1. THE ORIGIN OF THE KIRKWOOD GAPS: A MAPPING FOR ASTEROIDAL MOTION NEAR THE 3/1 COMMENSURABILITY

2. THE RESONANCE OVERLAP CRITERION AND THE ONSET OF STOCHASTIC BEHAVIOR IN THE RESTRICTED THREE-BODY PROBLEM

Thesis by

Jack Leach Wisdom

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1981

(Submitted on May 15, 1981)

To my parents

PREFACE

This thesis consists of two separate investigations which deal with nonlinear dynamical problems in the solar system. The first paper develops a new method for studying the long time evolution of asteroids near commensurabilities with Jupiter and presents a new hypothesis for the origin of the Kirkwood gaps. This study has been submitted in the present form for publication in the *Astronomical Journal*. The second paper is concerned with the long term stability of a test body moving in the gravitational field of two masses, one much larger than the other. In particular, it asks how close a test body in direct motion about the primary can come to the smaller secondary and still be stable. This study was published under the title "The Resonance Overlap Criterion and the Onset of Stochastic Behavior in the Restricted Three-Body Problem" in the *Astronomical Journal* **23**, 1122 (1980).

ABSTRACT

PAPER 1:

A mapping of the phase space onto itself with the same low order resonance structure as the 3/1 commensurability in the planar elliptic three-body problem is derived. This mapping is approximately one thousand (1000) times faster than the usual method of numerically integrating the averaged equations of motion (as used by Schubart, Froeschlé and Scholl in their studies of the asteroid belt). This mapping exhibits some very surprising behavior that might provide the key to the origin of the gaps. A test asteroid placed in the gap may evolve for a million years with low eccentricity (< 0.05) and then suddenly jump to large eccentricity (> 0.3) becoming a Mars crosser. The asteroid can then be removed by a close encounter with Mars. To test this hypothesis a distribution of 300 test asteroids in the neighborhood of the 3/1 commensurability was evolved for two million years. When the Mars crossers are removed the distribution of initial conditions displays a gap at the location of the 3/1 Kirkwood gap. While this is the first real demonstration of the formation of a gap, the gap is too narrow. The planar elliptic mapping is then extended to include the inclinations and the secular perturbations of Jupiter's orbit. The two million year evolution of the 300 test asteroids is repeated using the full mapping. The resulting gap is somewhat larger yet still too small. Finally the possibility that over longer times more asteroids will become Mars crossers is tested by studying the evolution of one test asteroid near the border of the gap for a much longer time. A jump in its eccentricity occurs after 18 million years indicating that indeed it may simply be a matter of time for the full width of the gap to open.

PAPER 2:

The resonance overlap criterion for the onset of stochastic behavior is applied to the planar circular-restricted three-body problem with small mass ratio (μ). Its predictions for $\mu = 10^{-3}$, $\mu = 10^{-4}$ and $\mu = 10^{-5}$ are compared to the transitions in the numerically determined Kolmogorov-Sinai entropy and found to be in remarkably good agreement. In addition, an approximate scaling law for the onset of stochastic behavior is derived.

TABLE OF CONTENTS

PAPER 1

THE ORIGIN OF THE KIRKWOOD GAPS: A MAPPING FOR ASTEROIDAL MOTION NEAR THE 3/1 COMMENSURABILITY

by Jack Wisdom

Submitted to the *Astronomical Journal*

I. INTRODUCTION

Over one hundred years have passed since the discovery of the Kirkwood gaps in the distribution of semimajor axes of the asteroids, yet there is still no adequate theory of their origin. Greenberg and Scholl (1979) give a review of the competing hypotheses and some of their difficulties. Briefly, there are four classes: 1) the gaps are only a statistical phenomenon, 2) the gaps are formed by purely gravitational forces, 3) the gaps form because asteroids near resonances tend to have larger eccentricities and hence larger probability of being removed by a collision with another asteroid away from the gaps, and 4) that no asteroids were ever in the gaps.

The statistical hypothesis is that asteroids near commensurabilities undergo large variations in their semimajor axes spending most of their time away from the commensurabilities, just as a pendulum spends most of its time away from the bottom of its swing. The time averaged distribution could then display a gap. Schweizer (1969) calculated the orbits of numbered asteroids near the gaps and found that most asteroids did not cross the gaps. Wiesel (1976) studied the statistical hypothesis theoretically within the averaged planar-circular restricted three-body problem. He obtained depressions in the distribution but nothing resembling the observed distribution. According to Wiesel the most serious limitation in his theory is the truncation of the disturbing function to the terms of lowest order in eccentricity. I believe that a more serious limitation is the neglect of Jupiter's eccentricity, which introduces qualitative changes in the types of motion possible (see Froeschlé and Scholl 1977), especially the possibility of "ergodic" trajectories (Giffen 1973, and Froeschlé and Scholl 1976). Although it is not the main thrust of this paper I will show below that the statistical hypothesis fails under much more general assumptions than considered

previously. I include the eccentricity of Jupiter, the inclinations, as well as the secular perturbations of Jupiter's orbit.

The gravitational hypothesis is that under the influence of gravitational forces alone, the semimajor axes of asteroids drift away from commensurabilities. The gravitational hypothesis has been studied analytically by Schubart (1964) in the averaged planar-circular restricted three-body problem. It is not too surprising that he finds that asteroids do not leave the gaps since in this approximation the problem is completely integrable. Thus all trajectories are quasiperiodic. Schubart's assumptions are clearly too restrictive, yet to lift any of the assumptions is to make the theory analytically intractable. The long time evolution of dynamical systems is in general a very difficult and unsolved problem. The Kolmogorov- Arnol'd-Moser (KAM) theorem (see e.g. Moser 1973) in certain special cases proves the existence of quasiperiodic trajectories. Though a tremendous advance in our exact understanding of dynamical systems, the KAM theorem is not a practical theorem in the sense that it applies only for sufficiently small perturbations where "sufficiently" is not precisely defined. The numerical experiments of Hénon and Heiles (1964) (and many others now) revealed a division of the phase space of simple Hamiltonian systems into quasiperiodic and "ergodic" regions for a range of perturbation strengths. The resonance overlap criterion (see Chirikov 1979, for a general review, and Wisdom 1980, for an application to the restricted three-body problem) provides some insight as to which regions of phase space will be "ergodic", but for detailed results there is no other recourse than numerical studies. Asteroidal motion at commensurabilities has been studied numerically by Scholl and Froeschlé (1974,1975) who integrated the averaged equations of motion for the planar elliptic problem as described by Schubart (1964,1968). None of their fictitious

asteroids were found to leave the gaps though some did alternate between circulation and libration. One possible objection to this work is that the integrations were continued for only 100,000 years which is very short compared to the age of the solar system. Is it possible that over much longer times asteroids do drift out of the gaps? Arnol'd has discovered that dynamical systems may exhibit a very slow diffusion through phase space (see Chirikov 1979). However, Arnol'd diffusion does not occur in autonomous systems with only two degrees of freedom and since the averaged planar elliptic problem has only two degrees of freedom it has no slow diffusion. If more degrees of freedom are added the possibility of diffusion must be considered. In particular, diffusion is possible in the unaveraged problem, i.e. the problem with high frequency terms, and in the three dimensional elliptic problem.

The validity of the collisional hypothesis was the primary topic addressed in the papers of Scholl and Froeschlé (1974,1975). Using the averaging procedure for the planar elliptic problem they studied a large number of fictitious asteroids with initial conditions near the gaps. Their integrations generally covered a time interval of less than or near 50,000 years, though a few were continued for 100,000 years. The important question is whether or not the orbits show large eccentricity variations. Scholl and Froeschlé found that test asteroids close to the 3/1, 5/2 and 2/1 commensurability with either small initial eccentricity or moderate eccentricity near the borders of the observed gaps did not develop large eccentricities. No test asteroids at the 7/3 resonance developed large eccentricity. Since the Kirkwood gaps are clear of asteroids of all eccentricities these results seem to militate against the collision hypothesis. The collision hypothesis is also doubtful because the collision probability depends weakly on the eccentricity, whereas the gaps have relatively sharp

boundaries. Furthermore an asteroid with low eccentricity near the 2/1 commensurability has a higher probability of collision than a high eccentricity asteroid because of the sharp drop in the number of asteroids beyond 3.2AU (Ip 1979).

Largely because of the failure of these studies to explain the Kirkwood gaps the possibility that asteroids were never in the gaps has started to be explored (e.g. Heppenheimer 1978). Since the dynamical hypotheses have not yet been fully disproved it seems premature to abandon them for the more speculative theories of origin. I will not discuss these further.

In this paper I introduce a significant new method for studying asteroidal dynamics near commensurabilities in the planar elliptic problem, namely, I derive algebraic mappings of the phase space onto itself that have the same low order resonance structure as the 3/1 commensurability. This means that the averaged equations of motion for the mappings are identical to the averaged equations of motion used by Froeschlé and Scholl (except that the mappings retain only the lowest order terms in the disturbing function). I choose to study the 3/1 Kirkwood gap because it is the largest of the gaps, except for the 2/1 gap, and is not confused with the boundaries of the asteroid belt as is the 2/1 gap. There are no fundamental difficulties in deriving mappings for the other commensurabilities. Mappings have two principal computational advantages over Schubart's method of numerically integrating the averaged equations of motion: 1) mappings are about one thousand (1000) times faster, the integrations that lasted forty minutes for Froeschlé and Scholl are reproduced in only a couple of seconds by a mapping, and 2) they are more accurate since they are purely algebraic and thus have the full accuracy of the computer. An important theoretical advantage is that the mappings have high frequency contributions,

which are not present in the averaged equations of motion. The presence of high frequency perturbations introduces some qualitative changes in the dynamics such as the "stochastic" separatrix and the possibility of the Arnol'd diffusion (see Chirikov 1979).

In Section II, I derive a Hamiltonian that approximates motion near the 3/1 commensurability. In the next section I derive the mapping for the planar elliptic problem and present some surprising results. In Section IV, I study a distribution of 300 test asteroids for two million years. I extend the mapping in Section V to include the secular perturbations of Jupiter's orbit and the inclinations. In section VI, I reexamine with the full mapping the evolution of the distribution of 300 test asteroids for two million years. My conclusions are stated in Section VII.

II. THE RESONANT HAMILTONIAN

In terms of the Poincaré canonical elements (see, e.g., Plummer 1960), the Hamiltonian for a zero mass test body moving in the field of a large central mass (the Sun) and perturbed by a smaller mass (Jupiter) whose orbit lies outside that of the test body is

$$H = -\frac{\mu_1^2}{2L^2} - \mu R(L,\ \rho_1,\ \rho_2,\ \lambda,\ \omega_1,\ \omega_2), \tag{1}$$

where R is the disturbing function, $\mu_1 = 1 - \mu$ and μ is the mass of the secondary. I have chosen units so that the product of the gravitational constant and the sum of the masses is unity and the separation of the two masses is also unity. Following Froeschlé and Scholl I take $\mu = 1/1047.355$. The Poincaré momenta may be written in terms of the usual osculating elliptic elements: $L \equiv \sqrt{\mu_1 a}$, $\rho_1 \equiv \sqrt{\mu_1 a}\,(1 - (1-e^2)^{1/2}) \approx \sqrt{\mu_1 a}\,\frac{e^2}{2}$ and $\rho_2 \equiv (\mu_1 a(1-e^2))^{1/2}(1-\cos i) \approx \sqrt{\mu_1 a}\,\frac{i^2}{2}$, where a is the semimajor axis, e is the eccentricity, and i is the inclination to the invariable plane. The conjugate coordinates are the mean longitude λ, minus the longitude of periapse ω_1, and minus the longitude of the ascending node on the invariable plane ω_2, respectively. The disturbing function may be written as the sum

$$R = \sum_{ijklmn} K^{ijklmn}(L,\ \rho_1,\ \rho_2)\cos(i\lambda + j\omega_1 + k\omega_2 + l\lambda' + m\omega_1' + n\omega_2') \tag{2}$$

which is constrained by the requirement that $i - j - k + l - m - n$ be zero and for definiteness i is restricted to be greater than or equal to zero. Jupiter's elements carry a prime. A resonance occurs when one of the cosine arguments is nearly stationary. Since the mean longitudes move much faster than the other angles this means that $i\dot{\lambda} + l\dot{\lambda}' \approx 0$, or $\dot{\lambda} \approx -\frac{l}{i} \equiv \frac{i+q}{i}$. The integer q is called

the order of the resonance; resonance strengths decrease with increasing order since for small e and i their terms in R are proportional to $e^a i^b$ where $a+b = q$. The most important 3/1 resonance terms have $i = 1$ and $q = 2$ and are quadratic in the eccentricities and inclinations. The $i = 2$ resonant terms are smaller by factors quadratic in the eccentricities and inclinations. In this paper all resonant terms with $i \geq 2$ are neglected. The terms whose arguments do not involve the mean longitudes, i.e. which have $i = l = 0$, are called secular terms and also affect the motion. In order to construct an algebraic mapping it is necessary to neglect all secular terms that are fourth order or greater in the eccentricities and inclinations. Besides the higher order 3/1 resonances there are no neighboring resonances in the range $0.47 < a < 0.49$ with order less than $q = 21$. The largest of these terms is proportional to the twenty first power of the eccentricity. It is thus a very good approximation to ignore all nearby resonances. The arguments of all other terms involve non-resonant combinations of the mean longitudes, and rotate at least as fast as the mean longitude of Jupiter. These non-resonant terms will be called high frequency terms. It is the high frequency terms which are removed by Schubart's numerical averaging procedure.

The explicit form of the disturbing function may be found in, for example, Leverrier(1855), Peirce(1849), or Brouwer and Clemence(1961). In Leverrier's notation

$$R_{secular} = (2)^{(0)}\left(\frac{e}{2}\right)^2 + (11)^{(0)}\eta_L^2 + (21)^{(-1)}\left(\frac{e}{2}\right)\left(\frac{e'}{2}\right)\cos(\omega_L - \varpi'). \qquad (3)$$

The primed quantities belong to Jupiter, $\eta_L^2 \equiv \sin^2\frac{J}{2}$, where J is the mutual inclination between the orbit of Jupiter and the orbit of the test body, ϖ is the longitude of the periapse and $\omega_L \equiv \varpi + \Omega' + \vartheta' - \Omega - \vartheta$ where ϑ and ϑ' are the angles

between the ascending nodes of the orbits on the invariable plane to the ascending node of the outer orbit on the inner orbit and Ω and Ω' are the longitudes of the ascending nodes on the invariable plane. The coefficients are defined as follows:

$$(2)^{(0)} \equiv a\frac{d}{da}b_{1/2}^{(0)}(a) + \frac{1}{2}a^2\frac{d^2}{da^2}b_{1/2}^{(0)}(a), \tag{4}$$

$$(11)^{(0)} \equiv \frac{1}{2}b_{3/2}^{(1)}(a) \tag{5}$$

and

$$(21)^{(-1)} \equiv -12b_{1/2}^{(2)}(a) - 2a\frac{d}{da}b_{1/2}^{(2)}(a) - a^2\frac{d^2}{da^2}b_{1/2}^{(2)}(a) \tag{6}$$

The Laplace coefficients, $b_j^{(i)}$, are defined by

$$\frac{1}{2}\sum b_j^{(i)}(a)\cos i(l_L' - \lambda_L) \equiv \frac{a^{j-1/2}}{(1-2a\cos(l_L' - \lambda_L)+a^2)^j}. \tag{7}$$

Similarly,

$$\begin{aligned} R_{resonant} = {} & (172)^{(3)}(\frac{e}{2})^2\cos(3l_L' - \lambda_L - 2\omega_L) \\ & + (182)^{(2)}(\frac{e}{2})(\frac{e'}{2})\cos(3l_L' - \lambda_L - \omega_L - \varpi') \\ & + (192)^{(1)}(\frac{e'}{2})^2\cos(3l_L' - \lambda_L - 2\varpi') \\ & + (212)^{(3)}\eta_L^2\cos(3l_L' - \lambda_L - 2\tau_L'), \end{aligned} \tag{8}$$

where l_L is Leverrier's mean longitude, $\lambda_L = l + \varpi + \Omega' + \vartheta' - \Omega - \vartheta$, and $\tau_L = \Omega + \vartheta$. The coefficients are:

$$(172)^{(3)} \equiv \frac{21}{2}b_{1/2}^{(3)}(a) + 5a\frac{d}{da}b_{1/2}^{(3)}(a) + \frac{1}{2}a^2\frac{d^2}{da^2}b_{1/2}^{(3)}(a), \tag{9}$$

$$(182)^{(2)} \equiv -20b_{1/2}^{(2)}(a) - 10a\frac{d}{da}b_{1/2}^{(2)}(a) - a^2\frac{d^2}{da^2}b_{1/2}^{(2)}(a), \tag{10}$$

$$(192)^{(1)} \equiv \frac{17}{2}b_{1/2}^{(1)}(a) + 5a\frac{d}{da}b_{1/2}^{(1)}(a) + \frac{1}{2}a^2\frac{d^2}{da^2}b_{1/2}^{(1)}(a) - \frac{27}{2}a \tag{11}$$

and

$$(212)^{(3)} = \frac{1}{2}b_{3/2}^{(2)}(a). \tag{12}$$

The coefficient $(192)^{(1)}$ has been corrected to include the indirect contribution (see Leverrier 1855). The coefficients vary little over the very small range of semimajor axes of the 3/1 Kirkwood gap so I simply evaluate them at exact resonance. In all terms not involving the mutual inclination, each $\Omega + \vartheta$ is cancelled by an $\Omega' + \vartheta'$ since they differ only by terms of order J^2. The terms involving the mutual inclination, J, must be written in terms of the individual inclinations and nodes. To order J^2 (see Plummer 1960)

$$\eta_L^2 \equiv \sin^2\frac{J}{2} = (\frac{i}{2})^2 + (\frac{i'}{2})^2 - 2(\frac{i}{2})(\frac{i'}{2})\cos(\Omega' - \Omega) \tag{13}$$

and

$$\eta^2\cos(3l_L' - \lambda_L - 2\tau_L) = (\frac{i}{2})^2\cos(3(l' + \varpi') - (l + \varpi) - 2\Omega)$$

$$- 2(\frac{i}{2})(\frac{i'}{2})\cos(3(l' + \varpi') - (l + \varpi) - (\Omega + \Omega'))$$

$$+ (\frac{i'}{2})^2\cos(3(l' + \varpi') - (l + \varpi) - 2\Omega'). \tag{14}$$

The disturbing function can now be written in terms of the canonical Poincaré elements:

$$\mu R_{secular} = -F_1 2\rho_1 - e'\overline{F}_1\sqrt{2\rho_1}\cos(\omega_1 + \varpi') - F_2 2\rho_2 - i'\overline{F}_2\sqrt{2\rho_2}\cos(\omega_2 + \Omega') \tag{15}$$

and

$$\mu R_{resonant} = C_1 2\rho_1 \cos(3t - \lambda + 2\omega_1) + e' D_1 \sqrt{2\rho_1}\cos(3t - \lambda + \omega_1 - \varpi')$$

$$+ e'^2 E_1 \cos(3t - \lambda - 2\varpi') + C_2 2\rho_2 \cos(3t - \lambda + 2\omega_2)$$

$$+ i' D_2 \sqrt{2\rho_2}\cos(3t - \lambda + \omega_2 - \Omega') + i'^2 E_2 \cos(3t - \lambda - 2\Omega'), \tag{16}$$

where the coefficients are

$$F_1 \equiv -\frac{\mu(2)^{(0)}}{4(\mu_1 a_{1/3})^{1/2}} \approx -0.2050694\mu, \tag{17}$$

$$\overline{F_1} \equiv -\frac{\mu(21)^{(-2)}}{4(\mu_1 a_{1/3})^{1/4}} \approx 0.1987054\mu, \tag{18}$$

$$F_2 \equiv -\frac{\mu(11)^{(0)}}{4(\mu_1 a_{1/3})^{1/2}} \approx 0.4266973\mu, \tag{19}$$

$$\overline{F_2} \equiv \frac{\mu(11)^{(0)}}{2(\mu_1 a_{1/3})^{1/4}} \approx -0.7103812\mu, \tag{20}$$

$$C_1 \equiv \frac{\mu(172)^{(3)}}{4(\mu_1 a_{1/3})^{1/2}} \approx 0.8631579\mu, \tag{21}$$

$$D_1 \equiv \frac{\mu(182)^{(2)}}{4(\mu_1 a_{1/3})^{1/4}} \approx -2.656407\mu, \tag{22}$$

$$E_1 \equiv \frac{\mu(192)^{(1)}}{4} \approx 0.3629536\mu, \tag{23}$$

$$C_2 \equiv \frac{\mu(212)^{(3)}}{4(\mu_1 a_{1/3})^{1/2}} \approx 0.2483461\mu, \tag{24}$$

$$D_2 \equiv -\frac{\mu(212)^{(3)}}{2(\mu_1 a_{1/3})^{1/4}} \approx -0.4134556\mu, \tag{25}$$

$$E_2 \equiv \frac{\mu(212)^{(3)}}{4} \approx 0.1720839\mu. \tag{26}$$

The resonant Hamiltonian is of the form

$$H_R = -\frac{\mu_1^2}{2L^2} - \mu R_{secular}(\omega_1, \rho_1, \omega_2, \rho_2) - \mu R_{resonant}(\omega_1, \rho_1, \omega_2, \rho_2, 3t - \lambda). \tag{27}$$

In order to remove the explicit time dependence from Eq.(27)[1] I perform a canonical transformation to the coordinate $\varphi = 3t - \lambda$ via the time dependent generating function $F = (3t - \lambda)(\Phi + L_R) + \omega_1\rho_1' + \omega_2\rho_2'$ where for convenience I have also translated the origin of momentum to be the exact resonance. $L_R \equiv \sqrt{\mu_1 a_{1/3}}$ is defined so that $\dot{\mu} = \frac{\partial H_R'}{\partial \Phi}$ is zero at $\Phi = 0$. This gives $L_R = (\frac{\mu_1^2}{3})^{1/3}$. Keeping only the quadratic term in Φ the resonant Hamiltonian is now

$$H_R = \frac{1}{2}\alpha\Phi^2 - \mu R_{secular}(\omega_1, \rho_1, \omega_2, \rho_2) - \mu R_{resonant}(\omega_1, \rho_1, \omega_2, \rho_2, \varphi), \quad (28)$$

where

$$\alpha = \left.\frac{\partial^2 H_R}{\partial \Phi^2}\right|_{\Phi=0} = -\frac{3\mu_1^2}{L_R^4} \approx -12.98851 \quad (29)$$

and I have dropped the primes.

Finally the eccentricity and inclination of Jupiter must be specified. Brouwer and van Woerkom (1950) provide a solution to the secular problem of planetary motion. The elements of Jupiter's orbit are given in the form of sums

$$\sum_k c_k \cos(-s_k t + \delta_k) = e'\cos\tilde{\omega}' \equiv \xi_1 \quad (30)$$

$$\sum_k c_k \sin(-s_k t + \delta_k) = e'\sin\tilde{\omega}' \equiv \eta_1 \quad (31)$$

$$\sum_k d_k \cos(-s_k' t + \delta_k') = i'\cos\Omega' \equiv \xi_2 \quad (32)$$

$$\sum_k d_k \sin(-s_k' t + \delta_k') = i'\sin\Omega' \equiv \eta_2. \quad (33)$$

---

1. N.B. The resonant Hamiltonian still contains the explicit time dependence of Jupiter's elements.

Two of the c_k dominate all the others, and I include only those terms. The values are: $c_1 = 0.0448188$, $c_2 = 0.0153546$, $s_1 = -3.97769\times10^{-5}$, $s_2 = -2.57167\times10^{-4}$. It is necessary to include three of the d_k: $d_1 = 0.0063064$, $d_2 = 0.0009571$, $d_3 = 0.0011689$, $s_1' = 2.382736\times10^{-4}$, $s_2' = 2.687651\times10^{-5}$, and $s_3' = 6.273352\times10^{-6}$.

In terms of the canonical momenta $x_i \equiv \sqrt{2\rho_i}\cos\omega_i$ and their canonical coordinates $y_i \equiv \sqrt{2\rho_i}\sin\omega_i$, the final form of the resonant Hamiltonian is

$$
\begin{aligned}
H_R = {} & \frac{1}{2}\alpha\Phi^2 + \mu F_1(x_1^2 + y_1^2) + \mu\overline{F}_1(x_1\xi_1 - y_1\eta_1) \\
& + \mu F_2(x_2^2 + y_2^2) + \mu\overline{F}_2(x_2\xi_2 - y_2\eta_2) \\
& - \mu[C_1(x_1^2 - y_1^2) + D_1(x_1\xi_1 + y_1\eta_1) + E_1(\xi_1^2 - \eta_1^2)]\cos\varphi \\
& - \mu[C_1 2x_1y_1 - D_1(x_1\eta_1 - y_1\xi_1) - E_1 2\xi_1\eta_1]\sin\varphi \\
& - \mu[C_2(x_2^2 - y_2^2) + D_2(x_2\xi_2 + y_2\eta_2) + E_2(\xi_2^2 - \eta_2^2)]\cos\varphi \\
& - \mu[C_2 2x_2y_2 - D_2(x_2\eta_2 - y_2\xi_2) - E_2 2\xi_2\eta_2]\sin\varphi \qquad (34)
\end{aligned}
$$

III. PLANAR ELLIPTIC MAPPING

The resonant Hamiltonian for the planar elliptic problem can be obtained from the full resonant Hamiltonian, Eq. (34), by setting the coefficients of all terms involving the inclinations to zero and fixing Jupiter's orbit by setting $\xi_1 = e_J$ and $\eta_1 = 0$. e_J is the constant value of Jupiter's eccentricity. The resulting resonant Hamiltonian is

$$H'_R = \frac{1}{2}\alpha\Phi^2 + F_1(x_1^2 + y_1^2) + \overline{F}_1 e_J x_1$$

$$- [C_1(x_1^2 - y_1^2) + D_1 e_J x_1 + E_1 e_J^2]\cos\varphi$$

$$- [C_1 2x_1 y_1 + D_1 y_1 e_J]\sin\varphi. \qquad (35)$$

In this form the derivation of a mapping is rather easy. To H'_R I add new high frequency terms,

$$H' = \frac{1}{2}\alpha\Phi^2 + F_1(x_1^2 + y_1^2) + \overline{F}_1 e_J x_1$$

$$- \sum_{i=-\infty}^{\infty} C_1(x_1^2 - y_1^2)\cos(\varphi - i\Omega(t - \zeta_1))$$

$$- \sum_{i=-\infty}^{\infty} D_1 e_J x_1 \cos(\varphi - i\Omega(t - \zeta_2))$$

$$- \sum_{i=-\infty}^{\infty} E_1 e_J^2 \cos(\varphi - i\Omega(t - \zeta_3))$$

$$+ \sum_{i=-\infty}^{\infty} C_1 2x_1 y_1 \cos(\varphi - \frac{\pi}{2} - i\Omega(t - \gamma_1))$$

$$+ \sum_{i=-\infty}^{\infty} D_1 e_J x_1 \cos(\varphi - \frac{\pi}{2} - i\Omega\,(t - \gamma_2)) \tag{36}$$

The frequency Ω is the same as Jupiter's mean motion when $\Omega = 1$, but for the time being it will be left as a parameter. These high frequency terms are of the same form as the real high frequency terms that were originally neglected, but are chosen so that each sum becomes a sum of delta functions. Consider just the first sum:

$$\sum_{i=-\infty}^{\infty} \cos(\varphi - i\Omega\,(t - \zeta_1)) = \cos\varphi \sum_{i=-\infty}^{\infty} \cos(i\Omega\,(t - \zeta_1))$$

$$= \cos\varphi \sum_{i=-\infty}^{\infty} 2\pi\delta(\Omega\,(t - \zeta_1) - 2\pi i)$$

$$\equiv \cos\varphi 2\pi\delta_{2\pi}(\Omega\,(t - \zeta_1)). \tag{37}$$

Thus the planar elliptic mapping Hamiltonian becomes

$$H'_M = \frac{1}{2}\alpha\Phi^2 + F_1({x_1}^2 + {y_1}^2) + \overline{F}_1 e_J x_1$$

$$- C_1({x_1}^2 - {y_1}^2)\cos\varphi 2\pi\delta_{2\pi}(\Omega\,(t - \zeta_1))$$

$$- D_1 e_J x_1 \cos\varphi 2\pi\delta_{2\pi}(\Omega\,(t - \zeta_2))$$

$$- E_1 {e_J}^2 \cos\varphi 2\pi\delta_{2\pi}(\Omega\,(t - \zeta_3))$$

$$- C_1 2x_1 y_1 \sin\varphi 2\pi\delta_{2\pi}(\Omega\,(t - \gamma_1))$$

$$- D_1 e_J y_1 \sin\varphi 2\pi\delta_{2\pi}(\Omega\,(t - \gamma_2)) \tag{38}$$

The constants ζ_i and γ_i determine the times at which the delta functions act and are arbitrary. I choose $0 \leq \zeta_1 < \zeta_2 < \zeta_3$, $\frac{\pi}{2\Omega} \leq \gamma_1 < \gamma_2$ and let $\zeta_3 \to 0$ and $\gamma_2 \to \frac{\pi}{2\Omega}$. The delta functions thus all act at either $t = 0$ or $t = \frac{\pi}{2\Omega}$, but they have a definite ordering. This ordering is a matter of choice and is certainly not the only one possible nor even necessarily the best. Between the delta functions the motion is determined entirely by the secular terms

$$H_{secular} = \frac{1}{2}\alpha\Phi^2 + F_1(x_1^2 + y_1^2) + \overline{F}_1 e_J x_1. \tag{39}$$

This Hamiltonian has the trivial solution:

$$x_1 = x_1(t_0)\cos(2F_1(t-t_0)) - y_1(t_0)\sin(2F_1(t-t_0)) - \frac{\overline{F}}{2F}[1 - \cos(2F(t-t_0))] \tag{40}$$

$$y_1 = x_1(t_0)\sin(2F_1(t-t_0)) + y_1(t_0)\cos(2F_1(t-t_0)) + \frac{\overline{F}}{2F}\sin(2F(t-t_0)), \tag{41}$$

$$\Phi = \Phi(t_0) \tag{42}$$

and

$$\varphi = \alpha\Phi_0(t - t_0) + \varphi_0, \tag{43}$$

To integrate across the delta functions I use a limiting procedure, the delta functions are replaced by

$$\delta(t - t_i) \leftarrow \lim_{\Delta\to 0} \begin{cases} \frac{1}{\Delta} & t_i \leq t < t_i+\Delta \\ 0 & \text{otherwise} \end{cases}. \tag{44}$$

Each delta function is integrated in turn. The Hamiltonian during the first delta function is

$$H_1 = -\frac{2\pi C_1}{\Delta\Omega}(x_1^2 - y_1^2)\cos\varphi, \tag{45}$$

which leads to the equations of motion:

$$\dot{x_1} = -\frac{\partial H}{\partial y_1} = -\frac{4\pi C_1}{\Delta\Omega}y_1\cos\varphi \equiv -\frac{R_1}{\Delta}y_1, \tag{46}$$

$$\dot{y_1} = \frac{\partial H}{\partial x_1} = -\frac{4\pi C_1}{\Delta\Omega}x_1\cos\varphi \equiv -\frac{R_1}{\Delta}x_1, \tag{47}$$

$$\dot{\Phi} = -\frac{2\pi C_1}{\Delta\Omega}(x_1^2 - y_1^2)\sin\varphi \tag{48}$$

and

$$\dot{\varphi} = 0. \tag{49}$$

The solutions to the first two of these equations are

$$x_1(t) = x_1(t_1)\cosh\frac{R_1(t - t_1)}{\Delta} - y_1(t_1)\sinh\frac{R_1(t - t_1)}{\Delta} \tag{50}$$

and

$$y_1(t) = y_1(t_1)\cosh\frac{R_1(t - t_1)}{\Delta} - x_1(t_1)\sinh\frac{R_1(t - t_1)}{\Delta}. \tag{51}$$

The changes in x_1 and y_1 across the first delta function are then calculated by setting $t = t_1 + \Delta$ (and letting Δ go to zero). They are

$$\Delta x_1 = x_1(t_1)[\cosh(\frac{4\pi C_1}{\Omega}\cos\varphi) - 1] - y_1(t_1)\sinh(\frac{4\pi C_1}{\Omega}\cos\varphi) \tag{52}$$

and

$$\Delta y_1 = y_1(t_1)[\cosh(\frac{4\pi C_1}{\Omega}\cos\varphi) - 1] - x_1(t_1)\sinh(\frac{4\pi C_1}{\Omega}\cos\varphi). \tag{53}$$

Having solved for $x_1(t)$ and $y_1(t)$, the equation for $\dot{\Phi}$ becomes

$$\dot{\Phi} = -\frac{2\pi C_1}{\Delta\Omega}(x_1^2(t) - y_1^2(t))\sin\varphi(t_1)$$

$$= -\frac{2\pi C_1}{\Delta\Omega}(x_1^2(t_1) - y_1^2(t_1))\sin\varphi(t_1), \tag{54}$$

which yields for the change in Φ across the delta function

$$\Delta\Phi = -\frac{2\pi C_1}{\Omega}(x_1^2(t_1) - y_1^2(t_1))\sin\varphi(t_1). \tag{55}$$

The Hamiltonian during the second delta function is

$$H_2 = -\frac{2\pi D_1}{\Delta\Omega}e_J x_1 \cos\varphi(t_2) \tag{56}$$

The equations of motion are

$$\dot{x_1} = 0, \tag{57}$$

$$\dot{y_1} = -\frac{2\pi D_1}{\Delta\Omega}e_J \cos\varphi(t_2), \tag{58}$$

$$\dot{\Phi} = -\frac{2\pi D_1}{\Delta\Omega}e_J x_1 \sin\varphi(t_2) \tag{59}$$

and

$$\dot{\varphi} = 0. \tag{60}$$

The changes in y_1 and Φ across the second delta function are

$$\Delta y_1 = -\frac{2\pi D_1}{\Omega}e_J \cos\varphi(t_2) \tag{61}$$

$$\Delta\Phi = -\frac{2\pi D_1}{\Omega}e_J x_1(t_2)\sin\varphi(t_2). \tag{62}$$

Only Φ changes across the third delta function. The change is

$$\Delta\Phi = -\frac{2\pi E_1}{\Omega}e_J^2 \sin\varphi(t_3). \tag{63}$$

During the fourth delta function the Hamiltonian is

$$H_4 = -\frac{4\pi C_1}{\Delta\Omega}x_1 y_1 \sin\varphi(t_4). \tag{64}$$

The equations of motion are

$$\dot{x_1} = \frac{4\pi C_1}{\Delta\Omega} x_1 \sin\varphi(t_4), \tag{65}$$

$$\dot{y_1} = -\frac{4\pi C_1}{\Delta\Omega} y_1 \sin\varphi(t_4) \tag{66}$$

and

$$\dot{\Phi} = \frac{4\pi C_1}{\Delta\Omega} x_1 y_1 \cos\varphi(t_4). \tag{67}$$

Their solutions are

$$x_1(t) = x_1(t_4)\exp[\frac{4\pi C_1 \sin\varphi(t_4)}{\Delta\Omega}(t - t_4)], \tag{68}$$

$$y_1(t) = y_1(t_4)\exp[-\frac{4\pi C_1 \sin\varphi(t_4)}{\Delta\Omega}(t - t_4)] \tag{69}$$

and

$$\Phi(t) = \Phi(t_4) + (t - t_4)\frac{4\pi C_1}{\Delta\Omega}\cos\varphi(t_4). \tag{70}$$

The changes across the fourth delta function are then

$$\Delta x_1 = x_1(t_4)[\exp(\frac{4\pi C_1 \sin\varphi(t_4)}{\Omega}) - 1], \tag{71}$$

$$\Delta y_1 = y_1(t_4)[\exp(-\frac{4\pi C_1 \sin\varphi(t_4)}{\Omega}) - 1] \tag{72}$$

and

$$\Delta\Phi = \frac{4\pi C_1}{\Omega} x_1(t_4) y_1(t_4)\cos\varphi(t_4). \tag{73}$$

The changes across the fifth delta function are evaluated in exactly the same manner. They are

$$\Delta x_1 = \frac{2\pi D_1}{\Omega} e_J \sin\varphi(t_5) \tag{74}$$

$$\Delta\Phi = \frac{2\pi D_1}{\Omega} e_J y_1(t_5)\cos\varphi(t_5). \tag{75}$$

The planar elliptic mapping is now complete. For clarity I will review one complete cycle of the mapping, starting at $t = 0$ and ending at $t = \frac{2\pi}{\Omega}$. Initially the elements are $x_1^{(0)}$, $y_1^{(0)}$, $\Phi^{(0)}$ and $\varphi^{(0)}$. There are seven steps:

(1)
$$x_1^{(1)} = x_1^{(0)}\cosh(\frac{4\pi C_1}{\Omega}\cos\varphi^{(0)}) - y_1^{(0)}\sinh(\frac{4\pi C_1}{\Omega}\cos\varphi^{(0)}) \tag{76}$$

$$y_1^{(1)} = y_1^{(0)}\cosh(\frac{4\pi C_1}{\Omega}\cos\varphi^{(0)}) - x_1^{(0)}\sinh(\frac{4\pi C_1}{\Omega}\cos\varphi^{(0)}) \tag{77}$$

$$\Phi^{(1)} = \Phi^{(0)} - \frac{2\pi C_1}{\Omega}((x_1^{(0)})^2 - (y_1^{(0)})^2)\sin\varphi^{(0)} \tag{78}$$

$$\varphi^{(1)} = \varphi^{(0)}, \tag{79}$$

(2)
$$x_1^{(2)} = x_1^{(1)} \tag{80}$$

$$y_1^{(2)} = y_1^{(1)} - \frac{2\pi D_1}{\Omega} e_J \cos\varphi^{(1)} \tag{81}$$

$$\Phi^{(2)} = \Phi^{(1)} - \frac{2\pi D_1}{\Omega} e_J x_1^{(1)}\sin\varphi^{(1)} \tag{82}$$

$$\varphi^{(2)} = \varphi^{(1)}, \tag{83}$$

(3)
$$x_1^{(3)} = x_1^{(2)} \tag{84}$$

$$y_1^{(3)} = y_1^{(2)} \tag{85}$$

$$\Phi^{(3)} = \Phi^{(2)} - \frac{2\pi E_1}{\Omega} e_J{}^2\sin\varphi^{(2)} \tag{86}$$

$$\varphi^{(3)} = \varphi^{(2)}, \tag{87}$$

(4)
$$x_1^{(4)} = x_1^{(3)}\cos\frac{\pi F}{\Omega} - y_1^{(3)}\sin\frac{\pi F}{\Omega} - \frac{\overline{F}}{2F}(1-\cos\frac{\pi F}{\Omega}) \tag{88}$$

$$y_1^{(4)} = x_1^{(3)}\sin\frac{\pi F}{\Omega} + y_1^{(3)}\cos\frac{\pi F}{\Omega} + \frac{\overline{F}}{2F}\sin\frac{\pi F}{\Omega} \tag{89}$$

$$\Phi^{(4)} = \Phi^{(3)} \tag{90}$$

$$\varphi^{(4)} = \varphi^{(3)} + \frac{\pi\alpha}{2\Omega}\Phi^{(3)}, \tag{91}$$

(5)
$$x_1^{(5)} = x_1^{(4)}\exp(\frac{4\pi C_1}{\Omega}\sin\varphi^{(4)}) \tag{92}$$

$$y_1^{(5)} = y_1^{(4)}\exp(-\frac{4\pi C_1}{\Omega}\sin\varphi^{(4)}) \tag{93}$$

$$\Phi^{(5)} = \Phi^{(4)} + \frac{4\pi C_1}{\Omega}x_1^{(4)}y_1^{(4)}\cos\varphi^{(4)} \tag{94}$$

$$\varphi^{(5)} = \varphi^{(4)}, \tag{95}$$

(6)
$$x_1^{(6)} = x_1^{(5)} + \frac{2\pi D_1}{\Omega}e_J\sin\varphi^{(5)} \tag{96}$$

$$y_1^{(6)} = y_1^{(5)} \tag{97}$$

$$\Phi^{(6)} = \Phi^{(5)} + \frac{2\pi D_1}{\Omega}e_J y_1^{(5)}\cos\varphi^{(5)} \tag{98}$$

$$\varphi^{(6)} = \varphi^{(5)}, \tag{99}$$

and

(7)
$$x_1^{(7)} = x_1^{(6)}\cos\frac{3\pi F}{\Omega} - y_1^{(6)}\sin\frac{3\pi F}{\Omega} - \frac{\overline{F}}{2F}[1-\cos\frac{3\pi F}{\Omega}] \tag{100}$$

$$y_1^{(7)} = x_1^{(6)}\sin\frac{3\pi F}{\Omega} + y_1^{(6)}\cos\frac{3\pi F}{\Omega} + \frac{\overline{F}}{2F}\sin\frac{3\pi F}{\Omega}. \tag{101}$$

$$\Phi^{(7)} = \Phi^{(6)} \tag{102}$$

$$\varphi^{(7)} = \varphi^{(6)} + \frac{3\pi\alpha}{2\Omega}\Phi^{(6)}. \tag{103}$$

These seven steps constitute one iteration of the mapping. In the next section I present the results of some calculations using this planar elliptic mapping.

IV. PLANAR ELLIPTIC MAPPING CALCULATIONS

Primarily as a test of the mapping I redid the work of Scholl and Froeschlé (1974). They study the range of variation of the orbital elements for a large number of initial conditions near the 3/1 resonance and the mapping gave variations in the eccentricities and semimajor axes which were in most cases very close to their results. The differences arise for several reasons. First, Scholl and Froeschlé do not specify exactly the initial conditions used nor the length of each integration. Some of their orbits were integrated for 12,000 years while others were integrated for 50,000 years. There were no qualitative differences in the variations over the two time intervals so the times were not specified. A more fundamental reason is that nearby orbits in the "ergodic" region separate at an exponential rate, so any small initial error quickly manifests itself. Because of the presence of the high frequency terms there is an inherent ambiguity about what starting values for the mapping correspond to the initial conditions of the averaged differential equations. In the "ergodic" region this uncertainty grows exponentially. An exact comparison of the two can be expected only in the quasiperiodic region of phase space, i.e. either for small eccentricities (< 0.1) near the resonance or outside the resonance region. Finally, when the variation in the eccentricity is large (> 0.3) the mapping can be only qualitatively correct because of the truncation of the disturbing function to the second order terms in the eccentricities.

In order to aid my comparison of the mapping and the averaged differential equations, Scholl and Froeschlé kindly provided the details of several orbits. Figure 1 shows a plot of the eccentricity versus time for one of these orbits, with $\Phi(t=0) = 0$, $\varphi(t=0) = \pi$, $y_1(t=0) = 0$ and $e(t=0) = 0.05$. For this calculation e_J was chosen to be 0.048. Even though this orbit is started at exact resonance it

looks remarkably regular. From the appearance of a surface of section one would conclude that this orbit is quasiperiodic. This is an example of the low eccentricity orbits that Greenberg and Scholl describe as behaving "non-resonantly". The integration step was three years with a relative accuracy per integration step of 10^{-12}. The calculation took 40 minutes on an IBM 360. Next comes the real surprise. Figure 2 shows the results of the mapping (with $\Omega = 1/2$) for the same initial conditions. For the first 100,000 years the mapping reproduces quite closely the very regular motion of Figure 1. The eccentricity then noticeably becomes more irregular and after nearly 240,000 years a huge spike occurs. This mapping "integration" was calculated in double precision (eight 8-bit bytes) on the Caltech IBM 370/3032. The calculation took only a couple of seconds. Naturally one wonders if this jump could somehow be due to an accumulation of roundoff error. In order to check this the calculation was repeated in quadruple precision (sixteen 8-bit bytes). Of course the orbits differed slightly but the eccentricity plots were indistinguishable. The jump still occurs, and at exactly the same time. What causes the jump? Is the jump an artifact of my special choice of high frequency terms or would a similar jump occur in the real unaveraged problem? The underlying philosophy of this work is that the presence of high frequencies is important but their exact form is not. There are several cases where this has been shown to be true. For example, in the presence of a high frequency perturbation the separatrix of a pendulum broadens into a narrow stochastic layer whose width depends exponentially on the frequency of the perturbation, but only linearly on its strength. In order to help isolate the effect of the high frequencies in the planar elliptic mapping I studied the evolution of this same orbit with several different mapping frequencies Ω. The times to first jump, T_{jump}, are shown in Figure 3. As the mapping frequency gets higher the time to first jump gets longer. On a logarithmic scale

the relationship between T_{jump} and Ω appears almost linear. One possible explanation of this is that the trajectories undergo a slow diffusion until they get to a point where jumps occur. If the diffusion rate depends exponentially on the high frequency, i.e.

$$D \propto \exp(-\alpha\Omega) \tag{104}$$

and the distance in phase space that must be covered in order to get the jump is Δx, then the time to first jump would obey the simple diffusion law

$$\Delta x = D T_{jump}{}^2 \propto e^{-\alpha\Omega} T_{jump}{}^2, \tag{105}$$

which gives

$$T_{jump} \propto (\Delta x)^{\frac{1}{2}} e^{\frac{\alpha\Omega}{2}}. \tag{106}$$

The line in Figure 3 is a linear least squares fit of the above expression (with $\alpha = 6.27$). It thus seems possible that diffusion is present, but it is certainly not proved. It is also interesting to see how this behavior is affected by varying Jupiter's eccentricity. The principal effect is that the height of the spike changes. Figure 4 shows the results. Finally, I verified that qualitatively the same results are obtained for several different choices of the arbitrary phases ξ_i and γ_i in Eq. (38).

The fact that this orbit undergoes such an extremely large increase in eccentricity introduces a new possibility for the origin of the 3/1 Kirkwood gap. Taking into account the secular variations of Mars' orbit, an asteroid at the 3/1 commensurability need only have an eccentricity of 0.3 to be a Mars crosser. According to Wetherill(1975) the typical half-life for a Mars crosser is only 200 million years. The half-life for removal from the Kirkwood gap is less than this since a smaller perturbation or sum of perturbations is required. Thus if large eccentricity increases are typical for orbits near the 3/1 commensurability, it is

possible that the gap is cleared by close encounters with Mars. To test this hypothesis, I "integrated" with the mapping a sample distribution of 300 test asteroids with initial conditions in the neighborhood of the 3/1 commensurability for two million years each. In order to get a distribution that would simulate the real asteroid distribution without a gap, I took the elements of the first 300 asteroids in the TRIAD file with $e < 0.3$ and with semimajor axes between 0.49 and 0.52 and shifted their semimajor axes so that they ranged from 0.4725 to 0.4875. The resulting distribution of eccentricities and semimajor axes is shown in Figure 5. Since the mapping frequency does not seem to affect the motion except to change the rate of diffusion I used the relatively small mapping frequency of $\Omega = 1/4$ to speed up any diffusion. I also set Jupiter's eccentricity to its maximum, $e_J = 0.0601734$ (see Plummer 1960). Figure 6 shows the initial conditions of those test asteroids that did not reach an eccentricity of 0.3 in two million years, i.e. it is the same as Figure 5 but with the Mars crossers removed. It is clear that a definite gap has been formed. Note that no secular drift of semimajor axes could be detected. Figures 5 and 6 are snapshots in time of the distribution. To see what these distributions look like when averaged over time I again "integrated" each of the test asteroids for 5,000 years and sampled the elements at every mapping period (48 years). The results are histogrammed in Figures 7 and 8. The histograms of Figure 7 show the relative probabilities of the various semimajor axes and eccentricities using all the initial conditions of Figure 5. This histogram of semimajor axes completely demolishes any remaining hopes that the statistical hypothesis is correct. There is only a faint hint of a dip in the distribution at the commensurability. Figure 8 shows the histograms for the initial conditions of Figure 6. Even after time averaging there is a definite gap in the semimajor axes.

Although this is the first real demonstration of the formation of a gap, the gap is too narrow. The actual distribution of asteroid elements is shown in Figure 9. The real Kirkwood gap is approximately twice as large. There seem to be two possible explanations for this discrepancy. Either the "integrations" were not continued for a long enough time (after all two million years is still short compared to the age of the solar system) or some important physics has been left out. There are two possibilities for the latter, the inclinations and the secular variations of Jupiter's orbit. In the next section, I will extend the derivation of the mapping to include both the inclinations and the secular perturbations. That is, I will derive a mapping for the full resonance Hamiltonian (34).

V. DERIVATION OF THE COMPLETE MAPPING

The mapping Hamiltonian for the complete resonant Hamiltonian, Eq. (34), is derived in exactly the same manner as was the planar elliptic mapping Hamiltonian. It is

$$H_M = \frac{1}{2}\alpha\Phi^2 + F_1(x_1^2 + y_1^2) + \overline{F_1}(x_1\xi_1 - y_1\eta_1)$$

$$+ F_2(x_2^2 + y_2^2) + \overline{F_2}(x_2\xi_2 - y_2\eta_2)$$

$$- C_1(x_1^2 - y_1^2)2\pi\delta_{2\pi}(\Omega\,(t - \zeta_1))\cos\varphi$$

$$- D_1(x_1\xi_1 + y_1\eta_1)2\pi\delta_{2\pi}(\Omega\,(t - \zeta_2))\cos\varphi$$

$$- E_1(\xi_1^2 - \eta_1^2)2\pi\delta_{2\pi}(\Omega\,(t - \zeta_3))\cos\varphi$$

$$- C_2(x_2^2 - y_2^2)2\pi\delta_{2\pi}(\Omega\,(t - \zeta_4))\cos\varphi$$

$$- D_2(x_2\xi_2 + y_2\eta_2)2\pi\delta_{2\pi}(\Omega\,(t - \zeta_5))\cos\varphi$$

$$- E_2(\xi_2^2 - \eta_2^2)2\pi\delta_{2\pi}(\Omega\,(t - \zeta_6))\cos\varphi$$

$$- C_1 2x_1y_1 2\pi\delta_{2\pi}(\Omega\,(t - \gamma_1))\sin\varphi$$

$$- D_1(x_1\eta_1 - y_1\xi_1)2\pi\delta_{2\pi}(\Omega\,(t - \gamma_2))\sin\varphi$$

$$+ E_1 2\xi_1\eta_1 2\pi\delta_{2\pi}(\Omega\,(t-\gamma_3))\sin\varphi$$

$$- C_2 2x_2y_2 2\pi\delta_{2\pi}(\Omega\,(t-\gamma_4))\sin\varphi$$

$$- D_2(x_2\eta_2 - y_2\xi_2)2\pi\delta_{2\pi}(\Omega\,(t-\gamma_5))\sin\varphi$$

$$+ E_2 2\xi_2\eta_2 2\pi\delta_{2\pi}(\Omega\,(t-\gamma_6))\sin\varphi. \tag{107}$$

Again, the ζ_i and γ_i are arbitrary. I choose $0 \leqq \zeta_1 < \zeta_2 < \zeta_3 < \zeta_4 < \zeta_5 < \zeta_6$ and $\frac{\pi}{2\Omega} \leqq \gamma_1 < \gamma_2 < \gamma_3 < \gamma_4 < \gamma_5 < \gamma_6$, and let $\zeta_6 \to 0$ and $\gamma_6 \to \frac{\pi}{2\Omega}$. The secular motion between delta functions is

$$x_1(t) = x_1(t_0)\cos(2F_1(t-t_0)) - y_1(t_0)\sin(2F_1(t-t_0))$$

$$- \overline{F}_1\sum_k \frac{d_k}{2F_1 - s_k}[\cos(-s_k t + \delta_k) - \cos(-s_k t_0 + \delta_k - 2F_1(t-t_0))] \tag{108}$$

$$y_1(t) = x_1(t_0)\sin(2F_1(t-t_0)) + y_1(t_0)\cos(2F_1(t-t_0))$$

$$+ \overline{F}_1\sum_k \frac{d_k}{2F_1 - s_k}[\sin(-s_k t + \delta_k) - \sin(-s_k t_0 + \delta_k - 2F_1(t-t_0))] \tag{109}$$

$$\Phi(t) = \Phi(t_0) \tag{110}$$

$$\varphi(t) = \varphi(t_0) + (t-t_0)\alpha\Phi(t_0) \tag{111}$$

The solutions for $x_2(t)$ and $y_2(t)$ are similar to those for $x_1(t)$ and $y_1(t)$. The Hamiltonians during the first and fourth delta functions are of the form

$$H_{1,4} = -\frac{2\pi C_i}{\Delta\Omega}(x_i^2 - y_i^2)\cos\varphi, \tag{112}$$

where $i = 1$ for H_1 and $i = 2$ for H_4. This Hamiltonian is the same as the Hamil-

tonian during the first delta function of the planar elliptic mapping. Thus the changes in the elements are

$$\Delta x_i = x_i[\cosh(\frac{4\pi C_i}{\Omega}\cos\varphi) - 1] - y_i\sinh(\frac{4\pi C_i}{\Omega}\cos\varphi) \tag{113}$$

$$\Delta y_i = y_i[\cosh(\frac{4\pi C_i}{\Omega}\cos\varphi) - 1] - x_i\sinh(\frac{4\pi C_i}{\Omega}\cos\varphi) \tag{114}$$

$$\Delta\Phi = -\frac{2\pi C_i}{\Omega}(x_i^2 - y_i^2)\sin\varphi \tag{115}$$

and

$$\Delta\varphi = 0. \tag{116}$$

During the second and fifth delta functions the Hamiltonians have the form

$$H_{2,5} = -\frac{2\pi D_i}{\Delta\Omega}(x_i\xi_i + y_i\eta_i)\cos\varphi. \tag{117}$$

Again, $i = 1$ for H_2 and $i = 2$ for H_5. Consider just the $i = 1$ case, i.e. the second delta function. The equations of motion are

$$\dot{x_1} = \frac{2\pi D_1}{\Delta\Omega}\eta_1\cos\varphi, \tag{118}$$

$$\dot{y_1} = -\frac{2\pi D_1}{\Delta\Omega}\xi_1\cos\varphi, \tag{119}$$

$$\dot{\Phi} = -\frac{2\pi D_1}{\Delta\Omega}(x_1\xi_1 + y_1\eta_1)\sin\varphi, \tag{120}$$

and

$$\dot{\varphi} = 0. \tag{121}$$

Let t_0 be the time at which the delta function acts. Expanding ξ_1 and η_1 about $t = t_0$

$$\xi_1(t_0 + \Delta t) = \xi_1(t_0) - \Delta t\sum c_k(-s_k)\sin(-s_k t_0 + \zeta_k) + o(\Delta t^2) \tag{122}$$

and

$$\eta_1(t_0 + \Delta t) = \eta_1(t_0) + \Delta t \sum c_k(-s_k)\cos(-s_k t_0 + \zeta_k) + o(\Delta t^2). \tag{123}$$

Substituting these into the equations of motion gives

$$\dot{x_1} = \frac{2\pi D_1}{\Delta\Omega}[\eta_1(t_0) + \Delta t \sum c_k(-s_k)\cos(-s_k t_0 + \zeta_k) + o(\Delta t^2)] \tag{124}$$

and

$$\dot{y_1} = -\frac{2\pi D_1}{\Delta\Omega}[\xi_1(t_0) - \Delta t \sum c_k(-s_k)\sin(-s_k t_0 + \zeta_k) + o(\Delta t^2)]. \tag{125}$$

These are trivially integrated to yield

$$x_1(t_0 + \Delta t) = x_1(t_0) + \Delta t \frac{2\pi D_1}{\Delta\Omega}\eta_1(t_0)\cos\varphi(t_0) + o(\frac{\Delta t^2}{\Delta}) \tag{126}$$

and

$$y_1(t_0 + \Delta t) = y_1(t_0) - \Delta t \frac{2\pi D_1}{\Delta\Omega}\xi_1(t_0)\cos\varphi(t_0) + o(\frac{\Delta t^2}{\Delta}). \tag{127}$$

Setting $\Delta t = \Delta$ and letting $\Delta \to 0$ gives the changes in x_1 and y_1:

$$\Delta x_1 = \frac{2\pi D_1}{\Omega}\eta_1\cos\varphi \tag{128}$$

$$\Delta y_1 = -\frac{2\pi D_1}{\Omega}\xi_1\cos\varphi. \tag{129}$$

Substituting Eqs. (122), (123), (126) and (127) into the equation of motion for Φ gives

$$\dot{\Phi} = -\frac{2\pi D_1}{\Delta\Omega}[x_1(t_0)\xi_1(t_0) + y_1(t_0)\eta_1(t_0)$$

$$- \Delta t \sum c_k(-s_k)(x_1(t_0)\sin(-s_k t_0 + \zeta_k) - y_1(t_0)\cos(-s_k t_0 + \zeta_k))$$

$$+ o(\Delta t^2)]\sin\varphi(t_0). \tag{130}$$

Integrating, setting $\Delta t = 0$, and letting $\Delta \to 0$ yields

$$\Delta\Phi = -\frac{2\pi D_1}{\Omega}[x_1\xi_1 + y_1\eta_1]\sin\varphi. \tag{131}$$

Since the third and sixth delta functions do not involve x_i, y_i or Φ they can be integrated without a limiting procedure. The resulting changes in Φ are

$$\Delta\Phi = -\frac{2\pi E_i}{\Omega}(\xi_i^2 - \eta_i^2)\sin\varphi, \tag{132}$$

for $i = 1$ and $i = 2$ respectively. The derivation of the changes across the other six delta functions are completely analogous. I give only the results. Across the seventh and tenth delta functions

$$\Delta x_i = x_i[\exp(\frac{4\pi C_i}{\Omega}\sin\varphi) - 1], \tag{133}$$

$$\Delta y_i = y_i[\exp(-\frac{4\pi C_i}{\Omega}\sin\varphi) - 1] \tag{134}$$

and

$$\Delta\Phi = \frac{4\pi C_i}{\Delta\Omega}x_i y_i\cos\varphi. \tag{135}$$

The changes across the eighth and eleventh delta functions are

$$\Delta x_i = -\frac{2\pi D_i}{\Omega}\xi_i\sin\varphi, \tag{136}$$

$$\Delta y_i = -\frac{2\pi D_i}{\Omega}\eta_i\sin\varphi \tag{137}$$

and

$$\Delta\Phi = -\frac{2\pi D_i}{\Omega}(x_i\eta_i - y_i\xi_i)\cos\varphi. \tag{138}$$

Finally, across the ninth and twelfth deltas

$$\Delta\Phi = -\frac{4\pi E_i}{\Delta}\xi_i\eta_i\cos\varphi \tag{139}$$

In the next section I display the results of some calculations using this mapping.

VI. FULL MAPPING CALCULATIONS

Using the full mapping ($\Omega = 1/4$) the same distribution of 300 test asteroids was "integrated" for two million years. The inclinations and nodes were taken from the TRIAD file as were the eccentricities before. So as not to prejudice the distribution with one particular set of the starting values for the phases of the secular terms (see Eqs. (30) through (33)), these phases were randomly chosen for each test asteroid. An asteroid is considered to be a Mars crosser if its eccentricity ever exceeds 0.3, thus the possibility of correlations among e, i, ϖ and Ω that prevent close encounters with Mars are ignored. The resulting distribution of asteroids that did not become Mars crossers is shown in Figure 10. The time averaged distributions are in Figures 11 and 12. The gap is only slightly larger than before (cf. Figures 6 and 8). The inclinations and secular terms do not seem to make much difference, at least in this two million year time interval. It remains a possibility that over much longer times more asteroids will show jumps in orbital eccentricity. The test asteroid that is circled in Figure 8 was found to have a spike in eccentricity after 18 million years. This result is encouraging and indicates that indeed the full width of the gap might be reproduced if integrations of much longer time could be carried out.

VII. SUMMARY AND CONCLUSIONS

The mappings in this paper are models for the motion of asteroids near the 3/1 commensurability. They have the correct secular and long period resonant terms, but possess false high frequency terms which are similar to high frequency perturbations by Jupiter which have been neglected in previous work.

Mappings are very useful for several reasons. First, they are a very fast (inexpensive) method for studying qualitatively what types of motion are possible near resonances. Second, within their approximations the mappings are more accurate since they are purely algebraic and hence use the full accuracy of the computer. "Integrations" using the mapping are thus valid for longer times than numerical integrations of the corresponding differential equations. Finally, and perhaps most importantly, the mappings include high frequency contributions which are usually ignored yet seem to introduce new types of motion, i.e. the slow diffusion to a region of phase space where large spikes occur in the eccentricity. The discovery of the spikes introduces a new possibility for the origin of the Kirkwood gaps.The hypothesis is that asteroids near the commensurabilities undergo large jumps in eccentricity thus becoming Mars crossers. They are subsequently removed from the gaps through collisions or close encounters with Mars. To test this hypothesis I studied the evolution over two million years of a distribution of 300 test asteroids with initial conditions near the 3/1 commensurability. If Mars crossers are removed, a gap is produced at the proper location but it is too narrow, both in the planar elliptic approximation and in the full three dimensional problem with the secular perturbations of Jupiter included. There remains the possibility that two million years of evolution is too short. One orbit near the boundary of the gap in the mapping distribution was "integrated" for a much longer time and after 18 million years it had a spike in

eccentricity. This implies that two million years is simply too short a time to see the full width of the gap open. The difficulty is that even for the mapping 18 million years is a long time and it is not clear to what extent the diffusion is due to the actual dynamics or to roundoff error. Ultimately these results cannot be believed until the jump and the diffusion to the jump are understood, at least approximately, in an analytic theory.

ACKNOWLEDGEMENTS

It is a pleasure to thank Peter Goldreich for his encouragement and for the uncountably many discussions and suggestions without which this work would never have been done. It is also a pleasure to thank Hans Scholl and Claude Froeschlé for providing details of their calculations and for their continued interest. The data from the TRIAD file were generously provided by Jim Williams. This work was partially supported by NASA Grant NGL 05-002-003 and NSF Grant AST-8020005.

REFERENCES

Brouwer, D., and Clemence, G.M. (1961). *Methods of Celestial Mechanics* (Academic, New York).

Brouwer, D., and van Woerkom, A.J.J. (1950). Astron. Papers of the American Eph. **13**, 81.

Chirikov, B.V. (1979). Phys. Rep. **52**, 263.

Froeschlé, C., and Scholl, H. (1976). Astron. Astrophys. **48**, 389.

Froeschlé, C., and Scholl, H. (1977). Astron. Astrophys. **57**, 33.

Giffen, R. (1973). Astron. Astrophys. **23**, 387.

Greenberg, R., and Scholl, H. (1979). *Asteroids* (University of Arizona Press, Tucson, Arizona), p. 310.

Hénon, M., and Heiles, C. (1964). Astron. J. **69**, 73.

Heppenheimer, T.A. (1978). Astron. Astrophys. **70**, 457.

Leverrier, U.-J. (1855). Annales de l'Obs. Paris, Mém. **1** (Observatory of Paris, Paris).

Moser, J. (1973). *Stable and Random Motions in Dynamical Systems* (Princeton University, Princeton, N.J.).

Peirce, B. (1849). Astron. J. **1**, 1.

Plummer, H.C. (1960). *An Introductory Treatise on Dynamical Astronomy* (Dover, New York).

Scholl, H., and Froeschlé, C. (1974). Astron. Astrophys. **33**, 455.

Scholl, H., and Froeschlé, C. (1975). Astron. Astrophys. **42**, 457.

Schubart, J. (1964). Smithsonian Astrophys. Obs. Spec. Rep. No. 149.

Schubart, J. (1968). Astron. J. **73**, 99.

Schweizer, F. (1969). Astron. J. **74**, 779.

Wetherill, G.W. (1975). Proc. Lunar Sci. Conf. 6th, 1539.

Wiesel, W.E. (1976). Celestial Mech. **13**, 3.

Wisdom, J. (1980). Astron. J. **85**, 1122.

FIGURE CAPTIONS

FIGURE 1 Eccentricity versus time using the averaged differential equations (in the planar elliptic approximation) for the initial conditions $\Phi = 0$, $\varphi = \pi$, $y_1 = 0$ and $e = 0.05$. Here t is measured in units of 10,000 $T_J \approx$ 240,000 years. T_J is Jupiter's period.

FIGURE 2 Eccentricity versus time using the planar elliptic mapping for the same initial conditions as in Fig. 1. As in Fig. 1 the time is measured in units of 10,000 $T_J \approx$ 240,000 years.

FIGURE 3 The times to first jump (in years) plotted versus the mapping frequency Ω. The initial conditions are the same as Figs. 1 and 2. Jupiter's actual frequency corresponds to $\Omega = 1$.

FIGURE 4 The height of the jump in eccentricity versus the value given to Jupiter's eccentricity e_J. Froeschlé and Scholl use $e_J = 0.048$. Jupiter has a maximum eccentricity of 0.0601734.

FIGURE 5 The eccentricities versus semimajor axes for the 300 test asteroids of the distribution.

FIGURE 6 The same as Fig. 5 but with the initial conditions of those test asteroids that became Mars crossers in two million years removed.

FIGURE 7 Histograms of the eccentricities and semimajor axes averaged over 5000 years for the full distribution of initial conditions.

FIGURE 8 Histograms of the averaged eccentricities and semimajor axes for those test asteroids that did not become Mars crossers in two million years.

FIGURE 9 The eccentricities versus semimajor axes for all asteroids in the TRIAD file with semimajor axes between 0.47 and 0.49.

FIGURE 10 The same as Fig. 6 but using the full mapping.

FIGURE 11 The same as Fig. 7 but using the full mapping.

FIGURE 12 The same as Fig. 8 but using the full mapping.

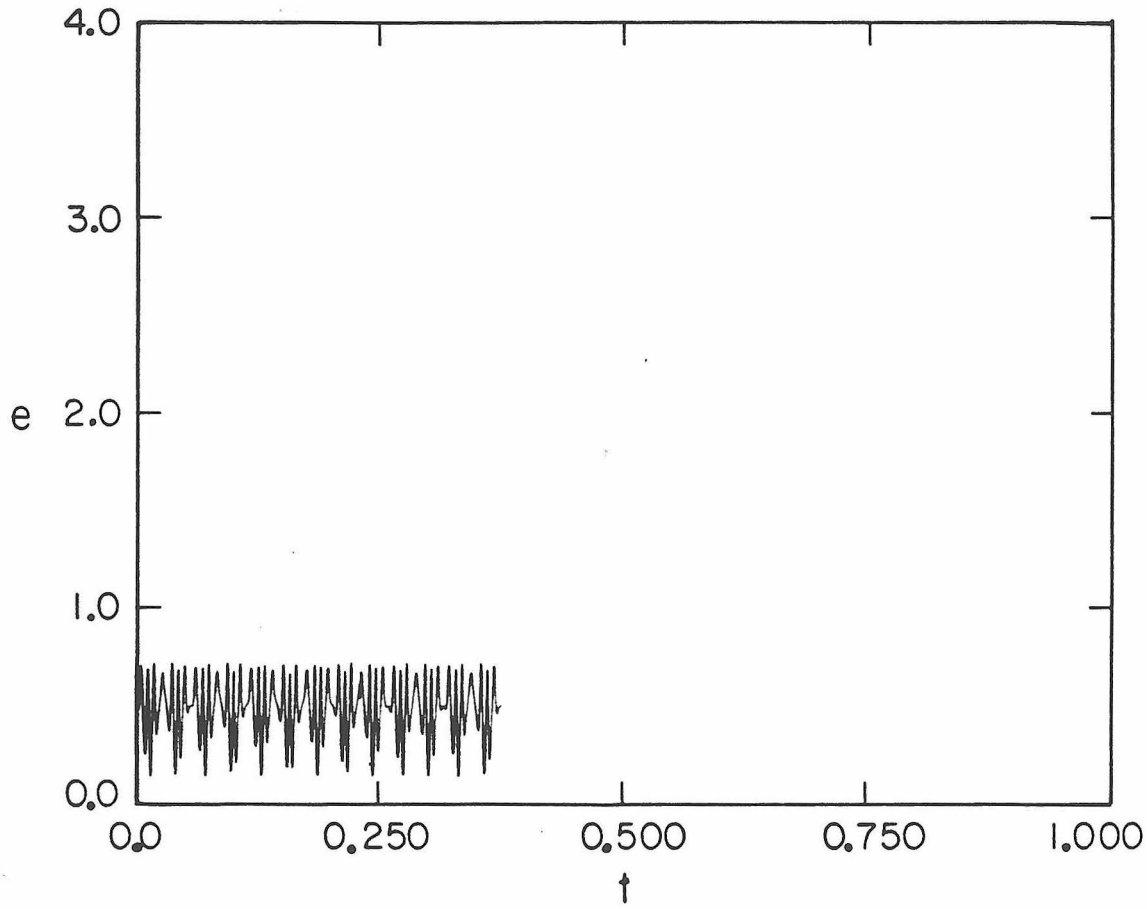


FIGURE 1

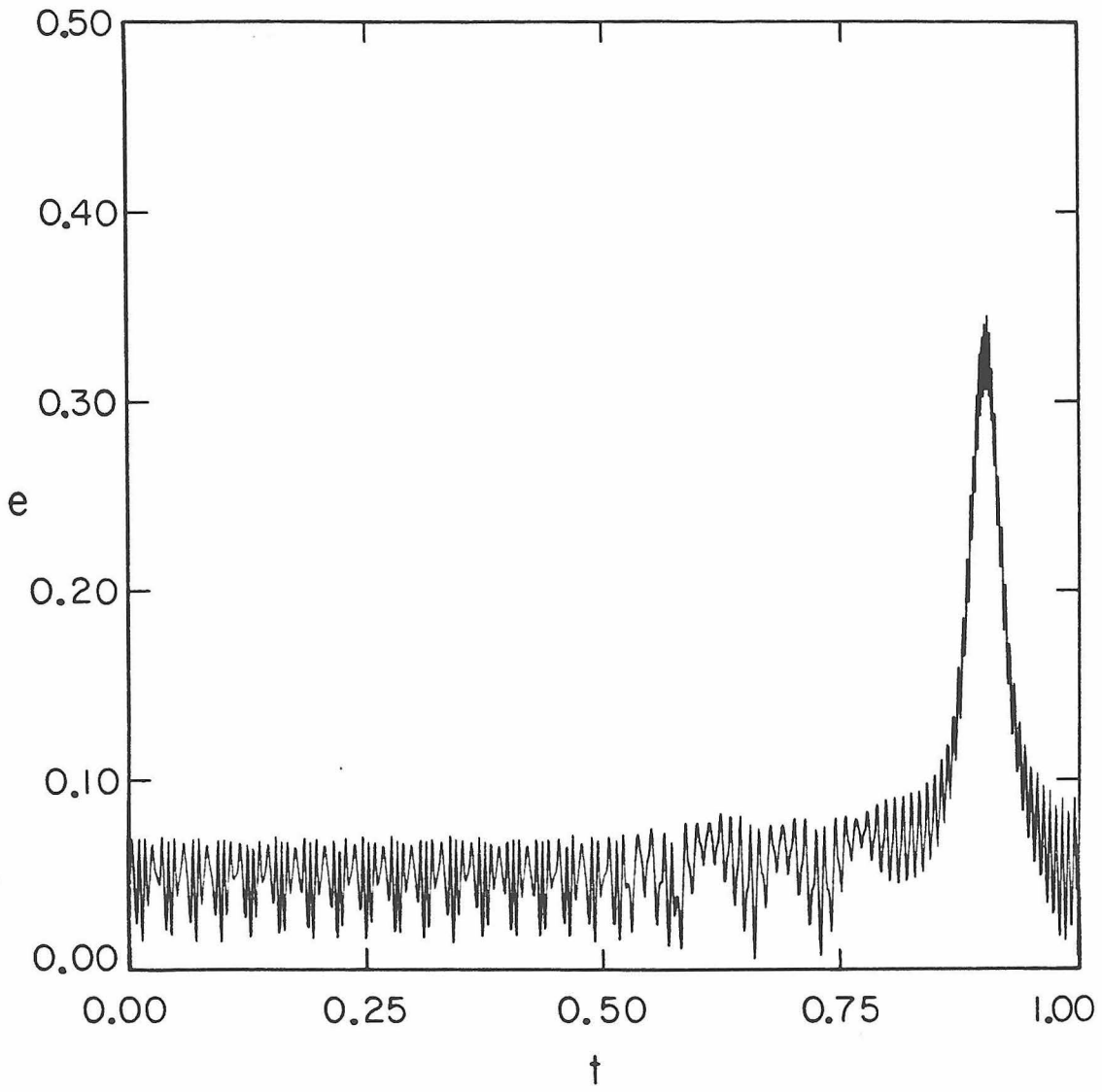


FIGURE 2

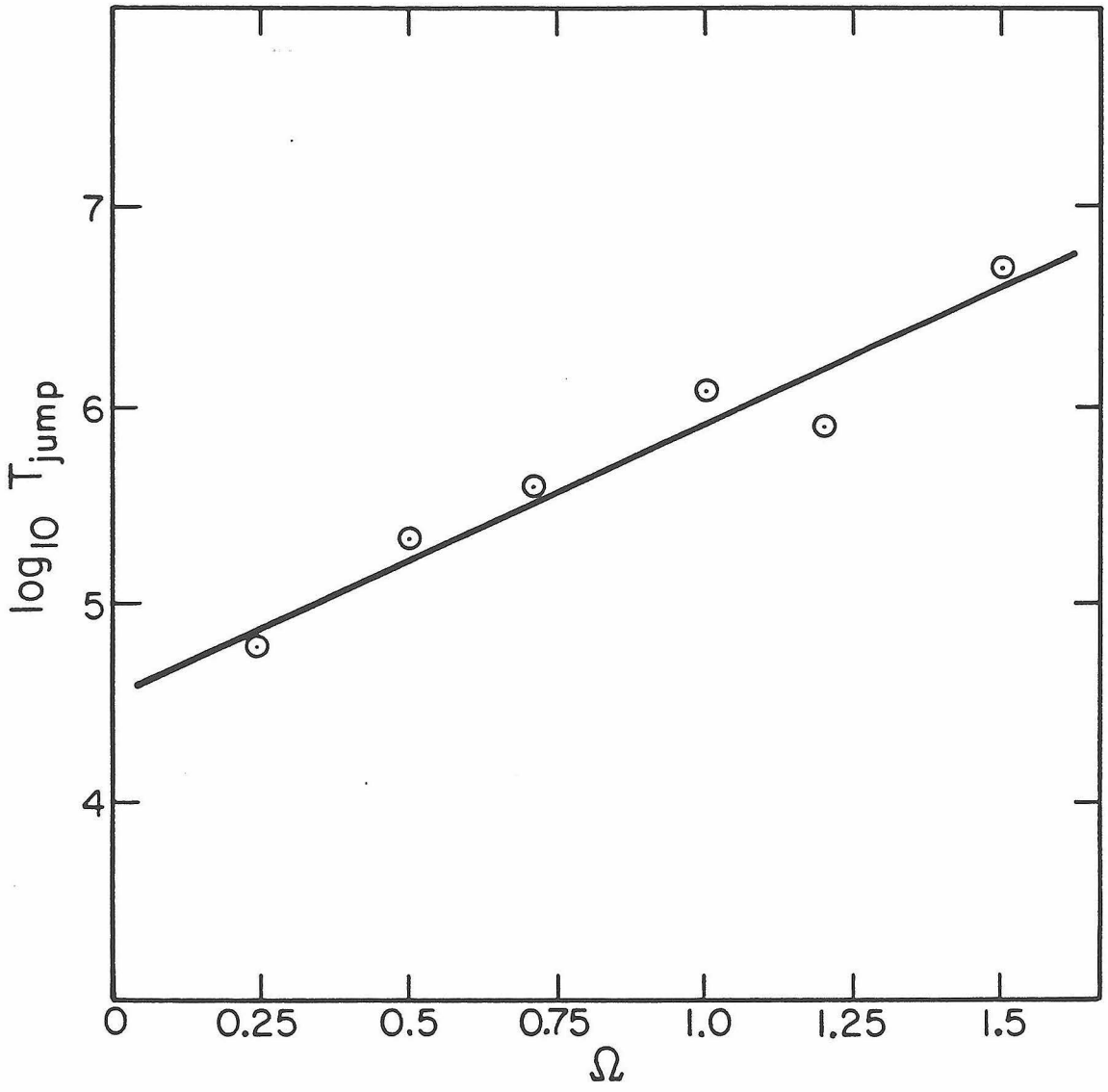


FIGURE 3

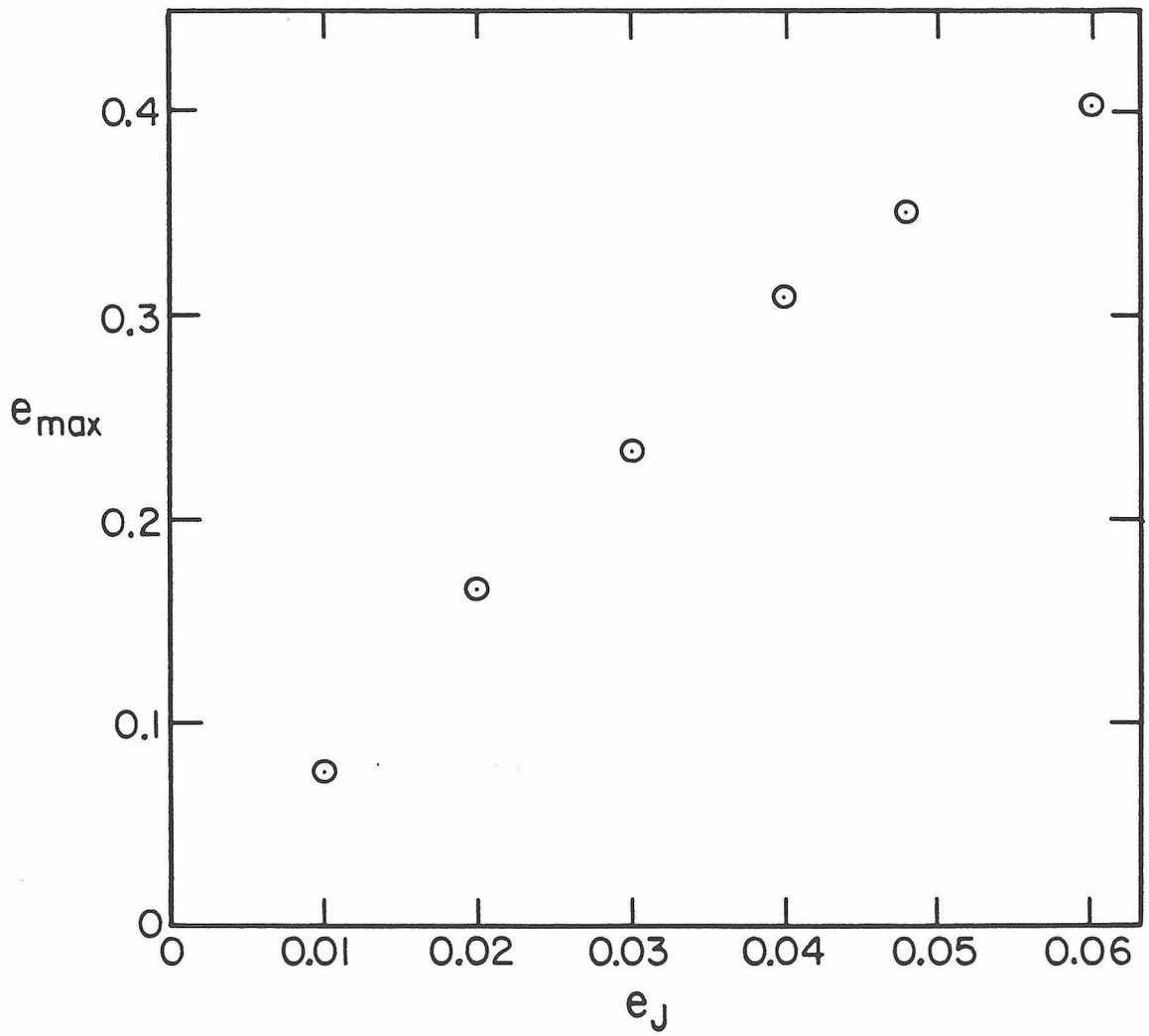


FIGURE 4

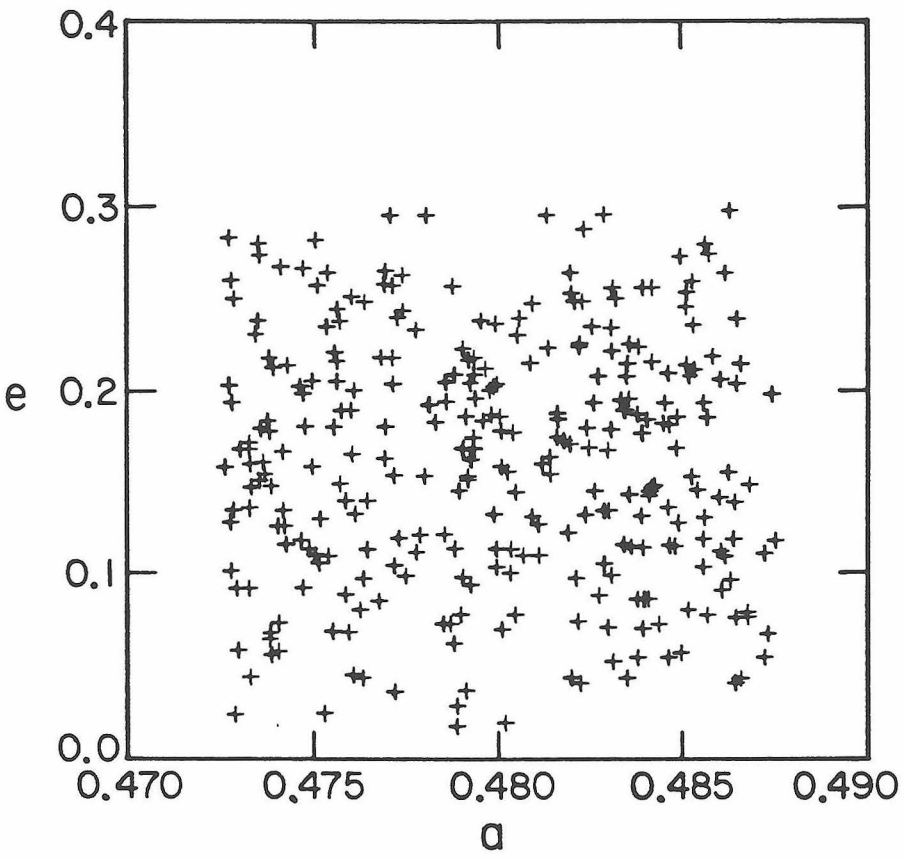


FIGURE 5

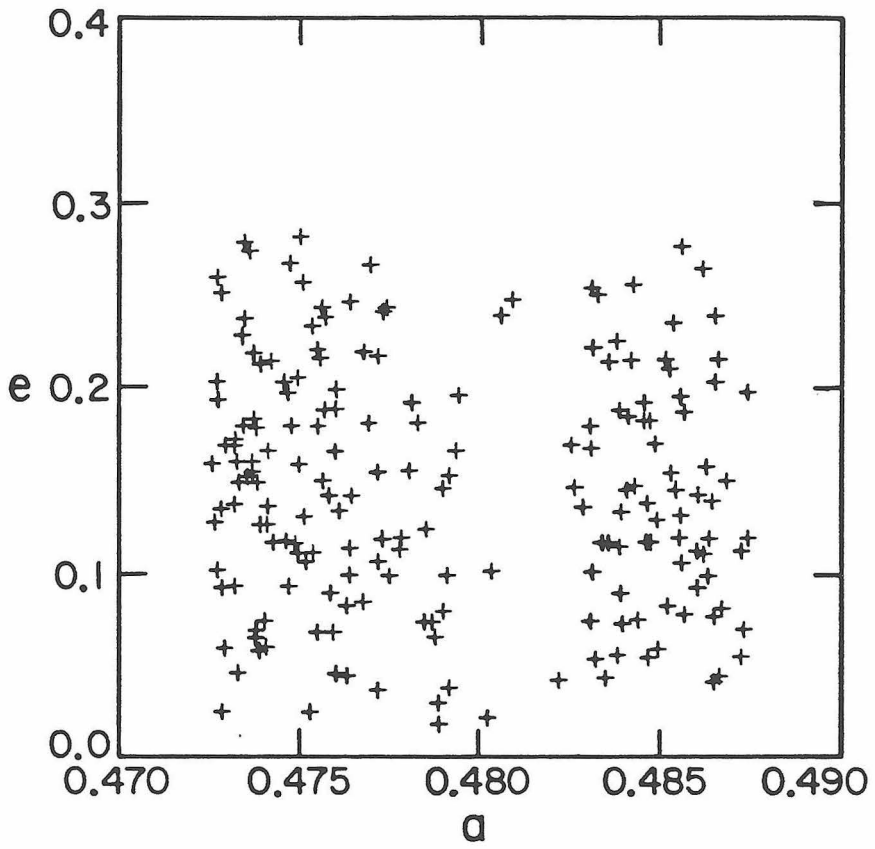


FIGURE 6

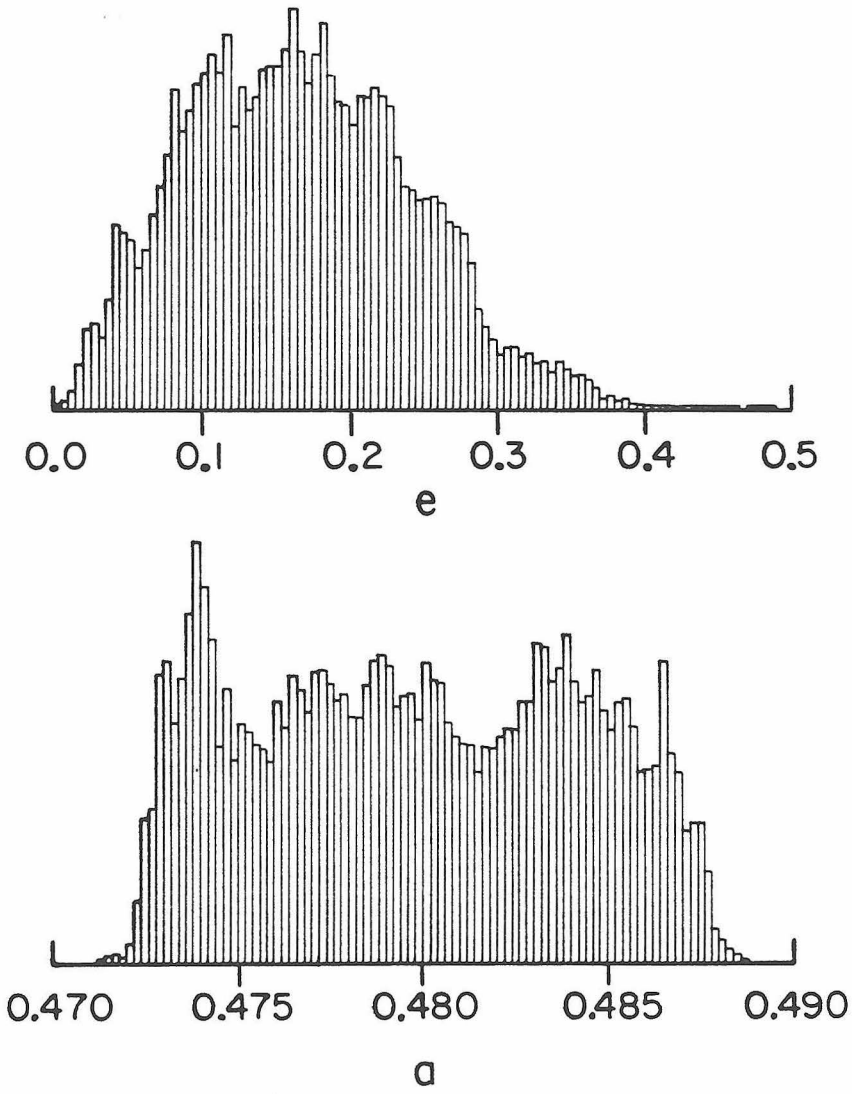


FIGURE 7

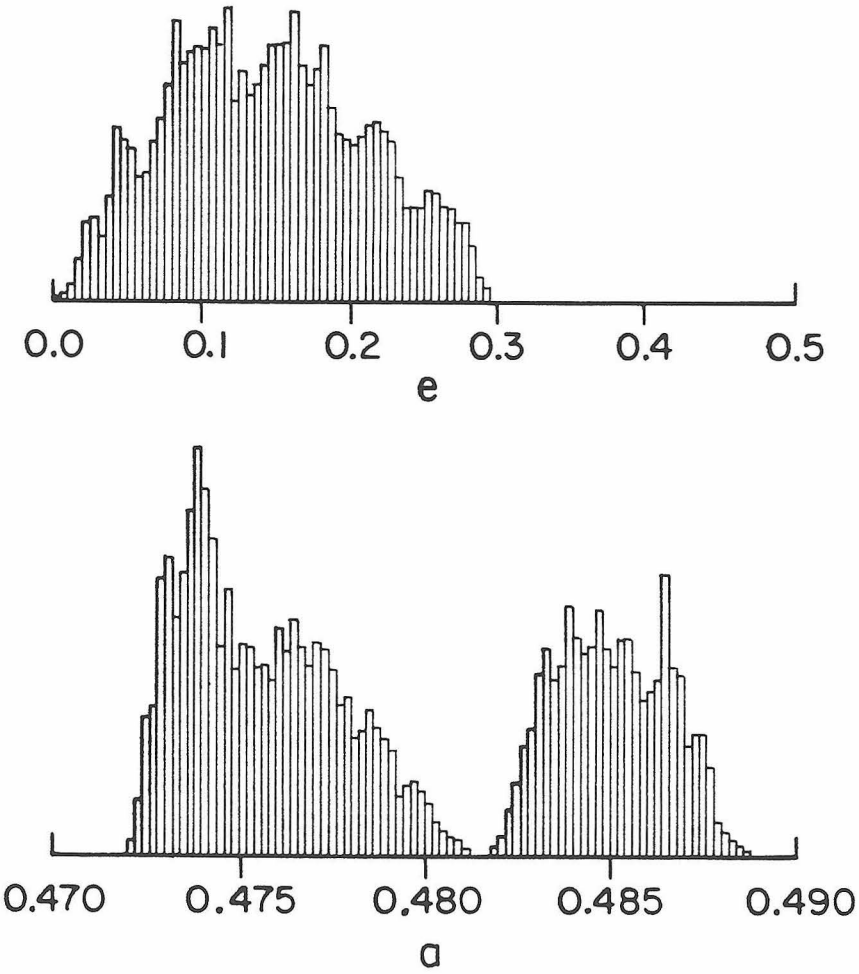


FIGURE 8

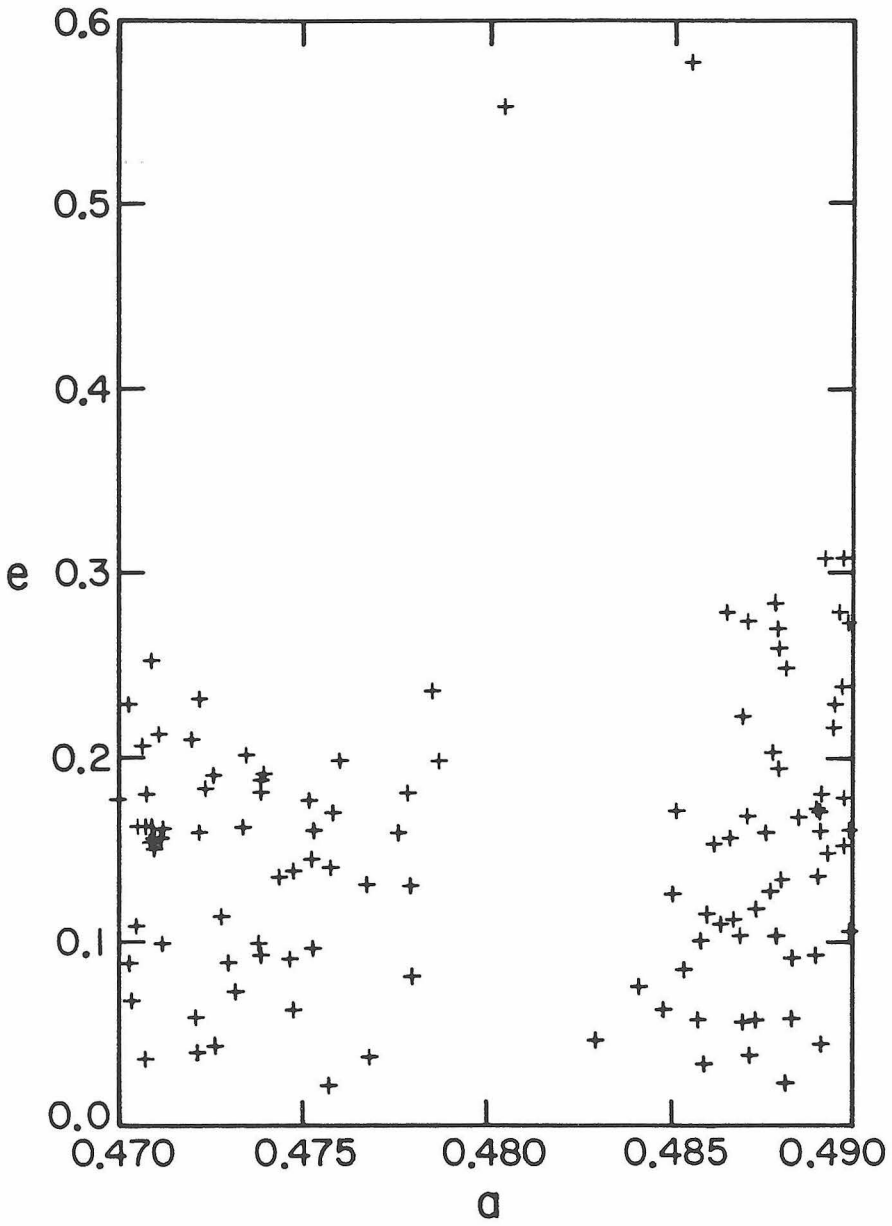


FIGURE 9

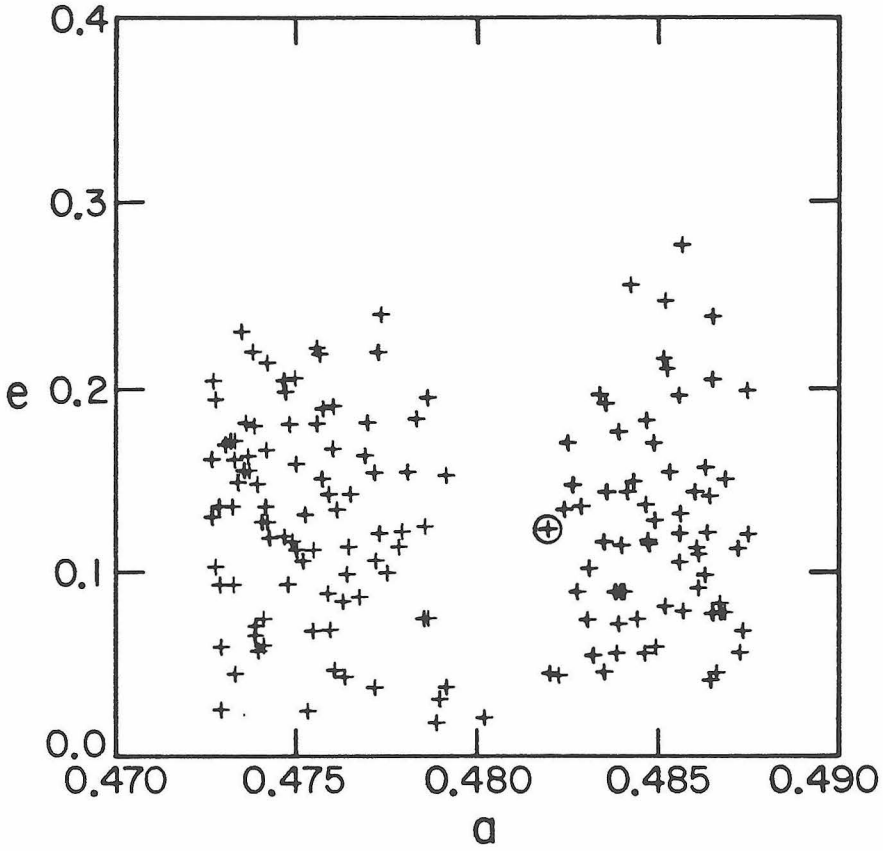


FIGURE 10

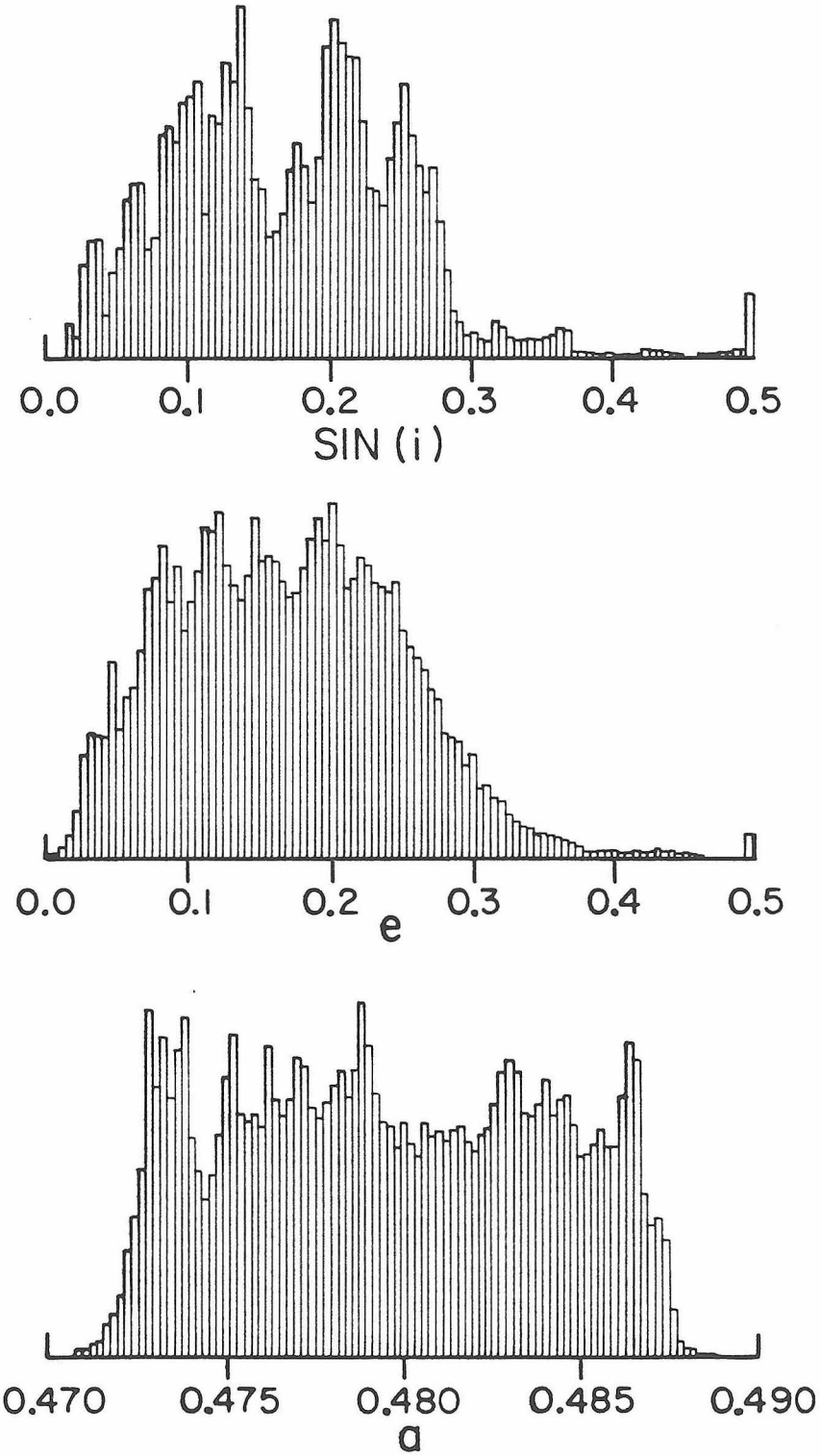
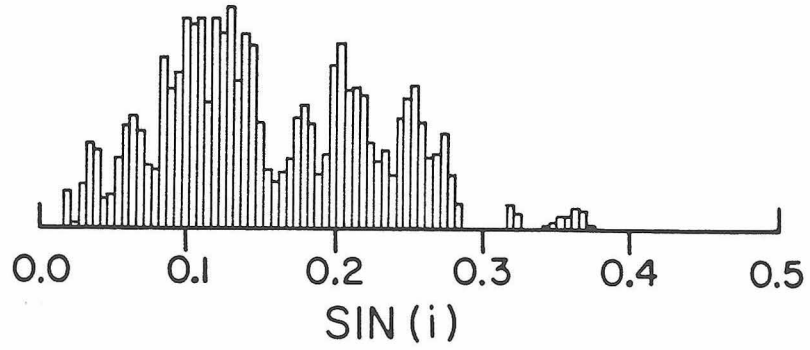
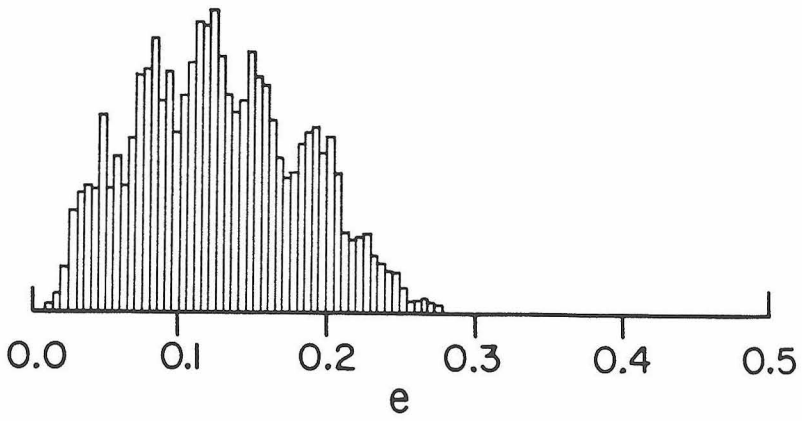

0.0
0.1
0.2
0.3
0.4
0.5
SIN (i)
0.0
0.1
0.2
0.3
0.4
0.5
e
0.470
0.475
0.480
0.485
0.490
a

FIGURE 11


0.0
0.1
0.2
0.3
0.4
0.5
SIN (i)


0.0
0.1
0.2
0.3
0.4
0.5
e

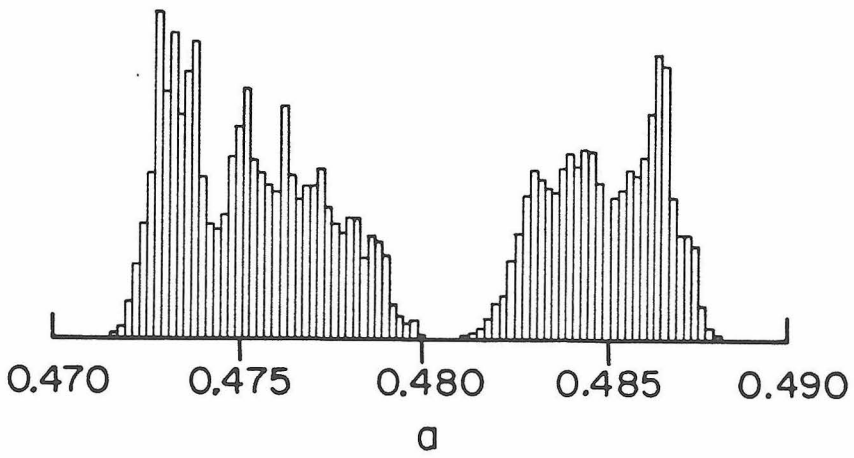

0.470
0.475
0.480
0.485
0.490
a

FIGURE 12

PAPER 2

THE RESONANCE OVERLAP CRITERION AND THE ONSET OF STOCHASTIC BEHAVIOR IN THE RESTRICTED THREE-BODY PROBLEM

by Jack Wisdom

Published in the *Astronomical Journal*

Volume 85, pages 1122-1133

Reprinted from:
THE ASTRONOMICAL JOURNAL VOLUME 85, NUMBER 8 AUGUST 1980

THE RESONANCE OVERLAP CRITERION AND THE ONSET OF STOCHASTIC BEHAVIOR IN THE RESTRICTED THREE-BODY PROBLEM[a)]

JACK WISDOM
California Institute of Technology, Pasadena, California 91125
Received 19 February 1980; revised 22 April 1980

ABSTRACT

The resonance overlap criterion for the onset of stochastic behavior is applied to the planar circular-restricted three-body problem with small mass ratio (μ). Its predictions for $\mu = 10^{-3}$, 10^{-4}, and 10^{-5} are compared to the transitions observed in the numerically determined Kolmogorov-Sinai entropy and found to be in remarkably good agreement. In addition, an approximate scaling law for the onset of stochastic behavior is derived.

I. INTRODUCTION

Is the solar system stable? There is as yet no answer to this notoriously difficult question, but some important steps have been made in the study of the stability of dynamical systems in the last two decades. Hénon and Heiles (1964) discovered in their now-classic study that the phase space of a simple nonlinear Hamiltonian system with two degrees of freedom (a model for the motion of stars in the Galaxy) was divided into regions that contain quasiperiodic trajectories and regions in which trajectories have a random character. Subsequent numerical experiments by Hénon (1966), Bozis (1966), and Jefferys (1966) then verified that the phase space of the planar circular-restricted three-body problem is similarly divided. The existence of regions of quasiperiodic trajectories is very important since all such trajectories possess long-term stability. While there is no rigorous way of predicting which regions will be stochastic, an approximate criterion involving the overlap of zero-order nonlinear resonances has been developed which has had considerable success in other problems (see Walker and Ford 1969 and the recent review in Chirikov 1979). The solar system is far too complicated, though, for a direct application of the resonance overlap criterion. To get our foot in the door of dynamical astronomy and gain confidence in the overlap criterion, I begin instead with the simplest of unsolved problems in dynamical astronomy. In this paper I apply the resonance overlap criterion to the planar circular-restricted three-body problem and compare its results to some numerical experiments.

In Sec. II, I review the resonance overlap criterion. I then apply the method to the restricted three-body problem in Sec. III. A comparison of its predictions with some numerical experiments is presented in Sec. IV. In Sec. V, I derive an approximate scaling law for resonance overlap, and in Sec. VI, I state my conclusions.

II. RESONANCE OVERLAP AND THE CHIRIKOV CRITERION

Consider a Hamiltonian of the form

$$H = H_0(J_1,J_2) + \mu \sum_{i=0}^{\infty} \sum_{j=-\infty}^{\infty} \times H_{ij}(J_1,J_2)\cos(i\theta_1 + j\theta_2), \quad (1)$$

where θ_l are the coordinates canonically conjugate to the momenta J_l and μ is a small parameter. One may attempt to solve this problem by first solving the zero-order Hamiltonian

$$H^{(0)}_{nm} = H_0(J_1,J_2) + \mu \sum_{k=0}^{\infty} \times H_{kn,km}(J_1,J_2)\cos[k(n\theta_1 + m\theta_2)] \quad (2)$$

and then perturbing the zero-order solutions with the remaining terms. This zero-order Hamiltonian will be useful whenever the resonance condition

$$n\omega_1(J_1,J_2) + m\omega_2(J_1,J_2) = 0 \quad (3)$$

is approximately satisfied, where

$$\omega_l(J_1,J_2) \equiv \frac{\partial(H_0 + \mu H_{00})}{\partial J_l}, \quad (4)$$

since the arguments of the cosines will then be approximately stationary. I assume n and m have no common divisors. In terms of the resonance variables

$$\psi = n\theta_1 + m\theta_2$$

and (5)

$$\varphi = -\theta_2/n,$$

this Hamiltonian assumes the simpler form

$$H^{(0)}_{nm} = H'_0(\Phi,\Psi) + \mu \sum_{k=0}^{\infty} H'_{kn,km}(\Phi,\Psi)\cos k\psi, \quad (6)$$

where

$$\Psi = J_1/n$$

[a)] Contribution No. 3390 of the Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125.

0004-6256/80/081122-12$00.90

and

$$\Phi = mJ_1 - nJ_2 \tag{7}$$

are the momenta canonically conjugate to ψ and φ, respectively. Since $H^{(0)}_{nm}$ has no explicit time dependence and is cyclic in φ, the system has two integrals of the motion, the Hamiltonian itself and Φ. Liouville's theorem (see Whittaker 1961) then guarantees that there is a canonical transformation to the system in which these integrals are the new momenta and the new Hamiltonian

$$\tilde{H}^{(0)}_{nm} = \tilde{H}^{(0)}_{nm}(H^{(0)}_{nm}, \Phi) \tag{8}$$

is cyclic in the new coordinates and thus trivially integrable. The character of the solutions is, however, most easily obtained by studying the contours of $H^{(0)}_{nm}$ on surfaces of constant Φ. To illustrate this, I make three simplifying assumptions. The first two are that we can ignore $H'_{kn,km}$ for all k except $k = 1$ and that $H'_{n,m}(\Phi,\Psi)$ is sufficiently well approximated near the resonance by $H'_{n,m}(\Phi,\Psi_R)$, where Ψ_R is defined implicitly by the resonance condition

$$\omega_\psi(\Phi,\Psi_R) \equiv \frac{\partial H'_0(\Phi,\Psi)}{\partial\Psi}\bigg|_{\Psi=\Psi_R} = 0. \tag{9}$$

The third simplifying assumption is that $H_0(\Phi,\Psi)$ is sufficiently well approximated by the quadratic terms in its Taylor series about $\Psi = \Psi_R$. Under these assumptions the resonance Hamiltonian is approximately

$$H^{(0)}_{nm} \simeq H'_0(\Phi,\Psi_R) + \frac{1}{2}\frac{\partial^2 H'_0}{\partial\Psi^2}\bigg|_{\Psi=\Psi_R} \times (\Psi - \Psi_R)^2 + \mu H'_{n,m}(\Phi,\Psi_R)\cos\psi, \tag{10}$$

where the linear term is absent because of the resonance condition (9). The level curves of this approximate $H^{(0)}_{n,m}$ are then explicitly

$$\Psi = \Psi_R \pm \left(\frac{H^{(0)}_{nm} - H'_0(\Phi,\Psi_R) - \mu H'_{n,m}(\Phi,\Psi_R)\cos\psi}{\frac{1}{2}\frac{\partial^2 H'_0}{\partial\Psi^2}\Big|_{\Psi=\Psi_R}}\right)^{1/2}. \tag{11}$$

Figure 1 illustrates these contours in the Cartesian coordinates $x = (2\Psi)^{1/2}\cos\psi$ and $y = (2\Psi)^{1/2}\sin\psi$. In drawing this figure, I have assumed $H^{(0)}_{nm} - H'_0(\Phi,\Psi_R)$ and $H'_{n,m}(\Phi,\Psi_R)$ are opposite in sign, and have arbitrarily restricted Ψ to be greater than zero. The extremum at the origin is an artifact of this restriction. It is clear that if a contour does not enclose the origin, then the angle ψ oscillates, whereas if a contour encircles the origin, ψ circulates. The oscillation region has been shaded in Fig. 1. The contours that form the boundary of the oscillation region are the so-called separatrices, which play an important role in the resonance overlap criterion. For the approximate Hamiltonian (10) the separatrices are

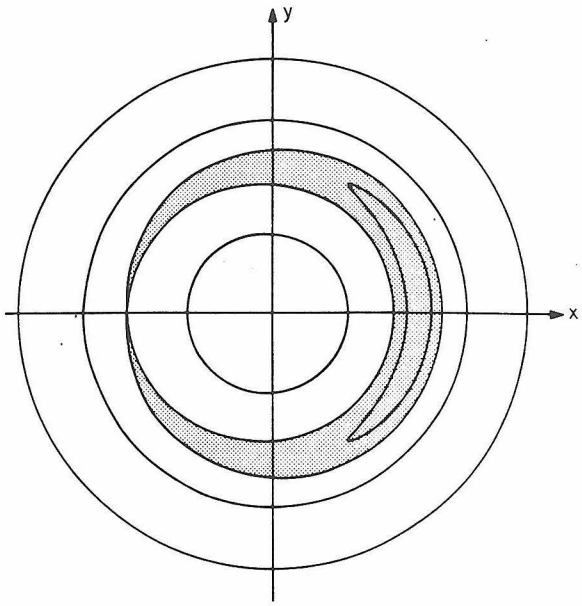


FIG. 1. Contours of the approximate Hamiltonian (10) on a surface of constant Φ. The Cartesian coordinates are $x = (2\Psi)^{1/2}\cos\psi$ and $y = (2\Psi)^{1/2}\sin\psi$.

$$\Psi_S = \Psi_R \pm \Delta\Psi\cos\frac{\psi}{2}, \tag{12}$$

where the resonance half-width $\Delta\Psi$ is defined by

$$\Delta\Psi \equiv 2\left(\frac{\mu H'_{n,m}(\Phi,\Psi_R)}{\frac{\partial^2 H'_0}{\partial\Psi^2}\Big|_{\Psi=\Psi_R}}\right)^{1/2}. \tag{13}$$

Now, if the three simplifying assumptions leading to the approximate Hamiltonian (10) are valid, the contours of the full resonance Hamiltonian (6) will be qualitatively the same as those in Fig. 1. For a more general Hamiltonian than that of Eq. (10), the contours can be quite different and require a detailed numerical mapping to determine the oscillation regions.

Having, in principle, completely solved the zero-order resonance Hamiltonian (6) by reducing it to the form of Eq. (8), I now ask what effect the other terms in the full Hamiltonian (1) will have on the unperturbed solutions. In the system with $H^{(0)}_{nm}$ and Φ as momenta, the full Hamiltonian has the form

$$H = \tilde{H}_{nm}(H^{(0)}_{nm}, \Phi) + \mu\sum_{i=0}^{\infty} \times \sum_{j=-\infty}^{\infty} \tilde{H}_{ij}(H^{(0)}_{nm}, \Phi)\cos(ih + j\varphi), \tag{14}$$

where h and φ are canonically conjugate to $H_{nm}^{(0)}$ and Φ, respectively. If the resonance conditions of this Hamiltonian analogous to the resonance condition (3) of the zero-order resonance Hamiltonian are sufficiently poorly satisfied and the H_{ij} are small enough, then the Kolmogorov-Arnol'd-Moser Theorem (see Moser 1973) assures us that the motion is still quasiperiodic and only slightly perturbed. On the other hand, if a second resonance of the original Hamiltonian is "sufficiently close," then the motion is more complicated. In fact, there has been no successful analytic attempt to solve for the motion under the simultaneous influence of two "close" resonances, nor is there a rigorous analytic estimate of what "sufficiently close" means. The basic idea of the resonance overlap criterion is that two resonances are "sufficiently close" when a separatrix of one resonance has crossed a separatrix of the other resonance, i.e., when the zero-order analysis indicates that two different resonance angles both oscillate. Ideally one would map all initial conditions in the four-dimensional phase space that lead to oscillation of each resonance angle and then look for the overlap of these regions. In practice, it is easier to specify initial values of θ_1 and θ_2 and then plot the separatrices for each resonance in the $J_1 - J_2$ plane. The simple example of Walker and Ford (1969) provides an excellent introduction to this method. Chirikov (see Chirikov 1979) has developed an approximate criterion. Chirikov first calculates the half-width of each resonance by Eq. (13), having made all three assumptions leading to the approximate Hamiltonian (10) and implicitly choosing θ_1 and θ_2 for each resonance to give the maximum width. He then calculates frequency half-widths by the approximate relation

$$\Delta\omega_l^q \simeq \left.\frac{\partial\omega_l}{\partial\Psi}\right|_{\Psi=\Psi_R} \Delta\Psi, \tag{15}$$

with the frequencies of Eq. (4). The superscript identifies the resonance under study. The half-widths of nearby resonances are then compared to the separation of the resonance centers,

$$\delta\omega_l^{qb} = |\omega_l^q - \omega_l^b|. \tag{16}$$

If both inequalities

$$\Delta\omega_1^q + \Delta\omega_1^b \geq \delta\omega_1^{qb}$$

and (17)

$$\Delta\omega_2^q + \Delta\omega_2^b \geq \delta\omega_2^{qb}$$

are satisfied, then there is resonance overlap. This then is Chirikov's approximate criterion. In Sec. III, I apply the resonance overlap criterion to the restricted three-body problem.

III. RESONANCE OVERLAP IN THE RESTRICTED THREE-BODY PROBLEM

In terms of the Delaunay canonical elements (see, e.g., Brouwer and Clemence 1961), the Hamiltonian for the planar circular-restricted three-body problem is

$$H = -\frac{(1-\mu)^2}{2L^2} - R, \tag{18}$$

where

$$R = \mu \sum_{i=0}^{\infty} \sum_{j=-\infty}^{\infty} K^{(i,j)} \cos N^{(i,j)} \tag{19}$$

and

$$N^{(i,j)} = il + j(t-g). \tag{20}$$

I have chosen units so that the product of the gravitational constant and the sum of the two masses is unity and the separation of the two masses is also unity. In these units the secondary has mass μ, which I assume is small compared to unity. In terms of the usual osculating elliptic elements, the canonical momenta are $L = [(1-\mu)a]^{1/2}$ and $G = [(1-\mu)a(1-e^2)]^{1/2}$, where a is the semimajor axis and e is the eccentricity. Their conjugate coordinates are the mean anomaly l and the angle of periapse g, respectively. $K^{(i,j)}$ is a function of L and G, and t is the time. A resonance occurs when one of the cosine arguments is nearly stationary. Since this Hamiltonian is time dependent, the resonance condition assumes a slightly different form,

$$0 = -s\omega_l(L,G) + (s+s')[1 - \omega_g(L,G)], \tag{21}$$

where s and s' are integers and the frequencies are defined in the usual way:

$$\omega_l \equiv \frac{\partial}{\partial L}\left(\frac{-(1-\mu)^2}{2L^2} - \mu K^{(0,0)}(L,G)\right)$$

and (22)

$$\omega_g \equiv \frac{\partial}{\partial G}[-\mu K^{(0,0)}(L,G)].$$

If we ignore the terms proportional to μ, we get the approximate resonance condition

$$0 = -s/L^3 + (s+s'), \tag{23}$$

or, in terms of the semimajor axis,

$$\alpha_{ss'} = [s/(s+s')]^{2/3}. \tag{24}$$

Poincaré (1902) was the first to study motion near a resonance in the restricted three-body problem by means of a zero-order resonance Hamiltonian of the type discussed in Sec. II. I perform a canonical transformation to the Poincaré resonance variables, as generalized by Woltjer (1923) and Hagihara (1943),

$$\varphi = l + g - t$$

and (25)

$$\psi = -sl + (s+s')(t-g)$$

via the generating function

$$F = [-sl + (s+s')(t-g)]\Psi + [l+g-t]\Phi. \tag{26}$$

Their respective conjugate momenta are

$$\Phi = \frac{(s+s')L - sG}{s'}$$

and (27)

$$\Psi = \frac{L-G}{s'}.$$

In terms of the osculating elliptic elements, these are

$$\Phi = [(1-\mu)a]^{1/2}\left(\frac{(s+s') - s(1-e^2)^{1/2}}{s'}\right)$$
$$= [(1-\mu)a]^{1/2}\left[1 + \frac{s}{2s'}e^2 + o(e^4)\right]$$

and (28)

$$\Psi = [(1-\mu)a]^{1/2}\,[1-(1-e^2)^{1/2}]$$
$$= [(1-\mu)a]^{1/2}\left[\frac{e^2}{2s'} + o(e^4)\right].$$

The new Hamiltonian is

$$H' = H + \frac{\partial F}{\partial t} = -\frac{(1-\mu)^2}{2(\Phi - s\Psi)^2} + (s+s')\,\Psi - \Phi - R, \quad (29)$$

where R is to be written in terms of the new variables. In particular,

$$N^{(i,j)} = i\left(\frac{(s+s')\,\varphi + \psi}{s'}\right) + j\left(\frac{\psi + s\varphi}{s'}\right). \quad (30)$$

Note that this new Hamiltonian is explicitly time independent and is thus an integral of the motion. In terms of the Poincaré variables, the resonance condition is

$$\omega_\psi \equiv \frac{\partial}{\partial \Psi}\left(\frac{-(1-\mu)^2}{2(\Phi - s\Psi)^2} + (s+s')\Psi - \Phi - \mu K^{(0,0)}\right) = 0. \quad (31)$$

The term $\mu K^{(0,0)}$ only shifts the position of the resonance by a quantity of order μ and will be ignored in the rest of this paper. The zero-order resonance Hamiltonian must contain all those terms with nearly stationary arguments, i.e., those independent of φ. Inspection of Eq. (30) reveals that the terms independent of φ satisfy

$$js = -i(s+s'), \quad (32)$$

and that for these terms,

$$\cos N^{[is,-i(s+s')]} = \cos i\psi. \quad (33)$$

If we define

$$K_i \equiv K^{[is,-i(s+s')]}, \quad (34)$$

then the zero-order resonance Hamiltonian is

$$H_{ss'}^{(0)} = \frac{-(1-\mu)^2}{2(\Phi - s\Psi)^2} + (s+s')\,\Psi - \Phi - \mu\sum_{i=1}^{\infty} K_i \cos i\psi. \quad (35)$$

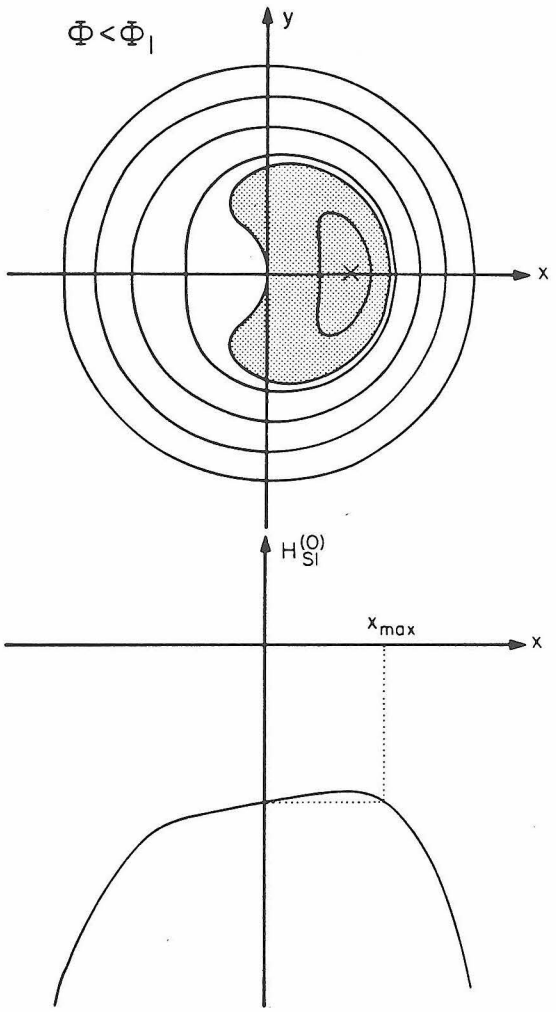


FIG. 2. Contours of constant $H_{s1}^{(0)}$ on a surface of constant Φ for $\Phi < \Phi_1$. The extremum is marked by a cross (×). The region in which Ψ oscillates is shaded. Also shown is a plot of $H_{s1}^{(0)}$ along the x axis which illustrates the definition of x_{max}.

I will be primarily interested in direct motion inside the secondary ($a < 1$) with small eccentricity ($e \leq 0.15$). Because $K_i \propto e^{i|s'|}$ and $s' \geq 1$ when $a < 1$ (see Brouwer and Clemence 1961), the most important resonances are those with $s' = 1$. I will consider only the $s' = 1$ resonances. This greatly simplifies the application of the resonance overlap criterion.

The contours of constant $H_{s1}^{(0)}$ on surfaces of constant Φ have been studied many times (see, for example, Schubart 1964, Message 1966, Jefferys 1966, and Wiesel 1976). Here I will only review the results of these discussions and mention some new features. For a given resonance there are three critical values of Φ which separate qualitatively different types of contours. Figures 2 through 5 illustrate the contours for these four regions. The Cartesian coordinates are $x = (2\Psi)^{1/2}\cos\psi$ and $y = (2\Psi)^{1/2}\sin\psi$. Extrema are marked by a cross (×); points at which contours cross are saddle points. Extrema

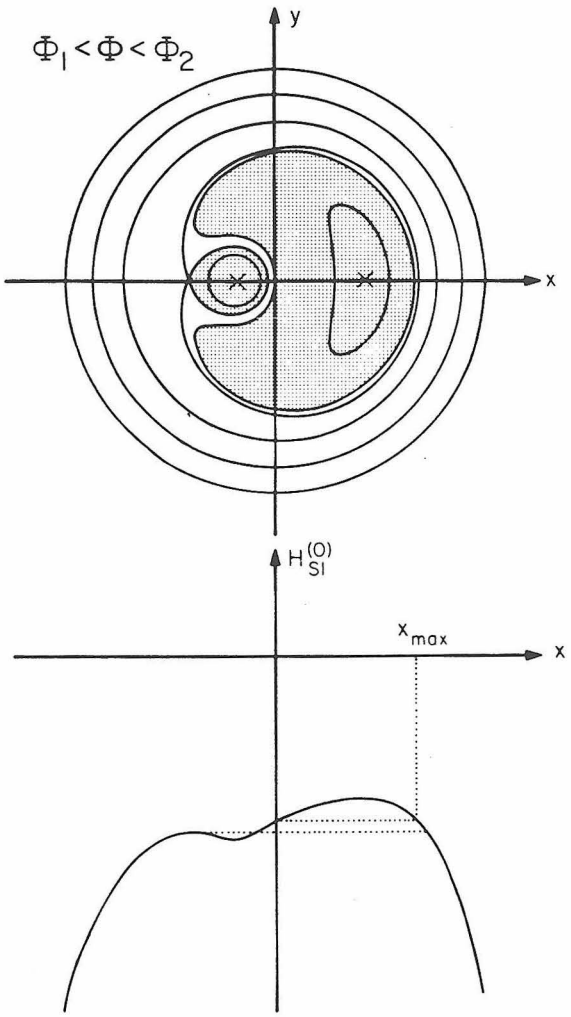


FIG. 3. The same as Fig. 2, but for $\Phi_1 < \Phi < \Phi_2$. There are now two oscillation regions, two extrema, and a saddle point. x_{max} is defined as before.

and saddle points in the zero-order Hamiltonian correspond to stable and unstable periodic orbits, respectively, in the full problem (see Message 1966 for a discussion). Though all the contours are symmetric about $y = 0$, it is not obvious from the form of the Hamiltonian that all the extrema and saddle points lie on the x axis. That this is in fact the case for $a < 1$ was shown by Message (1958). The behavior of the contours is thus completely characterized by plots of $H_{s1}^{(0)}$ along the x axis, which are also illustrated in Figs. 2 through 5. The regions in which ψ oscillates are shaded. For $\Phi < \Phi_1$ there is only one extremum for which $\psi = 0$, and only one oscillation region. The separatrix is that contour which passes through the origin. It crosses the x axis again at $x_{max} = (2\Psi_{max})^{1/2}$, where Ψ_{max} is obviously defined by

$$H_{s1}^{(0)}(\Phi,\Psi = 0) = H_{s1}^{(0)}(\Phi,\Psi_{max},\psi = 0). \quad (36)$$

This is illustrated in Fig. 2. As Φ is increased to Φ_1, a cusp appears which bifurcates into a saddle point and an extremum for $\Phi_1 < \Phi < (1-\mu)^{1/2}$. Φ_1 is near but not identical to Φ_s, which satisfies the modified "resonance condition"

$$0 = \omega_\psi|_{\psi=0} = \frac{-s}{\Phi_s^3} + (s+1). \quad (37)$$

This interval in Φ is further subdivided by Φ_2. Let Ψ_{sp} denote the value of Ψ at the saddle point. The interval $\Phi_1 < \Phi < \Phi_2$ is then characterized by the relation

$$H_{s1}^{(0)}(\Phi,\Psi_{sp},\psi = \pi) < H_{s1}^{(0)}(\Phi,\Psi = 0), \quad (38)$$

and is illustrated in Fig. 3. There are now two oscillation regions. The region that includes $\psi = 0$ is simply a continuation of the oscillation region for $\Phi < \Phi_1$. x_{max} is defined in the same way as before. When Φ is greater than Φ_2, inequality (38) is no longer satisfied and the contours change somewhat (see Fig. 4). The $\psi = 0$ oscillation region now has both an x_{min} and an x_{max}. The

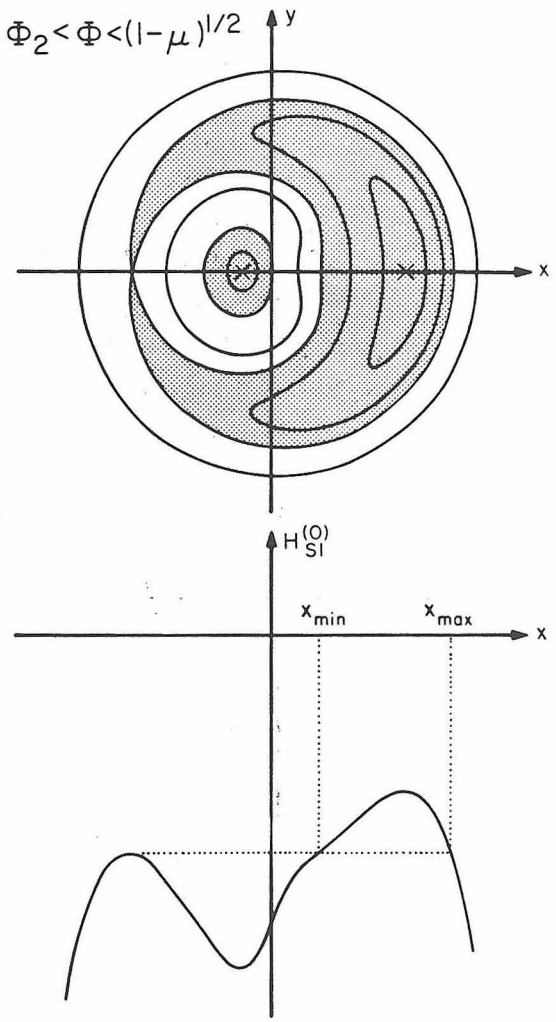


FIG. 4. The same as Figs. 2 and 3, but for $\Phi_2 < \Phi < (1-\mu)^{1/2}$. The oscillation region that includes $\psi = 0$ no longer extends to the origin. The definition of x_{min} is illustrated in the plot of $H_{s1}^{(0)}(y = 0)$ vs x.

contours for $\Phi > (1 - \mu)^{1/2}$ are illustrated in Fig. 5. They are complicated by the appearance of a singularity near the origin. The extremum with $\psi = \pi$ disappears and a new saddle point appears with $\psi = 0$. Near the singularity one does not expect the zero-order Hamiltonian to represent the motion accurately, so this region has been hatched.

The appearance of a new saddle point is quite interesting and has never been mentioned before. As I said above, saddle points correspond to unstable periodic orbits. Thus, this new saddle point corresponds to a new analytic family of periodic orbits in the restricted three-body problem. Colombo *et al.* (1968) have numerically traced out some families of periodic orbits. They found that for some periods there are two periodic orbits, whereas the usual perturbation theory (see, e.g., Message 1966) predicts only one. The appearance of this new family then completes their theoretical explanation. Figure 6 compares the analytic periodic orbits to those found numerically by Colombo *et al.*

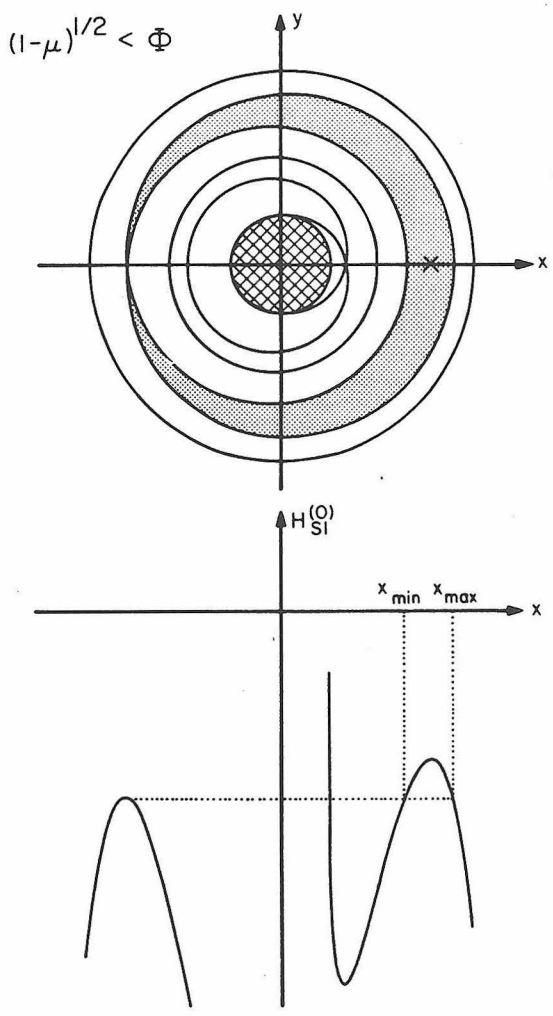


FIG. 5. The same as Figs. 2 through 4, but for $(1 - \mu)^{1/2} < \Phi$. x_{min} and x_{max} are defined as before. There is now a new saddle point with $x > 0$. Near the origin the zero-order resonance Hamiltonian does not represent the motion accurately, so this region has been hatched.

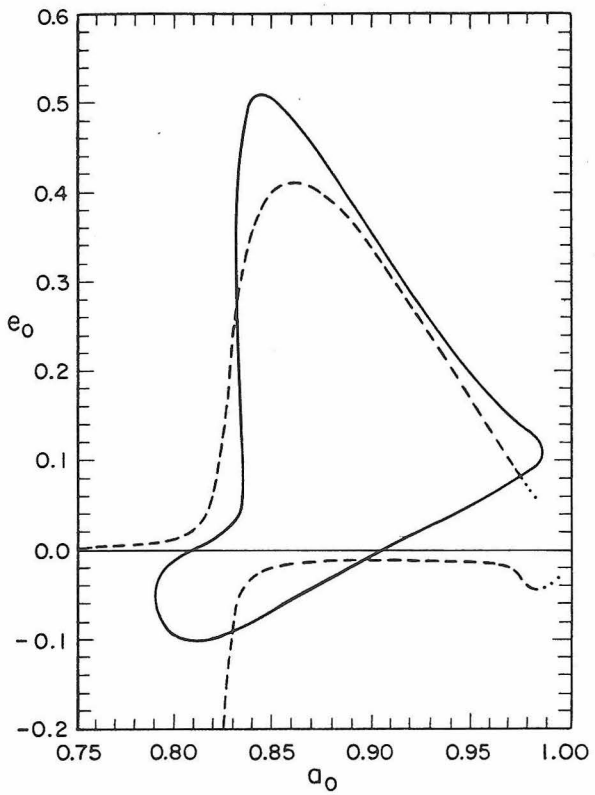


FIG. 6. A comparison of the analytic periodic orbits, including those of the new saddle point, to the periodic orbits found numerically by Colombo *et al.* a_0 and e_0 are the semimajor axis and eccentricity, respectively, at periapse on the line of inferior conjunction.

Rather than study the libration regions in the entire four-dimensional phase space of initial conditions, I restrict my attention to the initial angles $l_0 = 0$ and $g_0 = 0$, i.e., I study the motion of test particles started at periapse on the line of inferior conjunction. It is clearly possible, though, to repeat my analysis for any choice of initial angles. The initial angles $l_0 = g_0 = 0$ are especially easy to analyze since the initial resonance angle ψ_0 is then zero for all resonances. Given particular values of a_0 and e_0, one can calculate Φ_0 and Ψ_0 through Eqs. (28). The test for resonant oscillation when $\psi_0 = 0$ may then be summarized as follows: If $\Phi_0 < \Phi_2$, then ψ oscillates if $\Psi_0 < \Psi_{max}(\Phi_0)$; if $\Phi_0 > \Phi_2$, then ψ oscillates if $\Psi_{min}(\Phi_0) < \Psi_0 < \Psi_{max}(\Phi_0)$. I have used this test to solve numerically for the separatrices in the cases $\mu = 10^{-3}$, 10^{-4}, and 10^{-5}. I kept only the $i = 1$ term in the resonance Hamiltonian (35) and included in it all terms through cubic in the eccentricity (see Brouwer and Clemence 1961). The separatrices of the $s = 3$ and $s = 4$ resonances for $\mu = 10^{-3}$ are plotted in Fig. 7. The combination $1/(a_0^{-3/2} - 1)$ is equal to s when $a_0 = a_{s1}$. The two separatrices for each resonance have been

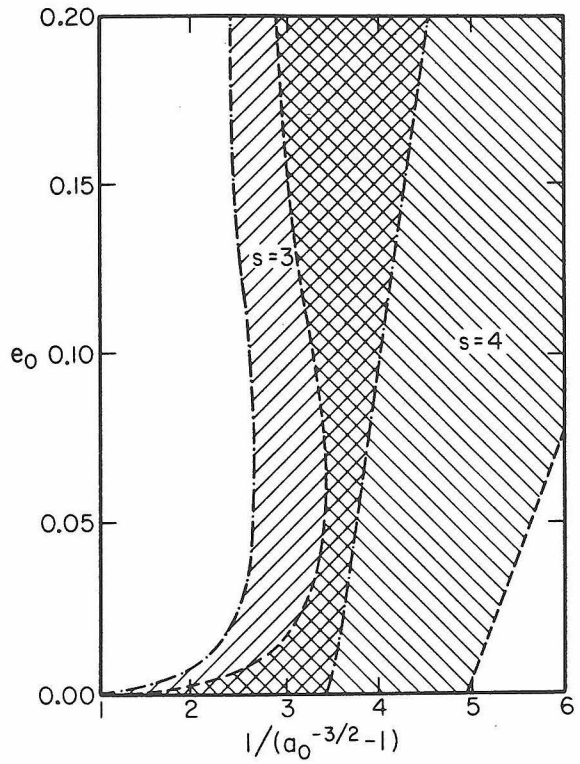


FIG. 7. The oscillation regions of the $s = 3$ and $s = 4$ resonances when $\mu = 10^{-3}$, illustrating the overlap of two resonances. The combination $1/(a_0^{-3/2} - 1)$ is equal to s when $a_0 = a_{s1}$.

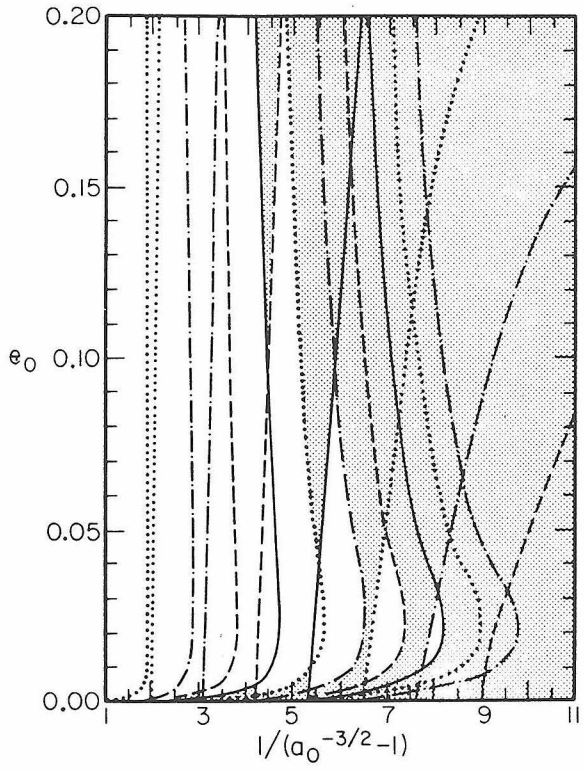


FIG. 9. The same as Fig. 8, but for $\mu = 10^{-4}$.

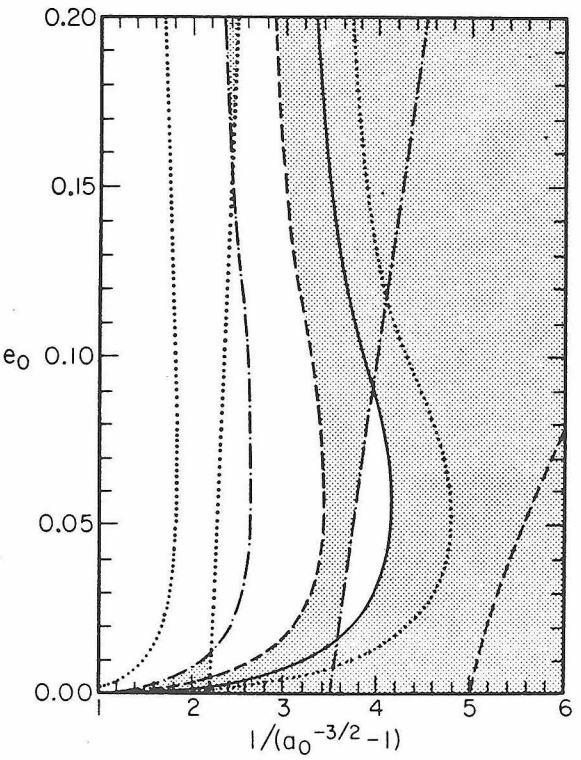


FIG. 8. The oscillation regions for $\mu = 10^{-3}$. The two boundaries of each region are drawn in the same line style. The regions in which two or more resonances overlap are shaded.

plotted in the same line style and the oscillation regions are labeled. The overlap of these two resonance oscillation regions is obvious. Figures 8–10 then show the complete diagrams for the secondary masses $\mu = 10^{-3}$, 10^{-4}, and 10^{-5}. While each oscillation region can be found by looking for two nearby lines with the same line style, for clarity only the overlap regions have been shaded. The resonance overlap criterion then predicts that initial conditions chosen from the unshaded regions will lead to quasiperiodic motion and those chosen from the shaded regions will lead to motion with a random character. In Sec. IV, I compare these predictions to some numerical experiments.

IV. EXPONENTIAL SEPARATION AND THE KOMOGOROV-SINAI ENTROPY

There are two numerical tools to determine whether or not motion is quasiperiodic. The most intuitive is the Poincaré surface of section (see Hénon and Heiles 1964). In this method a two-dimensional surface is chosen in the four-dimensional phase space. The equations of motion are then numerically integrated and each crossing of the surface is recorded. If the motion is quasiperiodic, there are two constants of the motion that constrain these crossings to lie on a "simple" curve. If the motion is not quasiperiodic, it is free to roam over some area of the surface. This was the method used by Hénon (1966), Bozis (1966), and Jefferys (1966) in their studies of the

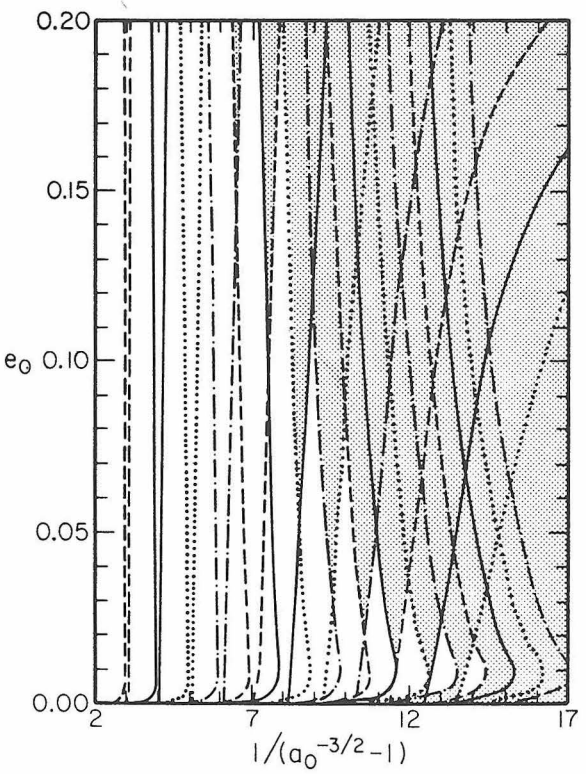


FIG. 10. The same as Figs. 8 and 9, but for $\mu = 10^{-5}$.

restricted three-body problem. This method suffers from the criticism that there is no way to tell whether the crossing points are "random" (i.e., unconstrained) or whether the calculation of more points would reveal that they all lie on a "simple" curve. A more quantitative method, requiring far less computer time, studies the separation of initially nearby orbits with the same value of the (time-independent) Hamiltonian. It has been found (see Chirikov 1979) that in phase space such orbits separate exponentially in the "stochastic" regions and approximately linearly when the motion is quasiperiodic. This is the method I used to study the restricted three-body problem. I examined the separation of nearby orbits for the same secondary masses, $\mu = 10^{-3}$, 10^{-4}, and 10^{-5}, as were studied in Sec. III. I chose the initial eccentricities $e_0 = 0.05$ and $e_0 = 0.10$. Initial semimajor axes a_0 were chosen to span the ranges studied in Sec. III. Of course l_0 and g_0 are zero in all cases. If we consider the system in a rotating frame of reference in which the two masses are stationary, the test particle is started at inferior conjunction with a velocity perpendicular to the line of conjunction. Its partner is also started at inferior conjunction but 10^{-7} closer to the larger mass, with a velocity perpendicular to the line of conjunction chosen so that the values of the Hamiltonian for the two particles are the same. The equations of motion in rotating (synodic) Cartesian coordinates (see Brouwer and Clemence 1961) were then numerically integrated using the algorithm of Bulirasch and Stoer (1966). As is customary for numerical integrations of the restricted problem, the accuracy of the solution was monitored by the variation of the Hamiltonian. In all cases presented here the Hamiltonian never varied by more than $\sim 10^{-10}$ from its initial value and usually varied by only $\sim 10^{-11}$ or 10^{-12}. Figure 11 presents a typical example of what the phase space separation as a function of time looks like in a quasiperiodic regime [$\mu = 10^{-4}$, $e_0 = 0.05$, $1/(a_0^{-3/2} - 1) = 6.5$]. Figure 12 is typical of the exponential separation in a "stochastic" region [$\mu = 10^{-4}$, $e_0 = 0.05$, $1/(a_0^{-3/2} - 1) = 7$]. Note the logarithmic scale in Fig. 12. The rate of divergence can be quantified by fitting in a least-squares sense, the form

$$d(t) = d_0 \exp(h_2 t) \tag{39}$$

to the phase space separation as a function of time, with d_0 equal to the initial separation. This leads to

$$h_2 = \frac{\sum_i t_i \ln(d(t_1)/d_0)}{\sum_i t_i^2} . \tag{40}$$

Every point calculated in each numerical integration up to $t = t_{\max}$ is included in the corresponding sum. The following values of $t_{\max}$ were used: for $\mu = 10^{-3}$, $t_{\max} = 200$; for $\mu = 10^{-4}$, $t_{\max} = 250$; and for $\mu = 10^{-5}$, $t_{\max} = 300$. In an exponential regime, h_2 should be almost independent of $t_{\max}$ (until the separation of the two particles is of order 1), whereas in a linear regime h_2 should decrease approximately as $\ln(t_{\max})/t_{\max}$. The quantity

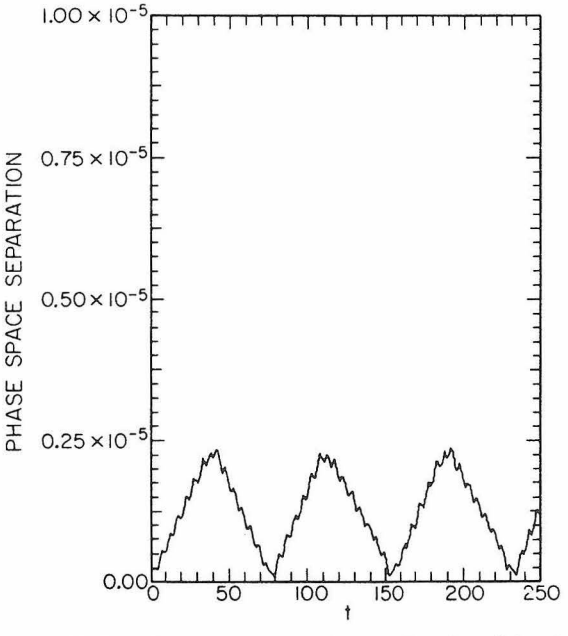


FIG. 11. Typical phase space separation as a function of time in a quasiperiodic regime [$\mu = 10^{-4}$, $e_0 = 0.05$, $1/(a_0^{-3/2} - 1) = 6.5$]. Note the small linear scale on the ordinate.

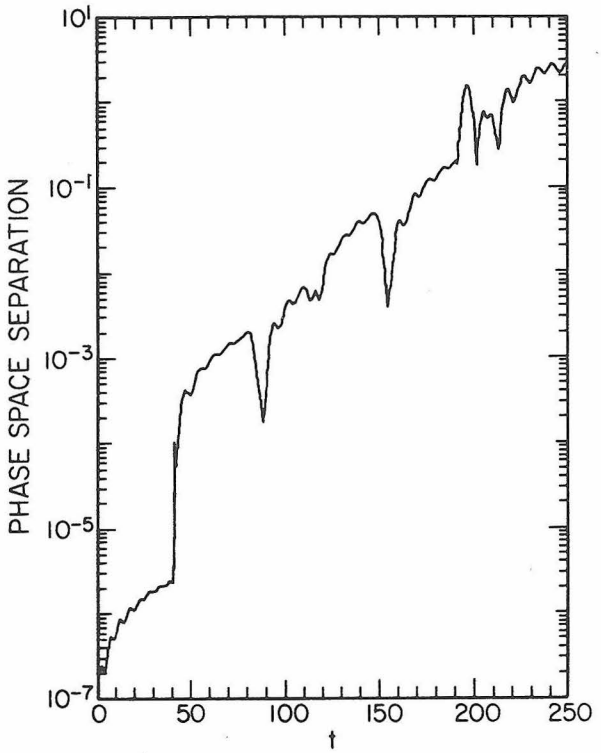


FIG. 12. Typical phase space separation as a function of time in a "stochastic" regime [$\mu = 10^{-4}$, $e_0 = 0.05$, $1/(a_0^{-3/2} - 1) = 7.0$]. Note the logarithmic scale on the ordinate.

h_2 is quite closely related to the Kolmogorov-Sinai entropy h, which is defined by (see Chirikov 1979)

$$h \equiv \left\langle \lim_{d \to 0} \frac{d}{dt} (\ln d) \right\rangle, \tag{41}$$

where the brackets denote an average over the trajectory and $d(t)$ is assumed to be infinitesimal. Figures 13–15 present the results of these calculations. Though the scatter in these plots is fairly large, there is quite clearly a critical value of a_0 in each such that for $a_0 \geq a_{\text{critical}}$, h_2 increases sharply. Figure 16 compares these critical values of a_0 to the predictions of Sec. III. The bars mark the predicted locations of the stochastic instabilities as given by the resonance overlap criterion. The left edge of each bar is the point at which overlap first occurs (for increasing a_0) and the right edge is the point beyond which there is only overlap. In each of the three cases studied, the observed instability occurs within the predicted region. Thus the resonance overlap criterion seems to work very well in the restricted three-body problem. In Sec. V, I derive an approximate criterion for resonance overlap which is valid for all $\mu \ll 1$.

V. SCALING LAW

Though the specific approximations leading to the Chirikov criterion are not valid in the restricted three-body problem, the intuition gained in the numerical study of the separatrices allows us to derive, in the spirit of Chirikov, an approximate criterion for resonance overlap. First, it is clear that no unique widths in semimajor axis can be assigned. For small eccentricity ($e_0 \leq \mu^{1/2}$) the oscillation regions are quite broad and there is overlap at most semimajor axes. For eccentricities greater than $\sim\mu^{1/2}$, the oscillation regions are somewhat more localized in semimajor axis, yet still not uniquely defined. We can characterize the half-widths, though, as the separation of the rightmost ($a_0 > a_{s1}$) boundaries at $e_0 = 0$ from a_{s1}. I call this semimajor axis a_2 since $\Phi(a_2, e_0 = 0) = \Phi_2$. I turn then to the calculation of an approximate expression for Φ_2. Expanding the resonance Hamiltonian (35) about $\Psi = 0$ and retaining only the quadratic terms, one gets

$$H_{s1}^{(0)} \simeq \left(-\frac{1}{2\Phi^2} - \Phi\right) + \left(\frac{-s}{\Phi^3} + (s+1)\right)\Psi + \frac{1}{2}\left(\frac{-3s^2}{\Phi^4}\right)\Psi^2 - \mu K_1 \cos\psi. \tag{42}$$

I have used the fact that for small Ψ (small eccentricity) the sum in Eq. (35) is well approximated by the single

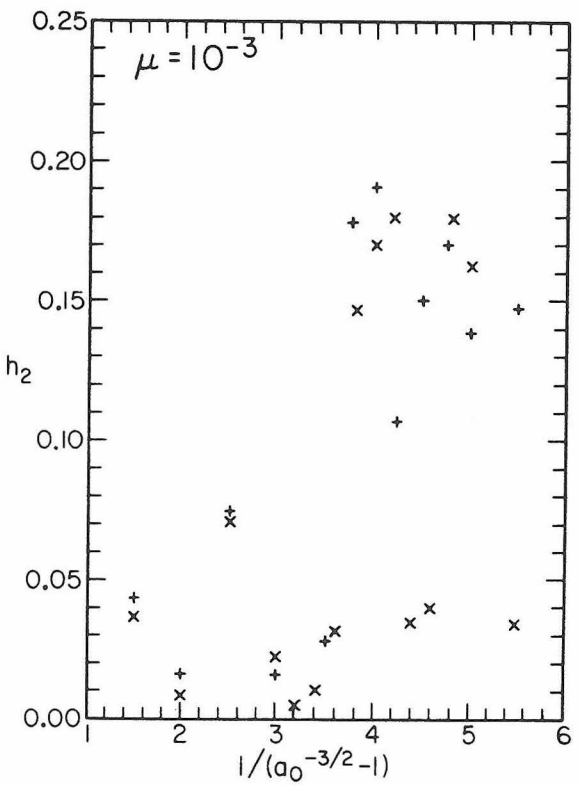


FIG. 13. The entropy h_2 for various initial conditions, with $\mu = 10^{-3}$. Points with initial eccentricity $e_0 = 0.05$ are marked by a cross (×), while points with initial eccentricity $e_0 = 0.10$ are marked by a plus (+).

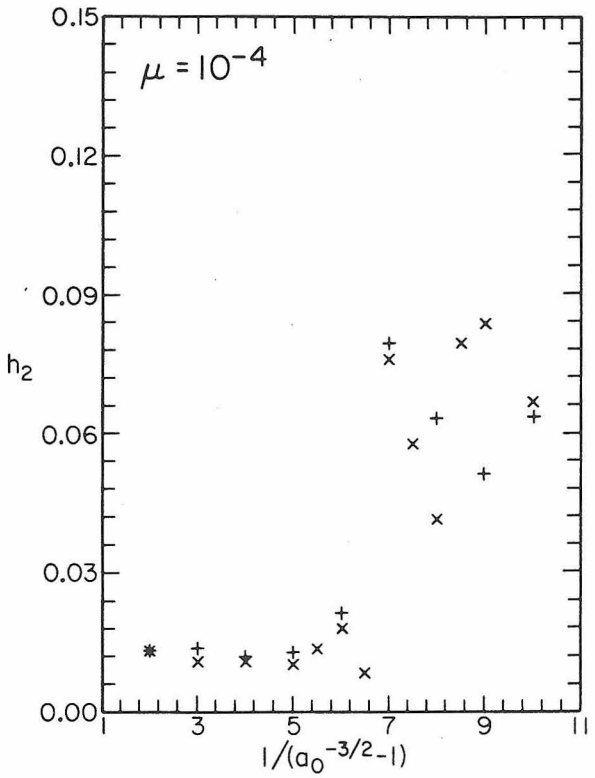


FIG. 14. The same as Fig. 13, but for $\mu = 10^{-4}$.

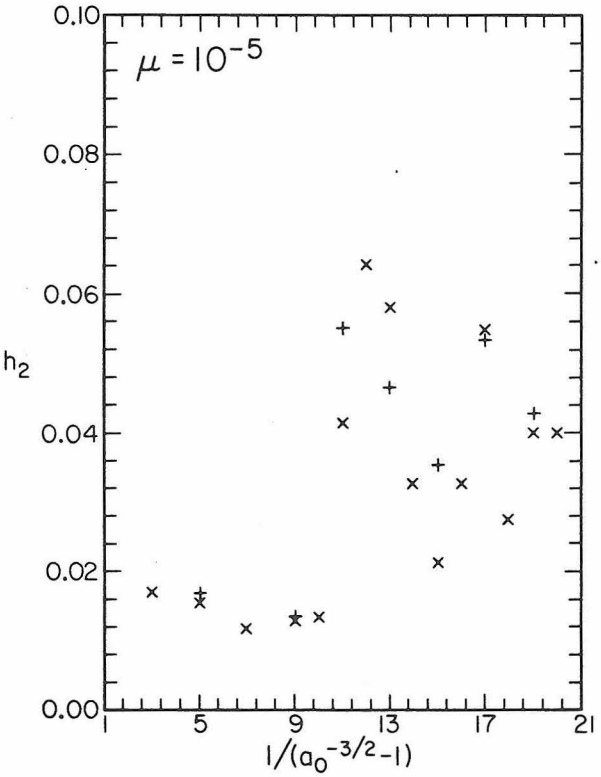


FIG. 15. The same as Figs. 13 and 14, but for $\mu = 10^{-5}$.

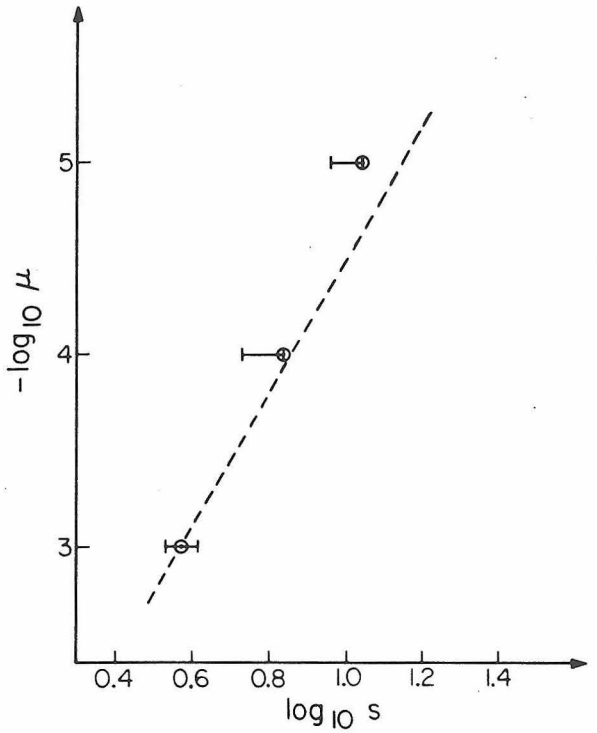


FIG. 16. A comparison of the predictions of the resonance overlap criterion (denoted by horizontal bars) with the location of the observed instability (denoted by a dot with a circle). The dashed line is the approximate theory of Sec. V.

term $i = 1$. Further, for sufficiently small Ψ and $\Phi < (1 - \mu)^{1/2}$, K_1 is well approximated by

$$K_1 \simeq (2\Psi)^{1/2} B_s(\Phi, \Psi = 0). \tag{43}$$

Changing to the Cartesian variables $x = (2\Psi)^{1/2} \cos\psi$ and $y = (2\Psi)^{1/2} \sin\psi$, the approximate Hamiltonian (42) becomes, for $y = 0$,

$$H_{s1}^{(0)}(y = 0) \simeq \left(-\frac{1}{2\Phi^2} - \Phi\right) + \left(\frac{-s}{\Phi^3} + (s+1)\right)\frac{x^2}{2} + \frac{1}{2}\left(\frac{-3s^2}{\Phi^4}\right)\frac{x^4}{4} - \mu x B_s(\Phi, \Psi = 0). \tag{44}$$

The extrema of $H_{s1}^{(0)}(y = 0)$ in this approximation can be found by solving the cubic equation

$$0 = \frac{\partial H_{s1}^{(0)}(y = 0)}{\partial x} = \frac{1}{2}\left(\frac{-3s^2}{\Phi^4}\right) x^3 + \left(\frac{-s}{\Phi^3} + (s+1)\right) x - \mu B_s(\Phi, \Psi = 0). \tag{45}$$

First note that for $\Phi < \Phi_1$ there is only one root; thus this Φ_1 approximates the Φ_1 of Sec. III. If we define

$$\delta\Phi_1 \equiv \Phi_1/\Phi_s - 1 \tag{46}$$

and ignore terms of order $\delta\Phi_1$ compared to unity, it is easy to show that for large s,

$$\delta\Phi_1 \simeq (|\mu B_s|)^{2/3} (3/8s)^{1/3}. \tag{47}$$

For a_0 near unity, B_s is well approximated by

$$B_s(a_0) \simeq \frac{-(s+1)}{\pi}\{2K_0[(s+1)(1-a_0)] + K_1[(s+1)(1-a_0)]\}, \quad (48)$$

where K_0 and K_1 are the modified Bessel functions. This expression is evaluated at $a_0 = a_{s1} \simeq 1 - 2/[3(s+1)]$. Equation (47) is then

$$\delta\Phi_1 \simeq 0.62\mu^{2/3} s^{1/3}. \quad (49)$$

This result justifies the statement in Sec. III that Φ_1 is near Φ_s. Φ_2 is determined by the condition

$$H_{s1}^{(0)}(\Phi_2, \Psi_{sp}, \psi = \pi) = H_{s1}^{(0)}(\Phi_2, \Psi = 0), \quad (50)$$

i.e.,

$$0 = \frac{1}{2}\left(\frac{-3s^2}{\Phi_s^4}\right)\frac{x^3}{4} + \frac{3}{2}(s+1)\,\delta\Phi_2\frac{x}{2} - \mu B_s(\Phi_s, \Psi = 0). \quad (51)$$

This equation, along with Eq. (45), allows us to solve for the two unknowns:

$$x_{sp}(\Phi_2) \simeq -1.02\mu^{1/3}s^{-1/3},$$
$$\delta\Phi_2 \simeq 0.79\mu^{2/3}s^{1/3}, \quad (52)$$

where $\Phi_2 = \Phi_s(1 + \delta\Phi_2)$. From the expressions (28) for the Poincaré momenta in terms of the osculating elliptic elements, we see that

$$\Phi \simeq (a_0)^{1/2}(1 + sx^2/2). \quad (53)$$

The half-width of a resonance is then characterized by

$$\delta a_2 \equiv a_2/a_s - 1 \simeq 2\delta\Phi_2. \quad (54)$$

Equating twice this width to the separation of resonances in a_0, which is approximately

$$\Delta a \simeq 2/3s^2, \quad (55)$$

I derive an estimate of when resonances should begin to overlap:

$$s_{\text{overlap}} \simeq 0.51\mu^{-2/7}. \quad (56)$$

Equation (56) is plotted in Fig. 13. Since the width of a resonance (in a_0) increases as e_0 increases, Eq. (56) is expected to overestimate s_{overlap}.

VI. CONCLUSION

A histogram of the number of asteroids versus semimajor axis (see, for example, Froeschlé and Scholl 1979) reveals a precipitous drop in the number of asteroids with semimajor axes outside the 2/1 resonance. This fact led Lecar and Franklin (1973) to hypothesize a dynamical origin for the absence of asteroids in this region. To test their hypothesis they integrated 260 test objects distributed uniformly between $0.55a_J$ and $0.85a_J$, where a_J is Jupiter's semimajor axis, with eccentricities between 0.0 and 0.3. They found that objects started exterior to $0.85a_J$ were "ejected" immediately and that within the 200 Jupiter revolution time span of their integrations, the region outside the 3/2 resonance was cleared, except for some objects at the 4/3 resonance. The region between the 2/1 resonance and the 3/2 resonance, however, remained well populated. They suggested that longer integrations might deplete this region. To test this hypothesis Froeschlé and Scholl (1979) performed a similar experiment covering a time span of 10^5 yr. The region was still not sufficiently depleted. In addition they found that after 60 000 yr no more objects escaped. Since the secular perturbations of the planets cause variations in the orbital elements with time scales of order 10^5 yr or longer (see Brouwer and van Woerkem 1950), it may be necessary to extend these numerical experiments to several million years before they capture all the dynamical features that are present. The extension of these calculations to much longer times, though, appears to be prohibitively expensive. Even if their numerical experiments had depleted the region outside the 2/1 resonance, we would have wanted a qualitative understanding of those dynamical features which led to the instability. The failure of the numerical experiments and the cost of extending them to the required time span heightens the need for a qualitative understanding of the instabilities in asteroidal motion. This paper constitutes a first step towards this qualitative understanding. I have applied the resonance overlap criterion to the planar circular-restricted three-body problem and compared its predictions to some numerical experiments. Since the predictions are in remarkably good agreement with my numerical experiments, great confidence has been gained in the usefulness of the resonance overlap criterion for obtaining a qualitative understanding of the instabilities in the solar system.

It is a pleasure to thank Peter Goldreich for valuable advice and helpful criticism. This work was partially supported by NASA Grant NGL 05-002-003.

REFERENCES

Bozis, G. (1966). Astron. J. **71**, 404.
Brouwer, D., and van Woerkom, A. J. J. (1950). Astron. Papers of the American Eph. **13**, 81.
Brouwer, D., and Clemence, G. M. (1961). *Methods of Celestial Mechanics* (Academic, New York).
Bulirasch, R., and Stoer, J. (1966). Numerische Math. **8**, 1.
Chirikov, B. V. (1979). Phys. Rep. **52**, 263.
Colombo, G., Franklin, F. A., and Munford, C. M. (1968). Astron. J. **73**, 111.
Froeschlé, C., and Scholl, H. (1979). Astron. Astrophys. **72**, 246.
Hagihara, Y. (1943). Jpn. J. Astron. Geophys. **21**, 29.
Hénon, M., and Heiles, C. (1964). Astron. J. **69**, 73.
Hénon, M. (1966). *The Theory of Orbits in the Solar System and in Stellar Systems*, IAU Symposium No. 25 (Academic, London),

p. 157.
Jefferys, W. H. (1966). Astron. J. **71**, 306.
Lecar, M., and Franklin, F. (1973). Icarus **73**, 422.
Message, P. J. (1958). Astron. J. **63**, 443.
Message, P. J. (1966). *The Theory of Orbits in the Solar System and in Stellar Systems,* IAU Symposium No. 25 (Academic, London), p. 197.
Moser, J. (1973). *Stable and Random Motions in Dynamical Systems* (Princeton University, Princeton, N.J.).
Poincaré, H. (1902). Bull. Astron. **19**, 289.
Schubart, J. (1964). Smithsonian Astrophys. Obs. Spec. Rep. No. 149.
Walker, G. H., and Ford, J. (1969). Phys. Rev. **188**, 416.
Whittaker, E. T. (1961). *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies,* 4th ed. (Cambridge University, Cambridge, England).
Wiesel, W. E. (1976). Celestial Mech. **13**, 3.
Woltjer, J. (1923). Bull. Astron. Inst. Neth. **1**, 219.