ICARUS 56, 51–74 (1983)

Chaotic Behavior and the Origin of the 3/1 Kirkwood Gap

JACK WISDOM

Department of Physics, University of California, Santa Barbara, California 93106

Received January 17, 1983; revised May 5, 1983

The sudden eccentricity increases discovered by J. Wisdom (*Astron J.* **87**, 577–593, 1982) are reproduced in numerical integrations of the planar-elliptic restricted three-body problem, verifying that this phenomenon is real. Maximum Lyapunov characteristic exponents for trajectories near the 3/1 commensurability are computed both with the mappings presented in Wisdom (1982) and by numerical integration of the planar-elliptic problem. In all cases the agreement is excellent, indicating that the mappings accurately reflect whether trajectories are chaotic or quasiperiodic. The mappings are used to trace out the chaotic zone near the 3/1 commensurability, both in the planar-elliptic problem and to a more limited extent in the three-dimensional elliptic problem. The outer boundary of the chaotic zone coincides with the boundary of the 3/1 Kirkwood gap in the actual distribution of asteroids within the errors of the asteroid orbital elements.

I. INTRODUCTION

It has long been known that the Kirkwood gaps in the distribution of semimajor axes of the asteroids are associated with mean motion commensurabilities of Jupiter (Kirkwood, 1867), yet no one has been able to demonstrate exactly why these commensurabilities should lead to gaps. The greatest obstacle is that there is no real understanding of the long-term dynamics near commensurabilities in the elliptic restricted three-body problem. There is no adequate analytic theory (as there is in the circular restricted problem), and the numerical studies, although numerous, have actually contributed very little toward the understanding that is required. The limited success of the numerical investigations is primarily due to the great amount of computer time required to perform the numerical integrations over times that are long enough to characterize the motion. This difficulty was largely overcome in Wisdom (1982), where a new method was presented for studying asteroid motion near commensurabilities that is approximately 1000 times faster (cheaper) than previous methods. Namely, algebraic mappings of the phase space onto itself with the same low-order resonance structure as the 3/1 commensurability were derived. Trajectories are followed by iteratively evaluating these mappings. Using this new method some very surprising behavior was discovered. Test asteroids placed near the 3/1 commensurability may evolve with low eccentricity ($e < 0.1$) for as much as a million years and then have a sudden increase in eccentricity to large values ($e > 0.3$). Since asteroids near the 3/1 commensurability with eccentricity greater than approximately 0.3 can be Mars crossers, it was hypothesized that the 3/1 Kirkwood gap was cleared by collisions with Mars. As a test of this hypothesis a distribution of 300 test asteroids near the commensurability was "integrated," with a mapping, for 2 million years. When the initial conditions of those test asteroids that became Mars crossing ($e > 0.3$) within this 2-my period were removed, the distribution of initial conditions showed a gap at the proper location, but it was too narrow when compared to the distribution of real numbered asteroids in the TRIAD file. The comparison was somewhat improved when the inclinations and the secular variations of Jupiter's orbit were included, but the predicted gap was still too narrow. A test asteroid near the boundary of the predicted

0019-1035/83 $3.00
Copyright © 1983 by Academic Press, Inc.
All rights of reproduction in any form reserved.

gap was found to have a sudden increase in its eccentricity after 18 my, indicating that the full width of the gap might open over longer times. I closed the paper with the cautionary remark that these results should not be trusted until the sudden increases in eccentricity are seen in numerical integrations or until it is understood how they can happen in an approximate analytic theory. In the present paper both of these are provided and the conflict between the predicted and observed distributions is resolved.

In general, it is expected that every commensurability will be accompanied by a chaotic zone (Chirikov 1979). This chaotic zone may be large or microscopically small depending on many factors. In the planar-elliptic restricted three-body problem it is known that chaotic behavior does occur near commensurabilities. Using Schubart's averaging procedure and the surface of section technique, Giffen (1973) discovered a chaotic zone near the 2/1 commensurability. Following Giffen, Scholl and Froeschlé (1974, 1976) found chaotic orbits near the 3/1 commensurability, but were unable to find any near the 5/2 or 7/3 commensurabilities. Having tested a number of orbits near the 2/1 and 3/1 commensurabilities. Scholl and Froeschlé (1976) concluded that ". . . isolating integrals exist in the main part of the phase space. Ergodic orbits occur seldom."

If indeed test asteroids "diffuse" through phase space until they reach a point where large eccentricity increases occur, as proposed in Wisdom (1982), then their trajectories must have a chaotic character. Quasiperiodic trajectories do not diffuse, except through computer roundoff error. Evidently, it is important to determine if the mappings accurately reflect whether trajectories are chaotic or quasiperiodic, as determined by numerical integrations of the differential equations. After a qualitative review of the mapping derivations in Section II and an explanation of the maximum Lyapunov characteristic exponent and its importance for determining the character of a trajectory in Section III, Section IV exhibits the results of a number of calculations of the maximum Lyapunov characteristic exponent using the differential equations for the unaveraged planar-elliptic restricted three-body problem, and compares them to equivalent calculations using a mapping. In all cases the two agree about whether the orbits are chaotic or quasiperiodic. Note that this is a more stringent test of the mappings than a direct comparison of orbits, as was done in Wisdom (1982). Orbits may look the same for a period of time, but have different characters when viewed over longer times. An exciting byproduct of these calculations is that in several cases the numerical solutions displayed jumps in eccentricity remarkably similar to those displayed by the mapping.

Once it has been verified that the mappings accurately reflect the character of a trajectory, they can be used to trace out the boundary of the chaotic zone. If these sudden eccentricity increases have any bearing on the origin of the 3/1 Kirkwood gap then chaotic behavior must be common, not "seldom." That this is in fact the case is shown in Section V, where the chaotic zones are traced out for the planar-elliptic problem and to a more limited extent in the three-dimensional elliptic problem. Evidently, the integrations of Scholl and Froeschlé were not long enough to detect this chaotic behavior.

Scholl and Froeschlé show that the extra, nonclassical, integral that is obtained in the planar-elliptic restricted problem by Schubart's averaging procedure constrains Giffen's chaotic orbit to eccentricities less than 0.18. There is an analogous quasi-integral for the 3/1 commensurability and it is important to understand to what extent it confines these chaotic orbits. This problem is discussed in Section VI, where it is shown that this quasi-integral does not prevent close encounters with Mars. In the three-dimensional problem the quasi-integral provides even less constraint, and the

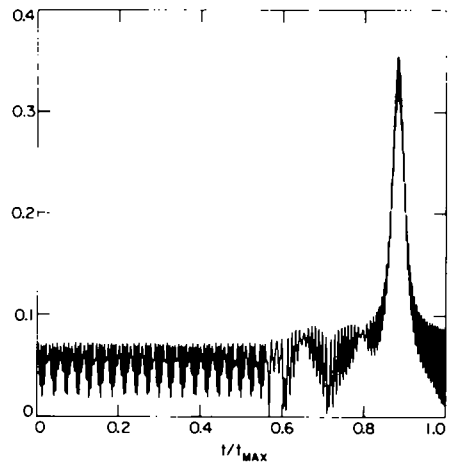

FIG. 1. Eccentricity versus time using the planar-elliptic mapping with initial conditions: $a = a_{3:1} \approx 0.48059a_J$, $e = 0.055001$, $l = \pi$, and $\tilde{\omega} = 0$. t_{max} is 22,000 Jupiter periods, T_J, or approximately 250,000 years.

orbital eccentricity can reach very high values. The quasi-integral may not significantly constrain the eccentricity of Giffen's orbit in the three-dimensional problem either.

The character of trajectories cannot change; a quasiperiodic trajectory cannot become a chaotic trajectory. Nevertheless, the trajectory shown in Fig. 1 begins with a long stretch of apparently quasiperiodic behavior, and then suddenly becomes more chaotic. A possible mechanism for this peculiar behavior is illustrated in Section VII with a simple two-dimensional system. More importantly, this simple system also exhibits apparently random transitions between several modes of behavior, analogous to the apparently random transitions from low-eccentricity behavior to high-eccentricity behavior that is observed for test asteroids.

Finally, it is shown in Section VIII that the outer boundary of the chaotic zone near the 3/1 commensurability agrees with the observed boundary of the 3/1 Kirkwood gap in the actual asteroid distribution within the errors of the asteroid orbital elements. Summary and conclusions follow in Section IX.

II. REVIEW OF THE METHOD

In general, it is possible to reduce the study of a Hamiltonian system to the study of a mapping of the phase space onto itself. This can be accomplished by looking at the successive intersections of a trajectory with a given surface in the phase space, or by looking at the system stroboscopically, i.e., at fixed intervals of time. The latter method is especially useful when the Hamiltonian depends periodically on the time. A mapping gives the state of a system at some intersection in terms of the state of the system at an earlier intersection. If this mapping is explicitly given as an algebraic expression, then the motion may typically be computed 1000 times faster than the corresponding differential equations (Hénon and Heiles, 1964). Thus, it is a tremendous advantage to have an explicit mapping for a problem. Unfortunately, explicit mappings for particular Hamiltonian systems are rarely found. Usually the only way to determine exactly where a point is mapped is to integrate the equations of motion until the section is crossed again. This may be useful, but it gives no computational advantage. However, in some cases where no exact, explicit mapping is known, the system can be approximated by an explicit mapping. A good example is the motion of a charged particle in an axially symmetric magnetic bottle. Chirikov (1979) has shown that the state of this system at successive intersections of a particular plane is well approximated by the mapping

$$I' = I + K \sin \vartheta,$$

$$\vartheta' = \vartheta + I'.$$

The variables are defined in Chirikov (1979); I is related to the magnetic moment and ϑ is a conjugate phase of gyration. Chirikov calls this mapping the standard mapping. The mappings for asteroidal motion near the 3/1 commensurability derived

in Wisdom (1982) are direct generalizations of Chirikov's mapping. As a simple illustration of how mappings can be derived, I will derive this standard map, and then outline the generalizations for asteroidal motion.

Consider the time-dependent Hamiltonian

$$H = \frac{I^2}{4\pi} + \frac{K_0}{2\pi}\cos\vartheta + \sum_{n\neq 0} K_n(I)\cos(\vartheta - nt), \quad (1)$$

where I is a momentum and ϑ is its canonically conjugate coordinate. Resonances occur whenever the argument of one of the cosines is nearly stationary, that is, whenever $I/2\pi \approx \dot{\vartheta} \approx n$. If the K_n are not too large, then the averaging principle (see Wisdom, 1982) says that when I is near zero the system is well approximated by the pendulum Hamiltonian

$$H_A = \frac{I^2}{4\pi} + \frac{K_0}{2\pi}\cos\vartheta. \quad (2)$$

The other terms may be ignored at first because they oscillate rapidly, and generally contribute only periodic oscillations about the motion determined by the averaged Hamiltonian. An important exception occurs for initial conditions near that infinite period trajectory which separates the circulating trajectories from the oscillating trajectories, the separatrix, which is replaced by a narrow chaotic band when the high-frequency terms are present.

If it is a good approximation to ignore the high-frequency terms, it is an equally good approximation to modify them. This is what is done to derive a mapping. Consider the Hamiltonian

$$H_M = \frac{I^2}{4\pi} + \frac{K_0}{2\pi}\cos\vartheta + \frac{K_0}{2\pi}\sum_{n\neq 0}\cos(\vartheta - nt), \quad (3)$$

which differs from Hamiltonians (1) and (2) only in the choice of the nonresonant terms. Since this Hamiltonian contains high-frequency terms the separatrix will again be chaotic, as it was in the original system. In this respect Hamiltonian (3) is superior to Hamiltonian (2), which is completely integrable and yields no chaotic behavior. Hamiltonian (3) may be rewritten in the forms

$$\begin{aligned} H_M &= \frac{I^2}{4\pi} + \frac{K_0}{2\pi}\cos\vartheta\sum_n\cos(nt) \\ &= \frac{I^2}{4\pi} + K_0\cos\vartheta\sum_l\delta(t - 2\pi l) \\ &= \frac{I^2}{4\pi} + K_0\cos\vartheta\,\delta_{2\pi}(t), \end{aligned}$$

where the second equality follows from the well known Fourier transform of the Dirac delta function, and the last equality implicitly defines $\delta_{2\pi}$ as a periodic delta function with period 2π. This Hamiltonian is now "locally" integrable. The equations of motion across a delta function are

$$\dot{I} = -\frac{\partial H_M}{\partial\vartheta} = K_0\sin\vartheta\,\delta(t),$$

$$\dot{\vartheta} = \frac{\partial H_M}{\partial I} = \frac{I}{2\pi}.$$

Since the delta function acts instantaneously and I is finite, there is no change in ϑ. The change in the momentum is

$$\Delta I = K_0\sin\vartheta.$$

Between delta functions the equations of motion are

$$\dot{I} = -\frac{\partial H_M}{\partial\vartheta} = 0,$$

$$\dot{\vartheta} = \frac{\partial H_M}{\partial I} = \frac{I}{2\pi},$$

which yields

$$\Delta\vartheta = I$$

for the change in the coordinate between delta functions. The state of the system just before a delta function can thus be written explicitly as a function of the state just before the previous delta function

$$I' = I + K_0 \sin \vartheta,$$
$$\vartheta' = \vartheta + I'.$$

Ecce, the standard map.

Essentially all studies of the long-term behavior of asteroids near commensurabilities have employed averaged Hamiltonians, Hamiltonians where the high-frequency terms were removed either analytically (following Poincaré, 1902) or numerically (following Schubart, 1964). Thus, in the past, asteroidal studies have concentrated on analogs of Hamiltonian (2). Now, the standard map was derived by adding new high-frequency terms to Hamiltonian (2) so that the system was locally integrable, yielding an analytic mapping of the phase space onto itself. Mappings for asteroidal motion near commensurabilities can be derived in a completely analogous manner. The Hamiltonian for the elliptic restricted three-body problem can be written

$$\begin{aligned} H = &-\frac{\mu_1}{2\alpha} + H_{\text{sec}}(a, e, i, \bar{\omega}, \Omega; \\ &\quad a_J, e_J, i_J, \bar{\omega}_J, \Omega_J) \\ &+ H_{\text{res}}(a, e, i, 3l_J - l, \bar{\omega}, \Omega; \\ &\quad a_J, e_J, i_J, \bar{\omega}_J, \Omega_J) \\ &+ H_{\text{hf}}(\cdots, 3l_J - 2l, \cdots), \end{aligned}$$

where a, e, i, l, $\bar{\omega}$, and Ω are the semimajor axis, the eccentricity, the inclination, the mean longitude, the longitude of the perihelion, and the longitude of the ascending node, respectively. Jupiter's elements carry the subscript J. No approximation has been made, but the terms have been grouped into three parts. The secular part of the Hamiltonian, H_{sec}, contains all those terms in the disturbing function which do not depend on the mean longitudes; the resonant part of the Hamiltonian, H_{res}, contains those terms which involve the resonant combination of mean longitudes (here $3l_J - l$); and the high-frequency part of the Hamiltonian, H_{hf} contains all other terms, the terms with nonresonant combinations of mean longitudes. In order to derive analytic mappings, I first remove all of H_{hf}, as do the disciples of Poincaré and Schubart, and then add new high-frequency terms so that the Hamiltonian is "locally" integrable. It is important to realize that this is essentially the same approximation as that used by Poincaré and Schubart; each uses Hamiltonians which differ from the actual Hamiltonian by high-frequency terms. Symbolically

$$H_{\text{M}} = -\frac{\mu_1}{2a} + H_{\text{sec}} + H_{\text{res}}\delta_{2\pi}(t).$$

Between delta functions the Hamiltonian is analytically integrable if terms of fourth order and higher in the eccentricities and inclinations in H_{sec} are ignored. It is also possible to give analytic expressions for the changes in the elements as the delta functions are crossed, if the high-frequency terms are chosen properly. Except for these few technical problems the derivation is completely analogous to the one above. The result is an analytic mapping of the phase space onto itself which approximates asteroidal motion near the 3/1 commensurability. See Wisdom (1982) for the complete derivations and more thorough discussion of the approximations involved. For completeness and because the mappings used in this paper have minor differences from the ones used in Wisdom (1982), I have included them in the Appendix.

III. MAXIMUM LYAPUNOV CHARACTERISTIC EXPONENTS

In chaotic regions of phase space two initially nearby trajectories separate roughly exponentially with time; in quasiperiodic regions neighboring trajectories separate roughly linearly with time (see Chirikov, 1979). The average rate of separation may be defined by

$$\gamma_2(t) = \frac{\ln\left(\frac{d(t)}{d(t_0)}\right)}{t - t_0}, \tag{4}$$

where d is the usual Euclidean distance between two initially nearby trajectories in their phase space and the subscript 0 refers

to initial values. If the trajectories are chaotic, then d will grow exponentially on the average and $\gamma_2(t)$ will approach some positive constant. If the trajectories are quasiperiodic, then d will grow linearly on the average and $\gamma_2(t)$ will approach zero as $\ln(t)/t$. These two cases are most easily distinguished on a plot of $\log(\gamma_2(t))$ versus $\log(t)$. In practice, this is an excellent method to determine the character of a trajectory, but some refinements are helpful. The worst aspect of this definition is that after some time (depending on the character of the trajectories) these two trajectories may be far apart, and will thus not be reflecting the local rate of separation. There are two ways to overcome this difficulty.

Consider the set of differential equations

$$\frac{d}{dt}\, x = f_i(x), \tag{5}$$

where x represents a point in an n-dimensional space and the subscript i refers to a component. Neighboring trajectories satisfy

$$\frac{d}{dt}\,(x_i + d_i) = f_i(x + d), \tag{6}$$

where d now represents a small n-dimensional displacement. When d is infinitesimally small it satisfies the differential equation

$$\frac{d}{dt}\, d_i = \sum_j d_j \frac{\partial}{\partial x_j} f_i(x), \tag{7}$$

which is found by Taylor expanding the right side of Eq. (6), keeping only the terms linear in d, and subtracting Eq. (5). Remember that the right side depends on x, which is determined through Eq. (5). Notice that since Eq. (7) is now linear in d, the absolute length of d is irrelevant. A new displacement vector, d', which is initially related to d by $d' = cd$, where c is a numerical scale factor, will always be related to d by the same relation, $d'(t) = cd(t)$. The original vector d must be infinitesimally small so that the linear approximation is valid. However, after linearization the scale factor c is arbitrary and may be chosen to make d' computationally convenient i.e., of order unity. The analog of Eq. (4) is

$$\gamma(t) = \frac{\ln\left(\dfrac{\|d\|}{\|d_0\|}\right)}{t - t_0}. \tag{8}$$

$\|d\|$ represents the usual Euclidean length of d. The character of the trajectory is determined as before, by plotting $\log(\gamma(t))$ versus $\log(t)$.

There is a simpler, though somewhat less satisfactory, way to overcome the difficulty. Since the displacement between two neighboring trajectories satisfies a linear differential equation if the displacement is not too large, the displacement can always be kept small by repeatedly restarting the test trajectory near the reference trajectory, as long as the displacement direction is left unchanged. Let r_k be the ratio of the length of d before and after this renormalization at time t_k, then

$$\gamma_l = \frac{\sum_{k=1}^{l} \ln r_k}{t_l - t_0}. \tag{9}$$

A plot of $\log(\gamma_l)$ versus $\log(t_l)$ determines the character of the reference trajectory. In some cases it is more convenient to use this method. Chirikov (1979) found that the two methods gave very nearly the same results.

The limit of $\gamma(t)$ as t approaches infinity is a Lyapunov characteristic exponent (LCE). It has been proven that for almost all initial conditions this limit can take at most n different values (LCE's) as the initial displacement is varied, where n is the dimension of the system. In fact, except for a set of measure zero, all initial displacements lead to the maximum LCE. Thus this method of determining the character of trajectories by plotting $\log(\gamma)$ versus $\log(t)$ is usually referred to as that of the maximum Lyapunov characteristic exponent. For a review of the mathematical results regarding LCE's, see Benettin *et al.* (1980).

As discussed above, there is no theoretical difficulty in defining the average rate of exponential separation of nearby orbits, but there are a number of purely numerical problems. The most serious is the effect of roundoff error on chaotic trajectories. Each time the position is specified some error is made; rather than specifying the original trajectory, a neighboring trajectory is specified. Now neighboring trajectories separate exponentially in a chaotic region of phase space, so the numerically followed trajectory separates exponentially from the originally specified trajectory. Moreover, this exponential splitting happens with each specification of the trajectory, not just at the beginning. The rate of exponential separation is determined by the maximum LCE. As an example, suppose the maximum LCE has the value $\lambda = 10^{-3.5}$/year, which is typical for asteroids near the 3/1 commensurability. If the initial error is, say, 10^{-12}, then the error at time t will be $10^{-12}\exp(\lambda t)$. If the final error is to be no more than 10^{-2} then t is limited to

$$t = \frac{\ln(10^{10})}{\lambda} \approx 80{,}000 \text{ years.}$$

If a numerical integration over 1 my is desired it is necessary to carry out the calculation with over 100 significant digits! Such a calculation is effectively impossible, it is already almost prohibitively expensive with 12 significant digits. Chaotic trajectories are difficult to compute precisely.

Nevertheless, calculations of chaotic trajectories appear frequently in the literature. With few exceptions (e.g., Channon and Lebowitz, 1980) this difficulty in numerically following chaotic orbits is ignored, and calculations are simply carried out with some convenient precision. Of course, the hope is that the computed trajectories are somehow related to actual trajectories. Can this be justified? No, not completely, but some insight is gained through the theorem of Anosov and Bowen (see Benettin *et al.*, 1978, for a discussion). Stated simply, this theorem proves for a very special class of dynamical systems, called Anosov systems, that it is always possible to find a true trajectory that lies near the computed trajectory for all time, if the computed trajectory is computed with sufficient (though not perfect!) precision. The precision required depends on the definition of "near." This is a remarkably useful result for those who want to study Anosov systems numerically. It means not only that trajectories may be computed without worrying about the precision of the computation, but that the statistical properties of the computed trajectory will be real. Unfortunately, most Hamiltonian systems are not Anosov systems. However, Benettin *et al.* (1978) exhibit a number of calculations which yield similar results for Hamiltonian and Anosov systems and argue that a similar theorem may hold for more general Hamiltonian systems. What is more important for this study, Benettin *et al.* (1980) found that the maximum LCE did not depend on the precision of their calculation. It appears likely then that as long as a certain minimum precision is kept, maximum LCE's may be accurately computed, even though it is not possible to precisely follow a specified trajectory for the required length of time. In this paper the most important question is simply whether the trajectory is chaotic or quasiperiodic; the precise value of the maximum LCE will play no role in my discussion of the origin of the 3/1 Kirkwood gap.

IV. THE CHARACTER OF TRAJECTORIES

The calculation of Lyapunov characteristic exponents by numerically integrating the differential equations for the restricted three-body problem is very time consuming. Consequently, very few cases have been studied, and for many of these the interpretation was difficult because the integrations were limited (see Froeschlé and Scholl, 1981). Mappings, on the other hand, are approximately 1000 times faster than other methods for studying asteroidal motion near the 3/1 commensurability. How-

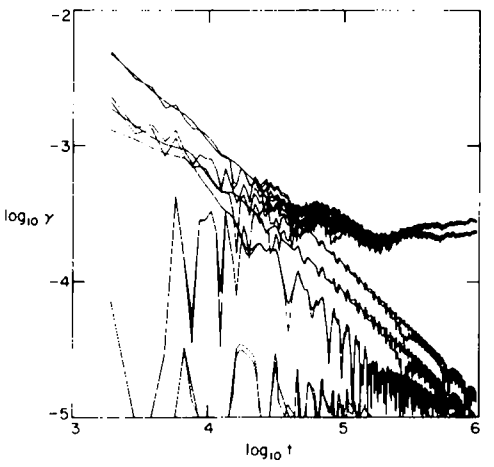

FIG. 2. A total of 16 calculations of the maximum Lyapunov characteristic exponents using the planar-elliptic mapping for two different trajectories, one chaotic and the other quasiperiodic, showing that the results are insensitive to the initial displacement vector, d, and to the precision of the calculation.

ever, before the mappings can confidently be used to explore the phase space for chaotic zones it is essential to first determine if the mappings accurately reflect whether the trajectories are chaotic or quasiperiodic by comparing calculations of maximum LCE's, i.e., $\gamma(t)$, using the differential equations and using a corresponding mapping. Because the mappings are so much faster, it is wise to use them first to get an idea of what to expect, and then verify the behavior with the differential equations.

To what extent does γ depend on the choice of the initial displacement? How long must the integrations be to unambiguously determine whether a trajectory is chaotic or quasiperiodic? How precisely must the trajectories be calculated? All of these questions are answered in Fig. 2. Figure 2 shows several calculations of $\gamma(t)$, using the planar-elliptic mapping and definition (8), for two different trajectories, one chaotic and the other quasiperiodic. The initial conditions for the chaotic trajectory are $a/a_J = 0.48$, $e = 0.15$, $l = \pi$, and $\tilde{\omega} = 0$. The perihelion of Jupiter is taken as the origin of longitudes. The mass of Jupiter divided by the sum of that of the Sun and Jupiter is taken to be 1/1047.355, and Jupiter's eccentricity has the value 0.048. The initial mean longitude of Jupiter is zero. The initial conditions of the quasiperiodic trajectory are the same except $a/a_J = 0.4795$. For each trajectory γ is computed for four independent choices of the initial displacement d, in both single (7 digits) and double precision (16 digits), a total of 16 calculations of γ. The eight γ's for the chaotic trajectory all appear to be approaching a constant value near $10^{-3.5}$; the eight γ's for the quasiperiodic trajectory all appear to be approaching zero with the characteristic $\ln(t)/t$ behavior. In all cases the behavior is independent of the initial displacement and of the precision of the calculation. Figure 2 is intentionally complex to illustrate that typically the two qualitatively different types of behavior only begin to separate after approximately 60,000 years.

How does this compare to the differential equations? In Fig. 3 are displayed calculations of $\gamma(t)$ for these same two trajectories using the differential equations for the planar-elliptic restricted three-body problem. The constants are all set as before. The co-

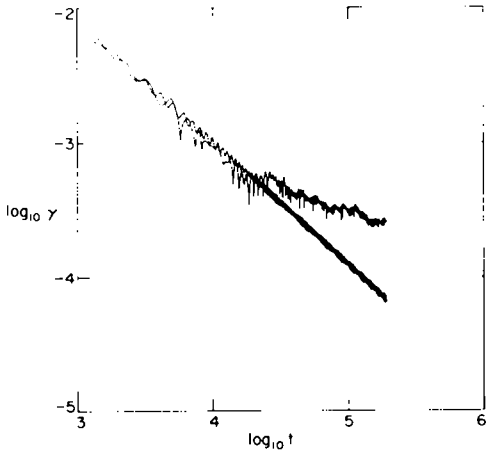

FIG. 3. Calculations of the maximum Lyapunov characteristic exponent for the same two trajectories as in Fig. 2, but using the unaveraged differential equations for the planar-elliptic restricted three-body problem, verifying that the mapping accurately reflects the character of these trajectories.

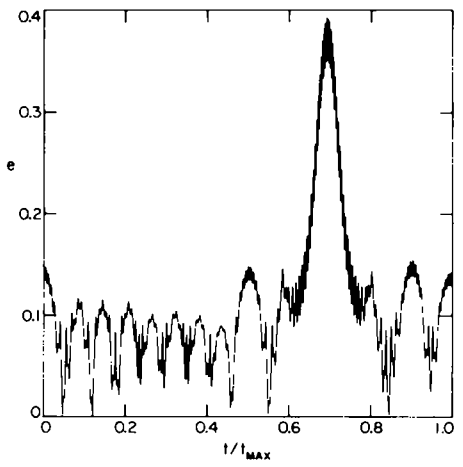

FIG. 4. Eccentricity versus time using the unaveraged differential equations for the planar-elliptic restricted three-body problem for the chaotic trajectory used in Figs. 2 and 3, showing a jump in eccentricity quite similar to those seen with the planar mapping. Here, t_{max} is $16.000T_J$ or approximately 200,000 years.

ordinates used in the integrations are those of Scheibner (1866), where the true longitude of Jupiter is used as the independent variable. This choice is convenient since the position of Jupiter is explicitly given in these coordinates, whereas for other coordinates, such as rotating or nonrotating Cartesian coordinates with time as the independent variable, the position of Jupiter must be computed. The algorithm of Bulirsch and Stoer (1966) was used to perform the integrations. For the mapping calculations presented above, the same result was obtained in both single and double precision. This indicates that the minimum precision required to capture the correct behavior, which was discussed in the last section, is no greater than 7 significant digits. I therefore chose 10^{-8} as the relative accuracy per integration step, which was on the order of 1 year. The mapping calculations indicated that integrations of less than 60,000 years would be of no value in determining the character of trajectories. These numerical integrations cover a time interval of approximately 200,000 years. The mapping calculations also showed that the behavior was independent of the initial displacement. Thus one arbitrary choice is sufficient. The behavior shown in Fig. 3 is essentially identical to that shown in Fig. 2; both trajectories have exactly the behavior expected of them, and even seem to be approaching the same values.

An exciting by-product of this calculation was that the chaotic trajectory showed a sudden increase in eccentricity, remarkably similar to the sudden increases found with the planar-elliptic mapping. Figure 4 shows the eccentricity versus time for this trajectory as calculated by the differential equations. Now this numerical integration was performed with fairly low precision to determine the character of the trajectory, not to follow it exactly. Since the maximum LCE is near $10^{-3.5}$ and only 8 digits were computed, validity can only be claimed for segments of approximately 50,000 years, even though it is still possible that some real trajectory shadows this computed trajectory over the whole 200,000-year interval. Nevertheless, 50,000 years is long enough to span both the jump and the transition from low-eccentricity behavior. To still further verify this behavior a 50,000-year segment surrounding the jump was recomputed with a relative accuracy of 10^{-10} per integration step, with identical results. The jump appears to be a real phenomenon of the differential equations, not an artifact of the mapping.

Froeschlé and Scholl (1981) calculated the maximum LCE for two trajectories near the 2/1 commensurability. They found very little difference in the behavior of γ for the chaotic orbit found by Giffen (1973) and a nearby quasiperiodic trajectory. Their calculation may be criticized on two points. First, they plot $\log(\gamma)$ versus time, t, rather than $\log(t)$. Frequently, for chaotic trajectories γ first tends to zero as $\ln(t)/t$ as it does in the quasiperiodic case, but then levels out near its limit value. It is very difficult to distinguish this leveling out from the $\ln(t)/t$ behavior on a plot that is linear in time. Second, they calculate γ over a time inter-

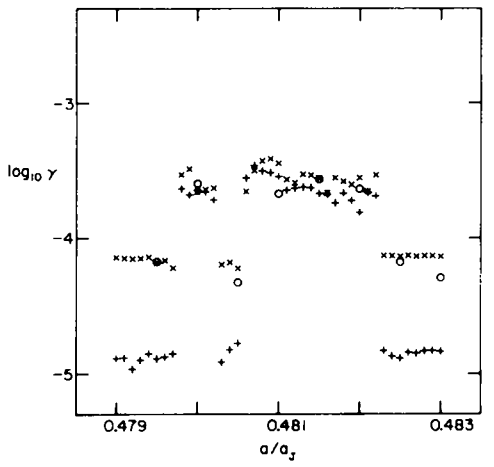

FIG. 5. Calculations of the maximum Lyapunov characteristic exponent for a set of initial conditions which span the 3/1 Kirkwood gap. The results using the differential equations are represented by circles (200,000 years); the mapping results are represented by crosses (200,000 years) and plus signs (1,000,000 years). The chaotic zones are readily apparent, and in all cases the agreement between the two methods is excellent, indicating that the mapping accurately reflects the width of the chaotic zone.

val of only 20,000 years. It is not known how long their calculations would have to have been to see an unambiguous difference in behavior, but judging from the difficulty of interpretation their calculations were simply not long enough. For Giffen's orbit, γ appeared to be approaching a value near 10^{-3}. This was smaller than Froeschlé and Scholl expected. They concluded that Giffen's trajectory should not be called "ergodic," the word commonly used at the time to describe chaotic behavior, but instead invented for it the new term "quasiergodic." However, a small value for the maximum LCE is simply an indication that one must look at the trajectory for a long time to become aware of its nonintegrable character. This value is comparable in magnitude to the value found above, $10^{-3.5}$, for a chaotic trajectory near the 3/1 commensurability. In that case a clear determination of the character of the trajectory could only be made after 60,000 years, but on a timescale of millions of years similar trajectories definitely appear chaotic (cf. Fig. 12). This time scale is still short when motion over the age of the solar system is being considered. All calculations of γ in this paper (except one) are carried out over at least 200,000 years.

For two trajectories it has been verified that the mapping accurately reflects whether the motion is chaotic or quasiperiodic. Does the planar-elliptic mapping accurately reflect the extent of the chaotic zone? To answer this question it is again necessary to compare calculations of γ using differential equations and using the corresponding mapping. As before, the mapping is studied first in order to see what range of initial conditions are interesting, and then the observed behavior is verified with the differential equations. Since the width of the chaotic zone is of fundamental importance, I have chosen to study a set of initial conditions on a line which crosses the commensurability. Namely, a is varied to span the commensurability, while $e = 0.15$, $I = \pi$, and $\tilde{\omega} = 0$. This eccentricity is near the mean eccentricity for asteroids near the 3/1 commensurability. The reason for choosing these particular values for the other two coordinates will be explained in the next section. Figure 5 displays the results of a number of calculations of γ, using definition (9) with an initial displacement of length 10^{-7} and a renormalization every 100 years. The values of $\log_{10}(\gamma)$ at $t = 200{,}000$ years (crosses) and $t = 1{,}000{,}000$ years (plus signs) are plotted versus the initial semimajor axis. If a trajectory is chaotic γ tends to a constant, while for a quasiperiodic trajectory it tends to zero as $\ln(t)/t$. Thus as long as enough time has passed, the two types of behavior may be distinguished simply by the value of γ. At a later time γ should be roughly the same for chaotic trajectories, but smaller for the quasiperiodic trajectory by approximately the ratio of the times. These features may be seen in Fig. 5, and readily distinguish the two types of behavior. In the chaotic zones, all values of γ

cluster near $10^{-3.5}$; in the quasiperiodic zones, γ takes smaller values with the ratio of the two γ's for each semimajor axis approximately the same as the ratio of the times, 1/5, as expected.

The results of several calculations of γ using the differential equations and definition (8) are also displayed in Fig. 5 (circles). The initial conditions are the same as before, but now there are only eight evenly spaced values of semimajor axis: 0.4795, 0.4800, 0.4805, . . . , and 0.4830. The details of the calculations were the same as the previously discussed calculations using the differential equations. In particular, each integration spanned approximately 200,000 years, except the one beginning with $a = 0.4815$ which spanned 160,000 years. It is tempting to use dollar (or pound) signs here, rather than the circles, since these were by far the most time consuming calculations in this paper, occupying a total of 100 hr of VAX 11/750 time. In every case the mapping and the differential equations agree on the character of the trajectories and in the magnitude of γ. Even the narrow quasiperiodic zone at $a = 0.4805a_J$ is reproduced. The agreement with the mapping results is better than might have been expected. The false high-frequency terms introduce variations in the semimajor axis of the same order as the ratio of the mass of Jupiter to the mass of the Sun. Thus the mappings might have been expected to make errors in the boundaries of the chaotic zone of order $0.001a_J$. I will just mention that several more examples of jumps in eccentricity were found in these integrations.

The planar-elliptic mapping has passed all tests with flying colors. In every case considered, the planar-elliptic mapping and the differential equations for the planar-elliptic restricted three-body problem give the same behavior for γ, both qualitatively and quantitatively. This clears the way for using the planar-elliptic mapping as a tool to explore the phase space near the 3/1 commensurability for chaotic zones.

V. THE CHAOTIC ZONE NEAR THE 3/1 COMMENSURABILITY

Having shown that the mappings reliably reflect the character of trajectories, it is now possible to take advantage of their great speed and systematically trace out the chaotic zone near the 3/1 commensurability. Immediately, though, there is a problem. The space of initial conditions for the planar-elliptic mapping is four-dimensional. It is impractical to attempt to systematically explore such a space, or to expect to be able to visualize the results of such a systematic exploration. Fortunately, there is an easier way. If it were true that all trajectories eventually cross some particular two-dimensional plane through the phase space, then a systematic study of initial conditions on this plane would be equivalent to a study of the entire four-dimensional space. Does such a plane exist for the planar-elliptic problem, and how can it be found? There are two facts to be considered. First, it is possible for $l + 2\bar{\omega} - 3l_J$ to librate about π. Thus, depending on the amplitude of the libration, this resonant argument can avoid values near zero but not those near π. This gives the first condition for this representative plane of initial conditions: $l + 2\bar{\omega} - 3l_J = \pi$. Second, far away from the exact commensurability, the motion of $\bar{\omega}$ is well described by the secular part of the Hamiltonian alone. The solution of this Hamiltonian is trivial (if fourth-order terms are ignored) and may be interpreted in terms of a forced and a free eccentricity, viz.

$$e\cos(\bar{\omega} - \bar{\omega}_J) = e_{\text{forced}} + e_{\text{free}}\cos(A(t - t_0)),$$

$$e\sin(\bar{\omega} - \bar{\omega}_J) = e_{\text{free}}\sin(A(t - t_0)).$$

The two constants e_{forced} and A are determined by the equations of motion; the constants e_{free} and t_0 are determined by the initial conditions. Notice that if e_{free} is smaller than e_{forced} then $\bar{\omega} - \bar{\omega}_J$ does not go through all values. However, for all initial conditions $\bar{\omega} - \bar{\omega}_J$ goes through zero. This gives

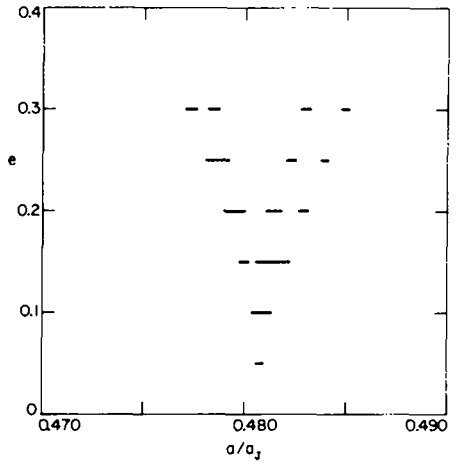

FIG. 6. The chaotic zone near the 3/1 commensurability using the planar-elliptic mapping with $e_J = 0.048$. The initial conditions were on the representative plane and evenly spaced by .0001 in a/a_J for the six values of the eccentricity: 0.05, 0.10, . . . and 0.30. A small cross is plotted whenever the maximum LCE indicated chaotic behavior. The chaotic zone is easily interpolated to other eccentricities.

the second condition for the representative plane: $\bar{\omega} - \bar{\omega}_J = 0$. The first condition must hold for librating trajectories; the second for trajectories away from the exact resonance. The simplest guess for a representative plane is the plane which is uniquely determined by taking these two conditions to hold everywhere. Experimentally, I have found that this is a good choice. I have tested several hundred initial conditions chosen at random in the four-dimensional phase space, as well as all the real numbered asteroids and the PLS asteroids with quality 1 or 2 orbits. In every case, this plane is crossed. I conclude that if there are trajectories which do not cross it they are rare. The existence of this representative plane enormously reduces the problem of systematically exploring the phase space.

Figure 6 shows the results of one systematic exploration using the planar-elliptic mapping, with $e_J = 0.048$. Initial conditions were chosen on the representative plane, $l - 3l_J = \pi$ and $\bar{\omega} = 0$, where the longitude of the perihelion of Jupiter is again taken as the origin of longitudes. The semimajor axes, measured in units of Jupiter's semimajor axis, were evenly spaced by 0.0001 from 0.47 to 0.49. The eccentricities were given the values 0.05, 0.1., 0.15, 0.2, 0.25, and 0.3. For each combination of semimajor axis and eccentricity $\gamma(t)$ was computed in double precision using definition (9) over a time interval of 300,000 years. The character of each trajectory was then determined by plotting $\log_{10}(\gamma(t))$ versus $\log_{10}(t)$. Wherever the trajectory showed a chaotic character, a small cross is plotted in Fig. 6; nothing is plotted for a quasiperiodic trajectory. Figure 6 must be interpolated for other eccentricities. There is a sizable chaotic zone.

This macroscopic chaotic zone is present in the elliptic-restricted problem, but in the circular-restricted problem only a very small chaotic zone near the boundary of libration is anticipated (Chirikov, 1979). Now, Jupiter's eccentricity varies from approximately 0.03 to 0.06. It might be expected that the chaotic zone would be larger with $e_J = 0.06$. Figure 7 shows the same calculation as Fig. 6, but with $e_J = 0.06$. The chaotic zone is not dramatically different. Some of the holes have been filled, but the overall width is actually somewhat smaller.

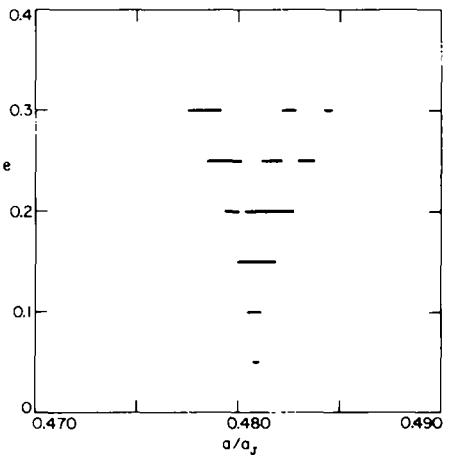

FIG. 7. The same as in Fig. 6, but with $e_J = 0.06$. The chaotic zone is more solid, but surprisingly somewhat narrower.

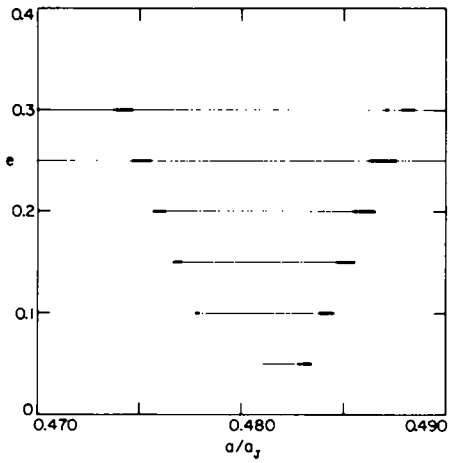

FIG. 8. The chaotic zone on the plane $l - 3l_J = \pi$ and $\tilde{\omega} = \pi$ which emphasizes the width of the region of quasiperiodic libration. A small cross again represents a chaotic trajectory, but this time a small dash is plotted for a quasiperiodic trajectory which reaches eccentricities above 0.3, i.e., can cross the orbit of Mars. Most quasiperiodic librators can cross the orbit of Mars.

To give some idea of how the chaotic zone looks on other planes through the phase space, Fig. 8 shows the chaotic zone for initial conditions on the plane $l - 3l_J = \pi$ and $\tilde{\omega} = \pi$. As before, a small cross is plotted where the trajectory was chaotic. This time a dash is plotted when the trajectory was quasiperiodic and potentially Mars crossing ($e > 0.3$), and nothing is plotted for a quasiperiodic trajectory which cannot cross the orbit of Mars. This plane of initial conditions emphasizes the width of the region of quasiperiodic libration, but also illustrates that most of the trajectories started within the libration region on this plane have large enough variations in eccentricity to become Mars crossers. Only a small region with initial eccentricity 0.05 is safe; most quasiperiodic librators cross the orbit of Mars. This conclusion is supported by the study of a random distribution of 300 test asteroids. Despite its large apparent size in Fig. 8, the quasiperiodic libration region occupies a much smaller volume in phase space than the chaotic zone. Of the 300 test asteroids, 89 were chaotic while only 11 were quasiperiodic librators. None of the quasiperiodic librators could avoid Mars.

In Wisdom (1982) I mentioned that trajectories seemed to have more freedom in the three-dimensional problem, i.e., the problem including the inclinations, compared to the planar problem. It is necessary to consider how the chaotic zone depends on the inclination, i, and the longitude of the node, Ω. Unfortunately, even with a representative plane this problem is again four dimensional. I can only give a taste of what happens. The three-dimensional mapping is presented in the Appendix. It has not been compared to numerical integrations as was the planar-elliptic mapping, but this mapping is a straightforward extension of the earlier one. There is no reason to expect that one should work and the other not work. I first studied the set of initial conditions used in Fig. 5 with several different choices for the inclination and node. The mean inclination of the asteroids is about 5 degrees. I studied inclinations of 5, 10, and 15 degrees, and longitudes of the node equal to 0 and 90 degrees. I found the somewhat curious result that when the node was at 90 degrees the chaotic zone was identical to the planar result, regardless of inclination. The results for $i = 10$ degrees and $\Omega = 0$ degrees are depicted in Fig. 9. The chaotic zone is now solid, and considerably wider at low eccentricities than in the planar problem.

In summary, chaotic behavior is common in the planar-elliptic restricted three-body problem, and even more common in the three-dimensional problem. There is also a small region of quasiperiodic libration, but most trajectories in this region can cross the orbit of Mars.

VI. CONSTRAINTS IMPOSED BY THE AVERAGED HAMILTONIAN

Chaotic behavior is common, but of what consequence is it? None, if there are other factors which constrain the motion to some

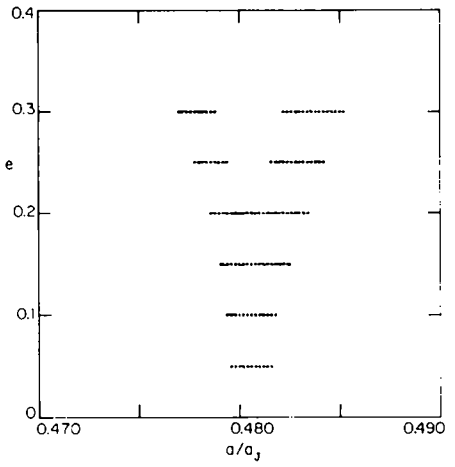

FIG. 9. The chaotic zone using the three-dimensional mapping. The initial conditions are the same as in Fig. 6 with $i = 10$ degrees and $\Omega = 0$ degrees. As the inclination increases the chaotic zone becomes broader and more solid.

small region of phase space. There are no known integrals in the planar-elliptic restricted three-body problem. Near the 3/1 commensurability, however, the averaged Hamiltonian provides a quasi-integral. If the high-frequency part of the Hamiltonian is ignored and the explicit time dependence of the averaged Hamiltonian is removed by a canonical transformation, then the resulting Hamiltonian is then a strict integral of the motion. When H_{hf} is included this time-independent version of the averaged Hamiltonian is no longer a strict integral, but exhibits periodic oscillations. Nevertheless, it is approximately conserved, and provides a constraint on the motion. For convenience, I will call the time-independent version of the averaged Hamiltonian simply the averaged Hamiltonian. To what extent are the chaotic trajectories near the 3/1 commensurability confined by the averaged Hamiltonian?

Some idea of the range of motion allowed can be found by restricting attention to the representative plane. Figure 10 shows lines of constant averaged Hamiltonian on the representative plane as a function of a/a_J and e, plotted with the chaotic zone as determined in Fig. 6. For many values of the averaged Hamiltonian there are two lines, labeled with the same number. It is expected that every time the trajectory crosses this plane it must cross near the line on which it started or near the other line with the same value of the averaged Hamiltonian. Now, quasiperiodic trajectories are not only distinguished from chaotic trajectories by the rate of separation, but also by the existence of a full set of integrals of motion. A quasiperiodic trajectory, except under very special circumstances, can only cross a particular plane at a finite number of points. On the contrary, chaotic trajectories are confined only by the near conservation of the averaged Hamiltonian; there are no extra integrals. A trajectory which is started in a chaotic zone is free to wander, as long as it always crosses the representative plane near one of its lines of constant averaged Hamiltonian. The range of possible intersections with the representative plane may be read directly off this plot, with a little interpolation of the chaotic

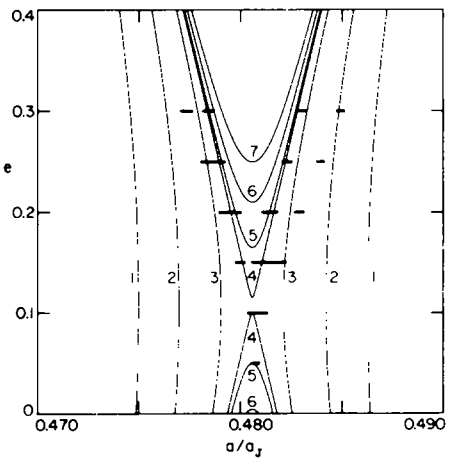

FIG. 10. Curves of constant averaged Hamiltonian on the representative plane, plotted with the chaotic zone found in Fig. 6. Trajectories are constrained to cross the representative plane near one of the curves with the same value of the averaged Hamiltonian as the curve on which the trajectory began. Curves with the same value of the averaged Hamiltonian are labeled with the same number.

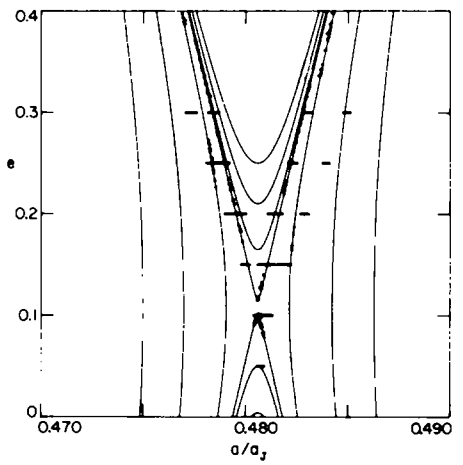

FIG. 11. Intersections of two trajectories with the representative plane, plotted with Fig. 10. Trajectories started in a chaotic zone seem to reach all points where their curves of constant averaged Hamiltonian cross the chaotic zone. All chaotic trajectories can eventually cross the orbit of Mars.

zone from the calculations. One simply looks for all places where the line of constant averaged Hamiltonian is within the chaotic zone. Is it possible to reach all such points? In particular, is it possible to jump from one curve of constant averaged Hamiltonian to the other curve with the same value of the averaged Hamiltonian? Whether or not this is possible depends on the properties of the trajectories in the full four-dimensional phase space. It is possible that a trajectory started on the lower curve labeled 3 would be forever confined to low eccentricity, even though the upper curve labeled 3 remains in the chaotic zone at very large eccentricity. This is an experimental question. On Fig. 11 I have replotted Fig. 10 along with the successive intersections with the representative plane of two trajectories for which curves were drawn. Their initial conditions are a = 0.4806, e = 0.1 and a = .4822, e = 0.15. Initially both trajectories have $l = \pi$ and $\tilde{\omega}$ = 0. Actually, it is not possible to wait for exact intersections of the representative plane, and a and e were plotted whenever $l - 3l_J$ and $\tilde{\omega}$ were both within 5 degrees of the representative plane. The results are not sensitive to the exact size of this window. First, Fig. 11 verifies that the averaged Hamiltonian is approximately conserved for chaotic trajectories. Second, these two trajectories illustrate that it is indeed possible to jump from one curve to the other, both between upper and lower curves and between left and right curves. Several other trajectories were studied, and in all cases the full range possible seems to be explored. However, in some cases the trajectory seems to show some reluctance to changing curves. For example, the first of the two trajectories just mentioned spends a long time on the lower curve before appearing on the upper one. The existence of two curves with the same value of the averaged Hamiltonian is the first clue to an explanation of how trajectories can spend large intervals of time with low eccentricity and then suddenly jump to large eccentricity.

It is difficult to extrapolate Fig. 6 to eccentricities much above 0.3, but it appears that in the planar problem these curves of constant averaged Hamiltonian all eventually leave the chaotic zone. That means that all chaotic trajectories have a maximum eccentricity above which they cannot go. However, meditation on Fig. 11 reveals that this maximum eccentricity is always greater than 0.3, the eccentricity for which Mars crossing becomes possible. Thus even in the planar problem, all chaotic trajectories can cross the orbit of Mars.

The chaotic zone is also present in the three-dimensional problem, and even seems to widen as the inclination increases. The averaged Hamiltonian still provides a constraint on the motion, but this is only one constraint in a six-dimensional phase space. The available phase space is much larger in the three-dimensional problem. This explains the apparent freedom of three-dimensional trajectories. Figure 12 illustrates the typical behavior of a test asteroid with a chaotic trajectory in the planar problem. The initial conditions are a =

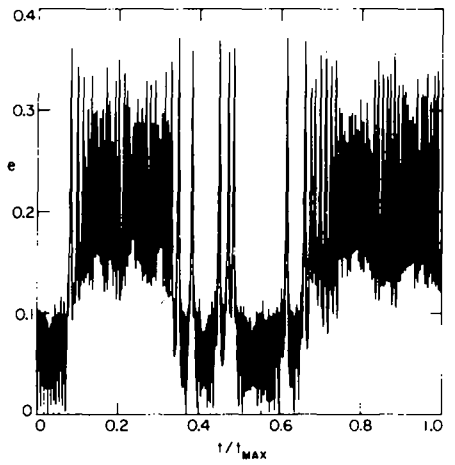

FIG. 12. Eccentricity versus time for a typical chaotic trajectory near the 3/1 commensurability using the planar-elliptic mapping. The eccentricity seems to be limited to values less than 0.4. t_{max} is $200,000T_J$ or approximately 2.4 my.

0.4806, $e = 0.098$, $l = \pi$, and $\tilde{\omega} = 0$. For periods of time the trajectory has low eccentricity and for other periods it has high eccentricity, with apparently random transitions between the two modes of behavior. In this 2.4-my interval, the eccentricity seems to be limited to values less than 0.4. Some planar trajectories display the particularly interesting behavior shown in Fig. 13. The initial conditions are $a = 0.4806$, $e = 0.05$, $l = \pi$, and $\tilde{\omega} = 0$. Here the eccentricity makes only jumps for a short time to high values, again at random intervals. Evidently the trajectory is so confined that the path to large eccentricity is very narrow. This phenomenon seems to be fairly rare, since no test asteroids in the random distribution displayed it. It also does not seem to occur in the three-dimensional problem where trajectories are less confined. Figure 14 shows a typical chaotic trajectory in the three-dimensional problem. The initial conditions are the same as in Fig. 13 with $i = 10$ degrees and $\Omega = 0$. Notice that the eccentricity scale is different. Much higher eccentricities are reached with the three-dimensional mapping. (The inclination seems to execute a random walk, and can also reach large values). Of course, when the eccentricity is so large the mapping is no longer valid since fourth-order terms in ec-

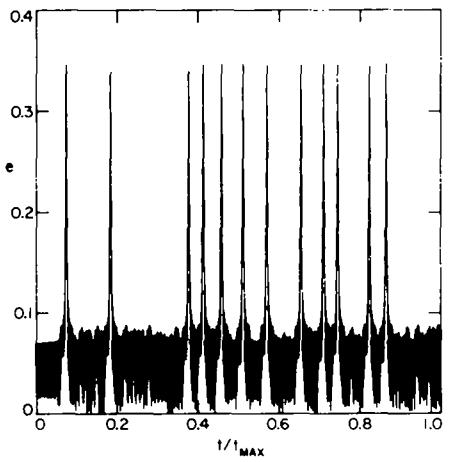

FIG. 13. Eccentricity versus time for a particularly interesting chaotic trajectory in the planar problem. This "intermittent" behavior seems to be fairly rare; initial conditions chosen at random do not display it. It also does not seem to occur in the three-dimensional problem. t_{max} is the same as in Fig. 12, approximately 2.4 my.

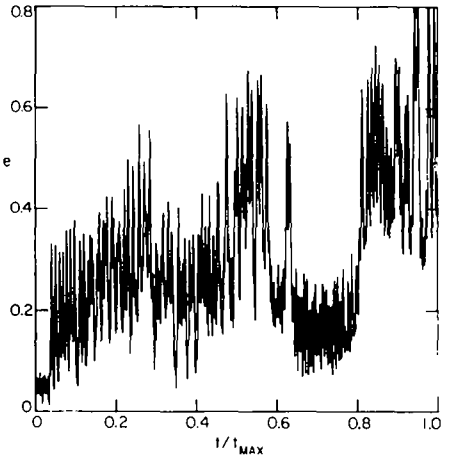

FIG. 14. Eccentricity versus time for the same chaotic trajectory as in Fig. 13, but computed using the three-dimensional mapping with an initial inclination of 5 degrees. Notice the eccentricity scale; much larger variations in eccentricity are obtained with the three-dimensional mapping. t_{max} is the same as in Fig. 12, approximately 2.4 my.

centricity were ignored in its derivation. Thus, it is not known exactly how large the eccentricity may grow in the actual three-dimensional problem. This is a vital question, for the halflife of an asteroid depends strongly on its eccentricity. The typical lifetime for a Mars crosser is on the order of 200 my (Wetherill 1975). The trajectory shown in Fig. 12 spends about 5% of the time as a Mars crosser. Thus the lifetime for a typical chaotic trajectory computed in the planar approximation is on the order of the age of the solar system. On the other hand the typical three-dimensional trajectory shown in Fig. 14 reaches eccentricities large enough to be Earth crossing ($e > 0.6$) and spends considerably more time as a Mars crosser (roughly 50%). While the asteroid is Earth crossing its halflife is considerably shorter, approximately 10 my (Wetherill, 1975). Thus it appears that collisions or close encounters with the Earth or Mars are more than adequate to remove all asteroids with chaotic trajectories within the age of the solar system.

Scholl and Froeschlé (1976) showed that Giffen's chaotic trajectory was confined by the averaged Hamiltonian to eccentricities less than 0.18. Except for the possibility that over very long times the trajectories escape their confinement through Arnol'd diffusion, their conclusion was that chaotic behavior has nothing to do with the origin of the Kirkwood gaps. However, their calculations were performed for the planar case. The above calculations showed that trajectories near the 3/1 commensurability reach much larger eccentricities in the three-dimensional problem than in the planar problem. I expect that this will also be true for trajectories near the 2/1 commensurability; the trajectories are confined in the planar problem, but probably not as confined in the three-dimensional problem.

All chaotic trajectories near the 3/1 commensurability can cross the orbit of Mars. In the three-dimensional problem very large eccentricities are reached, probably large enough for the asteroids to be Earth crossers. The collision probabilities with the Earth and Mars appear adequate to remove all of the asteroids with chaotic trajectories within the age of the solar system.

VII. JUMPING AND STICKING PHENOMENA

At first sight, Figs. 1, 12, and 13 display rather peculiar behavior of the eccentricity as a function of time. Do other Hamiltonian dynamical systems behave this way, and under what circumstances are similar phenomena to be expected? One of the most extensively studied area preserving mappings on the plane is the standard map of Chirikov, presented in Section II. It can exhibit similar phenomena.

In Fig. 15 a number of trajectories of the standard map for $K = 0.96$ are displayed. This shows the familiar mixture of chaotic and quasiperiodic behavior. Focus attention on the chaotic band surrounding the large central island. Successive iterations of the mapping move through this band in a number of modes. In one mode successive iterations appear near the lower boundary of the chaotic band with the angle ϑ rotating towards negative values. In another mode the trajectory moves near the upper

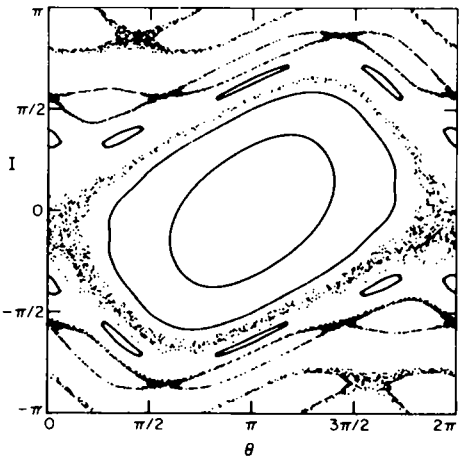

FIG. 15. 1000 successive iterations of the standard map with $K = 0.96$ for each of nine different initial conditions, showing the now familiar divided phase space, originally discovered by Henon and Heiles (1964).

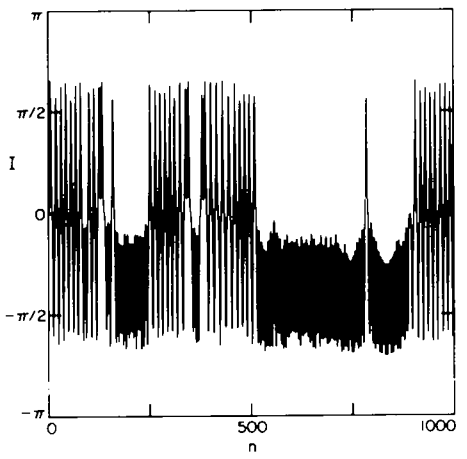

FIG. 16. I versus iteration number n for the chaotic trajectory which surrounds the large central island in Fig. 15, showing the apparently random transitions between the possible modes. This plot looks remarkably similar to Fig. 12. The standard map thus provides a model for understanding the behavior found with the planar-elliptic mapping.

boundary with ϑ rotating towards positive values. Finally, the trajectory may rotate around the large central island while ϑ oscillates. As a chaotic trajectory evolves each of these modes is entered; the transitions between modes occurring after apparently random intervals. Figure 16 illustrates these apparently random transitions between modes by plotting I versus iteration number, n. For convenience, successive iterations have been connected with a continuous line. This chaotic band may be viewed as resulting from the broadening of the separatrix of the pendulum Hamiltonian due to the high frequency contributions. In the pendulum, the two branches of the separatrix cross at the unstable equilibrium. This unstable equilibrium is still present in the standard map and provides the knife edge which separates the modes. Trajectories which approach this unstable equilibrium eventually fall away from it, but may go in two possible directions. Whether they stay in the same mode or enter a new one can depend very sensitively on the trajectory.

Figures 12 and 16 show similar behavior. It is possible that similar processes are at work. However, because of the large dimensionality of the planar-elliptic restricted three-body problem it is difficult to show an exact correspondence. As described above, one essential feature in this process is the presence of an unstable equilibrium. If this process is at work in the three-body problem, there must be an analogous unstable equilibrium with eccentricity near 0.1. In fact, there are two candidates. One was discovered by Hill (1902) with eccentricity near 0.08, and another by Sinclair (1970). The unstable periodic orbit found by Sinclair appears on Fig. 10 as a saddle point near $e = 0.11$. As expected, both of these orbits are fixed points of the planar-elliptic mapping. (Bien (1980) found a third periodic orbit, this one stable, at very large eccentricity, $e = 0.79$. This one is not a fixed point of the planar-elliptic mapping, but this is not surprising since, again, fourth-order terms in eccentricity were ignored in deriving the mappings.) The existence of these unstable equilibria with eccentricities near 0.1 support the picture that the process described above is operating in the restricted problem. In any case, the example of the standard map shows that such behavior occurs in other dynamical systems, even if the details of the process are different.

Figure 1 has one other unusual feature. In the beginning the trajectory appears very regular, almost quasiperiodic, and then suddenly becomes more irregular. Now, quasiperiodic trajectories cannot become chaotic trajectories; their character is fixed. However, it is possible for a chaotic trajectory to appear quasiperiodic for some finite interval of time. As a chaotic trajectory evolves, it eventually comes near the boundary of the chaotic zone, and can stay there for a fairly long time. It can "stick" to the boundary. While near the boundary it behaves in a very similar manner to the neighboring quasiperiodic trajectories. For periods of time chaotic trajectories may look quasiperiodic. Just below the large

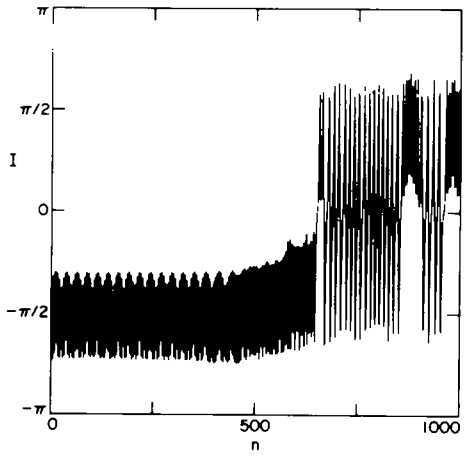

FIG. 17. The unusual behavior displayed in Fig. 1 has this analog in the standard map. Initially the trajectory is "stuck" to a chain of islands, making the trajectory appear almost quasiperiodic. Eventually the trajectory breaks free and begins to explore the full chaotic zone.

central island in Fig. 15 there is a chain of four smaller islands. If a trajectory is started very close to one of these islands, it can "stick" to them for a period of time before exploring the full chaotic zone. In Fig. 17, I is plotted versus number of iteration for such a trajectory. The resemblance to Fig. 1 is striking.

Of course, if the trajectory in Fig. 1 is "sticking" to some quasiperiodic region, it must be near one. The chaotic zone has now been traced out in Fig. 6, and indeed the initial conditions for the trajectory in Fig. 1 ($a/a_J = 0.48059$ and $e = 0.055001$) lie near the chaotic-quasiperiodic boundary. Since this behavior is associated with the boundary of the chaotic zone, and not the chaotic zone itself, initial conditions chosen at random should not display this behavior. It was only by chance that the first trajectory I studied with the planar-elliptic mapping had this behavior. Out of the 600 random initial conditions studied in Wisdom (1982) only one showed this "sticking" behavior.

All of the unusual features found in the behavior of the planar-elliptic mapping have analogs in the behavior of the standard map. It seems probable that similar mechanisms are at work.

VIII. COMPARISON WITH THE ACTUAL DISTRIBUTION OF ASTEROIDS

Figures 6 through 9 trace the character of trajectories near the 3/1 commensurability on certain planes through the phase spaces of the planar and three-dimensional elliptic-restricted three-body problem. Since the elements of the real asteroids are, naturally, not on these planes, a direct comparison is not possible. In order to make a comparison the trajectories of the real asteroids must be followed until they cross the appropriate plane through their phase space. Figure 18 shows the semimajor axes and eccentricities of the real numbered asteroids (circles) and the Palomar–Leiden asteroids with quality 1 or 2 orbits (plus signs) upon the first pass within two degrees of the representative plane. The trajectories were followed with the three-dimensional mapping. One trajectory near the small a border of the chaotic zone was also followed with the

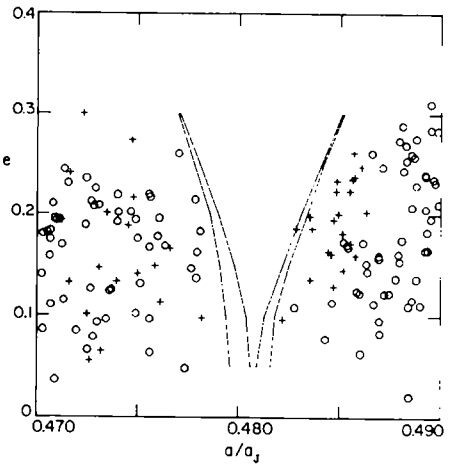

FIG. 18. The outer boundaries of the chaotic zones found in Fig. 6 (planar problem) and Fig. 9 (10 degree inclination) are displayed with the intersections of the real numbered asteroids (circles) and the PLS asteroids with quality 1 or 2 orbits (plus signs) with the representative plane. These boundaries mark quite well the boundaries of the 3/1 Kirkwood gap.

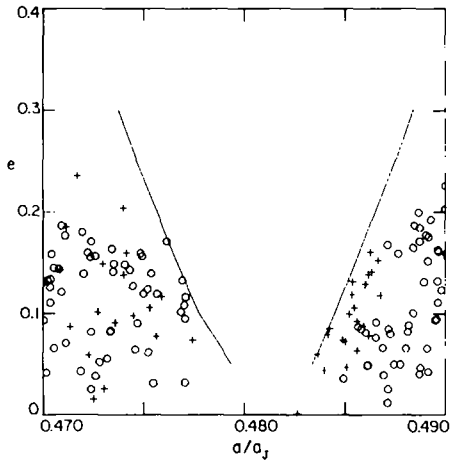

FIG. 19. The outer boundaries of the chaotic zone on the plane used in Fig. 8 are displayed with the intersections of the numbered asteroids (circles) and the PLS asteroids (plus signs). The precise agreement between the boundaries of the chaotic zone and the boundaries of the Kirkwood gap is even more apparent on this plane.

differential equations. The comparison with the mapping result was good; the error was smaller than the symbols which denote the positions. Figure 18 also shows the outer boundary of the chaotic zone as determined in Figs. 6 and 9, i.e., for zero inclination (planar problem) and 10 degree inclination. Simple linear interpolations of the boundaries have been made. The boundaries of the chaotic zone match the boundaries of the real 3/1 Kirkwood gap quite well.

Three real asteroids have been omitted from this plot. They are 887 Alinda, 1915 Quetzalcoatl, and PL 4917. 887 Alinda and 1915 Quetzalcoatl are the two 3/1 librators and have eccentricities which are too large to appear on the plot. (They are both Mars crossers.) The PL asteroid is not plotted for a different reason. When followed with the three-dimensional mapping this asteroid has a chaotic trajectory. The fit to the observational data for PL 4917 has particularly high residuals, with a formal error in the semimajor axis of 0.0246 AU (personal communication, Marsden, 1982). This error is not only large enough to bring the asteroid into the neighboring quasiperiodic region (+0.002 AU would have been sufficient), but it is large enough to take the asteroid across the Kirkwood gap to the opposite quasiperiodic region! The orbit of PL 4917 is too poorly known to be used in a study of the Kirkwood gaps. In comparing these errors with the figures remember that the abscissa is always a/a_J, not a measured in astronomical units. In general, the error expected of the PL asteroids with quality 1 and 2 orbits is only 0.003 AU in the semimajor axis (van Houten *et al.*, 1970).

Figure 19 is identical to Fig. 18, but with a different intersection plane, the plane used in Fig. 8. In this case, however, the chaotic zone has only been traced out in the planar problem. The correspondence of the outer boundary of the chaotic zone and the boundary of the actual Kirkwood gap is even more dramatic on this plane. This plane has the disadvantage, though, that the six asteroids numbered 189, 292, 619, 799, 1722, and 2273 do not cross it.

In Wisdom (1982) the predicted gap was too narrow when compared to the distribution of the real numbered asteroids. This discrepancy was caused by a combination of two factors. First, the initial conditions of test asteroids were removed when their eccentricity went above 0.3 in a 2-my interval. Over longer times more asteroids might be removed. This difficulty has now been solved, since it is possible to predict the long term behavior once the character is known. If a trajectory is quasiperiodic, its range of eccentricity variation may be determined after a fairly short interval. This range is fixed forever. If a trajectory is chaotic it will eventually reach large eccentricities, though this may not happen within the first 2 my. Thus those test asteroids which were chaotic but did not become Mars crossing should have been removed along with the Mars crossers. The second factor is more important and concerns the statistics. If too few objects are studied the actual gap may appear artificially large. This difficulty can be overcome by studying a

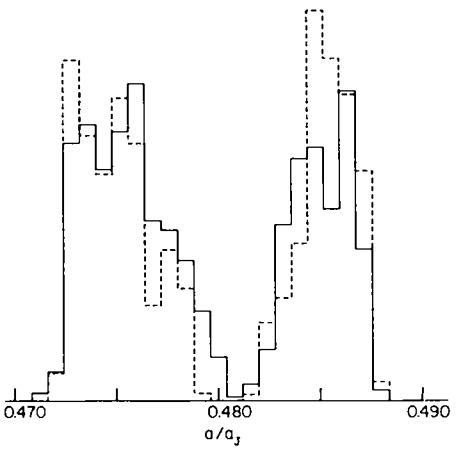

FIG. 20. Time-averaged histograms of the actual asteroids (dashed line) and of those test asteroids which were neither chaotic nor Mars crossing. The histograms are normalized to represent the same number of asteroids. Again, the predicted gap agrees well with the actual gap in the asteroid belt.

larger sample, namely, the PL asteroids with quality 1 and 2 orbits as well as the numbered asteroids.

The three-dimensional distribution calculation in Wisdom (1982) was repeated. As before 300 test asteroids were followed. The semimajor axes, eccentricities, and inclinations were the same as in Wisdom (1982); the longitudes were given new random values. This time each test asteroid was followed for only a million years, but the maximum LCE was computed to determine the character. Of the 300 test asteroids in the random distribution, 89 were found to have chaotic trajectories and only 11 were quasiperiodic librators. All but five of the chaotic trajectories became Mars crossing within 300,000 years and only one had not reached an eccentricity of 0.3 within 1 my. All 11 quasiperiodic librators could cross the orbit of Mars. Figure 20 shows a time-averaged histogram of the semimajor axes of those test asteroids which were neither Mars crossing nor chaotic (solid line). The semimajor axes were sampled every 12 years over a time interval of 6000 years. The same time average of the real numbered asteroids with the PL asteroids is also shown (dashed line). The histograms are normalized to have the same areas. The predicted gap is now in satisfactory agreement with the full distribution of real asteroids.

The removal of Mars crossers and chaotic trajectories is not only sufficient to explain the width of the 3/1 Kirkwood gap, but explains the detailed distribution of asteroids in phase space as well.

IX. SUMMARY

Calculations of the maximum Lyapunov Characteristic Exponent have shown that the mappings for asteroidal motion near the 3/1 commensurability presented in Wisdom (1982) accurately reflect whether trajectories are chaotic or quasiperiodic, as determined by numerical integrations of the corresponding differential equations. The sudden large increases in eccentricity which were first seen with a mapping have now been seen in numerical integrations of the differential equations. This peculiar behavior is thus not an artifact of the mapping technique, but a phenomenon which real asteroids near the 3/1 commensurability might exhibit. Similar phenomena occur in other dynamical systems, the behavior is not as unusual as it seemed at first. A systematic exploration of the character of trajectories has revealed three basic classes of trajectories: the chaotic trajectories, the quasiperiodic trajectories inside the chaotic zone, and the quasiperiodic trajectories outside the resonance region defined by the chaotic zone. Chaotic behavior is common, contrary to the conclusion of Froeschlé and Scholl (1976). In the planar-elliptic problem there is a quasiintegral which constrains the possible motions, but it does not prevent asteroids from becoming Mars crossing. In the three-dimensional problem trajectories exhibit greater freedom and reach very large eccentricities, probably large enough to become Earth crossers. Collision probabilities with the Earth and Mars appear adequate to remove the asteroids with chaotic

trajectories within the age of the solar system. Most quasiperiodic librators have large enough variations in eccentricity to be regular Mars crossers. (Fig. 8 also shows that some quasiperiodic trajectories outside the resonance region can become Mars crossers, viz. those with eccentricity greater than or equal to 0.25. The secular variations are primarily responsible for making these trajectories become Mars crossing.) In order to be perfectly safe asteroids must lie in the quasiperiodic region outside the chaotic zone. Figures 18 and 19 show that the outer boundary of the chaotic region does, in fact, correspond precisely to the boundary of the 3/1 Kirkwood gap within the experimental and theoretical uncertainties. The agreement is dramatic. Figure 20 confirms this agreement for the time averaged distribution of semimajor axes. Confronted with the evidence of Figs. 18–20 it is impossible to discount the importance of chaotic behavior in the formation of the gap. No extra hypotheses are needed beyond the dynamics of the three-dimensional elliptic-restricted three-body problem and the presence of Mars to obtain the 3/1 Kirkwood gap's precise size and shape.

APPENDIX

The three-dimensional mapping was not given explicitly in Wisdom (1982). For completeness and because the mappings used in this paper have one minor difference from those used before, the mappings are given here. In Wisdom (1982) the Hamiltonian was expanded in Φ about the resonant value Φ_R and only the quadratic term was kept. This was not a bad approximation, but since it was not necessary it is not made here.

Units are chosen such that $a_J = 1$ and $l_J = t$, the time. Thus the period of Jupiter is 2π. The mass of Jupiter divided by the mass of the Sun plus that of Jupiter is μ, and $\mu_1 = 1 - \mu$. The perihelion of Jupiter is taken as the origin of longitudes, and inclinations are measured relative to Jupiter's orbit plane. The mappings are expressed in terms of the variables:

$$x_1 = (\mu_1 a)^{1/4}[2(1 - (1 - e^2)^{1/2})]^{1/2} \cos \tilde{\omega} \approx e \cos \tilde{\omega},$$

$$y_1 = -(\mu_1 a)^{1/4}[2(1 - (1 - e^2)^{1/2})]^{1/2} \sin \tilde{\omega} \approx -e \sin \tilde{\omega},$$

$$x_2 = [\mu_1 a(1 - e^2)]^{1/4}[2(1 - \cos i)]^{1/2} \cos \Omega \approx i \cos \Omega,$$

$$y_2 = -[\mu_1 a(1 - e^2)]^{1/4}[2(1 - \cos i)]^{1/2} \sin \Omega \approx -i \sin \Omega,$$

$$\Phi = (\mu_1 a)^{1/2},$$

$$\varphi = l - 3l_J.$$

Initially, the elements are $x_i^{(0)}$, $y_i^{(0)}$, $\Phi^{(0)}$, and $\varphi^{(0)}$. There are four steps in the three-dimensional mapping:

Step 1.

$$x_1^{(1)} = x_1^{(0)} \cosh(4\pi C_1 \cos \varphi^{(0)}) - y_1^{(0)} \sinh(4\pi C_1 \cos \varphi^{(0)}),$$

$$y_1^{(1)} = y_1^{(0)} \cosh(4\pi C_1 \cos \varphi^{(0)}) - x_1^{(0)} \sinh(4\pi C_1 \cos \varphi^{(0)}) - 2\pi D_1 e_J \cos \varphi^{(0)},$$

$$x_2^{(1)} = x_2^{(0)} \cosh(4\pi C_2 \cos \varphi^{(0)}) - y_2^{(0)} \sinh(4\pi C_2 \cos \varphi^{(0)}),$$

$$y_2^{(1)} = y_2^{(0)} \cosh(4\pi C_2 \cos \varphi^{(0)}) - x_2^{(0)} \sinh(4\pi C_2 \cos \varphi^{(0)}),$$

$$\Phi^{(1)} = \Phi^{(0)} - (C_1[(x_1^{(0)})^2 - (y_1^{(0)})^2] + D_1 e_J x_1^{(1)} + E_1 e_J^2 + C_2[(x_2^{(0)})^2 - (y_2^{(0)})^2])2\pi \sin \varphi^{(0)},$$

$$\varphi^{(1)} = \varphi^{(0)}.$$

Step 2.

$$x_1^{(2)} = x_1^{(1)} \cos \pi F_1 - y_1^{(1)} \sin \pi F_1 - \frac{F_1 e_J}{2F_1}(1 - \cos \pi F_1),$$

$$y_1^{(2)} = x_1^{(1)} \sin \pi F_1 + y_1^{(1)} \cos \pi F_1 + \frac{F_1 e_J}{2F_1} \sin \pi F_1,$$

$$x_2^{(2)} = x_2^{(1)} \cos \pi F_2 - y_2^{(1)} \sin \pi F_2,$$

$$y_2^{(2)} = x_2^{(1)} \sin \pi F_2 + y_2^{(1)} \cos \pi F_2,$$

$$\Phi^{(2)} = \Phi^{(1)},$$

$$\varphi^{(2)} = \varphi^{(1)} + \frac{\pi\mu_1^2}{2(\Phi^{(1)})^3} - \frac{3\pi}{2}.$$

Step 3.

$$x_1^{(3)} = x_1^{(2)} \exp(4\pi C_1 \sin \varphi^{(2)}) + 2\pi D_1 e_J \sin \varphi^{(2)},$$

$$y_1^{(3)} = y_1^{(2)} \exp(-4\pi C_1 \sin \varphi^{(2)}),$$

$$x_2^{(3)} = x_2^{(2)} \exp(4\pi C_2 \sin \varphi^{(2)}),$$

$$y_2^{(3)} = y_2^{(2)} \exp(-4\pi C_2 \sin \varphi^{(2)}),$$

$$\Phi^{(3)} = \Phi^{(2)} - (2C_1 x_1^{(2)} y_1^{(2)} + D_1 e_J y_1^{(3)} + 2C_2 x_2^{(2)} y_2^{(2)}) 2\pi \cos \varphi^{(2)},$$

$$\varphi^{(3)} = \varphi^{(2)}.$$

Step 4.

$$x_1^{(4)} = x_1^{(3)} \cos 3\pi F_1 - y_1^{(3)} \sin 3\pi F_1 - \frac{\bar{F}_1}{2F_1}(1 - \cos 3\pi F_1),$$

$$y_1^{(4)} = x_1^{(3)} \sin 3\pi F_1 + y_1^{(3)} \cos 3\pi F_1 + \frac{\bar{F}_1}{2F_1} \sin 3\pi F_1,$$

$$x_2^{(4)} = x_2^{(3)} \cos 3\pi F_2 - y_2^{(3)} \sin 3\pi F_2,$$

$$y_2^{(4)} = x_2^{(3)} \sin 3\pi F_2 + y_2^{(3)} \cos 3\pi F_2,$$

$$\Phi^{(4)} = \Phi^{(3)},$$

$$\varphi^{(4)} = \varphi^{(3)} + \frac{3\pi\mu_1^2}{2(\Phi^{(3)})^3} - \frac{9\pi}{2}.$$

The constants are $F_1 = -0.2050694\mu$, $\bar{F}_1 = 0.1987054\mu$, $F_2 = 0.2050694\mu$, $C_1 = 0.8631579\mu$, $D_1 = -2.656407\mu$, $E_1 = 0.3629536\mu$, and $C_2 = 0.1193545\mu$. These four steps constitute one iteration of the three-dimensional mapping, and give the elements in terms of the elements one Jupiter period earlier. The planar-elliptic mapping is obtained by setting all variables with the subscript 2 equal to zero.

ACKNOWLEDGMENTS

It is a pleasure to thank Stanton J. Peale for frequent discussions and his support of this project. Thanks are also due to Brian Marsden for estimating the formal errors in several asteroid orbits. Most of the numerical calculations were performed at the Royal Greenwich Observatory. This was made possible through the generosity of D. B. Taylor. Some preliminary calculations were carried out at the Centre d'Etudes et de Recherches Géodynamiques et Astronomiques (France) through the kindness of François Mignard and at the Observatoire de Nice through Claude Froeschlé. Jörg Waldvogel suggested the use of Scheibnerian coordinates to accelerate the numerical integrations. J. Schubart provided references to the known periodic orbits. I am grateful to Carl Murray for detecting an error in an earlier version of this paper. This work was supported in part by NASA Grant NGR 05 010 062 and was begun while the author was supported under the NATO Postdoctoral Fellowship Program administered by the United States National Science Foundation.

REFERENCES

BENETTIN, G., M. CASARTELLI, L. GALGANI, A. GIORGILLI, AND J.-M. STRELCYN (1978). On the reliability of numerical studies of stochasticity. *Nuovo Cimento* **44,** 183–195.

BENETTIN, G., L. GALGANI, A. GIORGILLI, AND J.-M. STRELCYN (1980a). Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory. *Meccanica* March, 9–20.

BENETTIN, G., L. GALGANI, A. GIORGILLI, AND J.-M. STRELCYN (1980b). Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 2: Numerical application. *Meccanica* March, 21–30.

BIEN, R. (1980). Stationary solutions in simplified resonance cases of the restricted three-body problem. *Celest. Mech.* **21,** 157–161.

BULIRSCH, R., AND J. STOER (1966). Numerical treatment of ordinary differential equations by extrapolation methods. *Numerische Math.* **8,** 1.

CHANNON, S. R., AND J. L. LEBOWITZ (1980). Numerical experiments in stochasticity and homoclinic oscillation. In *Nonlinear Dynamics* (R. Helleman, Ed.), pp. 108–118. New York Academy of Sciences, New York.

CHIRIKOV, B. V. (1979). A universal instability of many-dimensional oscillator systems. *Phys. Rep.* **52,** 263–379.

FROESCHLÉ, C., AND H. SCHOLL (1976). On the dynamical topology of the Kirkwood gaps. *Astron. Astrophys.* **48,** 389–393.

FROESCHLÉ, C., AND H. SCHOLL (1981). The stochasticity of peculiar orbits in the 2/1 Kirkwood gap. *Astron. Astrophys.* **93,** 62–66.

GIFFEN, R. (1973). A study of commensurable motion in the asteroid belt. *Astron. Astrophys.* **23,** 387–403.

HÉNON, M., AND C. HEILES (1964). The applicability of the third integral of motion: some numerical experiments. *Astron. J.* **69,** 73–79.

HILL, G. W. (1902). Illustrations of periodic solutions in the problem of three bodies. *Astron. J.* **22,** 117.

KIRKWOOD, D. (1867). *Meteoric Astronomy*. Lippincott, Philadelphia.

POINCARÉ, H. (1902). Sur les planetes du type d'Hecube. *Bull. Astron.* **19,** 289–310.

SCHEIBNER, W. (1866). Satz aus der störungstheorie. *J. Reine Angew. Math.* **65,** 291–301.

SCHOLL, H., AND C. FROESCHLÉ (1974). Asteroidal motion at the 3/1 commensurability. *Astron. Astrophys.* **33,** 455–458.

SCHUBART, J. (1964). Long-period effects in nearly commensurable cases of the restricted three-body problem. In *Smithsonian Astrophys. Obs. Spec. Rep. No. 149.*

SINCLAIR, A. T. (1970). Periodic solutions close to commensurabilities in the three-body problem. *Mon. Not. Roy. Astron. Soc.* **148,** 325–351.

VAN HOUTEN, C. J., I. VAN HOUTEN-GROENEVELD, P. HERGET, AND T. GEHRELS (1970). The Palomar–Leiden survey of faint minor planets. *Astron. Astrophys. Suppl.* **2,** 339–448.

WETHERILL, G. W. (1975). Late heavy bombardment of the moon and terrestrial planets. *Proc. Lunar Sci. Conf. 6th,* 1539–1561.

WISDOM, J. (1982). The origin of the Kirkwood gaps: A mapping for asteroidal motion near the 3/1 commensurability. *Astron. J.* **87,** 577–593.