Fields Institute Communications
Volume 10, 1996

Symplectic Correctors

J. Wisdom and M. Holman
Earth Atmospheric and Planetary Sciences
Massachusetts Institute of Technology
Cambridge MA USA 02139

J. Touma
Applied Mathematics
Massachusetts Institute of Technology
Cambridge MA USA 02139

Abstract. Symplectic integration algorithms typically yield trajectories that exhibit spurious oscillation in energy and state variables. In the delta function formulation of symplectic integration these oscillations have a clear origin, and canonical transformations can be made to remove them. The accuracy of symplectic integrators is substantially improved when combined with these symplectic correctors. The methods developed here are generally applicable to the integration of perturbed dynamical systems, but illustrated here by applications to the planetary n-body problem.

1 Introduction

And so I turn to the abyss
Of necromancy, try if art
Can voice or power of spirits start,
To do me service and reveal
The things of Nature's secret seal,
And save me from the weary dance,
Of holding forth in ignorance.
Then shall I see, with vision clear,
How secret elements cohere,
And what the universe engirds.
– Goethe's *Faust*

In the delta function formulation of symplectic integration (Chirikov [1979], Wisdom [1982] [1983], Wisdom [1988], Wisdom and Holman [1991] [1992]), the

1991 *Mathematics Subject Classification.* Primary: 65L05; Secondary: 58F05, 70-08, 70F15.

mapping approximation is derived by adding high frequency terms to the Hamiltonian in such a way that periodic Dirac delta functions are formed; the evolution of the resulting mapping Hamiltonian is a composition of the evolutions governed by the separately integrable pieces of the original Hamiltonian. The added high-frequency terms are argued to be unimportant to the long-term evolution by the averaging principle, provided stepsize resonances do not overlap and are avoided. The delta function formulation leads to the same algorithms as the Lie series composition formulation, but the motivation is different. In the delta function formalism the mapping is explicitly derived from a mapping Hamiltonian, which is motivated by the averaging principle. The Lie series composition formulation is motivated by the desire to have symplectic algorithms with a high order match of the Taylor series of the solutions. Recent reviews of the symplectic integration literature are given by Sanz-Serna [1992] and McLachlan [1995a].

Analysis of the mapping Hamiltonian with techniques from non-linear dynamics (most notably the resonance overlap criteria) allow stability criteria to be developed for the integration algorithms (Wisdom and Holman [1992]). It is not at all apparent how a similar discussion of the non-linear stability of the integration algorithms could be carried out in the pure Lie series composition formulation. Stepsize resonances are clearly exhibited by the algorithms, but they have no clear origin within the Lie series formulation. On the other hand, stepsize resonances have a clear origin in the delta function formalism, and furthermore the delta function formalism gives a quantitatively correct description of the stepsize resonances (Wisdom and Holman [1992]).

Trajectories computed with symplectic integration algorithms typically display spurious oscillations in energy and state variables. These oscillations cloud the issue of determining the accuracy of the computed trajectories. The averaging principle and the oscillatory nature of the errors suggests that the mapping method is more accurate than a naive view of the apparent errors suggests. The Lie series formulation gives no clear insight into the origin of the oscillations nor provides any means of removing them except through going to higher order and/or smaller stepsize. The mapping Hamiltonian in the delta function formalism gives a clear understanding of the origin of the spurious oscillations in energy and state variables. In this paper, we show that the evolution of the mapping and the real evolution are more closely related than is suggested by the agreement of the Taylor series. We show that it is a mistake to identify the mapping variables and the actual state variables. In fact, we show how to relate the mapping variables to the real variables and consequently remove the spurious oscillations, and in many cases dramatically improve the accuracy of the integrations without going to higher order or smaller stepsize. A preliminary sketch of some of these ideas was presented in Wisdom [1988], and they have already been used to some extent in Tittemore and Wisdom [1988] [1989] [1990]. The results presented here are deeper and the methods are generally applicable.

2 Perturbation Theory of Symplectic Maps

Consider a general Hamiltonian of the form

$$H = H_0(J_1, J_2) + \epsilon H_1(J_1, J_2, \theta_1, \theta_2), \tag{2.1}$$

where, for convenience, we have chosen canonical coordinates (J_i, θ_i) for which the unperturbed problem governed by H_0 is cyclic in the coordinates. We shall focus on problems with moderately small ϵ. Two degrees of freedom are enough to exhibit the problems and our solutions to those problems. In the end our formulation is not only independent of the number of degrees of freedom, but coordinate independent as well.

A simple mapping Hamiltonian is generated from H by adding high-frequency terms so that the Hamiltonian becomes

$$H_{\mathrm{Map}} = H_0(J_1, J_2) + \epsilon H_1(J_1, J_2, \theta_1, \theta_2) 2\pi \delta_{2\pi}(\Omega t), \tag{2.2}$$

where

$$\delta_{2\pi}(t) = \sum_{n=-\infty}^{\infty} \delta(t - 2\pi n), \tag{2.3}$$

is a periodic sequence of Dirac delta functions with period 2π (Wisdom [1982]). Central to understanding the averaging motivation for the mapping method is the Fourier representation

$$\delta_{2\pi}(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \cos(nt). \tag{2.4}$$

Across each delta function the evolution is governed solely by H_1, and between the delta functions the evolution is governed solely by H_0. The mapping provides an approximation to the evolution of H by a composition of the evolutions governed by the pieces. If the component Hamiltonians are integrable and efficiently solvable, then the mapping provides an efficiently computable approximation to the evolution generated by the full Hamiltonian. As a conventional integration algorithm, the map presented above is only accurate to first order if the state of the system is recorded at integer multiples of the mapping period $t = n\Delta t = n2\pi/\Omega$. However if the state of the system is recorded halfway between the delta functions, the mapping is accurate to second order in the mapping stepsize. Construction of higher order mappings through the delta function formulation is discussed by Wisdom and Holman [1991].

To prepare for perturbation theory, we first express the perturbation as a multiply periodic Poisson series

$$H = H_0(J_1, J_2) + \epsilon \sum_{ij} a_{ij}(J_1, J_2) \cos(i\theta_1 + j\theta_2) \tag{2.5}$$

where, for convenience, we assume only cosine terms appear in the expansion of H_1. Using the Fourier representation of the periodic sequence of delta functions, the Poisson series for the mapping can be written

$$H = H_0(J_1, J_2) + \epsilon \sum_{ijn} a_{ij}(J_1, J_2) \cos(i\theta_1 + j\theta_2 + n\Omega t). \tag{2.6}$$

In this form it is apparent that the mapping Hamiltonian differs from the true Hamiltonian by the terms with $n \neq 0$. Now, we use ordinary perturbation theory to formally eliminate these terms and thus provide the connection between the mapping phase space variables and the real phase space variables.

We use Lie series in the extended phase space to facilitate the manipulations of canonical perturbation theory. Define the Lie derivative L_W by

$$L_W f = \{W, f\} = \sum_i \left\{ \frac{\partial W}{\partial q_i} \frac{\partial f}{\partial p_i} - \frac{\partial W}{\partial p_i} \frac{\partial f}{\partial q_i} \right\} \tag{2.7}$$

where, as indicated, the curly brackets denote the usual Poisson bracket. The operator $e^{\epsilon L_W}$ generates canonical transformations. This is easy to see by replacing ϵ by t and W by $-H$, then the exponential generates the Taylor series in time of any dynamical variable, and time evolution is canonical. A function of the phase space coordinates evolves only as a result of the evolution of the coordinates; thus the Lie transform of a function is the function of the Lie transformed coordinates. Alternately stated, the inverse Lie transform of a function evaluated at the transformed coordinates gives the original function of the original coordinates. Thus if a canonical transformation is carried out by a Lie transform, the Hamiltonian governing the evolution of the transformed coordinates is the inverse Lie transform of the original Hamiltonian. Keep in mind that the manipulations are formal and the resulting series may not converge. To use Lie series with time dependent Hamiltonians it is convenient to go to the extended phase space, we promote the time t to canonical coordinate status, with associated canonical momentum T, and introduce a new time τ. The Hamiltonian in the extended phase space is obtained by adding T to H: $H' = H + T$. The equation of motion for t is $dt/d\tau = \partial H'/\partial T = 1$, confirming that the original dynamics is preserved. For an introduction to Lie series see Steinberg [1985].

We ask: What generator W can be used to eliminate the $n \neq 0$ terms from the mapping Hamiltonian? Applying $e^{-\epsilon L_W}$ to H', we find, to first order in ϵ, that W must satisfy the equation

$$\{W, H_0 + T\} - \tilde{H}_1 = 0 \tag{2.8}$$

or

$$\omega_1 \frac{\partial W}{\partial \theta_1} + \omega_2 \frac{\partial W}{\partial \theta_2} + \frac{\partial W}{\partial t} - \sum_{ij, n\neq 0} a_{ij} \cos(i\theta_1 + j\theta_2 + n\Omega t) = 0, \tag{2.9}$$

where we have introduced the notation

$$\omega_i(J_1, J_2) = \frac{\partial H_0(J_1, J_2)}{\partial J_i}, \tag{2.10}$$

and represented the terms that must be added to H_1 to make the delta functions by $\tilde{H}_1 = H_1[2\pi\delta_{2\pi}(\Omega t) - 1]$. The solution of this equation is

$$W(J_1, J_2, \theta_1, \theta_2) = \sum_{ij, n\neq 0} \frac{a_{ij}(J_1, J_2) \sin(i\theta_1 + j\theta_2 + n\Omega t)}{i\omega_1 + j\omega_2 + n\Omega}. \tag{2.11}$$

Notice what has been accomplished. Formally, this transformation eliminates all terms that differ between the real Hamiltonian and the mapping Hamiltonian up to order ϵ^2. Contrast this with the Taylor series expansion for the generalized leap frog which contains error terms of order ϵ, the largest of which is generated by $\{H_0, \{H_0, H_1\}\}$. Even without further consideration of the perturbation theory, this is consistent with our suggestion that most of the error in the generalized leap frog is of an oscillatory nature and can be removed by a canonical transformation. However, we need a more practical representation of the generating function.

Note that n only appears in the sine function and in the denominator. This fact allows the sums over n to be carried out. By expanding the sine of the sum, we find

$$W = \sum_{ij,} a_{ij} \left[\sin(i\theta_1 + j\theta_2) F(i\omega_1 + j\omega_2, \Omega, t) + \cos(i\theta_1 + j\theta_2) G(i\omega_1 + j\omega_2, \Omega, t)\right] \quad (2.12)$$

where we have introduced the functions

$$F(\omega, \Omega, t) = \sum_{n \neq 0} \frac{\cos(n\Omega t)}{\omega + n\Omega} = \frac{2\omega}{\Omega^2} \sum_{n>0} \frac{\cos(n\Omega t)}{\left(\frac{\omega}{\Omega}\right)^2 - n^2} \quad (2.13)$$

and

$$G(\omega, \Omega, t) = \sum_{n \neq 0} \frac{\sin(n\Omega t)}{\omega + n\Omega} = -\frac{2}{\Omega} \sum_{n>0} \frac{n \sin(n\Omega t)}{\left(\frac{\omega}{\Omega}\right)^2 - n^2}. \quad (2.14)$$

These sums converge provided $\omega + n\Omega \neq 0$, for any n. The functions $F(\omega, \Omega, t)$ and $G(\omega, \Omega, t)$ are both periodic in time with period $2\pi/\Omega$, the mapping stepsize. Explicitly:

$$F(\omega, \Omega, t) = \frac{\pi}{\Omega} \frac{\cos(\omega t - \frac{\pi\omega}{\Omega})}{\sin \frac{\pi\omega}{\Omega}} - \frac{1}{\omega}, \quad (2.15)$$

for $0 \leq \Omega t \leq 2\pi$, and

$$G(\omega, \Omega, t) = -\frac{\pi}{\Omega} \frac{\sin(\omega t - \frac{\pi\omega}{\Omega})}{\sin \frac{\pi\omega}{\Omega}}, \quad (2.16)$$

for $0 < \Omega t < 2\pi$, with $G = 0$ for Ωt equal to integral multiples of 2π. The periodic extension of F is continuous but has discontinuous derivative for $\Omega t = 2\pi n$. The periodic extension of G is discontinuous at $\Omega t = 2\pi n$, and the sum converges to the midpoint of the discontinuity: $G(\omega, \Omega, 2\pi n/\Omega) = 0$. We have the limiting values $G(\omega, \Omega, 0) = 0$, $G(\omega, \Omega, 0^+) = \pi/\Omega$, $G(\omega, \Omega, 0^-) = -\pi/\Omega$, where 0^+ and 0^- indicate limits from above and below, respectively.

We note some further properties of F and G. First, both F and G have zero time average, as is evident from the original Fourier representation. We also note that

$$G(\omega, \Omega, t) = \frac{1}{\omega} \frac{d}{dt} F(\omega, \Omega, t). \quad (2.17)$$

This fact can be used to deduce the summed form of G from the form for F without explicitly performing the sum. The maximum value of F occurs at $\Omega t = \pi$:

$$F(\omega, \Omega, \pi/\Omega) = \frac{\pi}{\Omega} \frac{1}{\sin \frac{\pi\omega}{\Omega}} - \frac{1}{\omega}, \quad (2.18)$$

which, for small ω/Ω, is approximately

$$F(\omega, \Omega, \pi/\Omega) \approx \frac{\pi^2 \omega}{6\Omega^2}. \quad (2.19)$$

The minimum value of F occurs at $\omega t = 0$; for small ω/Ω it has the approximate value:

$$F(\omega, \Omega, \pi/\Omega) \approx -\frac{\pi^2 \omega}{2\Omega^2}. \tag{2.20}$$

Note that the magnitude of the function F is larger at $t = 0$ than at $t = \pi/\Omega$. The extreme values of G are given by the limiting values around $t = 0$. Recall that the stepsize of the mapping is $\Delta t = 2\pi/\Omega$, and note that the extreme values of F are proportional to Δt^2, whereas the extreme values of G are proportional to Δt. Both F and G have two zeros. $G(\omega, \Omega, 0) = G(\omega, \Omega, \pi/\Omega) = 0$. The zeros of F satisfy a transcendental equation.

The first lesson we haved learned by performing the sum over n is that, for small step size, the correction from map coordinates to dynamical coordinates is minimized at the midpoint of the interval between the delta functions, when the larger correction proportional to G is zero by virtue of the vanishing of G. Recording the state of the system midway between the delta functions minimizes the error of the mapping if the distinction between mapping variables and real state variables is ignored. Thus, from yet another point of view (see Wisdom and Holman [1991], for two others) we have discovered the generalized leap frog. The real lesson, though, is that the mapping variables and the actual variables are not the same, and that in fact they differ in an explicit calculable way. The leap frog is only the best solution if the distinction between these variables is ignored, and then the error is of order ϵ. Maintaining and accounting for the distinction between mapping variables and actual variables, the error of the mapping method is of order ϵ^2, and this is true at *any* time, not just at the time midway between the delta functions.

3 Illustration

A natural test application is to the pendulum. Consider the Hamiltonian

$$H = \frac{1}{2}J^2 + \epsilon \cos\theta. \tag{3.1}$$

We make a simple mapping approximation by introducing the delta functions:

$$H_{\mathrm{Map}} = \frac{1}{2}J^2 + 2\pi\delta_{2\pi}(\Omega t)\epsilon\cos\theta. \tag{3.2}$$

The equations of motion are just Hamilton's equations. Crossing a delta function changes the momentum:

$$J' = J + \Delta t \epsilon \sin\theta, \tag{3.3}$$

where $\Delta t = 2\pi/\Omega$, the mapping stepsize, and between delta functions the angle rotates uniformly

$$\theta' = \theta + \Delta t J'. \tag{3.4}$$

At intermediate times between the delta functions the angle is rotated appropriately to that time. Of course, with a simple change of notation this is just the standard map (Chirikov [1979]), but without the usually assumed periodicity in the momentum variable. The usual standard map parameter is given by $K = \Delta t^2 \epsilon$.

In this case, since the perturbation has a trivial Fourier expansion, the generator for the corrector can be given in closed form:

$$W = [\sin(\theta)F(\omega,\Omega,t) + \cos(\theta)G(\omega,\Omega,t)], \tag{3.5}$$

with frequency $\omega(J) = \partial H_0/\partial J = J$. The transformation from the mapping variables to the real variables is

$$J = \mathrm{e}^{\epsilon L_W} J_M \approx J_M + \epsilon \frac{\partial W}{\partial \theta}\bigg|_M \tag{3.6}$$

and

$$\theta = \mathrm{e}^{\epsilon L_W} \theta_M \approx \theta_M - \epsilon \frac{\partial W}{\partial J}\bigg|_M . \tag{3.7}$$

The subscript M indicates mapping variables. Using this Euler approximation to the evolution generated by W gives a corrector that at order ϵ is correct to all orders in Δt. The Euler formula is not symplectic, but this can hardly matter since the corrector is applied only to the output points and the corrected output points are not used further in the integration.

The maximum of the energy error for several methods is displayed in Figure 1. We compare the uncorrected first order map, the generalized leap frog, the corrected leap frog, and the fourth order method of Candy and Rozmus [1990], Yoshida [1990], and Forest and Ruth [1990]. Since the problem is artificial nothing guides the choice of parameters, so the parameters have been chosen to get the attention of the reader. Whether the idea of using correctors is useful will depend on application to problems of interest with appropriate parameters. On the principle that we should compare error of different methods for similar work, the stepsize used for the fourth order method is three times the stepsize for the leap frog, since three force evaluations are required for each fourth order step. Note the especially poor performance of the fourth order method. For this choice of parameters, the corrected map does substantially better than the other methods. Note also that the output time of the corrected map is not limited to the midpoint between the delta functions. That is, the corrected map is as accurate at any output phase.

4 Development of F and G

In general the perturbation will not have a simple Poisson series, so the practical use of the transformation to and from mapping coordinates will rest on the development of more practical expressions for the generating function. To this end, expand F and G as power series in the ratio ω/Ω (which we presume here to be less than 1). We find

$$F(\omega,\Omega,t) = \frac{2\omega}{\Omega^2}\sum_{n>0}\frac{\cos(n\Omega t)}{\left(\frac{\omega}{\Omega}\right)^2 - n^2} = -\frac{2\omega}{\Omega^2}\sum_{i=0}^{\infty} C_{2i+2}(\Omega t)\left(\frac{\omega}{\Omega}\right)^{2i} \tag{4.1}$$

and

$$G(\omega,\Omega,t) = -\frac{2}{\Omega}\sum_{n>0}\frac{n\sin(n\Omega t)}{\left(\frac{\omega}{\Omega}\right)^2 - n^2} = \frac{2}{\Omega}\sum_{i=0}^{\infty} S_{2i+1}(\Omega t)\left(\frac{\omega}{\Omega}\right)^{2i}, \tag{4.2}$$

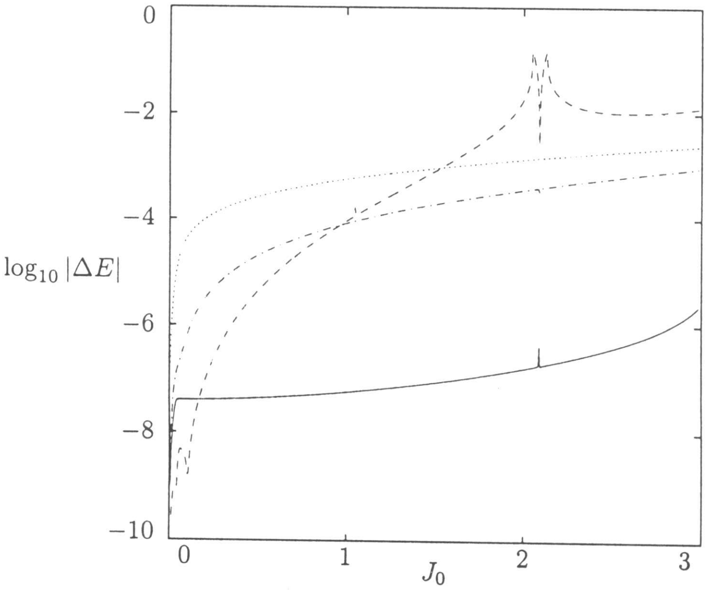

Figure 1 The maximum energy error for various methods is plotted as a function of the initial momentum. The dotted line is the uncorrected map. The dot-dashed line is the generalized leap-frog. The dashed line is the fourth order composition method. The solid line is corrected map. The performance of the corrected map is substantially better than that of the other methods. The parameters are: $\epsilon = 0.001$ with $\Delta t = 1$. The initial angle in all cases is $\theta = \pi$. The error plotted is the maximum over 10000 iterations.

where

$$C_i(\tau) = \sum_{n>0} \frac{\cos(n\tau)}{n^i} \tag{4.3}$$

and

$$S_i(\tau) = \sum_{n>0} \frac{\sin(n\tau)}{n^i}. \tag{4.4}$$

Both C_i and S_i are 2π periodic in τ, and they have zero average. The functions C_i and S_i have the important properties: $S_i(\tau) = -dC_{i+1}/d\tau$ and $C_i(\tau) = dS_{i+1}/d\tau$. Note that in F and G the index of C is always even and the index of S is always odd. We introduce the function E_i such that $E_{2i} = (-1)^i 2C_{2i}$ and $E_{2i+1} = (-1)^i 2S_{2i+1}$, then the E_i are a set of 2π periodic functions with zero average possessing the interesting property that $E_i(\tau) = dE_{i+1}/d\tau$. These properties determine all of the E_i given E_1. The sum for E_1 is recognized as the Fourier series for a line: $E_1(\tau) = \pi - \tau$ for $0 < \tau < 2\pi$, and $E_1(0) = 0$. Thus the E_n are in fact all polynomials. It turns out that the E_n can be written in terms of the Bernoulli polynomials

$$E_n(\tau) = -\frac{(2\pi)^n}{n!} B_n(\tau/2\pi). \tag{4.5}$$

The generating relation for the Bernoulli polynomials is

$$\frac{\mathrm{e}^{xt}}{\mathrm{e}^{t}-1} = \sum_{n=0} B_n(x)\frac{t^{n-1}}{n!} \tag{4.6}$$

The first few are:

$$\begin{array}{rcl} B_1(x) & = & x-\frac{1}{2} \\ B_2(x) & = & x^2-x+\frac{1}{6} \\ B_3(x) & = & x^3-\frac{3}{2}x^2+\frac{1}{2}x \\ B_4(x) & = & x^4-2x^3+x^2-\frac{1}{30}. \end{array} \tag{4.7}$$

The representation of E_i in terms of Bernoulli polynomials is only valid in the interval $0 < t < \Delta t$, and is otherwise periodically extended. For consistency, we must add

$$E_0(\tau) = 2\pi\delta_{2\pi}(\tau) - 1. \tag{4.8}$$

We return now to the development of the generator W. Substituting for F and G their expansions in terms of the C and S polynomials, we find

$$W = \sum_{ij,k\geq 0} a_{ij}\left\{\sin(i\theta_1+j\theta_2)\left[-\frac{2\omega}{\Omega^2}C_{2k+2}(\Omega t)\left(\frac{\omega}{\Omega}\right)^{2k}\right]\right.$$
$$\left. + \cos(i\theta_1+j\theta_2)\left[\frac{2}{\Omega}S_{2k+1}(\Omega t)\left(\frac{\omega}{\Omega}\right)^{2k}\right]\right\}. \tag{4.9}$$

Notice that each term can be written as a multiple Poisson bracket of H_0 with terms of H_1. Performing the sum over i and j first, the terms can be regrouped into H_1. We have then

$$W = \sum_{k>0}\frac{E_k(\Omega t)}{\Omega^k}L_{H_0}^{k-1}H_1 = -\Delta t\sum_{k>0}\frac{1}{k!}B_k\left(\frac{t}{\Delta t}\right)\Delta t^{k-1}L_{H_0}^{k-1}H_1 \tag{4.10}$$

This is a practical relation, as desired. The generator is written as a Taylor series in the stepsize the terms of which involve successively higher order Poisson brackets of H_0 with H_1 with rapidly decreasing coefficients. Notice that the expression is independent of the number of degrees of freedom, and coordinate free as well. The derivation is easily extended to n-degrees of freedom, with the same result.

5 Lie Series Correspondence

Lie series simplify the formulation of higher order composition methods (Forest and Ruth [1990], Yoshida [1990] and Wisdom and Holman [1991]). In this section we present the Lie series correspondence of the correctors introduced in the last section.

Let $A = \Delta t L_{H_0}$ and $B = \Delta t L_{H_1}$. Taylor series for the evolution of the actual system for one timestep are generated by the operator e^{A+B}. The generalized leap frog approximation to the evolution for one timestep is generated by $\mathrm{e}^{A/2}\mathrm{e}^{B}\mathrm{e}^{A/2}$. The error is third order in the timestep; the generalized leap frog is, in this sense, a second order integrator. More formally,

$$\mathrm{e}^{A/2}\mathrm{e}^{B}\mathrm{e}^{A/2} - \mathrm{e}^{A+B} = -\frac{1}{24}[A,[A,B]] + \frac{1}{12}[B,[B,A]] + o(\Delta t^4). \tag{5.1}$$

A gazillion such higher order integrators are known (see, for example, Koseleff [1993]). Note that even though the error is of order Δt^3 it is first order in ϵ. In fact, there are error terms which are first order in ϵ at all orders in Δt.

In the previous sections we found, however, that the mapping method, including the generalized leap frog as a special case, is actually better than the error in the Taylor series suggests. That is, the evolution generated by the mapping has calculable and removable high frequency oscillations. It is a mistake to identify the mapping variables and the actual variables. More properly, we must not only derive the mapping approximation to the evolution, but we must also derive the relationship of the mapping variables to the actual variables.

In the Lie series formulation, the operator that corresponds to the mapping evolution to a time after a delta function of $t = \alpha\Delta t$, where $0 \leq \alpha \leq 1$ is

$$\mathrm{e}^{(1-\alpha)A}\mathrm{e}^{B}\mathrm{e}^{\alpha A}. \tag{5.2}$$

Of course, this approximates the true evolution only to first order in the stepsize, the local truncation error is proportional to $\epsilon\Delta t^2$, unless the output point is midway between the delta functions ($\alpha = 1/2$) and then the local truncation error is of order $\epsilon\Delta t^3$. We now realize though that the mapping variables do not correspond to the true variables and before evolving with the mapping we must transform to mapping variables and after the mapping step we must transform back to the real variables. The generator W of the Lie transformation that transforms to mapping variables has been derived in previous sections. Let $C(\alpha) = L_W(\alpha)$ be the Lie derivative corresponding to the generator W, then

$$\mathrm{e}^{-C(\alpha)}\mathrm{e}^{(1-\alpha)A}\mathrm{e}^{B}\mathrm{e}^{\alpha A}\mathrm{e}^{C(\alpha)}, \tag{5.3}$$

generates an evolution which is correct at order ϵ to all orders in Δt. The fact that the corrector depends on the phase of the output α illustrates the high-frequency nature of the corrections. If only a single output phase is considered, as, for example, in the leap frog, the high-frequency oscillatory character of the error is not apparent. It is "strobed" away. We introduce a notation for the commutator with respect to the operator A: $\mathcal{L}_A B = [A, B] = AB - BA$. To carry out the corresponding canonical transformation we need the operator L_W. The Lie derivative with respect to the Poisson bracket of two generating functions is the commutator of two Lie derivatives with respect to the individual generating functions. Thus the expression for L_W can be immediately written down in terms of A and B:

$$C(\alpha) = L_W(\alpha) = -\sum_{k>0} \frac{B_k(\alpha)}{k!}\mathcal{L}_A^{k-1}B \tag{5.4}$$

For the special case of $\alpha = 1/2$ this is

$$C = \frac{1}{24}\mathcal{L}_A B - \frac{7}{5760}\mathcal{L}_A^3 B + \cdots \frac{2^{2n-1}-1}{(2n)!2^{2n-1}}B_{2n}\mathcal{L}_A^{2n-1}B, \tag{5.5}$$

where B_n are the Bernoulli numbers. This corrector eliminates the error terms that are first order in ϵ to all orders in Δt. Use of the corrector does not affect substantially the efficiency of the mapping if output is uncommon, because successive steps can be combined: the effect of the e^{-C} is reversed by the following e^{C}, and the two factors of $\mathrm{e}^{\alpha A}$ and $\mathrm{e}^{(1-\alpha)A}$ may be combined into a single factor of e^{A}. Only when output is desired do we actually have to carry out the e^{C} step. McLachlan (personal communication [1994]) has pointed out that Butcher [1969] developed a similar idea to extend the order of explicit Runge-Kutta methods.

Though, no doubt, the complete expression for C could be derived entirely with the commutator algebra of A and B, the delta function formalism has provided the motivation for the approach as well as given us the complete expression to all orders in Δt by simple analysis. Independent of α the third order error term is $-(1/24)\mathcal{L}_B^2 A$, which is second order in ϵ. At this point, this is the real error of the method, the other errors were only apparent errors because the mapping coordinates were being mistaken for the actual coordinates.

6 Lie Series Implementation

The generator for the transformation from mapping coordinates to real coordinates has been written in terms of successive Poisson brackets of H_0 and H_1. Since this transformation need be done only at the output points, the efficiency of the implementation of the corrector is unimportant. The most direct implementation is to numerically integrate Hamilton's equations corresponding to the generator. A more convenient approach is to represent the corrector as a composition of the readily computable evolutions governed by H_0 and H_1. In this section, we present a representation of the corrector as such a composition. We optimize efficiency of derivation rather than attempting to find the most efficient composition. That can be done later.

We note a key formula for manipulating Lie series involving correctors, the effect of a corrector on a kernel:

$$\mathrm{e}^{C}\mathrm{e}^{K}\mathrm{e}^{-C} = \mathrm{e}^{K+\mathcal{L}_C K+\frac{1}{2}\mathcal{L}_C^2 K+\cdots} = \mathrm{e}^{\mathrm{e}^{\mathcal{L}_C}K}. \tag{6.1}$$

This formula is used repeatedly in the following development. With the knowledge that there is only one independent n^{th} order bracket involving a single factor of K, it is easy to prove this relation by matching the coefficients of the $C^n K$ terms in the expansion (see, for example, Belinfante and Kolman [1989]). The corrector formula provides a convenient base form from which to generate Lie series products because of the simplicity with which higher order terms are determined.

We here assume output of the mapping has been taken midway between the delta functions. Thus we shall derive a product corrector for the leap frog. For this case, we recall that the corrector has only brackets involving an even number of factors of H_0 and H_1. A convenient form that has this property is the product of two corrector forms. Let

$$\mathrm{e}^{X(a_1,b_1)} = \mathrm{e}^{a_1 A}\mathrm{e}^{b_1 B}\mathrm{e}^{-a_1 A}. \tag{6.2}$$

All terms in X are proportional to B. Now let

$$\mathrm{e}^{Z} = \mathrm{e}^{X(a_1,b_1)}\mathrm{e}^{X(-a_1,-b_1)}. \tag{6.3}$$

Constructed in this way, Z has only terms with an even number of factors of A and B, at order ϵ. Composition of such forms maintains this property.

We write the corrector as a product of factors of Z:

$$\mathrm{e}^{C} = \mathrm{e}^{Z(a_1,b_1)}\mathrm{e}^{Z(a_2,b_2)}\cdots\mathrm{e}^{Z(a_n,b_n)}. \tag{6.4}$$

Let c_m be the coefficients of $\mathcal{L}_A^m B$ in the corrector. The constraint equations are easily written down using the corrector formula:

$$c_m = 2 \sum_{i=1}^{n} \frac{a_i^m b_i}{m!}. \tag{6.5}$$

As a set of linear equations for the b_i, the coefficients form a Vandermonde matrix, which has a simple explicit inverse. Given a set of a_i all of the $\mathcal{L}_A^m B$ constraint equations can be satisfied by choosing the b_i to be solutions of these linear equations. We do not want the corrector to introduce error terms of order ϵ^2, because this same corrector will be used subsequently with higher order methods. If we select $a_i = -a_{n+1-i}$ for $0 < i \le n$, then $b_i = -b_{n+1-i}$, then the coefficient of the $\mathcal{L}_B^2 A$ term is automagically zero as required. The coefficient of the $\mathcal{L}_B \mathcal{L}_A^2 B$ term is also zero. This form then satisfies all of the constraints for arbitrary a_i. Calculation of the inverse Vandermonde matrix is carried out with rational arithmetic to avoid any possible numerical difficulties. A sample solution is:

$$a_1 = 1\alpha \quad ; \quad b_1 = \frac{45815578591785473}{24519298961757600}\beta$$

$$a_2 = 2\alpha \quad ; \quad b_2 = \frac{-104807478104929387}{80063017017984000}\beta$$

$$a_3 = 3\alpha \quad ; \quad b_3 = \frac{422297952838709}{648658702692000}\beta$$

$$a_4 = 4\alpha \quad ; \quad b_4 = \frac{-27170077124018711}{112088223825177600}\beta$$

$$a_5 = 5\alpha \quad ; \quad b_5 = \frac{102433989269}{1539673404192}\beta$$

$$a_6 = 6\alpha \quad ; \quad b_6 = \frac{-33737961615779}{2641809989145600}\beta$$

$$a_7 = 7\alpha \quad ; \quad b_7 = \frac{26880679644439}{17513784972684000}\beta$$

$$a_8 = 8\alpha \quad ; \quad b_8 = \frac{682938344463443}{7846175667762432000}\beta$$

with $a_i = -a_{17-i}$, and $b_i = -b_{17-i}$, for $8 < i \le 16$, and $\alpha = \sqrt{7/40}$ and $\beta = 1/(48\alpha)$. The scaling of a_i and b_i by α and β was made to reduce the first two non-zero c_i (c_1 and c_3) to unit magnitude. Introduction of these scalings was convenient, but not necessary. This corrector has an error term of order $\mathcal{L}_A^{17} B$. Its performance for the pendulum for the parameters used above is nearly indistinguishable from that of the closed form expression for the corrector on the logarithmic scale. Correctors of lower order do not perform as well. We should keep in mind that different order correctors may be suitable for different problems. This implementation of the corrector is suitable for testing our ideas; optimal implementation of this and the other correctors derived below can be addressed in subsequent investigations.

7 Application to the N-Body Problem

The mapping method has been developed for the n-body problem by Wisdom and Holman [1991]. They tested the method with billion year integrations of the system of the outer planets (Jupiter to Pluto). Subsequently, Wisdom and Sussman [1992] used the mapping method to carry out 100 million year integrations of the whole solar system. These integrations confirmed the chaotic character of the evolution of the solar system (Laskar [1989]). Integrations of the whole solar system are extremely time-consuming and, practically speaking, would not have been possible without the speed of the mapping method. Unfortunately, the integrations were plagued by the high-frequency components. We now have a general prescription for removing them.

First consider the energy error. Most of the energy of the system is contributed by the massive planets, yet the small inner planets have the largest coordinate errors because for them the stepsize is a relatively larger fraction of the orbital period than for the outer planets. So the energy error is not a fair measure of the accuracy of an integration of the whole solar system. Nevertheless, it is interesting to examine it. Figure 2 shows the relative energy error (energy minus initial energy divided by initial energy) in the 100 million year integration of the whole solar system (Wisdom and Sussman [1992]). Early in the integration the energy error is of order a few parts in a billion, and there is no apparent increase in the magnitude of the error for the duration of the integration. Output was taken about every 20,000 years, and successive points are connected by lines for easy visibility. This presentation gives the false impression that the error varies on timescales of tens of thousands of years. The data are intact, and the corrector can be applied to the output even years after the integration was carried out. Figure 3 shows the energy error after the corrector has been applied to the output. Here the energy of the corrected output is compared to the energy of the corrected initial condition. Note that the scale has to be expanded by a factor of a hundred to see the variations in the energy after correction. The high frequency oscillation, which we have argued all along to be unimportant (Wisdom [1982] [1983], Wisdom [1988], Wisdom and Holman [1991] [1992], Sussman and Wisdom [1992]), has almost completely been removed. The remaining energy error shows a secular linear drift, and is probably real integration error. The slope of the relative energy drift is of order 2×10^{-11} per 100 million years, or about 4×10^{-21} per integration step. For comparison, the relative energy error in the Digital Orrery integration of the outer planets using a specially chosen "magic" stepsize was about 2×10^{-19} per integration step, and that was about a thousand times smaller than the error achieved in all previous long term integrations of the solar system (Sussman and Wisdom [1989]). The mapping achieves better energy conservation without the use of magic stepsizes. Both of these integrations use pseudo quadruple precision (compensated summation) in a small select subset of the calculation, but most operations were carried out in ordinary IEEE double precision with about 16 digits. On the average, in the mapping integration of the whole solar system it took about 40,000 integration steps to accumulate one bit of relative energy error! Of course, it is not completely fair to compare the errors achieved in the integration of different systems. The only other long-term direct integration of the whole solar system is the 3 million year integration of Quinn, Tremaine, and Duncan [1991] who used a high order symmetric integrator (Quinlan and Tremaine [1990]). These integrations also used select pseudo quadruple precision. The symmetric integrator is also subject to large energy oscillations, which may or may not be removable. On a 3 million year timescale, they estimate the energy error in their integration to already be 6×10^{-11}, which is larger than the energy error achieved by the mapping after 100 million years.

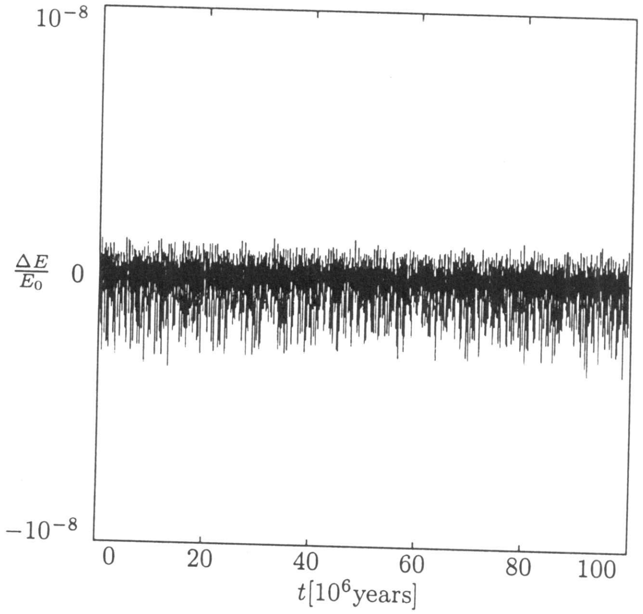

Figure 2 The energy error over 100 million years for the whole solar system. The energy error is large from the beginning, a behavior typical of symplectic integrators.

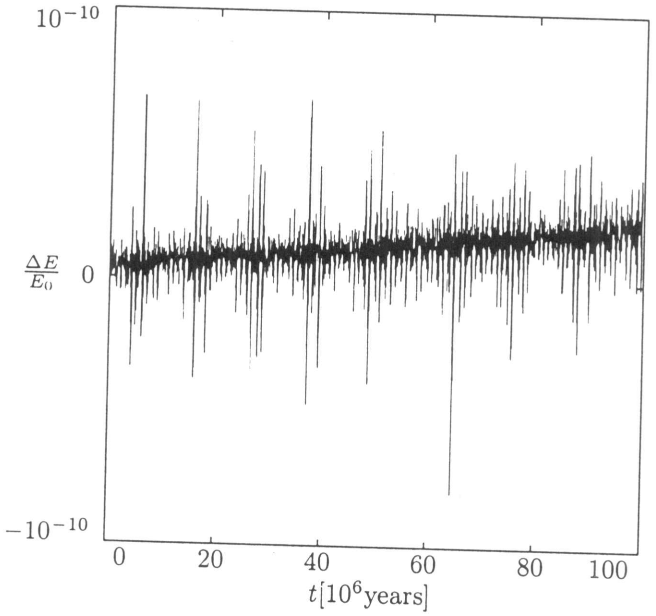

Figure 3 The energy error over 100 million years for the whole solar system after the corrector has been applied. The magnitude of the error is reduced by two orders of magnitude, and now a secular drift of the energy can be seen.

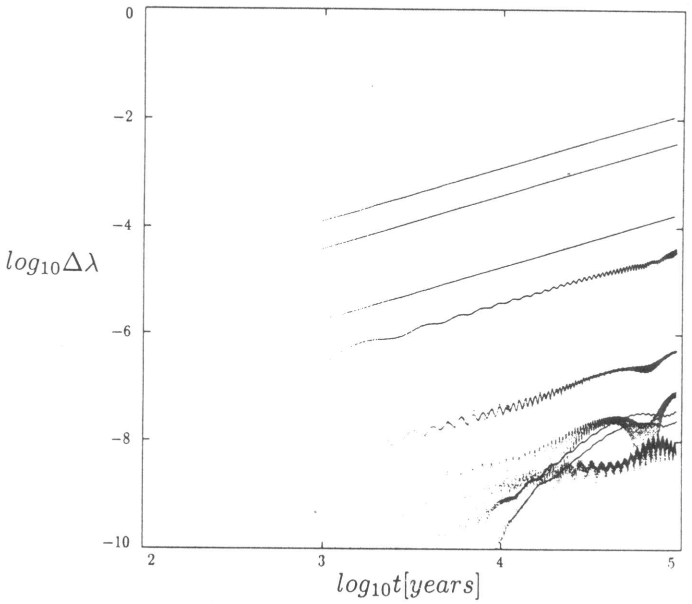

Figure 4 The longitude errors of the planets over 100,000 years using corrected initial conditions. The largest errors are in the inner planets, as expected. At the largest time displayed, the errors, from largest to smallest, are for the planets: Mercury, Earth, Venus, Mars, Saturn, Jupiter, Neptune, Pluto, Uranus.

Unfortunately, we now understand that the transformation from real coordinates to mapping coordinates should be carried out before the integration is begun. The adjustment of the initial conditions is small and does not substantially affect the conclusions of that study, since satisfactory agreement with other integrations was already observed, but the integration could have been better. The error in the initial conditions is most clearly manifest in the growth of the error of the longitudes of the planets, since small errors in the orbital frequencies produce secularly increasing longitude errors. Keep in mind that errors in longitude are not believed to be important to the long-term evolution of the solar system; indeed, Laskar removes the longitudes entirely through averaging. To examine the growth of longitude errors we compared a reference run with integrations with and without the corrector. In these runs we used a stepsize of 7.2 days, the same as the stepsize used in the 100 million year integrations, and no extended precision calculations were used. The reference run used the corrector, partial quadruple precision, and a stepsize reduced by a factor of 10 (from 7.2 days to 0.72 days). We found that the longitude errors grew linearly, both in the run with the corrector and in the run without the corrector. The errors in the run without the corrector are primarily due to using uncorrected initial conditions. The longitude errors in the corrected run are dominated by truncation error. The longitude errors in the corrected run are displayed in Figure 4. As expected, the largest longitude error is in the motion of Mercury. The longitude errors of the other planets are much smaller. The growth of the longitude error is linear on this timescale. Since the energy is experiencing a very small linear growth, we can expect that the growth of the longitude error is actually quadratic, but apparently the coefficient is too small to substantially affect the behavior. The effective relative error in the mean motion of Mercury using uncorrected initial conditions is 7×10^{-8}. With corrected initial conditions the effective mean motion error is 5×10^{-9}. For comparison, in the integration of Quinn, Tremaine, and Duncan, the longitude error of Mercury also grows linearly, with an effective mean motion error of about 4×10^{-9}. Keep in mind that the mapping integration took much larger steps: 7.2 day steps for the mapping, versus 0.75 day steps for the symmetric method. If the mapping step were reduced by this factor of 10 we could expect a factor of at least 100 improvement in the errors, assuming roundoff errors do not limit the improvement. The mapping calculations used for illustration here were carried out without extended precision, but the symmetric method used partial extended precision. Honestly, though this application of the corrector was very successful, we were surprised that the corrector did not reduce the longitude error of Mercury more than it did. Order arguments suggest there should have been much better improvement. We return to this issue in later sections.

We are speaking here of integration error. It should be noted that Sussman and Wisdom [1992] used a model for the general relativistic effects that represented the relativistic precession correctly, but made a small error in the mean motions. Though the model could have been improved (see the appendix in Saha and Tremaine [1992]), it was not considered important because the detailed longitudes are believed to be unimportant, and at the time it appeared that the mapping did not get the longitudes right anyway. Since the correctors have fixed the longitude errors, the modelling error is now larger than the integration error, though still probably unimportant physically.

The mapping method was derived as an efficient approximation to the evolution of the planets that might be only qualitatively correct because of the presence of the high-frequency terms. It turns out that the mapping is not only efficient, but, when used in conjunction with appropriate correctors, the mapping method surpasses the accuracy of competing integration schemes.

8 Extension to Higher Order

In this section we consider the problem of extending the method to higher order in ϵ.

Deriving mappings accurate to higher order in ϵ is facilitated by reinterpreting the transformation from mapping variables to real variables derived above. We shall call this the mapping transformation. Rather than asserting the Hamiltonian for the mapping and then deriving the correction that to some extent converts the mapping Hamiltonian back into the original Hamiltonian, we can use the inverse of the mapping transformation to go directly from the true Hamiltonian to the mapping Hamiltonian plus some error terms. We would like to keep track of the error terms and extend this process to higher order in ϵ.

We apply the transformation to mapping coordinates but now keep track of second order terms in ϵ. At this stage the mapping Hamiltonian is

$$\begin{aligned} H_{\mathrm{Map}} &= \mathrm{e}^{\epsilon L_W}(H_0 + T + \epsilon H_1) \\ &= H_0 + T + \epsilon\{W, H_0 + T\} + \tfrac{1}{2}\epsilon^2\{W, \{W, H_0 + T\}\} \\ &\quad +\epsilon H_1 + \epsilon^2\{W, H_1\} + o(\epsilon^3) \end{aligned}$$

The generating function W has been designed to insert high frequency terms so that perturbations become multiplied by delta functions to give local integrability. Let $\tilde{H}_1$ again denote the terms that must be added to H_1 to make $2\pi\delta_{2\pi}(\Omega t)H_1$. By construction we have $\{W, H_0 + T\} = \tilde{H}_1$ and

$$H_{\mathrm{Map}} = H_0 + T + 2\pi\delta_{2\pi}(\Omega t)\epsilon H_1 + \epsilon^2\{W, H_1 + \frac{1}{2}\tilde{H}_1\}. \tag{8.1}$$

We shall denote the terms proportional to ϵ^2 by H_2; these terms need to be brought into a more useful form. To facilitate this, introduce the notation F_n and G_n for the Fourier coefficients of F and G:

$$F(\omega, \Omega, t) = \sum_{n>0} F_n(\omega, \Omega)\cos(n\Omega t) \tag{8.2}$$

$$G(\omega, \Omega, t) = \sum_{n>0} G_n(\omega, \Omega)\sin(n\Omega t) \tag{8.3}$$

We have:

$$\begin{aligned} H_2 &= \Big\{\textstyle\sum_{ij,n>0} a_{ij}[\sin(i\theta_1 + j\theta_2)F_n\cos(n\Omega t) \\ &+ \cos(i\theta_1 + j\theta_2)G_n\sin(n\Omega t)], H_1\Big(1 + \frac{1}{2}\textstyle\sum_{m\neq 0}\cos(m\Omega t)\Big)\Big\}. \end{aligned}$$

Using the bilinearity of the Poisson bracket and the fact that T does not occur in this expression, we find

$$H_2 = \frac{1}{2}\{W, H_1\} + \frac{1}{2}2\pi\delta_{2\pi}(\Omega t)\{W(t=0), H_1\} \tag{8.4}$$

That is, the ϵ^2 terms are naturally written as a sum of oscillatory terms plus another set of terms multiplied by periodic delta functions.

The mapping Hamiltonian at this point is

$$\begin{aligned} H_{\mathrm{Map}} = H_0 + T \quad &+ \quad 2\pi\delta_{2\pi}(\Omega t)\left[\epsilon H_1 + \tfrac{1}{2}\epsilon^2\{W(t=0), H_1\}\right] \\ &+ \quad \tfrac{1}{2}\epsilon^2\{W, H_1\} + o(\epsilon^3) \end{aligned}$$

The oscillatory term can be pushed to higher order with a second corrector transformation. Let W_2 be the Lie series generator for this canonical transformation. Applying $e^{\epsilon^2 L_{W_2}}$ to H_{Map}, and requiring that the ϵ^2 oscillatory terms are killed, we find the determining condition for W_2:

$$\{W_2, H_0 + T\} + \frac{1}{2}\{W, H_1\} = 0. \tag{8.5}$$

This equation has the solution

$$W_2 = \frac{1}{2}\sum_{k>0}\frac{E_{k+2}(\Omega t)}{\Omega^{k+2}}\sum_{p=1}^{k} L_{H_0}^{k-p}L_{H_1}L_{H_0}^{p}H_1, \tag{8.6}$$

as may be verified by substitution. This corrector removes all oscillatory terms of order ϵ^2. Specializing to the case $\Omega t = \pi$ or $t = \Delta t/2$ the leading term in the second corrector is

$$W_2 = -\frac{7}{5760}\Delta t^3 L_{H_1}L_{H_0}^2 H_1, \tag{8.7}$$

where we have used the fact that $L_{H_0}L_{H_1}L_{H_0}H_1 = L_{H_1}L_{H_0}^2 H_1$.

After using the second corrector the mapping Hamiltonian is

$$H_{\mathrm{Map}} = H_0 + T + 2\pi\delta_{2\pi}(\Omega t)\left[\epsilon H_1 + \frac{1}{2}\epsilon^2\{W(t=0), H_1\}\right] + o(\epsilon^3) \tag{8.8}$$

In general the leading term in the energy error is of order ϵ^3 and is proportional to the fourth order bracket $L_{H_1}^2 L_{H_0} H_1$. If we specialize to $\Omega t = \pi$, then the leading error terms are still of order ϵ^3 but contain the two independent fifth order brackets $L_{H_0} L_{H_1}^2 L_{H_0} H_1$, and $L_{H_1} L_{H_0} L_{H_1} L_{H_0} H_1 = L_{H_1}^2 L_{H_0}^2 H_1$ that are of order ϵ^3.

McLachlan [1995b] introduces a useful notation to describe the order of integration methods for perturbed problems which distinguishes the order of the error at each order of ϵ. The order is described by the list of the powers of the highest matching term in the Taylor series, for successive powers of ϵ. If the order is described as $(n_1, n_2, ...)$ the terms $\epsilon^i \Delta t^{n_i}$ are correctly represented. Termination of the list indicates that all following terms are the same order as the last term. The methods developed in the previous section are, in this notation, $(\infty, 2)$, when the explicit closed form generator can be used. The particular Lie series composition formula we have presented for the corrector gives a composition integrator of order $(16, 2)$, but the method described can be used to achieve any desired order to first order in ϵ. When it is straightforward to extend the method to any desired order we shall continue to speak of the the order as ∞. In McLachlan's notation, the order of the mapping Hamiltonian developed in this section is generally $(\infty, \infty, 3)$, but with $\alpha = 1/2$ the order of the mapping Hamiltonian is extended to $(\infty, \infty, 4)$.

9 Lie Series Correspondence

In this section we translate the results of the last section into the Lie series language introduced earlier. For simplicity we consider only the case of $\alpha = t/\Delta t = 1/2$.

We have deduced

$$\mathrm{e}^{A+B} = \mathrm{e}^{-C}\mathrm{e}^{-C_2}\mathrm{e}^{K}\mathrm{e}^{C_2}\mathrm{e}^{C}, \tag{9.1}$$

where the kernel of the mapping is

$$\mathrm{e}^{K} = \mathrm{e}^{A/2}\mathrm{e}^{B'}\mathrm{e}^{A/2}, \tag{9.2}$$

where,

$$B' = B + \frac{1}{24}\mathcal{L}_B\mathcal{L}_A B - \frac{1}{1440}\mathcal{L}_B\mathcal{L}_A^3 B + \cdots \tag{9.3}$$

the corrector C is the same as before,

$$C = L_W = \frac{1}{24}\mathcal{L}_A B - \frac{7}{5760}\mathcal{L}_A^3 B + \cdots \tag{9.4}$$

and

$$C_2 = -\frac{7}{5760}\mathcal{L}_B\mathcal{L}_A^2 B + \cdots. \tag{9.5}$$

Recombining terms we can check that the integrator has the declared properties, and confirm that the error in the Hamiltonian is of order $\epsilon^3\Delta t^5$. We find that the leading error terms in representing $L_H = A + B$ are

$$\frac{1}{360}\mathcal{L}_B^2\mathcal{L}_A^2 B + \frac{11}{5760}\mathcal{L}_A\mathcal{L}_B^3 A + \frac{1}{720}\mathcal{L}_B^4 A. \tag{9.6}$$

In general it will not be possible to evaluate the kernel directly, and we are naturally led to develop approximations by composing Lie series. Whereas the efficiency of the corrector is not critical at all, it is particularly important to find expansions of the kernel with as few steps as possible, so that the mapping will be efficient. Here, however, we are interested in testing the ideas and leave the optimal expansion of the kernels for later investigation.

One key constraint on the expansion of the kernel is that B' must not have any terms of the form $\mathcal{L}_A^m B$ for any $m > 0$. One way to guarantee this (but perhaps not the only way) is to use the corrector formula again

$$\mathrm{e}^{B'} = \mathrm{e}^{Y} \mathrm{e}^{B} \mathrm{e}^{-Y}, \tag{9.7}$$

with all terms in Y containing at least one factor of B. The form of B' severely constrains Y. In fact, up to terms of order ϵ^3 we can write $Y = \beta B + L_{W(t=0)}/2$, where β is arbitrary and will not enter the final kernel. The task of representing the kernel is reduced to representing Y:

$$Y = \beta B + \sum_{k>0} \frac{B_k(0)}{k!} \mathcal{L}_A^k B. \tag{9.8}$$

Now, except for the term involving β, Y contains only odd powers of $\mathcal{L}_A$ multiplying B. One way to obtain this is to represent e^Y as a product of two exponentials with one exponential argument obtained from the other by reversing the sign of both A and B, thus preserving only the even terms in the product at order ϵ. The product will have ϵ^2 terms but these become ϵ^3 when wrapped around the central factor of e^B. One way for each of these factors to be represented as exponentials with all terms in the exponential proportional to B, is to again use the corrector formula. Using X as defined before

$$\mathrm{e}^{X(a_1,b_1)} = \mathrm{e}^{a_1 A} \mathrm{e}^{b_1 B} \mathrm{e}^{-a_1 A}. \tag{9.9}$$

All terms in X are proportional to B. In terms of X the simplest expression for Y satisfying the stated constraints is

$$\mathrm{e}^{Y} = \mathrm{e}^{X(a_1,b_1)} \mathrm{e}^{X(-a_1,-b_1)}. \tag{9.10}$$

Combining factors, we find that the coefficients a_1 and b_1 must satisfy $a_1 b_1 = -1/48$ to match the $k = 2$ term, and the $k = 1$ and $k = 3$ terms are zero as required, but the $k = 4$ term cannot simultaneously be matched. So with this Y the error in the kernel is of order $\epsilon^2 \Delta t^5$. A convenient choice is: $a_1 = 1/8$ and $b_1 = -1/6$. The full kernel is then:

$$\mathrm{e}^{K} = \mathrm{e}^{\frac{5}{8}A} \mathrm{e}^{-\frac{1}{6}B} \mathrm{e}^{-\frac{1}{4}A} \mathrm{e}^{\frac{1}{6}B} \mathrm{e}^{\frac{1}{8}A} \mathrm{e}^{B} \mathrm{e}^{-\frac{1}{8}A} \mathrm{e}^{-\frac{1}{6}B} \mathrm{e}^{\frac{1}{4}A} \mathrm{e}^{\frac{1}{6}B} \mathrm{e}^{\frac{3}{8}A} \tag{9.11}$$

We cannot say if the same order (including the absence of spurious $\mathcal{L}_A^m B$ terms) could have been obtained with fewer factors. We consider ours a "trial" solution, to test the ideas, and leave open the issue of whether there is a better product kernel at this order.

Consider next the problem of representing the correctors. We are interested in the problem of determining a corrector that is consistent in order with the kernel just derived. Thus, we require that the ϵ^2 terms up to fourth order in Δt be correctly represented, so that the error in the integrator remains of order $\epsilon^2 \Delta t^5$. The second corrector contributes terms of order $\epsilon^2 \Delta t^5$ to the integrator, so it may be ignored for the moment. The corrector we have already derived is thus of the required order.

Though the mapping Hamiltonian is correct to order $(\infty, \infty, 4)$, the practical implementation of this mapping in terms of the composition of Lie series that we have found is limited to order $(\infty, 4)$. It is a fourth order method with error proportional to ϵ^2. Implementation of the kernel to higher order as a composition is straightforward, but not pursued here. In this sense the method is still theoretically $(\infty, \infty, 4)$.

10 Special Cases

Some simplifications occur for Hamiltonians of the form $H_0 = T(p) + V_0(q)$ with quadratic T and $H_1 = V_1(q)$. The n-body problem has this form, where H_0 represents the Keplerian motion of each planet with respect to the Sun, and H_1 represents the planetary perturbations.

The Hamiltonian that governs the evolution across the delta functions is modified by a term proportional to $\{W(t=0), H_1\}$. The leading term of this series is proportional to $\{H_1, \{H_1, H_0\}\}$. For Hamiltonians of the special form just described this bracket depends only on the coordinates and can thus be integrated along with H_1 which also depends only on the coordinates. Unfortunately, the remaining terms in the expansion of $\{W, H_1\}$ are not zero, and not obviously integrable. Failure to include these terms results in a fourth order integrator (truncation error proportional to $\epsilon^2 \Delta t^5$) with an energy error of order $\epsilon^2 \Delta t^4$. Nevertheless, taking the special form of the first term into account provides a very efficient, very accurate fourth order integrator. The truncation error is of order ϵ^2, and there is only a single modified force evaluation per step. By comparison, the standard fourth order Runge-Kutta method has four force evaluations per step, and the Forest-Ruth fourth order symplectic scheme has three force evaluations; both of those methods have local truncation error that is first order in ϵ.

Explicitly, let

$$H_0 = \sum_i \frac{p_i^2}{2m_i} + V_0(q_1, ..., q_n), \tag{10.1}$$

and

$$H_1 = V_1(q_1, ..., q_n). \tag{10.2}$$

The Hamiltonian governing the evolution across the delta functions has the two leading terms

$$H_\delta = H_1 + \Delta t^2 \frac{1}{24}\{H_1, \{H_1, H_0\}\} = V_1 + \Delta t^2 \frac{1}{24} \sum_i \frac{1}{m_i} \left(\frac{\partial V_1}{\partial q_i}\right)^2 . \tag{10.3}$$

The modified "kick" is

$$\Delta p_j = -\Delta t \frac{\partial H_\delta}{\partial q_j} = -\Delta t \frac{\partial V_1}{\partial q_j} - \Delta t^3 \frac{1}{12} \sum_i \frac{1}{m_i} \frac{\partial V_1}{\partial q_i} \frac{\partial^2 V_1}{\partial q_i \partial q_j} . \tag{10.4}$$

For the n-body problem the extra terms are not expensive, since all the required square roots and inverses already have been computed to evaluate the original kick, so the modified kick involves no more than summing products of previously computed quantities.

Koseleff [1993] mentions that in the special case of $H = T(p) + V(q)$ there is a similar possibility of using a modified potential, but apparently, without using the correctors, order four can only be reached with two modified force evaluations. Here we reach fourth order with one modified force evaluation.

Wisdom and Holman [1991] raised the question of whether there were higher order methods that did not take backward steps. Their motivation was to maintain a Hamiltonian description of the algorithm. With negative steps, the representation as a delta function Hamiltonian is lost since the time order cannot be encoded in the scalar Hamiltonian. Yoshida [1993] subsequently answers the question negatively, by reference to a theorem by Suzuki [1991]. These theorems notwithstanding, we have here fulfilled both of the original goals we earlier set for ourselves. We have an accurate fourth order method that has a Hamiltonian representation, and furthermore involves only one force evaluation per step. It is interesting to note that the kernel involves no backward steps. In the Lie series implementation of the corrector, the backward steps required by the theorem are all hidden in the corrector. It is worth pointing out, though, that the corrector could also be implemented as a numerical integration of the generator, and then there would be no backward steps at all. Of course, in this case the theorems do not apply.

Next, we present an interesting alternate form of the modified kick. We note that the expression for the kick is the beginning of a multidimensional Taylor series expansion of the kick at a displaced point. Explicitly, denote the unmodified kick component functions by

$$k_j(x) = -\Delta t \left.\frac{\partial V_1}{\partial x_j}\right|_x \tag{10.5}$$

and let

$$x'_j = x_j + \frac{\Delta t}{12 m_j} k_j(x). \tag{10.6}$$

Then the modified kick is given simply by

$$p''_j = p_j + k_j(x'). \tag{10.7}$$

The coordinates are unchanged in this step

$$x''_j = x_j. \tag{10.8}$$

The intermediate values x'_j are discarded. Higher order terms are also generated, but they are of order ϵ^3, and so are ignorable at this stage. Strictly speaking, the error term lowers the order from fourth order to third order. Whether this matters depends of the magnitude of the $\epsilon^3 \Delta t^4$ error term compared to the $\epsilon^2 \Delta t^5$ error term. By combining three force evaluations we can keep the method fourth order. For the lazy implementer, this expression gives an approximation of the modified kick with very little extra programming effort. The cost is that two force evaluations are required this way. This version of the map is still symplectic since the change in p is only a function of x, and x is not modified. The two-force-evaluation scheme is of order $(\infty, 4, 3)$, the three-force-evaluation scheme is of order $(\infty, 4, 4)$. We do not have a direct generalization of this method to higher order.

These schemes look a lot like Runge-Kutta schemes in the sense that forces are evaluated at intermediate positions which are discarded in the final step. So in this sense we have an explicit symplectic generalization of the Runge-Kutta method that is third or fourth order in stepsize (depending on the form used), with error terms proportional to ϵ^2.

11 Application to the Restricted Three-Body Problem

In this section we illustrate the use of the higher order mapping derived in the previous three sections. This mapping is a fourth order integrator that has truncation error proportional to ϵ^2. We apply the mapping to the integration of the planar circular restricted three-body problem: an infinitesimal mass moving in the orbit plane of two massive bodies which themselves move in a circular relative orbit. In this illustration, the mass of one of the bodies is small: 0.001 times the total mass of the system. This is a crude model for the motion of asteroids perturbed by Jupiter. We split the Hamiltonian into two-body Kepler motion with respect to the "Sun" and perturbations from "Jupiter," as in the n-body problem (see Wisdom and Holman [1991]).

In an inertial frame, the circular restricted problem has two degrees of freedom, with explicit time dependence. In the frame rotating with the massive bodies the time dependence is removed, and the value of the Hamiltonian in the rotating frame is consequently conserved. The Hamiltonian in the rotating frame is proportional to what is traditionally called the "Jacobi" constant. It is natural to test integrators by the extent to which they preserve this integral.

To be specific, we choose to study an initially circular orbit with semimajor axis 0.63 times the distance between the massive bodies. The test particle is started on the line between "Jupiter" and the "Sun." This orbit is close to an orbital resonance, and consequently develops moderate eccentricity.

The relative error in the Jacobi constant is plotted against the number of steps per orbit in Figure 5. The integrations spanned 200 orbits. The curve at the top is the maximum relative error in the Jacobi constant using the uncorrected, unmodified second order mapping (M2 - for second-order map). The next curve down is the maximum relative error for the second order map used in conjunction with a corrector (CM2 - Corrected Map of $\underline{2}^{nd}$ order). The same corrector was used here as in the n-body example. The bottom curve is actually three curves superimposed. This is the relative error using the modified kick in conjunction with the corrector (CMM4 - Corrected Modified Map of $\underline{4}^{th}$ order). It is also the error obtained using the two-force-evaluation Runge-Kutta implementation of the modified kick as described in the last section. Finally, the bottom curve is also the error using the general purpose product kernel. The differences between the errors of these latter three methods are too small to be visible on this scale. Note that the Jacobi constant relative error of the second order map and the corrected second order map are indeed quadratic in step-size as expected. Also note that the error of the last three methods are all fourth order, even the two-kick method which could have have a third order component proportional to ϵ^3. Of course, among the higher order methods, the most efficient is the mapping with the explicitly modified kick, followed by the two-force-evaluation Runge-Kutta approximation, followed by the general purpose five kick Lie series product kernel. Note how ineffective it would be to try to eliminate error by reducing the stepsize alone. Whereas, using the corrector with some version of the modified kick is very effective at reducing the error.

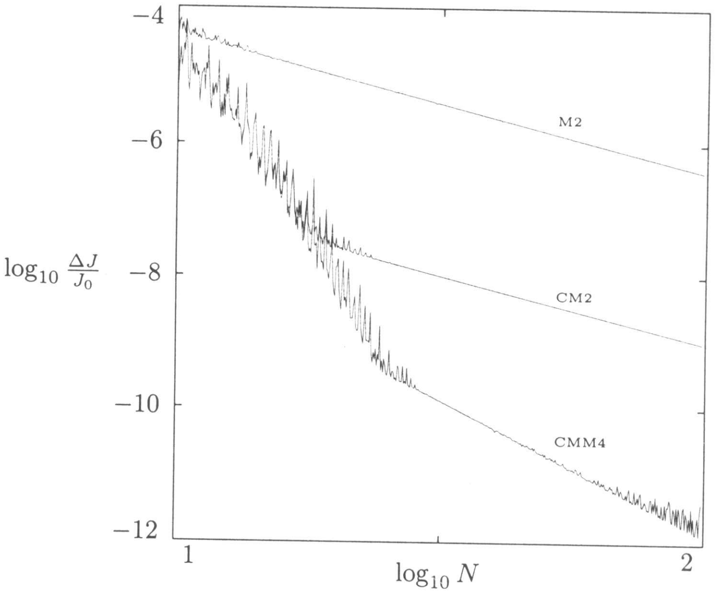

Figure 5 The relative error in the Jacobi constant versus number of steps per orbit, for several methods. The top curve is the error for the second order mapping (M2). The middle curve is the error for the second order mapping used in conjunction with a corrector (CM2). The bottom curve is the error for the fourth order corrected modified mapping (CMM4), the two-force-evaluation Runge-Kutta approximation to the modified kick with corrector, and the product kernel.

The most interesting result is not that the error behaved as expected, but rather how the observed error differs from that expected. The error of all the methods seem to be dominated by some "unexpected" error at large step sizes. This error could have several origins, and we are not sure which is the culprit. It could be a failure of the corrector to have high enough order. This does not seem to fit. The corrector we are using is correct to Δt^{17}, but the observed falloff is more like Δt^{13}. It could result from the fact that our derivation of the general purpose corrector formulas used here presumed that the mapping frequency is large compared to the frequencies present in the motion so that we could expand in the ratio of these frequencies. Thus we may be seeing what happens when that assumed condition is not satisfied. Another possibility is that we are seeing error introduced by stepsize resonances (Wisdom and Holman [1992]), which are most apparent at large stepsizes. We suspect that this is indeed a stepsize resonance problem.

We note that we have seen a similar high order falloff in a corresponding study of the energy error in the solar system as a function of stepsize, though the tests are much more limited. At the adopted stepsize of 7.2 days, the corrector seems to have less effect on the solar system integration error than we a priori expected (the corrector only improved the mean motion of Mercury by a factor of 10, and the energy by a factor of 100), but with a stepsize of 3.6 days the error dramatically drops and thereafter at smaller stepsizes the error falls off as a fourth order integrator should. The same phenomenon is seen here, but is more easily demonstrated since

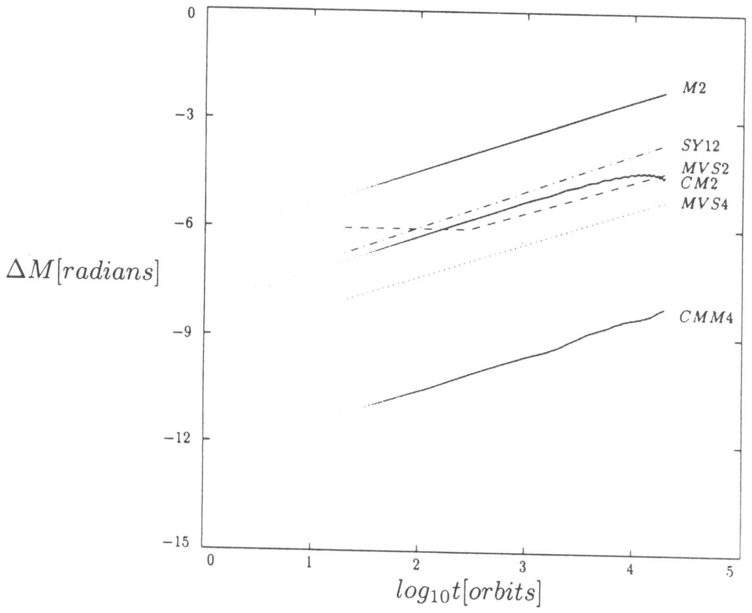

Figure 6 The error in the mean anomaly versus time of a fictitious asteroid in the three-dimensional circular restricted three-body problem for various methods.

the integrations take less time. In the case of the solar system we confirmed that higher order correctors did not solve the problem. Had we used only a factor of two smaller step in our 100 million year integration of the solar system, we would have had dramatically smaller errors!

Next, we consider the errors in longitude. Here we investigate the orbit studied by Saha and Tremaine [1992]. They were also interested in removing the effect of the spurious oscillations exhibited by the mapping trajectories. By "warming up" the trajectories they cleverly removed some of the ill effects of these oscillations without explicitly modelling the oscillations or removing them. By studying the same orbit we can compare the efficacy of our corrector method to the "warmup" method. The infinitesimal body orbits the "Sun" with a semimajor axis of about half the distance between the massive bodies. The orbit has moderate eccentricity $e = 0.25$, and moderate inclination $i = 0.2$ radians with respect to the plane of the massive bodies. They used roughly 100 force evaluations per orbit.

Figure 6 shows the error in mean anomaly of the test-particle for various methods as a function of time. In addition to the errors for the mapping methods, the errors for several other methods are also displayed. These are approximations to the data displayed in the paper by Saha and Tremaine. The method MVS2 refers to the second order Wisdom and Holman [1991] mapping with "warmup," the method MVS4 is the Forest and Ruth [1990] fourth order method with Wisdom and Holman [1991] components using "iterated start," and SY12 is a twelve step symmetric method (Quinlan and Tremaine [1990]). The performance of the corrector is comparable to that of the warmup procedure in eliminating spurious drift of the longitudes. Of course, the corrector does more than fix the longitude drift, it removes the spurious oscillations from the trajectory itself, so all aspects of the calculation are improved. Also shown is the error in longitude when the modified kick is used. Here the modified kick is programmed explicitly. The reduction in the error is dramatic. The modified kick reduces the error by roughly six orders of magnitude below the uncorrected second order map. The illustrations by Saha and Tremaine were all limited by truncation error rather than roundoff error; partial extended precision was used in this last calculation so that the error would continue to be dominated by truncation error rather than by roundoff error.

12 Suggestions for Future Research

The delta function formalism has suggested the use of correctors, and previous sections have directly explored this idea with a perturbative treatment of mapping Hamiltonians involving delta functions. This formalism leads to a particular set of integrators and a particular set of correctors. Of course, many other symplectic integrators are derived purely by matching Taylor series. Unfortunately, with integrators based solely on Taylor series matching there is no clear identification of which error terms are high-frequency and which would lead to secular growth of error. Nevertheless, we can also develop correctors for these integrators. In this case, since matching the Taylor series is the sole criterion for an integrator we could develop correctors that compensate for as many terms as possible. We may also develop composition integrators which are intended from the beginning to be used with a corrector. Presumably, the latter approach is the most economical path to high order, since the kernel must satisfy a much reduced set of constraints, only that the error terms be correctable, not that they be zero. Following this line of investigation, it would be interesting to determine whether the particular kernels suggested by the delta function formalism behave better or worse than kernels based purely on Taylor series order.

Another interesting avenue of investigation, in many ways related to the one presented here, is to look for composition methods that explicitly take into account the relative magnitudes of H_0 and H_1. Koseleff [1993] and McLachlan [1995b] have already pursued this line of investigation.

13 Summary and Exhortation

The spurious oscillations exhibited by symplectic integration schemes have a clear origin in the delta function formulation. We have developed canonical transformations that remove the spurious oscillations. For perturbations with simple Poisson series we give explicit closed form expressions for the generator of these canonical transformations. For more complicated problems, we have developed general expressions for the correctors that are coordinate independent and valid for any number of degrees of freedom.

We have illustrated the use of these correctors with applications to the n-body problem. Applied to the old data from the 100 million year integration of the solar system of Sussman and Wisdom [1992] the correctors have allowed us to reduce the apparent energy error in that integration by two orders of magnitude. Even more dramatic reduction of error is possible if the correctors are used with forethought.

The mapping method was originally based on first order averaging arguments. The mapping method is here extended to second order. A very efficient, very accurate fourth order mapping is obtained. Several different forms of the fourth order mapping are presented: a general purpose Lie series product formula, an explicit symplectic Runge-Kutta formula, and a single modified kick formulation. The latter two are special forms for problems such as the n-body problem. The modified kick mapping achieves fourth order with a single force evaluation. Application to the restricted three-body problem illustrates the dramatic decrease in the error in energy and longitude achieved by this higher order mapping.

The delta function formalism has been uniformly ignored in reviews of the symplectic integration literature, presumably because it is considered non-rigorous. Yet, the delta function formalism gives algorithms that are equivalent to the composition of Lie series approach. Furthermore, the delta function formalism gives clear insight into problems faced by symplectic integration schemes, and, more importantly, it provides solutions. Knowledge of their names gives power over demons. With the delta function formalism we can determine non-linear stability where other avenues lead only to linear results. With the delta function formalism we can quantitatively understand the origin of artifacts introduced by the algorithmic discretization. Practitioners of other methods simply shake their heads in despair. With the delta function formalism we understand the origin of the spurious oscillations induced by the discretization, and the delta function formalism has here provided a way to remove the oscillations and dramatically improve the accuracy of symplectic integrators. However it may appear to some, the delta function formalism is not necromancy. You might try whispering "delta functions" three times over your notes to see if the problems will disappear, but a better idea is to learn to use delta functions, and appreciate the wisdom they have to offer!

Acknowledgements

We thank Robert McLachlan for carefully reading the manuscript and making numerous insightful and scholarly suggestions. We also thank Presenjit Saha and Scott Tremaine for helpful discussions. This research was supported in part by grants from NASA and the National Science Foundation.

References

Belinfante, J. G. and Kolman, B. [1989], *A Survey of Lie Groups and Lie Algebras with Applications and Computational Methods.* SIAM, Philadelphia.

Butcher, J. C. [1969], In Lecture Notes in Mathematics, (Dold, A., Heidelberg, Z., and Eckmann, B., eds.), Springer-Verlag, New York.

Candy, J. and Rozmus, W. [1991], *A Symplectic Integration Algorithm for Separable Hamiltonian Functions*, J. Comp. Phys., **92**, 230-256.

Chirikov, B.V. [1979], *A Universal Instability of Many-Dimensional Oscillator Systems*, Phys. Rep., **52**, 263.

Forest, E. and Ruth, R. D. [1990], *Fourth-order Symplectic Integration*, Physica D, **43**, 105-117.

Hairer, E., Norsett, S. P., and Wanner, G. [1993], *Solving Ordinary Differential Equations I. Nonstiff Problems.* Springer-Verlag, Berlin.

Koseleff, P.-V. [1993], *Relations among Lie Formal Series and Construction of Symplectic Integrators*, In Applied Algebra, Algebraic Algorithms, and Error-correcting Codes (Cohen, G., Mora, T. and Moreno, O. eds.), AAECC-10, Springer-Verlag, New York.

Laskar, J. [1989], A *numerical experiment on the behaviour of the solar system*, Nature, **338**, 237-238.

McLachlan, R. I. [1995a], *On the numerical integration of ordinary differential equations by symmetric composition methods*, SIAM J. Sci. Comp., **16**, 151–168.

McLachlan, R. I. [1995b], *Composition methods in the presence of small parameters*, submitted to BIT.

Quinn, T. R., Tremaine, S. D. and Duncan, M. [1991], *A three million year integration of the Earth's orbit*, Astron. J., **101**, 2287-2305.

Quinlan, G. D. and Tremaine, S. D. [1990], *Symmetric multistep methods for the numerical integration of planetary orbits*, Astron. J., **100**, 1694-1700.

Saha, P. and Tremaine, S. D. [1992], *Symplectic integrators for solar system dynamics*, Astron. J., **104**, 1633-1640.

Sanz-Serna, J. M. [1992], *Symplectic integrators for Hamiltonian problems: an overview*, Acta Numer., **1**, 243-286.

Steinberg, S. [1988] *Lie Methods in Optics*, (Mondragon and Wolf, eds.). Springer-Verlag, New York.

Sussman, G. J. and Wisdom, J. [1988], *Numerical Evidence that the Motion of Pluto is Chaotic*, Science, 241:433.

Sussman, G. J. and Wisdom, J. [1992], *Chaotic Evolution of the Solar System*, Science, **257**, 56.

Suzuki, M. [1991], *General theory of fractal path integrals with applications to many-body theory and statistical physics*, J. Math. Phys., **32**, 400-407.

Tittemore, W. C. and Wisdom, J. [1988], *Tidal Evolution of the Uranian Satellites. Part I: Passage of Ariel and Umbriel through the 5:3 Mean-Motion Commensurability*, Icarus, **74**, 172.

Tittemore, W. C. and Wisdom, J. [1989]. *Tidal Evolution of the Uranian Satellites. Part II: Explanation of the Anomalously Large Inclination of Miranda*, Icarus, **78**, 63.

Tittemore, W. C. and Wisdom, J. [1990], *Tidal Evolution of the Uranian Satellites. Part III, Evolution through the Miranda-Umbriel 3:1, Miranda-Ariel 5:3, and Ariel-Umbriel 2:1 Mean Motion Commensurabilities*, Icarus, **85**, 394.

Wisdom, J. [1982], *The Origin of the Kirkwood Gaps: A Mapping for Asteroidal Motion Near the 3/1 Commensurability*, Astron. J., **87**, 577.

Wisdom, J. [1983], *Chaotic Behavior and the Origin of the 3/1 Kirkwood Gap*, Icarus, **56**, 51.

Wisdom, J. [1988], *Fire*, In Earth, Air, Fire, and Water (Roberts, P. and Schubert, G., eds.), Proceedings of the Fourth Annual University of California Summer School on Nonlinear Science. Institute of Geophysics and Planetary Physics, UCLA.

Wisdom, J. and Holman, M. [1991], *Symplectic Maps for the N-Body Problem*, Astron. J., **102**, 1528.

Wisdom, J. and Holman, M. [1992], *Symplectic Maps for the N-Body Problem: Stability Analysis*, Astron. J., **104**, 2022.

Yoshida, H. [1990]. *Construction of higher order symplectic integrators*, Phys. Lett. A, **150**, 262–268.

Yoshida, H. [1993], *Recent Progress in the Theory and Application of Symplectic Integrators*, Celest. Mech. Dyn. Ast., **56**, 27-43.