SYMPLECTIC MAPS FOR THE *N*-BODY PROBLEM

Jack Wisdom and Matthew Holman

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 8 January 1991; revised 14 May 1991

ABSTRACT

The mapping method of Wisdom [AJ, 87, 577 (1982)] is generalized to encompass all gravitational *n*-body problems with a dominant central mass. The method is used to compute the evolution of the outer planets for a billion years. This calculation provides independent numerical confirmation of the result of Sussman & Wisdom [Sci, 241, 433 (1988)] that the motion of the planet Pluto is chaotic.

1. INTRODUCTION

Long-term integrations are playing an increasingly important role in investigations in dynamical astronomy. The reason is twofold. First, numerical exploration is an essential tool in the study of complex dynamical systems which can exhibit chaotic behavior, and there has been a growing realization of the importance of chaotic behavior in dynamical astronomy (see, e.g., Wisdom 1987). Second, there has been a phenomenal increase in the capabilities of computers which is bringing many important problems in dynamical astronomy within reach. In particular, there has recently been considerable interest in the long-term evolution of the solar system. Long-term integrations of the solar system include the outer planet integrations of Cohen *et al.* (1973; 1 Myr), Kinoshita & Nakai (1984; 5 Myr), the first Digital Orrery integration (Applegate *et al.* 1986, 210 Myr), the LONGSTOP work (Roy *et al.* 1988; 100 Myr), the second Digital Orrery integration (Sussman & Wisdom 1988; 845 Myr), and the inner planet integrations of Richardson & Walker (1987; 2 Myr), Applegate *et al.* (1986; 3 Myr), and Quinn *et al.* (1991; 3 Myr). Long-term integrations have already produced startling results. Sussman & Wisdom (1988) found numerical evidence that the motion of the planet Pluto is chaotic, with a remarkably short timescale for exponential divergence of trajectories of only 20 million years. This massive calculation consumed several months of time on the Digital Orrery, a computer built specifically for the job which runs at about a third the speed of a Cray 1. Subsequently, Laskar (1989, 1990), in another massive computation, found numerical evidence that the motion of the inner planets is also chaotic, with a divergence timescale of only 5 million years. However, despite the phenomenal progress in computer technology, computers are still too slow for many important applications. For example, it is very important to test the sensitivity of the results concerning the chaotic character of the motions of the planets to uncertainties in initial conditions and parameters. It is also important to clarify the dynamical mechanisms responsible for the chaotic behavior to confirm that the positive Lyapunov exponents are not subtle numerical artifacts. The necessary calculations and those of many other problems of current interest in dynamical astronomy require orders of magnitude greater computing power than is currently available. Regardless of the speed of computers, better, faster algorithms for investigating the *n*-body problem are always welcome. This paper presents a new method for studying the long-term evolution of the *n*-body problem which is an order of magnitude faster than traditional methods of numerical integration. The method is a generalization of the "mapping" method introduced by Wisdom (1982, 1983) to study the motion of asteroids near the 3:1 mean-motion resonance with Jupiter. It is applicable to systems which are dominated by a large central mass such as planetary systems or satellite systems.

The mapping method of Wisdom (1982, 1983) was based on the averaging principle. It was noted that most studies of the long-term evolution of the *n*-body problem relied on the averaging principle in one way or another. This included both analytical and numerical studies. The intuition behind the averaging method is that rapidly oscillating terms tend to average out and give no net contribution to the evolution, while more slowly varying resonant or secular terms accumulate to give significant contributions to the evolution (see Arnold 1974). The intuition behind the mapping method was just the same: If the rapidly oscillating terms do not contribute significantly to the evolution then replacing them with other rapidly oscillating terms will have no ill effect. To get the mapping the rapidly oscillating terms are chosen so that they sum to give delta functions which can be locally integrated to give explicit equations specifying how the system changes from one step to the next. The mapping method was inspired by Chirikov's use of periodic delta functions to derive a Hamiltonian for the standard map (Chirikov 1979). The time step covered by the map is on the order of the period associated with the high-frequency terms. For the asteroid maps, the basic step was one full Jupiter period. The algebraic simplicity of the 3:1 map and the large step size combined to make it extraordinarily fast, about 1000 times faster than even the numerical averaging routines available at the time (Wisdom 1982). The great speed of the map allowed studies of the resonant asteroid motion over much longer times than were previously possible, and significant new phenomena were discovered. In particular, it was found that there was a large zone of chaotic behavior near the 3:1 resonance and that chaotic trajectories in these zones often displayed a peculiar phenomenon in which the eccentricity could remain at relatively low values for several hundred thousand years and then suddenly jump to much higher values. Over longer intervals of millions of years there were periods of low eccentricity behavior interspersed with bursts of high eccentricity behavior. These bursts in eccentricity were subsequently confirmed in traditional direct integrations of Newton's equations (Wisdom 1983; Murray & Fox 1984; Wisdom 1987), and explained perturbatively (Wisdom 1985a). The high eccentricities attained by the chaotic trajectories help explain the formation of the 3:1 Kirkwood gap (Wisdom 1983), as well as provide a mechanism for transporting meteoritic material directly from the asteroid belt to Earth (Wisdom 1985b; Wetherill 1985). Murray

0004-6256/91/041528-11$00.90

(1986) applied the mapping method to the 2:1 and the 3:2 resonances. Sidlichovsky & Melendo (1986) applied the method to the 5:2 resonance. Tittemore & Wisdom (1988, 1989, 1990) have applied the method to study the tidal evolution of the Uranian satellites through numerous mean-motion commensurabilities. The result of Tittemore & Wisdom (1989) that secondary resonances play a crucial role in determining the inclination of Miranda has been confirmed by Malhotra & Dermott (1990), also using the mapping method. Tittemore (1990) and Malhotra (1990) have recently used the mapping method to study the tidal evolution of the Gallilean satellites. The mapping method has been tremendously useful.

Unfortunately, the mapping method, as originally presented, has significant limitations. It is based on analytic representations of the averaged Hamiltonian near particular resonances. The only known explicit analytic representations of the averaged disturbing function are as expansions in the eccentricities and inclinations, or the canonical equivalents. Though the mapping method itself has no particular limitation to low eccentricities and inclinations, the use of a disturbing function which is truncated at some order in both eccentricity and inclination limits the applicability of any particular realization of the mapping method to low eccentricity and inclination. The original 3:1 mapping which included second-order terms in eccentricity and inclination (ignoring fourth-order terms) gave qualitatively correct trajectories even for eccentricities as large as 0.4. However, it could not be relied upon for the investigation of Earth-crossing meteoroid trajectories which have eccentricities above 0.6 (Wisdom 1985). Murray (1986) also used a disturbing function truncated after second-order terms in the eccentricity in his study of motion near the 2:1 and the 3:2 resonances (ignoring third-order terms). The eccentricity must be much smaller at the 2:1 and the 3:2 resonances than at the 3:1 resonance for a second-order disturbing function to accurately represent the motion. Comparison of his results with those obtained with unaveraged numerical integrations performed on the Digital Orrery (Wisdom 1987) shows that significant artifacts appear in Murray's maps above an eccentricity of only 0.1 at the 2:1 resonance, and the extent of the chaotic regions determined by Murray's map is qualitatively wrong for the 3:2 resonance even at low eccentricity. It is important to emphasize that the failure of Murray's maps is not a failure of the mapping method, but rather a failure of the truncated disturbing function to represent the averaged Hamiltonian. Another important limitation of the mapping method, as it has been used up to the present, is that it is limited to the vicinity of a particular resonance or group of resonances, again because the analytic representation of the averaged disturbing function can only be made for a particular set of commensurabilities. Thus the systematic investigation of the tidal evolution of the Uranian satellite system through a sequence of mean-motion commensurabilities (Tittemore & Wisdom 1988, 1989, 1990) required a separate derivation of the map appropriate to each resonance, an unbelievably tedious process!

The generalization of the mapping method presented here does not have these limitations. It is not limited to particular resonances nor is it limited to low eccentricities and inclinations. It is valid everywhere. Of course this comes at a cost. The new mapping method is not as fast as the original mapping method, but it still offers a significant advantage over conventional direct numerical integration.

The next section presents the rationale for the generalized mapping method. Details of the mapping for the n-body problem are then presented. Subsequent sections present some refinements of the method and show the relationship of the mapping method to other symplectic integration methods. The new n-body map has been used to compute the evolution of the outer planets for a billion years. The resulting evolution is compared to the 845 million year evolution of the outer planets performed on the Digital Orrery using standard numerical integration techniques (Sussman & Wisdom 1988).

2. MAPPING METHOD

In the original mapping method the Hamiltonian is first separated analytically into parts with different associated timescales,

$$H = H_{\text{Kepler}} + H_{\text{Orbital}} + H_{\text{Resonant}} + H_{\text{Secular}}, \quad (1)$$

where H_{Kepler} represents the interaction of each body with the central mass, H_{Orbital} represents rapidly oscillating terms which depend on the mean longitudes of the bodies but are not resonant in the region of interest, H_{Resonant} represents the terms which have resonant combinations of mean longitudes, and H_{Secular} represents the remaining terms which do not depend on mean longitudes. The averaging principle is used to argue that the terms in H_{Orbital} will not significantly affect the long-term evolution near resonance, and can thus be neglected (or removed by suitable Von Zeipel transformations). The original mapping method then added additional terms with the orbital frequency, which sum, together with H_{Resonant}, into terms involving periodic sequences of Dirac delta functions. The new terms, by the averaging principle, also play no important role in the long-term evolution. The resulting map Hamiltonian is

$$H_{\text{Map}} = H_{\text{Kepler}} + H_{\text{Secular}} + 2\pi\delta_{2\pi}(\Omega t)H_{\text{Resonant}}, \quad (2)$$

where $\delta_{2\pi}(t)$ represents a periodic sequence of delta functions with period 2π,

$$\delta_{2\pi}(t) = \sum_{n=-\infty}^{\infty} \delta(t - 2\pi n) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} \cos(nt), \quad (3)$$

and Ω is the mapping frequency, which is of the same order as the orbital frequencies. In the asteroid map, the mapping period was chosen to be the period of Jupiter. Hamiltonian (2) is only a sketch of the true mapping Hamiltonian because in the earlier applications it was convenient to break H_{Resonant} into several parts, each of which was multiplied by its own sequence of delta functions. Those details are not important here. Between the times when the delta functions act, the Hamiltonian is just given by the first two parts, the Kepler part and the secular part. Provided the secular Hamiltonian is truncated at second order in eccentricities and inclinations (ignoring fourth-order terms) Hamilton's equations can be solved analytically between the delta functions. The system can also be analytically integrated across the delta functions. The result is an analytic expression for the state of the system at the end of a mapping period in terms of the state of the system at the beginning of the mapping period. The time evolution of the system is obtained by iterating the mapping step. It is easily shown that the mapping step is a canonical transformation or, in other words, that the mapping is symplectic. Again, these mappings are limited to particular regions in which only certain resonant terms are important, and to low eccentricity and inclination

by the truncation of both H_{Secular} and H_{Resonant} to some manageable order.

The mappings presented in this paper are based on a simpler separation of the Hamiltonian for the n-body problem:

$$H = H_{\text{Kepler}} + H_{\text{Interaction}}, \qquad (4)$$

where again H_{Kepler} represents the basic Keplerian motion of the bodies with respect to the central body, and $H_{\text{Interaction}}$ represents the perturbation of the bodies on one another. Of course, this division of the Hamiltonian is very natural and is the starting point for most perturbation theory. Despite this, few numerical integration methods take advantage of this division of the problem. Encke's method makes use of the integrability of the Kepler problem by integrating the variations of the planetary trajectories with respect to fixed reference orbits. However, a serious problem with Encke's method is that as the system evolves new reference trajectories must be frequently chosen so that the variations are not too large. Nevertheless, Encke's method was successfully used in the LONGSTOP integrations (Roy *et al.* 1988) to reduce numerical error. The symplectic n-body maps introduced here are quite distinct from Encke's method and more fully exploit the integrability of the Kepler problem.

A mapping Hamiltonian for the n-body problem can be simply obtained by adding high-frequency terms to this Hamiltonian so that it becomes

$$H_{\text{Map}} = H_{\text{Kepler}} + 2\pi\delta_{2\pi}(\Omega t)H_{\text{Interaction}}. \qquad (5)$$

More refined versions of the mapping Hamiltonian will be presented in subsequent sections. In all of these mapping Hamiltonians high-frequency terms are introduced without first removing terms of corresponding frequency from the Hamiltonian. Nevertheless, by the averaging principle, the new high-frequency terms are unimportant. Our new n-body maps then consist of a sequence of steps alternating pure Keplerian evolution of the individual bodies between the delta functions, with periodic interaction kicks derived from integrating the whole system across the delta functions. The basic idea is remarkably simple.

The construction of an efficient mapping for any problem rests on the ability to separate the Hamiltonian into parts which are themselves not only integrable, but efficiently computable. This looks grim at first sight for this problem. Keplerian motion is integrable, but the solution is naturally expressed only in terms of Keplerian orbital elements or one of the canonical equivalents such as the Delaunay variables. The description of the gravitational interaction of two bodies in terms of Keplerian orbital elements leads again to the expansion of the disturbing function with all the attendant complications and limitations. A map of this form would be useless. However, there is no particular reason to insist on one single set of coordinates. In fact, since the evolution for each part of the Hamiltonian is computed separately, each can be evaluated in the coordinates most suitable for that part: the Kepler orbits can be advanced in canonical Keplerian elements, and the interactions can be evaluated in canonical Cartesian coordinates, with of course the appropriate intermediate canonical transformations. There is still a better solution. The Kepler orbits can be advanced directly in canonical Cartesian coordinates using Gauss' f and g functions (see Danby 1988) without ever having to convert to Keplerian elements. This can be naturally combined with

kicks resulting from the interaction Hamiltonian evaluated directly in canonical Cartesian coordinates. It is amusing that Cartesian coordinates appear to be the best coordinates to use to take full advantage of the fact that the basic motion is Keplerian.

3. n-BODY PROBLEM

The Hamiltonian for the n-body problem is

$$H = \sum_{i=0}^{n-1} \frac{p_i^2}{2m_i} - \sum_{i<j} \frac{Gm_im_j}{r_{ij}}. \qquad (6)$$

In order to make the n-body maps this must be separated into a Keplerian Hamiltonian and an interaction Hamiltonian. A Hamiltonian is Keplerian if it can be written in the form

$$H = \frac{p^2}{2m} - \frac{\mu}{r}, \qquad (7)$$

or as a sum of such forms. Unfortunately, the n-body Hamiltonian is not immediately in the desired form. For the two-body problem, the separation of the Hamiltonian into a Kepler Hamiltonian and a noninteracting center of mass Hamiltonian is achieved by transforming to relative coordinates and center of mass coordinates. For the n-body problem, it is easy to show that a similar transformation to coordinates relative to the central mass plus center of mass coordinates does not produce a Hamiltonian which is a sum of noninteracting planetary Kepler Hamiltonians, center of mass Hamiltonian, and an interaction Hamiltonian. The problem is that with this simple choice of relative coordinates the kinetic energy is no longer a diagonal sum of squares of the new momenta. Of course, the choice of variables which accomplishes the desired transformation of the n-body Hamiltonian is well known, and is just the Jacobi coordinates (see Plummer 1960). The Jacobi coordinates can be derived by writing them as a general linear contact transformation, then requiring that the kinetic energy remain a diagonal sum of squares of the new momenta, and also that the new Hamiltonian be cyclic in the center of mass coordinate. The latter condition means that all the distances between the bodies can be written in terms of $n - 1$ of the new "relative" coordinates. Despite this formal motivation, the resulting Jacobi coordinates turn out to have a simple interpretation. We take the first coordinate to be the position of the center of mass. The first relative coordinate is just the position of the first planet relative to the central mass. The second relative coordinate is the position of the second planet relative to the center of mass of the central mass and the first planet. In general, the ith relative coordinate is the position of the ith planet relative to the center of mass of the central mass and the planets with lower indices. It is not necessary, but increasing indices are usually taken to correspond to increasing semimajor axes.

Denoting the Jacobi coordinates by a prime, the first Jacobi coordinate $\mathbf{x}_0'$ is the center of mass. The remaining $n - 1$ Jacobi coordinates are ($0 < i < n$)

$$\mathbf{x}_i' = \mathbf{x}_i - \mathbf{X}_{i-1}, \qquad (8)$$

where $\mathbf{X}_i$ denotes the center of mass of bodies with indices up to i.

$$\mathbf{X}_i = \frac{1}{\eta_i} \sum_{j=0}^{i} m_j \mathbf{x}_j, \qquad (9)$$

with the definition

$$\eta_i = \sum_{j=0}^{i} m_i. \tag{10}$$

In terms of the $\mathbf{X}_i$, the first Jacobi coordinate is simply $\mathbf{x}_0' = \mathbf{X}_{n-1}$, the center of mass of the whole system. By virtue of the requirement that the new Hamiltonian is a diagonal sum of the squares of the new momenta, the momenta conjugate to the $\mathbf{x}_i'$ have the familiar form $\mathbf{p}_i' = m_i'\mathbf{v}_i'$, where $\mathbf{v}_i'$ is the time derivative of $\mathbf{x}_i'$. The new mass factors are given by $m_i' = \eta_{i-1} m_i/\eta_i$, for $0 < i < n$ and $m_0' = \eta_{n-1} = M$, the total mass of the system. It is only a matter of algebra to show that in terms of these Jacobi coordinates the Hamiltonian for the n-body problem becomes

$$H = \frac{p_0'^2}{2M} + \sum_{i=1}^{n-1} \frac{p_i'^2}{2m_i'} - \sum_{i=1}^{n-1} \frac{Gm_i m_0}{r_{i0}} - \sum_{0 < i < j} \frac{Gm_i m_j}{r_{ij}}, \tag{11}$$

where $r_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$, the distance between bodies i and j. By construction, r_{ij} does not depend on $\mathbf{x}_0'$, thus the total momentum $\mathbf{p}_0'$ is an integral of the motion. As expected, the center of mass moves as a free particle. Hereinafter the center of mass contribution to the Hamiltonian will be omitted. Adding and subtracting the quantity

$$\sum_{i=1}^{n-1} \frac{Gm_i m_0}{r_i'}, \tag{12}$$

where $r_i' = \|\mathbf{x}_i'\|$ the Hamiltonian becomes

$$H = \sum_{i=1}^{n-1} \left(\frac{p_i'^2}{2m_i'} - \frac{Gm_i m_0}{r_i'} \right)$$
$$+ \sum_{i=1}^{n-1} \left(\frac{Gm_i m_0}{r_i'} - \frac{Gm_i m_0}{r_{i0}} \right) - \sum_{0 < i < j} \frac{Gm_i m_j}{r_{ij}}. \tag{13}$$

The second sum, which we may call the indirect perturbation, contains differences of nearly equal quantities and is actually of the same order as the direct interaction terms. The Hamiltonian now separates into a sum of $n - 1$ noninteracting Kepler Hamiltonians, and a smaller interaction Hamiltonian, as desired:

$$H = H_{\text{Kepler}} + H_{\text{Interaction}}, \tag{14}$$

with

$$H_{\text{Kepler}} = \sum_{i=1}^{n-1} \left(\frac{p_i'^2}{2m_i'} - \frac{Gm_i m_0}{r_i'} \right) \tag{15}$$

and

$$H_{\text{Interaction}} = \sum_{i=1}^{n-1} Gm_i m_0 \left(\frac{1}{r_i'} - \frac{1}{r_{i0}} \right) - \sum_{0 < i < j} \frac{Gm_i m_j}{r_{ij}}. \tag{16}$$

It is common to expand the interaction Hamiltonian in terms of the small differences between the Jacobi coordinates and the heliocentric Cartesian coordinates, and keep only the first-order corrections in the ratio of the planetary masses to the mass of the central body. In this approximation the interaction Hamiltonian becomes

$$H_{\text{Interaction}} = - \sum_{0 < i < j} \left(\frac{Gm_i m_j}{r_{ij}'} - \frac{Gm_i m_j \mathbf{x}_i' \cdot \mathbf{x}_j'}{r_j'^3} \right)$$
$$+ o(m_i^3), \tag{17}$$

with $r_{ij}' = \|\mathbf{x}_i' - \mathbf{x}_j'\|$. We have found though that maps based on the exact Hamiltonian are nearly as efficient as those based on this approximate Hamiltonian, so the expanded form will not be considered further.

An important special case of the n-body problem is obtained if some of the bodies are given infinitesimal mass. These "test particles" are perturbed by the massive planets, but do not perturb them in return. The restricted three-body problem and all its variations such as the planar elliptic restricted three-body problem fall in this category. If the test particle is given the first relative Jacobi index, below those of the massive particles, then the test particle interaction Hamiltonian is given exactly by

$$H_{\text{TestParticle}} = - \sum_{j > 1} \left(\frac{Gm_j}{r_{1j}} - \frac{Gm_j \mathbf{x}_1' \cdot \mathbf{x}_{j0}}{r_{j0}^3} \right), \tag{18}$$

where $\mathbf{x}_{j0} = \mathbf{x}_j - \mathbf{x}_0$, the vector from the central mass to body j. There are several ways of deriving this. The most straightforward method is to expand the exact interaction Hamiltonian in the differences between the Jacobi coordinates and heliocentric Cartesian coordinates, then take the appropriate limit as the test particle mass goes to zero. There are several alternate routes. The most intuitive is to note that the acceleration of the vector from the central mass to the test particle is the difference of the direct acceleration of the test particle and the acceleration of the central mass due to the gravitational attraction of the other massive bodies. This immediately gives the same test particle interaction Hamiltonian. If the test particle is given any other Jacobi index the interaction Hamiltonian is more complicated, and will not be given here. Though the equations of motion are simpler if the test particle is given a Jacobi index below the massive bodies, the orbital elements are "cleaner" if the test particle is given the natural Jacobi index in order of increasing semimajor axis along with the massive planets (see Sussman & Wisdom 1988). The resulting orbital elements are then freed of the relatively rapid oscillations due to the motion of the central mass induced by those massive planets interior to the test particle.

4. n-BODY MAPS

A simple mapping Hamiltonian for the n-body problem is then just Hamiltonian (5). It involves two distinct operations: advancing the Kepler orbits between the delta functions, and integrating the system across the delta functions. The more refined n-body maps to be presented later use the same components.

Constructing an efficient map depends on being able to rapidly advance Keplerian orbits. A summary of methods for solving this classic initial value problem is given in Danby (1988). A key element in the solutions is that the motion can be determined without explicitly determining the orientation of the orbit plane. In particular, since two vectors determine a plane, the position and velocity at any time can be written as a time-dependent linear combination of the position and velocity at the initial epoch

$$\mathbf{x}(t) = f(t)\mathbf{x}(t_0) + g(t)\mathbf{v}(t_0) \tag{19}$$

and consequently

$$\mathbf{v}(t) = \dot{f}(t)\mathbf{x}(t_0) + \dot{g}(t)\mathbf{v}(t_0), \tag{20}$$

using Gauss' famous f and g functions. Refer to Danby for a derivation of the equations which determine f and g, and their time derivatives. An important step in the determina-

tion of f and g is the calculation of the change in the eccentric anomaly, ΔE. In this paper we concentrate on problems for which the time step is a fraction of an orbit period. For this case, it is efficient to determine ΔE through the solution of the difference form of Kepler's equation

$$\Delta M = n\Delta t = \Delta E - e \cos E_0 \sin \Delta E$$
$$+ e \sin E_0 (1 - \cos \Delta E), \qquad (21)$$

where e is the orbital eccentricity, M is the mean anomaly, n is the mean motion, and E_0 is the initial eccentric anomaly. A closed-form analytic solution is not known, but the solution can be found through a variety of iterative procedures. Danby recommends a generalization of Halley's iterative method with quartic convergence, and a particular initial guess. (Watch out though, there is a typo in Danby's code on p. 167 which reduces its convergence to cubic. Also, Danby's convergence criterion is not strict enough.) For problems in which the orbits may become hyperbolic, a solution of the initial value problem in terms of universal variables is probably preferred.

The integration of the system across a delta function is trivially accomplished in canonical Cartesian coordinates since in this system the interaction Hamiltonian depends only on the coordinates. The coordinates are unchanged since the interaction Hamiltonian does not depend on the momenta. The momenta each receive a kick proportional to the generalized force, which is derived in the usual way as minus the derivative of the interaction Hamiltonian with respect to the conjugate coordinate. The differentiation is straightforward and will not be presented here. We just mention a couple of key points. Note that with an appropriate rearrangement of terms the direct contributions to all of the disturbing accelerations can be evaluated in $o(n^2)$ operations, where n is the number of planets, and all of the indirect contributions can be evaluated in $o(n)$ operations. Also, some contributions to the forces arise as small differences of nearly equal quantities. The numerical inaccuracies which would be incurred by a straightforward evaluation of these expressions can be avoided by using the same trick used to avoid a similar difficulty encountered in Encke's method (see Danby 1988). Closed-form expressions for this trick are given by Battin (1987).

5. REFINEMENT OF THE MAPPING METHOD

There are some general refinements to the mapping method that can be made. We consider Hamiltonians of the general form

$$H = H_0 + H_1. \qquad (22)$$

Both H_0 and H_1 may depend on all of the coordinates and momenta, though to make a mapping it is necessary that each part in the absence of the other part be integrable. In this paper attention is focused on problems for which $H_0 \gg H_1$. The basic idea behind the refinement of the mapping method is that rather than using a single delta function per mapping period, one mapping period can consist of a series of delta functions with possibly different amplitudes and various phases, with the amplitudes and phases chosen so that the mapping will have better properties. The property to optimize is left to our discretion. There are two obvious choices. One choice is to optimize the "order" of agreement of the Taylor series of the actual solution in time with one step of the mapping, treating the mapping as if it were a symplectic numerical integration algorithm. The other

choice is to optimize the agreement of the mapping Hamiltonian with the true Hamiltonian, making the differences have as high a frequency as possible so that the average effect of the differences will be as small as possible. Curiously, the two choices are not equivalent. Since the averaging principle is at the core of our reasoning, the second choice, to make the map Hamiltonian agree as much as possible with the true Hamiltonian, will be considered first.

High-frequency terms are added so that the corresponding mapping Hamiltonian has the form

$$H_{\mathrm{Map}} = H_0 + \Phi(\Omega t) H_1, \qquad (23)$$

where

$$\Phi(t) = 2\pi \sum_{i=0}^{N-1} a_i \delta_{2\pi}(t - 2\pi d_i). \qquad (24)$$

There are N delta functions per mapping period, with amplitudes a_i, and phases d_i which are chosen in the interval $0 \leqslant d_i < 1$. Written as a Fourier series,

$$\Phi(t) = \sum_{i=0}^{N-1} a_i \sum_{n=-\infty}^{\infty} \cos[n(t - 2\pi d_i)] \qquad (25)$$

$$= \sum_{n=-\infty}^{\infty} A_n \cos(nt) + \sum_{n=-\infty}^{\infty} B_n \sin(nt), \qquad (26)$$

where

$$A_n = \sum_{i=0}^{N-1} a_i \cos(2\pi n d_i) \qquad (27)$$

and

$$B_n = \sum_{i=0}^{N-1} a_i \sin(2\pi n d_i). \qquad (28)$$

The coefficients A_n and B_n of each of the $\cos(nt)$ and $\sin(nt)$ terms provide constraint equations for the a_i and d_i. First, the average of $\Phi(t)$ over one mapping period must be unity for the average of the mapping Hamiltonian over a mapping period to equal the true Hamiltonian. The average is given by the $n = 0$ equation, which implies simply

$$A_0 = \sum_{i=0}^{N-1} a_i = 1. \qquad (29)$$

For the two Hamiltonians to agree the coefficients of all the terms involving the mapping frequency must be zero. This gives the set of equations which determine the coefficients for each $n \neq 0$:

$$A_n = \sum_{i=0}^{N-1} a_i \cos(2\pi n d_i) = 0, \qquad (30)$$

$$B_n = \sum_{i=0}^{N-1} a_i \sin(2\pi n d_i) = 0. \qquad (31)$$

Note that if the two coefficient equations for some $n > 0$ are satisfied, the corresponding two coefficient equations for $n < 0$ are also automatically satisfied. We would like to satisfy as many of the coefficient equations as is possible, beginning with those of lowest frequency (smallest $|n|$). For a given N there are $2N$ constants to be determined. Thus, it is expected that these $2N$ constants can be chosen to satisfy at least $2N$ coefficient equations. There is only a single coefficient equation for $n = 0$. Thus it should be possible to make the coefficients up to $n = N - 1$ equal to zero, plus one of the two coefficients for $n = N$. Of course, just making one coefficient equal to zero is not very useful since other terms of that frequency will remain.

Consider the $N = 1$ map first. In this case there is only a single delta function. The coefficient $a_0 = 1$, and no useful constraint is placed on d_0. In the $N = 2$ map, it is easy to see that the two a_i must be equal, and consequently $a_i = 1/2$, and the d_i must differ by $1/2$. No useful constraint is placed on the absolute phase. For general N, a solution of the coefficient equations is that the $a_i = 1/N$ and the d_i are evenly spaced with separation $1/N$. No useful constraint is placed on the absolute phases by the coefficient equations, since the one remaining constant cannot simultaneously satisfy both of the next set of coefficient equations for $n = N$. It can be easily shown that this solution does more than satisfy the $2N - 1$ equations which determined the coefficients, but in fact it satisfies all coefficient equations for all n that are not multiples of N. It seems likely that this is the only solution to the coefficient equations, though we have not been able to demonstrate this. The solution just mentioned is very simple indeed; it consists of N evenly spaced delta functions, all of equal amplitude $1/N$. Actually, this is nothing but the original single periodic delta function with a smaller mapping period.

6. HIGHER-ORDER MAPPINGS

Another possible approach to refine the mapping method is to use the flexibility introduced through the $2N$ constants a_i and d_i in $\Phi(t)$ to match the Taylor series of a mapping step to that of the actual solution. The mapping could then be viewed as a symplectic integrator. To match the Taylor series by brute force to high-order turns out to be a formidable task, even with the aid of computer algebra. There is, however, a more abstract representation of the generalized mapping method that makes the problem of the determination of the constants more tractable.

Consider in more detail the consequences of the mapping Hamiltonian

$$H_{\text{Map}} = H_0 + \Phi(\Omega t)H_1. \tag{32}$$

Again, both H_0 and H_1 may depend on all the coordinates and momenta. Between the times when the delta functions act the time evolution of the system is governed solely by H_0. On the other hand, the evolution across the delta functions is determined solely by H_1, even though the coordinates and momenta which are being affected may also appear in H_0. This may be seen through a simple limit process in which the delta function $2\pi\delta(\Omega t)$ is represented as the limit as $\Delta \to 0$ of a step function which is nonzero in the interval $0 < t < \Delta$ with constant magnitude $2\pi/\Omega\Delta$. Taylor expanding the solution across the delta function, including terms due to both H_0 and H_1, and then taking the limit of the result as Δ goes to zero, it is easily shown that the terms coming from H_0 do not contribute. In fact, it can also be seen that the evolution of the system through the delta function is the same as if the system evolved solely according to H_1 for a time $\Delta t = 2\pi/\Omega$. The evolution through the generalized map with H_1 multiplied by a sequence of weighted delta functions of amplitudes a_i at times $d_i\Delta t$ can then be interpreted in the following way. First, the system evolves according to H_0 for a time $d_0\Delta t$, then according to H_1 for a time $a_0\Delta t$, then according to H_0 for a time $(d_1 - d_0)\Delta t$ up to the next delta function, then according to H_1 for a time $a_1\Delta t$, and so on until after the last delta function, whereupon the system evolves to the end of the mapping period according to H_0 for a time $(1 - d_{N-1})\Delta t$.

Now, it is well known that the Taylor series of a function can be written formally as an exponential

$$f(t_0 + \Delta t) = \exp\left(\Delta t \frac{d}{dt}\right)f(t)\bigg|_{t = t_0}. \tag{33}$$

Also, the total time derivative of a function of the phase space coordinates (the n coordinates and n momenta for an n degree of freedom system), and possibly also the time, can be written in terms of a Poisson bracket with the Hamiltonian which governs the time evolution of the system

$$\frac{df}{dt} = [f,H] + \frac{\partial f}{\partial t}. \tag{34}$$

Let L_H represent the Poisson bracket operator $L_H f = [f,H]$, then the time evolution of a function which depends on time only through the phase space coordinates can be written

$$f(x,p)\big|_{t_0 + \Delta t} = e^{\Delta t L_H}f(x,p)\big|_{t = t_0}. \tag{35}$$

In particular, this is true for the individual coordinate functions

$$x\big|_{t_0 + \Delta t} = e^{\Delta t L_H}x\big|_{t = t_0}. \tag{36}$$

Furthermore, for a function which depends on time only through the phase space coordinates, the time-evolved function must equal the function of the time-evolved coordinates and momenta

$$e^{\Delta t L_H}f(x,p)\big|_{t = t_0} = f(e^{\Delta t L_H}x, e^{\Delta t L_H}p)\big|_{t = t_0}. \tag{37}$$

Of course these are nothing other than the basic equations governing Lie transformations (see, e.g., Steinberg 1988).

The operator that generates Taylor series for the generalized map can then be formally written

$$e^{b_N\Delta t L_{H_0}}\cdots e^{a_1\Delta t L_{H_1}}e^{b_1\Delta t L_{H_0}}e^{a_0\Delta t L_{H_1}}e^{b_0\Delta t L_{H_0}}, \tag{38}$$

where the $b_i = d_i - d_{i-1}$, with $b_0 = d_0$ and $b_N = 1 - d_{N-1}$. Of course the exponentials cannot be simply combined since the operators L_{H_0} and L_{H_1} do not commute. (If the operators commuted, the calculation of the time evolution would be trivial.) For time evolution which is governed by the true Hamiltonian the operator that generates Taylor series is just

$$e^{\Delta t L_H} = e^{\Delta t(L_{H_0} + L_{H_1})}. \tag{39}$$

Equating these two expressions up to some order in Δt gives constraint equations which must be satisfied by the a_i and b_i. The problem of determining the constants reduces to a problem in the algebra of exponentials of noncommuting operators. Though in principle some special property of L_{H_0} and L_{H_1} might be used to simplify the determination of the constants for a particular Hamiltonian, the solutions to date have treated them simply as general noncommuting operators. Replacing $\Delta t L_{H_0}$ by B and $\Delta t L_{H_1}$ by A, the coefficients are determined by requiring

$$e^{A + B} = e^{b_N B}\cdots e^{a_1 A}e^{b_1 B}e^{a_0 A}e^{b_0 B} \tag{40}$$

be satisfied to a specified order in the products of the noncommuting operators A and B.

The solution of these equations to second order is simply achieved for $N = 1$: $b_0 = b_1 = 1/2$, and $a_0 = 1$. The result is the "generalized leap frog": a half step following H_0, followed by successive whole steps alternately following H_1 and H_0, ending with a half step of H_0. Note that in this case, the delta functions are all equally spaced in time. Thus the agreement of the mapping Hamiltonian with the true Hamiltonian, as discussed in the previous section, is maximally re-

tained. The previously noted extra freedom of the phase of the evenly spaced delta functions is being used to make the map accurate to second order, at no extra cost.

The solution of these equations to fourth order is possible with $N = 3$. This has been independently accomplished by Forest & Ruth (1990) and Yoshida (1990), also using Eq. (40). The solution is the same as that found in a more restricted formalism by Candy & Rozmus (1990) and earlier by Neri (1988). Yoshida (1990) has found sixth- and eighth-order solutions for $N = 7$ and $N = 15$, respectively. Forest (1990) has also obtained sixth-order solutions. These authors were interested in extending the symplectic integration method of Ruth (1983) to higher order. Our formulation of the coefficient equations in terms of exponentials of noncommuting operators had a different motivation and was carried out before we became aware of the work of Forest & Ruth (1990) and Yoshida (1990).

It is interesting to note that in all the known solutions of the coefficient equations some of the b_i are negative, except the first- and second-order methods (the generalized leap frog). It appears that to get higher order it is necessary to take some steps backward in time. However, even though the coefficients solve the coefficient equations, these negative steps cannot simply be represented in the scalar Hamiltonian (32) from which the coefficient equations were derived, since in the Hamiltonian only the time at which the delta functions act is specified and not any additional order of application. A Hamiltonian could be written for the higher-order maps as

$$H_{\text{Map}} = \Phi_0(\Omega t)H_0 + \Phi_1(\Omega t)H_1, \tag{41}$$

where $\Phi_0(t)$ and $\Phi_1(t)$ are similar to $\Phi(t)$, but with different coefficients. Now between the delta functions the Hamiltonian is zero and there is no evolution. The higher-order maps can be generated if the delta functions for $\Phi_0(t)$ and $\Phi_1(t)$ are interleaved with the proper amplitudes. For Hamiltonians which have been separated into two parts H_0 and H_1 of comparable magnitude, this seems to be a satisfactory solution, though it is not clear that anything is gained by representing the maps in this manner. On the other hand, for Hamiltonians for which H_0 is much larger than H_1, a mapping Hamiltonian which introduces high frequencies proportional to H_0 seems like a bad idea. The first corrections to the evolution as deduced by, for example, Von Zeipel transformations will be large. Perhaps there is a better way to represent the Hamiltonian for the higher-order methods. An interesting possibility is to search for higher-order methods using Hamiltonian (32), but with the added constraint that all the b_i be positive. We have attempted to find two and three delta function maps of third order with all positive b_i, but were unsuccessful. It is important to determine whether or not such solutions exist in general.

At present then, there are two options available: (1) If the existence of an explicit Hamiltonian is desired, and the agreement of the map Hamiltonian with the true Hamiltonian is to be maximized, then the best solution is to take all the delta functions to be of equal amplitude and evenly spaced. Without any degradation of the agreement of the two Hamiltonians, the map can be made second order by choosing the phase of the delta functions so that the calculation begins and ends with a half step of H_0. (2) If the order of the mapping is to be increased beyond second order, any of a number of known solutions of the coefficient equations can be used. Apparently though, higher order comes at the cost

of losing, to some extent, our original motivation which connects a mapping Hamiltonian via the averaging principle to the true Hamiltonian, particularly for a system for which $H_0 \gg H_1$. Nevertheless, either of these alternatives can be used with the n-body mapping components described here.

The n-body mapping is then a composition of individual Kepler steps for each of the planets with kicks derived from the interaction Hamiltonian, each appropriately weighted to form either the generalized leap-frog version of the map or a higher-order version of the map. Note that since each of the Keplerian orbits is advanced independently of all the others, they can be advanced in parallel.

7. OTHER SYMPLECTIC INTEGRATORS

It is instructive to compare the generalized maps described here with other symplectic integrators. The symplectic integration scheme of Ruth (1983) (see also de Vogelaere 1956) is based on a time step given in terms of a mixed variable generating function,

$$F(x,p') = xp' + S(x,p'), \tag{42}$$

where the primed variables are the new variables after the time step. Nothing here depends on the dimension of the phase space, so for simplicity we use a single degree of freedom. Choosing $S(x,p') = \Delta t H(x,p')$, this generating function gives the canonical transformation of variables

$$p = \frac{\partial F}{\partial x} = p' + \Delta t \frac{\partial H(x,p')}{\partial x}, \tag{43}$$

$$x' = \frac{\partial F}{\partial p'} = x + \Delta t \frac{\partial H(x,p')}{\partial p'}. \tag{44}$$

This step is canonical and only approximates the evolution of the system under $H(x,p)$ to first order in Δt. Higher-order generating functions have been derived by Ruth (1983), Menyuk (1984), and Channel & Scovel (1988). In general, the transformation from x and p to x' and p' is only given implicitly, since it is based on a mixed variable generating function. In the special case where the Hamiltonian can be written in the form

$$H(x,p) = T(p) + V(x), \tag{45}$$

the transformation can be written explicitly:

$$p' = p - \Delta t \frac{\partial V(x)}{\partial x}, \tag{46}$$

followed by

$$x' = x + \Delta t \frac{\partial T(p')}{\partial p'}. \tag{47}$$

Ruth (1983) found higher-order symplectic integrators both by using higher-order generating functions for the step, and by composing low-order steps and adjusting constants to achieve higher order. The latter method is followed by Neri (1988) and Candy & Rozmus (1990) for Hamiltonians of the form of Hamiltonian (45). Forest & Ruth (1990) also describe their method in terms of a composition of steps achieved through a mixed variable generating function.

The mapping method has been described in terms of the averaging method which originally motivated it, and implemented in terms of mapping Hamiltonians containing periodic sequences of delta functions. This is quite different from the generating function description of the symplectic integrators following Ruth (1983). Nevertheless, mappings derived along the lines Wisdom (1982) are closely related to

the symplectic integration methods derived from Ruth (1983). In fact, for Hamiltonians of the form of Hamiltonian (45), they are identical. Consider a simple mapping Hamiltonian for Hamiltonian (45)

$$H_{\mathrm{Map}} = T(p) + 2\pi\delta_{2\pi}(\Omega t)V(q). \qquad (48)$$

Integrating across the delta function gives

$$p' = p - \Delta t \frac{\partial V(x)}{\partial x}, \qquad (49)$$

with $\Delta t = 2\pi/\Omega$. Then integrating between the delta functions gives

$$x' = x + \Delta t \frac{\partial T(p')}{\partial p'}. \qquad (50)$$

These are the same equations as those obtained with the mixed variable generating function. However, in the more general case in which the Ruth method will require the solution of implicit equations the mapping method presented here will give explicit equations. The two methods are then not equivalent.

It is clear that the mapping method based on periodic delta functions can also be thought of as a method of symplectic integration. It is sometimes the same as that derived from a generating function and sometimes distinct. In a sense, though, the mapping based on delta functions is more clearly related to the original system than is the symplectic integration step based on generating functions. In the Ruth school, the basic idea is that a finite difference method that is exactly canonical may be subject to fewer artifacts than a finite difference method that is not canonical. For instance, a noncanonical finite difference scheme may have attractors, in contradiction to the well-known fact that Hamiltonian systems do not have attractors. The finite difference scheme then has a possible behavior that cannot belong to the real system. The Ruth integrators are symplectic as the actual system they are modeling is symplectic, and for small enough time step the error in the step becomes arbitrarily small, but whether the symplectic integrator should give a good approximation to the long-term dynamics is not clear. This point was reiterated by Zhong & Marsden (1988). One method of deriving these symplectic integrators (see Forest & Ruth 1990) is to make a chain of canonical transformations as described above and then require that the Hamiltonian, when expressed in terms of the new variables, matches the actual Hamiltonian to some order. Since the match is not perfect, each individual step of a symplectic integrator makes some error in the Hamiltonian. There is no guarantee, nor any reason to expect, that the repeated composition of such steps will not lead further and further from the true Hamiltonian, even though the composition is canonical. However, this catastrophe does not seem to happen, though the reason remains unclear. On the other hand, the mappings derived from Wisdom (1982) are explicitly derived from a Hamiltonian. The differences between the true Hamiltonian and the mapping Hamiltonian are arguably unimportant to the long-term behavior on the basis of the averaging principle. Furthermore, the differences are explicit. In some cases it is possible to derive a correction from mapping variables to true variables by eliminating the extraneous high-frequency terms by Von Zeipel transformations (e.g., Tittemore & Wisdom 1988). The basic Ruth step is canonical, but it is given in terms of a generating function, not as the time evolution of a Hamiltonian [though Menyuk (1984) calls the generating function a "discrete Hamiltonian"]. It is

not clear to us whether or not a system Hamiltonian can be written for the implicit Ruth maps.

8. SIMPLER n-BODY MAPS

There is a simpler way to get symplectic maps for the n-body problem. Note that the basic n-body Hamiltonian (6) is in the form of Hamiltonian (45). Thus a particularly simple map for the n-body problem can be obtained by letting H_0 be the kinetic energy and H_1 be the potential energy. Then any of the mapping Hamiltonians we have discussed can be used. In particular, high-order mappings can be made with Hamiltonian (41). The second-order form of the map, the generalized leap frog, becomes in this case simply the ordinary leap-frog integrator.

There are two obvious disadvantages of these simple n-body maps. The most severe is that the basic Keplerian motion of the orbits is not taken into account. Thus the number of steps must be large enough to stably and accurately negotiate each passage of the planet around the Sun. A second disadvantage is that the motion of the central mass must also be integrated. It should be noted that Cowell's form of the equations of motion, which are written in terms of simple relative coordinates with respect to the central mass (see Brouwer & Clemence 1961), are not in Hamiltonian form. A map naively based on them would not have the desired property of being symplectic.

Higher-order maps of this simple form have already been applied to the n-body problem by Gladman & Duncan (1990) and Kinoshita et al. (1990). Note that in both of these applications the mappings are used in what may be called the "qualitative" mode of operation. The step size is relatively large and the truncation error is much larger than the machine precision. Of course, this may not matter for qualitative investigations, since the truncation error does not seem to accumulate in a bad way for these symplectic maps, and the energy error is observed to oscillate, at least locally. Gladman and Duncan take 300 steps per orbit; Kinoshita et al. take about 600 steps per orbit. The relative error in the value of the Hamiltonian is of order 10^{-9} for Kinoshita et al., and somewhat larger in the other study. In both cases the error is much larger than the available precision of the machine and is obviously dominated by truncation error. The simple fourth-order symplectic integrators used in these papers evaluate the accelerations three times per step. Thus they require about 2000 function evaluations per orbit to achieve a local relative energy error of 10^{-9}. These errors should be compared with the errors in a conventional numerical integration. For example, for the twelfth-order Stormer predictor, which was used in the Digital Orrery calculations, the truncation error is of order the machine precision with around 100 steps per orbit. The Stormer predictor evaluates the accelerations only once per step. Even after nearly a billion years the relative energy error in the Digital Orrery integration was only of order 10^{-10}, and locally the variation of the relative energy is much smaller. Even used in the qualitative mode of operation the simple symplectic integrators do not appear to be competitive with traditional integrators.

Consider the use of the simple n-body maps in a "high accuracy" mode of operation in which the step size is chosen to be small enough so that the truncation error is of order the machine precision. Assuming the truncation error is proportional to the fifth power of the step size for these fourth-order methods, for the relative energy error to reach machine pre-

cision (which we take to be about 10^{-16}) requires about 45 000 function evaluations per orbit. Of course, higher-order methods need to take significantly fewer steps per orbit. Suppose the relative energy error from truncation can be written $\Delta = C(h/N)^{o+1}$, where h is the step size divided by the orbital period, o is the order, N is the number of function evaluations per step, and C is an error constant. We presume that the error constant in this form is comparable for all of the higher-order methods; the factor of N is just a guess, and works in favor of the higher-order methods. Using this estimate we find that even for the eighth-order method of Yoshida (with 15 function evaluations per step), achieving a relative energy error of order the machine precision requires 400 function evaluations per orbit. Thus even the high-order versions of the simple n-body maps may still be inefficient compared to traditional high accuracy integrators. Of course, the relative inefficiency may be outweighed by a better long-term growth of error. To our knowledge the long-term growth of error for the simple symplectic n-body integrators has not yet been carefully examined, particularly in the "high accuracy" mode of operation where the truncation error is of order the machine precision.

Consider in the same manner the possibility of using the n-body maps described in this paper in the "high accuracy" mode. From a numerical integration point of view, the basic difference between these methods and the simple methods just described is that the error constant in these new maps may be expected to be smaller by about the ratio of the planetary masses to the central mass μ. For our solar system, μ is about 10^{-3}. The number of steps per orbit required to achieve the same truncation error as the simple maps is smaller by a factor of $\mu^{-1/(o+1)}$. For a fourth-order method with $\mu = 10^{-3}$, this factor is only about 4. For the eighth-order method, it is about 2. Considering the fact that the steps in the Kepler-based n-body maps are a little more expensive than those in the simple n-body maps, it is not clear that any advantage is gained by using the maps presented here over the simpler maps, at least in the "high accuracy" mode of operation. However, there may be an advantage to using the Kepler-based maps for orbits with high eccentricity. In this case, the simple n-body maps must take many more steps per orbit to stably and accurately execute the orbit, since the basic Kepler motion must be integrated as well. On the other hand, the n-body maps presented in this paper exactly represent a pure Kepler orbit at any eccentricity. Tests in the circular and elliptic restricted problems indicate the Kepler based n-body maps suffer no significant loss of stability or accuracy at high eccentricity. In this case there may be a significant advantage in using them over the simple maps even in the "high accuracy" mode.

On the other hand, consider the use of the n-body maps introduced in this paper in the "qualitative" mode of operation. Typically, efficient traditional integrators take on the order of 100 steps per orbit. We have found that in solar system integrations the qualitative behavior is reliably reproduced with as few as ten steps per orbit. Such a small number of steps per orbit is stable here because the Kepler motion is represented exactly and does not have to be rediscovered each orbit. The reduction in the number of function evaluations by a factor of 10 accounts roughly for the order of magnitude greater speed of the new mapping method over traditional integrators. The new n-body maps are the clear winners for qualitative studies.

Of course, the relative merits of the various methods in the two different modes of operation should be studied more thoroughly to check the estimates given here.

9. THE OUTER PLANETS FOR A BILLION YEARS

We have carried out numerous tests of the new n-body maps. First, a number of surfaces of section for the circular restricted three-body problem were computed with the new map and compared to sections computed with the conventional Bulirsch–Stoer numerical integration algorithm. The agreement was excellent and provided valuable initial experience with the new maps. These tests demonstrated the reliability and efficiency of the map at high eccentricity. The n-body maps have also been implemented for the planar elliptic restricted three-body problem. The numerical integrations reported in Wisdom (1983), which also used the conventional Bulirsch–Stoer algorithm, were all repeated with the map, with particular attention to whether the map would give the correct diagnosis of whether the trajectory was chaotic or quasiperiodic. In every case, the map agreed with the earlier results. Of course, the jumps in eccentricity were also recovered. Note that the codes for the various versions of the restricted three-body problem can be written to take advantage of the known fixed orbit of the two massive bodies. Rather than present these initial tests in detail, we present a much more stringent test. We have used the map to compute the evolution of the outer planets, including Pluto as a test particle, for about 1.1 billion years. For this problem the evolution has already been computed for 845 million years using conventional integration techniques on the Digital Orrery (Sussman & Wisdom 1988), and comparison can be made to those results.

We have chosen to use the second-order version of the mapping, which optimizes the agreement of the mapping Hamiltonian with the true Hamiltonian in accordance with our original motivation based on the averaging principle. We have used the exact form of the interaction Hamiltonian, and Pluto is given a Jacobi index below those of the massive planets. Of course, in order to make comparisons the initial conditions and parameters must be the same as those used in the Digital Orrery integrations (Applegate et al. 1986). The only parameter left to choose is the step size, or mapping period. The map is used in the "qualitative" mode and the step size is chosen to be relatively large. A number of preliminary tests indicate that the map does not work well for this problem if fewer than five steps are taken per Jupiter orbit period, which is about 12 yr. To add a margin of safety, a step size of 1 yr was chosen. This may be compared to typical steps of 40 days or less that have been used in other studies of the outer planets using conventional numerical integration techniques. The relative energy error oscillates as expected, and, using this step size, has a rather large peak to peak amplitude of about 10^{-5}. The map is remarkably fast. A billion year evolution of the outer planets takes only 14 days on a Hewlett-Packard HP9000/835 RISC workstation.

All of the principal results of Sussman & Wisdom (1988) are reproduced in the mapping evolution. For example, the argument of perihelion of Pluto again displays a 34 million yr modulation. The quantity $h = e \sin \varpi$, where ϖ is the longitude of perihelion, displays its strong 137 million yr period. This is illustrated in Fig. 1 which is to be compared with h as computed using the Stormer multistep predictor on the Digital Orrery, shown in Fig. 2. The two plots are not identical, but the similarity is astounding. The inclination of Pluto
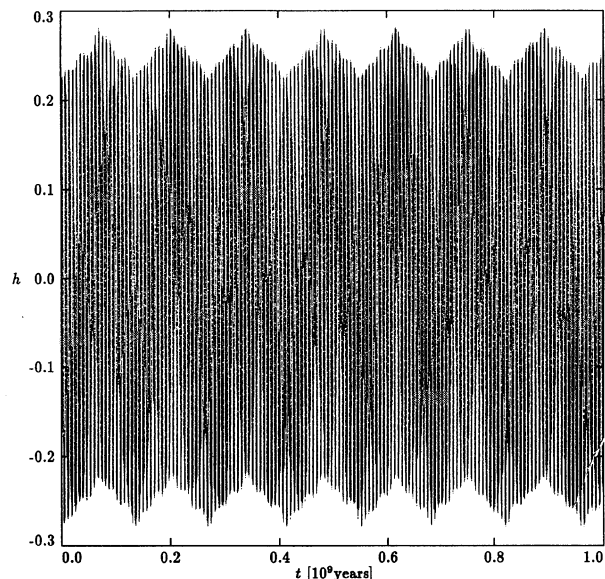
FIG. 1. The orbital element $h = e \sin \varpi$ for Pluto as computed with the mapping.
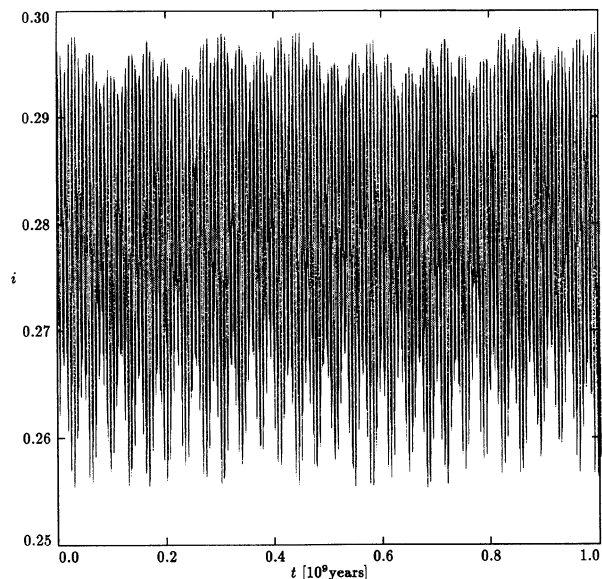

FIG. 3. The inclination of Pluto for a billion years as computed with the mapping.

again displays even longer periods. Figure 3 presents the inclination of Pluto for a billion years as computed with the map; and, for comparison, Fig. 4 presents the inclination of Pluto from the Digital Orrery computation. In the region of overlap the two plots are again remarkably similar. The observed differences can probably all be attributed to the different sampling times in the latter part of the Digital Orrery computation, and to slightly different frequencies in the two evolutions. Even used in the "qualitative" mode, these n-body maps are remarkably good. It is interesting that at the end of the mapping calculation, which was longer than the

Orrery integration, the inclination reaches a new maximum which gives the impression of a secular increase and at least indicates the presence of periods longer than a billion years in the motion of Pluto. The motion of Pluto appears to be inexhaustible, a property which is consistent with the numerical evidence that the motion of Pluto is chaotic.

Finally, the chaotic character of the motion of Pluto has been confirmed using the mapping method. The divergence of nearby trajectories has been computed using the two-trajectory method. The plots are qualitatively the same as the one shown in Sussman & Wisdom (1988). In that study the final Lyapunov exponent calculation was begun halfway
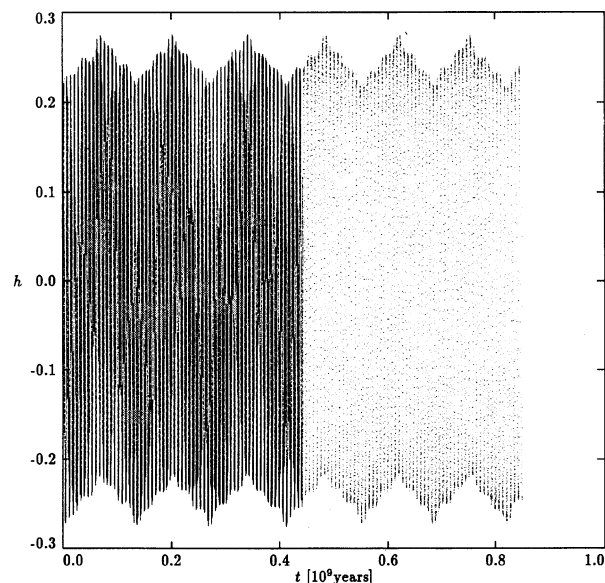

FIG. 2. The orbital element $h = e \sin \varpi$ for Pluto as computed with a conventional numerical integration method on the Digital Orrery. The sampling frequency was decreased in the latter part of the run.
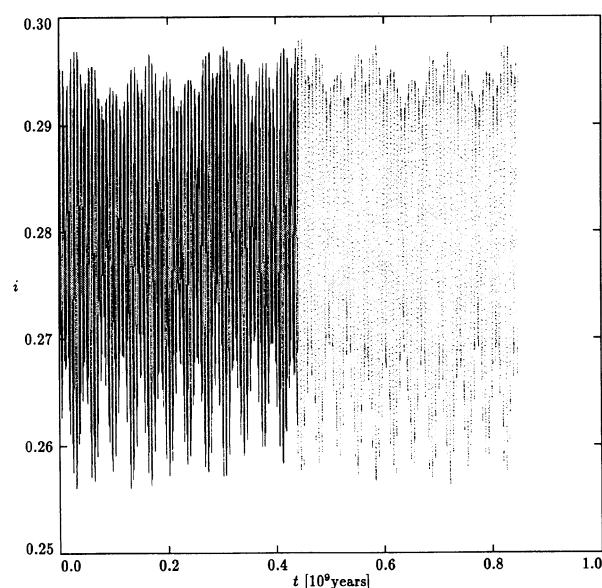

FIG. 4. The inclination of Pluto for 845 million years as computed with a conventional numerical integration method on the Digital Orrery.

through the computation; here the Lyapunov exponent calculation was started at the beginning of the run. Since the two calculations are not directly comparable, they will not be shown here. The timescale for exponential divergence is again about 20 million yr. This confirmation of the positive Lyapunov exponent on a different computer with a different length floating point mantissa, with such a different integration method, using a somewhat arbitrarily chosen step size which is an order of magnitude larger than the special step size used in the Digital Orrery integrations, considerably strengthens the conclusion of Sussman & Wisdom (1988) that the motion of Pluto is chaotic. Of course, it is probably wise to remain a little suspicious until the dynamical mechanisms are properly understood.

This is now the longest evolution of the outer planets to date. The remarkable agreement of this evolution with that computed with the Digital Orrery is a strong testimony to the validity of the averaging principle, and to the usefulness of the new *n*-body maps.

REFERENCES

Applegate, J. H., Douglas, M. R., Gursel, Y., Sussman, G. J., and Wisdom, J. 1986, AJ, 92, 176

Arnold, V. I. 1974, Mathematical Methods of Classical Mechanics (Springer, New York)

Battin, R. H. 1987, An Introduction to the Mathematics and Methods of Astrodynamics (American Institute of Aeronautics and Astronautics, New York)

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

Candy, J., and Rozmus, W. 1990, preprint

Channell, P. J., and Scovel, J. C. 1988, Los Alamos National Laboratory, LA-UR-88-1828

Chirikov, B. V. 1979, Phys. Rep., 52, 263

Cohen, C. J., Hubbard, E. C., and Oesterwinter, C. 1973, Astron. Papers Amer. Ephemeris, 22, 1

Danby, J. M. A. 1988, Fundamentals of Celestial Mechanics (Willmann-Bell, Richmond)

DeVogelaere, R. 1956, University of Notre Dame, Report 4

Forest, E. 1990, preprint

Forest, E., and Ruth, R. D. 1990, Physica D, 43, 105

Gladman, B., and Duncan, M. 1990, preprint

Kinoshita, H., and Nakai, H. 1984, Celest. Mech., 34, 203

Kinoshita, H., Yoshida, H., and Nakai, H. 1990, in Proceedings of the Twenty-Third Symposium on Celestial Mechanics, edited by H. Kinoshita and H. Yoshida (Kyoto), p.1

Laskar, J. 1989, Nat, 338, 237

Laskar, J. 1990, preprint

Malhotra, R. 1990, Icarus, 87, 249

Malhotra, R., and Dermott, S. 1990, Icarus, 85, 444

Menyuk, C. R. 1984, Physica, 11D, 109

Murray, C. D. 1986, Icarus, 65, 70

Murray, C. D., and Fox, K. 1984, Icarus, 59, 221

Neri, F. 1988, preprint

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

Quinn, T. R., Tremaine, S., and Duncan, M. 1991, preprint

Richardson, D. L., and Walker, C. F. 1987, in Astrodynamics 1987, Advances in the Astronautical Sciences, edited by J. K. Soldner, A. K. Misra, R. E. Lindberg, and W. Williamson (San Diego, Univelt), Vol. 65, p. 1473

Roy, A. E., et al. 1988, Vistas in Astronomy, 32, 95

Ruth, R. D. 1983, IEEE Trans. Nucl. Sci., NS-30, 2669

Sidlichovsky, M., and Melendo, B. 1986, Bull. Astron. Inst. Czechosl., 37, 65

Steinberg, S. 1988, in Lie Methods in Optics, edited by Mondragon and Wolf (Springer-Verlag, New York), p. 45

Sussman, G. J., and Wisdom, J. 1988, Sci, 241, 433

Tittemore, W. C. 1990, Sci, 250, 263

Tittemore, W. C., and Wisdom, J. 1988, Icarus 74, 172

Tittemore, W. C., and Wisdom, J. 1989, Icarus 78, 63

Tittemore, W. C., and Wisdom, J. 1990, Icarus 85, 394

Wetherill, G. 1985, Meteoritics, 18, 1

Wisdom, J. 1982, AJ, 87, 577

Wisdom, J. 1983, Icarus, 56, 51

Wisdom, J. 1985a, Nat, 315, 731

Wisdom, J. 1985b, Icarus, 63, 272

Wisdom, J. 1987, Icarus, 72, 241

Yoshida, H. 1990, Phys. Lett., A, 150, 262

Zhong, G., and Marsden, J. E. 1988, Phys. Lett. A, 133, 134