ESI The Erwin Schrödinger International
Institute for Mathematical Physics
Boltzmanngasse 9
A-1090 Wien, Austria

The Classical Three-Body Problem – where is Abstract Mathematics, Physical Intuition, Computational Physics Most Powerful?

H. A. Posch
W. Thirring

Vienna, Preprint ESI 843 (2000) February 17, 2000

Supported by Federal Ministry of Science and Transport, Austria
Available via http://www.esi.ac.at

The classical three-body problem – where is abstract mathematics, physical intuition, computational physics most powerful?

H. A. Posch
Institut für Experimentalphysik, Universität Wien, Boltzmanngasse 5, A-1090 Wien, Austria
E-mail: posch@ls.exp.univie.ac.at

W. Thirring
Institut für Theoretische Physik, Universität Wien, Boltzmanngasse 5, A-1090 Wien, Austria
E-mail: fwagner@ap.univie.ac.at

(February 17, 2000)

We propose a simple oscillator model for the reduced three-body problem to understand the stability of orbits with small eccentricity of the light planet. It models the main short-time features for small mass ratios of the other bodies. These results are confronted with the exact mathematical analysis for stability for all times, and with computer simulation results for bigger mass ratios, where chaotic features emerge.

I. INTRODUCTION

The three-body problem is very old (see Reference [1] for a historic review which starts even with the Babylonians) and an immense literature has accumulated over the centuries [2]. How can one think that one can make a new contribution to it? It is not that we possess new observational data, but the computer puts us in a better position than previous generations. Any idea which would have taken years to verify or falsify with a slide rule can now be settled within seconds. Furthermore, unlike astronomers we can change the mass ratios at will to understand the various mechanisms and to see when and why things become chaotic. Of course, a general solution is impossible and would also be too complicated to be of any use. So we concentrate on some limited but relevant questions mainly on the restricted three-body problem [3], where one body is so light that it does not influence the (circular) motion of the two others. The answers to these questions require different tools and we shall formulate them such that they make use of physical intuition, rigorous analysis and computational methods.

Question 1. Even if the second body is much lighter than the heaviest one, its influence on the third is much less than a naive estimate would tell us. For

instance, $M_{\text{Jupiter}}/M_\odot \sim 1/1000$, but without sun it would take Mars at rest only about 200 years to fall freely into Jupiter. But its near-Kepler orbit is stable for a much longer time, merely its eccentricity is about twice that of the Earth. What is exactly the mechanism which stabilizes the orbit?

Answer 1. The radial motion of nearly circular orbits is like a harmonic oscillator, and the influence of Jupiter is like periodic kicks (better pulls). From the kicked oscillator one knows that the amplitude of the induced oscillations gets damped again if one is not at a resonance, and the kicks get out of phase. We shall underpin this by an elementary calculation and illustrate it by computer simulations below. If resonance conditions apply, the amplitude increases linearly with time, but then one gets into the nonlinear domain and out of phase with the kicks. Whether this comes in time to quench the oscillations or whether the situation is already out of hand depends on the strength of the kicks, i.e. M_J.

Question 2. In general, for which initial conditions can one guarantee stability *ad aeternitatem*?

Answer 2. Since the orbits can become so complex, this question cannot be settled by naive models and computers cannot calculate to $t = \infty$. So this is the domain of mathematical proofs. Generally there are plenty of even periodic orbits, but the question is whether the stable sets - apart from that determined by the Jacobi constant - have finite measure or are even open. For a sweeping proof one has to be prepared for the worst situation, and any rational frequency ratio is a possible resonance. Though one can show that for small perturbations there are regions of finite measure (not open) which are stable, one had to cut out (perhaps unnecessarily) so many pieces in phase space that for the system sun + Jupiter + small planet one is still far away from a proof of stability of sets of finite measure for a mass ratio $M_J/M_\odot \sim 10^{-3}$.

Question 3. One has learned at school that if there is no other constant than the Hamiltonian, the system becomes ergodic. Computer studies show that for confining potentials $|x_i - x_j|^\nu$, $\nu > 0$, the orbits for several particles seem ergodic on the energy–angular momentum shell [4]. Is this still true here?

Answer 3. According to 2, for small perturbations this is not the case. But only the computer can give a hint how strong the perturbation has to be for ergodicity. (1) gives a clue for the mechanism of instability. If the kicks are too strong so that the planet will spill over and come near the sun or Jupiter before the quenching becomes effective it will be completely thrown out of its orbit and there is no stabilizing mechanism any more. A simple estimate shows that this happens for $M_J/M_\odot > 1/100$, and then the computer shows that there are large chaotic regions but they contain islands of regularity. They shrink with

increasing $M_J/M_\odot$ and look rather weird, not like a submanifold given by another constant of motion $K(x, y, p_x, p_y) = \text{const}$. Sometimes they are connected by a small bottleneck with other parts of the energy shell and the orbit fails to find the hole in a reasonable time.

The impression one gets from these considerations is that our solar system must be very cleverly constructed to be stable over such a long time [5,6]. Extensive computer-aided calculations [7] show that the Liapunov time in the planetary system is of the order of 10^7 years, much shorter than its age, destroying the hope of a general stability proof for 4×10^9 years. Jupiter is not too heavy but far enough from the sun to carry most of the angular momentum. This stabilizes its plane of motion, otherwise the inclination of the orbits would be random. Furthermore, all planetary orbits are nearly circular, and the two groups of outer and inner planets are fairly evenly spaced. Presumably, in the early solar system there were many more planets, but their orbits did not comply with the above stability specifications, so they collided, fell into the sun or were thrown out of the solar system. In the newly-discovered planetary systems [8,9], where the heaviest planet has about 1/10 of the mass of the central star the orbits of the other unseen planets must be so chaotic that they cannot provide a sufficiently well-tempered climate for life to exist.

II. INTUITIVE ARGUMENT

We consider here the situation where the two heavy bodies ("sun and Jupiter") make a circular orbit, and the third (the "planet") has a negligible mass (restricted 3-body problem). Furthermore, all move in the same plane. For the planet's motion the configuration space is 2-dimensional, the phase space is 4-dimensional, and there is one constant of the motion, the Hamiltonian in the rotating system (equivalent to the "Jacobi constant"). In those parts of phase space where the planet cannot escape, no other constant is known and we have the simplest situation of a non-integrable system. We shall start with an almost circular orbit of the planet, because in our solar system most eccentricities are small and these orbits are apparently the most stable ones. Without Jupiter the effective radial potential is

$$V_{\text{eff}}(r) = -\frac{1}{r} + \frac{L^2}{2r^2},$$

and the circular orbit is in the minimum of this potential. Throughout, we use reduced units for which the sum of the masses $M_\odot + M_J$ of the primaries, the sun - Jupiter distance, and the angular velocity of Jupiter are unity. r is the distance of the light planet from the sun, and L is its angular momentum. The potential is depicted in Fig. 1. Now we shall naively guess what the effect of Jupiter might

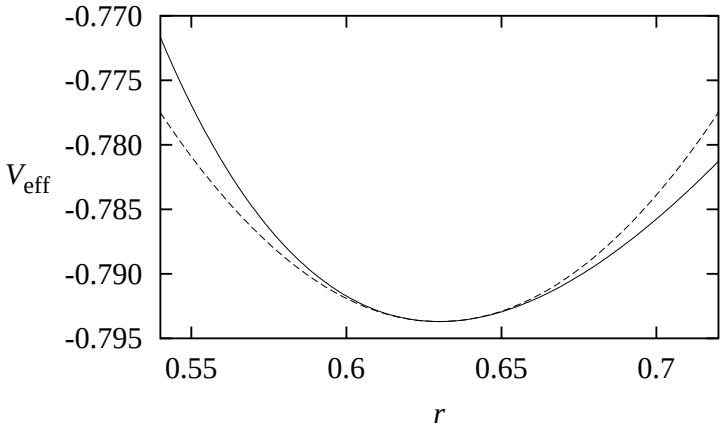

FIG. 1. Effective potential $V_{\rm eff}$ for the 2:1 resonance, $r_0 = 2^{-2/3}$. The dashed curve is the harmonic approximation for the kicked-oscillator model.

be on an orbit inside its circle. We are interested in mass ratios $M_J/M_\odot$ between $10{-}1$ and $10{-}3$, so Jupiter should not immediately throw the planet out of orbit. Since the force is $\sim M_J|x - x_J|^{-2}$ it should be most noticeable when the planet is on Jupiter's side of the sun and Jupiter pulls the planet outward of the minimum of $V_{\rm eff}$. Of course, there will also be an azimuthal force, but this will be first accelerating and then decelerating, so we think it will largely average out and forget about it. About the force f of Jupiter, we only assume that it is periodic with a period $\tau = 2\pi/(\omega - 1)$, where ω is the unpertubed angular velocity of the planet, that of Jupiter being unity in our units. τ is the time between successive conjunctions of Jupiter and the planet. Though the orbit of Jupiter is strictly periodic, the one of the planet is not, so $f(t) = f(t + \tau)$ is not quite correct. But we think it is a good approximation. Thus, if we concentrate on the radial motion of the planet, the complex coordinate $z = p_r + i\omega(r - r_0)$, $V'_{\rm eff}(r_0) = 0$, obeys

$$\dot{z}(t) = i\omega z(t) + f(t), \qquad V''_{\rm eff}(r_0) = \frac{\omega^2}{2} \tag{1}$$

near the minimum r_0. In the solution

$$z(t) = e^{i\omega t} z(0) + \int_0^t dt'\, e^{i\omega(t-t')} f(t') \tag{2}$$

the two terms have spectra $\{\omega\}$ and $\{\omega\} \cup (\omega - 1)Z$, respectively. In particular,

$$\int_0^\tau dt' f(t') e^{i\omega(\tau - t')} =: e^{i\omega\tau} K$$

shows that for all times the change of z during a period τ,

$$z(\tau) = e^{i\omega\tau}(z(0) + K), \tag{3}$$

depends only on $\omega = r_0^{-3/2}$ and the constant K. Since the detailed form of $f(t)$ does not enter, this gives us confidence that (3) might be a good guess, and we iterate it to the symplectic map

$$z(n\tau) = e^{in\omega\tau}\left(z(0) + K\frac{1 - e^{-in\omega\tau}}{1 - e^{-i\omega\tau}}\right), \qquad n \in Z. \tag{4}$$

To get an idea of the planetary motion, we have in Fig. 2 replaced the effect of Jupiter by periodic kicks, $f(t) = K\sum_n \delta(t - n\tau)$, where, in a generous mood, we have computed K as half of the total accumulated force of a planet passing Jupiter on a straight line with the correct minimal distance $1 - r_0$ and relative velocity $v = 1/\sqrt{r_0} - 1$,

$$K = \frac{M_J}{2}\int_{-\infty}^{\infty} \frac{dt(1 - r_0)}{[(1 - r_0)^2 + (vt)^2]^{3/2}} = M_J\frac{\sqrt{r_0}}{(1 - r_0)(1 - \sqrt{r_0})}. \tag{5}$$

We do not insist on this hair-raisingly crude approximation, but to our surprise it worked rather well as will be shown below.

What we learn from (4) is that the periodic pull of Jupiter excites radial oscillations of the planet, but unless there is a resonance, $\omega\tau = 2\pi g$, $g \in Z$, for which the denominator in (4) vanishes, these oscillations eventually get out of phase with the period of the pull. Thus, after some time there will be a "thrust reversal", and the oscillations will be damped again until one comes close to the original configuration. More in detail, the influence of Jupiter will be most noticeable near a resonance $\omega\tau = 2\pi g + \varepsilon$, $g \in Z$, $\varepsilon \ll 1$. For $n \ll 1/\varepsilon$, the relevant factor $(1 - e^{-in\omega\tau})/(1 - e^{-i\omega\tau})$ becomes about $n - i\varepsilon n^2/2$, whereas, for $n \sim 1/\varepsilon$, both p_r and $r - r_0$ become of order K/ε. For $n\varepsilon$ near π we get thrust reversal, and for $n\varepsilon$ near 2π p_r and $r - r_0$ go back to the order of K. Since K is of the order $M_J/M_\odot \sim 10^{-3}$, only a small region near $\omega \in Z$ is dangerous. However, even $\omega\tau = 2\pi g$ might not be catastrophic because the resonances have a built-in selfquenching mechanism. If we start, say, with $\omega = r_0^{-3/2} = 2$, $\tau = 2\pi$, then $r_{\max} = r_0 \max_{2\pi n < t < 2\pi(n+1)} \mathrm{Im}\, z(t)$ will determine the frequency after some time. The harmonic approximation to V_{eff} will break down and ω becomes $r_{\max}^{-3/2} \neq 2$. Hence, we will get thrust reversal and whether this comes in time before $r_{\max} \sim r_0 + nK$ is close to one depends on the strength of K. To follow this analytically by improving our crude model is very tedious and at this stage it is better to consult the computer to see what is going to happen.

For our numerical work in this section the equations of motion are derived from the Hamiltonian in the (synodic) center-of-mass frame rotating with Jupiter,

$$H = \frac{1}{2}(p_x^2 + p_y^2) - xp_y + yp_x - \frac{M_\odot}{[(x - M_J)^2 + y^2]^{1/2}} - \frac{M_J}{[(x + M_\odot)^2 + y^2]^{1/2}}, \tag{6}$$

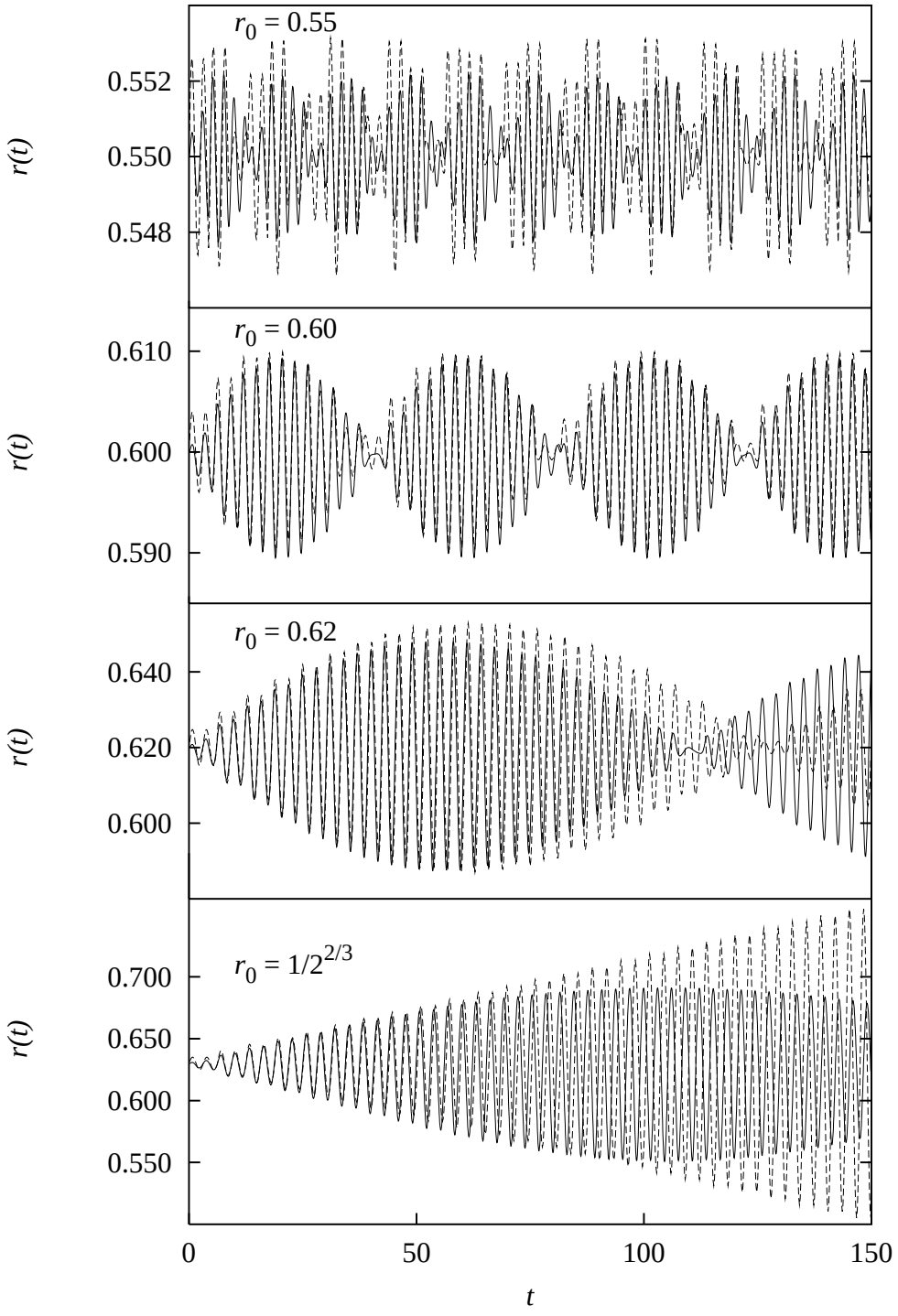

FIG. 2. Radial planetary motion, perturbed by Jupiter, for various unperturbed circular-orbit radii r_0. The mass ratio $M_J/M_\odot = 0.001/0.999$. $r(t)$ denotes the separation from the sun. The smooth lines are the "exact" computer-simulation results, and the dashed lines are for the kicked-oscillator model described in the text. From top to bottom: $r_0 = 0.55$, 0.60, 0.62, and $1/2^{2/3} = 0.62996$.

where the sun and Jupiter are located at $(M_J, 0)$ and $(-M_\odot, 0)$, respectively, and where $M_\odot + M_J = 1$. They are integrated with a variable-step-size Runge-Kutta algorithm of fourth order, keeping the energy constant to 10 significant digits for 30,000 Jupiter periods. Since in this section only slightly perturbed cirular orbits are considered, no regularization of the equations of motion is required [3]. In all cases, the planet is initially located on the x-axis at $x(0) = r_0 - M_J$, with a velocity in y-direction corresponding to the respective unperturbed circular orbit $(M_J = 0)$ with radius r_0.

In Fig. 2 we compare the "exact" simulation results (smooth lines) with the predictions of the kicked-oscillator model (dashed lines) for a perturbed orbit near and at the 2:1 resonance. The mass ratio $M_J/M_\odot = 0.001/0.999$. As before, $r(t)$ denotes the radial distance from the sun. The unperturbed radius r_0 corresponds, from top to bottom, to 0.55, 0.60, 0.62, and $2^{-2/3} = 0.62996$, and is indicated by the labels. According to this model, $r(t)$ oscillates between the kicks occuring at the times $n\tau$, $n = 0, 1, 2, \ldots$, with the unperturbed angular velocity ω and with an amplitude determined from (4). It is surprising that away from the major 2:1 resonance, which occurs at $r_0 = 2^{-2/3}$, this simple model gives a rather good description of the eccentricity of the orbit. Not unexpectedly, the model breaks down at the resonance, for which it predicts an undisturbed linear increase of the amplitude with time, wheras the exact oscillations are damped by self quenching as mentioned above. From the different scales in Fig. 2 we infer that the oscillations are much less pronounced when one moves away from the resonance.

For a given r_0 close to the resonance, the oscillation amplitudes are proportional to $M_J/M_\odot$. This is demonstrated in Fig. 3 for $r_0 = 0.60$, where $M_J/M_\odot$ is varied between $1/999$ and $5/995$. For smaller mass ratios $\leq 2/998$ the kicked oscillator model provides a reasonable description of the "exact" simulation results. It fails for $M_J/M_\odot = 5/995$ due to the dephasing induced by the frequency changes in the nonlinear regime of the effective potential.

To study this phase mismatch between the orbit and the periodic pull in more detail, we show in Fig. 4 the radial oscillations at the 2:1 resonance, $r_0 = 2^{-2/3}$. The perturbed amplitude starts to grow linearly with time, until it reaches the nonlinear regime of the effective radial potential depicted in Fig. 1, and the trajectory gets out of phase with Jupiter. As a consequence, the radial displacement is quenched again and the whole process repeated. This phase mismatch becomes apparent also in Fig. 5, where the time intervals Δ between successive maxima of $r(t)$ in Fig. 4 are plotted at the end of each interval. Δ differs significantly from π, which is the unperturbed period of the planet in this case, equal to half the period of Jupiter. For most of the time, $\Delta < \pi$, and the phase shift accumulates until the force exerted by Jupiter damps the motion again. We also deduce from Fig. 4 that the radial oscillations of this $\omega = 2/1$ resonance are not symmetrical around r_0. The largest amplitudes occur for $r < r_0$, for which the effective potential $V_{\rm eff}$ increases more steeply than for $r > r_0$. This

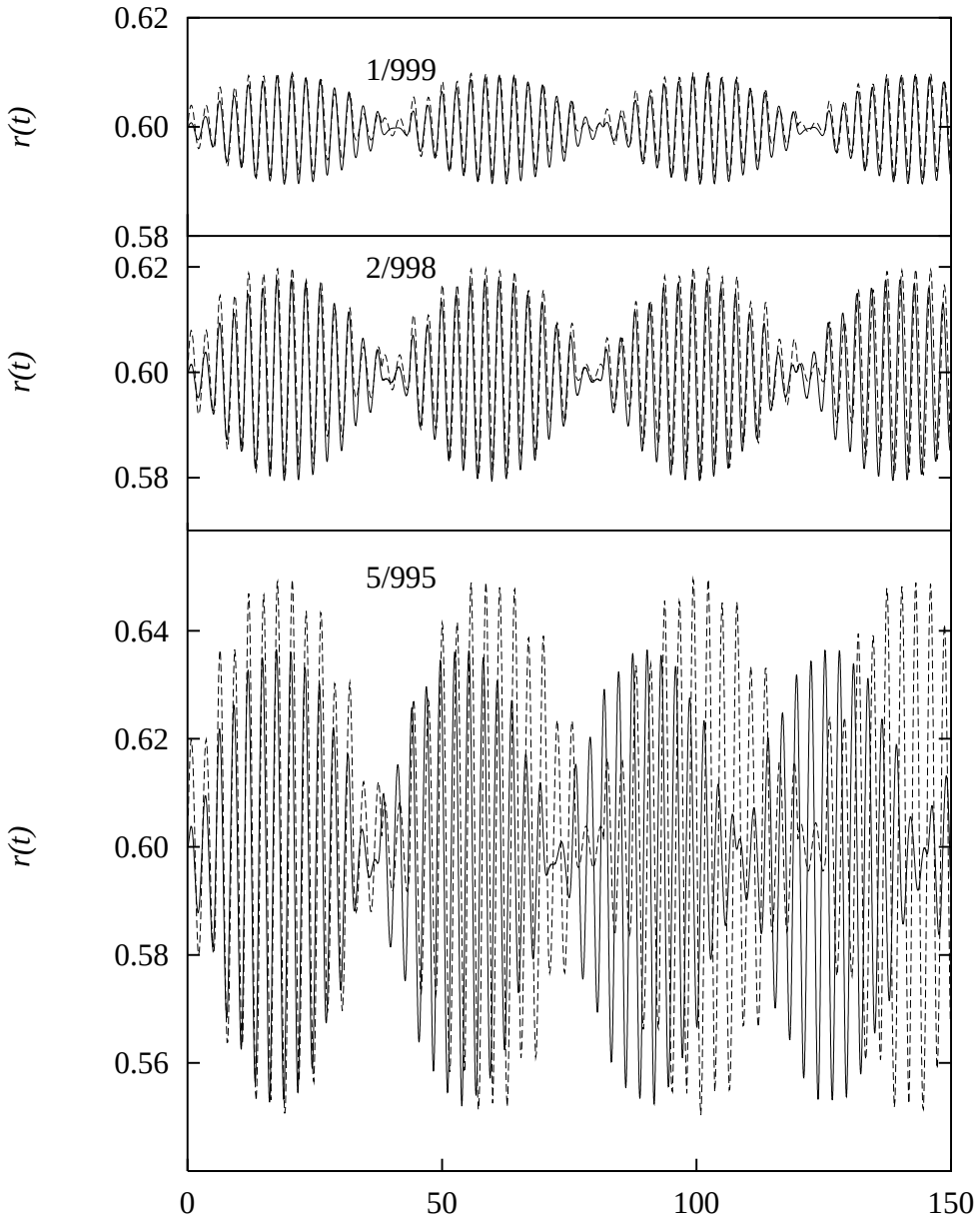

FIG. 3. Radial planetary motion, perturbed by Jupiter, for various mass ratios $M_J/M_\odot$ as indicated by the labels. The unperturbed radius $r_0 = 0.60$.The smooth lines are the "exact" computer-simulation results, and the dashed lines are for the kicked-oscillator model described in the text.

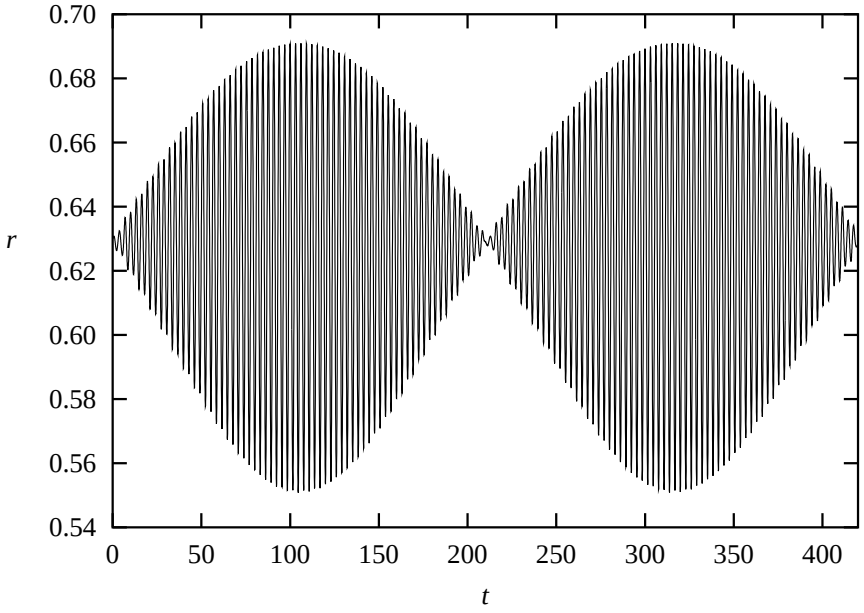

FIG. 4. Radial oscillations of the perturbed planetary orbit for the 2:1 resonance with Jupiter. The mass ratio $M_J/M_\odot = 0.001/0.999$. r is the distance from the sun.

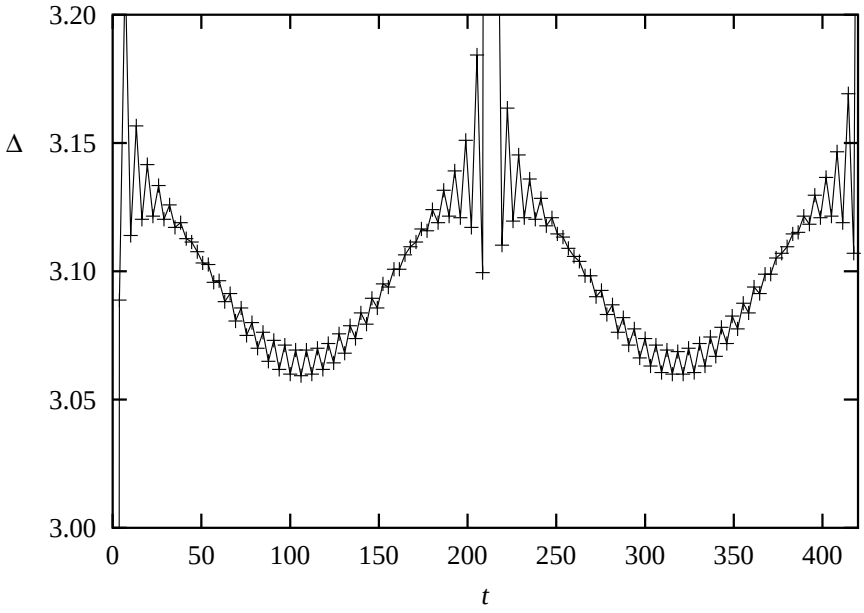

FIG. 5. Time difference Δ between successive maxima for the perturbed orbit shown in Fig. 4, plotted at times t at the end of each interval. The unperturbed planetary period is π.

subtlety cannot be captured by the kicked-oscillator model and severely limits our intuition. A closer look at the exact computer-generated trajectories reveals that the largest amplitudes for $r(t)$ mainly occur in a direction not aligned with Jupiter in our co-rotating frame.

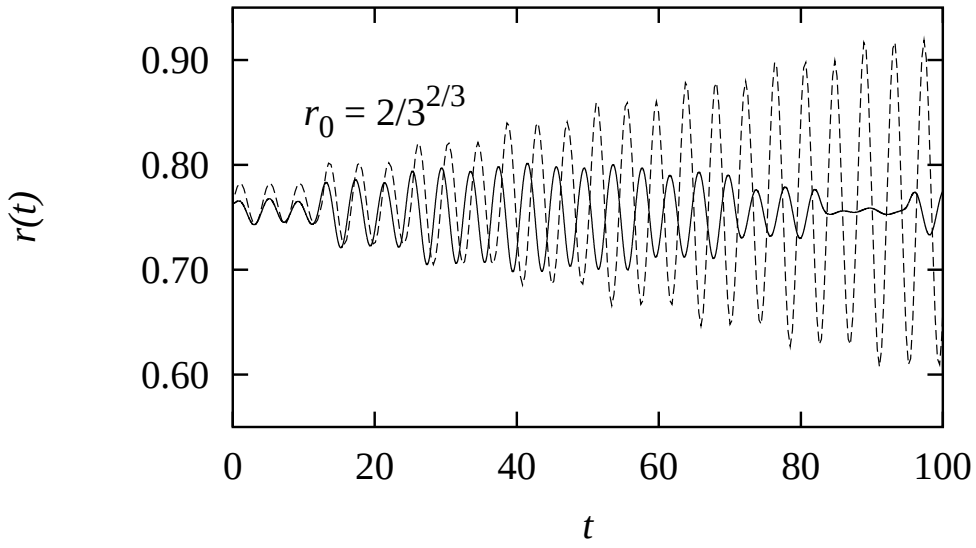

FIG. 6. Radial oscillations of the perturbed planetary orbit for a fractional resonance, $\omega = 3/2$, with Jupiter. The mass ratio $M_J/M_\odot = 0.001/0.999$, and the unperturbed radius $r_0 = (3/2)^{-2/3} \simeq 0.7631$. The smooth lines are the "exact" computer-simulation results, and the dashed lines are for the kicked-oscillator model.

For comparison we show in Fig. 6 also the next fractional resonance for which $\omega = 3/2$, $r_0 = \omega^{-2/3} \simeq 0.7631$, and $\tau = 4\pi$. Now the planet is closer to Jupiter, and K is bigger. The nonlinear regime is reached sooner, and the quenching time is shorter than before. In spite of the rather complicated structure of $r(t)$, the orbit appears to be quasiperiodic with a smooth and ring-shaped Poincaré map.

Unless $\omega\tau = 2\pi g$, the maximum amplitude in (4) is bounded for all n. Interesting phenomena appear for fractional resonances such as the $\omega = 2/5$ resonance of Jupiter - Saturn (see Fig. 7). There, the first conjunction occurs when Saturn is at the angle ϕ_1, $\phi_1/(\phi_1 + 2\pi) = 2/5 \Rightarrow \phi_1 = 4\pi/3$, and the next at $\phi_2 = 8\pi/3$. For $\phi_3 = 12\pi/3 \sim 0$ we are back again. Thus, $\omega\tau = 4\pi/3$, the force f is periodic with period 3τ and the amplitude is periodic in n with period 3.

Also fractional resonances like the Saturn-Jupiter 2:5 resonance depicted in Fig. 7 are not contained in (4). As the figure shows, however, the radial oscillations are small and show an interesting double periodicity. This orbit is not bound by the Jacobi constant (see Fig. 13) to a finite region in configuration

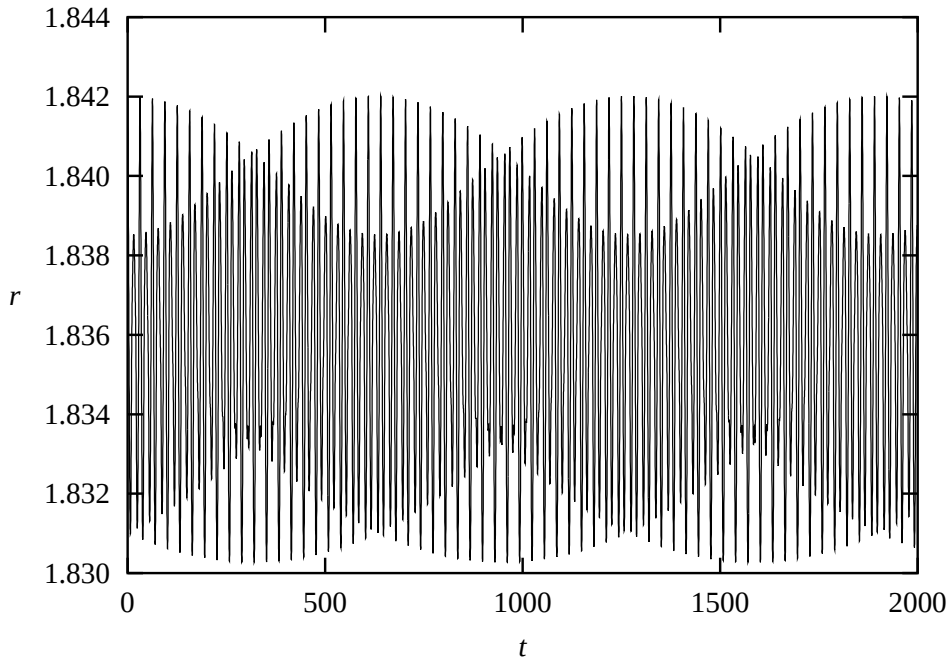

FIG. 7. Radial oscillations of the perturbed planetary orbit for a fractional 2:5 resonance with Jupiter. The mass ratio $M_J/M_\odot = 0.001/0.999$, and the unperturbed radius $r_0 = (2/5)^{-2/3} \simeq 1.8420$.

space. Nevertheless it is stable for a long time due to the action of the Coriolis forces in the rotating frame.

III. RIGOROUS MATHEMATICS

One of the dogmas of classical statistical mechanics is that even if a system is not in equilibrium since in addition to H there are some other constants of the motion a little speck of dust ("Staubkörnchen") will break them and render the system ergodic. Many great scientists tried to prove that, or even thought that they could prove it, but finally light was shed on this question by Kolmogorov, Arnold and Moser (KAM theorem) [14]. What they proved was not that some constants persist for small perturbations but that in regions of phase space with a finite measure the orbit stays on a submanifold homeomorphic to a torus. Thus, for small perturbations the system does not become ergodic. The proof proceeds as follows. If we have an integrable system with action variables I_j and an unperturbed Hamiltonian $H_1(I_j)$ and add $\lambda H_1'(I_j, \varphi_i)$, φ_j the angular variables and λ the perturbation parameter, then by a canonical transformation $I, \varphi \to \bar{I}, \bar{\varphi}$ we try to cast $H = H_1 + \lambda H_1'$ into the form $H_2(\bar{I}_j) + \lambda^2 H_2'(\bar{I}_j, \bar{\varphi}_j)$. If

successful, we iterate the procedure to get $H = H_3 + \lambda^4 H_3'$ and keep on to finally reach, for small λ, $H = H_\infty(I(\infty))$ for which the orbit remains on an invariant torus. Thus we are faced with three problems:

a) Do we succeed in the first step, and if not, why not?

b) Under which conditions do we keep succeeding?

c) Does the procedure converge to an H_∞?

a) Standard perturbation theory proceeds as follows: For the transformation $(I, \varphi) \to (\bar{I}, \bar{\varphi})$ we use a generator $S(\bar{I}, \varphi)$: $I_j = \bar{I}_j + \lambda\, \partial S/\partial\varphi_j$, $\bar{\varphi}_j = \varphi_j + \lambda\, \partial S/\partial \bar{I}_j$ such that for some value of I, say $I_j = 0$, the system remains integrable up to $O(\lambda^2)$. With

$$\omega_j = \left.\frac{\partial H_1}{\partial I_J}\right|_{I=0}, \qquad H_1'(I,\varphi) = \sum_{k \in Z^m} \widetilde{H}_k(I) e^{i(k\cdot\varphi)}$$

we set

$$\omega_j \frac{\partial S}{\partial \varphi_j} + H'(\bar{I}, \varphi) = \widetilde{H}_{k=0}(\bar{I}) \tag{7}$$

then

$$H = H_0(\bar{I}) + \lambda \widetilde{H}_{k=0}(\bar{I}) + \lambda^2 H_2'(\bar{I}, \bar{\varphi}, \lambda) \tag{8}$$

with $H_2'(\bar{I}, \bar{\varphi}, 0) < \infty$ if everything is sufficiently differentiable. (7) is solved in Fourier space by

$$S(\bar{I}, \varphi) = -\sum_{k \neq 0} \frac{e^{i(k\cdot\omega)}}{i(\omega \cdot k)}\, \widetilde{H}_k(I), \tag{9}$$

and we fail if

a_1) $(\omega \cdot k) = 0$ for some $0 \neq k \in Z^m$, or

a_2) $\sum_{k\neq 0}$ diverges.

a_1) means that the ω_j are not linearly independent, $\exists\ 0 \neq k \in Z^m$, $\omega_1 k_1 + \omega_2 k_2 + \ldots + \omega_m k_m = 0$ and we have the resonance situations considered in Section II. Although a term in (9) becomes infinite in this case, this does not mean that in the orbit something becomes infinite. It only means that it cannot be described by (8). To see this more explicitly consider a simplified "Jupiter–Saturn" resonance

$$H = 2I_1 + 5I_2 + \lambda \sin(5\varphi_1 - 2\varphi_2): \tag{10}$$

$$\dot{\varphi}_1 = 2, \qquad \dot{\varphi}_2 = 5, \qquad \dot{I}_1 = 5\lambda \cos(5\varphi_1 - 2\varphi_2) = -\frac{5}{2}\dot{I}_2 \Longrightarrow$$

$$\varphi_1(t) = \varphi_1(0) + 2t, \qquad \varphi_2(t) = 5\varphi_2(0) + 5t$$

$$I_1(t) = I_1(0) + 5\lambda tc, \qquad I_2(t) = I_2(0) - 2\lambda tc, \qquad c = \cos(5\varphi_1(0) - 2\varphi_2(0)).$$

Thus, nothing drastic happens except that the action variables increase linearly in time. Mathematically, this is harmless, since the group structure of the time evolution tells us that the worst case is exponential growth. In reality this would be catastrophic if it were to go on forever, but we have seen in Section II that the linear increase of the amplutide of oscillation is quenched by nonlinear effects, which break the resonance. Nevertheless, in our strategy we have to be prepared for the worst and stay away from points in phase space where the frequencies are rationally related. In fact, in our restricted three-body problem we seem to be in trouble right at the beginning because in the two-body Kepler problem the angular and radial frequencies $(\omega_\varphi, \omega_r)$ are not only rationally related on some points but equal in all of phase space where $H < 0$. This difficulty is spurious since we have to go into the frame rotating with Jupiter and there (Ref. [13], 4.4.12) the Hamiltonian becomes ($M_J = \mu$, $M_\odot = 1$, $\vec{x}_J = (1,0)$)

$$H_1 = \frac{1}{2}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) - p_\varphi - \frac{1}{r}$$

and $\omega_\varphi = \omega_r - 1$. (Jupiter is now fixed and for circular orbits $\omega_r = r_0^{3/2}$, so for $r_0 = 1$ we have $\omega_\varphi = 0$.) However, the perturbation $H_1' = \mu(r^2 - 2r\cos\varphi + 1)^{-1/2}$ is not a polynomial in the exponentials of the angle variables since r is rather complicated when expressed by action-angle variables (Ref. [13],5.3.15,2). Thus all $\widetilde{H}_{k_1,k_2}(I)$ will be different from zero and to avoid $\omega_r k_1 + \omega_\varphi k_2 = 0$ we have to delete all rational $\omega_\varphi/\omega_r = 1 - r_0^{-3/2}$. Since this set is dense in phase space, $H = H^\infty(I_k)$ cannot hold in an open set and we still seem to be in trouble. One might cherish some hope because this set has no interior points and is of measure zero. This hope is destroyed by

a_2). For the series (9) to converge we need not only $(\omega \cdot k) \neq 0$ but it has to stay sufficciently far away from zero. However, since the rationals are dense in $\mathbf{R}$ we can approximate ω_r/ω_φ closely by k_2/k_1 if the k's are sufficiently big. So the situation can be saved only if the $\widetilde{H}_k$ decrease sufficiently with increasing k. It is known that if H' is r-times differentiable $\widetilde{H}$ decreases with a power r, and if $\widetilde{H}$ is analytic it decreases exponentially. Away from $r = 1$, $\varphi = 0$ we have the latter situation, so $\widetilde{H}_k$ can beat any power. Thus, if

$$|\widetilde{H}_k| < c\, e^{-|k|\rho}, \qquad |k| = \sum_{j=1}^{m} |k_j|,$$

in the regions of phase space where for some n we have

$$(\omega \cdot k) \geq \frac{\varepsilon}{|k|^n} \quad \forall\, 0 \neq k \in Z^m, \tag{11}$$

there is no problem with the convergence in (9) since

$$\sum_k |k|^n \; e^{-\rho|k|+i(\varphi\cdot k)} < \infty.$$

We even have analyticity for $|\mathrm{Im}\ \varphi_j| < \rho$. But are there ω's which satisfy (11)? The good set G is in our planar case

$$G = \left\{ (\omega_r, \omega_\varphi) : \forall\ k \neq 0 \ \left| \frac{\omega_\varphi}{\omega_r} - \frac{k_1}{k_2} \right| \geq \frac{\varepsilon/\omega_r}{|k_2||k_1+k_2|^n} \right\}, \tag{12}$$

so its complement does not only contain all the rationals. It even contains an open neighbourhood of each of them. To some extent this agrees with our previous experience where it did not make much difference whether one is exactly on the resonance or just close, but now we learn that the bad set G^c is not only dense but also open. It is surprising that there is still something left over for G, and people with a brilliant physics intuition thought that it is not. Yet simple consideration shows that the measure of G^c goes with ε to zero. We may consider in our case $0 \leq \omega_\varphi/\omega_r \leq 1$, so k_1 and k_2 have the same sign (say positive) and $k_1 < k_2 + 1$. Now we just add the length of the dangerous intervals around ω_φ/ω_r given by (12). Since they might overlap we get an inequality, which, however, goes in the right direction,

$$\mu\left(\frac{\omega_\varphi}{\omega_r} \in D\right) \leq 2 \sum_{\substack{k_2+1>k_1>0 \\ k_2>0}} \frac{\varepsilon/\omega_r}{k_2^{1+n}} \leq \frac{2\varepsilon}{\omega_r} \sum_{k_2>0} \frac{k_2+1}{k_2^{1+n}} < \frac{2\varepsilon}{\omega_r}\left(\frac{1}{n-1}+\frac{1}{n}\right). \tag{13}$$

Thus for n sufficiently big and for small ε, there is a lot left over for G where first order perturbation theory works.

b) The Iteration.

If we include $\lambda\widetilde{H}_{k=0}$ into H_1 then the Hamiltonian regains its original form except that λ is replaced by λ^2. Before starting the same procedure again we have to check whether the resonance condition holds. In fact, the new term will add $\lambda\frac{\partial}{\partial I}\widetilde{H}_{k=0}$ to the frequencies and may break a resonance in H_1, the effect we encountered in Section II. However, by the same token it may also throw us into a resonance and we have to be able to avoid that by moving a little with the action variables. This would not help in the simple example (10) where the frequencies are fixed. One needs at least some quadratic terms in the action variables such that the Hessian

$$C := \det\left(\frac{\partial}{\partial I_j}\frac{\partial}{\partial I_j} H_1\right) \neq 0.$$

If H_1 is quadratic in the I_j one can manage this with some effort [13], for the general case we recommend Ref. [18], or for more courageous people the original paper by Arnold [15].

c) The Convergence.

In the terminology of physicists we have carried out a renormalization group transformation, and now we have to prove that it leads to a fixed point. What one needs is that for some norm $\| \ \|$ at each step $\|H'_n\|$ gets smaller than the square of the previous one, since the recursive relation

$$\|H'_n\| \leq \|H'_{n-1}\|^2 \gamma \delta^n$$

implies

$$\|H'_n\| \leq \frac{(\gamma \delta^3 \|H'_1\|)^{2^{n-1}}}{\gamma \delta^{n+2}}. \tag{14}$$

Thus, if $\gamma \delta^3 \|H'_1\| < 1$ for $n \to \infty$ $\|H'_n\|$ converges to zero, we have reached our goal. The constant γ contains among other things the perturbation parameter λ, and by making it sufficiently small we can always satisfy this inequality. The estimate of δ is very cumbersome and contains also $\|C^{-1}\|$ and, alas, the price for stability from now to eternity is high. An estimate by Hénon [17] limits λ to be $< 10^{-50}$. Celletti and Chierchia [16] have truncated $\sum_k \widetilde{H}_k e^{i(\omega \cdot k)}$ and got the limit down to 10^{-6}, but what one would need is $M_J/M_\odot \sim 10^{-3}$. This truncation is not mathematically rigorous but physically reasonable since in our planetary system there are more important influences. This is like in music where a consonant interval contains higher overtones which are strongly dissonant, but they are only faintly excited and do not bother us.

Since the question of stability to eternity does not seem to be amenable to a complete understanding, we turn to another feature of the problem which can be easily deduced and understood. It is the problem of energy gain by a test planet interacting with a rotating system. This is of obvious importance as a fuel saving measure for space travel and leads, in the extreme, to the celebrated example of a planet gaining so much energy between two rotating binaries that it can push them to infinity within a finite time [10]. The simple general rule is expressed by the following

Theorem: (15)

Let a body with coordinates $X(t)$ rotate around the origin, $|X(t)| =$ constant, and interact with the test planet (coordinates x) through a central potential $V(r)$. Denote by the future (resp. past) half space the half space bounded by the plane perpendicular to $\dot{X}(t)$ and going through $X(t)$ and the origin, into which $X(t)$ moves (resp. which it leaves). Then, if $V(r)$ is attractive ($V' > 0$), the test planet gains energy if it is located in the past half space, and loses energy when it is in the future half space. For repulsive V, it is the other way round.

Proof:

For a potential depending explicitely on the time t, the change of the test-planet

energy $E(t)$ in the fixed frame is simply given by $dE/dt = \partial V/\partial t$. In our case, $V(r) = V(|X(t) - x|)$ and $\partial|X(t) - x|^2/\partial t = -2(\dot{X}(t) \cdot x)$. Thus,

$$\frac{\partial V}{\partial t} = -\frac{(\dot{X}(t) \cdot x)}{|X(t) - x|} V',$$

which, in the attractive case, is positive in the past half space, $(\dot{X} \cdot x) < 0$, and negative otherwise. For repulsive potentials, $V' < 0$, it goes the other way.

Remarks:
1) To accumulate energy, x has to remain in the past half space for some time and has to follow $X(t)$. This gives some intuitive basis for the theorem since it means that the test planet gains energy as it is dragged along.
2) In our two-dimensional example with the Newton potential $V = -\mu/r$, and with $X(t) = (\cos t, \sin t)$, we have with polar coordinates (r, ϕ) for x: $|X(t) - x| = (1 - 2r\cos(\phi - t) + r^2)^{1/2}$ and $dE/dt = \mu r \sin(t - \phi)(1 - 2r\cos(t - \phi) + r^2)^{-1/2}$. If we assume in a first approximation that x follows a circular orbit $(r, \omega t)$ with $\omega = r^{-3/2}$, the accumulated energy gain between $t = 0$ and $t = \pi/(1-\omega)$ becomes

$$\mu \int_0^{\pi/(1-\omega)} dt \frac{r \sin(t(1-\omega))}{(1 - 2r\cos t(1-\omega) + r^2)^{3/2}} = \frac{\mu}{1-\omega} \begin{cases} 2/(r^2 - 1) & \text{for } r > 1 \\ 2/(1 - r^2) & \text{for } r < 1. \end{cases}$$

Since $\omega < 1$ for an outside orbit, and $\omega > 1$ for an inside orbit, the former gains energy (the test planet is dragged along) and the latter looses energy (it is pulled back). Since $\omega = r^{-3/2}$, the planet is ejected after one swing if $\mu(r-1)^{-2} \sim 1$.
3) In our system actually both the sun and Jupiter rotate around their center of gravity, and there is a contribution to (15) also from the sun with the opposite sign, since the sun goes up when Jupiter moves down. Although the force of the sun is $\sim M_\odot$, its velocity is $\sim M_\odot^{-1}$ and the mass dependence cancels out. Only the distance matters. If sun and Jupiter are located at $(1/2, 0)$ and $(-1/2, 0)$, respectively, then the quadrants II and IV in the co-rotating frame are energy increasing, the others energy decreasing.

We illustrate this behavior in Fig. 8 by a trajectory of a test particle in the synodic, co-rotating frame, for which Jupiter is located at (-1/2,0), and the sun at (1/2,0). Initially, the particle is at $(x, y) = (-\cos(a) + 0.5, \sin(a))$, $a = \pi/18$, with a synodic velocity $(-0.9\cos(a), 0.9\sin(a))$ pointing away from the origin. In the absence of Jupiter, the particle is trapped by the sun with a negative fixed-frame energy $E = -0.095$, and follows the dashed trajectory in Fig. 8. For the same initial conditions but with a mass ratio $M_J/M_\odot = 0.1/0.9$, the test particle traces out the smooth line in the Figure. As predicted by (15), between A and B in the past half space it is attracted by M_J, and the fixed-frame energy increases from -0.6465 to +0.0668. As follows from Remark 3), the energy E starts to decrease slightly again between the points B and C in the first quadrant as is shown in Fig. 9. For small mass ratios this second-order correction to (15)

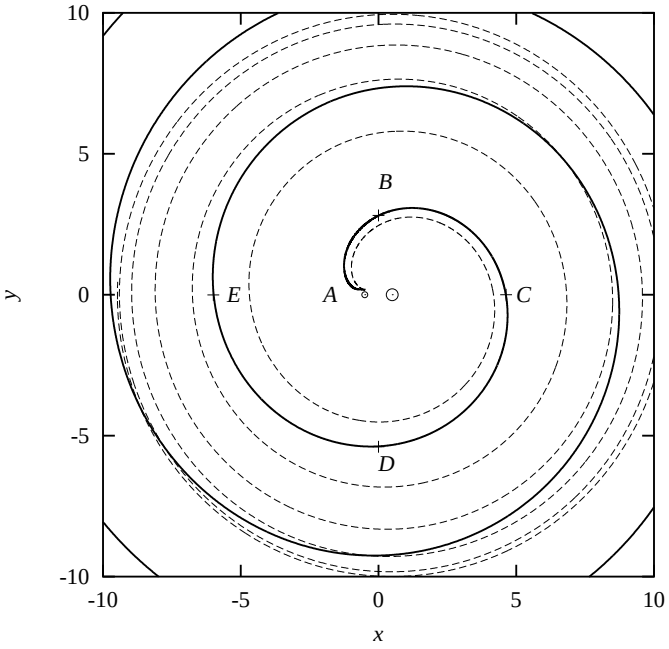

FIG. 8. Example of an unperturbed (dashed) and a perturbed trajectory (full line, mass ratio $M_J/M_\odot = 1/9$) in a synodic, co-rotating coordinate system for an initial position $-\cos(a)+0.5, \sin(a))$, $a = \pi/18$, and an initial synodic velocity $(-0.9\cos(a), 0.9\sin(a))$. The position of the sun at $(1/2, 0)$ and of Jupiter at (-1/2, 0) are indicated by dots.

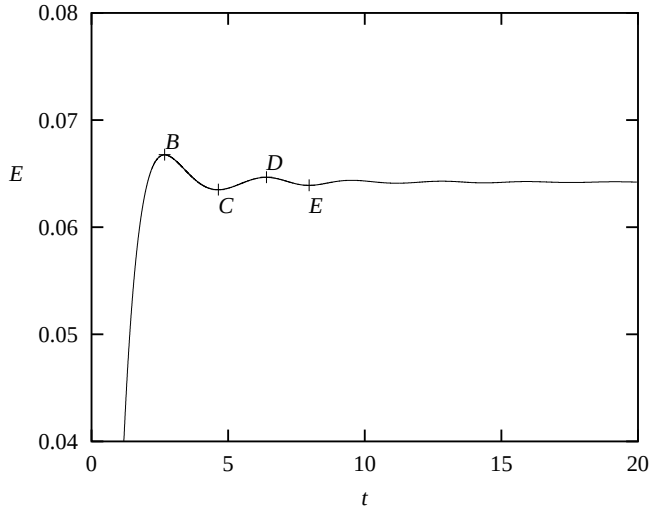

FIG. 9. Time dependence of the total fixed-frame energy for the perturbed test particle of Fig. 8. The labels correspond to the positions in Fig. 8.

becomes neglegible. The final energy in our example remains positive, and the particle escapes to infinity. Since the trajectory is represented here in the co-rotating frame, this path to infinity appears as a spiral. It is interesting to note that some comets actually perform such loops around Jupiter.

In Fig. 10 we show another example of a trajectory in the synodic frame, for which the test particle is in the past half space between the points A and B and, as a consequence, gains considerable energy during that time. The time dependence of the total fixed-frame energy is given in Fig. 11. Between the points B and C the particle is in the future half plane and loses so much energy that the subsequent energy gain in the past half space beyond the point C still leaves the energy negative asymptotically. The particle remains trapped and does not escape. Also in this example the mass ratio $M_J/M_\odot = 0.1/0.9$, and the synodic locations of the sun and of Jupiter are indicated by the big and small circles, respectively. The large variation of the energy in Fig. 11 is reminiscent of the energy oscillations in the Sitnikov problem [11,12].

IV. THE COMPUTER

Let us return for a moment to the self-quenching phenomenon which led to the introduction of the kicked-oscillator model in Section II. If the perturbation $M_J/M_\odot$ becomes bigger than about $1/100$, then the orbit gets out of its harmonic shelter too soon for the stabilizing factor to become effective, and the orbit will come close to the sun or Jupiter. Then the test body will be thrown out of its original circle and the orbit becomes chaotic. This is demonstrated in Fig. 12 for a few "exact" perturbed trajectories distinguished by the mass ratio $M_J/M_\odot$. In terms of the kicked-oscillator model, for the quenching mechanism to be effective it is essential that the beat it strictly observed. If the planet is thrown out too far of its harmonic regime and the frequency of the radial motion becomes strongly dependent on the amplitude, it never gets the rhythm. Thus, it is bound to happen that a few kicks will throw the planet beyond the point of no return. Then there is no stabilizing mechanism, and chaos prevails.

For these large perturbations the analysis of Sect. 3 certainly does not apply, and the first guess is that then the system becomes ergodic. For this to be true one first has to make sure that the orbit remains in a compact region in phase space. If it escapes to infinity then with probability one it has also come from infinity and one has a scattering situation. In this case one even has the maximal number of constants of the motion (three in our case) and one is in the opposite extreme of ergodicity. However, in the synodic rotating frame the Hamiltonian (6) can be written [13,3]

$$H = \frac{1}{2}\left[(p_x + y)^2 + (p_y - x)^2\right] + \Omega(x, y),$$

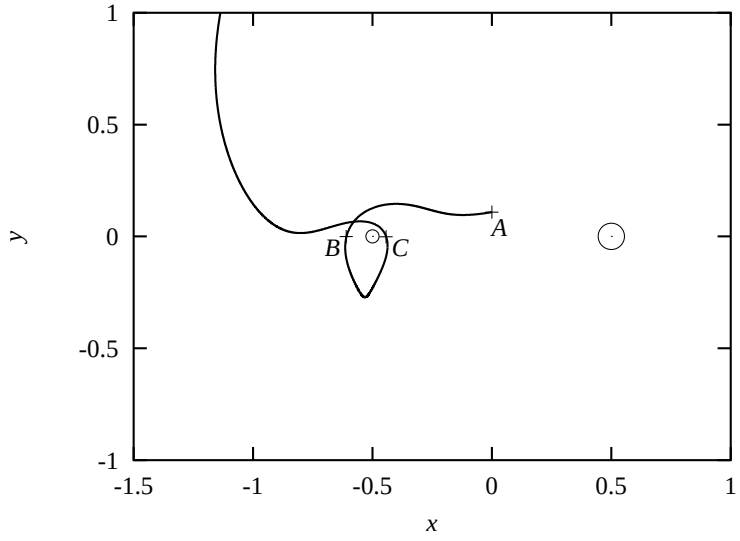

FIG. 10. Perturbed test-particle trajectory in the co-rotating frame displaying motion in the past half plane (between the points A and B), in the future half plane (between the points B and C), and in the past half plane again (beyond C).

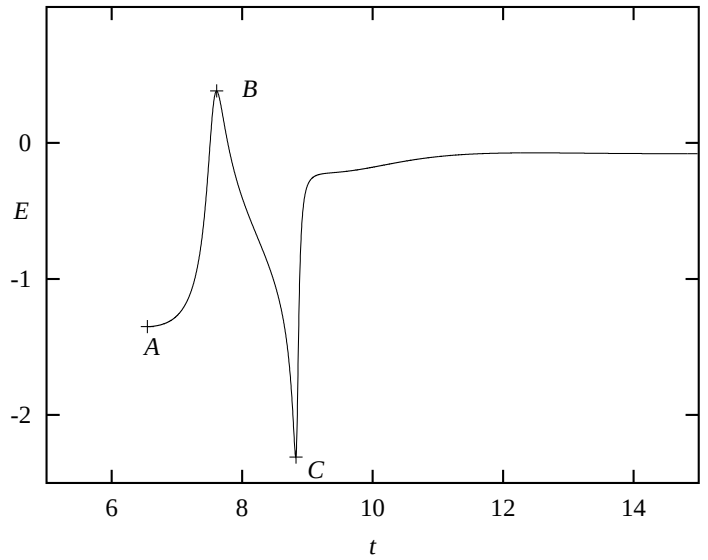

FIG. 11. Time dependence of the total energy in the fixed (inertial) frame for the trajectory in Fig. 10. The labelled points refer to the corresponding locations in Fig. 10.

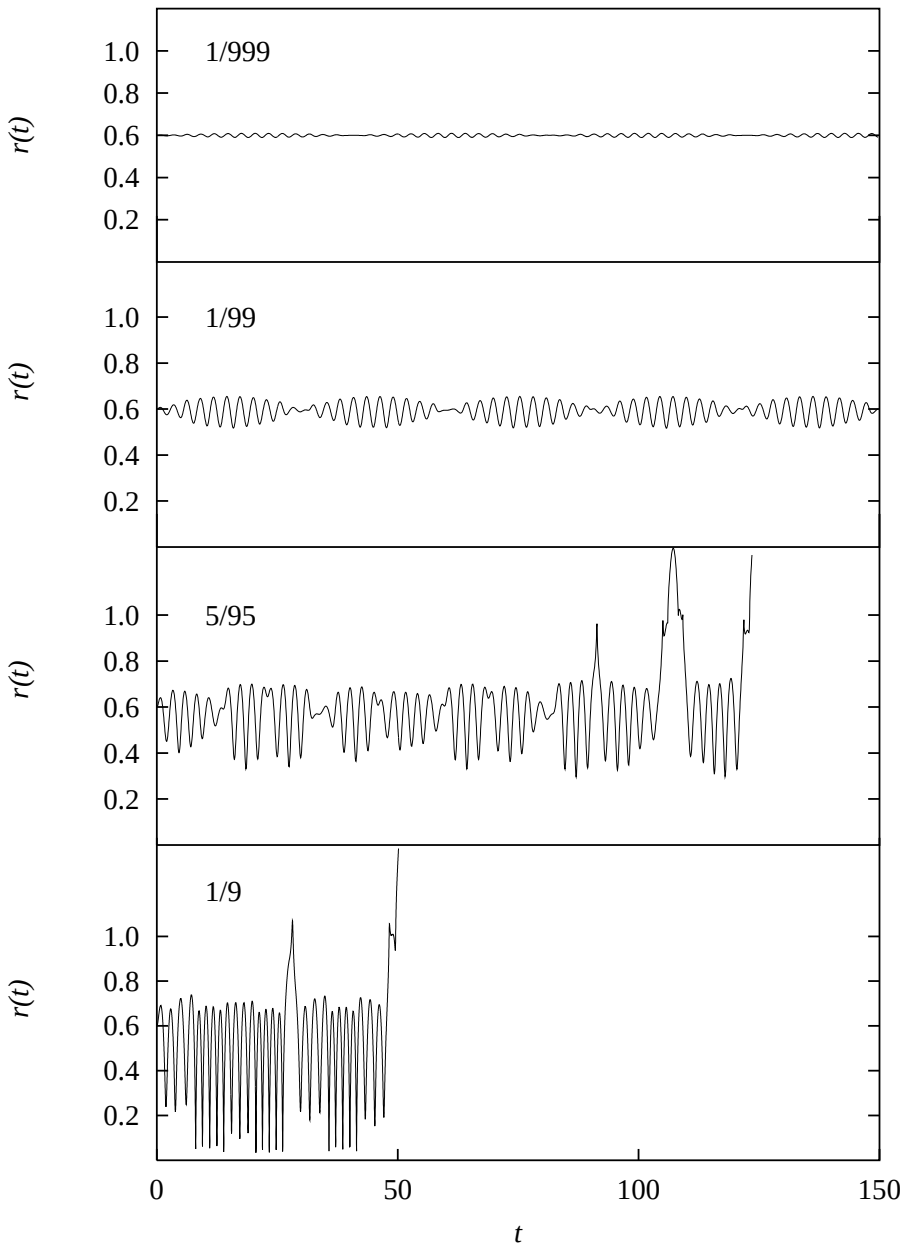

FIG. 12. Radial planetary motion, perturbed by Jupiter, for the unperturbed circular-orbit radius $r_0 = 0.6$. The mass ratio $M_J/M_\odot$ increases from top to bottom as indicated by the label.

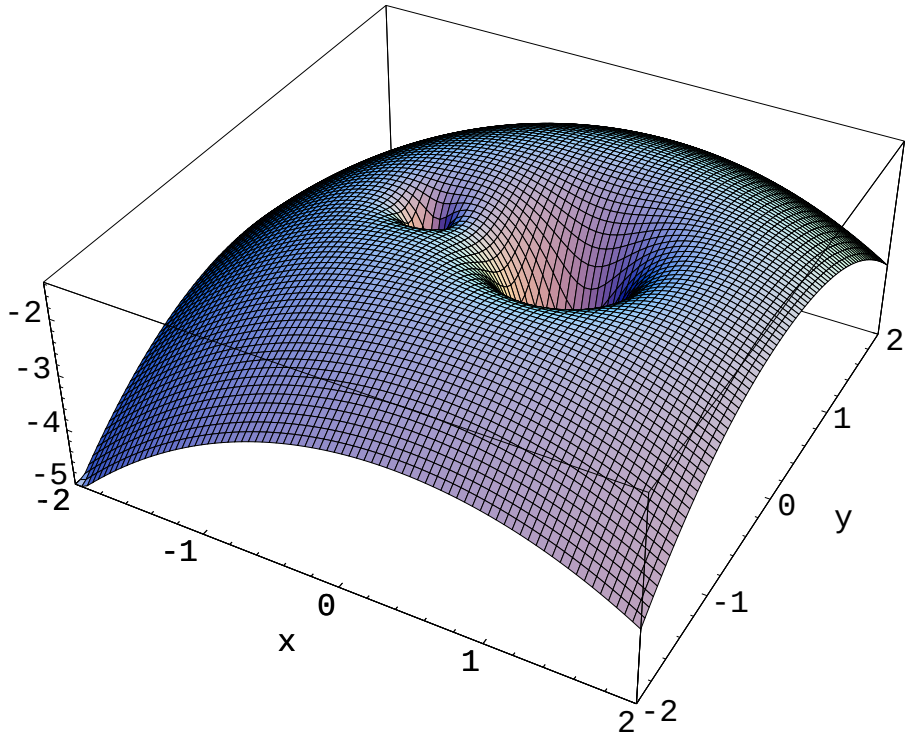

FIG. 13. The surface $\Omega(x, y)$ for a mass ratio $M_J/M_\odot = 1/9$. The sun is located at (0.5,0), Jupiter at (-0.5,0).

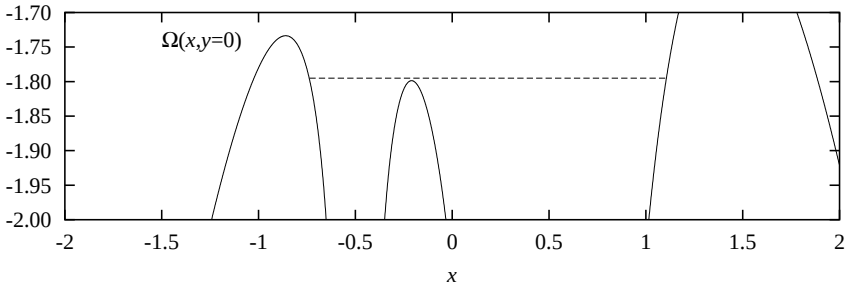

FIG. 14. Cut through the Ω-surface of Fig. 13 along the x axis. The mass ratio $M_J/M_\odot = 1/9$. The horizontal line corresponds to an energy $E_0 = -1.795$ in the synodic frame.

thus like for a particle in a constant magnetic field perpendicular to the plane of motion and subjected to a potential

$$\Omega(x,y) = -\frac{1}{2}(M_\odot r_\odot^2 + M_J r_J^2) - \frac{1}{2}\left(\frac{M_\odot}{r_\odot} + \frac{M_J}{r_J}\right) + \frac{M_\odot M_J}{2}$$

(see Fig. 13). Here, $r_\odot = [(x-1/2)^2 + y^2]^{1/2}$ and $r_J = [(x+1/2)^2 + y^2]^{1/2}$ are the distance of the planet from the sun and from Jupiter, respectively, where we use for convenience a co-rotating frame in which the sun is located at (1/2,0) and Jupiter at (-1/2,0). The regions $\Omega < E_0$ are time invariant and are compact in configuration space for sufficiently low synodic energy E_0. So the question is in this case whether the energy shell $H = E_0$ is covered uniformly by the orbit or whether it is divided further by hitherto not-discovered constants. We shall see that neither seems to be the case. Since neither physical intuition nor rigorous mathematics are in a position to answer this question we have to avail ourselves of modern computer technology.

Ergodicity means that the time average of the orbit gives a homogeneous density on the energy shell. The former we have to calculate on the computer and the latter, $\delta(H(x,y;p_x,p_y) - E_0)$, becomes particularly simple when projected onto configuration space as follows from the more general Bohr-van Leeuwen-type (see Ref. [21], 2.5.39,1)

Theorem: (16)
In two dimensions the microcanonical density in configuration space of a particle in an arbitrary potential and arbitrary magnetic field is constant in the energetically allowed region.

Proof:

$$H = \frac{1}{2}\left[(p_x - A_x(x,y))^2 + (p_y - A_y(x,y))^2\right] + V(x,y),$$

$$\begin{aligned}\rho(x,y) &= \int dp_x dp_y \delta(H - E_0) = \int dv_x dv_y \delta\left(\frac{1}{2}(v_x^2 + v_y^2) + V(x,y) - E_0\right) \\ &= 2\pi\Theta\left(V(x,y) - E_0\right)\end{aligned}$$

with $v_i = p_i - A_i$ and Θ the step function.

To study this chaotic behaviour in more detail we have followed the dynamical evolution on the computer. Since we are concerned with long chaotic trajectories, a regularization procedure according to Birkhoff is used to remove the singularities at the position of both primaries [20,3]. In combination with a Runge-Kutta 4-th order algorithm with variable time step we ascertain that the energy is conserved to 10 significant digits over the whole length of the simulation. In Fig. 15 a stroboscopic map reflecting the probability density in configuration space is shown. The energy $E_0 = -1.795$ was chosen to allow for a narrow channel

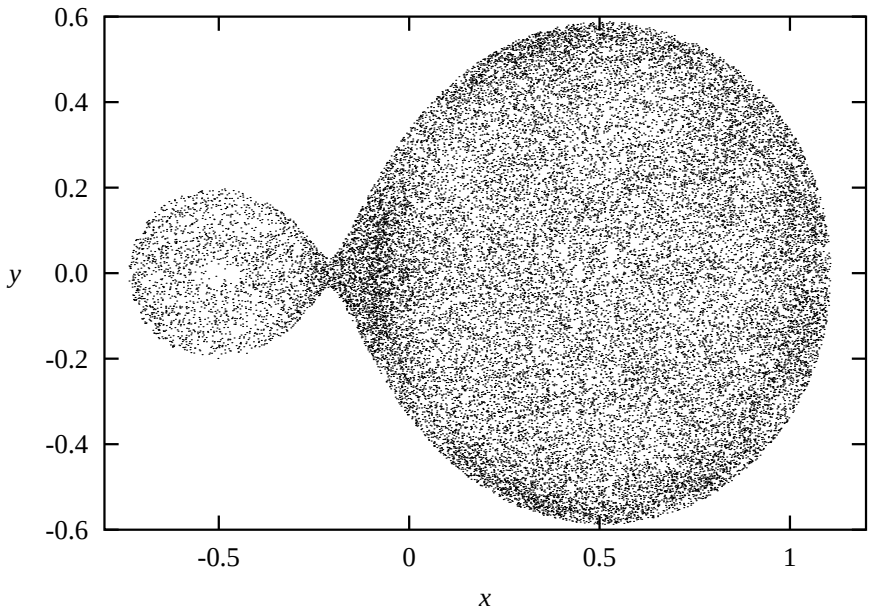

FIG. 15. Probability density in configuration space for a synodic energy $E_0 = -1.795$ corresponding to the horizontal dashed line in Fig. 14. The points of this stroboscopic map are taken from a single chaotic trajectory lasting for about 30000 Jupiter years. The mass ratio $M_J/M_\odot = 1/9$. The sun in this rotating frame is located at (0.5,0), Jupiter at (-0.5,0).

between the sun and Jupiter, and corresponds to the dashed horizontal line in Fig. 15. The initial configuration for this trajectory, which is followed for 30000 Jupiter years, is at the position of the central saddle point between the sun and Jupiter in Fig 13, with the planet velocity pointing towards the sun. Clearly, the distribution of points in Fig. 15 is almost homogeneous. The fact that the theorem does not strictly apply and some accumulation of points appear at the boundary is a consequence of the fact that the system is not ergodic.

This may be seen more clearly by looking at other phase-space projections, say onto the (x, v_x)-plane, which are harder to treat theoretically. In Fig. 16 a double-sided Poincaré map in the $(x - v_x)$-plane is shown for the same chaotic trajectory as in Fig. 15. The plotted points correspond to states for which the velocities $\dot{y} \equiv v_y$ in y-direction may be positive or negative. The results show that there are large islands of regularity in the chaotic sea, so the system is not ergodic. Nevertheless, the consequences of Theorem (16) are quite well satisfied, and the density in configuration space is nearly homogeneous. In Fig. 16 the sections of a few regular tori are also shown in some of the regularity islands for the same energy, $E_0 = -1.795$.

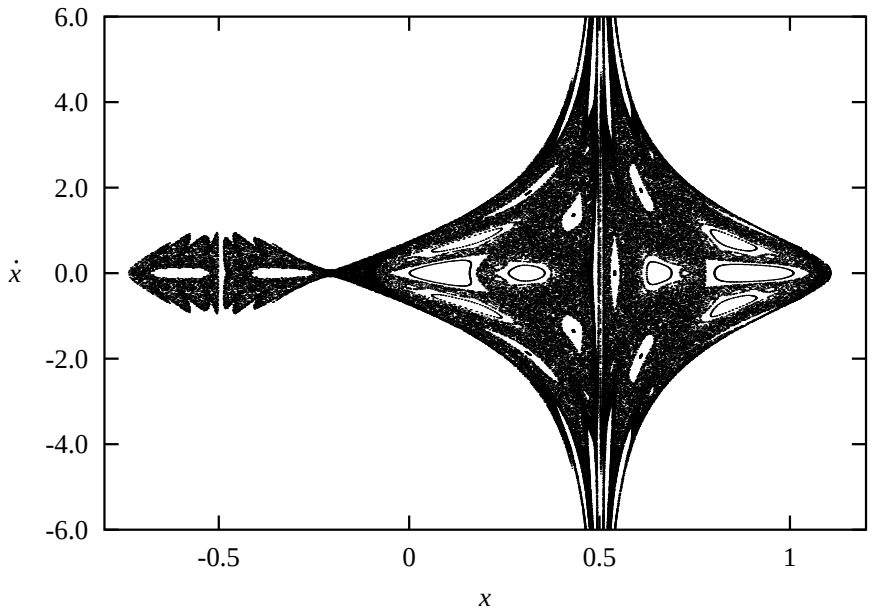

FIG. 16. Double-sided Poincaré map for mass ratio $M_J/M_\odot = 1/9$ and a synodic energy $E_0 = -1.795$ corresponding to the horizontal dashed line in Fig. 14. The sun in the rotating frame is located at (0.5,0), Jupiter at (-0.5,0). In some of the larger regularity islands the closed sections of regular tori for the same total energy are also shown.

V. SUMMARY

The (reduced) three-body problem, the cradle and for more than a century a paradigm of the science of chaotic dynamical systems [22,23], still provides new and surprisingly simple results. For weakly-perturbed circular orbits one gets away with a very crude and intuitive kicked-oscillator model, which provides a good representation of the eccentricity of the test-particle motion. In Section II we have explored the strengths and limitations of this model. For a small-enough mass ratio $M_J/M_\odot$, even resonance conditions may not be catastrophic for the orbit of the test particle due to a selfquenching mechanism supplied by the nonlinearity of the effective potential in the co-rotating frame. For large perturbations, however, the model is not applicable and only computer simulations may provide test-particle trajectories for a reasonable length of time.

If one considers conditions for which the test particle may visit the neighborhood of the sun and of Jupiter but is still bounded by the Jacobi constant in the frame co-rotating with Jupiter, the phase space consists of a chaotic sea with regularity islands enbedded. The existence of these islands demonstrates that the system is not ergodic in spite of considerable nonlinearities in the potential Ω.

According to Theorem (16) of Section IV, ergodicity requires that the probability density of the test particle in configuration space is constant in the allowed domain. Computer simulations in Section IV demonstrate that this is not strictly the case due to nonergodicity.

What one really wants to know is what is the measure associated with the regular domains in phase space, in which the trajectory stays on a toroidal submanifold, for which the characteristic function is an additional constant of the motion. Only rigorous mathemats is capable of answering this question, if at all. Impressive progress has been achieved recently by Celletti and Chierchia [16], although one is still a few orders of magnitude away from this goal. However, analytical theory in Section III provides a surprisingly simple answer to a less-ambitous question concerning the gain and loss of the test-particle energy $E(t)$ in the fixed, inertial frame. Let us consider a synodic co-rotating frame in which Jupiter is located at $(-1/2, 0)$ and the sun at $(1/2, 0)$. Then we conclude from Theorem (15) and the following remarks in Section III that $\dot{E} > 0$ whenever the particle is in the second or fourth quadrant of that frame of reference, and $\dot{E} < 0$ whenever it is in the first or third quadrant. This simple result has been confirmed by numerical simulations.

ACKNOWLEDGMENTS

We gratefully acknowledge helpful discussions with Rudolf Dvorak, William G. Hoover, Heide Narnhofer, and Karl Wodnar, and support from the Fonds zur Förderung der wissenschaftlichen Forschung, Grant No. P11428-PHY.

[1] M. Gutzwiller, Rev. Mod. Phys. **70**, 589 (1998).
[2] M. Hénon, *Generating Families in the Restricted Three-Body Problem* (Springer, Berlin, 1997).
[3] V. Szebehely, *Theory of Orbits: The Restricted Problem of Three Bodies* (Academic Press, New York, 1967).
[4] Lj. Milanović, H.A. Posch, and W. Thirring, Phys. Rev. E **57**, 2763 (1998).
[5] J. Laskar, Icarus **88**, 266 (1990).
[6] J. Laskar and Ph. Robutel, Nature **361**, 608 (1993).
[7] N. W. Evans and S. Tabachnik, Nature **399**, 41 (1999).
[8] M. Mayer and D. A. Queloz, Nature **378**, 355 (1995).
[9] A. C. Cameron, K. Horne, A. Penny, and D. James, Nature **402**, 751 (1999).
[10] Z. Xia, Ann. of Mathematics **135**, 411 (1992).
[11] K. Wodnar, Sitzungsberichte d. Österr. Akad. d. Wiss. Abt. II, Mathematische, Physikalische und Technische Wissenschaften, **202**, 133 (1993).
[12] R. Dvorak and Y. S. Sun, Celestial Mechanics and Dynmical Astronomy, **67**, 87 (1997).
[13] W. Thirring, *Classical Mathematical Physics: Dynamical Systems and Field Theories* (Springer, New York, 1997).
[14] J. Moser, *Stable and Random Motions in Dynamical Systems*, (Princeton University Press, Princeton, 1973).
[15] V. Arnold, Dokl. Akad. Nauk. **142**, 758 (1962).
[16] A. Celletti, and L. Chierchia, Planet. Space Sci. **46**, 1433 (1998).
[17] M. Hénon, Bulletin Astronom. Soc. **3**, 49 (1966).
[18] G. Gallavotti, *The Elements of Mechanics* (Springer, New York, 1983).
[19] *The Dynamical Behaviour of our Planetary System*, R. Dvorak and J. Henrard (eds.) (Kluwer, Dordrecht, 1997).
[20] G. D. Birkhoff, Rend. Circ. Mat. Palermo **39**, 1 (1915). also *Collected Mathematical Papers*, Vol. 1, p.628 (Am. Math. Soc., New York, 1950).
[21] W. Thirring, A Course in Mathematical Physics Vol. IV: Quantum Mechanics of Large Systems (Springer, New York, 1982).
[22] H. J. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, Vols. 1 - 3 (Gauthier-Villars, Paris, 1892, 1893, 1899); reprinted by Librairie Albert Blanchard, 1987.
[23] F. Diacu and Ph. Holmes, *Celestial Encounters: The Origin of Chaos and Stability* (Princeton University Press, Princeton, 1996).