Resonances and chaos in the asteroid belt

M. Šidlichovský

Astronomical Institute of the Academy of Sciences of the Czech Republic,
Boční II 1401, 141 31 Prague 4, Czech Republic

Abstract. The present paper reviews the evolution of our understanding of the
effect of resonances on the distribution of asteroids in the asteroid belt. The
history of this problem goes back to the Kirkwood's discovery (1867) of the
Kirkwood gaps located at resonances with Jupiter. We started to understand
the mechanism of their origin only in last decades. It seems that only gravita-
tional effects are sufficient for the depletion. It is now clear that the overlap of
secular resonances inside the orbital resonance is the most effective mechanism
leading to large chaos and variation of orbital elements. This results in the final
removal of asteroids from the gaps by collisions with the inner planets. Chaos,
however, does not always mean fast removal of the body. The question of the
so called stable chaos will be discussed together with the offered explanations
(the high order resonances and the so called three-body resonances). Recently
it was shown that chaotic diffusion can play an important role for the 2/1 re-
sonance where the aforementioned explanation for other gaps fails. Basic facts
will be reviewed but we will not go into this problem as the importance of
chaotic diffusion in dynamics of asteroids (and comets) will be the subject of
invited lecture at this conference given by Morbidelli and Nesvorný.

Key words: asteroids – resonances – chaos – celestial mechanics

1. Introduction

The first demonstration of the effect of resonances on the structure of the as-
teroid belt was found by Kirkwood (1867) when he discovered the Kirkwood
gaps. Fig. 1 shows the semimajor axis–eccentricity diagram of 7541 numbered
asteroids. In 1867 only 95 asteroids were known. The position of gaps is clearly
associated to the 3/1, 5/2, 7/3, 2/1 resonances but the mechanism of their ori-
gin was not understood for a long time. The substantial progress was made in
last decades when it became clear that the Kirkwood gaps can be explained
just by gravitational interaction. The first mathematical formulation and im-
portant result were obtained by Poincaré (1902) and Andoyer (1903). Schubart
(1964) generalized the Poincaré approach and, using averaging techniques on
computer, presented nice graphs giving the behaviour of an asteroid in the re-
stricted, resonant, circular, averaged, and planar problem of three bodies (the
Sun–asteroid–Jupiter). The important requirement is that each theory should

Eds.: J. Svoreň, E.M. Pittich, and H. Rickman,
explain not only the origin of the gaps but also the existence of the Hilda group at the 3/2 resonance (Fig. 1). When the simpler models failed, more sophisticated models were constructed. Qualitatively new features of the system appeared with each new model. We will present the arguments where each model failed and the new features that appeared with new models. The next section will introduce the basic formalism and variables in the three-body problem.

Figure 1. The Kirkwood gaps as seen in $a$–$e$ diagram made for 7541 numbered asteroids

2. Basic formalism of the resonant three-body problem

The Hamiltonian of the restricted three-body problem (3D elliptic) is

$$\mathcal{H} = L' - \frac{1 - \mu}{2a} - \mu \left( \frac{1}{|r - r'|} - \frac{rr'}{r'^3} \right),$$

(1)

where $L'$ is conjugate to the mean longitude of Jupiter $\lambda'$, $r$ is the vector from the Sun to asteroid, $r'$ is the vector from the Sun to Jupiter, $a$ is the semimajor axis of asteroid, $\mu$ is the mass of Jupiter (the gravitational constant $G=1$, sum of Jupiter and the Sun masses=1, the semimajor axis of Jupiter orbit $a'=1$). In the non-resonant case there is a very useful set of the variables – the Delaunay canonical variables: $L = \sqrt{(1 - \mu)a}$, $G = L\sqrt{1 - e^2}$, $H = G \cos I$, $l$ is the mean anomaly, $g$ is the argument of pericenter, $h$ is the longitude of the ascending node, $e$ is the eccentricity and $I$ is the inclination of the asteroid. In the resonant case when the critical argument $(p+q)\lambda' - p\lambda$ varies very slowly, we can introduce the following set of canonical variables:

$$\sigma = \frac{p + q}{q} \lambda' - \frac{p}{q} \lambda - \tilde{\omega}, \quad S = L - G,$$
\[
\sigma_z = \frac{p+q}{q} \lambda' - \frac{p}{q} \lambda - \Omega, \quad S_z = G - H,
\]
\[
-\nu = \frac{p+q}{q} \lambda' - \frac{p}{q} \lambda - \dot{\omega}', \quad N = \frac{p+q}{p} L - H,
\]
\[
\lambda', \quad \lambda' = L' + \frac{p+q}{p} L.
\]

The elliptic 3D three-body problem has four degrees of freedom.

The fast variable \( \lambda' \) may be removed by averaging the disturbing function

\[
R = \left( \frac{1}{|r - r'|} - \frac{rr'}{r'^3} \right). \tag{3}
\]

\[
\bar{R}(\sigma, \sigma_z, \nu, S, S_z, N) = \frac{1}{2\pi p} \int_0^{2\pi p} R(\sigma, \sigma_z, \nu, \lambda', S, S_z, N) d\lambda', \tag{4}
\]
keeping all remaining canonical variables constant. After removal of short-period terms we reach at 3 degrees of freedom autonomous Hamiltonian.

In the planar, averaged problem \( \sigma_z \) is not defined but \( \bar{R} \) is independent of \( \sigma_z \). It follows that \( S_z = 0 \) is constant (as \( G = H \)). After averaging we have 2 degrees of freedom

\[
\bar{R}(\sigma, \nu, S, N) = \frac{1}{2\pi p} \int_0^{2\pi p} R(\sigma, \nu, \lambda', S, N) d\lambda', \tag{5}
\]
and, e.g. the Poincaré surface of section method is available.

In the planar, circular, averaged problem \( \bar{R} \) is independent of \( \nu \). After averaging we have 1 degree of freedom problem

\[
\bar{R}(\sigma, S, N) = \frac{1}{2\pi p} \int_0^{2\pi p} R(\sigma, \nu, \lambda', S, N) d\lambda', \tag{6}
\]
as \( \bar{R} \) is independent of \( \nu \). Hence

\[
N = \frac{p+q}{p} L - G \tag{7}
\]
is an integral of motion yielding the relation between \( a \) and \( e \):

\[
\sqrt{a} \left( \frac{p+q}{p} - \sqrt{1-e^2} \right) = \text{const.} \tag{8}
\]

and the averaged Hamiltonian

\[
\bar{H} = \lambda' - \frac{p+q}{q} (N - S) - \frac{q^2(1-\mu)^2}{2p^2(N-S)^2} - \mu \bar{R}(\sigma, S, N) \tag{9}
\]
is integrable having 1 degree of freedom. For each constant parameter \( N \) the motion of \( \sigma \) and \( S \) takes place along the curves of constant \( \bar{R} \). Such curves were in principle first drawn by Schubart (1964) for various resonances in \( x, y \) plane, where \( x = \sqrt{2S} \cos \sigma, \ y = \sqrt{2S} \sin \sigma \).
3. Schubart method

The Schubart results for restricted, resonant, circular three-body problem were improvement of the Andoyer’s (1902) work without restriction to small eccentricities. Schubart applied his averaging to 17 members of the Hilda group (the 3/2 resonance) and found them all on trajectories well protected from close approaches to Jupiter. The observed small amplitude libration of $\sigma$ about 0 means that conjunction is never far from the asteroid perihelion. This protecting mechanism is very important for the stability of the Hilda group. On the other hand, similar $x-y$ graphs were obtained for the 2/1 resonance indicating that similar protected asteroids should exist there. The circular model failed to explain the 2/1 gap.

Schubart (1968) generalized his approach to the elliptic case. He constructed integrator solving the resulting averaged equations (at least 2 degrees of freedom problem – in the planar case). The averaging made the integration much faster allowing for larger integration step. Giffen (1973) used this integrator and found the first interesting differences between the 2/1 and 3/2 cases, applying the sort of surface of section method. In $e, \sigma$ plane he plotted points for which $\dot{a} = 0$ (and $a$ is at maximum). For initial eccentricity $e_0 \in (0.1 - 0.3)$, $a_0 \in (3.920 - 4.115)$ he found only regular trajectories in the 3/2 resonance. For the 2/1 resonance he found for $e_0 \in (0.1 - 0.34)$ chaotic trajectories. This was also the first demonstration of chaos in the asteroid belt. Froeschlé and Scholl (1976, 1981) confirmed the existence of chaotic trajectories for small eccentricities and showed that the chaotic region is confined to low eccentricities. It means that the elliptic three-body problem cannot explain the 2/1 Kirkwood gap. This conclusion was not changed by using 3D elliptic problem. The origin of Giffen’s chaotic region was explained (Lemaitre and Henrard, 1990) by overlap of the secondary resonances (between $\sigma$ and $\omega$).

On the other hand the 3/1 Kirkwood gap could be explained by elliptic problem, even if still more effective mechanism was found later. The first indication was found by Scholl and Froeschlé (1974) who observed peculiar orbits starting in chaotic region located at small eccentricity for which eccentricity could increase to 0.3. The interval of integration was about 50 000 years so it could not be clear that these eccentricity jumps take place for much wider set of trajectories (sometimes as long as 1Myr of small eccentricity mode).

4. Wisdom mapping

Wisdom (1982) used a mapping, which is actually a first order symplectic integrator. It made possible to integrate the equations for truncated elliptic problem very fast. Wisdom (1987) claims the mapping to be several hundred times faster than the Schubart’s averaging program. Wisdom could calculate the evolution of a hypothetical asteroid located at the 3/1 resonance (for planar-elliptic
problem) for more than 1 Myr very fast and he could prove the existence of a large region where initially small eccentricity goes sooner or later to the Mars crossing values. Wisdom (1985) explained this behaviour by existence of a large chaotic zone extending to large eccentricities. Adiabatic approximation was used to separate $\sigma$ and $\varpi$ oscillations, yielding the so called guiding trajectories. The existence of chaotic region was related to the critical curve where this approximation failed. The Wisdom results gave the first satisfactory explanation of one of the gaps purely on the basis of gravitational forces (no cosmogonical or statistical hypothesis were used). The Wisdom mapping was generalized to the $2/1$ and $3/2$ resonance by Murray (1986), to the $5/2$ by Šídlíchovský and Melendo (1986), and general approach to all resonances was presented by Šídlíchovský (1992). The failure of the truncation of the Hamiltonian at low powers of eccentricities used in formulating the mapping was studied in many papers (e.g. Henrard and Lemaitre, 1987 or Šídlíchovský, 1993). This failure lead Ferraz-Mello and Sato (1989) to very high eccentricity expansion of the disturbing function near resonances. Using this expansion Ferraz-Mello and Klaake (1991) calculated the surfaces of section for the planar, elliptic, three body problem valid to large eccentricities. Hadjidemetriou (1992) introduced a heuristic correction term of the sixth degree in $e$ to the Hamiltonian, to obtain fixed points corresponding to the $3/1$ resonant families of periodic orbits in elliptic problem. Many other interesting papers in the frames of restricted, elliptic, resonant three-body problem were published. However, we will not go into these results because it was finally shown that for the origin of the Kirkwood gaps the secular resonances are important.

5. Secular resonances inside the mean motion resonance

While the elliptic three-body problem seemed to be a promising model for the $3/1$ resonance, it certainly failed to explain the $2/1$ resonance. It was necessary to construct more complicated models. Morbidelli and Moons (1993) took into account the secular variations of Jupiter’s orbit due to perturbations from Saturn. The inclusion of Saturn’s indirect effects introduces three new frequencies: main frequencies of the Jupiter’s and Saturn’s perihelion $g_5$, $g_6$ and main frequency of Saturn’s node $s_8$ into the problem. These three frequencies appear in the decomposition of the Jupiter’s elements. The three corresponding secular resonances are: $\nu_5$ with $\dot{\omega} = \dot{\omega}_J$, $\nu_6$ with $\dot{\omega} = \dot{\omega}_S$, $\nu_{16}$ with $\dot{\Omega} = \dot{\Omega}_S$. The subscripts $S$ and $J$ stand for Saturn and Jupiter, respectively. The position of these resonances inside the $2/1$ and $3/2$ resonances was studied by Morbidelli and Moons (1993). They found large overlap (Fig. 2) for the planar case of the $\nu_5$ and $\nu_6$ resonances. The chaotic zone for $e < 0.15$ corresponds to the overlap of secular resonances. These two chaotic zones are separated by “white zone” which is for the $2/1$ practically depleted while for the $3/2$ this is the region where the Hilda asteroids are concentrated (points in Fig. 2). The corresponding chaotic zone
Figure 2. The overlap of the secular resonances $\nu_5$ and $\nu_6$ in the planar 2/1 resonance (left) and 3/2 resonance (right) Moons (1996). The chaotic zone resulting from this overlap extends to very high eccentricities and is separated from the chaotic zone of overlapping secondary resonances with “white zone” where Hildas can stay in the 3/2 resonance (points).

extends to very large eccentricities. For the 4/1, 3/1, 5/2 and 7/3 resonances (Moons and Morbidelli, 1995) the chaotic zone extends over nearly all the resonant region. The eccentricity of asteroids in this region increases very rapidly to values 0.8 and in time period 1 Myr they become not only Mars crossers but even Earth crossers. Thus the mechanism for fast depletion of these gaps is provided. It is much faster than for models without secular resonances. There is still one problem with this mechanism. The central region of the “white zone” for the 2/1 resonance does not contain asteroids at variance with similar zone in the 3/2 resonance. The study of changes of the main frequency in $\omega$ variations with time (which quantify the chaotic diffusion as in regular cases the frequency would be constant) was performed by Ferraz-Mello et al. (1999) and showed that the diffusion in the “white zone” of the 2/1 resonance is ten times faster than in the 3/2 resonance. Moreover, the diffusion in the 2/1 resonance is sensitive to the value of the Jupiter’s Great Inequality (GI). Small change in initial conditions of Jupiter can lead to reduction of GI-period from 880 days to about 450 yrs. Ferraz-Mello et al. (1999) showed that the diffusion is then much faster. Similarly, the numerical calculations of maximum Lyapunov exponent (MLCE) showed that they are larger in the 2/1 “white zone”. These results prove that the solution of the remaining problem, namely depletion of the “white zone” in the 2/1 resonance, is the faster diffusion than in the corresponding zone in the 3/2 resonance.
6. The higher order resonances in the outer asteroid belt

The importance of the higher order resonances was first indicated by Milani and Nobili (1992) in their study of asteroid 522 Helga. This asteroid has a very short Lyapunov time $T_L$ therefore, its trajectory is quite chaotic. On the other hand it shows no significant long-term evolution and no macroscopic instability over a time period longer than 1000 $T_L$. For such behavior Milani and Nobili (1992) introduced label stable chaos. They showed that the strongest chaotic effects

![Graphs showing Lyapunov time and removal time as functions of semimajor axis](image)

**Figure 3.** Left: Lyapunov time as a function of the semimajor axis in the elliptic restricted three-body problem. Right: Removal time as a function of semimajor axis. Solid points are numerical measurements, open points analytical estimates and the solid triangles in the 12/7 resonance were obtained by integration including the effects of four giant planets. (From Murray and Holman (1997).)

on Helga are caused by the interaction of the intermediate order resonance 12/7 with Jupiter. Morbidelli and Froeschlé (1996) investigated the relation between the Lyapunov time and macroscopic instability time, as the observed stable chaos seemed to contradict the power low suggested by some results (obtained for much enhanced Jupiter’s mass). They explained this contradiction by existence of two possible regimes. In the so called Nekhoroshev regime the macroscopic diffusion time is exponentially long with respect to Lyapunov times. There is an interesting question: Is there an analogy to the Kirkwood gaps at the higher order resonances? The Helga example suggests that it is not the case, at least not for the 12/7 resonance. For the outer belt outside 3.4 AU this problem was studied by Holman and Murray (1996). They studied the variation of Lyapunov time as a function of the initial semimajor axis throughout the outer belt taking into account the giant planets perturbers. They observed dips in the $a - T_L$ diagrams corresponding to resonances and verified that the resonant angle makes transitions between circulation in one sense to circulation in other sense,
interspaced with brief periods of libration. The width of these dips increased with the initial eccentricity. The relation of the higher degree of chaos to resonances was clearly established. Holman and Murray (1996) concluded that although most of the outer belt asteroids have chaotic trajectories, they lie outside of the chaotic zones with the shortest Lyapunov time (i.e. outside resonances). The observed gaps were at the 11/6, 9/5, 7/4, 5/3 and 8/5 Jupiter mean motion resonances.

Murray and Holman (1997) succeeded in making analytical estimates of the Lyapunov time for various outer resonances (Fig. 3a) and even in the estimates of removal time $T_R$ necessary for removal of the asteroid from resonance. Though these estimates were made in the frames of three-body problem they confirm the theoretical expectation of the observed gaps. The bodies in the 12/7, 13/8, and possibly 11/7 resonances, respectively, appear to have removal times greater or comparable to the lifetime of the solar system.

Figure 4. Log of the MLCE as a function of semimajor axis (MLCE is in yr$^{-1}$). The effect of the outer planets is included. The position of the Jupiter mean motion resonances and the three-body resonances Jupiter–Saturn–asteroid are indicated. The three-body resonance is characterized by $m_J+m_S+m$. (From Nesvorný and Morbidelli, 1999a.)

7. Resonances in the inner asteroid belt, three-body resonances

Morbidelli and Nesvorný (1998a) calculated MLCE for 5700 test particles set on a regular grid in semimajor axis that ranges from 2.1 to 3.24 AU. The initial eccentricity was 0.5 and outer planets were taken into account. Fig. 4 shows a part of their $a-\log MLCE$ diagram. In the outer belt the chaos is correlated with the position of the Jupiter mean motion resonances, but there are many
Table 1. The twenty most chaotic asteroids ($T_L > 32000$ yr) of the first one hundred numbered asteroids. Integrations are performed with 7 (column 3) and 4 planets (column 4). Column 5 shows the resonance with the strongest chaotic effect. The three body resonances are schematically written as $m_J + m + m_S$.

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>$L_T[10^3$ yr]</th>
<th>$L_T[10^3$ yr]</th>
<th>Resonance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>7 planets</td>
<td>4 planets</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Pallas</td>
<td>10</td>
<td>23</td>
<td>18:7 with Jupiter</td>
</tr>
<tr>
<td>7</td>
<td>Iris</td>
<td>17</td>
<td>$&gt;1000$</td>
<td>25:49 with Mars</td>
</tr>
<tr>
<td>8</td>
<td>Flora</td>
<td>30</td>
<td>$&gt;1038$</td>
<td>19:33 with Mars</td>
</tr>
<tr>
<td>10</td>
<td>Hygiea</td>
<td>16</td>
<td>15</td>
<td>8-3-4 Jupiter–asteroid–Saturn</td>
</tr>
<tr>
<td>12</td>
<td>Victoria</td>
<td>33</td>
<td>$&gt;994$</td>
<td>29:55 with Mars</td>
</tr>
<tr>
<td>15</td>
<td>Eunomia</td>
<td>25</td>
<td>$&gt;980$</td>
<td>7:16 with Mars</td>
</tr>
<tr>
<td>23</td>
<td>Thalia</td>
<td>25</td>
<td>$&gt;1233$</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Polyhymnia</td>
<td>10</td>
<td>14</td>
<td>22:9 with Jupiter</td>
</tr>
<tr>
<td>35</td>
<td>Leukothea</td>
<td>20</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Atlante</td>
<td>4</td>
<td>5</td>
<td>4-2-3 Jupiter–asteroid–Saturn</td>
</tr>
<tr>
<td>41</td>
<td>Daphne</td>
<td>14</td>
<td>$&gt;360$</td>
<td>9:22 with Mars</td>
</tr>
<tr>
<td>46</td>
<td>Hestia</td>
<td>30</td>
<td>$&gt;1039$</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Virginia</td>
<td>10</td>
<td>12</td>
<td>11:4 with Jupiter</td>
</tr>
<tr>
<td>53</td>
<td>Kalypso</td>
<td>19</td>
<td>14</td>
<td>6-2-1 Jupiter–asteroid–Saturn</td>
</tr>
<tr>
<td>60</td>
<td>Echo</td>
<td>27</td>
<td>$&gt;1159$</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Panopea</td>
<td>24</td>
<td>36</td>
<td>2-1-2 Jupiter–asteroid–Saturn</td>
</tr>
<tr>
<td>75</td>
<td>Eurydike</td>
<td>16</td>
<td>$&gt;932$</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>Diana</td>
<td>13</td>
<td>149</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>Eurynome</td>
<td>32</td>
<td>497</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>Semele</td>
<td>6</td>
<td>6</td>
<td>13:6 with Jupiter</td>
</tr>
</tbody>
</table>

Other peaks which were identified with three-body mean motion resonances. These resonances correspond to the relation between mean motions of Jupiter, asteroid and Saturn:

$$m_J \lambda_J + m \lambda + m_S \lambda_S \sim 0,$$

where $m$, $m_J$ and $m_S$ are integers and subscripts $J$ and $S$ denote Jupiter and Saturn.

The analytical model of the three-body resonances was presented by Nesvorný and Morbidelli (1999b). The effect of Mars was studied by Nesvorný and Morbidelli (1999a). They observed many new peaks in $a - \log MLCE$ diagram corresponding to resonances with Mars and even to the three-body resonances Mars–Jupiter–asteroid. All observed peaks become wider with the increasing initial eccentricity.
8. The most chaotic asteroids

Froeschlé et al. (1999) calculated the so-called Fast Lyapunov Indicators for 5400 asteroids without the inner planets. Šidlichovský and Nesvorný (1999) calculated MLCE for the first one hundred of the numbered asteroid with the effect of inner planets. By comparing our results with the list of 848 most chaotic asteroids (Froeschlé et al., 1999) we found (Šidlichovský, 1999) that we have more chaotic asteroids with the inner planets included. Tab. 1 shows the twenty most chaotic asteroids of the studied sample. Column 3 shows the Lyapunov time with the inner planets, column 4 without them. Fig. 5a shows the filtered

![Graph](image_url)

**Figure 5.** a) The filtered semimajor axis of Iris  b) The Iris and Mars distance

(low-pass digital filter is used for removing the periods shorter than 82 years) semimajor axis $a$ of asteroid 7 Iris. The abrupt variation of $a$ takes place exactly when close approach to Mars is possible. This is a typical behaviour for asteroids where chaos is caused by inner planets (in our sample it is always Mars). From the value of $a$ the responsible resonance may be found. In the case of Iris we found that it is the Mars mean motion 25/49. Fig. 6 shows the behaviour of filtered $a$ and the corresponding critical argument $\sigma$. Such behaviour is the evidence that this very high order resonance is really responsible for abrupt
variations of filtered $a$. Usually more subresonances are librating which indicates their overlap. Šidlichovský (1999) could determine by this method many resonances responsible for the chaos of the studied set of asteroids. These resonances are shown in column 5 of Tab. 1.

![Graph](image)

**Figure 6.** a) The filtered semimajor axis of Iris b) The critical argument of Iris: $\sigma = 49\lambda - 25\lambda_M - 24\varpi$

9. Diffusion

Morbidelli and Nesvorný in their lecture at this conference demonstrated in more detail and also on videos how the eccentricity of asteroids in certain diffusion tracks can slowly diffuse to Mars crossing values on the time scales of 100 Myr (see also Migliorini et al., 1998). As the median dynamical lifetime of Mars crossers is about 25 Myr, this is also a way from the main belt to the NEA’s. The main diffusion traces are again related to resonances such as the 7/2 with Jupiter, 5/9 with Mars etc (Nesvorný and Morbidelli, 1999a). Thus the resonances are the very important force leading to the evolution of the structure of the asteroid belt.
Acknowledgements. This work has been supported by a Grant 205/98/1102 of the Grant Agency of the Czech Republic.

References

Schubart, J.: 1964, *SAO rept* 149,