LONG TERM EVOLUTION OF THE SOLAR SYSTEM

JACK WISDOM

Massachusetts Institute of Technology

Abstract. The mapping method of Wisdom (1982) has been generalized to encompass all n-body problems with a dominant central mass (Wisdom and Holman, 1991). The new mapping method is presented as well as a number of initial applications. These include billion year integrations of the outer planets, a number of 100 million year integrations of the whole solar system, and a systematic survey of test particle stability in the outer solar system.

1. Introduction

A new mapping method for the numerical integration of planetary orbits is presented in Wisdom and Holman (1991). This mapping method is typically about an order of magnitude faster than other methods of integration which have been used in studies of the long-term evolution of planetary orbits. Here, the basic idea of the new n-body mappings is first presented. This is followed by a presentation of new billion year integrations of the outer planets carried out with the new mapping. The new integrations are compared to the 845 million year integration of the outer planets performed with a conventional integration technique on the Digital Orrery. Preliminary results of a number of 100 million year integrations of the whole solar system are then presented, followed by a quick summary of a systematic survey of test particle stability in the outer solar system.

2. Mapping Method

A complete presentation of the new mapping method can be found in Wisdom and Holman (1991). The new mapping method is a generalization of the mapping method of Wisdom (1982). In the original mapping method the Hamiltonian was analytically separated into parts with different associated timescales

$$H = H_{Kepler} + H_{Orbital} + H_{Resonant} + H_{Secular}. \tag{1}$$

The "Kepler" terms describe the basic Keplerian motion of the planets with respect to the Sun. The "Secular" terms include those terms which do not depend on the mean longitude. The "Resonant" terms contain resonant combinations of the mean longitudes. The "Orbital" terms represent all other terms that depend on the mean longitudes. The mapping method is motivated by the averaging principle. It is well known that the long-term evolution of planetary systems is often well described by the averaged system, obtained analytically by removing the "Orbital" terms, those terms which are fast, but not resonant. The idea of the mapping method is that if terms with periods of order of the orbital period can be removed without significantly affecting the long-term evolution then, equally well, other terms with those frequencies can be innocuously added. So new high-frequency terms are added so the resonant terms become multiplied by periodic delta functions. Symbolically, the Hamiltonian for the original mapping method is then

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

17

where $\delta_{2\pi}(t)$ represents a periodic sequence of delta functions with period 2π and Ω is the mapping frequency. The introduction of delta functions allows the equations of motion to be solved locally. The "Kepler" terms and the "Secular" terms can be analytically integrated between the delta functions provided that "Secular" fourth order terms in the eccentricities and inclination are ignored. The "Resonant" terms can be integrated across the delta functions. Thus, analytical expressions can be derived for the local evolution of the system through one mapping period. The evolution of the system is obtained by iterating this mapping. Refer to Wisdom (1982, 1983) for more details. The mapping is very fast and has, now, been successfully used in a large number of studies of resonant planetary and satellite evolution.

Despite the success and widespread use of these resonance mappings, they have significant limitations. First, they are not valid at high eccentricity and inclination, because the derivation of the mappings rely on an analytic representation of the disturbing function which must be truncated at some order in the eccentricities and inclinations. Also, they are valid only near the particular commensurabilites or group of commensurabilites included in the "Resonant" terms. Thus, the systematic study of the motion near a large number of commensurabilities requires a new derivation of the mapping near each set of commensurabilities, a very tedious job indeed. The new mapping method does not have these limitations. The cost is that the new mapping method is not as fast as the resonant mapping method, but it is still significantly faster than traditional methods of numerical integration.

There are two key ideas in the new n-body mapping method. First, the new mapping method is based on a simpler separation of the Hamiltonian for the n-body problem:

$$H = H_{Kepler} + H_{Interaction}. \tag{3}$$

Here again, the "Kepler" terms are the terms which represent the basic Keplerian motion of the bodies relative to the dominant central mass. The "Interaction" terms are the other terms. There are several different forms for the new mapping method. The simplest form of the new mapping method is obtained by adding new high-frequency terms to the n-body Hamiltonian so that it becomes

$$H_{Map} = H_{Kepler} + H_{Interaction} 2\pi \delta_{2\pi}(\Omega t) \tag{4}$$

The Kepler Hamiltonian is integrable between the delta functions and the interaction terms are integrable across the delta functions. This time no high-frequency terms are first removed, but by the same reasoning as before, the new high-frequency terms do not affect the long-term evolution.

This idea alone does not make an efficient mapping. The reason is that the Kepler problem is naturally solved only in terms of Keplerian orbital elements or some canonical equivalent such as the Delaunay variables. Expressing the interaction in terms of the Keplerian orbital elements would lead again to all the troubles associated with the expansion of the disturbing function. The interaction terms are simple only in one of the Cartesian coordinate systems. A possible solution is to evaluate the Kepler evolution in terms of the canonical Delaunay variables, make a canonical transformation to canonical heliocentric or Jacobi coordinates

and evaluate the interaction, then make another canonical transformation back to Delaunay variables. The intermediate transformations are cumbersome and the resulting mapping method is far from optimal. The second key idea in the new mapping method is to evaluate the Kepler evolution using Gauss' f and g functions directly in Cartesian coordinates. The solution of the initial value problem in terms of Gauss' functions is another classic problem (for a review, see Danby, 1988). The Kepler evolution in Cartesian coordinates is trivially made canonical. Intermediate canonical transformations are thus avoided.

The basic idea of the new n-body mappings is remarkably simple. The mapping consists of alternating steps of Kepler evolution of the individual orbits interspersed with interaction kicks. The mapping is efficient if Cartesian coordinates are used in both phases of the calculation.

Accomplishing the separation of the Hamiltonian for the n-body problem into the Keplerian parts and the interactions parts is a classic problem, and can be accomplished in either Jacobi coordinates or "canonical heliocentric" coordinates. Wisdom and Holman (1991) used Jacobi coordinates, erroneously dismissing the canonical heliocentric coordinates. The relative merits of the mappings in these different coordinate systems have yet to be carefully examined. Jacobi coordinates have been used in the applications presented here.

Several improvements can be made in the simple mapping method just described. For example, the mapping method can be made accurate to arbitrarily high order in the mapping stepsize. This and other refinements are discussed in Wisdom and Holman (1991). In the applications presented here, the second order version of the n-body mapping is used. Second order is achieved by evolving the system with the Kepler Hamiltonian for a half mapping step, followed by an alternating succession of full interaction kicks and whole Keplerian steps, but ending with a half Kepler step.

Though the new mapping method has been found as a generalization of the Wisdom (1982) mapping method, it is clear that it is also an example of a symplectic integration scheme. In fact, the mapping method of Wisdom (1982) can also be viewed as a symplectic integrator for the averaged resonance problem. It was not called this because the term "symplectic integration" is a later invention.

3. Billion Year Evolution of the Outer Planets

A large number of tests of the new n-body mapping method have been carried out. The tests include evolutions of the circular and elliptic restricted three-body problems. Surfaces of section for the circular restricted problem have been compared with surfaces of section computed with traditional integration schemes. Lyapunov exponents in the planar elliptic problem have been compared with Lyapunov exponents reported in Wisdom (1983), computed with a traditional integration scheme. In all cases the agreement was excellent and confirmed the usefulness of the new mapping method, even at high eccentricity. The most stringent test has been the calculation of the evolution of the outer planets for a billion years. The stepsize or mapping period in the new integration is about 1 year. By any measure, the longest and most accurate existing integration of a planetary system to date is the

845 million year evolution of the outer planets computed on the Digital Orrery reported by Sussman and Wisdom (1989). The nearest competitor is the ±50 million year integration of project LONGSTOP (Roy, 1990), which has approximately three orders of magnitude greater rate of growth of energy error. The evolution of the outer planets computed with the mapping agrees astonishingly well with the Digital Orrery evolution. The 137 million year period in the evolution of $h = e \sin \varpi$ is recovered. The longer period variations in the inclination of Pluto are reproduced in detail. That the new mapping can accurately reproduce such long period variations is very impressive. The most astounding feature of these evolutions is however the relative computation time: the calculation took about three months of time on the Digital Orrery, a computer built specifically for the task running at about a third the speed of a Cray 1, whereas the new mapping evolution was computed in less than two weeks on an ordinary Hewlett-Packard RISC workstation.

The Lyapunov exponent for the motion of Pluto found in the new calculation is consistent with the one found in the Digital Orrery calculation. This considerably strengthens the earlier result that the motion of the planet Pluto is chaotic. The new calculation was performed with a different integration method, with a drastically different stepsize which was not specially chosen, and on a different computer with a different word size.

4. 100 Million Year Evolution of the Solar System

Following the calculation of Sussman and Wisdom (1989) that provided evidence of the chaotic motion of Pluto, Laskar (1989, hereafter La89) integrated the whole solar system (excluding Pluto) for ±100 million years and found numerical evidence that the evolution is chaotic. The timescale for exponential divergence is only 5 million years. Laskar integrated a set of averaged equations with a very large number of terms (over 100,000) determined through extensive computer algebra. Laskar's system is perturbative and necessarily truncated in eccentricity, inclination, and mass. It is obviously quite important to confirm Laskar's results in direct integrations.

Direct integrations are however very time consuming, the longest direct integration of the whole solar system, performed earlier this year by Quinn, Tremaine, and Duncan (1991, hereafter QTD91) is only three million years. With the goal of testing Laskar's results, the mapping method has been used to integrate the evolution of the whole solar system for nearly 100 million years (see Sussman and Wisdom, 1992, hereafter SW92, for full details). This calculation was carried out on the successor to the Digital Orrery, the Supercomputer Toolkit, another special purpose parallel computer designed by Gerald J. Sussman (MIT), with William McAllistor (Hewlett Packard) and colleagues. The Toolkit is about 20 times faster than the Digital Orrery, and roughly has the scalar speed of the Cray X/MP. Eight simultaneous but independent calculations of the solar system evolution for 100 million years, using the new n-body mapping method, took about 1000 hours on the Toolkit. In order to have a point of comparison, the physics and initial conditions were chosen to be the same as those in the three million year integration of QTD91, which included relativistic and Earth-Moon quadrupole corrections, and took initial conditions from DE102. Even though the stepsize in this new mapping

TABLE I
Maximum differences in the eccentricities.

		La89 – QTD91			QTD91 – SW92	
Mercury	0.0041	0.000018
Venus	0.0020	0.000065
Earth	0.0024	0.000059
Mars	0.0041	0.000132
Jupiter	0.0038	0.000047
Saturn	0.0081	0.000162
Uranus	0.0051	0.000008
Neptune	0.0026	0.000002
Pluto	–	0.000001

integration is an order of magnitude larger than that of QTD91 the agreement is quite good. The maximum difference in the argument of perihelion of Mercury between QTD91 and SW92 over this three million year interval is of order 0.0001 radians; for comparison, the precession of the argument of perihelion due to general relativity is about 2π radians over three million years. As another comparison, Table 1 lists the maximum differences between the eccentricities of the planets in the three integrations.

The table illustrates that the mapping agrees quite well with a conventional direct integration. The comparison between La89 and QTD91 is from Laskar, Quinn, and Tremaine (1991). Obviously, QTD91 and SW92 are mutually more consistent than either is to La89, though it is not clear whether this discrepancy is due primarily to model differences or to the effects of truncation in La89. The evolution of Pluto over the full 100 million years has characteristics quite similar to those found in the outer planets integrations such as the 34 million year oscillation of the argument of perihelion of Pluto.

Several different estimates of "Lyapunov exponents" have been computed. The divergence of individual planets gives an exponent for that planet. The exponent for Pluto is again consistent with the Digital Orrery calculation and the new billion year mapping calculations, and gives an inverse Lyapunov exponent in the range of 15-20 million years. This confirms that the evolution of Pluto is chaotic, and furthermore that the chaotic character of the evolution of Pluto is not a sensitive function of the model. For the whole solar system, the conventional Lyapunov exponent of the full phase space does not give a clear indication of chaos. Rather, it seems to be dominated by a quasiperiodic component in the divergence, and only at the end of the calculation does exponential divergence begin to dominate. Exponential divergence is clearer in the individual planet divergences. In order to reduce the quasiperiodic component in the divergence, a new "secular Lyapunov exponent" is introduced. The secular Lyapunov exponent is calculated by first converting the positions and momenta into Keplerian elements, then the standard shape and orientation variables $h = e \sin \varpi$, $k = e \cos \varpi$, $p = \sin i/2 \sin \Omega$, and

$q = \sin i/2 \cos \Omega$ are formed. The full set of these variables are then used to define the "secular phase space," thus ignoring the mean longitude and semimajor axis dimensions. The new secular phase space divergence appears to be more diagnostic of the dynamics. The secular divergence appears to have two distinct exponential timescales. There is an initial exponential divergence with a timescale of about 20 million years, and another which appears later with a timescale of about 5 million years. The two exponential timescales suggests that there may be more than one dynamical mechanism responsible for the chaotic evolution of the solar system.

Laskar (1990) has found two resonance angles which appear to be alternately librating and circulating. In our calculation they also alternately circulate and librate, though of course the detailed behavior is different since the evolution is chaotic.

The fact that essentially the same results are found in two such strikingly different calculations as those of La89 and SW92, strongly suggests that the solar system is indeed chaotic.

5. Test Particle Surveys

Several thousand 20 million year integrations of test particles in the outer solar system have been carried out. Initial conditions studied include circular orbits between the planets as well as circular or nearly circular orbits at the triangular Lagrange points of the major planets. The basic results of this 20 million year study are (1) by and large most orbits between the planets are unstable and (2) there exist extensive stable regions at all the Lagrange points of the major planets, though the stable region is not in all cases simple. Details of these results will be published elsewhere (Holman and Wisdom)

6. Conclusions

The new mapping method for the n-body problem of Wisdom and Holman (1991) is a powerful new tool for studying the long term evolution of planetary and satellite systems.

New billion year evolutions of the outer planets confirm the result of Sussman and Wisdom (1989) that the motion of the planet Pluto is chaotic with a divergence timescale of 15-20 million years.

Hundred million year integrations of the whole solar system confirm Laskar's (1989) result that the evolution of the solar system is chaotic with a divergence timescale of about 5 million years. Alternate circulation and libration of two resonance angles provide additional evidence that the evolution is chaotic, and suggests that the chaos may be a consequence of competing secular resonances among the inner planets.

Acknowledgements

M. Quinn, M. Duncan, and J. Laskar. The research reviewed here was supported in part by grants from NSF and NASA.

References

Danby, J.M.A. *Fundamentals of Celestial Mechanics*, (Willmann-Bell, Richmond, 1988).
Laskar, J.: 1989, *Nature* **338**, 237.
Laskar, J.: 1990, *Icarus* **88**, 266.
Quinn, T.R., Tremaine, S.D. and Duncan, M.: 1991, *Astron. J.* **101**, 2287.
Roy, A.E., Walker, I.W., Macdonald, A.J., Williams, I.P., Fox, K., Murray, C.D., Milani, A., Nobili, A.M., Message, P.J., Sinclair, A.T. and Carpino, M.: 1988, *Vistas in Astronomy* **32**, 95.
Sussman, G.J. and Wisdom J.: 1988, *Science* **241**, 433.
Sussman, G.J. and Wisdom, J.: 1992, submitted.
Wisdom, J.: 1982, *Astron. J.* **87**, 577.
Wisdom, J.: 1983, *Icarus* **56**, 51.
Wisdom, J., and Holman, M.: 1991 *Astron. J.* **102**, 1528.

Discussion

P.Goldreich – I have the same question for you as for the previous speaker. What do your integrations imply about qualitative changes of the planetary orbits over the age of the solar system?

J.Wisdom – The integration to date do not give any indication of a gross instability of the solar system. Of course, since the system is apparently chaotic, we cannot rule out the possibility of gross instability. Recall some chaotic asteroid trajectories have been seen to evolve chaotically for 100 Lyapunov times at low eccentricity and then suddenly jump to large eccentricity. It will be very interesting to see a number of integrations of the whole solar system for the age of the solar system and longer.

S.F.Dermott – At the recent "Asteroids, Comets, Meteors" meeting in Flagstaff (June, 1991), E.Shoemaker argued that the Hirayama asteroid families may be young as he expects planetary perturbations to disperse the orbital elements. Have you found any evidence for this in your test-particle calculations? Would you expect any dispersion to occur?

J.Wisdom – I have not studied that problem.

Cl.Froeschlé – (1) What about the close approaches which are so important for cometary transfer through the solar system? (2) Do you see any limitation to your mapping integrator? .

J.Wisdom – (1) As I have programmed the new map, it can only handle the case when the Keplerian motions are elliptical. Thus, in very close approaches, in which the osculating elements momentarily become hyperbolic, the program, as it stands, will fail. Presumably, this could be fixed by using universal variables for the solution of the Kepler motion. Not so close approaches, even temporary satellite capture, are handled satisfactorily even by the present program. (2) I am not aware of any significant limitations of the method. It has worked far better than I ever expected.

G.Tancredi – Milani and Nobili (1990), in the frame of the Longstop project, criticized your conclusion about the chaotic behaviour of Pluto, pointing out that: (1)

the saturation distance of 45 AU is an indication that the mean motion libration is preserved; (2) The mass of Neptune in your integration did not correspond to the latest determination and they claimed that this discrepancy could lead to a different behaviour of some resonant argument that could induce the observed chaos. Do you have any comments on these points?

J. Wisdom – (1) Their comment on the saturation distance of 45 AU was not a criticism, just an explanation of an observed fact. We computed the divergence of Pluto also by the linearized variational method which does not saturate and agrees with the two particle method before it saturates. (2) We have carried out time integrations which show that neighboring Pluto's diverge exponentially with a time scale of 20 Myr. In the last calculations, the masses and initial conditions are different from the first two. In particular, it had a large Pluto mass. The first calculations gave Pluto zero mass. The agreement of these calculations shows that the result is not a very sensitive function of masses and initial positions, as they conjecture.

W. Landgraf – I like to pass a general remark and suggestion w.r.t. the long term motion of the planets and the question of their stability or any chaotic motion.

In all investigations published until now, the authors have integrated only the motion of the objects, $\mathbf{x}(t)$, using some initial values $\mathbf{x}_0(t) = \mathbf{x}(t_0)$. Then, they comment that, after very long time, because of the uncertainty of the initial values, σ_{x_0}, the integration will have only statistic value, or even could happens a chaotic motion.

My suggestion is to integrate the full matrix of deviations $\mathcal{G}(t) = \frac{\partial \mathbf{x}(t)}{\partial \mathbf{x}_0}$, for all planets. This can be performed with full accuracy (not necessarily by numeric variation), and it is not necessary to integrate, besides of this, the motion also (because $\mathbf{x}(t) = \mathcal{G}(t) \cdot \mathbf{x}_0$). Instead of $\mathcal{G}(t)$ referred to the rectangular coordinates, can also be integrated $\mathcal{G}_E = \frac{\partial \mathbf{E}}{\partial \mathbf{E}_0}$ where $\mathbf{E}(t)$ are the elements at t and $\mathbf{E}_0$ at t_0. The behaviour of $\mathcal{G}_E$ is *much more informative* than the integration of the motion itself, *and do depend only minutely from the precise values of the initial values, or from their precision.*

$\mathcal{G}_E(t)$ shows immediately, when happens a transformation of the motion into a chaotic one (large values of $\mathcal{G}_E$) or an unstable one (differences of $\mathcal{G}_E$ from $\mathcal{I}$ change quickly with time), as normally $\mathcal{G}_E$ is close to the unit matrix $\mathcal{I}$. The integration of $\mathcal{G}_E$ is only a little more effort than that of $\mathbf{E}(t)$, but gives much more informations.

J. Wisdom – Integration of the full variational system is too computationally demanding. The integration of the planetary system alone, over the timescales of current interest, is barely possible.