Pramāna, Vol. 8, No. 5, 1977, pp. 438–446. © Printed in India.

The asteroidal belt and Kirkwood gaps—I. A statistical study

R PRATAP
Physical Research Laboratory, Ahmedabad 380009

MS received 2 February 1977

Abstract. In this paper we have made a spectral analysis study of matter distribution in the asteroidal belt. We have Fourier analysed this distribution and obtained the autocorrelation and power spectrum, and have identified the ratios from the resonance theory. We have shown that the Kirkwood gaps cannot be satisfactorily interpreted as due to mere resonance between the asteroid and Jupiter orbital motions. We propose that they may be regarded as a consequence of density waves generated in the gas disc in the ecliptic plane in the neighbourhood of the Sun. We have also shown that the process is non-Marcovian and hence cannot be subjected to a hydrodynamical analysis.

Keywords. Density waves; Non-Marcovian; Kirkwood gaps.

1. Introduction

It was observed by Daniel Kirkwood (1867) that the asteroidal belt had a ring structure, and since then, resonance theory was invoked to explain the rings of minimal matter-gaps. The total mass distribution in these rings has been estimated to be about $0{\cdot}1\ M_\oplus$. Recently Ovenden (1972) tried to estimate as to what would have been the mass in this region at the origin of the solar system. He invoked a "principle of least interaction action" and using this, he calculated the semi-major axes of the various planets, trying to examine the physical content of Titius-Bode's law. He has shown that a best fit is obtained if he assumes a planet of $90 \pm 5\ M_\oplus$ in the asteroidal belt in the beginning, and then got dissipated at $1{\cdot}6 \pm 0{\cdot}4 \times 10^7$ years ago. This would imply that, if we assume a continuous process, then an average mass loss of 3×10^9 kg/year has taken place during $1{\cdot}6 \times 10^7$ years to get to the present value. If the Belt is considered as being confined between 2 and 4 AU then this implies an average mass loss of 35 kg/sq.km/year. It might not have been a uniform process as we have assumed here, but it could have been a more rapid dissipation in the beginning and from the outer rings, and then decreased in time. It can be seen from the next paper that the outer rings are comparatively less stable than the inner ones. Nevertheless there would have been more frequent collisions, and hence the system should be considered as a highly nonequilibrium one.

In this paper we propose to undertake a spectral analysis study of the existing matter distribution (radial—$\tilde{\omega}$) from the known data $\rho\ (\tilde{\omega})$ in the asteroidal belt. We do not make any assumption as regards the dynamical process that brought the system to the present state. We shall do a kinematical study in the next paper. In this, we shall try to find out the various bidden symmetries and frequencies. This would probably give us a better insight into the dynamical processes

that have given the present state (as it is known today). We shall on a later occasion try to develop an evolutionary theory with the present as the final state.

In the next section, we have made a Fourier analysis of the known data. We have shown that the given distribution is indeed due to the summation of two *distinct series*. In the third section, we have discussed the autocorrelation function and have made a study of power spectrum (PS). We made a PS analysis, by two methods: (1) by the maximum entropy method (MEM), and (2) by the Blackman and Tuckey (B and T) method. We see that the dominant frequencies from these two methods agree fairly well. We have made a comparison of the PS obtained from the two methods, to show the relative merits. We have also shown that the PS analysis which would give the real frequency distribution do not betray the frequencies that the earlier authors have discussed based on the resonance theory. The final section gives a discussion and the way of looking at the problem from the theory of density waves as propounded by Jeans (1961), Lindblad (1962), Lin and Shu (1964) and others.

2. Fourier analysis

The plot of the number density against distance is given in figure 1. We now write a Fourier expansion for the above as

$$\rho(\tilde{\omega}) = a_0 + \frac{2}{N}\sum_{n=-N/2}^{N/2}\left[a_n \cos\left(\frac{2\pi n\tilde{\omega}}{L}\right) + b_n \sin\left(\frac{2\pi n\tilde{\omega}}{L}\right)\right]$$

$$= a_0 + \frac{2}{N}\sum_{n=-N/2}^{N/2} R_n \cos\left(\frac{2\pi n\tilde{\omega}}{L} + \phi_n\right) \qquad (1)$$

where a_0 is a constant (sometimes called the DC component) and N (= 48) denotes the number of divisions made of the interval 2 AU $\leqslant \tilde{\omega} \leqslant$ 4 AU. a_n and b_n are Fourier coefficients defined in the usual way. ϕ_n is the phase factor in a polar representation, such that

$$a_n = R_n \cos\phi_n$$

$$b_n = -R_n \sin\phi_n. \qquad (2)$$

Figure 2 gives a polar plot of the vector R_n defined by (2), and the numbers written against each vector is the index n. The following features are very remarkable.

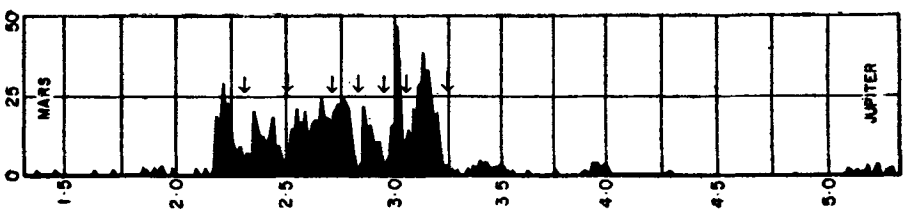

Figure 1. A plot of number density against the radial distance from the Sun. The positions of Mars and Jupiter are marked by the vertical lines. The abscissa is in AU and ordinates denote the number of asteroids > 1·6 km in diameter.

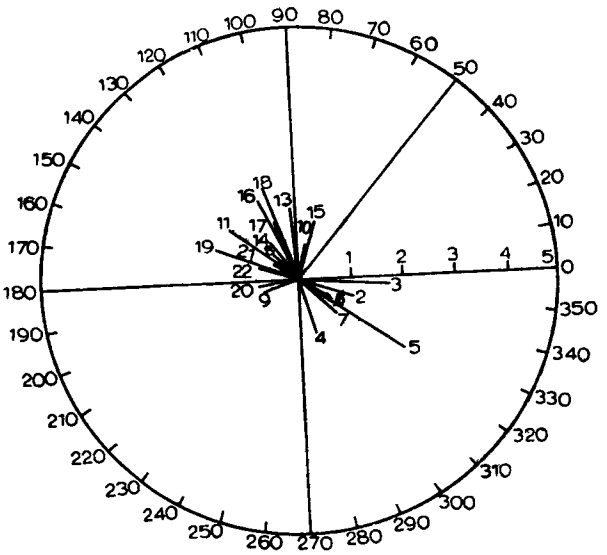

Figure 2. Fourier space representation of the distribution in figure 1. The vectors represent amplitude and the corresponding phase of the Fourier components. The number against each vector represents the number n of the Fourier component.

2.1 *Amplitude* $|R_n|$

The absolute values of the vector R_n vary in length. The vector $n = 1$ is the most dominant one with $\phi_1 = 50°$ and length 4·96 units. In the interval $1 \leqslant R_n \leqslant 2$, we have vectors in decreasing sequence in length at $n = 18$, 16, 3, 19, 11, 13, 17, 15. The remaining vectors are all of magnitude < 1. One can see that the convergence of the series in the beginning is slow, but beyond 18, the values consistently decrease and hence the convergence is good. It is obvious from the diagram. But it is also obvious that the major part of the series is indeed the first harmonic.

2.2. *Phase* ϕ_n

The distribution of these vectors in the Fourier space is very striking. It could be seen that while the first harmonic is in the first quadrant, the harmonics 2 to 7 are all in the fourth quadrant. Most of the other harmonics are in the second quadrant with a very few in the first and the third ones. There are *no* vectors in the intervals $355° < \phi_n < 50°$ and $200° < \phi_n < 286°$. Thus the Fourier space is clearly divided into two distinct domains. In taking the inverse Fourier transform, therefore, the resultant series would be a sum of *two* distinct series developed in these two different domains. This would therefore imply that the original distribution of matter as given in figure 1 is really a resultant of sum of two harmonic series. We can therefore consider the Kirkwood gaps as the minima produced by the mutual interference of two distinctly separate density waves. The two wave trains are built out of various frequencies and hence are not purely sinusoidal. This manifests itself in the fact that maxima and minima are not equally spaced, and they are not of the same intensity. This is a clear indication of density waves in astrophysical context, which was invoked by Jeans (1961) to

define a characteristic length, and which was again worked out by Lindblad (1962) and Lin and Shu (1964) and in a series of papers by Lin to explain the formation of spiral arms in galaxies.

3. Autocorrelation and power spectrum

In figure 3 we have given the autocorrelation function evaluated for the density distribution, plotted against the lag ($\triangle\tilde{\omega}$) at equal intervals of 1/24. This has some very special features. We get all the positive correlations between the lag 0 to 11, and from 11 to 35 all correlations are negative. Beyond 35, there is a change of sign, but are all small and negligible. This implies that in the first segment, while the particles are highly correlated, amongst themselves, the correlations die down as we approach the lag 11·5, and there is another segment in which they are highly correlated. But the particles in the first and second segments are not mutually correlated. If we consider this autocorrelation function as a series, then it has a dominant first harmonic, which is evident from the diagram for PS in figure 4. It would be interesting to compare the frequencies here with those obtained from the resonance theory. The major resonances which are the ratios of the orbital periods of Jupiter and the asteroidal particles are the commensurability ratios 2:1, 7:3, 5:2 and 3:1. We can calculate the corresponding lag points on the diagram as follows: using Kepler's law $a^3/T^2 = \text{constant}$, wherein a is the semi-major axis of the orbit, and T the orbital period, we can calculate the semi-major axis of the asteroidal particle. These are given in table 1. From the table the

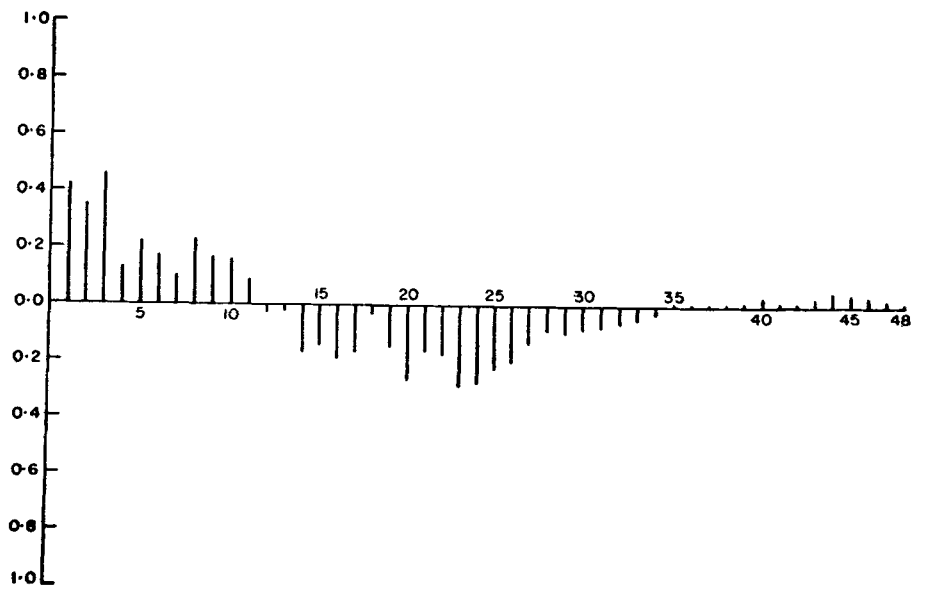

Figure 3. This figure represents the autocorrelation function at equal lag points (= 48) on the x-axis, and the length of the correlations on the y-axis. The maximum, *viz.*, 1, is at the origin by definition. The correlation is consistently positive till 11·5, and then negative till 34·5. The correlation then changes sign and then dies down. This shows that the particles could be divided into three groups mutually correlated amongst themselves, but uncorrelated groupwise.

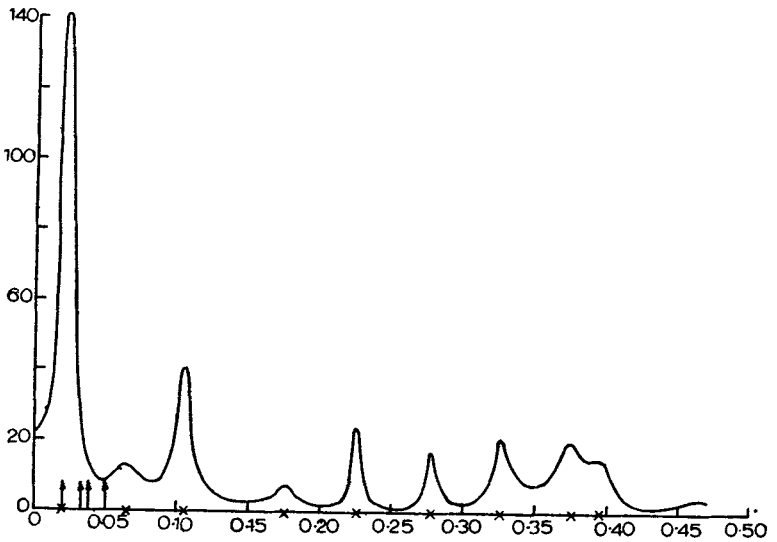

Figure 4. This gives the power spectrum based on MEM. *X*-axis gives the frequency and *y*-axis the power. The crosses on the abscissa depict the frequency at which maximum power resides. The vertical arrows are the frequencies given by the resonance theory. The frequency 0·02 corresponds to the ratio of 3:1 as given by resonance theory. The other ratios are not natural frequencies of the problem.

Table 1. The frequencies given by the theory of resonance and its corresponding points in figure 4. The ratio of the orbital periods of Jupiter and the asteroidal particle is converted to the ratio of the semi-major axes using Kepler's law.

Sl. No.	Commensurability ratios	Semi-major axis of the gaps a	$a-2$	lag no.	Frequencies on the PS
1	2 : 1	3·27 AU	1·27 AU	30·5	0·05
2	7 : 3	2·95 AU	0·95 AU	22·8	0·038
3	5 : 2	2·82 AU	0·82 AU	19·7	0·033
4	3 : 1	2·50 AU	0·50 AU	12·0	0·020

ratio 3:1 gives the lag point 12, which corresponds to the point where the positive correlation changes to negative correlation. This therefore is the most significant ratio and this can be seen even from the PS. The maximum power is in the frequency 0·02 by the MEM and 0·021 by the B and T method. This is given in table 2. There are no other inversions of significance at any other lag points. Further the other ratios are such that they all fall in the negatively correlated domain and hence they are really insignificant. This again is apparent from the PS. Hence we like to assert that the other ratios are not real in the problem. This result casts a serious doubt on the theory based on the resonance phenomena. This clearly shows that while Jupiter is a factor in the determination of density waves, the source is much more complex. It could be that Mars also may have a significant role in the formation of Kirkwood gaps, or it could also be that a collective effect is responsible for the formation of these gaps.

We now come to the analysis of PS. The autocorrelation analysis is really a preliminary effort to study the PS. Figure 4 gives the plot for MEM. The results

Table 2. A comparative study of the MEM based on the Burg method and B and T. Power at maximum frequency is given with power at the two adjacent frequencies on either side. This indicates the width of the peak. MEM width is narrow and the power falls to about 1% on either side of the maxima showing the superiority of the MEM over the B and T.

	MEM						B and T					
Sl. No.	Maxima	k	AU	Power at the peak	Power at the neighbourhood L	Power at the neighbourhood R	Maxima	k	Power at the peak	AU	Power at the neighbourhood L	Power at the neighbourhood R
1	0·0200	0·50	2	3030	28·82	30·09	0·021	0·525	1·91	13·09	13·05	13·06
2	0·1040	2·6	0·385	199·1	38·37	20·38	0·104	2·6	0·385	3·874	3·87	3·86
3	0·2210	5·53	0·181	177·4	6·91	9·39	0·380	9·5	0·105	2·86	2·86	2·85
4	0·3810	9·53	0·105	111·0	34·85	11·06	0·329	8·23	0·122	2·298	2·296	2·291
5	0·0600	1·5	0·667	97·55	13·80	25·52	0·224	5·6	0·179	1·550	1·543	1·548
6	0·2740	6·85	0·146	81·94	10·49	8·62	0·278	6·95	0·144	1·456	1·449	1·454
7	0·4610	11·53	0·087	39·16	3·47	1·87	0·065	1·63	0·614	1·845	1·843	1·840
8	0·4040	10·1	0·099	34·11	5·55	2·04	0·176	4·4	0·227	0·913	0·910	0·912
9	0·3180	7·95	0·126	27·35	19·30	16·01	0·461	11·5	0·087	0·527	0·525	0·527
10	0·1780	4·45	0·225	21·04	2·47	7·45						
11	0·1340	3·35	0·299	20·96	13·20	5·30						
12	0·2330	5·83	0·172	18·15	8·28	4·93						
13	0·3540	8·85	0·113	15·51	13·31	11·97						

are given in table 2. The superiority of MEM in locating the significant frequencies correctly was recently discussed at length by Radofko *et al* (1975). The most dominant frequencies are 0·02, 0·105, 0·225, etc., which correspond to the vectors 1, 5, 11, etc. (these can be obtained by multiplying these numbers by the number of intervals, *viz.*, 48). We have not listed all the frequencies here, but we see that these agree well with those inferred from the Fourier analysis. Small discrepancies are bound to appear, but MEM is the most reliable method of getting the dominant frequencies. We have marked the frequencies predicted by the resonance theory (and given in table 1) by arrows in figure 4. It is evident that the ratio 3:1 is the most dominant frequency. The other ratios given by the resonance theory do not have a corresponding one on the PS and hence these ratios are spurious.

In figure 5, the PS due to MEM and B and T are plotted side by side. We took a closer interval of 0·001 so as to pinpoint the maxima. The MEM estimate plotted here is after the method developed by Burg (see Radofko *et al* 1975), and this narrows down the width of the curve and thereby obtains exact frequencies. One observes the following features: (1) The MEM reveals a large variety of structures which are completely obliterated by the B and T method. Hence the superiority of this method over the B and T. (2) Two or three maxima appearing in the MEM spectra merge together in the B and T method and hence show an apparent shift in the maxima. (3) The power is largely concentrated in the lower modes, thereby exhibiting the nonstationary character of the distribution. This may be interpreted as the process being non-Marcovian. This is a very significant observation as regards the dynamics of the system. If the process is non-Marcovian in space, one can extend the same to time evolution as well, and then this would imply that the present state is due to long chain of events with cumulative effects. Hence the dynamical equations characterising the evolution should also be non-Marcovian. (4) The non-stationary character of the curve reveals the fact that if we consider the process of formation of these gaps as an evolutionary one, then the process can be inferred as a quasilinear or nonlinear one. We shall discuss this fact in more detail in the next paper.

In table 2 we have given the peaks according to descending power in each mode. We have also given the maximum power in each mode along with the power at

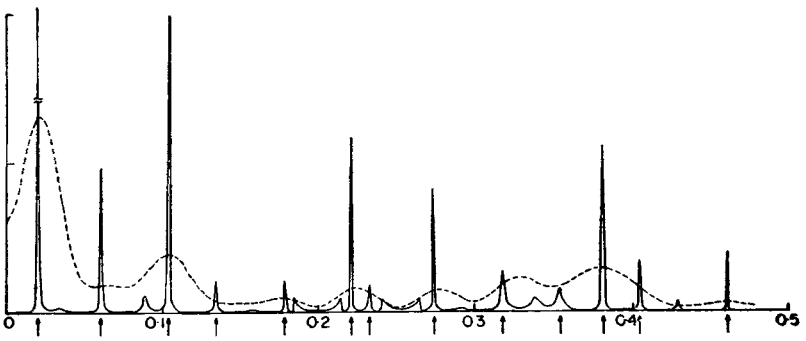

Figure 5. This gives a comparative study of PS by MEM (solid line) based on the Burg method, and the B and T (dashed line). The arrows denote the frequencies at which power peaks by the MEM. B and T method gives a much flatter and smoother curve, and do not give the frequencies accurately. The various peaks in the MEM spectra depict the non-Marcovian nature of the process.

adjacent points differing by 0·001 unit. This gives the estimates of the half width of the peak. One can see that in general, the power falls to 1% of the peak value at the neighbourhood points. The peaks are very sharp and unmistakable. Power is not distributed evenly in the various members and the modes in which maximum power is concentrated coincide in both methods.

4. Conclusions

The above analysis indicates that the Kirkwood gaps are essentially the *interference pattern* due to two spatial density waves and not merely due to the three body resonance phenomena between Jupiter, the Sun and the asteroid. This is probably the first clear unmistakable evidence of density waves. This also shows that this phenomenon is due to collective effects between the particles of the system together with the gravitational field provided by the Sun, all the planets, and probably by the stars as well.

Are these matter tubes spirals? It is very difficult to answer this question. It could be a tube spiralling between 2 AU and 4 AU and closing on itself, thereby forming a bounded system or it could also be a tube in which particles spiral out and take matter beyond the orbit of Jupiter. If we assume them to be spiral tubes, then the diameter of these tubes are the distances between the successive minima. To ascertain as to whether they are spiral tubes closed on themselves or not, one should make these density measurements in two siderial longitude directions and see whether there are two extra maxima in one direction over the other. Since we have observed only prograde motions in the asteroids, this possibility seems to be remote. On the other hand, if we have only one extra maxima, we can infer this to be an open spiral throwing matter out into space beyond Jupiter. If we have only the same number of maxima, we can then conclude that the tubes are either circular or elriptical. In the first two cases, the density waves could have a differential rotation effect, while in the last case the waves are purely radial. From the analysis in the next paper we have shown that there can exist differential rotation in the medium, and hence the tubes can be spiral in nature. If the spiral tubes close on themselves, then the motion could have a complex topology.

This analysis reveals yet another striking feature. As we have shown in table 2, the maximum power in the PS lie in low frequency modes and this is a very clear case of a nonstationary process, and therefore a non-Marcovian one. Hence in the study of density waves, one *cannot* resort to hydrodynamical equations, since these equations characterise a Marcovian process. What should be the macroscopic equations which would represent a non-Marcovian situation? Obviously it cannot be the moment of any kinetic equation which describes a Marcovian process. One should evolve a non-Marcovian kinetic equation and then take the moment of this equation to obtain a non-Marcovian hydrodynamics. An attempt has already been made in this direction (Pratap 1975). It may be noted that this feature is evident in every self-gravitating problem, and hence in a system wherein gravitational forces are the most dominant ones, the use of hydrodynamical equations (either Euler or Navier-Stokes equations) can be highly questionable. The case at hand is a direct proof for the same.

In the next paper, we shall develop a kinematical theory for the asteroidal belt using the techniques of stellar dynamics.

Acknowledgements

I am greatly indebted to Suhasini Ravipati for introducing me to the power spectrum analysis, and in particular to the MEM and B and T techniques. I am also thankful to Ranna A Patel for helping me with the computer program. Without these people's help, this work would not have taken the present form.

References

Jeans J H 1961 *Astronomy and Cosmogony* (Dover Reprint)
Kirkwood D 1867 *Meteoric Astronomy* (Chapter 13)
Lin C C and Shu F H 1964 *Ap. J.* **140** 646
Lindblad P O 1962 *Interstellar Matter in Galaxies* (ed. L Woltjer, W A Benjamin, NY)
Ovenden M W 1972 *Nature* **239** (October 27) 508
Pratap R 1975 *General Many Body Problem* [ed. R Pratap and Jitendra C Parikh, Hindusthan Publishing Corporation (India) Ltd., Delhi]
Radofko H R, Fougere P F and Zawalick E J 1975 *JGR* **80** 619