

Statistical analysis of asteroidal and cometary polarization phase curves

A. Penttilä¹, K. Lumme¹, E. Hadamcik², and A.-C. Levasseur-Regourd²

¹ Observatory, University of Helsinki, PO Box 14, 00014 University of Helsinki, Finland
e-mail: Antti.I.Penttila@helsinki.fi

² Université Paris VI/Aéronomie CNRS-IPSL, BP 3, 92371 Verrières, France

Received 7 October 2004 / Accepted 15 November 2004

Abstract. We use an empirical polarization phase curve model at various wavelengths to predict some statistical properties of comets and asteroids. First, we show how our model with Bayesian MCMC numerical methods predicts polarization behaviour at larger phase angles when only the smaller angle data are used. Our empirical model incorporates both the phase and colour dependency on the same footing. Second, a comparison between the existing taxonomy of asteroids and the clustering of asteroid classes based on polarization suggests a new method for the classification. We also heavily concentrate on the derivations of quantitative errors in our analysis.

Key words. polarization – methods: statistical – techniques: polarimetric – astronomical data bases: miscellaneous – comets: general – minor planets, asteroids

1. Introduction

Photopolarimetry has been for a long time one of the key methods to retrieve physical information about the Solar System objects. These include the planets, their aerosols, satellites and rings, the asteroids, the cometary comae and the interplanetary dust. Roughly speaking, they can be subdivided into objects with a surface (planets, satellites, asteroids) and clouds of particles (aerosols in planetary atmospheres, rings, coma, interplanetary dust). The solar light scattered by such media is actually partially linearly polarized.

The linear polarization, as defined in e.g. Gehrels (1974) or Hapke (1993), is the ratio of the difference to the sum of the two polarized components of the brightness, respectively perpendicular and parallel to the scattering plane. One of the most interesting points is indeed that, since the polarization is a ratio (remaining between -1 and $+1$), no normalization in brightness is required to compare data obtained on different objects (or on the same object at different times). This is especially important for comets, since the brightness variations with the distances to the Sun and to the Earth depend not only upon the distances but also upon changes in the activity of the comet and upon the size of the coma.

Both types of objects have been widely modelled using the ever-increasing amount of observational data. It is fairly obvious that the models for bodies with an atmosphere have been successful and have produced unique knowledge of the physical structure. The basic reason for this is that light-scattering models are much easier to apply to atmospheres because the basic constituents are much smaller than the wavelengths

used. Then the Rayleigh-type light scattering approach is quite adequate.

Seelinger's work (1887) on Saturn's rings was obviously the first attempt to gain quantitative information of the ring particles. He introduced the basic idea of the mutual shadowing concept which since then has been widely used for the regolith studies to explain the classical opposition effect. The fundamental difficulty in the shadowing scheme is the question how to separate the shadowing contribution and that originating from a single particle because both depend on the same scattering angle or phase angle. The implicit and completely unjustified assumption normally made was that small ring or regolith particles do not have their own strong opposition brightening. It was not until much after Seelinger's work when the dynamists realized that the packing density of Saturn's rings is about ten times as high as predicted by the shadowing mechanism (Dones et al. 1993; Mishchenko 1993).

At about the same time computer capabilities had increased so much that exact wave-optical calculations became possible for small non-spherical particles and their aggregates. This allowed a better interpretation of polarimetric data. Indeed, the question of negative polarization, sometimes wrongly called anomalous polarization, has recently been a widely studied subject in planetary research (see e.g. Muinonen et al. 2002; Muinonen 2004; Shkuratov et al. 2004). Here the fundamental question is why the inversion angle is so insensitive to parameters like albedo or wavelength.

An increasing amount of observational data has been compared to numerical data for modelled particles to obtain clues

about the dust physical properties. These comparisons need a continuous curve to fit the data and if possible predict the values for phase angles where data are missing. For very small particles in an atmosphere or for liquid (and thus spherical) aerosols, the Mie scattering theory can be used. However, such an approach is not always adequate for solid particles (see Hadamcik et al. 2003d, for soot particles). Lumme et al. (2003) started a systematical study of light scattering by various aggregates of wavelength-sized constituents. These studies unequivocally show that both nonlinear backward scattering and negative polarization naturally follow without any restrictive assumptions. Unfortunately computer capabilities still limit the aggregate sizes to a few hundred constituents. Calling these aggregates a single particle allows radiative transfer calculations for large aggregates and regoliths.

The structure of the paper is such that in Sect. 2 we will present the asteroidal and cometary data to be used. In Sect. 3 we will introduce the empirical model for polarization phase curves and study the parameter estimation and error analysis of the model. In Sect. 4 we will use the model in analyzing and predicting polarization. Furthermore, we will show the link between asteroid taxonomy and polarization. In Sect. 5 we will extend the empirical polarization phase curve model to multi-wavelength data. In Sect. 6 we will study the wavelength effect on polarization using our multi-wavelength model.

2. Polarization data for asteroids and comets

Numerous polarimetric measurements for asteroids and comets can be found in published papers. Some compilations were attempted for asteroids (available on the Internet at the Small Bodies Node (SBN)¹). Except for some near-Earth objects, unfortunately, the explored phase angles are smaller than 30°. For cometary data, a general electronic database does not yet exist and it is thus necessary to use the published papers on polarimetric observations of comets.

For all these objects (asteroidal surfaces and cometary comae), phase curves at least within a given wavelength range present similar smooth shapes with a small negative branch, an inversion region, and a wide positive branch with a maximum near 90°. Data retrieved near opposition (phase angle equal to 0°) correspond to extremely small polarizations. Although no data have ever been obtained at 180°, with a current maximal phase angle for comets equal to 121° (Hadamcik & Levasseur-Regourd 2003c), polarization seems to significantly decrease near forward scattering.

2.1. Asteroidal database

The polarimetric asteroidal data used in this paper come mainly from the SBN database². They include 1635 entries for 137 asteroids (number, date of observation, filter, phase angle, polarization, measurement error, position angle etc.). More than 85 percent of the measurements are in

the U (362 nm), B (435 nm) and V (559 nm) and only 7 percent in the R (685 nm) wavelength domains. The phase angle range is from 0.1° to 120°, but in most cases the phase angles are limited to less than 30°. Only some near-Earth objects were explored at large phase angles. A classification based on spectra obtained in the visual and near infrared domains, and data on the physical parameters of the asteroids are also available at the same address. Further data, not yet included in the SBN, have been published by e.g. Goidet-Devel et al. (1995), Mukai et al. (1997), Kiselev et al. (1999), Kiselev et al. (2002).

2.2. Cometary database

Polarimetric observations of cometary dust are not easy. First, it is necessary to use filters to avoid the gaseous emissions. If the filters are not correctly chosen or the measured intensities are not corrected from the gaseous emissions by a study of their spectrum close to the period of observations in polarization, the measurements are not adequate for a comparison with other comets. The second difficulty is related to the variation of the polarization with the aperture due to differences of dust properties in the coma (Jockers 1997; Hadamcik & Levasseur-Regourd 2003b,c). To build up a database for cometary dust it is thus necessary to use large enough apertures (depending on the comet). The variation in the observations at a fixed phase angle is mainly due to aperture differences. With the imaging polarimetry technique it is possible to observe the coma regions and to better choose an aperture that includes the main structure (Hadamcik & Levasseur-Regourd 2003b; Hadamcik & Levasseur-Regourd 2003c). However, polarimetric observations do not require normalization and give fundamental information on the physical properties.

Levasseur-Regourd et al. (1996) used all the data already published by various groups (see references therein) to derive a first classification of comets from their polarimetric properties. The data from new observations were later added to this database (Hadamcik & Levasseur-Regourd 2003b; Hadamcik & Levasseur-Regourd 2003c, and references therein). Finally, S. Kikuchi (personal communication) added data, mainly for the comet 109P/Swift-Tuttle. The database includes 36 comets with more than 1000 data points between the near ultra-violet and near-infrared spectral domains, but mainly in the green and red domains. Two comets have been extensively observed, comet 1P/Halley at its 1985–1986 apparition (for which the Vegas and Giotto space probes provided some “ground truth”) and the bright comet C/1995 O1 Hale-Bopp in 1995–1997. The generally used continuum filters are from the International Halley Watch program (IHW) or later from NASA (see wavelengths and eventual contamination in Jockers 1997). The phase angle range is from 0.6° to 121°.

The cometary data that we analyze in this article consist of observations of comets C/1995 O1 Hale-Bopp, 1P/Halley, C/1996 B2 Hyakutake, 109P/Swift-Tuttle and C/1975 V1 West. Comets will be referred to in the text without the catalogue identifier. We will also refer to the asteroids 3 Juno and 4179 Toutatis without the identifier.

¹ <http://pdssbn.astro.umd.edu>

² http://pdssbn.astro.umd.edu/sbnhtml/asteroids/polar_radar.html

3. Model for polarization phase curve

Lumme & Muinonen (1993) have suggested the following function as an empirical model for a polarization phase curve

$$P(\alpha) = b \sin^{c_1}(\alpha) \cos^{c_2}\left(\frac{\alpha}{2}\right) \sin(\alpha - \alpha_0). \quad (1)$$

This function has some clear advantages as an empirical model. The most important advantage is that it can describe polarization throughout the phase angle range $[0^\circ, 180^\circ]$. Furthermore, the values of the function are limited to the range $[-1, 1]$ when the parameter ranges are correctly defined. This guarantees that the function is always physically reasonable, as polarization by definition is also in that range. Also noticeable is that an important polarization feature, the inversion angle α_0 , is explicitly included in the expression.

This function has been used by numerous authors to fit the observational data (e.g. Goidet-Devel et al. 1995; Levasseur-Regourd et al. 1996; Hadamcik & Levasseur-Regourd 2003b,c; Kiselev et al. 2002) but nevertheless the meaning of its formalism has not been published in detail before.

3.1. Model parameters and their effect

The phase curve model in Eq. (1) has four parameters: b , α_0 , c_1 and c_2 . The parameter b is mainly connected to the amplitude of polarization. The physically reasonable range for b is $[0, 1]$. With a choice of $\alpha_0 = 0^\circ$, $c_1 = 3$, $c_2 = 0$ and $b = 1$ we approximately get the polarization phase curve of a single Rayleigh particle with a maximum polarization of 1 at phase angle $\alpha = 90^\circ$. Parameter b also affects the slope of the phase curve at α_0 .

The parameter α_0 is the inversion angle. The phenomenon of a negative branch of polarization, which will turn positive in the neighbourhood of 20° , is common for Solar System dust where multiple scattering and interactions between the constituent grains inside aggregates play an important role. A physically reasonable range for α_0 is obviously $[0^\circ, 180^\circ]$, although in observations the inversion angle seems to stay below $\sim 30^\circ$.

The powers c_1 and c_2 have an influence on the shape of the phase curve. The parameter c_1 mainly affects the position of the minimum and the second derivative of the curve, while c_2 has influence on the maximum and on the asymmetry of the curve, moving the angle for maximum polarization away from 90° . These two parameters should have positive values.

The collection of these four parameters offers a wide variety of different, realistic shapes for phase curves. The effects of the parameters are studied in Fig. 1.

3.2. Model fitting by nonlinear Bayesian regression

The common problem with empirical phase curve fitting is that, especially for asteroids, there is a lack of good datasets. For a single asteroid, there might be just a few data points from a quite limited phase angle range. Phase angles above $\sim 30^\circ$ are rare, and the measurement errors can be noticeable. This heavily affects the goodness and reliability of the phase curve fit,

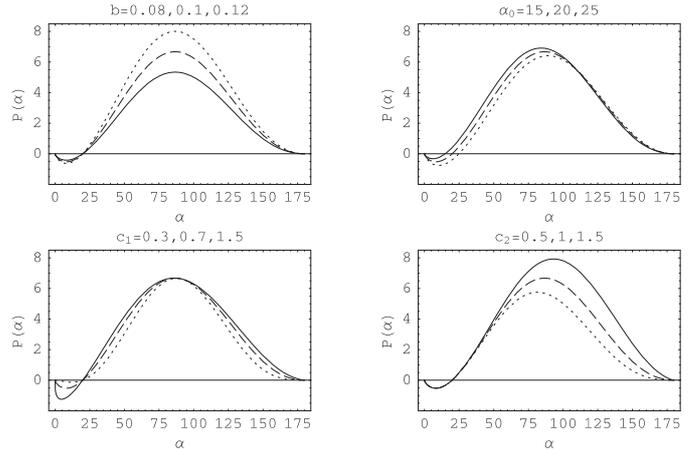


Fig. 1. The effect of the parameters on the polarization phase curve model (given in percents). In all figures only one parameter changes, which is denoted in the individual figure label. The solid curve is for the first value, dashed for the second, and dotted for the third. The other parameters have values $b = 0.1$, $\alpha_0 = 20^\circ$, $c_1 = 0.7$, $c_2 = 1$.

which is a nonlinear regression problem for our model. For the four parameters we need at least four observations, but in practice only datasets starting from, say eight or ten observations, can be considered useful. Furthermore, the model parameters are not totally independent from each other, so there is multicollinearity in the model, i.e. a unique solution to the model fit is hard to find.

In many nonlinear regression procedures, the problems mentioned can result in physically unrealistic estimates, for example negative values for the power parameters c_1 and c_2 . While the fit is usually good in that small phase angle range where our observations are, any kind of extrapolation from that range is dangerous, as polarization might have values above one etc. In some applications this is not a problem, but in our study we want to examine the polarization from the whole phase angle range and consider also prediction of polarization. In some regression procedures it might be possible to introduce limits for parameters, but in practice we have noticed that this is not always reliable.

A very elegant solution to the problems of the standard nonlinear regression is the framework of Bayesian analysis and regression, where the posterior distribution of model parameters is a product of the prior distribution of the parameters and the likelihood of the parameters given by the regression model. In Bayesian regression we can use our prior knowledge and physical limitations of the model parameters in the prior distribution. Physical limitations for parameter values are handled by using distributions that have probability densities greater than zero only in the reasonable parameter range. Furthermore, since we have been working with the asteroid and comet data for quite a while, we have acquired a good insight of the typical phase curves for different types of objects. Using this knowledge by assigning larger probabilities to the expected values of the parameters, we can guide the model fit towards a good, robust and realistic estimate. For a more detailed description of the Bayesian paradigm see e.g. Box & Tiao (1992).

Table 1. Hyperparameters for prior distributions for the phase curve model of asteroids and comets. The probability distribution function (PDF) of Beta distribution is $x^{\alpha-1}(1-x)^{\beta-1}/B(\alpha,\beta)$, and the PDF of Weibull distribution is $\alpha^{-\beta}\beta x^{\beta-1}e^{-(x/\alpha)^\beta}$.

parameter	distribution	hyperparameters for asteroids	hyperparameters for comets
b	Beta	$\alpha = 3.70,$ $\beta = 32.23$	$\alpha = 1.43,$ $\beta = 2.39$
α_0	Weibull	$\alpha = 0.42,$ $\beta = 2.69$	$\alpha = 0.38,$ $\beta = 2.38$
c_1	Weibull	$\alpha = 0.47,$ $\beta = 1.74$	$\alpha = 0.64,$ $\beta = 1.97$
c_2	Weibull	$\alpha = 1.41,$ $\beta = 2.20$	$\alpha = 1.02,$ $\beta = 1.56$

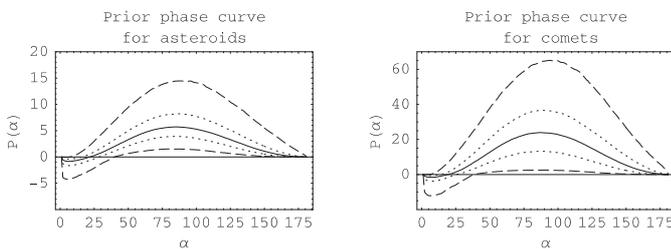


Fig. 2. Prior distributions for the polarization phase curves (given in percents) of asteroids and comets. The solid line is the median, dotted lines are the 50% confidence level, and dashed lines are the 90% confidence level. Confidence levels are empirical confidence levels, calculated by simulating observations of the parameter vector (b, α_0, c_1, c_2) from the prior distribution and using the simulated parameter values in the polarization phase curve model (Eq. (1)).

We choose to use a beta distribution as a prior distribution for parameter b , and Weibull distributions for α_0 , c_1 and c_2 . In theory, the distribution for α_0 should be limited to $[0^\circ, 180^\circ]$, but any reasonable choice for the shape of the Weibull distribution ensures that the probabilities are practically zero long before 180° . The parameters for the prior distributions, called hyperparameters, can be different for asteroids and comets, thus taking into account the basic differences in their polarization behaviour, mainly the greater values of polarization of comets, especially on the positive branch. After an intensive analysis we ended up with a suggestion for the hyperparameters, presented in Table 1. Prior distributions are always based on a subjective decision, and thus other choices for distributions and hyperparameters could be possible. The most important things to consider are that the expected shape of the phase curve (Fig. 2) reflects the overall behaviour of polarization observations, and that the prior model is sensitive to the actual data and will not dominate the estimate. We study this sensitivity further in Sect. 4.2. The actual regression procedure can only be done numerically. Bayesian regression models are estimated with Markov Chain of Monte Carlo (MCMC) simulation algorithms (see e.g. Gilks et al. 1995). We have used a free software package WinBUGS³ for this task.

³ BUGS project in the Internet, see <http://www.mrc-bsu.cam.ac.uk/bugs/>

3.3. Important features of polarization

There are a few important, widely used features of polarization that are interesting and are often reported in polarization studies. These include the inversion angle α_0 where the negative branch turns to positive, the values of maximum negative and positive polarization P_{\max} and P_{\min} , and the phase angles where these are reached α_{\max} and α_{\min} , and the slope s of the curve at α_0 . α_0 is included in our model, but other features have to be calculated from the estimated parameters. For the slope s there is a simple expression, $s = b \sin^{c_1}(\alpha_0) \cos^{c_2}(\frac{1}{2}\alpha_0)$. The parameters α_{\max} and α_{\min} can also be derived analytically with the help of a symbolic mathematical software package, but the resulting formulas are too complicated to be useful. In practice, P_{\max} , P_{\min} , α_{\max} and α_{\min} can be derived either numerically or analytically.

Within the Bayesian framework one can derive analytical results for only a limited family of parameter distributions, and we do not see any good reason to limit ourselves to that family of conjugate prior distributions. Instead, in our general case, the numerical MCMC-algorithm converges to sampling from the posterior distribution of model parameters and after converging can produce an unlimited number of samples. We base our further analysis on these samples. Our point-estimates for the model parameters are the medians of parameters sampled from the posterior distribution. Also error estimates and different confidence levels can be calculated from these samples. In the framework of standard nonlinear regression, the errors for different derived features of polarization could be estimated by the propagation of errors-method. In the MCMC-case, it is more straightforward to use the same posterior samples to numerically estimate the errors.

4. Applications to phase curve modelling

Next we present some case studies of the possible uses of the phase curve model in practice.

4.1. Phase curves for asteroid Juno and comet Halley

Figure 3 gives the estimated phase curve for the main belt asteroid Juno. We have a total of 46 observations of Juno's polarization with phase angles ranging from 2° to 28.7° . These observations have been made at different wavelengths from 333 nm to 952 nm, but the wavelength effect at these phase angles is weak, and the number of observations for each wavelength is too small, so we treat these data as one phase curve. We will discuss the wavelength effect later in Sect. 5. We can see from Fig. 3 that the fit is very good and accurate in the range where we have the data. However, the confidence intervals of the fit grow noticeably at large phase angles where we do not have any data. Nevertheless, it seems that some extrapolation can be done with a reasonable accuracy, for example to something like 60° . In Table 2 we present the key features of the polarization and their confidence intervals. We can notice that features which are supported by the data: α_0 , s , α_{\min} and P_{\min} can be estimated quite accurately, while the extrapolated features α_{\max} and P_{\max} have very limited accuracy.

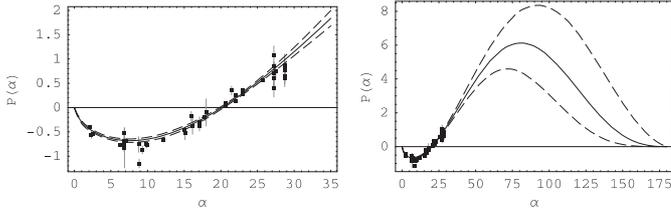


Fig. 3. Phase curve fit for Juno. In the left the fit is plotted for the range for which we have observations, and in the right for the whole phase angle range. The solid curve is the fit, and the dashed curves are the 90% confidence intervals. Observations are plotted as boxes and their measurement errors as lines. Polarization is given in percents.

Table 2. Polarization features for Juno. The 5% column is the lower limit of the 90% confidence interval, and the 95% column is its upper limit. Median is the actual estimate of the feature.

feature	5%	median	95%
α_0	19.8°	20.1°	20.4°
s	0.0895	0.0963	0.103
α_{\min}	7.20°	7.63°	8.05°
α_{\max}	71.6°	80.3°	88.8°
P_{\min}	-0.715%	-0.681%	-0.647%
P_{\max}	4.65%	6.08%	8.25%

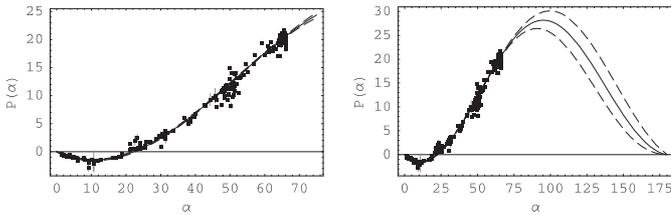


Fig. 4. Phase curve fit for Halley. Notations as in Fig. 3.

In Fig. 4 and Table 3 the phase curve and its features are presented for comet Halley (see references in Levasseur-Regourd et al. 1996). For Halley there are observations for phase angles up to 66°. In that range the wavelength effect can be noticeable, so instead of all wavelengths we take only red wavelengths, 630–760 nm. This gives us a total of 186 observations. Because of the larger phase angle range, the polarization near the maximum can be predicted with good accuracy.

Estimates for the measurement errors in polarization observations are given in the asteroid and comet databases. These errors are generally instrumental errors and the rotation of asteroids whose surface may be inhomogeneous (variegation effects) or the spatial variations in a cometary coma are not taken into account. For these reasons the errors are not realistic as deviations between the observations and a unique and smooth polarization phase curve of the object. Our analysis of Juno and Halley gives as an estimate that the deviations between observations and the fitted phase curve are approximately two times as large as the measurement errors for Juno, and 5.8 times as large as those for Halley.

Table 3. Polarization features for Halley. Notations as in Table 2.

Feature	5%	Median	95%
α_0	21.6°	22.2°	22.9°
s	0.234	0.246	0.259
α_{\min}	10.1°	10.7°	11.3°
α_{\max}	80.3°	84.9°	89.1°
P_{\min}	-1.59%	-1.47%	-1.35%
P_{\max}	26.4%	27.0%	27.4%

4.2. Prediction of polarization

We use a Bayesian approach for regression, because we need to use our prior knowledge of the phase curve parameters to obtain a good fit. With any kind of Bayesian analysis, though, one needs to confirm that the prior distributions do not dominate the posterior ones. This is especially important if we want to fit the model with just a few observations. If the prior distributions carry too much information, the fit is biased towards priors. To check the possible bias in the model, we do some simulation studies. We simulate a small number of observations from a phase curve that is different from the prior phase curve, and we fit the model. If we do this many times, we can see whether the fits are concentrated properly around the phase curve they were simulated from. In our case it seems that priors are flexible enough, and even a data set consisting of only five observations contains adequate information, and the fit is unbiased.

We can further investigate our model with simulation studies. From the observational point of view, it is important to have some idea of how the number of observations and the phase angles they are taken from affect the accuracy of the fit. This is extremely important if we want to predict or extrapolate polarization, or if we need to have a given accuracy for some polarization feature. We show here some results of the fit done with a different number of observations and phase angle ranges.

We focus on the prediction of a typical asteroid phase curve. We choose a phase curve that is slightly different from the prior distribution. This allows us to check at the same time the bias of the fit mentioned earlier. As we simulate observations from this phase curve, we add a typical measurement error calculated from the SMB asteroid data, and multiply it by two to obtain a realistic error. The multiplication by two was found realistic in the study of Juno in Sect. 4.1. By simulating observations and fitting the model over and over again, we can set confidence limits to the phase curve estimate. In Fig. 5 this is done in the case where we have observations only at phase angles below 30°. Figure 6 shows how the fit and its capabilities for accurate prediction are improved if we could have observations from the range [0°, 50°]. In that same figure there is also an example of a quite typical phase curve for a comet, and the confidence intervals of the situation where we have observations from the range [0°, 111°]. It can be clearly seen that if we want to predict the polarization near its maximum, then the phase angle range available for us is the most important factor for an accurate fit.

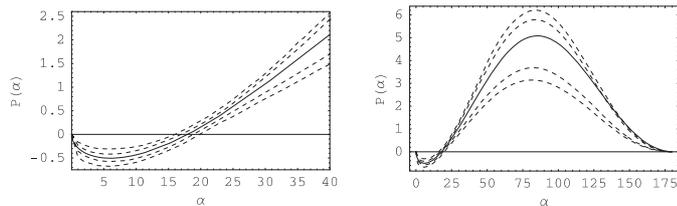


Fig. 5. Simulated 90% confidence intervals for the phase curve estimate, when 10 and 50 observations are done of the phase curve marked with a solid line. The observations are simulated from evenly distributed angles in the range $[0^\circ, 30^\circ]$. On the left there is the more detailed plot for that area, and on the right the plot for the whole phase angle range $[0^\circ, 180^\circ]$. The larger confidence interval (outermost dashed lines) is for 10 observations, and the smaller interval (innermost dashed lines) is for 50 observations.

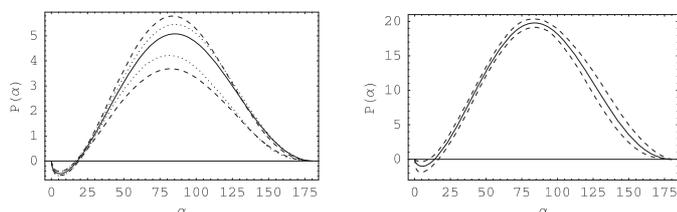


Fig. 6. The effect of a larger phase angle for the observations. On the left there is the 90% confidence interval for the phase curve estimate when 50 observations are simulated from the range $[0^\circ, 30^\circ]$ (dashed lines), and when the same number of observations is retrieved from the range $[0^\circ, 50^\circ]$ (dotted lines). On the right, there is a confidence interval for the estimate for a typical comet phase curve (solid line) when 50 observations are retrieved from the range $[0^\circ, 111^\circ]$.

The phase angle range of the observations has different effects on the four parameters in our model and on their accuracy. The shape parameters c_1 and c_2 are more sensitive. c_1 affects the curvature of the phase curve. If the observations can be done only with small angles, we have limited information on the curvature, and thus also a limited accuracy in the estimate of c_1 . The estimation of c_2 is even more difficult, since the main effect of c_2 is to make the phase curve asymmetric and define the angle of maximum polarization. If our observations are from such a range that the second derivative of the curve after the inversion angle is still positive in that range, then we have very poor information of the possible place of the maximum. Thus, of all four parameters, c_2 has the lowest accuracy.

We can also study the behaviour of our model with real observations instead of simulations. We choose to use our model for a subgroup of our comet data, the so-called high- P_{\max} comets (Levasseur-Regourd et al. 1996). In the data, comets Halley, Hyakutake and West belong to this group. These comets have similar polarimetric properties at large phase angles, so it is reasonable to fit a single polarization phase curve for all the comets in this group.

We use observations in the red wavelength domain resulting in 229 observations in the phase angle domain $[1.57^\circ, 111.3^\circ]$ for two experiments. In the first experiment, only the observations at phase angles up to 70° are used to fit the polarization phase curve. In the second experiment, only the observations

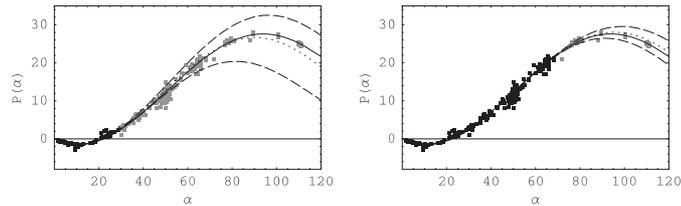


Fig. 7. Prediction of polarization for the high P_{\max} comets. On the left, there is the polarization phase curve fit using observations up to 30° (the black points). On the right is the fit using observations up to 70° (the black points). In both figures, the fit using all the observations (the black and the grey points) is marked with a solid line, the fit using the limited range of observations is marked with a dotted line, and its 90% confidence intervals with dashed lines.

at phase angles up to 30° are used. The results of these experiments are shown in Fig. 7, where the resulting fits and the corresponding confidence intervals are compared to the polarization phase curve fit using all the observations. In both experiments the fit is very close to the fit obtained by using all observations. The confidence interval for the polarization phase curve is significantly smaller at large phase angles for the fit using observations up to 70° , but the fit using observations up to 30° has also enough accuracy to be useful at phase angles up to, say, 60° – 70° .

4.3. Taxonomy of asteroids using polarization

Asteroid taxonomy is a classification system based on spectrometry. Asteroids are classified to classes which correspond more or less with the surface material of asteroid. There are several competing classification methods, and Tholen & Barucci (1989) describe these methods in more detail.

Polarization is also sensitive to the surface material of the object. While spectrometry is a far more powerful tool for material analysis, we should see some of the effects also in the polarization data. For example, Dollfus et al. (1989) show a relation between the polarization features α_0 and P_{\min} , and the asteroid class. Goidet-Devel et al. (1995) established that the phase curves are comparable for asteroids that belong to the same taxonomic type (e.g. S-type and C-type), and that there are strong analogies between the phase curves for asteroids of S, M, E type (so-called igneous) on the one hand, and C, G, P type on the other hand. Here we present some interesting results that also relate polarization to the taxonomy.

In the SBN database, there are data on 100 asteroids that belong to an identified taxonomic class, including the “unidentified” class X. We analyze the similarities in the polarization among the classes using clustering methods similar to Tholen’s taxonomy (Tholen & Barucci 1989).

We estimate one polarization phase curve for each class, obtaining 14 phase curves. For all curves we use only the range $\alpha \in [0^\circ, 30^\circ]$, since there are many classes where there are no observations beyond that. By using the fitted curve instead of just a few selected points, we use the observed information in a more efficient way. To classify the curves, we use a method identical to spectroscopic analysis: we discretize the curves

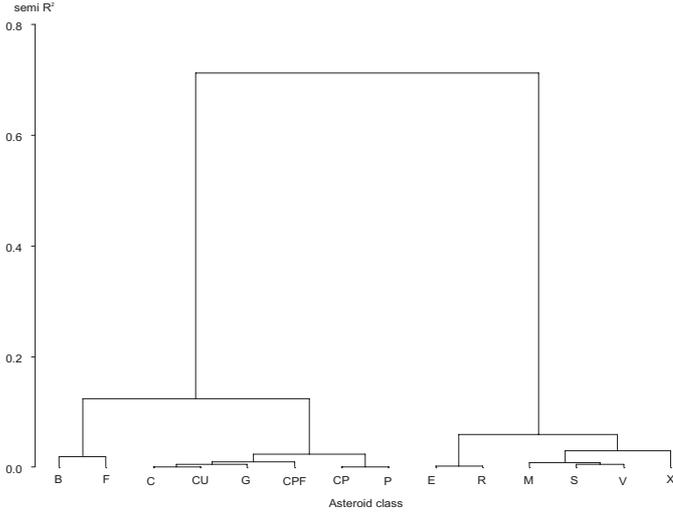


Fig. 8. Tree graph of the clustering analysis of the asteroid class polarization phase curves. The vertical scale describes how the variability inside the group (semipartial R^2) diminishes as the objects are divided into sub-groups.

with a fine grid, and use principal component analysis (PCA) to extract the most important features from the data. With only three principal components we are able to reproduce 99.8% of the variability in the data, which can be considered adequate. With these three PCA-variables for the 14 taxonomic classes, we use clustering analysis to find the relationships between classes. In Fig. 8 the result of the clustering is presented as a tree graph.

The tree-form describes the division of all classes into smaller groups based on their polarization. At the top level, all the classes belong to the same group. The first branch in the tree separates this group into two clearly distinguishable subgroups. This procedure continues until at the lowest level all the classes form their own group. At each division, a measure for the variation inside the groups will be minimized. The lowest level of the tree is not interesting, but the phases before that are, because they show the similarities between the classes. There are several competing clustering methods and measures for the inside-group-variation, and the results in Fig. 8 are obtained with Ward’s minimum-variance method and semipartial R^2 measure for the variation (SAS Institute 1999).

What is quite interesting is that the separation of the different groups closely reflects the historical evolution of the taxonomical system. At the second highest level where the classes are divided into two large groups, the first group consist of the C-class and its subtypes along with B, F and P, and the other group consists mainly of the S-type, together with E, R, M, V and X. Historically the C- and S-types are the first two classes. On the other hand, at the lowest levels classes and their subclasses are very close to each other: C and CU, CP and P, and S and V. This shows that the polarization phase curve carries similar information on the surface properties as the spectrometric observations.

5. Model for multi-wavelength polarization phase curves

The question of the wavelength dependency on asteroid and comet polarization is important. Studying it can be quite tricky, since in many cases the different errors, like the measurement error of the instrument or the error due to the evolution of the comet over observation time, can be of the same magnitude as the possible wavelength effect. As there are so many error sources, the phase curve fit should be as robust and reliable as possible. We extend here the model for one wavelength (Eq. (1)) to the case of multi-wavelength data.

Our approach is similar to multiple regression analysis with different groups and effects associated with the groups. The observations at different wavelengths will be divided into groups according to the filter, e.g. red wavelengths will form one group, blue wavelengths another etc. The (continuous) wavelength variable in the data will be replaced with a group indicator variable λ , a vector of size $n - 1$. An arbitrary group will be the “reference group” with $\lambda = \mathbf{0}$, other groups k from $k = 1$ to $k = n - 1$ will be coded with λ having components $\lambda_k = 1$, $\lambda_{\neq k} = 0$.

Our suggestion for the multi-wavelength polarization phase curve model is

$$P(\alpha, \lambda) = (b + \lambda \mathbf{b}^*) \sin^{c_1}(\alpha) \cos^{c_2} \left(\frac{\alpha}{2} \right) \sin(\alpha - (\alpha_0 + \lambda \alpha_0^*)), \quad (2)$$

where the parameters determining the wavelength dependency \mathbf{b}^* and α_0^* are also $n - 1$ sized vectors, and components b_k^* and α_{0k}^* indicate the difference to the reference group for the corresponding group k .

The most important advantage in our multi-wavelength model is that it uses all the observations of one object in one model using different wavelengths as groups inside the model. One wavelength is the reference group, and the same parameters as in the one-wavelength model describe the phase curve for that group. For other wavelengths, the parameters b_k^* and α_{0k}^* introduce deviations from that “main” curve. By using suitable prior distributions for parameters \mathbf{b}^* and α_0^* we can force the phase curves for different wavelengths to have similar shapes. This is a very important feature because we know from practice that phase curves at different wavelengths of the same object do correlate strongly. If we were to use completely separate modelling for each wavelength, we would lose quite a lot information, and that would make the already difficult analysis of the wavelength effect even more cumbersome.

We decided to use multiple parameters only for b and α_0 , and not for c_1 and c_2 . This is because the shape parameters c_1 and c_2 are the most difficult to estimate, and in most cases the shapes of different wavelength phase curves are very similar. Thus, the wavelength effect is only studied through the inversion angle α_0 and the magnitude parameter b .

For all parameters other than the vectors \mathbf{b}^* and α_0^* , the prior distributions and their hyperparameters are the same as in the one-wavelength model. The prior distributions for both vectors are uncorrelated multinormal distributions with zero expectations and small standard deviation for the components, $N_{n-1}(\mathbf{0}, \sigma^2 \mathbf{I})$. This means that the expected shape of the phase curve is the same for all wavelengths, and the small standard

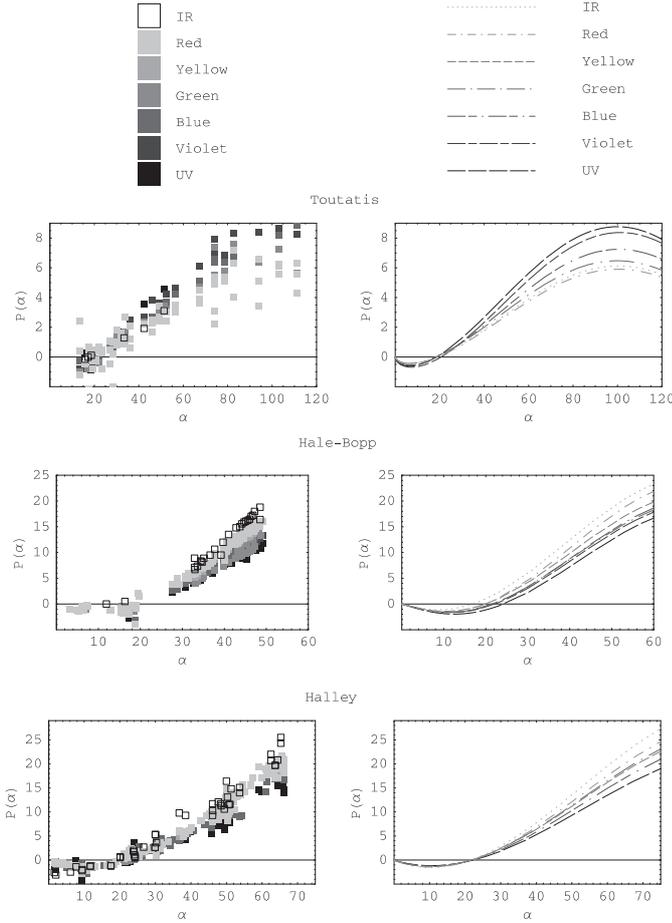


Fig. 9. The observations (*left column*) and the fitted model (*right column*) for Toutatis, Hale-Bopp and Halley. At the top, there is an explanation of symbols and line types.

deviations in the prior will keep the posterior phase curves well correlated, unless the observations strongly indicate otherwise. In our analysis all the components in the magnitude vector \mathbf{b}^* have standard deviations σ of 0.015 for asteroids and 0.03 for comets. The components in the vector α_0^* have standard deviations σ of 2° for asteroids, and 3° for comets.

6. Applications to multi-wavelength phase curve modelling: Analysis of the wavelength effect on polarization

We have included in this analysis the comets Hale-Bopp, Halley, Hyakutake, Swift-Tuttle and West, because these objects can be studied in different wavelengths with data at phase angles larger than 40° . For asteroids, the only good dataset for this purpose is available from Toutatis. When asteroid datasets are divided into different wavelengths, there are usually just a few observations per wavelength. Furthermore, these observations tend to be at small phase angles, where the wavelength effect is nearly inexistent between the blue and green domains compared to the measurement errors. There are 174 observations of Toutatis ranging from 13.55° to 111° , which makes it possible to use it in the analysis.

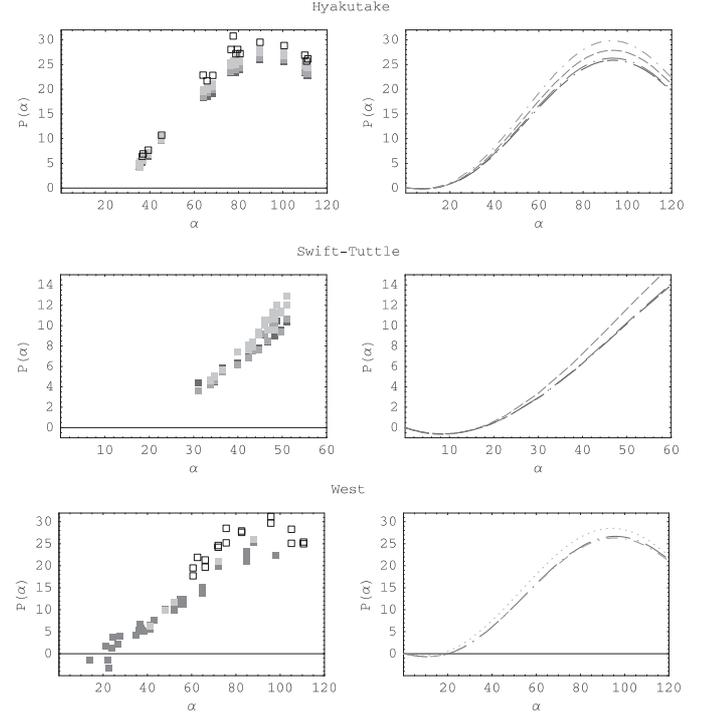


Fig. 10. The observations (*left column*) and the fitted model (*right column*) for Hyakutake, Swift-Tuttle and West. For symbols and line codes see Fig. 9.

The wavelength filters used in polarimetric measurements can be slightly different between instruments, and also it seems possible that the practice of reporting the filters can vary inside the database from one observer to another. The filters for cometary measurements have to be adapted to avoid the gaseous emission lines and, when possible, narrow band filters are used and the results eventually corrected by the knowledge of the emission spectrum. Therefore, we have used a rougher division and grouped the wavelengths into visible colours violet (380–450 nm), blue (450–490 nm), green (490–560 nm), yellow and orange (560–630 nm) and red (630–760 nm), and into UV (below 380 nm) and IR (760–2200 nm).

Figures 9 and 10 represent the observations and the fitted multi-wavelength models for our data. Our model seems to behave rationally, the phase curves for the different wavelengths tend to be roughly the same, and the differences are quite logical, for example the curves are more or less ordered according to the wavelength.

From the fitted models in Figs. 9 and 10 we can also have the estimates and confidence intervals for polarization features as a function of the wavelength. There are some notable wavelength effects, but in most cases these effects are not similar for all the comets, which implies that these are not global effects, but only applicable for that particular comet. An example of this is in Fig. 11, where the wavelength effect for Hale-Bopp is presented for polarization features α_0 , α_{\min} and P_{\min} . As seen in Fig. 11, Hale-Bopp shows a clear wavelength effect in the inversion angle. On the other hand, the inversion angle for comet Halley is practically the same for all the wavelengths. Comets Hyakutake and Swift-Tuttle have no observations near the

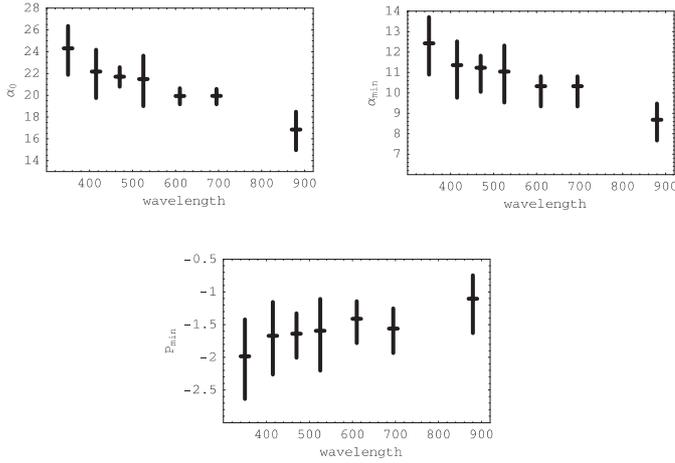


Fig. 11. Wavelength effects for comet Hale-Bopp in α_0 , α_{\min} and P_{\min} . The vertical lines are the 90% confidence intervals, and the horizontal bars are the point-estimates.

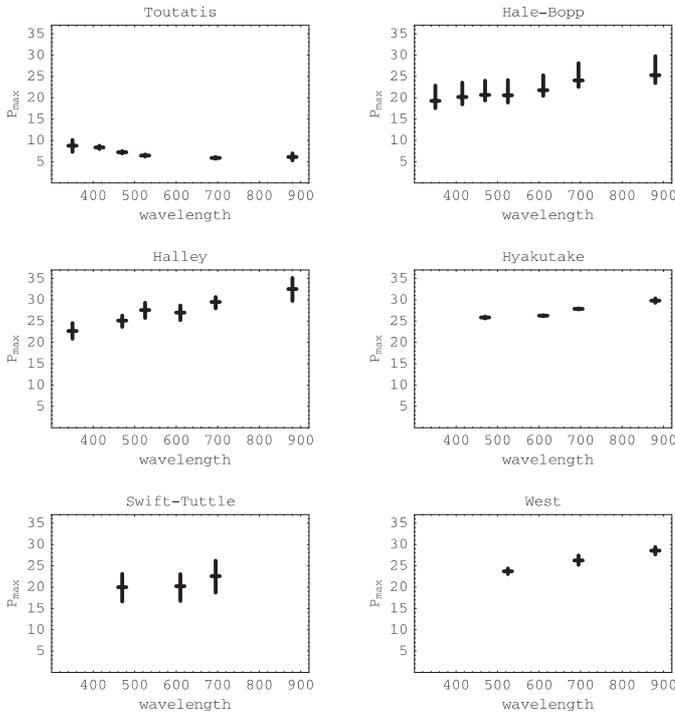


Fig. 12. Wavelength effects for all objects in maximum polarization P_{\max} . The vertical lines are the 90% confidence intervals, and the horizontal bars are the point-estimates.

inversion angle and therefore it is impossible to draw conclusions about the wavelength effect in the inversion angle, and the same applies for comet West, for which there is a lack of data in some wavelengths.

The only clear effect is seen either in the minimum value of polarization, or much clearer in the maximum value of polarization (P_{\max} in Fig. 12). For all the comets, P_{\max} increases with wavelength, and an inverse effect is observed for the S-type asteroid Toutatis.

7. Discussion

The dataset used here provides fundamental information from polarimetry of asteroids and comets. The two-dimensional (phase and colour) study permits a comparison between the objects in different asteroidal and cometary classes and between the two kinds of objects. The phase curves and their features in different colours have to be compared to those obtained by computation of modelled particles (regoliths and clouds) to improve our knowledge of the physical properties of the dust

As previously discussed by Goidet-Devel et al. (1995), the taxonomy using polarimetry provides a complementary approach for classification and could be used in tandem with spectroscopy. In a way it is rather surprising that such a small phase angle range from 0° to 30° , as is the case for most asteroid, can produce taxonomic classes so close to the spectroscopic taxonomy. This is a good example of a situation where large datasets are valuable. In the clustering classification scheme we have not derived any quantitative errors for categorizing a given object into a certain class, but we are interested in carrying on in future some further analysis of the possibilities and quantitative errors of using only polarimetric data to categorize an object into a taxonomic class.

Multi-wavelength study makes it possible to improve the fits of data sets, eventually extrapolated to a larger or smaller phase angle range for comparison with other objects and finally to approach the physical properties of the dust by remote observations.

Acknowledgements. We would like to thank Karri Muinonen for his comments and the Magnus Ehrnrooth foundation for their financial support to the first author.

References

- Box, G. E. P., & Tiao, G. C. 1973, Bayesian Inference in Statistical Analysis (Addison-Wesley Pub Co.)
- Dollfus, A., Wolff, M., Geake, J. E., Lupishko, D. F., & Dougherty, L. M. 1989, in Asteroids II, ed. R. Binzel, T. Gehrels, & M. Matthews (The University of Arizona Press), 594
- Dones, L., Cuzzi, J. N., & Showalter, M. R. 1993, *Icarus*, 105, 184
- Gehrels, T. 1974, *Planets, Stars & Nebulae studied with Photopolarimetry* (Tucson: University of Arizona Press)
- Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. 1995, *Markov Chain Monte Carlo in Practice* (CRC Press)
- Goidet-Devel, B., Renard, J.-B., & Lvasseur-Regourd, A.-C. 1995, *Planet. Space Sci.*, 43, 779
- Hadamcik, E., & Lvasseur-Regourd, A.-C. 2003a, *A&A*, 403, 757
- Hadamcik, E., & Lvasseur-Regourd, A.-C. 2003b, *Icarus*, 166, 188
- Hadamcik, E., & Lvasseur-Regourd, A.-C. 2003c, *JQSRT*, 79, 661
- Hadamcik, E., Renard, J.-B., Lvasseur-Regourd, A.-C., & Worms, J.-C. 2003d, *JQSRT*, 79, 679
- Hapke, B. 1993, *Theory of reflectance and emittance spectroscopy* (Cambridge: Cambridge University Press)
- Jockers, K. 1997, *Earth, Moon, Planets*, 79, 221
- Kiselev, N. N., Rosenbush, V. K., & Jockers, K. 1999, *Icarus*, 140, 464
- Kiselev, N. N., Jockers, K., & Rosenbush, V. K. 2002, *Earth, Moon, Planets*, 90, 167

- Kiselev, N. N., Rosenbush, V. K., Jockers, K., et al. 2003, *ESA-SP*, 500, 887
- Levasseur-Regourd, A.-C., Hadamcik, E., & Renard, J.-B. 1996, *A&A*, 313, 327
- Levasseur-Regourd, A.-C., & Hadamcik, E. 2003, *JQSRT*, 79, 903
- Lumme, K., & Muinonen, K. 1993, A two-parameter system for linear polarization of some solar system objects. *Asteroids, Comets, Meteors*, IAU Symp., 160, 194
- Lumme, K., Green, K., Muinonen, K., & Penttilä, A. 2003, In Proc. of the 7th Conference on electromagnetic and Light Scattering by Nonspherical Particles: Theory, Measurements and Applications, ed. T. Wriedt (Germany: Bremen), 220
- Mishchenko, M. I. 1993, *ApJ*, 411, 351
- Muinonen, K. 2004, *Waves in Random Media*, 14, 1
- Muinonen, K., Piironen, J., Shkuratov, Yu. G., Ovcharenko, A., & Clark, B. E. 2002, in *Asteroids III*, ed. W. F. Bottke, A. Cellino, P. Paolicchi, & R. P. Binzel (University of Arizona Press), 123
- Mukai, T., Iwata, T., Kikuchi, S., et al. 1997, *Icarus*, 127, 452
- SAS Institute 1999, *SAS/STAT User's Guide*, Version 8 (SAS Publishing)
- Seeliger, H. 1887, *Abhandlung der Bayerische Akademie Wissenschaftliche Mathematik-Naturwissenschaft, Klasse II*, 16, 405
- Shkuratov, Yu., Ovcharenko, A., Zubko, E., et al. 2004, *Icarus*, 159, 396
- Tholen, D., & Barucci, A. 1989, in *Asteroids II*, ed. R. Binzel, T. Gehrels, & M. Matthews (The University of Arizona Press), 298