).

In 1994, Professor J. X. Zhang with his colleagues did the prediction work of comet Shoemaker–Levy 9 (SL9) impacting on Jupiter. In the situation of lacking the last key observational data, the precision of impact time we predicted reached high level in the world compared with the actual impact time observed by Galileo spacecraft.

After the event of comet SL9 impacting on Jupiter, Chinese promised to build a NEO Search Telescope to take part in the international NEO joint survey at the meeting for discussion how to safeguard our earth in 1995.

2. 1.0/1.2 m Schmidt Telescope for Asteroid Survey

Our NEO search telescope construction was started in 1999 and has been finished. It is a 1.0/1.2 m Schmidt telescope with 4 K × 4 K shift scanning CCD detector. The telescope can also inspect space debris.

Figure 1 is the main building for the telescope and Fig. 2 is the telescope mounted on the dome.

2.1. *Scientific objectives*

The scientific objectives of the telescope are as follows:

(i) survey NEOs,
(ii) determine the orbits of asteroids and comets, so that
(iii) to predict the possible collision events,
(iv) research the dynamical evolution of the orbits of asteroids and comets.

Fig. 1. The main building for the NEO telescope.

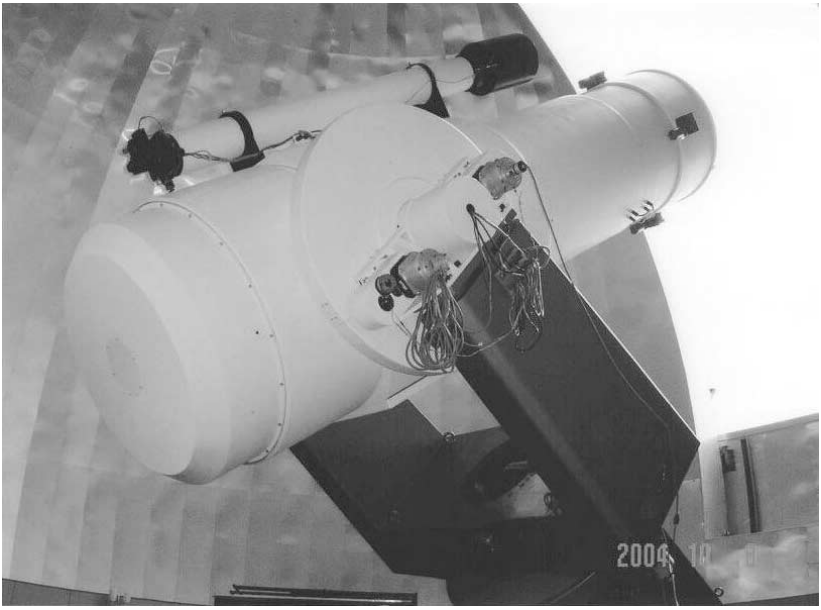


Fig. 2. The NEO telescope.

2.2. *Main parameters of the Schmidt telescope*

(i) Diameter of the correcting lens: 1.0 m.
(ii) Diameter of the primary mirror: 1.2 m.
(iii) Focal ratio: $F = f/D = 1.8 \quad (f = 1.8\,\text{m})$.
(iv) Effective field: 3.14° (linear diameter: 100 mm).
(v) Limiting magnitude of the telescope: 20.7.
(vi) Center wavelength of the correcting lens: 656.3 nm.
(vii) Light power distribution: 80% of light in less than 2″ (linear diameter is less than 20 μm).
(viii) Distortion caused by optical designing and machining: less than 15 μm.
(ix) Bearing of the tube: bending to focal plane less than 0.02 mm.
(x) Tracking precision: 1″/4 min.
(xi) Pointing precision: less than 10″.

2.3. *Site of the observational station*

The NEO search telescope is mounted at Xuyu station. The site is from the north to Nanjing in 120 km. Its position is at N 32°44′.2 ± 0.5′, E 118°27′.9± 0.5′ and sea level is 180.9 m. The survey covering sky is the northern area from latitude S 20°. By historical record, the observing nights from 1995 to 1999 were averagely 214.6 per year and 207.5 per year from 1990 to 1999. We did observation from August 1999 to July 2000 for the whole year, got the astronomical observing nights being 194. In the year, the number of astronomical observing hours was, 1652.

Seeing Observation Results: from June to December in 1999, we got total 1493 seeing, the average result is $\theta = 0.81''$.

Figure 3 shows that seeing changes symmetry to midnight.

3. "Auto-Navigation for NEO Exploration" Proposal

In addition to ground-base NEO survey and investigation, we have tabled a proposal "auto-navigation for NEO exploration" and hope to have international cooperation.

The aims of asteroids exploration are to study asteroids and their origin, by which we can learn the origin and evolution of the solar system and look for near-Earth Asteroids, investigate their spatial distributions and trace the Potentially Hazardous Objects (PHO).

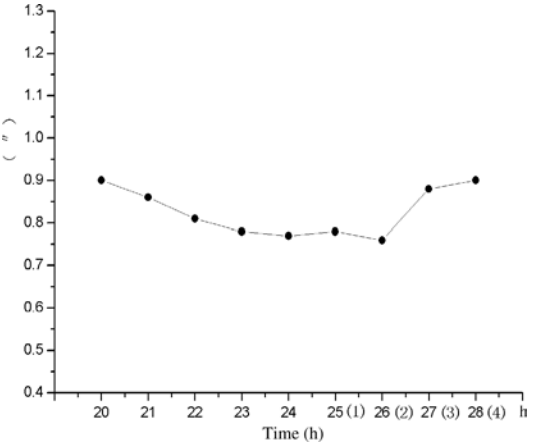


Fig. 3. Seeing changes symmetrically to midnight.

Orbital parameters and spectrum of an asteroid can be obtained by large ground-based telescopes. But it is only applicable to big and bright objects. Asteroid observation and investigation in space has been realized by several flybys and single body missions, e.g., NEAR,[1] Hayabusa (Muses-C).[2] Recent years, based on advanced instruments and auto-navigation system, space missions for asteroid flux, category, and size distribution are proposed. Bering[3,4] proposed by Danmark aims at main belt asteroids. The orbit is 0.72–3.5 AU. The EUNEOS[5] proposed by France is to explore NEOs and PHOs. The orbit is 0.5–0.72 AU.

In our ASTROD I mission[6] proposal for studying precision astrodynamics, the mission orbit of the spacecraft is designed to reach between the orbits of Venus and Mercury after two Venus encounters. From the experience in the ASTROD I orbit design, there could be various choices for the orbit of the asteroid exploration spacecraft to combine the two goals of Bering and EUNEOS in one space exploration. Figure 4 is the designed orbit of ASTROD I.

Our asteroid exploration scheme can be: after two encounters with Venus to reach an orbit between that of Venus and Mercury to observe inner asteroids for some time, the spacecraft encounters with the Venus for the third and fourth times and Earth once to reach the asteroid belt. Before reaching asteroid belt, the spacecraft can discover and explore the asteroids in the area from Mercury orbit to asteroid belt, especially NEOs and asteroids across Mars. After reaching the main asteroid belt, its main object is to

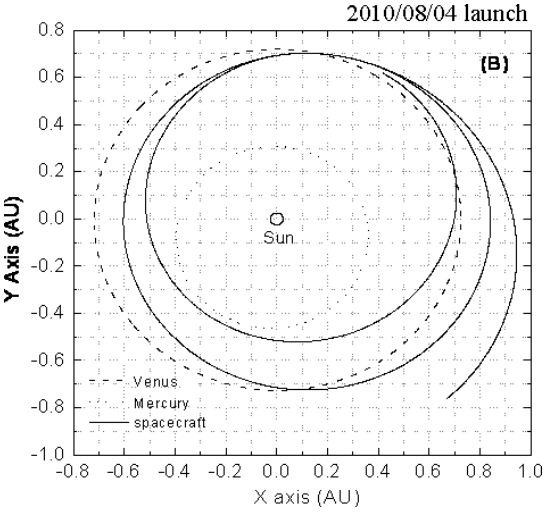


Fig. 4. Designed orbit of ASTROD I.

explore main-belt asteroid flux, spatial distributions, composition, and the relationship to meteorites.

Acknowledgments

This work was supported by National Natural Science Foundation of China (No. 10273024) and the Minor Planetary Foundation of Purple Mountain Observatory. We thank Prof. Ni Wei-tou and Yao Dazhi for valuable discussions on the early drafts.

References

1. D. K. Yeomans, P. G. Antreasian, J.-P. Barriot *et al.*, *Science* **289** (2000) 2085.
2. J. Kawaguchi, H. Kuninaka, A. Fujiwara and T. Uesugi, *Proceedings of the 5th IAA International Conference on Low-Cost Planetary Missions, ESTEC* (Noordwijk, the Netherlands, 24–26 September 2003, ESA SP-542, November 2003), p. 25.
3. A. C. Andersen, R. Michelsen, H. Haack, J. L. Jorgensen, M. Betto and P. S. Jorgensen, *Proceedings of the 5th IAA International Conference on Low-Cost Planetary Missions, ESTEC* (Noordwijk, the Netherlands, 24–26 September 2003, ESA SP-542, November 2003), p. 307.

4. J. L. Jorgensen, P. S. Jorgensen, M. Betto, T. Denver and L. J. Tunon, *Proceedings of the 5th IAA International Conference on Low-Cost Planetary Missions, ESTEC* (Noordwijk, the Netherlands, 24–26 September 2003, ESA SP-542, November 2003), p. 321.
5. V. Martinot and A. Morbidelli, *Proceedings of the 5th IAA International Conference on Low-Cost Planetary Mission, ESTEC* (Noordwijk, the Netherlands, 24–26 September 2003, ESA SP-542, November 2003), p. 47.
6. W.-T. Ni, S. Shiomi and A.-C. Liao, *Class. Quantum. Grav.* **21** (2004) S641.

This page intentionally left blank

ENA SIGNALS COMING FROM THE HELIOSPHERIC BOUNDARIES

K. C. HSIEH
Department of Physics, University of Arizona
Tucson, AZ 85721, USA
hsieh@physics.arizona.edu

Being the only article in this volume concerning energetic neutral atoms (ENA) coming from the heliospheric boundaries, this will be a mini-review, beginning with a brief introduction to the subject, and then highlighting the current status on observation and modeling in this field. We conclude with a view for future investigations.

1. Introduction

Our Sun moves inside a local interstellar cloud (LIC).[1] Pressure balance between the radially expanding supersonic solar wind (SW) and the impinging LIC plasma creates boundary layers that separate the plasmas, each carrying its own imbedded magnetic fields.[2] The region of supersonic SW is like a bubble bounded by the termination shock (TS), at which SW becomes subsonic and the magnetic field more intense. The subsonic solar plasma fills the heliosheath (HS), the region bounded by TS on the inside and the heliopause (HP) on the outside, separating the solar plasma from the LIC plasma. Sun's motion relative to LIC compresses HS in the apex or upwind direction, i.e., the direction of Sun's relative motion, and stretches HS into a tail in the anti-apex or downwind direction. The subsonic solar plasma in HS is convected from the apex to the tail region, while the LIC plasma flows outside HP in the same direction. Away from the apex of HP, the LIC plasma may form a bow wave or bow shock, depending on the ratio of Sun's speed relative to the local interstellar medium (LISM) to the speed of sound therein. We note that terms "LIC" and "LISM" are used interchangeably in the literature. The regions and boundaries of the heliosphere described above are shown schematically in Fig. 1.

Unlike the supersonic SW in the bubble and the subsonic solar plasma in the HS, which are predominantly charged particles, LIC plasma has a significant neutral component.[3] Unaffected by magnetic field, neutral atoms of

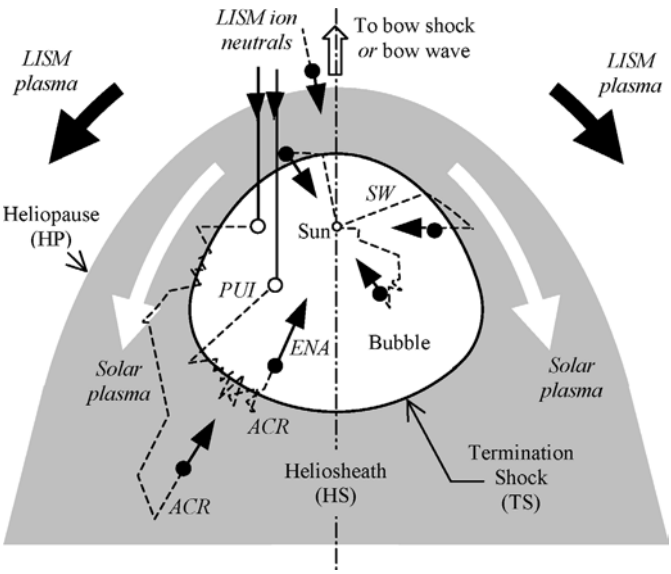


Fig. 1. Heliosphere and heliospheric ENA. Regions and boundaries are in plain text, and particles in italics. Dashed lines are ion trajectories, while solid lines those of neutral atoms. Solid circles signify neutralization and open circles ionization. LIC ions charge exchange with LIC atoms outside the HP can enter the heliosphere. Penetrating LIC neutrals become pickup ions (PUI) after ionization; and PUI accelerated at the termination shock become anomalous cosmic rays (ACR).

LIC, mainly H, He, C, N, and O atoms, penetrate into the heliosphere with lose, as they approach the Sun, due to ionization by solar photons, ions and electrons. Depending on their ionization potentials, each atomic species has its specific spatial distribution in the SW bubble. Most pronounced is the gravitationally focused LIC He or the "He cone", stretching from 0.3 AU outward in the anti-apex direction.[4] Direct measurement of the distribution of interstellar He shows Sun's relative velocity of 26.3 ± 0.4 km/s in the direction of ecliptic longitude $74.7 \pm 0.5°$ and latitude $-5.2 \pm 0.2°$.[5] Interstellar atoms are the partners in charge exchange with energetic ions of different origins and in different regions to produce ENA, as well as the main source of PUI. Figure 1 shows ENA created in and near HS, and can reach 1 AU for our detection, thus serving as signals from the heliospheric boundaries.

Crudely speaking, ENA from the boundaries group into three: (1) those from interstellar ions just outside the heliopause, especially those in the hydrogen wall[6] in the apex region; (2) those from subsonic solar ions near

TS; (3) those from ACR in HS. Each group of ENA is a direct sample of the ion population of its place of origin, while ions of that region cannot reach us and be detected. The ballistic trajectories of ENA across magnetic field afford imaging of remote ion populations. Group 3 has the most circuitous journey in the heliosphere and carries most information concerning particle acceleration at TS, since ACR are believed to be PUI accelerated at TS, and PUI are penetrating LIC neutrals ionized in and picked up by the supersonic SW. Also represented in Fig. 1 are ENA from ions accelerated by interplanetary shocks.[7] This group of ENA may mask the detection of ENA from the boundaries.[8]

The detection of ENA requires dedicated instruments and opportunities, in order to take the unique advantage of imaging space plasmas in regions not easily accessible to space missions; e.g., heliospheric boundaries. The first detection of ENA, in 1951, was in the aurora by ground-based optical means in the blue shifted H-Balmer lines.[9] First detection of ENA (1–20 keV H) in space was performed on a rocket flight, by an ion mass spectrometer using a thin carbon foil to ionize penetrating incident ENA, so they could be analyzed by an electrostatic ion mass analyser.[10] Earnest testing and development of space-borne ENA detectors revived in the 1970s. Thin carbon foils are used as filter against intense EUV radiation as well as generator of secondary electrons for "start" signals needed for the newly evolving time-of-flight (TOF) mass spectrometers.[11–15] At lower energies, say, ENA from LIC neutrals, the technique of surface conversion is employed and yielded the first direct detection of LIC He atom on the Ulysses mission.[16] A thorough review of techniques in ENA imaging of space plasma, including the use of transmission grating as EUV filter is given by Gruntman.[17] For a more recent exposition, see Ref. 18.

2. Current Status

Indeed, the study of the heliospheric boundaries via ENA began with modeling,[19,20] and modeling continues to lead the way. We shall mention observations and their role in modeling. Page limitation prohibits more coverage of many worthy works.

2.1. *Observation*

Of the three groups of ENA from heliospheric boundaries, only ENA of ACR origin can claim to have been observed. Even that is only the beginning and the results from HSTOF of CELIAS on SoHO (55–88 keV H and

85–141 keV He) remain singular.[21–24] High background dashed the hope of harvesting from INA of MIMI[25] on *Cassini* and from HENA[26] on IMAGE (D. Mitchell, private communication, 2005). Whether ASPERA-3[27] on Mars Express, which has two ENA instruments, NPI (∼0.1–60 keV) and NPD (∼0.1–10 keV), could yield any results in heliospheric ENA, including imaging the He cone, remains to be seen. Meanwhile, we pin our hopes on Interstellar Boundary Explorer (IBEX), which is to be launched in 2008 to image the heliosphere in 0.01–6 keV H.[28] The SoHO results and IBEX expectations are best seen in light of existing models.

2.2. *Modeling*

There is a host of good modeling work on ENA in the outer heliosphere; we will, however, only highlight two that make use of the only available ENA observations, i.e., that of SoHO's, and one that predicts what IBEX may expect. Most models use H and ENA interchangeably, because the predominance of H and the lack of mass resolution in some ENA instruments. Up to date, HSTOF of CELIAS on SoHO is the only ENA instrument that can identify H and He separately.[23,24]

2.2.1. *Learning from SoHO observation*

By assuming ACR's mode of propagation in a HS of certain size and shape and ACR's spectral form at TS, with its parameters constrained by the demodulated ACR spectrum based on observations of Voyagers 1 and 2, a self-consistent model fits SoHO's measured H and He fluxes independently and yields a common power index of the differential spectra, and a shock strength of 2.89.[29] Figure 2 is a sample of their results.

In a comprehensive model studying the "injection problem" in acceleration of superthermal PUI to become ACR at TS and in HS, the energetic H fluxes from the apex and the heliotail direction measured by SoHO are compared with H spectra resulting from stochastic acceleration by different magnitudes of turbulence in each region.[30]

The five curves in Fig. 3 arise from different degrees of stochastic acceleration in HS.[30] Curve 1 stands for the flux coming from outside TS in the absence of stochastic acceleration. Curve 2 stands for the flux coming from both sides of TS, also without stochastic acceleration. Curves 3–5 stand for fluxes from PUI stochastically accelerated in different degrees, characterized by a turbulence parameter of 0.012, 0.12 and 1.2, respectively.

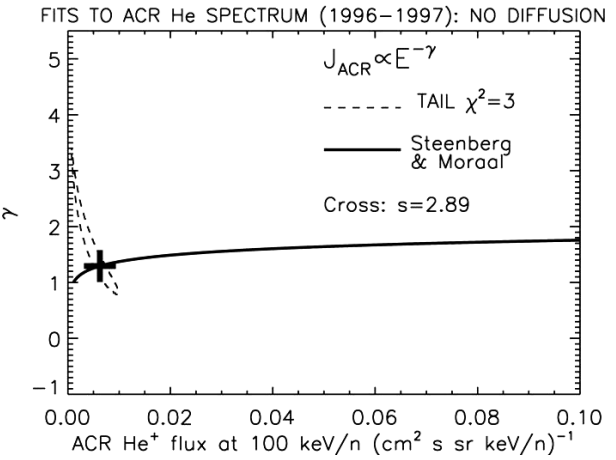


Fig. 2. Sample of an attempt in finding the yet to be observed ACR spectrum at the termination shock and the compression ratio of that shock by adjusting the parameters of a model to fit the H and He neutral atoms flux observed by SoHO at 1 AU. Fitting H and He data independently yields the same result. The shock strength found is 2.89 (figure by courtesy of A. Czechowski).

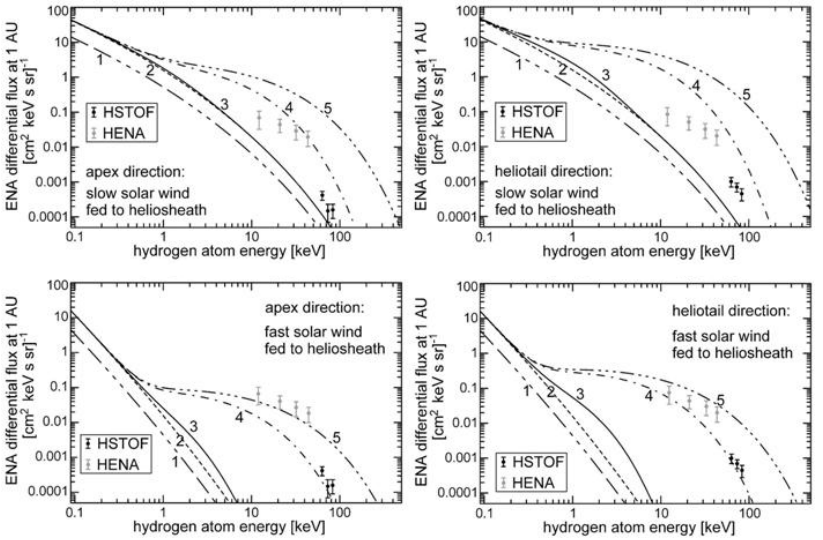


Fig. 3. Neutral H fluxes from the apex and from the anti-apex direction as measured by HSTOF of CELIAS on SoHO are compared with a model investigating the role stochastic acceleration of the superthermal PUI plays in HS. The numbers beside the curves refer to different magnitudes of turbulence in the region. See text for definition. The SoHO measurements in two opposite regions of the heliosheath are in agreement with the same set of modeling parameters. The data points from HENA on IMAGER are upper limits of ENA flux averaged over all directions[32] (figure by courtesy of R. Kallenbach).

As curves 3–5 show, to determine the turbulence parameter with better confidence, measurements at energies >10 keV are much desired. Extending the modeling energy to beyond 100 keV is especially useful, when linking ACR acceleration at heliospheric boundaries via the detection of ENA with solar modulation of cosmic rays via the detection of ACR, both at 1 AU.[31]

Both models[29,30] cited above show the usefulness of ENA observation at 1 AU, in both cases, by HSTOF on SoHO at L1, to modeling the ions in far-away heliospheric boundaries, even when the instrument has a limited energy window and can only detect ENA under quiet conditions set by Sun's activities. These models and the SoHO measurements focus on the acceleration of superthermal PUI to become ACR at the heliospheric boundaries. The upcoming IBEX has a different goal and requires different models.

2.2.2. *In anticipation of the IBEX mission*

Expected H fluxes from the heliospheric boundaries anticipated in IBEX's energy range (0.01–6 keV) depend on the nature and structure of TS and, also, on the evolution of PUI in the supersonic solar wind, at TS and in HS. Gruntman *et al.*[33] have considered several scenarios and calculated H fluxes in different directions as functions of energy. More exact multi-component models have been published[3,34,35] based on the heliospheric interface modeled by Baranov and Malama.[6] The latest model[34,35] treats different groups of PUI separately from the bulk solar wind, and takes into account of all the participating particles ranging from galactic and anomalous cosmic rays to different species of ions in the magnetized interstellar and solar plasmas, in a multi-component and kinetic-continuum computational approach to set the location, shape, and particle distributions of the boundary layers, thus providing the most comprehensive model of the heliosphere. Protons from different origins are injected to produce the different H fluxes. The dominant components are presented in Fig. 4.[34,35] Curve 0 shows the flux coming from solar-wind protons in the inner HS. Curves 1 and 2 correspond to fluxes originated from two types of pickup protons in HS (see details in Ref. 34). Curve 3 represents secondary interstellar H flux. For comparison, curve BM is the result of the one-component model.[6] Two-dimensional images of the entire heliosphere in neutral H flux based on a similar but simpler model have been produced.[33]

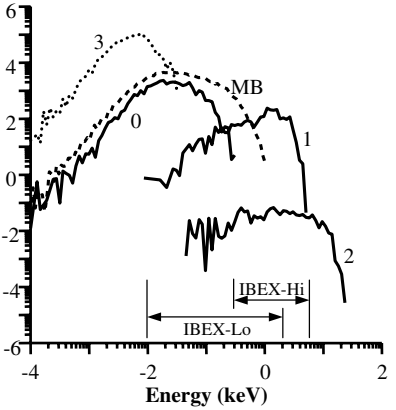


Fig. 4. Predicted differential spectra of H at 1 AU from the upwind direction. This log-log plot's vertical axis is H flux in $(\mathrm{cm}^2\,\mathrm{s}\,\mathrm{keV})^{-1}$. Number indices denote different populations of protons that gave birth to the H flux. See text for definition. This figure is adapted from the one kindly provided by V. Izmodenov and modified by the author to indicate the energy windows of IBEX's ENA instruments, IBEX-Hi and IBEX-Lo.

3. Conclusion

In this mini-review, we tried to provide a background in the study of the heliospheric boundaries by measuring ENA flux at 1 AU, and the current status in modeling efforts in utilizing existing results from SoHO and in anticipation of IBEX set for launched in 2008. In the process, we give a basic list of references for interested readers. Due to page limitation, many topics were not discussed; e.g., time dependence of heliospheric conditions and the information that may be carried in ENA.[36]

Now that Voyager 1 has crossed the termination shock, new data should help modelers to refine their evermore-comprehensive models of the heliosphere and the ENA flux coming from the outer regions. More immediately, ion flux in the 50–80 keV interval measured by Voyager 1 in the heliosheath should be used to test the consistency of the ACR interpretation of the ENA flux from the upstream direction measured by SoHO, with the reservation that the thickness of the heliosheath upstream will not be determined for some time to come.

While SoHO's emphasis was on ACR in the higher energies, IBEX will be concentrating on the lower energy pre-ACR PUI. We wish IBEX great success. Meanwhile, effort to seek opportunity to image the heliospheric boundaries in energetic H and He in the interval 20–200 keV should be

pursued, because it is in this energy window, through which the acceleration of UPI to become ACR is best studied, especially, in direct comparison with ACR measurements at 1 AU, to yield a self-consistent study of charged-particle acceleration and cosmic-ray modulation in the heliosphere.[31] The technology is available, only the opportunity is lacking.[37]

Acknowledgments

Hearty thanks go to A. Czechowski, M. Gruntman, V. Izmodenov, R. Kallenbach, and E. Roelof for helpful discussions and for illuminating figures used for the invited talk, on which this paper is based.

References

1. J. D. Salvin and P. C. Frisch, *Astrophys. J.* **565** (2002) 364.
2. H. Fichtner and K. Scherer, in *The Outer Heliosphere: Beyond the Planets*, eds. K. Scherer, H. Fichtner and E. Marsch (Copenrnicus Gesellschaft e. V., Katlenburg-Lindau, 2000), p. 1.
3. V. Izmodenov *et al.*, *Space Sci. Rev.* **97**, 1/4 (2001) 413.
4. G. Gloeckler *et al.*, *Astron. & Astrophys.* **426** (2004) 845.
5. M. Witte, *Astron. & Astrophys.* **426** (2004) 835.
6. V. B. Baranov and Y. G. Malama, *J. Geophys. Res.* **98** (1993) 15,157–15,163.
7. J. Kóta *et al.*, *J. Geophys. Res.* **106**, 24 (2001) 907.
8. E. C. Roelof, in *COSPAR Colloquia Ser. Vol. 3*, eds. E. Marsch and R. Schwenn (Pergamon, Oxford, 1992), p. 385.
9. A. B. Meinel, *Astrophys. J.* **113** (1951) 50.
10. W. Bernstein *et al.*, *J. Geophys. Res.* **74** (1969) 3601.
11. K. C. Hsieh, C. Y. Fan and J. J. L'Heureux, *Proceedings of the 15th International Cosmic Ray Conference, Conference Papers*, Vol. 9 (Plovdiv, 1977), p. 163.
12. K. C. Hsieh, E. Keppler and G. Schmidtke, *Applied Optics* **18** (1979) 3732.
13. G. Gloeckler and K. C. Hsieh, *Nucl. Instr. Meth.* **165** (1979) 537.
14. M. Gruntman and V. A. Morozov, IKI preprint-667 (1981). Available at http://astronauticsnow.com/mg_pubs/guntman_preprint-667-iki-1981.pdf.
15. G. Gloeckler, *Rev. Sci. Instrum.* **61**, 11 (1990) 3613.
16. M. Witte *et al.*, *Astron. Astrophys. Suppl. Ser.* **92** (1992) 333.
17. M. Gruntman, *Rev. Sci. Instrum.* **68**, 11 (1997) 3617.
18. P. Wurz, in *The Outer Heliosphere: Beyond the Planets*, eds. K. Scherer, H. Fichtner and E. Marsch (Copenrnicus Gesellschaft e. V., Katlenburg-Lindau, 2000), p. 251.
19. K. C. Hsieh *et al.*, *Astrophys. J.* **393** (1992) 756.
20. K. C. Hsieh and M. Gruntman, *Adv. Space Res.* **13**, 6 (1993) 131.
21. D. Hovestadt *et al.*, *Solar Phys.* **162** (1995) 441.

22. M. Hilchenbach *et al.*, *Astrophys. J.* **503** (1998) 916.
23. M. Hilchenbach *et al.*, in *COSPAR Colloquia Ser. Vol. 11*, eds. K. Scherer, H. Fichtner, H.-J. Fahr and E. Marsch (Pergamon, Amsterdam, 2001), p. 273.
24. A. W. Shaw, *et al.*, in *COSPAR Colloquia Ser. Vol. 11*, eds. K. Scherer, H. Fichtner, H.-J. Fahr and E. Marsch (Pergamon, Amsterdam, 2001), p. 219.
25. S. M. Krimigis *et al.*, *Space Sci. Rev.* **114** (2004) 233.
26. D. G. Mitchell *et al.*, *Space Sci. Rev.* **91** (2000) 67.
27. S. Barabash *et al.*, in *Mars Express: The Scientific Payload*, ed. A. Wilson (European Space Agency, Noordwijk, 2004), p. 121.
28. D. McComas *et al.*, in *Physics of the Outer Heliosphere: Third Annual IGPP Conference,* AIP Conference Proceedings, CP 719 (American Institute of Physics, New York, 2004), p. 162.
29. A. Czechowski, M. Hilchenbach and K. C. Hsieh, *Astron. & Astrophys.* **431** (2005) 1061.
30. R. Kallenbach *et al.*, *Astron. & Astrophys.* **439**, 1 (2005) 1.
31. K. C. Hsieh, A. Czechowski and M. Hilchenbach, in *Acceleration and Transport of Energetic Particles Observed in the Heliosphere: ACE 2000 Symposium*, eds. R. A. Mewaldt *et al.*, *AIP Conference Proceedings, CP 528* (American Institute of Physics, New York, 2000), p. 325.
32. E. C. Roelof, R. Demajistre, D. G. Mitchell and P. C. Brandt, private communication (2005).
33. M. Gruntman *et al.*, *J. Geophys. Res.* **106** (2001) 15,767.
34. Y. G. Malama, V. V. Izmodenov and S. V. Chalov, in *Proccedings of Solar Wind 11/SOHO 16*, eds. T. Zurbuchen and B. Fleck, ESA Conference Proceedings SP-592, in press (2005).
35. Y. G. Malama, V. V. Izmodenov and S. Chalov, *Astron. & Astrophys.* **445** (2006) 693–701.
36. A. Czechowski, M. Hilchenbach and K. Scherer, poster presented at EGU General Assembly in Vienna, 24–29, April, 2005.
37. K. C. Hsieh *et al.*, *Adv. Space Sci.* **34** (2004) 213.

This page intentionally left blank

HABITABLE ZONES FOR EARTH-LIKE PLANETS IN THE 47 UMa PLANETARY SYSTEM

JIANGHUI JI*,†,§ and LIN LIU‡,§

†*Purple Mountain Observatory, Chinese Academy of Sciences Nanjing 210008, China*

‡*Department of Astronomy, Nanjing University, Nanjing 210093, China*

§*National Astronomical Observatory Chinese Academy of Sciences, Beijing 100012, China*

**jijh@pmo.ac.cn*

The Habitable zones are usually believed to be appropriate environment for terrestrial planets that can provide the liquid-water, subtle temperature, atmosphere components of CO_2, H_2O, and N_2 [Kasting *et al.*, *Icarus* **101** (1993) 108], supporting the development and biological evolution of life on their surfaces. In this work [see an accompanied paper, Ji *et al.*, *Astrophysical Journal* **631** (2005) 1191 for details], we investigated the dynamical architecture of 47 UMa with the planetary configuration of the best-fit orbital solutions by Fischer *et al.* [*Astrophysical Journal* **586** (2003) 1394], to study the existence of the Earth-like planets in the region for $0.05\,\mathrm{AU} \leq a \leq 2.0\,\mathrm{AU}$ for 47 UMa by numerical simulations. In the study, we found that the "hot Earths" at $0.05\,\mathrm{AU} \leq a < 0.4\,\mathrm{AU}$ can dynamically survive at least for 1 Myr. The Earth-like planets can eventually remain in the system for 10 Myr at the areas involved in mean motion resonance (MMR) (e.g., 3:2 MMR and 9:5 MMR) with the inner companion. Moreover, we showed that the 2:1 and 3:1 resonances could be marginally stable, but the 5:2 MMR is unstable. In a dynamical sense, we point out that the most possible candidate habitable environment is that the Earth-like planets may bear the orbits of $0.8\,\mathrm{AU} \leq a < 1.0\,\mathrm{AU}$ and $1.0\,\mathrm{AU} < a < 1.30\,\mathrm{AU}$ (except 5:2 MMR) for relatively lower eccentricities. We also conducted similar studies in other multi-planet systems and found the potential existence of the Earth-like planets in habitable zones.

1. Introduction

The discovery of the first Jupiter-mass planet to the solar-type star 51 Peg[1] dates back 10 years ago. Such historical scientific finding not only uncovers the prelude to search for other planets or alien worlds outside our own solar system, but confirms that planets can be formed anywhere about their parent stars in the circumstellar disks (e.g., HD 141569 and Beta Pictoris).

*Corresponding author.

These flat disks enshrouding young stars are thought to be a common feature of stellar evolution and of planetary system formation. Primordial protoplanetary disks contain gas and dust and supply the raw ingredients from which the new planetary systems can form. With the help of Doppler radial velocity measurements and other means, a fruitful of diverse planetary systems come to our sight from the fact of observations, rather than the mind of theorists. To date, 145 planetary systems (October 18, 2005, see http://vo.obspm.fr/exoplanetes/encyclo/index.php), containing 169 extrasolar planets, have been detected about solar-type stars, among which there are 18 multiple-planet systems, 14 two-planet systems (HD 82943, GJ 876, HD 168443, 47 UMa, HD 38529, HD 12661, HD 128311...), three-planet systems (Upsilon And, HD 160691, HD 37124)[2–4] and one four-planet system 55 Cnc,[5] where the studies of the dynamics or the formation of these systems are essential to understand how two (or more) planets originate and evolve there.

In Table 1, some characteristics (the planetary masses, the orbital periods P, the semi-major axes a and the eccentricities e) of seven multi-planet systems are listed. According to the ratios of two orbital periods of the planets, the planetary systems are usually classified into resonant systems with $P_1/P_2 < 5{:}1$ (where $P_{1,2}$ are, respectively, the orbital periods for the

Table 1. Properties of seven multi-planet systems.[a]

Planet	$M_{\rm star}(\mathrm{M}_\odot)$	$M \sin i (M_{\rm Jup})$	Period P(days)	a(AU)	ecc.
GJ 876 b	0.32	1.89	61.02	0.207	0.10
GJ 876 c	0.32	0.56	30.12	0.130	0.27
HD 128311 b	0.84	2.18	458.6	1.099	0.25
HD 128311 c	0.84	3.21	928.3	1.76	0.17
47 UMa b	1.03	2.86	1079.2	2.077	0.05
47 UMa c	1.03	1.09	2845.0	3.968	0.00
HD 82943 b	1.05	1.63	444.6	1.159	0.41
HD 82943 c	1.05	0.88	221.6	0.728	0.54
HD 12661 b	1.07	2.30	263.6	0.823	0.35
HD 12661 c	1.07	1.57	1444.5	2.557	0.20
Ups And b	1.30	0.64	4.617	0.058	0.01
Ups And c	1.30	1.79	241.16	0.805	0.27
Ups And d	1.30	3.53	1276.15	2.543	0.25
55 Cnc b	0.95	0.784	14.67	0.115	0.02
55 Cnc c	0.95	0.217	43.93	0.240	0.44
55 Cnc d	0.95	3.912	4517.4	5.257	0.33
55 Cnc e	0.95	0.045	2.808	0.038	0.17

[a]The data are adopted from Refs. 2, 4, and 5.

inner and outer planets),[2] e.g., the 2:1 resonant systems: GJ 876, HD 82943 and HD 128311 (see Table 1), and hierachical systems with $P_1/P_2 \gg 5{:}1$, for example, HD 168443 and HD 38529. In addition, we point out that the above mentioned planets all bear the mass of Jupiter. Due to the limitation of the ground-based observation, the much smaller planets (e.g., the terrestrial planets) cannot be detected at the present. Hence, a direct question is that whether there exist Earth-like planets or terrestrial planets, or other small bodies in these multi-planet systems, just as the situation in our solar system. Because, on the base of planetary formation theory, it is now believed that μm-size dust grains in the planetary debris disks where most of the gas has been dissipated, are presumably accreted to form km-size rocks or planetesimals from physical collisions, then these planetesimals can further grow up to larger bodies as 10^3 km-size planetary embryos by gravity focusing, and finally form the planets.

To study the possibility of the existence of the Earth-like planets in the multi-planet systems, we carried out a systematic numerical explorations of the dynamical structure in several systems. However, herein we mainly concentrate on the investigation of the 47 UMa planetary system (see also Ref. 6), because this system is usually considered to be a close analog of the solar system, for example, the mass ratio of the two giant companions in 47 UMa is $\sim$2.62 (see Table 1), as compared to that of Jupiter–Saturn of 3.34; and the ratios of two orbital periods are close to each other. Several pioneer works[7–10] were concentrated on the structure of the system and presented a preliminary understanding of this issue according to some earlier solutions,[11] where the dynamical model was treated as a restricted multi-body problem. Nevertheless, as the terrestrial planets possess significant masses, they can interact with the two giant planets by mutual gravitation, which may result in secular effects for the planetary system. Accordingly, we should take into account the masses of terrestrial bodies in the model when exploring the dynamical architecture. In this paper, we performed extensive simulations to examine the dynamical architecture in both the HZ and extended areas, for Earth-like planets (with masses from $0.1\,M_\oplus$ to $10\,M_\oplus$) of 47 UMa with stable coplanar planetary configuration, based on the best-fit orbital parameters given by Paper I. On the other hand, in the extended study, we also explored such low-mass planets in the region $0.05\,\mathrm{AU} \leq a < 0.4\,\mathrm{AU}$ and we found that the secular resonance arising from the inner giant planet can render the eccentricity excitations for the Earth-like planets.

2. Simulations Results

In the study, we use N-body codes[12] of direct numerical simulations with the RKF7(8) and symplectic integrators.[13] We always take the stellar mass and the minimum planetary masses from Table 1. The adopted time stepsize is usually ∼1%–2.5% of the orbital period of the innermost planet, which is sufficiently small for the integration. Additionally, the numerical errors were effectively controlled over the integration timescale, and the total energy is conserved to 10^{-6}–10^{-8} for the integrations (see also Ref. 6). In the dynamical study, there are other numerical tools used to explore stability zones for Earth-like planets in planetary systems, e.g., the Fast Lyapunov Indicator (FLI) (e.g., Ref. 14) and the Mean Exponential Growth Factor of Nearby Orbits (MEGNO) (e.g., Ref. 15). Our main results now follow.

2.1. *Terrestrial planets in Habitable zones*

The Habitable Zones are generally conceived as places where the biological evolution of life is able to develop on planetary surfaces of the environment of liquid-water, subtle temperature and atmosphere components of CO_2, H_2O, and N_2.[16] The HZ could be considered to be centered at ∼1 AU $(M_c/M_\odot)^2$ (see Ref. 16). For 47 UMa, the inner and outer boundaries of HZ range from 0.7 to 1.3 AU,[17] however, in our practical simulations, we extended the HZ to other areas for the purpose of a more comprehensive study.

In this series of runs, we extensively investigated the case of two giant companions with one terrestrial planet in the HZ. The mass of the assumed terrestrial planet ranges from 0.1 to 10 $M_\oplus$. And the adopted initial orbital parameters are as follows: numerical scanning was performed for the $[a, e]$ space by direct integrations, where the low-mass bodies were placed at equal intervals of 0.01 AU from 0.05 to 2.0 AU in a, the eccentricities e were taken every 0.01 from 0.0 to 0.2 (and up to 0.1 for 0.05 AU $\leq a <$ 0.4 AU), the inclinations are $0^\circ < I < 5^\circ$, and the other angles were randomly distributed between 0° and 360°. Thus, over 3,000 simulations were exhaustively performed for typical integrations from 1 to 10 Myr.

2.1.1. $0.05\,AU \leq a < 0.4\,AU$

In these runs, we explored the secular evolution of 385 "hot Earths" or "hot Neptunes" for a time span of 1 Myr. All the simulations are dynamically stable for 10^6 years, and 96% of the orbits bear $e_{\rm final} < 0.20$. However,

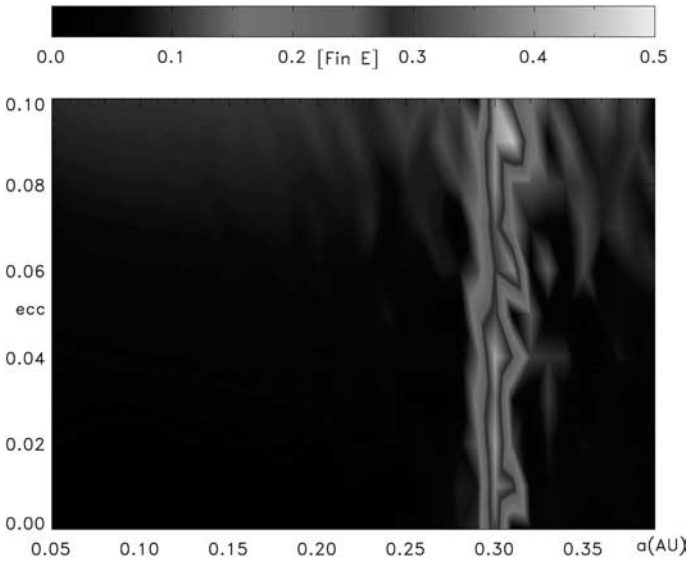


Fig. 1. The contour of status of the final eccentricities for Earth-like planets, the vertical axis for the initial e. Case for 0.05 AU $\leq a <$ 0.4 AU for 1 Myr. Notice ν_1 secular resonance at ∼0.30 AU pumps up the eccentricities.

Fig. 1 shows that the eccentricities for bodies at ∼0.30 AU are excited to ∼0.40, where the secular resonance ν_1 (41″.11/years) of the inner companion (similar to ν_5 for Jupiter) is responsible for the excitation of eccentricity. The debris disk at ∼0.30 AU is also shown by Malhotra,[18] who presented similar results of the eccentricity excitation of massless bodies by nonlinear analytic theory for secular resonance. Nevertheless, one can easily see that this region is not a good location for habitability, due to extra high temperature. G. W. Marcy (2004, private communication) pointed out that the improvement of precision of the ground-based observations will lead to the discovery of additional low-mass planets (∼10 $M_\oplus$) in other known planetary systems. In addition, it is hopeful to detect such planets with $a \geq 0.05$ AU in future space missions (e.g., COROT, KEPLER, TPF).

2.1.2. $0.4\,AU \leq a < 1.0\,AU$

We carried out 1,260 integrations in this region for 5 Myr and we found that none of the orbits escaped during this time span and 94% of them were in the resulting $e < 0.25$, see the final status shown in Fig. 2. We find that the eccentricities of the orbits with 0.70 AU $< a <$ 0.78 AU can be

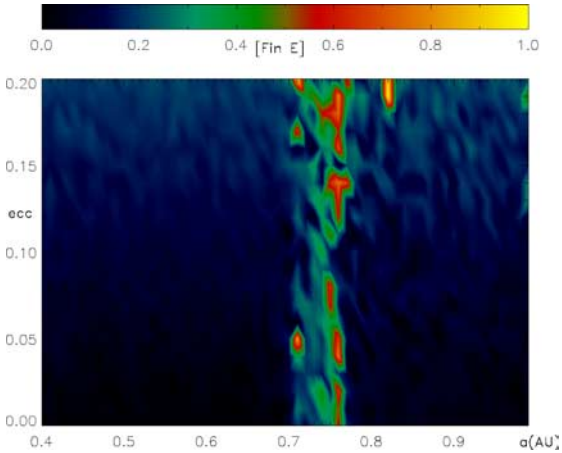


Fig. 2. Case for 0.4 AU $\leq a <$ 1.0 AU for 5 Myr. The e of the orbits with 0.70 AU $< a <$ 0.78 AU can be excited and in the 2:9 MMR at ∼0.76 AU, e can reach ∼0.90.

pumped up and may reach ∼0.90 in the 2:9 MMR at ∼0.76 AU, indicating that there may exist a gap near this resonance. Most of the Earth-like planets about 1:4 MMR at ∼0.82 AU move stably in bounded motions with low-eccentricity trajectories, except for two cases where the eccentricities eventually grow to high values. Paper II pointed out that the secular resonance[b] ν_2 arising from the outer companion (similar to ν_6 for Saturn) can remove the test bodies. Would the ν_2 also influence the Earth-like planets in this system? Nevertheless, we did not find this mechanism at work at about ∼0.85 AU (see Paper II) when we examined the results, because the terrestrial planets under study that all bear finite masses that may change the strength of this resonance; on the other hand, the location of the secular resonance is changed due to the orbital variation of the outer companion. For a terrestrial planet with a mass of 10 $M_\oplus$, the region for ν_2 secular resonance is now shifted to ∼0.70 AU, where two eigenfrequencies for the terrestrial body and outer giant planet given by the Laplace–Lagrange secular theory are, respectively, 211″.37/yr and 225″.48/yr. This indicates that both planets almost have the same secular apsidal precession rates in their motion. At the new location, the ν_2 resonance, together with the mean motion resonance, can work at clearing up the planetesimals in the disk

[b]This means two bodies may bear almost the same precession frequencies in secular orbital motion, Bois *et al.*[15] suggested to use the item "apsidal synchronous precession".

(see Fig. 2) by the excitation of the eccentricity; qualitatively, our results are in accord with those of Paper II.

While the inner edge of HZ marks out a narrow unstable area, it is very possible to discover Earth-like planets in this wider area, $0.4\,\mathrm{AU} < a < 1.0\,\mathrm{AU}$, in future surveys, which are to be the best candidate of habitable places for biological evolution of intelligent beings.

2.1.3. $1.0\,AU \leq a < 1.3\,AU$

There were 630 simulations in this region for 10 Myr and we found that 88% survived the integration, confirming the results given in Paper II that most of the test particles with $a < 1.3\,\mathrm{AU}$ can eventually remain in the system. Here, for the 3:1 resonance at ∼1.0 AU, Fig. 3 shows that there are stable orbits in the zones about 1 AU with $e \leq 0.1$, which agrees with the work by Rivera and Haghighipour[19] who showed that a test particle can last 100 Myr at 1 AU in 47 UMa; while for $0.1 < e \leq 0.2$, the orbits tend to be in unstable state owing to the excitation of the eccentricities. The simulations may imply that the Earth-mass planets near 3:1 resonance are possibly on the edge of stability. However, the previous studies[7,10] on this system showed that the 3:1 resonance is a gap with no survivors. Let us mention that such differences may arise from the adopted initial planetary configurations, and here we adopt the reliable best-fit orbital solutions given

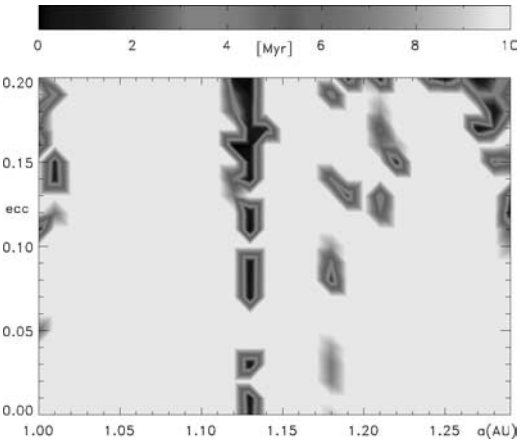


Fig. 3. The surviving time for Earth-like planets for the integration of 10 Myr, the vertical axis for the initial e. Case for $1.0\,\mathrm{AU} \leq a < 1.3\,\mathrm{AU}$, see the gap for the 5:2 MMR at ∼1.13 AU.

in Paper I that can describe the exact motions for the two giant planets. Hence, a comparative run was carried out to examine this, again we ran 630 simulations for 10 Myr but with the earlier orbital elements for the two massive planets to reproduce the previous results at the 3:1 resonance. Our results with the earlier data show that most of the Earth-mass planets about 3:1 resonance are unstable for the investigated time, and their eccentricities can be pumped up to ∼1 through resonance; besides, the inclinations are excited to high values ranging 90°–180°, indicating that the orbits of the Earth-mass planets become retrograde in the dynamical evolution and cross those of the prograde giant planets before they terminate their dynamical lifetimes. Thus, we may safely conclude that the stability of the terrestrial planets depends on the initial planetary configuration.

In Fig. 3, a narrow unstable stripe appears at the 5:2 MMR at ∼1.13 AU, although several of them can be luckily left behind, most of the Earth-size planets are removed within 1 Myr due to the perturbation of 47 UMa b, and in this sense it is analogous to the situation in the solar system. However, a wider area between 3:1 and 5:2 MMR is assumed to be a qualified candidate habitable environment where the Earth-mass planet will not encounter the problem of dynamical stability. This is also almost true for the region (1.13 AU and 1.30 AU) with $e \leq 0.1$, except several unstable islands near 7:3 MMR at 1.18 AU. The smaller eccentricity (near-circular orbits) may not cause dramatic variations of temperature on the planet's surface, so favoring habitability. Therefore, in a dynamical sense, if the 47 UMa system can be adopted as a candidate target for SIM, it is also possible to detect other Earths with stable orbits about 1 AU.

2.1.4. $1.3\,AU \leq a \leq 1.6\,AU$

We performed 651 simulations for 10 Myr, and found the dynamical structure in this regime to be quite complicated: 14% of them can finally survive for this time span, and 86% are lost by ejection into hyperbolic trajectories, indicating the chaotic nature for these bodies. In Fig. 4, we can notice that the 2:1 MMR region is at ∼1.31 AU and also close to the outer edge of HZ, and the orbits with $0.0 < e < 0.10$ are unstable, while for $0.10 \leq e \leq 0.20$, there are stable islands where fictitious planets can remain in bounded motions in the final system. We observe that the 2:1 resonance marks out a remarkable boundary between chaotic and regular orbits, indicating that orbits with $a < 1.31$ AU can have much larger surviving rates than those of $a > 1.31$ AU. However, there are wider stable region

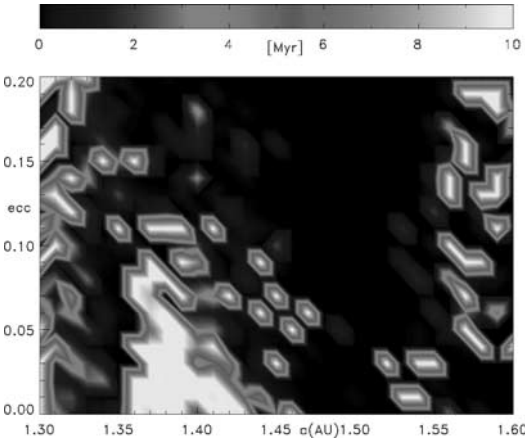


Fig. 4. Case for 1.3 AU $\leq a \leq$ 1.6 AU, a population of the terrestrial planets is about 9:5 MMR at ∼1.40 AU for low eccentricities.

about the 9:5 MMR at ∼1.40 AU for low eccentricities $0.0 < e \leq 0.05$. Most of the unstable orbits are in the region 1.43 AU $< a <$ 1.56 AU, using resonance overlapping criterion,[20] the separation in semi-major axis $\Delta a \approx 1.3\ (M_1/M_c)^{2/7} a_1 \doteq 0.496$ AU (a_1, the semi-major axis of Companion B), thus, the inner boundary $R_O = a_1 - \Delta a$ for 47 UMa b is at ∼1.58 AU, and the orbits in this zone become chaotic during the orbital evolution because the planets are both within $3R_H$ ($R_H = [M_1/(3M_c)]^{1/3} a_1$ is the Hill radius; M_c and M_1 are, respectively, the masses of the host star and the inner planet) and also close to R_O. And the characterized ejecting time $\tau \sim$ 1 Myr, which means the apparent gap (e.g., 5:3 MMR at ∼1.48 AU) in the system, except for several stable islands. Another possible population for terrestrial planets is located at 3:2 resonance ∼1.59 AU for $0.04 < e < 0.20$, and 18 small bodies can last for 10 Myr and confirm the results of Paper II. The 3:2 MMR zone in 47 UMa is reminiscent of the Hilda asteroids in the solar system moving in a stable region.

2.1.5. $1.6\,AU < a \leq 2.0\,AU$

About 840 Earth-like planets are placed in this region for integration for 10 Myr, and the simulation revealed that 98% are removed from the system within a typical ejection time $\tau < 3 \times 10^4$ yr, which is much shorter than for 1.3 AU $\leq a \leq$ 1.6 AU, implying a thoroughly chaotic situation in this

area. Such situation can be well understood by adopting the $3R_H$ stability criterion.

3. Summary and Discussions

In this work, we have systematically studied the existence of Earth-like planets in the region for $0.05\,\mathrm{AU} \leq a \leq 2.0\,\mathrm{AU}$ for 47 UMa by numerical simulations. We now summarize the main results as follows:

(1) The Earth-like planets can eventually remain in the system for 10 Myr in the areas associated with mean motion resonance (e.g., 3:2 MMR) with the inner detected giant companion. We also showed that the 2:1 and 3:1 resonances could be on the fringe of stability, but the 5:2 MMR is unstable and the bodies can be ejected as "extrasolar comets". And this may sketch out an asteroidal belt structure similar to the solar system. Moreover, the 2:1 MMR (near the outer boundary of HZ at 1.30 AU) marks out a significant barrier between chaotic and regular motions, implying that a large fraction of the orbits inside this resonance can be survive, while most of them are lost in the simulations outside the 2:1 resonance. Again, considering the inner boundary of HZ and dynamical stability, we point out that the most likely candidate for habitable environment is terrestrial planets with orbits in the ranges $0.8\,\mathrm{AU} \leq a < 1.0\,\mathrm{AU}$ and $1.0\,\mathrm{AU} < a < 1.30\,\mathrm{AU}$ (except 5:2 MMR, and several unstable cases) with low eccentricities.

(2) However, in our own solar system there are no terrestrial planets from the 1:4 MMR out to Jupiter, although there are stable orbits there. This may suggest that although some orbits are stable, conditions are such that terrestrial planets cannot form so close to giant planets. Perhaps this is because runaway growth is suppressed due to the increased eccentricities from the perturbations of the giant planet. In 47 UMa, the corresponding region runs from 0.82 AU on out, almost completely covering the HZ. Hence, it would be reasonable to conclude that the only proper place to find habitable planets in this system would be at about 0.8 AU. But this should be carefully examined by forthcoming space measurements (e.g., SIM) capable of detecting low-mass planets.

Finally, we can say that the 47 UMa planetary system may bear a similar dynamical structure to our solar system. As a comparison, we also conducted similar studies in other multi-planet systems (e.g., HD 12661, HD 169830, and GJ 876) and found the potential existence of the Earth-like

planets in habitable zones. These studies will be addressed in our forthcoming works.

Acknowledgments

We thank two anonymous referees for useful comments that helped to improve the original contents. This work is financially supported by the National Natural Science Foundations of China (Grant 10573040, 10203005, 10233020), the joint project by Academy of Finland and NSFC and the Foundation of Minor Planets of Purple Mountain Observatory.

References

1. M. Mayor and D. Queloz, *Nature* **378** (1995) 355.
2. D. Fischer *et al.*, *ApJ* **586** (2003) 1394 (Paper I).
3. G. W. Marcy, *Prog. of Theor. Phy. Suppl.* **158** (2005) 24.
4. S. S. Vogt *et al.*, *ApJ* **632** (2005) 638.
5. B. E. McArthur *et al.*, *ApJ* **614** (2004) L81.
6. J. Ji *et al.*, *ApJ* **631** (2005) 1191.
7. B. W. Jones, P. N. Sleep and J. E. Chambers, *A&A* **366** (2001) 254.
8. G. Laughlin, J. Chambers and D. Fischer, *ApJ* **579** (2002) 455 (Paper II).
9. K. Gozdziewski, *A&A* **393** (2002) 997.
10. N. Asghari *et al.*, *A&A* **426** (2004) 353.
11. D. Fischer *et al.*, *ApJ* **564** (2002) 1028.
12. J. Ji, G. Li and L. Liu, *ApJ* **572** (2002) 1041.
13. J. Wisdom and M. Holman, *AJ* **102** (1991) 1528.
14. E. Pilat-Lohinger, R. Dvorak, E. Bois and B. Funk, *ASP Conf. Ser. 321: Extrasolar Planets: Today and Tomorrow* **321** (2004) 410.
15. E. Bois *et al.*, in *Evolution of Habitable Planets*, ISSI, Bern, H. Lammer (ed.), to be appeared (2005).
16. J. F. Kasting *et al.*, *Icarus* **101** (1993) 108.
17. K. Menou and S. Tabachnik, *ApJ* **583** (2003) 473.
18. R. Malhotra, Locating sub-jovian planets and debris in exo-planetary systems, in *AAS, DPS Meeting* **36** (2004) #42.04.
19. E. J. Rivera and N. Haghighipour, *ApJ*, submitted (astro-ph/0406429) (2004).
20. C. D. Murray and S. F. Dermott, *Solar System Dynamics* (Cambridge University Press, New York, 1999).

This page intentionally left blank

JAXA FUTURE PROGRAM FOR SOLAR SYSTEM SCIENCES

MASATO NAKAMURA*, MANABU KATO and YASUMASA KASABA†
Institute of Space and Astronautical Science (ISAS)
Japan Aerospace Exploration Agency (JAXA)
Sagamihara, Kanagawa 229-8510, Japan
**mnakamur@stp.isas.jaxa.jp*
†*kasaba@isas.jaxa.jp*

This paper reviews the proposal of future solar system sciences by JAXA and Japanese science communities, based on the discussion associated with the proposal for the future space programs, "JAXA Vision — JAXA 2025 —", released in March 2005.

1. Introduction

Japan Aerospace Exploration Agency (JAXA) was formed in October 2003 by the unification between three aerospace organizations: Institute of Space and Astronautical Science (ISAS), National Aerospace Laboratory (NAL), and National Space Development Agency (NASDA), which had different backgrounds and roles. In September 2004, the JAXA members started discussion to establish the future vision for Japanese space activities expected for the new agency. In March 2005, it was released as "JAXA Vision — JAXA 2025 —".[1] The Vision contains the proposal of future space activities for the next 20 years. The part for space science activities consists of two blocks, Astrophysics and Solar system. This paper reviews the proposal for latter part from JAXA and Japanese space science communities.

Solar system science missions in Japan have been promoted and organized under ISAS (now ISAS/JAXA). Since the launch of the first Japanese satellite 'Ohsumi' in February 1970, scientists in Japan were actively involved in space projects to create an ambitious future for solar system sciences. Sun and Sun–Earth system were the main targets from the beginning. Shinsei (launched in 1971), Denpa (1972), Kyokko (1978), Jikiken (1978), and Ohzora (1984) made preliminary observations of Earth's ionosphere and magnetosphere. Taiyo (1975) and Hinotori (1981) blazed the trail for the study of the Sun. Those missions led to the success of Akebono

(1989–) for auroral studies, Yohkoh (1991–2004) for solar X-ray imagings, and Geotail (1992–) for the first detailed magnetotail survey.

Deep space missions started with "Sakigake" and "Suisei" which went Comet P/Halley (1985), followed by "Hiten" to the Moon (1990) and "Nozomi" to the Mars (1998–2003). Unfortunately, the last one could not be inserted into the Mars orbit, but those heritages have been expanded to "Hayabusa (Muses-C)" mission to the asteroid Itokawa (2003–), which arrived at Itokawa in the autumn 2005.

The solar system science program stated in the "JAXA Vision" is based on those heritage, in order to propose possible next steps in these 20 years.

2. JAXA Vision — JAXA 2025 —

"JAXA Vision" is not an officially established program, but "a proposal" to the public, for the provision of a clear picture and an ideal future for aerospace activities in Japan. The activities in the proposal are not only limited to the ones inside JAXA but also the ones with various organizations and academic communities.

"JAXA Vision" consists of five main categories listed below, and the space science is described in category (2):

(1) Contributing to building a secure and prosperous society.
(2) Contributing to the advancement of knowledge and the expansion of human frontier: for people's hope and for better future.
(3) Developing indigenous capability to independently carry out space activities.
(4) Facilitating the growth of space industry: toward self-sustained and competitive space industry.
(5) Facilitating the growth of aviation industry and breakthroughs for future air transportation.

They are described on an assumption that the financial resource will be as "the twice as that in 2004" (as about 4/3 of the one in 1999) to realize the Vision. (The activities which require much resource, e.g., independent human space program and reusable vehicles, are not included in first 10 years. To start these or not will be decided in 2015, 10 years from now.)

Category (2) of the Vision consists of two medium targets, "Space science" and "Lunar utilization". The former is the result of extensive discussion by young members, not only inside JAXA but also in space science

and engineering communities in Japan, who will have the responsibilities for the activities in 2025.

Exploration of the solar system and observations of the universe will be made in cooperation with many academic and research institutions in Japan and overseas. It is an ultimate goal for JAXA to make Japan to be one of the topmost space science centers in the world.

3. Vision for Solar System Sciences

We consider the future solar system sciences described in "JAXA Vision" to be very much a practical expectation, based on our heritage, resource, and potential capabilities. All activities shall be directed to our final target, "*Raison d'etre* of the planet(s) with life", through the studies of:

(a) the future of the Sun and the Earth,
(b) the conditions for habitable planetary system(s),
(c) arrival at the unexplored region of the solar system, and
(d) the universality of solar system(s) in the Cosmos.

3.1. *Phase-1 (2005–2015): ongoing missions and the missions in preparation*

The first 10 years (2005–2015, Phase-1) of the Vision consist of currently ongoing and approved missions. The main targets of them are the detailed studies of Sun and Sun–Earth system, the initiation of the surveys of terrestrial objects (Moon, Venus, and Mercury), and the first sample return from an asteroid.

3.1.1. *Sun and Sun–Earth system*

Detailed studies for the solar and Sun–Earth system are expected to be active areas in the context of space weather research and the *in situ* physical laboratories for energetic plasma processes universally found in the space.

Based on the success of the Yohkoh mission (cf. Ref. 2), Solar-B (a joint mission with NASA, PPARC, and ESA) will be launched in 2006 (cf. Ref. 3). Solar-B will, for the first time, provide quantitative measurements of the full vector magnetic field in sufficiently small scales to resolve elemental flux tubes. The main objectives are to study (1) the creation and destruction of the Sun's magnetic field, (2) the modulation of the Sun's luminosity, (3) the generation of UV and X-ray radiations, and (4) the

eruption and expansion of the Sun's atmosphere. SOLAR-B has visible, X-ray and EUV telescopes with highly leveraged international participation, and will greatly advance our understanding of the crucial first link in the Sun–Earth connection.

For Sun–Earth system studies, Akebono (1989–; in collaboration with Canada) and Geotail (1992–; a joint mission with NASA) are still in good health. The primary objective of Akebono (cf. Refs. 4 and 5) is to study auroral acceleration processes, using its highly elliptic polar orbit (75° inclination, 275 km perigee, and 10,500 km initial apogee). Geotail (cf. Ref. 6) was the first spacecraft to make detailed observations of the magnetotail, the most important region of the Sun–Earth connection. Geotail was initially in an $8 \times 210\,R_E$ equatorial orbit. Its apogee was then lowered to $30\,R_E$ in early 1995, so as to study substorm processes in the near-tail region. Geotail continues to be a unique spacecraft in its present survey of the equatorial magnetosphere.

As a new mission "Reimei" was launched in August 2005. It is a piggyback microsatellite, in a 610 km Sun-synchronous orbit near 1 LT. The scientific objectives of Reimei are fast and fine-scale auroral structure studies. It carries a suite of CCD cameras, electrostatic electron and ion analyzers, and a Langmuir probe. Reimei will undertake extensive collaborative studies with ground observatories and other polar satellites. Meanwhile, National Institute of Information and Communications Technology (NiCT) is planning another small satellite program, "SmartSat". The SmartSat weighs about 150 kg, and will be a collaborative program of government agencies (NiCT & JAXA) and the private sector (Mitsubishi Heavy Industries). The space weather experiment of SmartSat consists of WCI (wide field imager for CME tracking) and SEDA (advanced high-energy particle sensor). These satellites are under development for launch around 2008. And their payloads will be the principal components of the L5 mission for space weather research and operational forecasting experiment proposed by the NiCT.[7] The spacecraft will be located at the L5 point of the Sun–Earth system for remote sensing of the Sun and the interplanetary space and for *in situ* measurement of the solar wind and high-energy solar particle events.

3.1.2. *Initiation of the surveys of terrestrial objects: Moon, Venus, and Mercury*

Moon was the first target of Japanese missions orbiting around solar system objects other than the Earth, by "Hiten" (1990). JAXA is preparing

the lunar mission "Selene" (cf. Ref. 8) (Launch: 2007). It will study lunar origin, evolution, and environment. The main orbiter will be in a 100-km altitude polar circular orbit around the Moon, while its two sub-satellites will be in elliptical orbits. The main orbiter will carry: X-ray/Gamma-ray spectrometers, a multi-band imager/spectrometer for studies of elemental and mineralogical composition; a terrain camera, a laser altimeter and a radar sounder for topological studies of the lunar surface and subsurface; a plasma imager for monitoring the Earth's aurora and plasmasphere; a magnetometer for measuring the remnant magnetic field; and a charged particle spectrometer, a plasma analyzer, and plasma wave detector for *in situ* studies of the magnetosphere and lunar environment. The S- and X-band telemetry will also be used for radio occultation studies of the tenuous lunar ionosphere. And the two sub-satellites will act as differential VLBI radiosources and relay satellites for precise determination of lunar gravity field.

Beyond the disappointing results of the Nozomi orbiter to Mars,[9] JAXA is aiming two orbiters toward our neighborhood planets, PLANET-C to Venus, and BepiColombo to Mercury. PLANET-C, known as the Venus Climate Orbiter (VCO)[10] (Launch: 2010), will focus on the exploration of Venus's atmospheric dynamics. The VCO will be in a highly elliptic near-equatorial orbit ($300\,\mathrm{km} \times 13\,\mathrm{R_V}$, with an inclination of 172°). The orbit is selected to lock onto the angular motion of Venus's cloud layer when VCO is at its apoapsis, to make possible continuous imaging and visualization of the global circulation in 3-dimensions, so as to address the questions of (a) how the super-rotation is generated, (b) how the meridional circulation contributes to the super-rotation, and (c) how the clouds on Venus are generated. In order to address these questions, the VCO will carry a suite of short, medium, and long-wavelength IR imagers as well as a UV imager and a lightning/airglow camera.

As part of the joint ESA-JAXA BepiColombo mission to Mercury (Launch: 2012), the Mercury Magnetospheric Orbiter (MMO) will be led by JAXA.[11] The main objectives of MMO are (1) the first detailed survey of the magnetic field of the Mercury's magnetized body, (2) the first detailed survey of the unique magnetosphere of Mercury, (3) the first detailed study of Mercury's tenuous and unstable exosphere, and (4) detailed monitor of the inner heliosphere, in close collaboration with the Mercury Planetary Orbiter (MPO). MMO will have a magnetometer, plasma, and energetic neutral particle package, a plasma, wave instrument, an imaging spectroscopic for sodium exosphere, and a dust counter. The MPO led by ESA

will study the interior and the surface of the planet. Both payload teams consist of scientists from the Japanese and European communities.

Those data will be opened to the public, and the Japanese space science community will welcome the participation of planetary science activities in our projects.

3.1.3. *The first sample return from a near-Earth asteroid*

In order to unveil the primitive materials in the solar system, the Hayabusa spacecraft (cf. Ref. 12) was launched in May 2003. It made a rendezvous with the near-Earth asteroid "Itokawa" (named after one of the founders of the ISAS space program) from September to November 2005. Hayabusa is an engineering test spacecraft, and is designed to make the first sample return from asteroids: its goal is to rendezvous and dock with the asteroid, then take the samples from the asteroid surface and bring them back. In November 2005, Hayabusa tried to take samples from the asteroid's surface. Hayabusa will return to the earth with first asteroid materials in 2007.

3.2. *Phase-2 (2015–2025): to learn about the Sun–Planetary connection*

The main targets of Phase-2 (2015–2025) are the common cosmic physical processes in the all magnetospheres, detailed studies of the neighborhood planets Mars and Venus, and the expansion of our fields to the outer solar system, e.g., Jupiter and asteroids further than Mars. We consider that some of the missions should be initiated and launched before 2015.

3.2.1. *Sun and Sun–Earth system*

The main targets for the Sun and the Sun–Earth system in this decade will be advanced studies of energetic processes in the magnetosphere related to the basis of space weather and of physical process in the magnetosphere that are common cosmic processes. For these objectives, two new solar terrestrial missions are proposed: cross Scale COupling in Plasma universE (SCOPE), a mission to study the cross-scale coupling in the plasma universe, and Energization and Radiation in Geospace (ERG), a mission to investigate energization and radiation in geospace.

The ERG mission (cf. Ref. 13) is planned as a small-sized satellite, aiming to study the acceleration and loss mechanisms of relativistic particles in the radiation belts during magnetic storms. Its proposed orbit has a

perigee of 250 km, an apogee of 6.6 R_E, and an inclination of 10°. Its proposed instrument payload includes B- and E-field sensors, electron and ion analyzers with continuous energy coverage from a few eV to MeV, to capture the instability and acceleration processes in the radiation belt directly. The launch date is not fixed yet, but we hope to realize this mission in early 2010s.

The SCOPE mission will be a full-scale "Post-MMS (Smart)" formation flight satellite mission, aiming to probe processes of a variety of spatial and temporal scales, from the scales of electron and ion gyro-radius and inertial lengths, to MHD scales. It is conceived as a constellation of a mother spacecraft and four daughter spacecraft, one of which will be separated by 5–100 km from the mother spacecraft and the other three 5–5,000 km away. The orbit will have a perigee near 7,000 km and an apogee of 30 R_E. This will provide ample opportunities to study regions of dayside/nightside magnetic reconnection, shocks, and many kinds of highly non-linear processes.

3.2.2. *Detailed studies for Mars and Venus*

The mission following BepiColombo is not fixed yet. A "Nozomi recovery mission" to Mars is debated.

Since the unexpected termination of the Nozomi mission in 2003, no other spacecraft with a primary focus on the planetary atmosphere's interaction with the solar wind has been sent to Mars. Therefore, numerous intrinsic questions about the Martian upper atmosphere are still open. The next Mars mission is now being extensively discussed mostly from a scientific point of view, including the objectives proposed in the Nozomi mission. In particular, elucidation of the atmospheric escape process and its impact on atmospheric evolution will be one of the most important subjects.

Studies of the Venusian atmosphere "Post-VCO" are also under study, involving the combination of imaging using wide field of view cameras like VCO and *in situ* observations by airplane or balloon experiments. The possible Venus balloon missions will explore the lower atmosphere of Venus below the main cloud layer. The balloon under consideration is a superpressure type using water vapor as the lifting gas, aiming to float at an altitude lower than 40 km for more than a few weeks. Such a balloon will enable meteorological measurements, identification of haze particles below clouds, and optical sounding of the ground surface at wavelengths in the near-infrared atmospheric windows. Key technologies will be high-temperature electronics including local oscillators, batteries, and solar cell arrays.

Those activities might be connected to 'Exploration' regime (see Sec. 3.2.5).

3.2.3. *Initiation of the surveys of the outer solar system*

One candidate in this category is the Solar Sail project officially proposed in 2004, for proofing "Solar Sail" technologies in deep space exploration. It is also expected to act as a multi-scientific platform, including (a) cosmic IR background survey, (b) gamma-ray burst detection, (c) dust measurement with huge-sized counter, (d) main-belt and Torojan asteroid flyby, and (e) Jupiter system observation. For the last objectives, the mother spacecraft will carry the daughter "orbiter" and small-sized "entry-probe". In the base plan, the launch is expected in early 2010s.

Although the payload capability in this project is much limited, they will lead to the outer planet exploration by JAXA. In the decade 2015–2025, we will expect a full-scale mission to Jupiter and the Jovian system, for the study of the largest planets and most complex satellites system, and the multiple sample return mission to the asteroids and comets in outer solar system to unveil the primitive materials of the life precursors.

3.2.4. *Small spacecraft*

Even though ISAS and ISAS/JAXA has been developing only medium size science spacecraft in recent years, there are many requests from university scientists that JAXA develops common small satellite bus systems for quick and focused science. The first good example was the piggy-back mission "Reimei".

Such request is urgent because the medium class science spacecraft is launched only once every 1.5–2 years after 5–10 years of preparation. Such cycle does not match well the life cycle of the scientists. In 2005, JAXA sent out questionnaires to the space communities on the kinds of common bus needed. There were 23 responses. Most of them were categorized into the class of 50–100 kg science payload. The accuracy and stability requirement for the attitude control is not very severe, such as $0.01^\circ\,s^{-1}$. We believe that it is not difficult to design and build such a bus. However, the question is what kind of launcher should be assumed for the design. And human and financial resource to start and keep small satellite programs is not resolved yet. It will likely take some time before we can come to a conclusion.

3.2.5. *Exploration: Moon and beyond, linked to the utilization of space*

Even though science will be main driving force, lunar programs will be categorized in the "Utilization" regime. Establishing the advanced engineering requested for new *in situ* studies and sample return missions, we will develop the technologies for penetrators and landers in several engineering missions to Moon.

Lunar-A (cf. Ref. 14) is a lunar penetrator mission. Its aim is to study the lunar interior, by deploying two penetrators on the near and far side of the lunar surface, respectively, and using seismometers and heat-flow probes in the penetrators. Selene-2 is also planned, as a lander mission with much more precise location selection capability. Since the technical development is not established yet for these, so its launch date is not fixed for the time being.

These will not be limited to the lunar science, but will directly expand the region we can access, including neighborhood planets and beyond.

4. Summary

In the past, the missions of ISAS and ISAS/JAXA were planned and proposed by working groups through grassroots level discussions, and developed by scientists and engineers in academic communities. A mission was selected every 1 or 2 years, purely based on science and engineering merit. Such an approach worked well, because there were so many science targets to be tackled by small-sized communities. We selected the best one at the day of the selection each time. However, the long-term planning and categories of science targets were not considered insufficient.

Now, the development of the long-term vision is required due to the increasing technical difficulty and resources for each mission for next decades. The planetary program especially requires much longer mission period, much more advanced technologies, and larger resources. Therefore, we cannot justify such a program without a long-term vision and a change of the development framework.

Now, scientists and engineers in JAXA have started the discussion of the detailed strategic plan based on the road map from the Vision. The most important point will be the establishment of a "new exploration program" (see Sec. 3.2.5). At this stage, the assumed target is Moon. But there are also other alternatives being forcefully advocated. For example, sample return missions of solid, liquid, and atmospheric components of Mars and its

satellites. Venus and minor objects will also be realistic themes beyond the initial engineering programs based on Moon. The decision will be expected in the near future.

We need extensive discussions among scientists and engineers, not only in JAXA but also in Japan and in foreign communities. The JAXA space science programs should be discussed, decided, and promoted by ourselves.

Acknowledgments

The authors would like to thank all the members who have contributed to past and current space science programs of ISAS and ISAS/JAXA. They also acknowledge the contributions of young scientists in the Japanese space communities who contributed to the intense discussions for future space programs.

References

1. JAXA, "JAXA Vision — JAXA 2025 —", http://www.jaxa.jp/about/vision_missions/long_term/index_e.html, March 2005.
2. T. Watanabe, T. Kosugi and A. C. Sterling (eds.), *Observational Plasma Astrophysics: Five Years of Yohkoh and Beyond* (Kluwer, Dordrecht, 1998).
3. T. Sakurai and T. Sekii (eds.), The SOLAR-B mission and the forefront of solar physics — the fifth solar-B science meeting —, *ASP Conference Series* **325** (2004)
4. H. Oya and K. Tsuruda, Introduction to the Akebono (EXOS-D) satellite observations, *J. Geomag. Geoelectr.* **42** (1990) 367–370.
5. K. Tsuruda and H. Oya, Introduction to the EXOS-D (Akebono) project, *Geophys. Res. Lett.* **18** (1991) 293–295.
6. A. Nishida, D. N. Baker and S. W. H. Cowley (eds.), New perspectives on the Earth's magnetotail, *Geophys. Monograph* **105** (1998).
7. M. Akioka, T. Nagatsuma, W. Miyake, K. Ohtaka and K. Marubashi, The L5 mission for space weather forecasting, *Adv. Space Res.* **35** (2005) 65–69.
8. M. Kato, S. Sasaki, Y. Iijima, K. Tanaka, H. Mizutani and K. Tsuruda, Science instruments and their development in SELENE mission, *Proc. ICEUM-4* (*ESA SP-462*) (2000) 119–123.
9. T. Yamamoto and K. Tsuruda, The planet-B mission, *Earth, Planets and Space* **50**, 3 (1998) 175.
10. T. Imamura, M. Ueno, N. Iwagami, T. Satoh, M. Nakamura, S. Watanabe, M. Taguchi, Y. Takahashi, M. Suzuki, T. Sakanoi, M. Yamada, S. Okano, G. L. Hashimoto, Y. Kasaba, J. Yoshida, T. Abe, N. Ishii, T. Yamada and K. Oyama, Venus climate orbiter mission (PLANET-C) of Japan, *Planet. Space Sci.*, submitted, 2005.

11. T. Mukai, H. Yamakawa, H. Hayakawa, Y. Kasaba and H. Ogawa, Present status of the BepiColombo/Mercury magnetospheric orbiter, *Adv. Space Res.*, in press, 2005.
12. A. Fujiwara, J. Kawaguchi and K. T. Uesugi, Role of sample return mission MUSES-C in asteroid study, *Adv. Space Res.* **34** (2004) 2267–2269.
13. K. Shiokawa, K. Seki, Y. Miyoshi, A. Ieda, T. Ono, M. Iizima, T. Nagatsuma, T. Obara, T. Takashima, K. Asamura, Y. Kasaba, A. Matsuoka, Y. Saito, H. Saito, M. Hirahara, Y. Tonegawa, F. Toyama, M. Tanaka, M. Nose, Y. Kasahara, K. Yumoto, H. Kawano, A. Yoshikawa, Y. Ebihara, A. Yukimatsu, N. Sato, S. Watanabe and the Inner Magnetosphere Subgroup in the Society of Geomagnetism and Earth, Planetary and space sciences, ERG — a small-satellite mission to investigate the dynamics of the inner magnetosphere, *Adv. Space Res.*, in press, 2005.
14. H. Mizutani, A. Fujimura, S. Tanaka, H. Shiraisi and T. Nakajima, LUNAR-A mission: Goals and status, *Adv. Space Res.* **31** (2003) 2315–2321.