Automated Classification of Asteroids into Families at Work

Zoran Knežević[1], Andrea Milani[2], Alberto Cellino[3], Bojan Novaković[4], Federica Spoto[2], and Paolo Paolicchi[5]

[1] Astronomical Observatory, Volgina 7, 11060 Belgrade, Serbia
email: `zoran@aob.rs`

[2] Dept. of Mathematics, University of Pisa, Largo Pontecorvo 5, 56127 Pisa, Italy
email: `milani@dm.unipi.it`

[3] INAF-Osservatorio astrofisico di Torino, 10025 Pino Torinese, Italy
email: `cellino@oato.inaf.it`

[4] Dept. of Astronomy, University of Belgrade, Studentski trg 16, 11000 Belgrade, Serbia
email: `bojan@matf.bg.ac.rs`

[5] Dept. of Physics, University of Pisa, Largo Pontecorvo 3, 56127 Pisa, Italy
email: `paolicchi@df.unipi.it`

Abstract. We have recently proposed a new approach to the asteroid family classification by combining the classical HCM method with an automated procedure to add newly discovered members to existing families. This approach is specifically intended to cope with ever increasing asteroid data sets, and consists of several steps to segment the problem and handle the very large amount of data in an efficient and accurate manner. We briefly present all these steps and show the results from three subsequent updates making use of only the automated step of attributing the newly numbered asteroids to the known families. We describe the changes of the individual families membership, as well as the evolution of the classification due to the newly added intersections between the families, resolved candidate family mergers, and emergence of the new candidates for the mergers. We thus demonstrate how by the new approach the asteroid family classification becomes stable in general terms (converging towards a permanent list of confirmed families), and in the same time evolving in details (to account for the newly discovered asteroids) at each update.

Keywords. asteroids, asteroid families, automated classification

1. Introduction

The classification of asteroids into families becomes increasingly challenging as the asteroid datasets continue to grow at an unprecedented rate. The problem is not only with the need to handle huge amounts of data, but also with different and rapidly changing outcomes of classification attempts which tend to create confusion with the users.

In order to cope with these problems we have recently proposed a new approach to the asteroid family classification by combining the classical HCM method with an automated procedure to add new members to existing families (Milani *et al.* 2014). The basic idea is to setup a classification which can be revised only once in a while, but being automatically updated every time the dataset is significantly increased. In practice, our approach consists of several steps to segment the problem and handle the very large amount of data in the most efficient and accurate manner. We use the proper elements first, thus defining dynamical families, then use information from absolute magnitudes, albedos and colors as either confirmation or rejection.

In the following we briefly present all these steps. We then show some results of the application of our procedure in three subsequent updates making use of only the automated step of attributing the newly numbered asteroids to the known families.

2. Method

Our procedure for family identification is based on the classical Hierarchical Clustering Method (HCM), used in most previous families searches since the pioneering work by Zappalà *et al.* (1990) and later improved in a number of papers (Zappalà *et al.* 1994, Zappalà *et al.* 1995, Milani *et al.* 2010, Novaković *et al.* 2011, Masiero *et al.* (2011), Carruba *et al.* (2013)).

We begin our classification with segmenting the problem, that is by dividing the catalog of asteroid osculating elements into parts that can be considered separately. We first divide the asteroid belt into zones corresponding to different intervals of heliocentric distance, delimited by the Kirkwood gaps wide enough to exclude family classification across the boundaries. Next, we split the most populous zones in the central part of the asteroid belt by the value of proper sin I, between a moderate inclination region $\sin I < 0.3$ and a high inclination region $\sin I > 0.3$. Finally, we split the sample in the low-inclination regions of the same central zones by the absolute magnitude, so that in the first step of our procedure we consider only objects having absolute magnitudes H brighter than H_{comp}, with H_{comp} roughly corresponding to the local completeness limit (see Table 2 in Milani *et al.* 2014). Thus, after this first step, in the central zones of the belt we get the "core" families that consist of only the brightest/largest members (red points in Fig. 1 of Milani *et al.* 2014). They represent the inner skeletons of larger families, whose other members are to be identified in the following steps of the procedure. Note, that in the zones with less objects (Hungarias, zones beyond 2:1 mean motion resonance with Jupiter, and high-inclination zones) we identified families by the direct application of the HCM procedure, without the multistep approach.

The second step of the procedure in the low-I portions of the populous central zones is the classification of faint asteroids not used in the first step, that is attaching them to the previously established family cores. We allowed only single links for this attachment, to avoid chaining which would result in merging most families together. Consequently, in step 2 we attribute to the core families the asteroids having a distance from at least one member of the same core family not larger than the critical (threshold) distance. The result is that the families are extended in the absolute magnitude/size dimension, but not much in proper elements space, especially not in proper semimajor axis (green points in Fig. 1 of Milani *et al.* 2014).

As an input to the third step we use the intermediate background asteroids, defined as the set of all the objects not attributed to any family in steps 1 and 2. Families identified at this step are formed by the population of asteroids left after removing from the proper elements data set the already identified family members. We can distinguish two possible outcomes of this step: families can either be fully independent new families having no relation with the families identified previously, or they may be found to overlap step 1+2 families and form satellite families of smaller objects surrounding family cores (yellow points in Fig. 1 of Milani *et al.* 2014).

With the same algorithm of step 2, in step 4 we repeat a single-link attribution to all the families in the extended list of families formed by adding the step 3 families to the list of core families of step 1. Note that with this procedure a small number of asteroids with double classification is unavoidable; if an asteroid is found to be attributed to more than one family, it belongs to an intersection. The multiple intersections between particular

families could be due to the occurence of families at the boundaries between high and low inclination regions in central zones where there is no natural gap (due e.g. to a secular resonance) between these regions, like in the case of family of (729) Watsonia. This is an artifact of our decomposition in zones and needs to be corrected by merging the intersecting families.

More importantly, family intersections occur also when a new family appears as an extension of a family already identified at steps 1 and 2, with intersections near the mutual boundary. Again, a remedy for such a situation is merging of "satellite" families, where, in general, for the merging of two families we require multiple intersections. Visual inspection of the three planar projections of the intersecting families in terms of the proper elements is used to assess the ambiguous cases.

The merging of families constitutes step 5 of our procedure. As an example, let us quote results of the first attempt to family classification by means of our procedure, as given in Milani *et al.* (2014): out of 77 families generated in step 3, 34 have been considered to be satellite families (even 2 core families of step 1 have been found to be satellite of other core families and thus merged); the other 43 families have been left as independent families, consisting mostly of smaller asteroids. There were of course dubious cases, with too few intersections to perform merger. In principle, as the list of asteroids attached to established families grows, the intersection can increase. In some cases the new intersections will support merge previously not implemented, some will certainly open new problems. In any case to add a new merger is a delicate decision which at the moment remains the only step of the procedure we are unable to automatize.

The final step (step 6) of our procedure of asteroid family classification is motivated by the rapid growth of the proper elements database, which results in any family classification becoming quickly outdated. Thus we devised an automatic update of the current family classification, which consists in repeating the attribution of asteroids to the existing families every time the catalog of synthetic proper elements is updated. What we repeat is actually step 4, thus the lists of core families members (found in step 1), of members of smaller families (from step 3), and also the list of already implemented mergers (from step 5) are kept unchanged.

Let us emphasize here that the purpose of this final step is to maintain the general validity of the classification for many years, without the need for repeating the entire procedure. With time, the new data will require to repeat also step 5, to reconsider the list of small families, confirm some of them as statistically significant, discard others as statistical flukes, to decide on pending mergers if intersections increase enough or otherwise new data give enough reason for such a decision, and so on. In brief, we must monitor as the classification is updated and perform non-automated changes whenever we believe there is enough evidence to justify them.

In the next section we shall show how our classification upgrade works, by reviewing the results of the non-automated step 5 application after 3 iterations of automated step 6, during which the catalog of proper elements increased by 48,117 objects or $\sim 14.3\%$.

3. Classification upgrade

As a result of the automatic attribution step in the previous upgrades, the number of family members increased by 10 355 members to 97 440. In the same time, the number of intersections among the 128 families increased from 29 to 48. Thus we had to analyse these intersections, to decide if in some cases the number of intersections among two families has increased enough to suggest a merge.

The case in which a merge was suggested most convincingly was that of the families of (1040) Klumpkea and of (3667) Anne-Marie. In the initial classification there were

10 intersections, now they have grown to 19, which means that out of the total of 30 members of 3667, the majority of members of the smaller family has been attached also to the larger one. Since the inspection of projections in the three proper elements planes did not contradict the proposed merger, we merged the two families and removed 3667 from the list. Note that there is also a new intersection between families 1040 and 29185, but this is not enough to perform a merger at this update.

Another case we found significant enough for a merger involves the families of (375) Ursula and (2967) Vladisvyat, now with 3 intersections. In the original classification there were no intersections, but this had already been pointed out as a future satellite family case by the overlapping box method (see Milani *et al.* 2014, [Section 4.3.2]). The third case is that of the small satellite family 6138 with the family of (135) Hertha: the intersections have grown from 2 to 4, which in this case was enough to decide on merger.

As a result of the 3 mergers described above, the number of families has decreased to 125 and the number of intersections to 22. The remaining ones are mostly due to possible satellite families of the large families 15, 221, 135, 10, and 2076.

An interesting case is an intersection of the family of (5) Astraea with the small family 4945, which was already suggested by the overlapping box method. This latter family appears to form, together with some other small families, a structure that extends along the secular resonance $g - 2\,g_6 + g_5$ in which most of the big family is locked, as predicted by Milani & Knežević (1994) [Figure 9].

An increase of the number of members in a proposed small family can be considered as confirmation of its statistical significance. This growth occured for almost all the families in our classification, with only 5 exceptions. The worst case in this sense is family 3460, which remained unchanged at its original 52 members: this family could be a satellite of family of (24) Themis at higher e and strongly affected by the 2/1 mean motion resonance with Jupiter. Four tiny families (with < 30 members) have also failed to grow: 20494, 1101, 6355, 10654; however, they all belong to the high inclination region in which the total number of asteroids is growing slowly, thus the significance of the lack of increase is dubious. Therefore, we are not removing them from the list yet.

In conclusion, we would like to stress that through the above described procedures and example results we demonstrate the automated classification at work and show how, by converging to a permanent list of confirmed families, the asteroid family classification becomes stable in general terms with the new approach, while evolving in the same time in details (to account for the newly discovered asteroids) at each update.

Acknowledgement

References

Carruba, V., Domingos, R. C., Nesvorný, D., Roig, F., Huaman, M. E., & Souami, D. 2013, *MNRAS*, 433, 2075

Masiero, J. R., Mainzer, A. K., Grav, T., *et al.* 2011, *Astrophys. J.* 741, 68

Milani, A. & Knežević 1994, *Icarus*, 107, 219

Milani, A., Knežević, Z., Novaković, B., & Cellino, A. 2010, *Icarus*, 207, 769

Milani, A., Cellino, A., Knežević, Z. *et al.* 2014, *Icarus*, 239, 46

Novaković, B., Cellino, A., & Knežević, Z. 2011, *Icarus*, 216, 69

Zappalà, V., Cellino, A., Farinella, P., & Knežević, Z. 1990, *Astron. J.*, 100, 2030

Zappalà, V., Cellino, A., Farinella, P., & Milani, A. 1994, *Astron. J.*, 107, 772

Zappalà, V., Bendjoya, P. H., Cellino, A., Farinella, P., & Froeschlé, C. 1995, *Icarus*, 116, 291