Asteroid family classification from very large catalogues

Anne Lemaitre
Department of Mathematics, University of Namur
Rempart de la Vierge, 8 - B-5000 Namur Belgium
email: anne.lemaitre@fundp.ac.be

Abstract. The paper presents a review of the recent contributions and open questions concerning the families of asteroids. Due to the availability of very large catalogues (synthetic and analytical proper elements of the asteroids and large observational surveys of their spectra) and to the introduction of non gravitational forces in their determination, the concept of static family has disappeared, to be replaced by this of dynamical families. The proper elements are not constant anymore but are ageing on very long timescales. The size distributions of the populations of asteroids, in and out the families, their ages, the ejection velocities of the fragments after an impact, have been reconsidered by several teams of research, with this new approach. Parallel numerical simulations of collisions and fragmentations of bodies have showed that most of the asteroids are likely rubble piles or agglomerates than monolithic blocks. The methods of classification have been refined and combine, in their newest versions, the dynamics and the observations, working now on 5 dimensional space instead of 3. A series of sub families of the large well-known families have been recently identified, using catalogues with more than 100 000 asteroids (the cluster Karin for example).

Keywords. Asteroid family, proper elements, surveys, catalogues, micro family.

1. Introduction

The so-called families of asteroids were detected by Hirayama (1918), on a set of a few hundreds of planets. The problem of the missing planet between Mars and Jupiter seemed then to have found a partial answer; several quite big bodies had first the opportunity of growing up in this region, but repetitive collisions between them or systematic bombardments led to their fragmentation, and reduced them, chock after chock, to bodies with a much smaller size and volume. The asteroids are obviously a population of bodies whose evolution is deeply and extensively influenced by collisional processes, and families are direct proofs that these collisions really occur in the asteroid belt. However their origin is not a single large body destroyed by collisions.

Up to the last decade, and before the new large catalogues, the research in that particular field was very well organized, in three separate steps, concerning different people and knowledge. Let us summarize these three sequential steps and their main past characteristics:

• The calculation of proper elements for the asteroids.

Based on a double elimination of periodic terms, of short (orbital scale) and long (secular scale) periods, the purpose is to calculate three invariants of motion, supposed to be very close to each other for the members of a same family. This calculation is very technical; after a pioneer work of Kozai (1962) let us mention two classical approaches: a completely analytical one, developed by Yuasa, Knežević and Milani (see Yuasa 1973,
and Milani and Knežević 1994), and a semi-numerical one, introduced by Williams (1989) and adapted by Lemaitre and Morbidelli (1994). Specific regions, crossed by mean motion or secular resonances require a special treatment, by adding the resonant angle to the slow variables. The dynamics is therefore more complex; the definition of the corresponding invariants of motion is purely local, and not compatible with the non resonant one. The work of Schubart for the Hilda’s (jovian resonance $3/2$), is summarized on the website http://www.rzuser.uni-heidelberg.de/s24/hilda.htm. Let us also mention the paper of Morbidelli (1993) on the calculation of secular proper elements.

- The application of a clustering method.

On the three dimensional set of proper elements and taking into account a local background density of population, clusters and groups of asteroids are identified and considered as candidates for a common origin, unless the mineralogical compositions show incompatibilities. Two approaches were developed and compared: a Hierarchical Clustering Method (HCM) developed by Zappala et al. (1990) and a Wavelet Analysis Method (WAM) due to Bendjoya et al. (1991). The principle of HCM is the use of a specific and adequate metric in the proper elements space to search the nearest neighbour of each body:

$$d = na(C_a(\delta a/a)^2 + C_e(\delta e)^2 + C_i(\delta \sin i)^2)^{1/2}$$

where $C_a = 5/4$, $C_e = 2$ and $C_i = 2$ is a possible and usual choice for these coefficients; $d$ is given here in $m/sec$, represents the “distance”, while $\delta a$, $\delta e$ and $\delta \sin i$ are the differences in semimajor axis, eccentricity and sine of inclination between the two bodies.

A diagram called dendrogram connecting all the objects with increasing distances is drawn, the results are given in the form of stalactites. Systematic comparisons with stalactites obtained from quasi-random simulated populations are performed and a cut off of about 100 m/sec is used to identify the main families, on a sample of a few thousands objects.

WAM idea is to superimpose a network to the set of data (in 2 and 3 dimension proper elements space); at each node of the network, a wavelet coefficient is calculated, as an indicator of local increase in density in the vicinity of the node. Comparisons with a random distribution of proper elements (with the same local density) are performed, and the significant coefficients are used to model the cluster responsible for these values. Bendjoya et al. (1993) showed that both methods (HCM and WAM) were able to identify families (even with potentially odd structures like filaments in proper elements space): they used simulations in which synthetic families were put inside background populations of increasing spatial density.

- The spectral analysis of the candidates to a common origin.

From the beginning, the specialists in proper elements and families have claimed that it was necessary to test the mineralogical composition of the different bodies, to be sure of their common origin. A catalogue of spectra was then elaborated by the observers, mapping the position, brightness, and the 5-color CCD photometry of the concerned objects (see Cellino et al. 2002 for references).

A classification of the asteroids was proposed by Tholen (1984) on a sample of 405 objects; 14 classes were distinguished, called: C (for dark carbonaceous objects), S for Silicate, A, B, D, F, G, T for some spectrally distinct curves, E, M, P, three classes reduced to X in the absence of the albedos, and 3 classes with a unique object : Q (Apollo), R (Dembowska) and V (Vesta).

The definition of a family, given in the glossary of Asteroids III in 2002 corresponds to this static and sequential approach of the phenomenon : a family is a statistically
significant cluster of asteroids in proper elements space that may share a common origin perhaps by the collisional disruption of a larger body.

The situation has evolved quite rapidly, thanks to the availability of very large catalogues of asteroids, and spectacular new surveys, in parallel with several bright theoretical results. The proper elements are not really constant, once non gravitational forces are taken into account, the clustering methods have to mix dynamics and observations so to recover obvious new families (since the recent available results of large surveys). Moreover, as a consequence of the use of CCD cameras, a new taxonomy is introduced for the classification, more adapted to the data.

In the following sections those important contributions of the last years in each of the mentioned fields, are listed, still considered as independent. In the last section I shall show how all these features are related to each other and should be considered together in the future.

2. The new synthetic proper elements

A new systematic way of calculating the proper elements has been developed by Knežević and Milani (2003); it consists of a purely numerical procedure, based on time integrations over periods of 2 Myr (10 Myr for some cases). The elimination of the short-periodic effects is performed through digital filters. A Fourier analysis allows to identify proper modes. A whole portrait of the belt can be drawn, with systematic and individual checking of the accuracy (standard deviations and maximum excursions). The maximum Lyapunov characteristic exponent is given so to measure the chaotic diffusion. The relative velocities at breakup are estimated with an error of about 5 m/s (through the Gauss equations - see for example Zappala et al. 1996) and the accuracy has been globally improved by more than a factor 3.

Full information is available on the site http://hamilton.dm.unipi.it/cgi-bin/astdys/astibo, with complete catalogues regrouping more than 100 000 planets, with analytical and synthetic proper elements.

Proper elements are also provided (using the synthetic technique) for very specific groups like the Trans-Neptunian Objects (TNO) or the Trojans.

3. The Trojan and Hilda specific proper elements and corresponding families

Besides the synthetic catalogues, an interesting semi-analytical work has been developed for the Trojans by Beaugé and Roig (2001). They provided a set of proper elements for the two Lagrange points, L4 and L5, and identified several obvious families (or groups) about L4, while the situation about L5 is less structured, with only one or two potential groups. Let us also mention the agreement with the purely numerical approach of Burger et al. (1998).

For the Hilda group, a new semi-numerical and promising approach has been proposed by Miloni et al. (2004).

4. The surveys

Many more data concerning lightcurves are now available, due to systematic surveys of regions of the sky. Besides the Space Watch Survey and several Subaru japanese surveys, let us describe three of the most often quoted in the search for families:

• SMASS : Small Belt Asteroid Spectroscopy Survey divided in two phases : SMASS I (Xu et al. 1995) and SMASS II (Bus and Binzel 2002a); they catalogue the spectra of
1447 asteroids of small and medium size, mainly from the inner main belt, or from the NEA population for SMASS I, and from the middle-region of the belt, for SMASS II.

- S3OS2: Small Solar System Objects Spectroscopy Survey: Lazzaro et al. (2001) with more than 734 objects
- SDSS MOC: Sloan Digital Sky Survey Moving Object Catalog (Ivezic et al. 2000 or Stoughton et al. 2002) which has already covered one eighth of the entire sky, which means 125 283 objects (amongst them, 35 401 known asteroids were identified).

5. The official families

The specialists of the clustering methods (see Zappala et al. 1995) carried out a new comprehensive search for families based on the largest catalogues ever used for this purpose, and using both methods of identification developed by Torino and Nice teams. A review of these methods and the results of these calculations is also summarized in Bendjoya and Zappala (2002).

The simulations were based on a much larger set of proper elements (12 487 minor planets) than in their first comparison. 20 official families were identified with the criterion of sharing at least 75% of objects, and less than 10 families could be added to this list if we accept a criterion of 50% of common members for the two theories. Except for a few exceptions, those families are not very different from the previous ones; the number of planets in each family is just much larger, gathering a quasi constant percentage of the total number of objects (about 35% of the whole asteroid population belong to families).

6. The taxonomy by Bus

Based on the SMASS survey, and due to the intensive use of CCD cameras, a new taxonomy was recently proposed by Bus (1999) or Bus and Binzel (2002b) based on 1447 asteroids, and respecting the following properties:

- It is still connected to Tholen’s classification, easy to use and applicable by others
- It is based only on spectral features and on multivariate analysis
- It shows agreement with natural groups, for the sizes and boundaries of the taxonomic classes
- It provides 3 spectral principal components: the spectral slope, and two others, mixing the data coming from the five colors observations with different weights (see Bus and Binzel 2002a for complementary information).

The result gives 26 classes, in which we find Tholen’s classes; new subdivisions of the classes C, S, X appear (C, Cg, Cgh, Ch for example). This taxonomy is already adopted by other groups, as shown by the results of Carvano et al. (2003) or Mothé-Diniz et al. (2003).

7. The background and family density of population

A very significant discussion for the understanding of the families concerns the densities of population inside and outside the families. Since official families have been identified, the size distribution of families have been analyzed and similar curves have been drawn for the background population of asteroids (objects not belonging to a family).

These analyses (for example, Zappala and Cellino 1996) concluded that families likely dominated the asteroid inventory, due to their steep size distribution. A simple extrapolation law was used to extend this result to very small objects, not yet observed, and it tended to prove that most of the objects should belong to families. Different authors
completed this theory by adding new elements to the discussion: for example, the geometry of the family offers an explanation of this steep distribution of sizes (see Tanga et al. 1999).

A possible explanation was recently given in a paper of Morbidelli et al. (2003); the idea is to debias the distribution of families relative to that of the background, using the data coming from the new surveys as calibration parameters. Their conclusion is that, after treatment of the samples by this procedure, for the objects with a magnitude $H > 13$ (which roughly means a radius less than 10 km), most families have a magnitude distribution even shallower than that of the background. The population of 1 km size objects could then be dominated by background bodies and not by families members.

8. The collisions and the fragmentation phase

The birth of a family is usually due to a collisional event, with another asteroid or a comet. This kind of mechanisms were simulated in laboratories, where projectiles of different size and compositions were crashed on meter size rocks. The fragments were described, followed, classified, the ejection velocities were measured and several scenarios were build, following all these features. At that point occurs a problematic step: are we sure that everything really observed in laboratories corresponds effectively to the reality in the space, after application of a simple scale factor? An affirmative answer means that the fragmentation models do not depend on the size of the bodies but only on their relative scales. The values obtained by this technique for the fragment dispersion velocities were systematically smaller than the ones calculated by the theory (for a review, see Holsapple et al. 2002).

Fortunately, the collisional dynamics was considerably improved and understood those last years, in particular through the work of Michel and collaborators (Michel et al. 2001, 2002 and 2003). These authors performed a very technical and complete numerical simulation of collisions, introducing a fragmentation phase in the process, thanks to an hydrocode (software 3DSPH) inspired by the fluid dynamics concepts. They simulated the dispersion and reaccumulation of the fragments (1 km size minimum) after a collision. The process also allows diffusion, and for an asteroid size body, the catastrophic event (from the impact to the end of the partial reaccumulation) takes a few hours or days.

Tests have been made on the classical families Eunomia, Koronis, Flora and for the new cluster Karin.

We can summarize the conclusions as follows: the reaccumulation process differs from one case to the other one, due to the internal structures and masses of the bodies. For Eunomia (284 km size body) the largest fragment after reaccumulation has reassembled 70% of the initial mass before impact. For Flora (154 km size body) the percentage is of 56%, and for Koronis (119 km size body) it is only of 4%. It seems quite obvious that the largest family members must be made of reaccumulated fragments and be closer to rubble piles than to monolithic rocks. Again, as in the laboratories experiments, the ejection velocities are much smaller than the 100 m/sec predicted by independent estimates of Dell’Oro et al. (2004).

9. The ejection of the fragments and the ageing of the proper elements

As already mentioned in the previous sections, the ejection velocities of the fragments after an impact seem to be sur-estimated by the theories, in comparison with the data coming from the laboratories experiments and the hydrocode numerical simulations.
The theoretical calculation of these velocities refers to papers of Zappala et al. (1996) and Zappala et al. (1997) and aims to rebuild the velocities field at the impact with the present orbital motions of the members of the family. The method was based on the likely existence of an axial symmetry.

The first step is the determination of the proper elements of the different members, the second one is the calculation, in the proper elements space, of the parameters \( \delta a, \delta e \) and \( \delta i \) (respectively the differences in semi major axis, eccentricity and inclination between the member and the estimated center of mass of the group). The third step is to replace the quantities \( \delta a, \delta e \) and \( \delta i \) in the lefthand side of the Gauss equations, so to extract, from the righthand side, the three components of the velocity, tangential, \( V_T \), radial \( V_R \) and perpendicular \( V_W \). The principle is simple, except that the equations also depend on \( f \), the true anomaly, for \( V_T \) and \( V_R \) and on \( f \) and \( \omega \) (the argument of pericenter), for \( V_W \). The proper elements sets is three dimensional; the values of \( f \) and \( \omega \) at the impact time, are not known.

They are then determined through an optimization process, so to maximize the axial symmetry of the reconstructed field. The results (the reconstructed fields) were convincing, except that the ejection velocities took very high values, often greater than 100 m/sec.

Different authors completed this theory, for example (Cellino et al. 1999) and (Pisani et al. 1999).

However, these classical studies consider that proper elements are constants for ever; it is now commonly accepted that it is not the case. Several features could play a role in the process. Dell’Oro et al. (2004) introduced a systematic noise on the proper elements, so to simulate this ageing of the constants of motion; they re-applied the same technique for the fragment velocities, corrected the estimators and obtained different results, in better agreement with the data.

10. The non gravitational forces and evolving families

In the paper of Dell’Oro et al. (2004) a list of effects potentially responsible for the ageing of the families is given :

- The intrinsic noise in proper elements
- The presence of resonances
- The incompleteness of the families (there are biases in the observations)
- The close encounters with large asteroids like Ceres
- The Low-velocities collisions between asteroids
- The presence of non gravitational forces, like Yarkovsky
- The family structure itself

The non-gravitational forces, combined with resonances, have already proved their efficiency; a series of papers from Vokrouhlický et al. (2001), Nesvorný et al. (2002a) and Bottke et al. (2001) showed how the addition of the thermal Yarkovsky effect could affect the stability of the proper elements.

Let us recall the principle of the Yarkovsky force : the surface of an asteroid is heated by the Sun during its day and cools off during its night; the asteroid emits more heat from its afternoon sight. This unbalanced thermal radiation produces a tiny acceleration and acts on any asteroid with a diameter smaller than 20 km. The numerical simulations make use of synthetic families (less dispersed than the real ones by a factor 3), and do not consider planets with diameters D smaller than 2 km. They are based on isotropic ejection fields of velocities (not exceeding 100 m/sec at infinity).
The following families were selected by the authors mentioned above: Flora, Maria, Eunomia and Koronis for the S-Type, Themis and Dora for the C-type; the thermal parameters are chosen with respect to the taxonomy because they are linked to the conductivity of the matter. Integrations were performed over periods of time between 300 and 500 Myr, with the integrator SWIFT of Levison and Duncan (2000), with the perturbation of 4 outer planets (except for Flora, where a more complete set of planets was used). Numerical proper elements were calculated for each member of the synthetic family. The results show an excellent concordance between the simulations and the expected positions of real bodies. The role of the resonances, even the three-body ones, appears clearly, cutting some families in a sharp way and exciting the eccentricities (see also Tsiganis et al. 2003). Some of the most spectacular figures can be found on the following web sites:

- http://www.boulder.swri.edu/davidn/yarko/yarko.html

The Yarkovsky effect introduces a kind of random walk, pushing the objects in or close to the resonances; it allows the members of the families to diffuse in the proper elements space, giving a non uniform spreading shape to the cluster. The concept of dynamical families, in opposition with static, is really introduced here. The results nicely reproduce the proper semi-major axis distributions, but a fit of the eccentricities and of the inclinations of the members of the different families is more difficult to achieve.

With the same philosophy, the close encounters with the largest asteroids (like Ceres) were tested for some families, like Gefion and Adeona (more concerned, because of their relative position) by Carruba et al. (2003). Their simulations showed a slightly modification of the proper elements due to Ceres, but nothing comparable with Yarkovsky effect, for which they reproduced similar qualitative results as the previous studies, even with different parameters and adjustment of constants.

The Yarkovsky effect was measured for the first time thanks to several radar observations of the NEA asteroid 6489 Golevka. This 0.5 km size asteroid was observed in 1991, 1995, 1999 and 2003, and a shift of 15 km on the semi major axis was clearly detected on this period of 12 years by Chesley et al. (2003).

This gives a first real measurement of the Yarkovsky effect on the semi-major axis: more than $10^{-4}$ AU per Myr, for a 1 km size body, which is, as mentioned by Dell’Oro et al. (2004), smaller than the values used by Bottke et al. (2001) or Morbidelli and Vokrouhlický (2003), but larger than those by Spitale and Greenberg (2001) or Carruba et al. (2003).

11. A new cluster : Karin

A new cluster was discovered by Nesvorný et al. (2002b) using the numerical integration methods mentioned previously, including Yarkovsky forces; they started from a subtle filament-like structure appearing in the Koronis family, identified by looking at the results obtained by the extended database of asteroids proper elements. It is called Karin and could be composed of 39 potential members, with confirmation for 13 of them.

The family contains two large bodies, respectively of 19 and 14 km size: Karin (832) and 1990FV (4507,) and a few others in the range of 2 to 7 km size. Backwards integrations of the 13 members of the family converged to one single orbit, this of the pre-breakup parent object. A collision between an initial body of 25 km with another one of 3 km, at about 5 km/sec, could be at the origin of the breakup, 5.8 Myr ($\pm$ 0.2 Myr) ago. Karin is then a young family and represents a very interesting event, especially for space weathering, because it is quite recent event, precisely dated.
12. The new families

A very convincing approach to the classification has been introduced very recently by Nesvorný et al. (2004). This is a real interaction between dynamical proper elements, mixed clustering methods, taxonomic components coming from new surveys and large catalogues, so to get a sample of 106,264 minor planets in which 40 families are clearly identified, regrouping 38,625 asteroids (about 36.3% of the whole population).

The authors use the data coming from the SSDC MOC survey (more than 125,000 objects), classified in the 5 colors bands (u, g, r, i, z). They run the PCA algorithm (Principal Component Analysis), to calculate (among all the combinations of the 5 possibly correlated variables), the 2 principal components, maximizing the de-correlation. They obtain the following linear combinations:

\[
PC_1 = 0.396(u - g) + 0.553(g - r) + 0.567(g - i) + 0.465(g - z)
\]

\[
PC_2 = -0.819(u - g) + 0.017(g - r) + 0.090(g - i) + 0.567(g - z)
\]

(PC standing for Principal Component).

They generalize the notion of distance used in the HCM method; the distance between two objects A and B is now calculated in a 5 dimensional space and is given by the expression:

\[
d(A, B) = na(C_a(\delta a/a)^2 + C_e(\delta e)^2 + C_i(\delta \sin i)^2 + C_{PC}(\delta PC_1)^2 + C_{PC}(\delta PC_2)^2)^\frac{1}{2}
\]

always expressed in \(m/sec\) where \(C_a = \frac{5}{4}\), \(C_e = 2\), \(C_i = 2\) and \(C_{PC} = 106\).

The cut off values are chosen between 0 and 150 \(m/sec\). The numerical coefficients are given here as examples; other choices are also relevant. Depending on the chosen values, large families are re-discovered or smaller sub-structures appear; for example, with a cut off of 23 \(m/sec\), Koronis is still a unique block, but with 12 \(m/sec\), Karin is identified. Of course, to distinguish the sub-structures of the main families (the micro families) a large number of bodies is necessary and such studies have to be performed with a minimum of 100,000 bodies to give significant results.

13. The age of the families and the space weathering

Concerning the attempts to estimate the age the families, let us recall the precursor approach concerning the family of Veritas, performed by Milani and Farinella (1994), using a chaotic chronology. For the first time, the formation of a family appeared as a very recent event: between 50 and 100 Myr for the first approach, confirmed by Knežević and Pavlović (2002). Following Nesvorný et al. (2004) we can list a short review of past and present methods for dating a family: the SFD (Size - Frequency Distribution) developed by Marzari et al. (1995), the Global Main Belt SFD, by Durda and Dermott (1997), the family spreading via thermal forces and numerous tiny resonances, the collisional ejection of fragments corrected by Dell’Oro et al. (2004) and the backward numerical integration of orbits (for young families like Karin) used by Nesvorný et al. (2002b).

Many of the catastrophic events at the origin of the breakup of a family can now be dated and, surprisingly, some of them are very young (less than 1 Myr). These families should be used as a fundamental constraint in order to derive the original ejection velocities, since in one Myr, Yarkovsky has not had the time to modify, in a relevant way, the family structure.

The age of the families has a very important consequence: it allows to measure the traces of space weathering, by comparative study of the surfaces of members of old
and young families. By space weathering we refer to all the processes that alter optical properties of surfaces of airless bodies.

The first studies have shown (Chapman 2004) evidence of space weathering, with first estimations of the timescales; it is now recognized that an age dependent component alters asteroid colors.

14. The spin vectors

Let us mention a very interesting result obtained by the observers and experts in photometry.

In the Koronis family, systematic higher lightcurves amplitudes were observed, in comparison with the other regions. Mechanisms were proposed to explain this situation: either it was a mechanism linked to the process of formation of the family, or it was due to the alignment of the spin vectors of the different bodies. Slivan (2002) and Slivan et al. (2003) showed a bimodality in the spin orientation of Koronis family members, and proved that at least 10 members of Koronis had aligned spin vectors (using only the photometric data of the family). Morbidelli and Vokrouhlický (2003) presented a first tentative explanation of this peculiar situation, based on Yarkovsky effect and the presence of a resonance with Saturn.

15. Conclusions

The important point to mention is that dynamical and observational features can now be treated together and not separately; the new clusters or the new members of the known families will be probably discovered with those combined approaches.

Acknowledgements

I would like to acknowledge Alessandro Morbidelli and David Nesvorný for useful discussions, precious documents and preprints, and very kind help. I also thank A. Cellino for his comments and complementary information.

References

Bus, S.J. and Binzel, R.P. 2002b, *Icarus* 158, 146
Durda, D.D. and Dermott, S.F. 1997 *Icarus* 130, 15
Hirayama, K. 1918 *Astron. J.* 31, 185