UNIVERSITY OF HELSINKI REPORT SERIES IN ASTRONOMY

No. 6

Asteroid astrometric and photometric studies using Markov-chain Monte Carlo methods

Dagmara Anna Oszkiewicz

Academic dissertation

Department of Physics
Faculty of Science
University of Helsinki
Helsinki, Finland

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Auditorium PHY E204 on June 27, 2012, at 12 o'clock noon.

Helsinki 2012

Cover:
Markov-chain Monte Carlo methods and Bayesian statistics have been used to infer dynamical and physical properties of asteroids.

ISSN 1799-3024 (printed version)
ISBN 978-952-10-7092-1 (printed version)
Helsinki 2012
Helsinki University Print (Unigrafia)

ISSN 1799-3032 (pdf version)
ISBN 978-952-10-7093-8 (pdf version)
ISSN-L 1799-3024
http://www.ethesis.helsinki.fi
Helsinki 2012
Electronic Publications @ University of Helsinki
(Helsingin yliopiston verkkojulkaisut)

Dagmara Anna Oszkiewicz: **Asteroid astrometric and photometric studies using Markov-chain Monte Carlo methods**, University of Helsinki, 2012, 49 p.+63 p. original publications, University of Helsinki Report Series in Astronomy, No. 6, ISSN 1799-3024 (printed version), ISBN 978-952-10-7092-1 (printed version), ISSN 1799-3032 (pdf version), ISBN 978-952-10-7093-8 (pdf version), ISSN-L 1799-3024

Classification (INSPEC): A9510J, A9575D, A9510C, A9510E, A9575P, A9630H
Keywords: Astrometry and spherical astronomy, Astronomical photographic and electronic imaging, and photometry, Celestial mechanics, Orbit determination and improvement, Mathematical and computer techniques in astronomy, Asteroids.

Abstract

Planetary Science belongs to the so-called fundamental sciences, which do not have to have immediate practical applications or implications. The recent decades have however shown that the study of asteroids may have direct implications on our life. Studies of asteroid dynamics have shown that some of those objects can collide with the Earth. Studies of asteroid mineralogy suggest that some of them contain minerals and elements important for industry. For both of those topics, determining physical and dynamical properties is crucial.

Markov-chain Monte Carlo methods and algorithms such as the Metropolis-Hastings algorithm are growing in popularity and becoming important tools in deriving model parameters in many branches of science today. In this thesis, Bayesian statistics along with the above-mentioned numerical methods have been used to infer dynamical and physical properties of asteroids.

First, a new Markov-chain Monte Carlo ranging method is developed for computing asteroid orbits. The method is applicable for asteroids with short observational time intervals and/or small number of observations. The method is particularly useful in deriving orbits for new asteroid discoveries and computing collision probabilities for such objects. The Markov-chain Monte Carlo ranging method is applied to a number of asteroids including a recent Earth impactor - asteroid 2008 TC_3. Markov-chain Monte Carlo ranging is available through the open-source orbit-computation package called OpenOrb and is implemented into the Gaia satellite data processing pipeline, where it will be heavily used in the daily data processing.

Second, Markov-chain Monte Carlo and Monte Carlo methods are used to assess phase curve photometric parameters and their uncertainties. Absolute magnitudes and photometric parameters are derived for half a million asteroids by fitting phase curves to the Lowell Observatory photometric database. Asteroid phase curves depend on physical properties of regolith and absolute magnitudes are useful in computing sizes and albedos. Fitting the phase functions to a large number of asteroid families suggests homogeneity of photometric parameters in asteroid families. The derived photometric parameters are also found to correspond to asteroid taxonomic complexes and colors.

Acknowledgements

I owe gratitude to very many who contributed directly or indirectly to this thesis and supported me during my studies. In chronological order:

It is with immense gratitude that I acknowledge the support and help of my supervisor Prof. Karri Muinonen whose enthusiasm, patience and encouragement was crucial for this work. Without his scientific guidance and persistent help this dissertation would not have been possible.

I would like to thank Dr. Jenni Virtanen, Dr. Antti Penttilä and Dr. Mikael Granvik for their research assistance and suggestions. Furthermore, I thank all the colleagues from the Helsinki Observatory and later the Division of Geophysics and Astronomy who made my stay in Finland a pleasant experience.

It gives me great pleasure in acknowledging the Gaia community and the ELSA network. I greatly appreciated the possibility to follow the development of the Gaia satellite and the data analysis pipeline. Thank you for all the meetings, workshops and schools, for inspiration and your insight. I thank the ELSA students and fellows for all the fun, adventures and activities. In particular I would like to acknowledge Dr. Michael Weiler whose great sense of humor helped me survive many of the steering committee meetings.

I am really grateful to the Lowell Observatory, especially Dr. Edward Bowell and Dr. Larry Wasserman, and Northern Arizona University staff, in particular, Prof. David Trilling and Dr. Cristina Thomas for making my stay in the US a fruitful and inspiring experience. I cannot find words to express my gratitude to Dr. Edward Bowell whose wisdom and scientific guidance is greatly appreciated. It was a huge pleasure and an honor to work with and benefit from his experience and scientific understanding. Thank you both Ted Bowell and Ann-Marie Malotki for all your help, taking care of us, all your advice, sharing your life experience, creating an exceptional atmosphere and making me feel at home many miles away from home. Flagstaff truly is one of the best places on earth and the time spent there will long stay in my memory. I would also like to thank Brian Skiff and the volleyball team at the Lowell Observatory!

I would like to show my gratitude to the staff and students from the Nordic Optical Telescope (and other telescopes at the Roque de los Muchachos Observatory). It was a pleasure to learn the observing techniques, telescope operations and life cycle of a telescope at the NOT. Most of all thank you for having the patience to observe my asteroids and comets and for Bonanza.

I am grateful to all the Planetary Science community for the useful comments and inspirations. Particularly I thank the Poznań Observatory staff and students (especially Prof. Tadeusz Michałowski, Dr. Magdalena Polińska, Dr. Anna Marciniak, Dr. Regina Rudawska, Dr. Tomasz Kwiatkowski and Dr. Agnieszka Kryszczyń ska) whose presence and kind attitude at many of the international conferences and meetings made my life easier. Thank you also for all the help and discussions and for having your door always open on my visits to Poznań.

Last but not the least, thank you to all the family and friends for moral support, following my trips and hours spent on skype. Especially, I would like to thank my

parents who were always there to support me and waited for my visits home [In Polish]:

Z głębi serca chciałabym podziękować moim Rodzicom, za miłość i wsparcie jakiego mi udzielali w przeciągu lat spędzonych poza granicami kraju. Dziękuję za wszystkie dobre rady, godziny przegadane na skypie oraz za miejsce do którego wiem że zawsze mogę powrócić - dom rodzinny.

Finally, I would like thank everybody who crossed my way during those years.

This thesis was financially supported from the EU Marie Curie TMR project ELSA, Academy of Finland, Magnus Ehrnrooth foundation, Northern Arizona University, Lowell Observatory and Nordic Optical Telescope.

List of papers

I **Oszkiewicz D. A.**, Muinonen K., Virtanen J., Granvik M., Asteroid orbital ranging using Markov-Chain Monte Carlo. *Meteoritics & Planetary Science* **44**, 1897-1904 (2009).

II Granvik M., Virtanen J., **Oszkiewicz D. A.**, Muinonen K., OpenOrb: Open-source asteroid orbit computation software including statistical ranging. *Meteoritics & Planetary Science* **44**, 1853-1861 (2009).

III **Oszkiewicz D. A.**, Muinonen K., Bowell E., Trilling D., Penttilä A., Pieniluoma T., Wasserman L. H., Enga M.-T., Online multi-parameter phase-curve fitting and application to a large corpus of asteroid photometric data. *Journal of Quantitive Spectroscopy & Radiative Transfer* **112**, 1919-1929 (2011).

IV **Oszkiewicz D. A.**, Bowell E., Wasserman L. H., Muinonen K., Trilling D., Penttilä A., Pieniluoma T., Asteroid taxonomic signatures from photometric phase curves. *Icarus* **219**, 283-296 (2012).

V **Oszkiewicz D. A.**, Muinonen K., Virtanen J., Granvik M., Bowell E., Modeling collision probability for Earth-impactor 2008 TC_3. *Planetary and Space Science, in press.*

Acronyms used in the text

AU	Astronomical Unit
MBO	Main-belt Object
MBA	Main-belt Asteroids
NEO	Near-Earth Object
NEA	Near-Earth Asteroids
PHO	Potentially Hazardous Object
SSO	Solar System Object
TNO	Transneptunian Object
KBO	Kuiper Belt Object
MC	Monte Carlo
M-H	Metropolis-Hastings (algorithm)
MCMC	Markov chain Monte Carlo
O–C	Observed minus Computed (residuals)
p.d.f.	probability density function
WAM	Wavelet Analysis Method
HCM	Hierarchical Clustering Method
YORP	Yarkovsky-O'Keefe-Radzievskii-Paddack (effect)
CCD	Charge-coupled device
WISE	Wide-Field Infrared Survey Explorer
IRAS	Infrared Astronomical Satellite
SMASS	Small Main-Belt Asteroid Spectroscopic Survey
SDSS	Sloan Digital Sky Survey
NOT	Nordic Optical Telescope
DPAC	Data Processing and Analysis Consortium
CU4	Coordination Unit 4
ESA	European Space Agency
NASA	National Aeronautics and Space Administration
JPL	Jet Propulsion Laboratory
MPC	Minor Planet Center

Symbols and units

$$1 \text{ AU} \quad = \quad 149.598.000 \text{ km}$$

a	semi-major axis
e	eccentricity
i	inclination
Ω	longitude of ascending node
ω	argument of perihelion
M_0	mean anomaly at epoch t_0
p_p	a posteriori probability density function
p_{pr}	a priori probability density function
p_ϵ	likelihood function
$\mathbf{P}$	model parameters
A	Bond albedo
p	geometric albedo
q	phase integral
α	phase angle
D	diameter
H	absolute magnitude
V	apparent V-band magnitude
G_{12}	photometic parameter of the H, G_{12} phase function
G_1	photometic parameter of the H, G_1, G_2 phase function
G_2	photometic parameter of the H, G_1, G_2 phase function
G	photometic slope of the H, G phase function
k	photometic slope of the H, G_{12} and H, G_1, G_2 phase functions

Contents

1 Introduction **1**

1.1 Discovery of asteroids 1

1.2 Dynamical populations of asteroids 2

1.3 Physical properties 4

1.4 Gaia mission and asteroids 7

1.5 Outline of this work 8

2 Inverse problems theory **9**

2.1 Bayesian solution to inverse problems 9

2.2 Advantages of Bayesian method 10

2.3 Markov-chain Monte Carlo technique for solving inverse problems . 11

2.3.1 Convergence diagnostics 13

2.3.2 Markov-chain Monte Carlo vs. Monte Carlo 14

3 Application: Asteroid orbits **15**

3.1 Brief history of orbit computation 15

3.2 Orbital inverse problem 16

3.3 MCMC ranging technique 17

3.4 Selected results 18

4 Application: Asteroid photometric parameters **22**

4.1 Obtaining phase curves from traditional observations 22

4.2 Calibration of Minor Planet Center photometry 23

4.3 Phase functions 24

4.4 MC and MCMC error computation 27

4.5 Selected results 27

5 Relevance to Gaia mission **30**

6 Summary of papers **31**

6.1 Paper I: Oszkiewicz D. A., Muinonen K., Virtanen J., Granvik M., Asteroid orbital ranging using Markov-Chain Monte Carlo. *Meteoritics & Planetary Science* **44**, 1897-1904 (2009) 31

6.2 Paper II: Granvik M., Virtanen J., Oszkiewicz D. A., Muinonen K., OpenOrb: Open-source asteroid orbit computation software including statistical ranging. *Meteoritics & Planetary Science* **44**, 1853-1861 (2009). 31

6.3 Paper III: Oszkiewicz D. A., Muinonen K., Bowell E., Trilling D., Penttilä A., Pieniluoma T., Wasserman L. H., Enga M.-T., Online multi-parameter phase-curve fitting and application to a large corpus of asteroid photometric data. *Journal of Quantitive Spectroscopy & Radiative Transfer* **112**, 1919-1929 (2011) 31

6.4 Paper IV: Oszkiewicz D. A., Bowell E., Wasserman L. H., Muinonen K., Trilling D., Penttilä A., Pieniluoma T., Asteroid taxonomic signatures from photometric phase curves. *Icarus* **219**, 283-296 (2012) 32

6.5 Paper V: Oszkiewicz D. A., Muinonen K., Virtanen J., Granvik M., Bowell E., Modeling collision probability for Earth-impactor 2008 TC3. *Planetary & Space Science, in press.* 32

6.6 Author's contributions 32

References **34**

1 Introduction

1.1 Discovery of asteroids

Today we know over half a million asteroids (MPC accessed 2012), orbiting the Sun. However the history of asteroid discovery begins back in the 16th century, when Johannes Kepler, discovered the laws of planetary motion. Kepler's laws are (Shu 1982):

- The orbit of a planet forms an ellipse with the Sun at one focus.
- The Sun-planet radius vector sweeps out equal areas in equal times.
- The square of the period of revolution of a planet is proportional to the cube of the semimajor axis of its elliptical orbit.

Kepler's laws are restricted to so-called two-body motion (two point-like sources not affected by gravity of other planets), which is a useful first approximation to calculating the orbits and positions of planets and small bodies. Kepler's laws led the past astronomers to realize that the distance between Mars and Jupiter is not proportional to distances between other planets and that there must be another, yet undiscovered planet between the orbits of Mars and Jupiter. The hypothesis of a new planet was additionally strengthened by the discovery of an empirical law by Johann Titius (Jaki 1972). The law describes the planetary distances as a series of numbers:

$$y = 0.4, 0.4 + (0.3 \times 2^n), \tag{1}$$

where $n = 0, 1, 2, 3, ...$ The resulting y values give the approximate heliocentric distances of the major planets in Astronomical Units (AU). The first planet is Mercury at 0.4 AU. The law was then popularized by J. E. Bode and quickly became the "Titius-Bode law". Discovery of Uranus with semimajor axis at 19.2 AU in the 18^{th} century by William Herschel seemed to confirm the empirical law and strengthen the belief in the existence of the yet undiscovered planet between the orbits of Mars and Jupiter. Inspired by the discovery of Uranus (Alexander 1965) and the Titius-Bode law astronomers began a search for the new planet near 2.8 AU. Finally, Giuseppe Piazzi discovered (now a dwarf planet) Ceres on 1 January 1801 (Serio et al. 2002). The discovery was then published in the September 1801 issue of the Monatliche Correspondenz, but by this time Ceres moved into conjunction with the Sun making it impossible to observe. By the end of the year the new planet should have been visible again, but after such a long time it was impossible to predict its exact position with the methods available at that time. To recover Ceres, C. F. Gauss developed an orbit-computation method based on three observations and succesfully predicted the position of Ceres (Teets & Whitehead 1999). Subsequent observations were made to improve the orbit of Ceres and allow for more precise follow-up. In 1802, H. W. M. Olbers found a second moving object in the vicinity of Ceres, later named Pallas (Lardner & Dunkin 1860). Pallas posed a potential problem since now there were two objects where the Titius-Bode law predicted only

one. J. F. W. Herschel computed sizes of Ceres and Pallas and showed that they are much smaller than any of the other planets (Bus 1999). Herschel proposed a new class of objects called asteroids (star-like) (Armitage 1962) and hypothesized that they are fragments of a planet that broke into pieces and suggested that additional fragments might be found. As predicted, soon asteroid Juno was discovered in 1804 by K. L. Harding and (Lardner 1854) by Olbers. By the end of 1890, over 300 asteroids were known, and the region was called the "asteroid belt" (now known as the asteroid main belt). One of the asteroids from the main-belt is pictured in Fig. 1.

Figure 1: Asteroid (4) Vesta as seen by NASA's Dawn spacecraft. Source: `http://dawn.jpl.nasa.gov/`

The next big revolution in the discovery of asteroids happened with the discovery of photographic techniques, which allowed for more efficient searches and more precise and accurate positional measurements. Some progress was made with improving photographic emulsions, light amplification etc., but starting in 1970s after the invention of the CCD, photographic plates have given way to electronic imaging in professional observatories (Bus 1999). Development of CCD techniques, giant multi-mirror, segmented mirror telescopes, adaptive optics and space-based telescopes led to a great improvement in the asteroid research field. Today asteroids are considered to be remnants of the protoplanetary disk, and they are not only found in the asteroid belt, but also in other parts of our Solar System.

1.2 Dynamical populations of asteroids

The vast majority of asteroids orbit the Sun in the asteroid main-belt (located between the orbits of Mars and Jupiter between 1.8 and 4.5 AU from the Sun). Those asteroids are often referred to as main belt asteroids or objects (MBAs or MBOs). They are believed to be remnants of the proto-planetary disc prevented from accreting to a planet by gravitational perturbations from Jupiter. Large numbers of MBAs are grouped in so-called asteroid families. An asteroid family is a population

of asteroids sharing similar dynamical properties. Members of asteroid families are believed to be fragments of a larger parent body disrupted in a collisional break-up event (Marzari et al. 1999). The existence of asteroid families was first suggested by Kiyotsugu Hirayama (Hirayama 1918, 1928, 1933), who identified the first few asteroid families in 1918. Currently there are at least 64 statistically significant families (Zappalà et al. 1995, Morbidelli et al. 1995). Some of the most prominent asteroid families include the Flora, Eunomia, Koronis, Eos, and Themis families.

Family membership is usually derived based on statistical cluster analysis in the proper orbital elements (proper semimajor axis, proper eccentricity, proper inclination) phase space. The two most common methods include wavelet analysis method (WAM) and hierarchical clustering method (HCM) (Zappalà et al. 1995). In agglomerative hierarchical clustering, each data point is first considered as its own cluster, in next steps clusters are combined into higher-order clusters based on a measure of dissimilarity (usually using some metric, for example, a measure of distance between pairs of data). Wavelet analysis method relies upon examination of densities in some parameter space. Until recently asteroid families were only derived based on dynamical classification (using proper orbital elements). However, recent developments suggest that physical classification can also be very helpful. In some cases physical classification is the only way to separate intermixed families such as the Nysa-Polana family (Parker et al. 2008, Oszkiewicz et al. 2011). For example, Parker et al. (2008) showed that based on the Sloan Digital Sky Survey (SDSS) colors it was possible to split the Nysa-Polana group into two separate families. Figure 2 shows the separation of the Nysa-Polana family into two physically distinct groups. Most of the other asteroid families seem to be quite homogeneous when it comes to their physical properties, such as spectral properties (Mothé-Diniz et al. 2005), color indices (Ivezić et al. 2001, Jurić et al. 2002), and photometric parameters (Oszkiewicz et al. 2011).

Asteroids in the main belt are considered the main source of near-Earth asteroids (NEAs). NEAs are asteroids that have orbits that pass close to that of the Earth (perihelion distance $q < 1.3$ AU and aphelion distance $Q > 0.983$ AU). NEAs are usually discussed in the light of possible Earth impacts that may result in more catastrophic consequences than any other known natural disasters (Schulte et al. 2010). An asteroid impact is considered the most probable cause for the extinction of dinosaurs and many other species (Schulte et al. 2010).

Some NEAs are also of high interest as they could potentially be explored due to their low relative (to the Earth) velocity and as potential sources of valuable minerals, limited on our planet. NEAs are traditionally subdivided into Apollos ($a \geq 1.0$ AU; $q \leq 1.0167$ AU), Atens ($a < 1.0$ AU; $Q \geq 0.983$ AU) and Amors (1.0167 AU $< q \leq 1.3$ AU). Most Atens and all Apollos have orbits that cross the Earth's orbit, posing a possible impact threat to the Earth on their current orbits. Amors do not cross the Earth's orbit but their orbits may evolve into Earth-crossing orbits in the future. Among mechanisms bringing asteroids in the vicinity of the Earth are mean-motion and secular resonances (Morbidelli et al. 2002), encounters with planets (mainly Jupiter), collisions and the Yarkovsky effect (which affects asteroids below few meter size). It was shown that about 23% of NEAs come from the 3:1 resonance, 25% from the intermediate Mars-crossers population (asteroids crossing the orbit of Mars), 37% from the ν_6 secular resonance, 8% from the outer

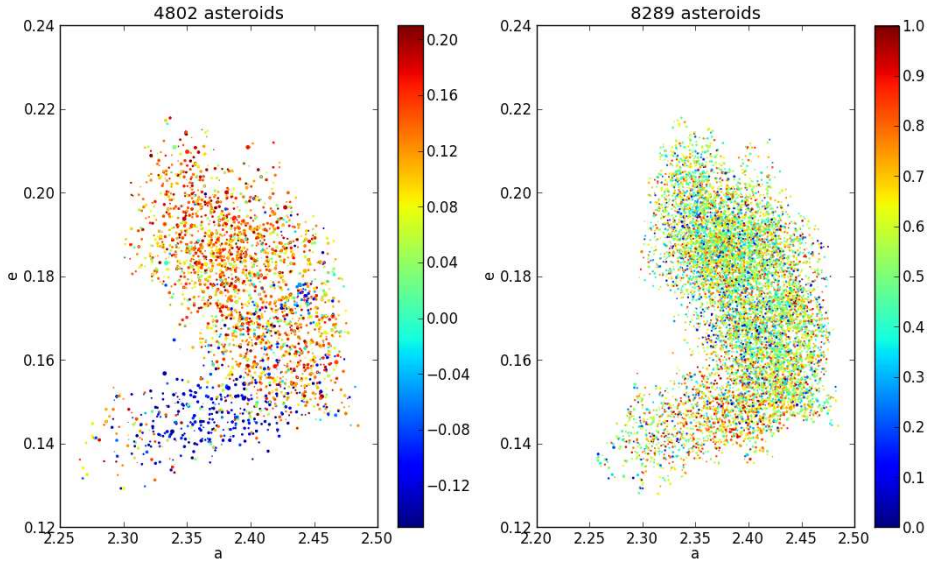

Figure 2: Distribution of proper elements for asteroids in the Nysa-Polana region, color coded according to the SDSS a^* color (left) and the photometric parameter G_{12} (right). The sizes of the points correspond to the errors in G_{12} and SDSS a^*.

main belt and about 6% from the transneptunian region (Bottke et al. 2000).

Transneptunian objects (TNOs) are orbiting the Sun at a greater average distance than Neptune. The first discovered TNO was the dwarf planet Pluto. TNOs are often divided into two large groups: the Kuiper-belt objects (roughly between 30 AU and 55 AU) and scattered-disc objects (roughly beyond 55 AU). The Kuiper-belt objects have also been subdivided into smaller dynamical groups: resonant KBOs (in mean motion resonance with Neptune, the most occupied ones are 3:2 – with prime example being Pluto – 2:1, and 5:2), Centaurs (non resonant objects whose perihelia lie inside the orbit of Neptune), classical KBOs (non-resonant, non-planet-crossing objects), scattered KBOs (non-classical, non-resonant objects whose perihelion distances remain outside the orbit of Neptune).

Some other asteroid populations worth mentioning are the so-called trojans. Trojans are asteroids locked in one of the two Lagrangian points of stability, L4 and L5 which lie approximately 60° ahead of and behind a larger body. Trojans of Jupiter, Neptune and Mars have been observed. Recently the first candidate for an Earth trojan (Connors et al. 2011) has been found.

1.3 Physical properties

Large-scale exploration of asteroid physical properties began in the modern era of automated sky surveys and space missions. Ground-based surveys such as for example the Small Main-Belt Asteroid Spectroscopic Survey (SMASS) and the SDSS (Ivezić et al. 2002) provided taxonomic classification for a few thousand and colors for some hundreds of thousands of asteroids. Several space-based missions, such as for example the Infrared Astronomical Satellite (IRAS) (Tedesco 1989), the Spitzer

Space Telescope (Trilling et al. 2007), and WISE (Wright et al. 2010) delivered colors and geometric albedos (and subsequently sizes) for some hundreds of thousands of asteroids. The distribution of asteroid proper elements color-coded with geometric albedo extracted from the WISE mission is shown in Fig. 3. The visible clumps of distinct albedo correspond to asteroid families. Much more: masses, sizes, spins, shapes, and classifications are expected from the upcoming Gaia mission (Mignard et al. 2007).

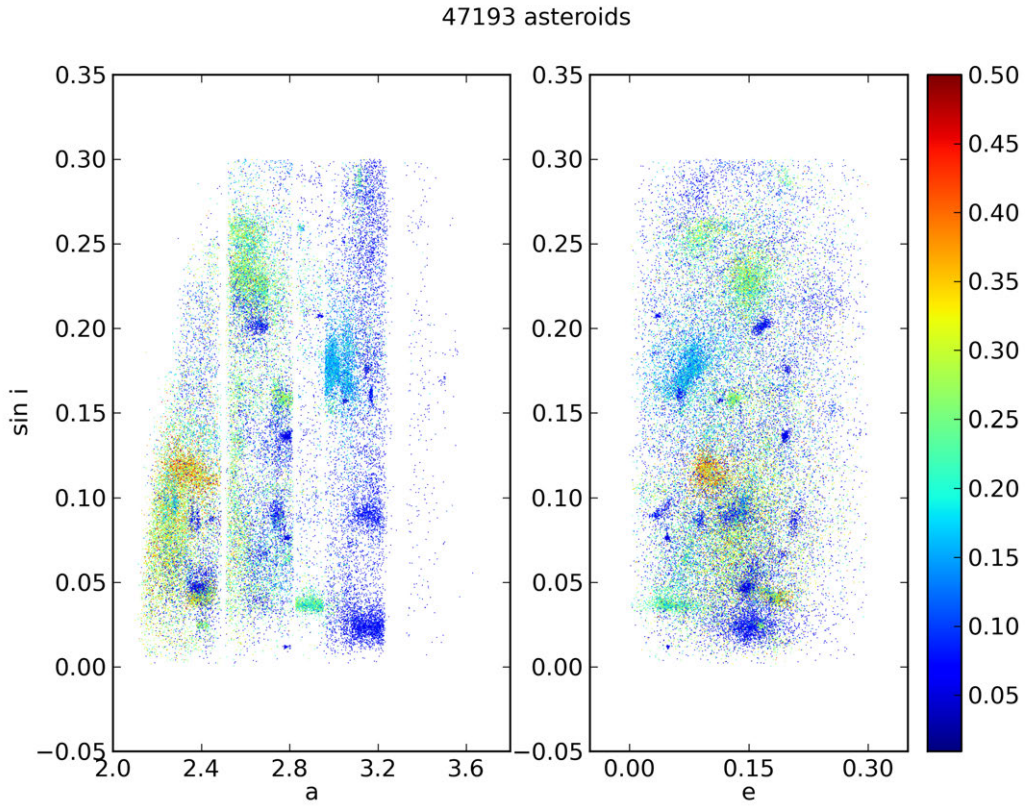

Figure 3: Distribution of asteroid proper elements color-coded with geometric albedo (extracted from WISE), for all asteroids with albedo below 0.5 (there are less than 300 asteroids in the dataset with higher albedo).

Asteroid soil and meteorite samples are also available. Tiny regolith particles from asteroid Itokawa were brought to the earth by the Hayabusa mission (Tsuchiyama et al. 2011). Some meteorites have been reported linking to particular asteroids (for example (Binzel & Xu 1993) or (Burbine et al. 2002)). However, to what extent are they representative of asteroids is still under question (Farinella et al. 1993). Beside Earth contamination, processes such as collisional metamorphosis and space weathering could have affected the meteorite samples traveling through space making them less representative of their origin.

About a dozen asteroids have been visited by spacecraft: (951) Gaspra (1991, Galileo), (243) Ida/Dactyl (1993, Galileo), (254) Mathilde (1997, NEAR), (9969) Braille (1999, Deep Space 1), (443) Eros (2000, NEAR), (5535) Annefrank (2002, Stardust), (25143) Itokawa (2005, Hayabusa), (2867) Steins (2008, Rosetta), (21) Lutetia (2010, Rosetta), (4) Vesta (2011, Dawn). These visits and flybys have re-

vealed a variety of surfaces, structures, and geological processes on asteroids. The first asteroid satellite (named Dactyl) was discovered by the spacecraft Galileo visiting asteroid (243) Ida on its way to Jupiter. Varying asteroid structures from monolithic through fractured to so-called gravitational aggregates were revealed together with processes such as impact communition, impact melting, formation of agglutinates, solar-wind sputtering, impact vaporization, impact-vapor condensation, and shock and thermal welding of grains, thermal fatigue or surface refreshing.

For most of the asteroids, however, we still need to rely on the observations made from Earth-based observatories. Asteroid spins and shapes are mostly derived based on lightcurve inversion techniques (some other are derived based on radar observations or satellite fly-bys). Masses are determined based on observations of an asteroid's gravitational effect on another body. Diameters are mostly derived based on albedo and absolute magnitudes.

The main physical parameters discussed in this thesis are the absolute magnitude, the photometric parameter(s), and the albedo. The absolute magnitude for an asteroid is defined as the apparent V-band magnitude that the object would have if it were 1 AU from both the Sun and the observer and at zero solar phase angle. The absolute magnitude relates directly to asteroid size and is useful in computing the geometric and Bond albedos. The geometric albedo of an object is the ratio of its actual brightness at zero phase angle to that of an idealized flat, fully reflecting, diffusively scattering (radiation is reflected isotropically with no memory of the location of the incident light source) Lambertian disk with the same cross section. The Bond albedo (A) is related to the geometric albedo (p) by $A = pq$ where q is the phase integral:

$$q = 2 \int_0^{\pi} \frac{I(\alpha)}{I(0)} \sin \alpha \, d\alpha, \tag{2}$$

where $I(\alpha)$ is the directional scattered flux (averaged over all wavelengths and azimuthal angles) and α is the phase angle. The diameter of an asteroid (D) can be computed using:

$$D = \frac{1329}{\sqrt{p}} 10^{-0.2H}. \tag{3}$$

Albedos are available for a few thousand asteroids from missions such as IRAS (Tedesco 1989), the Spitzer Space Telescope (Trilling et al. 2007), and WISE (Wright et al. 2010). Asteroid albedos vary depending on taxonomic type (Thomas et al. 2011). Also, asteroid color indices are shown to correspond to asteroid taxonomy. Based on SDSS data, Ivezić et al. (2002) showed that asteroid families tend to have similar optical colors which could indicate chemical composition. Surface color homogeneity of asteroids in families supports the idea of common origin of those groups. This has also been confirmed by the discovery of photometric parameter homogeneity in asteroid families (Oszkiewicz et al. 2011). Photometric parameters of the phase curves relate to the physical properties of an asteroid's surface, such as porosity, roughness, and grain size distribution. Steep phase curves are characteristic of bodies with an exposed regolith. The steepness of the phase curve also relates

to asteroid composition (Oszkiewicz et al. 2012). Most asteroids seem quite homogeneous (Degewij et al. 1979) when it comes to their physical properties over the surface such as for example aldebo, color indices and spectra. In particular, most asteroids display no or very small (comparable with the noise level) variations in spectrum or albedo during their rotational cycles (Degewij et al. 1979). Some slight color variations with rotation have been noted for several large asteroids. Asteroid (4) Vesta shows variations in both spectra and polarization with rotation. Those variations were attributed to changes of albedo with rotation (Degewij et al. 1979).

1.4 Gaia mission and asteroids

Gaia is an astrometric space mission (ESA) and a successor to the Hipparcos mission (Perryman et al. 1997). It will be launched in 2013, and will operate around the Sun-Earth L_2 Lagrangian point. Gaia will deliver a catalogue of approximately one billion stars down to magnitude 20. It is going to obtain astrometric observations, determine the positions, distances, and proper motions of stars (Perryman et al. 1997) with an accuracy of about 20 µas at 15 mag, and 200 µas at 20 mag, spectrophotometric observations, and radial velocity observations spanning over 5 years. The main objective of Gaia is to create an extremely precise three-dimensional map of stars throughout the Milky Way galaxy (and beyond), and map their motions in order to provide insights to the origin and evolution of the Galaxy. The spectrophotometric observations performed by Gaia will provide luminosity, effective temperature, gravity and elemental composition of each star observed. However, a large number of other objects like quasars, galaxies, extrasolar planets and Solar System bodies will be observed as well (Mignard et al. 2007).

Assuming the limiting magnitude $V = 20$ mag and estimating the completeness level of the current sample of asteroids, it has been evaluated that Gaia will observe about 3×10^5 asteroids (Mignard et al. 2007). Most of those objects will be known by the time of Gaia launch, but there exists a small possibility for the satellite to discover asteroids with orbits lying in a region that is difficult to observe for Earth-based telescopes (heliocentric distances less than 1 AU) (Mignard 2002). The limiting solar elongation for Gaia is 45°, and for ground-based surveys it is around 60°. The Gaia spacecraft will therefore be ideally situated to probe the asteroid blind spot between the Sun and the Earth and discovering so-called inner Earth objects (orbiting the Sun between Earth's orbit and the Sun). Gaia will scan solar elongations from approximately 45° to around 140°.

The Gaia data-processing pipeline will deliver the physical and dynamical properties of the observed small Solar System objects (SSOs) (Mignard et al. 2007). It will obtain sizes of about 1000 objects (currently well known for about 500 objects), masses of about 150 (currently well known for about 40 objects), spin and shapes of more than 10000 objects (currently well known for about 200 objects), derive improved orbits for many more and devise a Gaia-based asteroid taxonomy (Mignard et al. 2007). The final results will also include direct measurements of tiny radiative effects on small bodies, in particular of the Yarkovsky effect acting on near-Earth objects, and the measurement of tiny relativistic effects on the motion of some of these bodies. An impressive improvement of orbital accuracy for know objects will also become possible (by at least two orders of magnitude). Additionally improved

stellar catalogues will lead to improvements in differential astrometry and photometry. Exact impact of the Gaia mission on Solar System science is hard to predict, but it will definitely be a major science inspiration for the upcoming years if not decades.

1.5 Outline of this work

This thesis presents new findings on asteroid dynamical and physical properties based on Markov-chain Monte Carlo (MCMC) methods. Chapter 2 describes the inverse problem theory and numerical methods used in deriving dynamical and physical properties of asteroids. A particular attention is put to the Metropolis-Hastings algorithm which is then applied in deriving asteroid orbits, computing Earth-impact probabilities and asteroid photometric properties. Chapter 2 also describes the advantages and disadvantages of the methods used. Chapter 3 provides an introduction to the asteroid-orbit-computation problem and outlines the new MCMC ranging orbit-computation technique applicable to asteroids having short observational time intervals and/or small numbers of observations. Orbits obtained from the method can then be used in a variety of other dynamical problems such as for example computing collision probabilities. Chapter 4 gives an overview of phase functions and methods for deriving asteroid photometric properties (absolute magnitudes and photometric parameter(s)). A short introduction to obtaining phase curve data is also given. Chapter 5 describes importance of this thesis for the Gaia mission. Chapter 6 describes the main scientific findings of this work and provides a summary of the attached articles.

2 Inverse problems theory

2.1 Bayesian solution to inverse problems

Inverse problems arise in many branches of today's science, and particularly astronomy. As emphasized by Brown (1995) *"Astronomy is by its very essence an inverse problem and can be regarded as the ultimate one, both in scale and difficulty, in the sense that astronomers are attempting to determine the structure of the universe from data arriving at one point in it"*. This thesis deals with much smaller scales than the Universe, nevertheless the key properties of our observations remain the same. That is most of our data are remotely sensed and the whole process of data collection is restricted to passive observation (except few special cases where laboratory experiments are possible). Usually the properties of observed objects cannot directly be observed/measured and have to be inferred from the observational data.

In general, in inverse problems, the values of model parameter(s) must be inferred from the observed data (Tarantola 2005). The physical relation between the model parameters and the data can be expressed as:

$$\psi = f(\mathbf{P}) + \epsilon, \tag{4}$$

where ψ are the observations or measurements, $\mathbf{P}$ are the unknown model parameters, f is a function returning theoretical computed values of observations for a given $\mathbf{P}$, and ϵ denotes the observational errors. Because each observation can contain systematic and random errors, our measurement is just one possible outcome of an observation. Observations can therefore be treated as random variables (Lehtinen 2010, Menke 1989). What follows is that the model parameters must also be treated as random variables following an unknown distributions of our interest. Many inverse problems concern ill-posed models, where the solution to the problem is not unique, and there might exist a large (or even infinite) number of solutions. Inverse problems can be defined using the Bayesian formulation (Tarantola 2005). A Bayesian solution to an inverse problem is expressed as a conditional a posteriori probability density function (p.d.f.) which combines a priori information on model parameters and the physical relation between the data and model parameters in the form of a likelihood function:

$$p_p(\mathbf{P}) = C \quad \underbrace{p_{pr}(\mathbf{P})}_{\text{a priori}} \quad \underbrace{p_\epsilon(\Delta\psi(\mathbf{P}))}_{\text{likelihood}},$$

where $\mathbf{P}$ denotes the unknown model parameters, $p_\epsilon(\Delta\psi(\mathbf{P}))$ is the likelihood function, dependent on observed-minus-computed residuals $\Delta\psi$. The likelihood function is often assumed to be Gaussian, i.e., $p_\epsilon \propto \exp\left(-\frac{1}{2}\chi^2(\mathbf{P})\right)$, where $\chi^2(\mathbf{P}) = \Delta\psi^T(\mathbf{P})\Lambda^{-1}\Delta\psi(\mathbf{P})$, Λ is the covariance matrix for the observational errors, $\Delta\psi(\mathbf{P})$ is a vector of observed-minus-computed residuals. The normalization constant C is computed as

$$C = \left(\int p(\mathbf{P}, \psi) d\mathbf{P}\right)^{-1}.$$

The physical relation between the data and the model parameters is present in the observed-minus-computed residuals $\Delta\psi$, as is needed to obtain the theoretical

computed values of the data given some specific values of model parameters. That relation solves the so-called forward (prediction) problem of computing/predicting the data given some particular values of the parameters. The a priori p.d.f. $p_{pr}(\mathbf{P})$ can be chosen for example based on previous knowledge on the model parameters or some physical constrains. One convenient choices for the a priori p.d.f. is Jeffreys' a priori p_J, which has the property of keeping the a posteriori p.d.f. invariant in transformations between different model parameters sets (Jeffreys 1946):

$$\begin{aligned} p_J(\mathbf{P}) &\propto \sqrt{\det \Sigma^{-1}(\mathbf{P})}, \\ \Sigma^{-1} &= \Phi(\mathbf{P})^T \Lambda^{-1} \Phi(\mathbf{P}), \end{aligned}$$

where Σ^{-1} is the information matrix evaluated for local parameters $\mathbf{P}$, Λ is the covariance matrix for the measurement errors, and Φ is the partial derivatives matrix, containing derivatives of observed quantities with respect to model parameters. Jeffreys' a priori belongs to the so-called weakly informative priors (not supplying substantial information) as opposed to the so-called informative priors (supplying crucial information that might come from previous knowledge such as for example earlier data analysis, expert knowledge or some physical constraints) and so-called non-informative priors (when we know very little and require only the data to drive the a posteriori p.d.f.).

The total a posteriori p.d.f. can then be expressed as:

$$p_p(\mathbf{P}) = C\, p_{pr_i}(\mathbf{P}) \sqrt{\det \Sigma^{-1}(\mathbf{P})}\, \exp\left(-\frac{1}{2}\chi^2(\mathbf{P})\right), \tag{5}$$

where $p_{pr_i}(\mathbf{P})$ is an informative a priori. An analytical solution to the above equation is however rarely found. In fact, only a limited number of simple mathematical problems have an analytical a posteriori p.d.f. . Thus, Monte Carlo algorithms are often used to obtain a numerical solution (more in Sec. 2.3). In this kind of treatment, instead of a unique solution, we arrive at a large number of discrete estimates, each of which is probable and yields observed-minus-computed residuals which lie within the observational error ϵ.

2.2 Advantages of Bayesian method

Bayesian methods and the deterministic (classical) approach to inverse problems both have some advantages and disadvantages. Bayesian methods allow for the incorporation of past information on the parameters in the a priori p.d.f.s. and are proven to deal well with nonlinear or poorly constrained cases. Other advantages of the probabilistic approach include the possibility of dealing with noise assumptions other than Gaussian. Bayesian analysis can estimate any functions of the model parameters directly from the distributions of the parameters.

In the classical approach usually a single-point estimate together with some error is computed. In the probabilistic approach this is equivalent to the situation when the solved p.d.f. is summarized with some parameters describing the distribution obtained (Virtanen 2005) (for example, the maximum a posteriori point, moments,

confidence intervals etc.). In many cases the two approaches are equivalent and produce the same results, for example in the cases of well-constrained p.d.f.s.

The situation changes for problems with complicated and not-well-constrained p.d.f.s. In such cases the classical approach may fail (for example resulting in a non-physical solution or no solution at all). There might be too few or too noisy observations for a classical approach to return meaningful results. The solution might not be unique. The Bayesian approach can deal with some of these problems.

There are also disadvantages, for example there is no rule for choosing the a priori p.d.f. and it has to be possible to translate the past information on parameters into mathematical formulations. In some cases, Bayesian analysis can produce posterior distributions that are influenced by the priors and the methods are sometimes computationally expensive. A complete discussion of statistical versus deterministic inversion methods can be found in for example Virtanen (2005).

2.3 Markov-chain Monte Carlo technique for solving inverse problems

Markov-chain Monte Carlo (MCMC) algorithms are a set of methods used to approximate a desired distribution of unknown model parameters. The main idea relies upon constructing a Markov chain (chain of consequent samples that may be correlated and having so-called Markov property — the next state depends only on the current state and not on the past states), which converges to the desired target distribution as it reaches the equilibrium (stationary) state starting from any point in the parameter phase space. When reaching the stationary distribution, the current state no longer depends on the initial starting point. After this first, so-called burn-in phase the chain will eventually end up sampling the same parameter subspace, no matter what the starting point was. The resulting sample solutions are identically (each state has the same probability distribution as the others), but not independently, distributed according to the target density. A proof of convergence can be found for example in Roberts & Rosenthal (2004).

The key characteristic of Markov chains ensuring stationarity is reversibility on a phase space with respect to a probability distribution. If $\mathbf{X}$ is the phase space and π the probability distribution, P is a transition probability, then

$$\pi(dx)P(x,dy) = \pi(dy)P(y,dx) \text{ , where } x, y \in \mathbf{X} \tag{6}$$

is the reversibility condition (also called detailed balance condition or detailed balance criteria). If a Markov chain is reversible with respect to π, then π is stationary for this chain. This can be shown mathematically as (Roberts & Rosenthal 2004):

$$\int_{x\in X} \pi(dx)P(x,dy) = \int_{x\in X} \pi(dy)P(y,dx) = \pi(dy)\int_{x\in X} P(y,dx) = \pi(dy). \tag{7}$$

Other important characteristics are reducibility, periodicity, recurrence and ergodicity. A Markov chain is said to be irreducible if it is possible to get to any state from any other state, i.e., there is non-zero probability that a chain started from any point in the phase space will transit to any other point in that state at some time

(Roberts & Rosenthal 2004). A chain is called periodic when any return to state i must occur in multiples of k time steps. Recurrence (Roberts & Rosenthal 2004) refers to the idea of a transient state, that is a state with a non-zero probability of returning to that state, given that we have started the chain in that state. The markov chain is called ergodic if it is aperiodic, irreducible, and positive recurrent.

MCMC gained popularity in 1990s, and currently various algorithms relying on Markov chains exist. The most commonly used are the Metropolis-Hasting algorithm (based on the proposal probability density and acceptance criteria), Gibbs sampling (requires knowledge of conditional distributions), slice sampling (based on uniform sampling of regions under p.d.f.), multiple-try Metropolis (used in problems with large number of dimensions). More sophisticated algorithms include for example successive over-relaxation, hybrid Monte Carlo, Langevin MCMC, and reversible jump. The Metropolis-Hastings (M-H) algorithm is the most widely used. M-H was invented by Nicholas Metropolis in 1950s and then further developed by W. Keith Hastings. The M-H algorithm relies upon drawing samples from the so-called proposal distribution $P_c(\mathbf{P})$ (other names: candidate, transit, jump distribution), which can be arbitrary and depends only on the current (last accepted) sample in a chain. However, the closer the proposal distribution is to the unknown target parameter distribution, the faster the algorithm converges to the stationary state. In each iteration new candidate parameters $\mathbf{P}'$ are proposed to be added to the chain with the help of the transition probability. Those candidate parameters are next accepted or rejected from the chain based on the acceptance criteria:

$$\begin{array}{ll} \text{If } a \geq 1, & \text{then } \mathbf{P}_{t+1} = \mathbf{P}'. \\ \text{If } a < 1, & \text{then } \begin{cases} \mathbf{P}_{t+1} = \mathbf{P}', \text{ with probability } a, \\ \mathbf{P}_{t+1} = \mathbf{P}_t, \text{ with probability } 1 - a, \end{cases} \end{array} \tag{8}$$

where

$$a = \frac{p_p(\mathbf{P}')}{p_p(\mathbf{P}_t)} \frac{P_c(\mathbf{P}_t \mid \mathbf{P}')}{P_c(\mathbf{P}' \mid \mathbf{P}_t)} \tag{9}$$

and $\mathbf{P}_t$ denotes last accepted parameters and $P_c(\mathbf{P}_t \mid \mathbf{P}')$, $P_c(\mathbf{P}' \mid \mathbf{P}_t)$ are transition probabilities from parameters $\mathbf{P}'$ to $\mathbf{P}_t$ and in the reverse direction, respectively. The posterior probabilities $p_p(\mathbf{P}')$, $p_p(\mathbf{P}_t)$ are defined as in Eq. 5. If a symmetric proposal density is used (such as the Gaussian density) the acceptance coefficient a is simplified to:

$$a = \frac{p_p(\mathbf{P}')}{p_p(\mathbf{P}_t)}. \tag{10}$$

In practice, the candidate parameters are always accepted if they produce a better fit to the data than the last accepted ones in the chain. The trial parameters are sometimes accepted if the fit is worse than the last accepted solution. If we have a choice between jumping to 'bad fit' and a 'very bad fit', the 'bad fit' jump is more probable. If candidate parameters are rejected, then the last accepted solution is repeated in a chain.

If the size of the proposal distribution differs significantly from the size of the target distribution the chain might mix (sample the stationary distribution) slowly.

If the selected proposal distribution is too small (small step size), the chain acceptance ratio will be high, but the chain will not sample the entire phase space in a reasonable time. If the size is too large, the acceptance ratio will be very small and it might take a long time to reach the steady-state distribution. Both of those problems may be identified through convergence diagnostics.

Based on MCMC sampling of a one-dimensional Gaussian distribution, it was shown that an ideal acceptance ratio is around 50%. This theoretical value decreases with increasing dimensions and complexity of considered problems. For very complex and multi-dimensional problems the acceptance ratio may be low.

The initial phase of the sampling, the burn-in phase relates to the fact that the chain has to take some time before it starts sampling the target distribution (if started in a random spot in the phase space). The samples drawn in the burn-in period have to be removed from the resulting chain(s). To shorten the burn-in phase a non-random point (based on expert knowledge, previous results, etc.) in the parameter phase-space can be selected as a starting point. To find a reasonable proposal density and avoid manual tuning, some 'pre-runs' to actual MCMC can be run. Haario et al. (1999) recommend an adaptive MCMC sampling method, which does not require pre-runs and tunes the proposal during the sampling, not affecting the Markov property of the chains.

2.3.1 Convergence diagnostics

Convergence diagnostics is a critical part of all MCMC algorithms and includes testing the equilibrium of the sampled distribution. The M-H algorithm converges to the stationary distribution (equilibrium) given large enough number of samples Roberts & Rosenthal (2004). In practice, we are limited by the number of iterations and time, and thus we need to perform additional checks on obtained samples to make sure that the convergence was reached. Typical approaches include monitoring "stability" of a single long chain and/or number of shorter chains.

The most common convergence testing methods found in the literature include history plots (plot of sampled parameters against number of iterations, also called trace plots), kernel plots, and correlograms. Gelman & Rubin (1992) proposed a statistical indicator, so called "shrink factor" R, involving the average within-chain variances, and the variance between the means obtained from m parallel over-dispersed chains. Once convergence is reached, the between and within variances should be approximately the same. Raftery & Lewis (1992) proposed a method for estimating the total numbers of iterations of the burn-in period based on chosen quantile, required accuracy, required probability to obtain specified accuracy and convergence tolerance. Gewke (1992) indicated the use of methods from spectral analysis in testing for convergence. The method is graphical and based on running means of parameters and can be used in, for example, checking for the length of the burn-in period. Zellner & Min (1995) proposed a criterion based on conditional probability to test for convergence to a correct distribution. Extensive reviews of MCMC convergence diagnostics methods can be found in, for example, Mengersen et al. (1999) and Cowles & Carlin (1996).

2.3.2 Markov-chain Monte Carlo vs. Monte Carlo

Monte Carlo (MC) is based on random sampling and MCMC on Markov chains. The goal of MC is to sample a multidimensional region uniformly and then summarize the posterior using the resulting samples and normalized a posteriori p.d.f. as weights. The goal of MCMC is not to sample uniformly, but to visit a point x in the parameter phase space $\mathbf{X}$ with certain probability. The resulting samples have equal weights. Additionally, some of the samples might be repeated in a chain.

One of the advantages of MCMC is that the normalization constant in Eq. 5 does not have to be known or computable to obtain information about the resulting a posteriori p.d.f. and model parameters. Another advantage is that in MCMC there are proportionally more samples in high a posteriori p.d.f. regions than in low a posteriori p.d.f. regions. This is especially important for high-dimensional, computationally expensive problems, where MCMC often proves to be more efficient. Furthermore, in MC, one usually has to specify the domain of possible parameter inputs, whereas in MCMC, such a specification is not necessary, as MCMC can move freely through the parameter phase space. Additionally MCMC comes with vast convergence diagnostics methods.

3 Application: Asteroid orbits

3.1 Brief history of orbit computation

From the beginning of the 19th century, when the first asteroid was found, asteroid discovery methods have dramatically improved; particularly encouraged by the era of space exploration. With the technical improvements and development of observation techniques came a corresponding improvement in orbit computation methods. An increasing number of space and ground-based surveys have resulted in an exponential growth in the number of discoveries. The need to search for new asteroids has recently been strengthened by the awareness of the possibility of Earth impacts by these objects. Asteroid surveys discover a great number of asteroids, for which an initial orbit computation is needed in light of the collision probability problem. Conventional, deterministic methods of preliminary orbit computation often come with vast restrictions for their usage (for example, number of requested observations), are based on many assumptions (for example, observations made in perihelia) and can still simply fail to produce results for the object. Therefore a need for novel, reliable and fast methods has become apparent.

Orbit determination has a long history, starting with the discovery of the planets and the following attempts to predict their motions (Dubyago 1961). The first insight into the orbit computation problem appeared with Kepler's laws describing planetary motion and later Newton's law of gravitation. However, the first correct orbit determination and subsequent recovery attempts were made much later.

The first analytical comet orbit computation method was proposed in 1744 by Euler (Dubyago 1961). The method required a minimum of four observations. At the end of the 18^{th} century, Lagrange proposed an orbit determination method relying on the condition that the three observations of the position of the object have to be situated in one plane with the Sun (Dubyago 1961). Lagrange's ideas were later further developed in the method proposed by Gauss (Dubyago 1961). Gauss's method became invaluable to the recovery of Ceres, which would have been essentially lost otherwise. It is worth mentioning that Gauss also introduced the method of least squares, which allowed the first statistical treatment of orbit computation. Among other methods known at the time was also the one by Olbers solving the equations for geocentric distances by means of trials (Dubyago 1961). Various new methods were developed afterwards, most of them based on previously known principles.

In the more recent decades novel methods were also developed. For example methods of Marsden (Marsden 1985, 1991), GEM (Gauss-Encke-Merton), MVC (Moulton-Väisälä-Cunningham). New, more reliable methods of orbit computation have also been developed for asteroids with small numbers of observations and short observational arcs. Two-observation methods by Väisälä (1939) and Orlov (1939) follow the assumption that new asteroids are often discovered near perihelion. These methods also require the specification of either geocentric distance or eccentricity.

Use of maximum likelihood and Bayesian solution-like methods was suggested in the beginning of the 1990s (Kristensen 1992) by various authors (Muinonen & Bowell 1993, McNaught 1999, Tholen & Whiteley 1998, Chodas & Yeomans 1996).

McNaught (1999) and Tholen & Whiteley (1998) developed methods of deriving a set of possible orbits from two observations. A large set of different methods were subsequently invented, most of which are based on Monte Carlo sampling of the orbital element phase space. However, the majority of these methods cannot be readily applied to asteroids having sparse observational data or short arcs as the linearization of the inversion is usually inappropriate. Nonlinear methods for orbital uncertainty were considered by, for example, Milani (1999), Milani et al. (2004) and Chesley & Milani (1999) towards the end of 1990s and start of 2000s. In late 1990s, Virtanen et al. (2001) and Muinonen et al. (2001) presented the method of statistical ranging, for the treatment of poorly observed asteroids. Their method follows from the statistical inversion theory outlined by Lehtinen (2010) and Menke (1989). In 2009, Oszkiewicz et al. (2009) further developed the ranging method introducing Markov chains. A variation of the MCMC ranging method was later utilized in linking of geosynchronous orbital debris tracks (Schneider 2011).

3.2 Orbital inverse problem

Orbital inversion comprises a problem of computing orbital parameters from astrometric observations. An orbit can be described using different parameter sets $\mathbf{P}$ (for example, Keplerian, Cartesian, or Delaunay orbital elements). The most commonly used are the Keplerian elements:

- semi-major axis a - half the major axis of an orbit's ellipse,
- eccentricity e - describes the ellipticity of the orbit and can be expressed as $e = \sqrt{1 - \frac{b^2}{a^2}}$,
- inclination i - angle between the orbital plane and the ecliptic plane,
- argument of pericenter ω - the angle from the ascending node to the object, measured in the orbital plane,
- longitude of the ascending node Ω - angle between the line of nodes and the reference direction (equinox point),
- mean anomaly M - angle describing position of an object in the orbital plane.

The inclination i, argument of the pericenter ω and the longitude of the ascending node Ω are illustrated in Fig. 4. The reference plane is the plane of the ecliptic and the zero point of longitude is the equinox point. All the different orbital parameter sets are equivalent and can be transformed into each other. For example, the transformation between the Keplerian elements and the Cartesian elements can be found in Vallado & McClain (2001). Given the orbital parameters, an asteroid's future position can be predicted.

To find the orbital parameters, first the orbital inverse problem has to be solved for the given astrometric positions $\psi = (\alpha_1, \delta_1; ...; \alpha_N, \delta_N)^T$. The solution to the orbital inverse problem can be written as an a posteriori probability, using Bayes' theorem:

$$p_p(\mathbf{P}) = C\, p_{pr}(\mathbf{P})\, p(\psi \mid \mathbf{P}) = C\, p_{pr}(\mathbf{P})\, p_\epsilon(\Delta\Psi(\mathbf{P})), \tag{11}$$

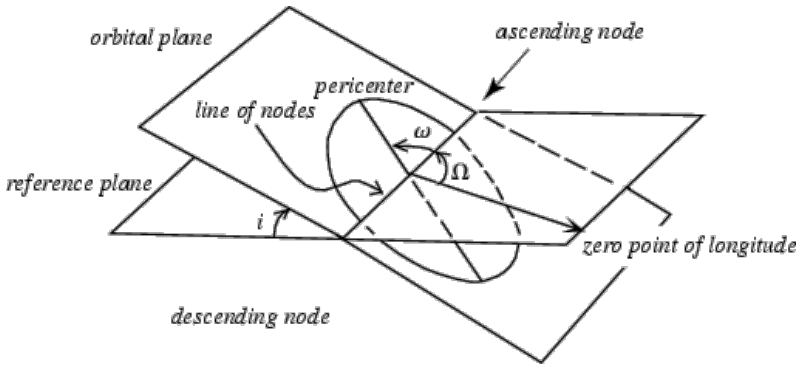

Figure 4: Illustration of Keplerian elements. Source: `http://scienceworld.wolfram.com`

where:

- $p_{pr}(\mathbf{P})$ is the a priori probability density function,
- $p_\epsilon(\Delta\Psi(\mathbf{P}))$ is the observational error p.d.f. (usually being assumed Gaussian), evaluated for the observed-minus-computed (O-C) residuals $\Delta\Psi(\mathbf{P})$,
- $C = (\int p(\mathbf{P}, \psi) d\mathbf{P})^{-1}$ is a normalization constant,
- $\mathbf{P}$ denotes osculating elements $(x, y, z, \dot{x}, \dot{y}, \dot{z})^T$ or $(a, e, i, \Omega, \omega, M_0)^T$ at epoch t_0,
- ψ denotes a set of astrometric observations, consisting of R.A., Dec. pairs: $\psi = (\alpha_1, \delta_1; ...; \alpha_N, \delta_N)^T$.

Extensive review of the orbital inverse problem can be found in Virtanen (2005), Virtanen et al. (2001), and Muinonen et al. (2001). To solve the inverse problem expressed in terms of Eq. 11, Markov-chain Monte-Carlo methods are used. The resulting distributions of orbital elements can then be summarized in terms of maximum-likelihood orbital parameters and corresponding non-symmetric error estimates.

3.3 MCMC ranging technique

MCMC ranging is a non-linear orbital sampling technique originating from the statistical ranging technique which was developed at the University of Helsinki (Virtanen et al. 2001, Muinonen et al. 2001). In MCMC ranging, a pair of spherical coordinates, that is R.A.s, Dec.s, and topocentric distances (ranges) $\mathbf{Q} = (\rho_A, \alpha_A, \delta_A, \rho_B, \alpha_B, \delta_B)$ at two chosen observation dates A and B are used as inversion parameters. Next with the help of a multivariate Gaussian proposal distribution $p_t(\mathbf{Q}', \mathbf{Q_t})$ centered on the last accepted spherical coordinates $\mathbf{Q_t}$, new candidate spherical coordinates $\mathbf{Q}'$ are iteratively sampled producing candidate orbital solutions. Each candidate orbit is accepted or rejected based on the M-H acceptance criteria (Eq. 8), where the acceptance coefficient is defined as:

$$a = \frac{p_p(\mathbf{P}')}{p_p(\mathbf{P}_t)} \frac{\mid J_t \mid}{\mid J' \mid}. \tag{12}$$

J' and J_t are the Jacobians from topocentric coordinates to orbital parameters for the candidate and the last accepted sample, respectively. $\mathbf{P}'$ and $\mathbf{P}_t$ denote a set of orbital elements, and p_p is the a posteriori probability density function. The generation of sample orbits is repeated until the initial starting point is "forgotten" and no longer influences the chain (the burn-in phase). The samples obtained in the burn-in period are then removed from the final set. To shorten the burn-in phase the starting orbit is picked to produce a good fit to the data (for example least square orbit or MC orbit with a high p.d.f. value). In order to have better chances of recovering multi-modal distributions, it is recommended to run more than one Markov chain. Convergence diagnostics are performed after the sampling is completed to ensure that the steady-state distribution has been reached. To avoid manual tuning of the proposal density covariance, we have implemented an automated version of the MCMC algorithm. In the automated version, we first run 'pre-runs' (which could be considered as a burn-in phase) to the actual sampling, in order to find adequate proposal density. That proposal density is computed based on empirical samples from the pre-runs, and then updated in a manner similar to that described in Haario et al. (1999):

$$p_t^{(i)}(\mathbf{Q}' \mid \mathbf{Q_1}, ..., \mathbf{Q_n}) \quad \sim \quad N(\mathbf{Q_t}, c_d^2 \mathbf{R}^{(i)} + c_d^2 \epsilon \mathbf{I}), \tag{13}$$

where $\mathbf{R}^{(i)}$ is the empirical 6×6 covariance matrix obtained from pre-run i from n samples $\mathbf{Q_1}, ..., \mathbf{Q_n}$; $c_d^2 = 2.4/\sqrt{d}$ is a scaling factor that depends on the number of dimensions d in the problem, and in our case $d = 6$, $c_d \approx 0.97$. This scaling factor is optimized for M-H in the case of Gaussian targets and Gaussian proposals, for more details see Gelman et al. (1996).

Convergence diagnostic have to be performed after the sampling is completed to ensure that the steady-state distribution has been reached. Some additional checks, e.g., difference in the maximum likelihood value obtained or in the obtained covariance matrices between subsequent pre-runs can also be carried out. Convergence diagnostics were designed to detect possible problems with the convergence to the stationary distribution. It is recommended to proceed with caution and remember that convergence diagnostics may detect possible problems but cannot guarantee that the convergence has been reached.

3.4 Selected results

MCMC ranging was applied to a number of asteroids (paper I), and also used in Earth impact probability computation (paper V). As an example application of the method in Fig. 5 we plot the time evolution of the collision probability computed based on short observational time intervals (maximum 6 astrometric positions, 0.6 day observational time interval) for asteroid 2008 TC_3. Asteroid 2008 TC_3 impacted the Earth on October 7, 2008 and was the first asteroid to be observed and tracked before the impact. We used the MCMC method to obtain orbital element distributions for sets of observations with increasing observational time intervals (that is including subsequently from 2 to 6 observation in a set). The collision probability was then computed as:

$$P_c(\tau) = \int_0^R d\rho \;\; p_p(\rho, \tau), \tag{14}$$

where R is the radius of the planetary body and $p_p(\rho, \tau)$ is defined as:

$$p_p(\rho', \tau) = \int d\mathbf{P} p_p(\mathbf{P}) \delta_D(\rho' - \rho(\mathbf{P}, \tau)), \tag{15}$$

where $\rho(\mathbf{P}, \tau)$ is the minimum distance of the small body from the center of the large (planetary) body, δ_D is Dirac's function, $\mathbf{P}$ denotes orbital elements, p_p is the a posteriori p.d.f. of the orbital elements leading to a collision. τ is the impact interval (time interval when collisions are possible) – here we used an impact interval of Oct. 6-8, 2008. Including more than 6 observations led to collision probability essentially equal to 1.0. The MCMC ranging led to obtaining very high impact probabilities for the Earth-impactor 2008 TC_3 based on very few observations and a short time interval. The MCMC ranging method was further extended to sample the astrometric noise together with the orbital elements for cases where the observational uncertainty is not known or cannot be assumed. The extended method allows for treating each of the astrometric positions and relating uncertainties individually and could be utilized in deriving astrometric biases of individual observatories in the future. The improved method for collision-probability estimation could potentially become useful in monitoring the Earth-impact probabilities for new discoveries.

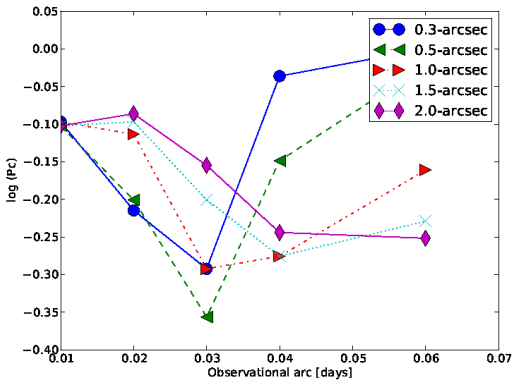

Figure 5: Time evolution of collision probabilities for 2008 TC_3 using MCMC ranging generated orbits with Jeffreys' a priori (n-body approach).

The MCMC ranging method was next implemented into OpenOrb package (paper II). The average computing time for a set of orbits using the MCMC ranging method on an ordinary dual core laptop is of the order of a few to a few tens of minutes. The computation of collision probability is a bit more computationally expensive and may take up to few days on an ordinary dual core laptop. This is however not a problem given the modern technology and the possibility to perform computation on small to medium computer clusters.

MCMC ranging is now also a part of the Gaia satellite data processing pipeline (more in Sec. 5). In Fig. 6 the MCMC ranging method was used to illustrate the influence of the extreme Gaia accuracy on the resulting orbital-elements distributions. We have selected a part of simulated observations (one transit data) for asteroid (4) Vesta to demonstrate the collapse of orbital-element distributions with increasing accuracy. As expected, the resulting orbital element p.d.f. collapses with improving accuracy of the astrometric positions. A variation of the MCMC ranging method

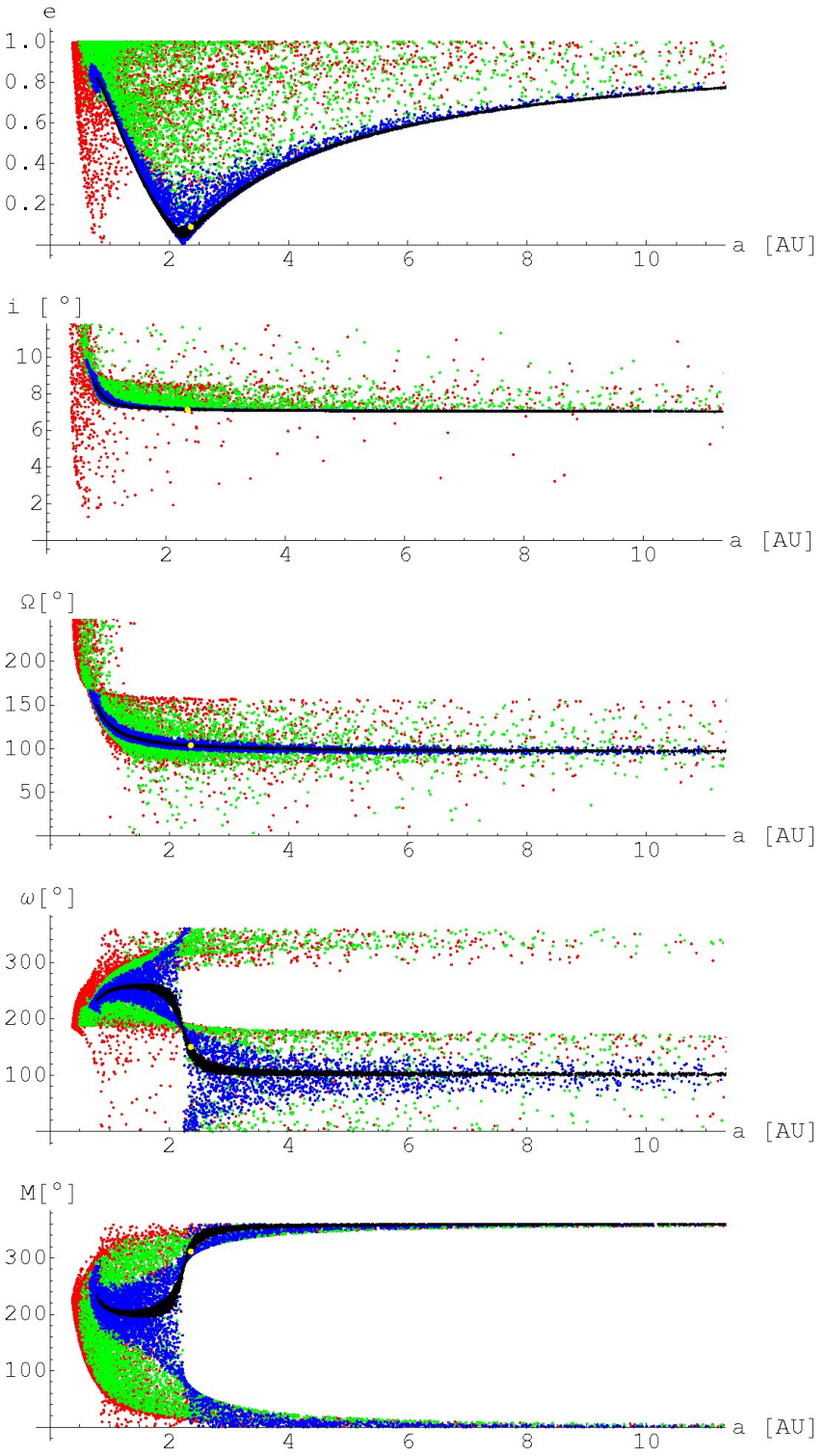

Figure 6: Sets of distributions each composed of 5000 possible orbit solutions for the main-belt object (4) Vesta obtained from simulated Gaia data for epoch 2010 August 5.0 TDT. The distributions were obtained using 4 observations from a single scan (observational time interval of 0.32 d). Different colors correspond to different assumptions about the observational error (red $\sigma_{\alpha,\delta} = 1.0$ arcsec, green $\sigma_{\alpha,\delta} = 0.1$ arcsec, blue $\sigma_{\alpha,\delta} = 0.01$ arcsec and black $\sigma_{\alpha,\delta} = 0.001$ arcsec). The yellow dot indicates the end-of-mission least-squares solution computed with the expected Gaia data astrometric accuracy.

was also utilized in the linking of geosynchronous orbital debris tracks (Schneider 2011).

4 Application: Asteroid photometric parameters

4.1 Obtaining phase curves from traditional observations

An asteroid's photometric parameters can be obtained from its phase curve. The phase curve describes an asteroid's apparent V magnitude at a distance of one astronomical unit from the Earth and the Sun (also called the reduced magnitude) as a function of solar phase angle α (the Sun-object-observer angle). The phase curve is mostly linear, except at small phase angles, where the so-called opposition effect occurs. At small phase angles, pores and pits introduce constructive interference in multiple scattering among regolith particles, which results in brightness enhancement due to coherent backscattering. Fitting a phase function to a set of reduced apparent magnitudes and corresponding phase angles results in the determination of the absolute magnitude and photometric parameter(s). The absolute magnitude is the reduced magnitude at zero solar phase angle, and the photometric parameter describes how steep the phase function is. The absolute magnitudes and photometric parameters are connected to a number of other parameters such as asteroid diameter, taxonomic type (Harris 1989, Lagerkvist & Magnusson 1990), geometric albedo, porosity, packing density, particle size distribution, and other physical properties of the surface. The phase curve can therefore be useful for characterizing an object's regolith or atmosphere. Given an asteroid's absolute magnitude, slope parameter(s), and observing geometry, the apparent magnitude can be calculated.

Obtaining a phase curve is both time-consuming and complicated. It requires the determination of an asteroid's rotational lightcurves over a wide range of solar phase angles (this typically means tens of observational nights spread over two months (Buchheim 2010)) with high photometric accuracy (typically ± 0.03). Moreover, the data reduction process requires linking of an asteroid's brightness between observing nights (linking of comparison stars or all sky photometry) (Buchheim 2010). Broad and dense phase-curve coverage is not easy to obtain. Ground-based observatories are limited in the possible phase-angle α ranges. Main-belt asteroids do not reach solar phase angles greater than 20–30°, Jupiter Trojans display an even smaller range of phase angles and Kuiper-belt objects are only observable at phase angles smaller than 2° (Buchheim 2010). Asteroids possible to observe at large phase angles are the NEAs during their close approaches to the Earth. Also satellites can provide observations at large phase angles. The brightness of an asteroid is affected by the phase angle, the asteroid's shape, rotation and distances to the Earth and the Sun. To obtain the phase curve influence of the shape and distances has to be removed. The influence of distances can be removed by using the reduced magnitude (Buchheim 2010):

$$V(\alpha) = V - 5\log(R_{Sun}R_{Earth}), \tag{16}$$

where $V(\alpha)$ is the reduced magnitude, and R_{Sun} and R_{Earth} are the asteroid-Sun and asteroid-Earth distances. Those distances change very slowly and can be assumed constant for a single observational night. The most significant changes in asteroid brightness during a single observation night come from irregular asteroid shape and its rotation. As the asteroid spins, it reflects different amounts of sunlight, to a large part due to different cross-sections exposed to the observer. To remove the influence of asteroid shape and rotation a single averaged measurement is reported

for each observing night. The asteroid lightcurve is obtained using the technique of differential photometry.

4.2 Calibration of Minor Planet Center photometry

We have made use of Minor Planet Center (MPC) photometry. MPC contains sparse photometry in various filters. Most of the photometric data are of low precision (generally rounded to 0.1 mag) and low accuracy (rms magnitude uncertainties of $\pm$ 0.3 mag are typical). The MPC data comprise photometric observations from many sources (each having different systematic and random errors, sometimes varying with time). The photometric data are very numerous: about 73,000,000 individual, largely independent brightness measurements exist. For most asteroids there exist hundreds of photometric measurements at different phase angles, geometries, and rotational states.

The photometric data from ten observatories have been calibrated using accurate broad-band photometry of asteroids observed in the course of the Sloan Digital Sky Survey (SDSS) (Ivezić et al. 2001). The SDSS data were converted to the V-band using equations derived by Rodgers et al. (2006). Most photometric observations have been reported in the V and R bands. The latter can be converted, with good accuracy using a mean $V - R$ color index for asteroids (of +0.4 mag). Other observations, such as those reported in Rodgers et al. (2006) cannot be so accurately converted to the V band without knowledge of asteroids' spectral characteristics. By adopting mean color index corrections, we have introduced small systematic errors into some of the photometric observations.

SDSS spans the brightness range 15 mag $< V <$ 21 mag. Therefore photometry of bright asteroids has not been calibrated as well as photometry of fainter objects. Also, observations made near the magnitude limit of telescopes will tend to be made when asteroids are brighter than average. In other words, asteroids exhibiting significant lightcurve variation might only be observed near maximum brightness. Because such observations also tend to be made when an asteroid is far from opposition (and near the largest observable phase angle), some of the phase curves we compute could be less steep than the true phase curves. We noted that photometric observations from a given observatory exhibited time-dependent shifts, sometimes exceeding 1 mag, that have probably arisen from programmatic changes in brightness calibration methods. After correction of this effect, we computed O-C magnitude residuals as functions of magnitude and rate of sky-plane motion. For each of the observations, we established relationships between the O-C magnitude residuals and the accurate V-band measurements (derived from SDSS). Then we applied corrections to the raw data to place them on the accurate V-band system. The procedure was iterated to convergence.

Hitherto, photometric observations have not been weighted. However, it is clear that different observers achieve different photometric accuracy, and that photometric accuracy declines as a telescope's limiting magnitude is approached. To compound the problem, it is known that small asteroids tend to be less spherical than large ones, and thus exhibit lightcurves having greater amplitudes. Using the large quantity of data we have in hand, it will be possible to separate the compo-

nents of the photometric error budget, and thereby assign a realistic weight to each observation. In this work we fitted phase functions to calibrated MPC photometry for about half a million asteroids.

4.3 Phase functions

A number of different mathematical formulations for asteroid phase curves have been developed. The H,G phase function was developed by Bowell et al. (1989) and adopted by the International Astronomical Union in 1985.

In the H,G magnitude phase function, the reduced apparent magnitudes can be obtained from:

$$\begin{aligned} 10^{-0.4V(\alpha)} &= a_1\Phi_1(\alpha) + a_2\Phi_2(\alpha) \\ &= 10^{-0.4H}\left[(1-G)\Phi_1(\alpha) + G\Phi_2(\alpha)\right], \end{aligned} \tag{17}$$

where α is the phase angle, and $V(\alpha)$ is the reduced magnitude. The basis functions Φ_1, Φ_2 are defined in terms of trigonometric functions:

$$\begin{aligned} \Phi_1(\alpha) &= w\left(1 - \frac{0.986\sin\alpha}{0.119 + 1.341\sin\alpha - 0.754\sin^2\alpha}\right) \\ &+ (1-w)\left(\exp(-3.332\tan^{0.631}\frac{1}{2}\alpha)\right), \\ \Phi_2(\alpha) &= w\left(1 - \frac{0.238\sin\alpha}{0.119 + 1.341\sin\alpha - 0.754\sin^2\alpha}\right) \\ &+ (1-w)\exp\left(-90.56\tan^{1.218}\frac{1}{2}\alpha\right), \\ w &= \exp\left(-90.56\tan^2\frac{1}{2}\alpha\right). \end{aligned} \tag{18}$$

The coefficients a_1 and a_2 are estimated from observations using linear least squares. The absolute magnitude H and slope parameter G, can then be obtained from:

$$H = -2.5\log_{10}(a_1 + a_2), \tag{19}$$

$$G = \frac{a_2}{a_1 + a_2}. \tag{20}$$

The H,G phase function is not valid for phase angles greater than 120°. Different authors criticized the H, G system for its inability to accurately describe asteroid phase curves for low and high albedo asteroids (Belskaya & Shevchenko 2000). Couple of other, less popular phase functions were developed, for example Akimov (1988) and Shevchenko (1996) (Belskaya & Shevchenko 2000).

Two new phase functions (H,G_1,G_2 and H,G_{12}) were developed recently by Muinonen et al. (2010) to better fit asteroid phase curves and better predict apparent magnitudes. H,G_1,G_2 is targeted at very accurate phase curves, and H,G_{12} can

be used for asteroids with noisy photometric data. It was also shown in our treatment of half a million asteroids (paper IV), that the H,G_1,G_2 often fails for noisy data. The H,G_1,G_2 function will however be very useful for example for the Gaia satellite, which will produce highly accurate asteroid photometry. Fitting H,G_{12} phase function is a robust method for obtaining absolute magnitude and slope parameter(s) for noisy data. Both of those functions give equal or better fits to the data than the H,G function. In the H,G_1,G_2 magnitude phase function, the reduced apparent magnitudes $V(\alpha)$ can be obtained from (Muinonen et al. 2010):

$$\begin{aligned} 10^{-0.4V(\alpha)} &= a_1\Phi_1(\alpha) + a_2\Phi_2(\alpha) + a_3\Phi_3(\alpha) \\ &= 10^{-0.4H}\left[G_1\Phi_1(\alpha) + G_2\Phi_2(\alpha) + (1 - G_1 - G_2)\Phi_3(\alpha)\right], \end{aligned} \tag{21}$$

where the absolute magnitude H and slope parameters G_1, and G_2 are:

$$H = -2.5\log_{10}(a_1 + a_2 + a_3), \tag{22}$$

$$G_1 = \frac{a_1}{a_1 + a_2 + a_3}, \tag{23}$$

$$G_2 = \frac{a_2}{a_1 + a_2 + a_3}. \tag{24}$$

The coefficients a_1, a_2, a_3 are estimated from the observations using linear least squares. The basis functions $\Phi_1(\alpha)$, $\Phi_2(\alpha)$, and $\Phi_3(\alpha)$ are defined as:

- For $0° < \alpha \leq 7.5°$:
  - $\Phi_1(\alpha) = 1 - \frac{6}{\pi}\alpha$;
  - $\Phi_2(\alpha) = 1 - \frac{9}{5\pi}\alpha$;
  - $\Phi_3(\alpha)$ is defined using a cubic spline as defined in Table 2.
- For $7.5° < \alpha \leq 30°$:
  - $\Phi_1(\alpha)$ is defined using a cubic spline as defined in Table 1;
  - $\Phi_2(\alpha)$ is defined using a cubic spline as defined in Table 1;
  - $\Phi_3(\alpha)$ is defined using a cubic spline as defined in Table 2.
- For $30° < \alpha \leq 150°$:
  - $\Phi_1(\alpha)$ is defined using a cubic spline as defined in Table 1;
  - $\Phi_2(\alpha)$ is defined using a cubic spline as defined in Table 1;
  - $\Phi_3(\alpha) = 0$.

The first derivatives (per radian) at the ends of splines are: $\Phi_1'(\frac{\pi}{24}) = -\frac{6}{\pi}$, $\Phi_1'(\frac{5\pi}{6}) = -9.1328612 \times 10^{-2}$, $\Phi_2'(\frac{\pi}{24}) = -\frac{9}{5\pi}$, $\Phi_2'(\frac{5\pi}{6}) = -8.6573138 \times 10^{-8}$, $\Phi_3'(0) = -0.10630097$, $\Phi_3'(\frac{\pi}{6}) = 0$.

In the H,G_{12} magnitude phase function, the G_1 and G_2 from the three-parameter phase function are replaced by a single slope parameter G_{12} which relates to the G slope parameter in the H,G system (though there is not an exact correspondence).

Table 1: Knots for splines used in Φ_1 and Φ_2.

α (deg)	Φ_1	Φ_2
7.5	7.5×10^{-1}	9.25×10^{-1}
30.0	3.3486016×10^{-1}	6.2884169×10^{-1}
60.0	1.3410560×10^{-1}	3.1755495×10^{-1}
90.0	5.1104756×10^{-2}	1.2716367×10^{-1}
120.0	2.1465687×10^{-2}	2.2373903×10^{-2}
150.0	3.6396989×10^{-3}	1.6505689×10^{-4}

Table 2: Knots for spline used in Φ_3.

α (deg)	Φ_3
0.0	1
0.3	8.3381185×10^{-1}
1.0	5.7735424×10^{-1}
2.0	4.2144772×10^{-1}
4.0	2.3174230×10^{-1}
8.0	1.0348178×10^{-1}
12.0	6.1733473×10^{-2}
20.0	1.6107006×10^{-2}
30.0	0

The reduced flux densities can be obtained from (Muinonen et al. 2010):

$$10^{-0.4V(\alpha)} = L_0(G_1\Phi_1(\alpha) + G_2\Phi_2(\alpha) + (1 - G_1 - G_2)\Phi_3(\alpha)) \tag{25}$$

where:

$$G_1 = \begin{cases} 0.7527G_{12} + 0.06164, & \text{if } G_{12} < 0.2; \\ 0.9529G_{12} + 0.02162, & \text{otherwise;} \end{cases}$$
$$G_2 = \begin{cases} 0.9612G_{12} + 0.6270, & \text{if } G_{12} < 0.2; \\ 0.6125G_{12} + 0.5572, & \text{otherwise;} \end{cases} \tag{26}$$

$$H = -2.5 \log_{10} L_0; \tag{27}$$

and L_0 is the disk-integrated brightness at zero phase angle. The basis functions are as in the H,G_1,G_2 magnitude phase function. Coefficients L_0 and G_{12} are estimated from observations using non-linear least squares. Phase curve fitting for H,G_1,G_2 and H,G phase functions can be reduced to linear least square fitting via shifting to a flux domain. The flux for the i^{th} observation is computed using:

$$\begin{aligned} L_i &= 10^{-0.4V_i}, \\ \sigma_i^{(L)} &= L_i(10^{0.4\sigma_i^{(V)}} - 1), \end{aligned} \tag{28}$$

where $\sigma_i^{(V)}$ are the standard deviations of the magnitude measurements. The χ^2-value to be minimized here with respect to the parameters $\mathbf{a}$ is

$$\chi^2(\mathbf{a}) = \sum_{i=1}^{N} \frac{[L_i - L_i(\alpha_i, \mathbf{a})]^2}{[\sigma_i^{(L)}]^2}. \tag{29}$$

The computed disk-integrated brightnesses are expressed via N_a basis functions $\Phi_1(\alpha), \Phi_2(\alpha), \ldots, \Phi_{N_a}(\alpha)$:

$$L_i(\alpha_i, \mathbf{a}) \;=\; \sum_{l=1}^{N_a} a_l \Phi_l(\alpha_i). \tag{30}$$

In the case of H,G_{12}, fitting cannot be reduced to linear least squares and nonlinear techniques have to be employed.

4.4 MC and MCMC error computation

Errors in the absolute magnitude and slope parameters are known to be non-Gaussian. Absolute magnitude and slope parameter(s) error estimation in the case of H,G and H,G_1,G_2 is carried out using Monte-Carlo sampling. Equations 17 and 21 are linear with respect to a_1, a_2, or a_1, a_2, a_3 and result in Gaussian errors in those parameters. Converting it to absolute magnitude and slope parameters with the help of non-linear Eqs. 19, 20, 22, 23, 24 results in non-linear errors in the H,G or H,G_1,G_2 parameters. The least squares solution and error covariance matrix of the parameters a can be utilized to obtain errors in the H,G or H,G_1,G_2 parameters.

With the help of multi-normal distribution with a mean equal to the previously obtained a parameters and covariance equal to error covariance of a parameters from the least squares solution, we generate a sample of a_1, a_2 or a_1, a_2, a_3 using a multi-normal random number generator. The a_i samples are then converted to H,G or H,G_1,G_2 samples using either Eqs. 19 and 20 or 22, 23 and 24. Next, the samples are ordered in descending goodness of fit and then 68.27% (equivalent to 1-σ) and 99.73 % (equivalent to 3-σ) error cut-offs are computed, resulting in a list of subsamples. The limiting (maximum and minimum) H,G or H,G_1,G_2 parameters are selected from that list and the two-sided errors are computed.

To estimate absolute magnitude and slope parameter errors in the case of H,G_{12}, we use a Markov-chain Monte Carlo method. General Monte Carlo is not applicable here as Eq. 25 is non-linear and the errors in the fitted parameters, that is in this case of L_0 and G_{12} cannot be assumed Gaussian. We start a single Markov chain in the least-squares point for L_0, G_{12}, and make use of a multivariate Gaussian proposal distribution, where an error covariance matrix for L_0, G_{12} is taken from the least-squares solution. After obtaining a substantial number of different solutions, two-sided errors are computed based on the equivalents of the 1-σ and 3-σ cut-offs as in the H,G and H,G_1,G_2 phase functions.

4.5 Selected results

Markov-chain Monte Carlo and Monte Carlo methods were used to assess phase curve photometric parameters and their uncertainties. Absolute magnitudes and photometric parameters were derived for half a million asteroids by fitting phase curves to Lowell Observatory photometric database. The Lowell Observatory photometric database contains the largest set of asteroid phase curves currently known. Phase curves are fitted to all the data available without making an assumption on

any of the photometric parameters (in other databases slope parameter G is often assumed to be $G = 0.15$ or $G = 0.18$). In particular we made use of the novel phase functions, which are an improvement over the old H, G phase function. Fitting the phase functions to a large number of asteroid families suggests homogeneity of photometric parameters in asteroid families. This is yet another piece of evidence (next to color and taxonomic homogeneity) supporting the idea of asteroid families originating from homogeneous parent bodies disrupted by collisions. One exception from this rule is the Nysa-Polana family, which was previously indicated to be composed of two different classes of asteroids, perhaps originating from different parent bodies. In the current analysis, the Nysa-Polana family also shows separation into two distinct groups (see Fig. 2).

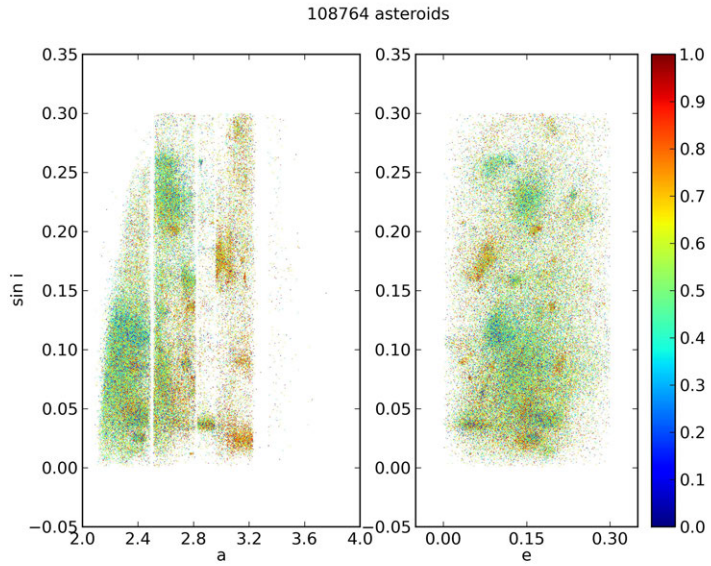

Figure 7: Proper elements for main-belt asteroids color-coded with the G_{12} parameter.

Figure 7 presents the distribution of proper elements for the main-belt asteroids color-coded with the G_{12} parameters. The visible clumps correspond to asteroid families, which tend to have similar G_{12} photometric parameters.

The derived photometric parameters are also found to correspond to asteroid taxonomic complexes and colors. This correlation is investigated and suggest that some C-complex asteroid families can be identified using this correlation and Bayesian statistics. As an example, we indicate the Dora family having the mean $G_{12} = 0.7$ and standard deviation $\sigma_{G_{12}} = 0.18$, which results in a high C-complex preponderance probability. The distribution of the photometric parameter G_{12} for the Dora family is plotted in Fig. 8. The over plotted probability functions correspond to C, S, and X taxonomic complexes. Most of the Dora family asteroids must have originated from the C-complex distribution. It is clear that phase curve parameters correspond to many of an asteroid's surface properties. The correlations between those parameters are however poorly known today and will have to be investigated further in the future.

The H,G_1,G_2 and H,G_{12} are to be proposed to the IAU to replace the H,G function together with a list of absolute magnitudes and slope parameters for half

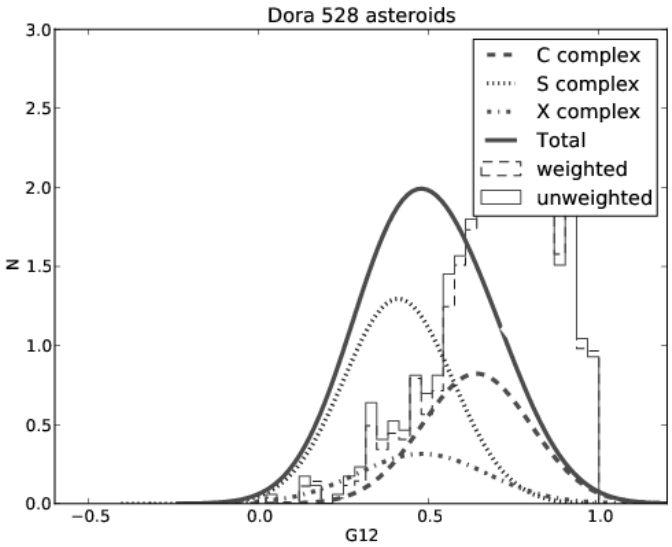

Figure 8: Distribution of the G_{12} photometric parameter for the Dora family.

a million asteroids. We have also developed an online open-use Java applet, named "Asteroid Phase Function Analyzer" computing absolute magnitudes and slope parameter(s) of all the three phase functions. This software will be further developed for the Gaia processing pipeline.

5 Relevance to Gaia mission

The MCMC ranging described in Chapter 3 has been implemented into the Gaia data-processing pipeline. The method is applicable to poorly observed, single-apparition asteroids (for example, new discoveries) and will be used in short-term processing of the Gaia data. Gaia will provide sparse high accuracy data sets requiring orbital analysis based on Bayesian statistics in the early processing. The end-of-mission (after 5 years of operation) average number of observations per asteroid is going to be around 60, each single scan containing 3-4 observations. Extreme Gaia astrometric accuracy ($20 - 200$ μas compared to current ground based accuracy of 0.5 as) will lead to much improved orbits. For single-scan observations, it has been shown that Gaia observations will lead to a substantial collapse of orbital-element p.d.f.s based on methods such as MCMC ranging. MCMC ranging also has a potential of contributing to other problems such as for example performing dynamical classification, identifying asteroids, and aiding the recovery of lost objects after the mission is over. For extensively observed asteroids during the course of the Gaia mission (long-term processing), the standard linear approximation based on nonlinear least squares will provide the differentially corrected orbital elements and their covariance matrices based on partial derivatives. It is expected that Gaia will produce around hundred times better orbits than currently available leading to end-of-mission semimajor axis accuracy as good as $\sigma_a = 10^{-9}$ AU.

The photometric phase curves and software described in chapter 4 will be utilized in the Gaia pipeline. The H,G_1,G_2 function will be very useful for the Gaia satellite, which will produce highly accurate asteroid photometry. Expected phase angle coverage for Gaia is from about 8° to 36° for the main-belt objects and from 8° to 90° for the near-Earth objects. This should result in a decent phase curve coverage and estimation of photometric parameters especially for the NEAs. Most probably the sparse photometry (corresponding to random points on asteroids lightcurve) obtained by Gaia will require simultaneous spin, shape, and phase curve modeling.

6 Summary of papers

6.1 Paper I: Oszkiewicz D. A., Muinonen K., Virtanen J., Granvik M., Asteroid orbital ranging using Markov-Chain Monte Carlo. *Meteoritics & Planetary Science* 44, 1897-1904 (2009)

In paper I, we introduce a new orbital inversion method, based on the Metropolis-Hastings algorithm. Similarly to MC ranging, the method is targeted at asteroids with short observational arcs (such as for example new discoveries) and two or more observations. The basics of orbital inverse problem expressed in Bayesian statistics are reviewed. The method is then applied to three objects from different dynamical groups, near-Earth, main-belt, and transneptunian objects. Possible issues related to tuning the algorithm are also described. We compare effectiveness of MCMC ranging versus MC ranging method. The method is specially important for NEAs requiring impact probability computation based on short observational arcs.

6.2 Paper II: Granvik M., Virtanen J., Oszkiewicz D. A., Muinonen K., OpenOrb: Open-source asteroid orbit computation software including statistical ranging. *Meteoritics & Planetary Science* 44, 1853-1861 (2009).

Paper II introduces an open-source orbit computation package called OpenOrb. The tool includes the MCMC ranging method.

6.3 Paper III: Oszkiewicz D. A., Muinonen K., Bowell E., Trilling D., Penttilä A., Pieniluoma T., Wasserman L. H., Enga M.-T., Online multi-parameter phase-curve fitting and application to a large corpus of asteroid photometric data. *Journal of Quantitive Spectroscopy & Radiative Transfer* 112, 1919-1929 (2011)

Paper III introduced an online java applet called Asteroid Phase Function Analyzer for computing absolute magnitudes and slope parameters using three different phase functions: the H,G phase function, the H,G_1,G_2 phase function, and the H,G_{12} phase function. The tool also includes non-Gaussian error analysis using Monte Carlo methods. We give an overview of phase functions, and describe numerical methods used for the Least Squares solution and Monte-Carlo error estimation. We describe the software and subsequent packages. We apply the software to about half a million of asteroids. We report photometric parameters for the first ten numbered ones and show typical phase curves obtained from the Lowell Observatory data set. We discuss possible homogeneity of slope parameters within dynamical families and taxonomic classes. We find a correlation of family derived slope parameters G_1 and G_2 with median family albedo. We plan to present the obtained list of absolute

magnitudes and slope parameters for the half a million of asteroids to the IAU at the General Assembly. The results will also be made publicly available at that time.

6.4 Paper IV: Oszkiewicz D. A., Bowell E., Wasserman L. H., Muinonen K., Trilling D., Penttilä A., Pieniluoma T., Asteroid taxonomic signatures from photometric phase curves. *Icarus* 219, 283-296 (2012)

Paper IV investigates the correlation of the photometric parameter G_{12} and taxonomy. Photometric parameters computed in paper IV are used along with spectral classification of asteroids to derive G_{12} distributions for the main taxonomic complexes (C, S, and X). Based on the obtained G_{12} distributions for taxonomic complexes and G_{12} values for individual asteroids, we compute probabilities of individual asteroids of belonging to C, S, and X complexes. Taxonomic complex preponderance in asteroid families is then predicted and compared to earlier findings in the literature. We continue presenting compositional fractions of the taxonomic complexes in the main belt. This paper constitutes the widest exploration of the taxonomy-slope parameter correlation so far.

6.5 Paper V: Oszkiewicz D. A., Muinonen K., Virtanen J., Granvik M., Bowell E., Modeling collision probability for Earth-impactor 2008 TC3. *Planetary & Space Science, in press.*

In paper V, we further develop the MCMC method and propose an automatization procedure for tuning of the proposal density. We review convergence statistics. Further, we apply MCMC ranging method to asteroid 2008 TC_3 to obtain a distribution of possible orbital elements based on maximum 6 observations (0.06 day time interval). Based on the orbits obtained, we study the evolution of the Earth-impact probability with increasing observational arc and different assumptions for the observational noise. We conclude that we could predict the impact of 2008 TC_3 asteroid with very high (approx. 95%) probability based on 6 observations and 0.5 arc seconds noise assumption. We highlight the importance of the collision-probability methods based on Bayesian statistics for asteroids with short observational arcs and short impact notice. We investigate the sensitivity of our collision probability computation to the noise assumption and propose a new sampling technique, which includes astrometric uncertainty as an inversion parameter.

6.6 Author's contributions

The author of this thesis developed the MCMC ranging technique originating from MC ranging. The author investigated applications of this method (paper I and V) to the orbital computation problem and collision probability estimation. The author implemented the method first in Fortran95 in OpenOrb package (described in paper II) and then initiated implementation in Java for the needs of the Gaia

mission. In papers I and V, the author is responsible for the application of the orbit computation techniques to the example cases of the selected asteroids as well as for illustrating and describing the results. In paper III the author was responsible for the implementation (Java) of the online software for fitting asteroid phase curves. The author has also been responsible for applying the software and processing the vast photometric database (containing data for about half a million of asteroids) at 96-core computer cluster "Javelina" at the Northern Arizona University. In papers III and IV, the author was also responsible for applying the software to asteroid families and taxonomic classes. In paper IV the author performed the taxonomic analysis of asteroid families based on the G_{12} slope parameters. The author is responsible for performing the taxonomic complex probability computations, and illustrating and describing the results. Papers I, III, IV, V were written by the author except for section 5 (Photometric data and data calibration) and parts of the conclusions in paper III.

References

Alexander, A. (1965), *The planet Uranus: a history of observation, theory, and discovery*, American Elsevier Pub. Co.

Armitage, A. (1962), *William Herschel*, British men of science, Nelson.

Belskaya, I. N. & Shevchenko, V. G. (2000), 'Opposition effect of asteroids', *Icarus* **147**, 94–105.

Binzel, R. & Xu, S. (1993), 'Chips off of asteroid 4 vesta: Evidence for the parent body of basaltic achondrite meteorites', *Science* **260**(5105), 186–191.

Bottke, W. F., Rubincam, D. P. & Burns, J. A. (2000), 'Dynamical evolution of main belt meteoroids: Numerical simulations incorporating planetary perturbations and yarkovsky thermal forces', *Icarus* **145**(2), 301 – 331.
URL: *http://www.sciencedirect.com/science/article/pii/S0019103500963619*

Bowell, E., Hapke, B., Domingue, D., Lumme, K., Peltoniemi, J. & Harris, A. W. (1989), Application of photometric models to asteroids, *in* 'Asteroids II', University of Arizona Press, pp. 524–555.

Brown, J. C. (1995), 'Overview of topical issue on inverse problems in astronomy.', *Inverse Problems* **11**, 635–638.

Buchheim, R. (2010), Methods and lessons learned determining the hg parameters of asteroid phase curves, *in* 'Symposium on Telescope Science', p. 101.

Burbine, T., McCoy, T., Meibom, A., Gladman, B. & Keil, K. (2002), 'Meteoritic parent bodies: Their number and identification', *Asteroids III* **653**.

Bus, S. (1999), Compositional Structure in the Asteroid Belt: Results of a Spectroscopic Survey., PhD thesis, Massachusetts Institute of Technology.

Chesley, S. R. & Milani, A. (1999), Nonlinear methods for the propagation of orbital uncertainty, *in* 'Proceedings of the AAS/AIAA Space Flight Mechanics Meeting, Girkwood, Alaska, 1999', pp. 99–418.

Chodas, P. W. & Yeomans, D. K. (1996), *The Collision of Comet Shoemaker-Levy and Jupiter*, Cambridge Univ. Press.

Connors, M., Wiegert, P. & Veillet, C. (2011), 'Earth's Trojan asteroid', *Nature* **475**(7357), 481–483.

Cowles, M. K. & Carlin, B. P. (1996), 'Markov chain monte carlo convergence diagnostics: A comparative review', *Journal of the American Statistical Association* **91**, 883–904.

Degewij, J., Tedesco, E. & Zellner, B. (1979), 'Albedo and color contrasts on asteroid surfaces', *Icarus* **40**(3), 364 – 374.

Dubyago, A. D. (1961), *The determination of orbits*, The Macmillan company.

Farinella, P., Froeschlé, C. & Gonczi, R. (1993), 'Meteorites from the asteroid 6 hebe', *Celestial Mechanics and Dynamical Astronomy* **56**(1), 287–305.

Gelman, A., Roberts, G. & Gilks, W. R. (1996), *Bayesian Statistics V*, Oxford University Press.

Gelman, A. & Rubin, D. B. (1992), 'Inference from iterative simulation using multiple sequences', *Statistical Science* **7**, 457–472.

Gewke, J. (1992), *Bayesian Statistics 4*, Oxford University Press.

Haario, H., Saksman, E. & Tamminen, J. (1999), 'Adaptive proposal distribution for random walk metropolis algorithm', *Computational Statistics* **14**, 375–395.

Harris, A. W. (1989), The H-G asteroid magnitude system: Mean slope parameters, *in* 'Abstracts of the Lunar and Planetary Science Conference', p. 20.

Hirayama, K. (1918), 'Groups of asteroids probably of common origin.', *Astron. J.* **31**, 185–188.

Hirayama, K. (1928), 'Families of asteroids.', *Jap. J. Astron. Geophys.* **5**, 137–162.

Hirayama, K. (1933), 'Present state of the families of asteroids', *Proc. Imp. Acad. Japan* **9**, 482–485.

Ivezić, Ž., Lupton, R., Jurić, M., Tabachnik, S., Quinn, T., Gunn, J., Knapp, G., Rockosi, C. & Brinkmann, J. (2002), 'Color confirmation of asteroid families', *The Astronomical Journal* **124**, 2943.

Ivezić, Ž., Tabachnik, S., Rafikov, R., Lupton, R. H., Quinn, T., Hammergren, M., Eyer, L., Chu, J., Armstrong, J. C., Fan, X., Finlator, K., Geballe, T. R., Gunn, J. E., Hennessy, G. S., Knapp, G. R., Leggett, S. K., Munn, J. A., Pier, J. R., Rockosi, C. M., Schneider, D. P., Strauss, M. A., Yanny, B., Brinkmann, J., Csabai, I., Hindsley, R. B., Kent, S., Lamb, D. Q., Margon, B., McKay, T. A., Smith, J. A., Waddel, P. & York, D. G. (2001), 'Solar system objects observed in the Sloan digital sky survey commissioning data', *The Astron. J.* **122**(5), 2749.

Jaki, S. L. (1972), 'The early history of the Titius-Bode law', *American Journal of Physics* **40**, 10–14.

Jeffreys, H. (1946), An invariant form for the prior probability in estimation problems, *in* 'Proceedings of the Royal Society of London', Vol. 186, pp. 453–461.

Jurić, M., Ivezić, Ž., Lupton, R., Quinn, T., Tabachnik, S., Fan, X., Gunn, J. E., Hennessy, G. S., Knapp, G. R., Munn, J. A., Pier, J. R., Rockosi, C. M., Schneider, D., P., Brinkmann, J., Csabai, I. & Fukugita, M. (2002), 'Comparison of positions and magnitudes of asteroids observed in the Sloan digital sky survey with those predicted for known asteroids', *The Astron. J.* **124**(3), 1776.

Kristensen, L. K. (1992), 'The identification problem in asteroid surveys.', *Astron. Astroph.* **262**, 606–612.

Lagerkvist, C.-I. & Magnusson, P. (1990), 'Analysis of asteroid lightcurves. II - Phase curves in a generalized HG-system', *Astronomy and Astrophysics Supplement Series* **86**, 119–165.

Lardner, D. (1854), *Hand-books of natural philosophy and astronomy*, number v. 2 *in* 'Hand-books of Natural Philosophy and Astronomy', Blanchard and Lea.

Lardner, D. & Dunkin, E. (1860), *Handbook of astronomy*, Oxford University.

Lehtinen, M. S. (2010), *Theory and Applications of Inverse Problems*, Pitman Res. Notes in Math. Ser. Longman Scientific and Technical.

Marsden, B. G. (1985), 'Initial orbit determination - the pragmatist's point of view.', *Astron. J.* **90**, 1541–1547.

Marsden, B. G. (1991), 'The computation of orbits in intermediate and uncertain cases.', *Astron. J.* **102**, 1539–1552.

Marzari, F., Farinella, F. & Davis, D. R. (1999), 'Origin, aging and death of asteroid families', *Icarus* **142**, 63–77.

McNaught, R. H. (1999), 'Letter to the editor: Unusual object 1998 DK_{36}', *J.Br. Astron. Assoc.* **109**, 229.

Mengersen, K. L., Robert, C. P. & Guihenneuc-Jouyaux, C. (1999), *Bayesian Statistics 6*, Oxford University Press.

Menke, W. (1989), *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press.

Mignard, F. (2002), 'Observations of solar system objects with GAIA. I detection of neos', *Astronomy and Astrophysics* **393**(2), 727–731.

Mignard, F., Cellino, A., Muinoen, K., Tanga, P., Delbò, M., Dell'Oro, A., Granvik, M., Hestroffer, D., Mouret, S., Thuilot, W. & Virtanen, J. (2007), 'The Gaia mission: Expected applications to asteroid science', *Earth Moon Planets* **101**, 97–125.

Milani, A. (1999), 'The asteroid identification problem. I. Recovery of lost asteroids', *Icarus* **137**, 269–292.

Milani, A., Gronchi, G., Vitturi, M. & Knežević, Z. (2004), 'Orbit determination with very short arcs. I admissible regions', *Celestial Mechanics and Dynamical Astronomy* **90**(1), 57–85.

Morbidelli, A., Bottke, W. F., Froeschle, C. & Michael, P. (2002), Origin and evolution of near-Earth objects, *in* 'Asteroids III', Univ. of Arizona Press, Tucson, pp. 409–422.

Morbidelli, A., Zappala, V., Moons, M., Cellino, A. & Gonczi, R. (1995), 'Groups of asteroids probably of common origin', *Icarus* **118**, 132–154.

Mothé-Diniz, T., Roig, F. & Carvano, J. (2005), 'Reanalysis of asteroid families structure through visible spectroscopy', *Icarus* **174**, 54–80.

MPC (accessed 2012), 'IAU Minor Planet Center'. `http://www.minorplanetcenter.net/iau/mpc.html`.

Muinonen, K., Belskaya, I. N., Cellino, A., Delbò, M., Levasseur-Regourd, A.-C., Penttilä, A. & Tedesco, E. F. (2010), 'A three-parameter magnitude phase function for asteroids', *Icarus* **209**, 542–555.

Muinonen, K. & Bowell, E. (1993), 'Asteroid orbit determination using Bayesian probabilities', *Icarus* **104**, 255–279.

Muinonen, K., Virtanen, J. & Bowell, E. (2001), 'Collision probability for Earth-crossing asteroids using orbital ranging', *Celestial Mechanics and Dynamical Astronomy* **81**, 93–101.

Orlov, B. A. (1939), 'Determination of the preliminary orbit of minor planet by two observations', *Pulkovo Obs. Circ.* **26**, 55–63.

Oszkiewicz, D., Muinonen, K., Bowell, E., Trilling, D., Penttilä, A., Pieniluoma, T., Wasserman, L. & Enga, M.-T. (2011), 'Online multi-parameter phase-curve fitting and application to a large corpus of asteroid photometric data', *Journal of Quantitative Spectroscopy and Radiative Transfer* **112**(11), 1919–1929.

Oszkiewicz, D., Muinonen, K., Bowell, E., Trilling, D., Penttilä, A., Pieniluoma, T., Wasserman, L. & Enga, M.-T. (2012), 'Asteroid taxonomic signatures from photometric phase curves', *Icarus* **219**, 283–296.

Oszkiewicz, D., Muinonen, K., Virtanen, J. & Granvik, M. (2009), 'Asteroid orbital ranging using Markov-Chain Monte Carlo', *Meteoritics & Planetary Science* **44**, 1897–1904.

Parker, A., Ivezić, Ž., Jurić, M., Lupton, R., Sekora, M. D. & Kowalski, A. (2008), 'The size distributions of asteroid families in the SDSS moving object catalog 4', *Icarus* **198**, 138–155.

Perryman, M., Lindegren, L. & Turon, C. (1997), The scientific goals of the gaia mission, *in* 'Hipparcos-Venice'97', Vol. 402, pp. 743–748.

Raftery, A. & Lewis, S. (1992), *Bayesian Statistics 4*, Oxford University Press.

Roberts, G. & Rosenthal, J. (2004), 'General state space markov chains and mcmc algorithms', *Probability Surveys* **1**, 20–71.

Rodgers, C., Canterna, R., Smith, J., Pierce, M. & Tucker, D. (2006), 'Improved $u'g'r'i'z'$ to $UBVR_CI_C$ transformation equations for main-sequence stars', *The Astronomical Journal* **132**, 989.

Schneider, M. (2011), 'Bayesian linking of geosynchronous orbital debris tracks as seen by the large synoptic survey telescope', *Advances in Space Research* .

Schulte, P., Alegret, L., Arenillas, I., Arz, J., Barton, P., Bown, P., Bralower, T., Christeson, G., Claeys, P., Cockell, C. et al. (2010), 'The chicxulub asteroid impact and mass extinction at the cretaceous-paleogene boundary', *Science* **327**(5970), 1214–1218.

Serio, G. F., Manara, A., Sicoli, P. & Bottke, W. F. (2002), *Giuseppe Piazzi and the discovery of Ceres*, University of Arizona space science series, University of Arizona Press.

Shu, F. (1982), *The physical universe: An introduction to astronomy*, Univ Science Books.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics.

Tedesco, E. (1989), Asteroid magnitudes, UBV colors, and IRAS albedos and diameters, *in* 'Asteroids II', University of Arizona Press, pp. 1090–1138.

Teets, D. & Whitehead, K. (1999), 'The discovery of Ceres: How Gauss became famous', *Mathematics Magazine* **72**, 83–93.

Tholen, D. & Whiteley, R. J. (1998), 'Results from NEO searches at small solar elongation', *Bull. Am. Astron. Soc.* **30**, 1041.

Thomas, C. A., Trilling, D. E., Emery, J. P., Mueller, M., Hora, J. L., Benner, L. A. M., Bhattacharya, B., Bottke, W. F., Chesley, S., Delbò, M., Fazio, G., Harris, A. W., Mainzer, A., Mommert, M., Morbidelli, A., Penprase, B., Smith, H. A., Spahr, T. B. & Stansberry, J. A. (2011), 'ExploreNEOs. V. Average albedo by taxonomic complex in the near-Earth asteroid population', *The Astronomical Journal* **142**(3), 85.

Trilling, D., Bhattacharya, B., Blaylock, M., Stansberry, J., Sykes, M. & Wasserman, L. (2007), The Spitzer asteroid catalog: Albedos and diameters of 35,000 asteroids, *in* 'Bulletin of the American Astronomical Society', Vol. 38, p. 484.

Tsuchiyama, A., Uesugi, M., Matsushima, T., Michikami, T., Kadono, T., Nakamura, T., Uesugi, K., Nakano, T., Sandford, S., Noguchi, R. et al. (2011), 'Three-dimensional structure of hayabusa samples: Origin and evolution of itokawa regolith', *Science* **333**(6046), 1125–1128.

Väisälä, Y. (1939), 'Eine einfache methode der bahnbestimmung', *Suomalainen Tiedeakatemia* **1**, 1–32.

Vallado, D. & McClain, W. (2001), *Fundamentals of astrodynamics and applications*, Vol. 12, Springer.

Virtanen, J. (2005), Asteroid orbital inversion using statistical methods, PhD thesis, University of Helsinki.

Virtanen, J., Muinonen, K. & Bowell, E. (2001), 'Statistical ranging of asteroid orbits', *Icarus* **154**, 412–431.

Wright, E., Eisenhardt, P., Mainzer, A., Ressler, M., Cutri, R., Jarrett, T., Kirkpatrick, J., Padgett, D., McMillan, R., Skrutskie, M. et al. (2010), 'The wide-field infrared survey explorer (WISE): Mission description and initial on-orbit performance', *The Astronomical Journal* **140**, 1868.

Zappalà, V., Bendjoya, P., Cellino, A., Farinella, P. & Froeschlé, C. (1995), 'Asteroid families: Search of a 12,487-asteroid sample using two different clustering techniques', *Icarus* **116**, 291–314.

Zellner, A. & Min, C.-K. (1995), 'Gibbs sampler convergence criteria', *Journal of the Americal Statistical Association* **90**, 921–927.