CITA-95-3

Is the solar system stable?

Scott Tremaine

Canadian Institute for Theoretical Astrophysics,
University of Toronto and Institute of Astronomy, Cambridge

"Rosseland Centenary Symposium of Astrophysics," Oslo,
June 16 - 17, 1994.

Is the solar system stable?

Scott Tremaine

Canadian Institute for Theoretical Astrophysics, University of Toronto
and Institute of Astronomy, Cambridge

The question in the title dates back to Newton's *Opticks*: 'the Planets move one and the same way in Orbs concentric, some inconsiderable Irregularities excepted, which may have arisen from the mutual Actions of...Planets upon one another, and which will be apt to increase, till this System wants a Reformation'. In more modern language,

A point mass M is surrounded by N masses $m_i \ll M, i = 1, \ldots, N$ on nearly circular, nearly coplanar orbits. Is the configuration stable over very long times?

For the solar system $N = 9$, $\max(m_i/M) \simeq 0.001$, and 'very long times' means 10^{10} orbits. This problem has attracted many famous mathematicians over the past three centuries (Laplace, Lagrange, Poincaré, Arnold, etc.) and has played a central role in the development of nonlinear dynamics and chaos theory. The present review focuses on its astronomical implications: in this context, the principal tool is numerical experiment and the mathematical theorems mainly provide a conceptual framework for interpreting the experimental results. For other recent reviews see Ferraz-Mello (1992) and Duncan and Quinn (1993).

Equations of motion and initial conditions

We wish to solve the equations of motion

$$\ddot{\mathbf{x}}_i = \sum_{\substack{j=0 \\ j\neq i}}^{N} \frac{Gm_j}{|\mathbf{x}_j - \mathbf{x}_i|^3}(\mathbf{x}_j - \mathbf{x}_i) + \mathbf{a}_i, \tag{1}$$

where $i =$ (Sun, Mercury, ..., Pluto) and $\mathbf{a}_i$ represents any small extra acceleration arising from sources other than Newtonian gravitational interactions between the planets and the Sun. These include:

(i) Satellites. In general, satellites can be lumped in with their parent planet, so the masses m_i in equation (1) are the total masses of planets plus their satellites. This approximation mainly neglects the solar attraction on the quadrupole moment of the planet-satellite system; the fractional size of this error relative to the solar attraction is $\sim (m_s/m_p)(r_s/r_p)^2$ where m_s and r_s are the satellite's mass and orbital radius and m_p and r_p are the analogous parameters for the planet. The fractional error is

$\sim 10^{-7}$ for the Moon and far smaller for all other satellites. An approximate formula for the extra acceleration of the Earth-Moon system is given by Quinn et al. (1991) and included in their equations of motion; Touma and Wisdom (1994) have carried out a long integration that explicitly follows the lunar orbit. Estimates of the lunar contribution to the dynamics are plagued by the uncertain tidal evolution of the lunar orbit.

(ii) General relativity. The dominant corrections to the equations of motion have fractional size $GM_\odot/(c^2 r) = 10^{-9}(1\,\text{AU}/r)$ and are given in Quinn et al. (1991) and Saha and Tremaine (1994). More accurate corrections are given by Will (1981) and Newhall et al. (1983).

(iii) Asteroids. The total mass of the asteroids is $\sim 10^{-9} M_\odot$, much of it in the few largest. Most integrations simply neglect the asteroids, although the gravitational forces from the largest few bodies are included in the most accurate ephemerides (e.g. Newhall et al. 1983). Unmodeled asteroids provide the dominant source of 'noise'.

(iv) The Galaxy. The dominant component of the Galactic tidal acceleration is along the direction $\mathbf{e}_z$ normal to the Galactic plane and equals $-4\pi G\rho z \mathbf{e}_z$ where $\rho \simeq 0.15 M_\odot\,\text{pc}^{-3}$ is the local Galactic density; its amplitude relative to the solar acceleration is $\sim 4\pi\rho r^3/M_\odot \simeq 10^{-11}(r/40\,\text{AU})^3$, which is negligible.

(v) Passing stars. In the lifetime of the solar system, the nearest approach of a passing solar-type star has been $r \sim 1\times 10^3$ AU; at this distance it exerts a tidal acceleration on the planets of fractional amplitude $\sim 10^{-4}(r/40\,\text{AU})^3$, which lasts for $\sim 100\,\text{yr}$. The effects of such perturbations should be insignificant, especially since the timescale is generally long enough that the semi-major axes of most planets will be adiabatic invariants.

(vi) Solar mass loss. The Sun loses mass at a rate $dM_\odot/dt \simeq 10^{-13} M_\odot\,\text{yr}^{-1}$, mostly in photons but also through the solar wind. Mass loss causes the planetary orbits to expand slowly. To a first approximation, the expansion preserves the relative frequencies of the planetary orbits and hence should not have a strong qualitative effect on the orbit evolution.

To solve the equations of motion (1) we need initial conditions and planetary masses. These are obtained from optical observations of the outer planets (which give angular positions but not ranges), radar observations of the inner planets (which give ranges but not angular positions), and radio tracking of spacecraft that orbit or fly by the planets. In particular, the *Voyager* spacecraft has provided us with the first accurate masses for all four of the giant planets (fractional errors $\lesssim 10^{-5}$). The mass of the Pluto-Charon system is much less well-known (fractional error 0.5%) but is so small that its effects on the other planets are negligible.

The parametrization of the 'accuracy' of a solution is a subtle issue. Certainly we cannot hope that our initial conditions, planetary masses, and numerical methods are good enough that the errors in planetary positions after 10 Gyr will be a small fraction of a radian, even if the planetary trajectories were regular (which they are not). A more appropriate criterion is based on the frequencies of the lines in the power spectra of the planetary trajectories $\mathbf{x}_i(t)$. A realistic goal is frequencies that are accurate to $\sim 10^{-7}$

to $10^{-8}\,\mathrm{yr}^{-1}$; with this accuracy, the positions of the planets are unreliable beyond 10 to 100 Myr, but our integrations should provide useful insight into the evolution of the solar system for much longer times.

Numerical methods

Recent integrations of the solar system take up to 10^{10} timesteps and by this measure are probably the most ambitious numerical solutions of ordinary differential equations ever made; thus it is essential to use an accurate and efficient algorithm. A principal consideration is to minimize the number of expensive force evaluations per unit time—for nine planets there are 45 forces to be calculated. One simplification is that a variable timestep is not needed, because the planetary orbits are nearly circular and well-spaced.

Multistep integrators The traditional algorithm is the Störmer multistep method: if the acceleration is $\ddot{x}$ then the position at timestep $n+1$ is given by

$$x_{n+1} = 2x_n - x_{n-1} + h^2 \sum_{j=0}^{k-1} \beta_{k-1-j}\ddot{x}_{n-j}, \tag{2}$$

where h is the timestep. The coefficients β_i are chosen so that the difference equation is satisfied by a polynomial $x(t)$ of as high an order as possible. Because the acceleration varies smoothly, a high-order scheme can be used (typically $k=13$); the strength of the algorithm is that each force evaluation is re-used k times in subsequent timesteps.

The principal defect of this scheme is most easily seen by considering motion in a harmonic potential $\ddot{x} = -\omega^2 x$. In this case (2) becomes a linear difference equation with constant coefficients, solved by a linear combination of functions of the form $x_n = z_l^n$, $l = 1,\ldots,k$, where $z_l(h)$ are the roots of the characteristic equation. The two principal roots are close to the true solution $\exp(\pm i\omega h)$, while the others represent spurious oscillations. The problem is that in general $|z_l|$ is not quite unity for the principal roots, so that the solution slowly spirals inward or outward. In other words the Störmer algorithm introduces numerical dissipation, a crippling problem for very long integrations.

Dissipation is absent in symmetric multistep methods (Lambert and Watson 1976, Quinlan and Tremaine 1990), which have the form

$$x_{n+1} = -\sum_{j=0}^{k-1} \alpha_{k-1-j}x_{n-j} + h^2 \sum_{j=0}^{k-1} \beta_{k-1-j}\ddot{x}_{n-j},$$

where

$$\alpha_i = \alpha_{k-i}, \quad \alpha_0 = 1 \quad ; \quad \beta_i = \beta_{k-i}, \quad \beta_0 = 0. \tag{3}$$

The restrictions (3) ensure that all of the roots of the characteristic equation lie on the unit circle; or, interpreted more broadly, that the method is time-reversible: if the initial conditions $x_1, \ldots, x_k$ generate a trajectory $x_1, \ldots, x_M$ then the initial conditions $x_M, \ldots, x_{M-k+1}$ generate a trajectory that returns to x_1 (apart from roundoff error). One defect of high-order symmetric multistep methods is that they behave badly near certain resonant values of the timestep (Quinlan and Toomre 1994); however, these should be easy to avoid in practice.

Symplectic integrators The equations of motion (1) can be derived from a Hamiltonian. The resulting trajectories $\mathbf{w}(t) \equiv [\mathbf{x}(t), \mathbf{v}(t)]$ therefore conserve the Poincaré invariants (the simplest of which is phase-space volume) because the map $\mathbf{w}(0) \to \mathbf{w}(t)$ is canonical or symplectic (Arnold 1978). If an integrator is a symplectic map as well, then it automatically conserves all the Poincaré invariants; integrators boasting this feature are known as symplectic integration algorithms (SIAs; see Channell and Scovel 1990 and Yoshida 1993 for reviews).

For example, consider the Hamiltonian $H(q,p) = \frac{1}{2}p^2 + U(q)$, which leads to the equations of motion

$$\dot{p} = -\frac{\partial U(q)}{\partial q}, \qquad \dot{q} = p.$$

The simplest SIA for this system, with timestep h, is the modified Euler method:

$$q_{n+1} = q_n + hp_n, \qquad p_{n+1} = p_n - h\frac{\partial U}{\partial q}(q_{n+1}). \tag{4}$$

This map is clearly symplectic since it can be derived from the Hamiltonian $H(q,p) = \frac{1}{2}p^2 + hU(q)\delta_h(t)$, where $\delta_h(t) = \sum_{j=-\infty}^{\infty} \delta(t - jh)$ is the periodic delta function, by identifying step n with time $t = nh + 0^+$. We denote the map (4) by $L_1(h) : (q_n, p_n) \to (q_{n+1}, p_{n+1})$, the subscript '1' indicating that it is first-order accurate, i.e. it agrees with the Taylor expansion of the exact solution to first order in h. The map is not time-reversible: $L_1(-h)L_1(h) \neq 1$.

A better algorithm is

$$q_{n+1/2} = q_n + \tfrac{1}{2}hp_n, \quad p_{n+1} = p_n - h\frac{\partial U}{\partial q}(q_{n+1/2}), \quad q_{n+1} = q_{n+1/2} + \tfrac{1}{2}hp_{n+1}, \tag{5}$$

which can be derived from the Hamiltonian $H(q,p) = \frac{1}{2}p^2 + hU(q)\delta_h(t + \frac{1}{2}h)$. This map, denoted $L_2(h)$, is the familiar 'leapfrog' integrator, which is symplectic, second-order accurate, and time-reversible ($L_2(h)L_2(-h) = 1$).

Higher order time-reversible SIAs can be obtained by the composition of leapfrog operators with different timesteps. The simplest example is the fourth-order integrator (Forest 1987)

$$L_4(h) = L_2(ah)L_2(bh)L_2(ah), \tag{6}$$

with $a = 1/(2 - 2^{1/3})$, $b = -2^{1/3}/(2 - 2^{1/3})$. Sixth- and eighth-order integrators are described by Yoshida (1990).

Leapfrog can be viewed more generally. Consider a Hamiltonian of the form

$$H = H_A + H_B \tag{7}$$

where H_A and H_B are separately integrable. Let the operator $A(h)$ advance the phase-space trajectory over a time h under the influence of H_A, with a similar definition for $B(h)$; formally

$$A(h)z = \exp(h\{z, H_A\}), \qquad B(h)z = \exp(h\{z, H_B\}),$$

where $z = (q, p)$ and $\{\ ,\ \}$ is the Poisson bracket.

In the examples we have examined so far

$$H_A = \tfrac{1}{2}p^2, \qquad H_B = U(q).$$

Thus

$$A(h) : (q, p) \to (q + hp, p), \quad B(h) : (q, p) \to (q, p - h\partial U/\partial q).$$

The maps in equations (4), (5), and (6) are the operators

$$L_1(h) = B(h)A(h); \quad L_2(h) = A(\tfrac{1}{2}h)B(h)A(\tfrac{1}{2}h);$$

$$L_4(h) = A(\tfrac{1}{2}ah)B(ah)A(\tfrac{1}{2}[a+b]h)B(bh)A(\tfrac{1}{2}[a+b]h)B(ah)A(\tfrac{1}{2}ah);$$

thus L_2 and L_4 are examples of the numerical technique known as operator splitting.

Since the planets travel on nearly Keplerian orbits, we would like to have an integrator that follows Keplerian orbits exactly. One way to achieve this goal is to set H_A in equation (7) equal to the part of the Hamiltonian that is first order in the planetary masses (the kinetic energy, and the potential energy of the interactions of the planets with the Sun) and H_B equal to the part that is second order (the potential energy of the mutual planetary interactions). Then H_A is integrable because it is the sum of independent two-body Hamiltonians, and H_B is integrable because it depends only on coordinates, not momenta (Wisdom and Holman 1991, Kinoshita et al. 1991). To implement this approach we must in effect perform a canonical transformation from Kepler elements (in which we evaluate the operator A) to Cartesian coordinates (to evaluate B) and back at every timestep—hence such integrators are sometimes called 'mixed variable symplectic' or MVS methods. Wisdom and Holman show that the gains from integrating the Keplerian orbit exactly far outweigh the computational cost of the transformations (in part because the transformations scale only as N, while the number of gravitational interactions scales as N^2).

An additional advantage of MVS methods is that they follow exactly the dynamics of a 'surrogate' Hamiltonian $H_{\rm surr} = H + H_{\rm err}$, with $H_{\rm err}/H = \mathrm{O}(\epsilon)$, so that the errors can be analyzed using Hamiltonian perturbation theory (Wisdom and Holman 1992). For example, $H_{\rm err}$ has no secular terms at $\mathrm{O}(\epsilon)$, so the relation between actions and frequencies is the same in H and $H_{\rm surr}$. Thus the dominant long-term errors in planetary positions are removed by ensuring that the actions are the same in the actual and surrogate systems, which can be done by a startup procedure that changes adiabatically from H to $H_{\rm surr}$—for example, integrate backwards using a very small timestep, while slowly reducing the interplanetary interactions to zero, then integrate forward using the timestep desired while reviving the planetary interactions at the same rate (Saha and Tremaine 1992, 1994). Wisdom et al. (1994) describe a related technique called 'symplectic correction' that employs a canonical transformation to annihilate the $O(\epsilon)$ part of $H_{\rm err}$. They apply the transformation to the initial conditions, integrate as usual, and invert the transformation only when output is desired.

A final refinement is to allow individual timesteps for each planet (Saha and Tremaine 1994): the orbital periods of the planets range over a factor of 10^3 so there is no point in following Pluto with the same small timestep as Mercury.

A general lesson is that it is better for integrators to reflect the important features of the real dynamical system—time-reversibility, symplecticity, a Hamiltonian, the integrability of Kepler orbits—than to minimize the errors over a single timestep.

A state-of-the-art integration of the nine planets can achieve integration errors of $\lesssim 0.03$ radians in the planetary longitudes after 1 Myr using a 7-day timestep for Mercury, and runs at a rate of 50 Myr per week on a 1993 workstation. The current method of choice is probably a fourth-order MVS integrator with individual timesteps that employs warmup or a symplectic corrector to determine the initial conditions.

Despite impressive progress in integration methods over the past five years, several issues remain to be explored:

(i) Techniques for controlling roundoff error in multistep integrations are described by Applegate et al. (1986) and Quinn and Tremaine (1990), but these cannot be applied directly to MVS methods, where the dominant roundoff arises in the repeated conversion between Kepler elements and Cartesian coordinates. With the optimistic assumption that roundoff leads to a random walk in energy, we expect energy diffusion of order $\Delta E/E \approx \epsilon(T/h)^{1/2} \approx 10^{-10}(T/1\,\mathrm{Gyr})^{1/2}$, where $\epsilon = 2^{-53} \simeq 10^{-16}$ is the precision of IEEE REAL*8 variables, T is the integration time, and $h \simeq 7$ days is the timestep. With the pessimistic assumption of linear drift in energy, $\Delta E/E \approx \epsilon(T/h) \approx 10^{-5}(T/1\,\mathrm{Gyr})$. It is important to understand how to control roundoff error in MVS integrations.

(ii) Leapfrog-based methods use each force evaluation only once, in contrast to multistep methods which achieve high order by recycling force evaluations several times in successive timesteps. Perhaps multistep mixed-variable methods are more efficient than leapfrog-based methods.

(iii) A related question is the relative importance of symplecticity and time-reversibility in integration algorithms—symmetric multistep methods are time-reversible but not obviously symplectic, while the modified Euler map (4) is symplectic but not reversible. Symplecticity is no guarantee of good behaviour: for example, symplectic methods with variable timestep lead to energy drift when integrating eccentric Kepler orbits (Yoshida 1993). Perhaps time-reversibility is more fundamental.

(iv) Is there an efficient way to parallelize or distribute long solar system integrations?

(v) The solar system is nearly integrable, with fractional perturbations $\epsilon \lesssim 10^{-3}$. Despite the smallness of the perturbations, we are restricted to timesteps that are a fraction of the smallest planetary orbital period. Are there integration schemes that permit us to take timesteps larger than an orbital period when $\epsilon \ll 1$? One such approach is described next, but there may be others.

Secular perturbation theory

In an integrable potential, trajectories are confined to a 3-dimensional torus in the 6-dimensional phase space. Such trajectories are quasiperiodic: the power spectrum of $\mathbf{x}(t)$ is non-zero only at frequencies that are integer combinations $\sum_{i=1}^{3} k_i \Omega_i$ of three independent frequencies Ω_i.

In a spherical potential one of the three frequencies, say Ω_1, can be chosen to be zero (the orbit lies in a fixed plane so the nodal procession rate is zero). In a Kepler potential two of the three frequencies—Ω_1 and Ω_2—can be chosen to be zero (the apsidal precession rate is also zero). Thus all of the frequencies in the power spectrum are multiples of Ω_3.

Now consider a Kepler orbit that is subjected to perturbations of fractional strength ϵ. The power spectrum will contain components in two well-separated frequency ranges: 'fast' components with frequency $k_3\Omega_3 + \mathrm{O}(\epsilon)$, $k_3 \neq 0$—in the solar system the corresponding periods range from a fraction of a year to hundreds of years—and 'slow' components, which arise from terms with $k_3 = 0$ and have frequencies that are $\mathrm{O}(\epsilon)$.

The strengths of the perturbing potentials are comparable at the fast and slow frequencies, but the effects of the slow perturbations are much larger, because they act for a longer time. The approach of secular perturbation theory is therefore to average over the fast perturbations and treat only the slow perturbations. Unfortunately, secular theory requires expansions in powers of the eccentricity e and inclination i of the planetary orbits, so the algebra is formidable in all but the simplest approximations (an incidental consequence is that Pluto cannot easily be included—since it crosses Neptune's orbit the standard expansions do not converge).

The simplest theory, due to Lagrange, evaluates the secular Hamiltonian of the planetary system to $\mathrm{O}(e^2, i^2, m)$. To this order the Hamiltonian is that of two sets of N coupled harmonic oscillators: one set describing the eccentricity variations and the other describing the inclination variations (the semi-major axes are constant). The eigenfrequencies are all real, which implies that the solar system is stable for all time in this approximation; thus Lagrange's work provided the first significant conclusion about the long-term stability of a simplified model solar system.

The most sophisticated secular theory, due to Laskar (1985, 1989, 1990), is based on an expansion of the secular Hamiltonian to $\mathrm{O}(e^6, i^6, m^2)$, which contains 150,000 terms. The resulting equations of motion must be integrated numerically; the advantage over a direct N-body integration is that the timestep can be much larger. Comparisons of Laskar's secular theory with N-body integrations show very good agreement (Laskar et al. 1992, Sussman and Wisdom 1992).

Table. Numerical integrations of the solar system

authors	timestep	integration time (Myr)	model	machine
Eckert et al. (1951)	40 d	0.0004	outer 5 planets	mainframe
Cohen & Hubbard (1965)	40 d	0.12	outer 5 planets	mainframe
Cohen et al. (1973)	40 d	1.0	outer 5 planets	mainframe
Newhall et al. (1983)	0.25 d	0.0044	9 planets+Moon	mainframe
Kinoshita & Nakai (1984)	40 d	5.0	outer 5 planets	mainframe
Applegate et al. (1986)	40 d	217	outer 5 planets	special-purpose
Applegate et al. (1986)	–	3	8 planets (no Mercury)	special-purpose
Roy et al. (1988)	40 d	100	outer 5 planets	vector supercomputer
Richardson & Walker (1989)	0.5 d	2.0	9 planets	mainframe
Sussman & Wisdom (1988)	32.7 d	845	outer 5 planets	special-purpose
Laskar (1989,1990)	500 yr	200	secular theory	vector supercomputer
Quinn et al. (1991)	0.75 d	3.0	9 planets	workstation
Sussman & Wisdom (1992)	7.2 d	100	9 planets	special-purpose
Wisdom & Holman (1992)	1 yr	1100	outer 5 planets	workstation
Laskar (1994)	250 yr	25,000	secular theory	workstation

Results and implications

The Table summarizes some numerical investigations of the long-term evolution of the solar system. Many follow only the outer five planets (Jupiter to Pluto) since: (i) the masses of inner planets are so small that the outer planets form an independent dynamical system; (ii) the large masses of the outer planets suggest that interesting effects are more likely in this region; (iii) the orbital periods of the outer planets are longer so it is easier to follow the system for a given time.

A variety of machines has been used for these projects. Solar system integrations cannot (yet) be vectorized efficiently, so there is little advantage to a vector or parallel machine; thus fast workstations that can be dedicated to an integration for weeks or months are presently the best choice. Impressive results have also been obtained with two special-purpose machines constructed at MIT (Applegate et al. 1986, Sussman and Wisdom 1992).

The maximum timespan over which such calculations may be relevant is 4.5 Gyr backward (the age of the solar system) and 7.7 Gyr forward (the time until the Sun swallows Mercury and loses a significant portion of its mass; Sackmann et al. 1993). Although calculations based on secular theory now extend for up to 25 Gyr, the longest N-body integration is only 100 Myr, or 2% of the age of the solar system. Thus the conclusions described below must be treated cautiously.

The first important result is that all the planets are still there: none has been ejected, fallen into the Sun, or collided with another planet, and the overall configuration of the planetary system remains quite similar.

Nevertheless the behaviour of the planets is not boring. Sussman and Wisdom (1988) discovered that the trajectory of Pluto is chaotic: small changes grow exponentially, with an e-folding time (Liapunov time) of 20 Myr. Despite this chaotic behaviour, Pluto's semi-major axis, eccentricity and inclination appeared to vary fairly regularly over the 845 Myr integration. This apparent regularity is impressive, since small disturbances were amplified by a factor of $\exp(845/20) \approx 10^{18}$ over the integration, and suggests that the trajectory is restricted—at least for the timespan of the integration—to a narrow chaotic zone in phase space.

Pluto has the largest inclination and eccentricity of any planet and is trapped in a complicated set of resonances with Neptune, and it is tempting to dismiss Pluto's chaotic behaviour as a curiosity arising from one or more of these unique features. This view was shown to be false by Laskar (1989), who discovered that the motions of the four terrestrial planets are also chaotic, with an even shorter Liapunov time of 5 Myr. Although Laskar's result is based on secular theory, the chaotic behaviour is confirmed by N-body integrations (Sussman and Wisdom 1992). Sussman and Wisdom showed that the orbits of the four giant planets are chaotic as well, with Liapunov times of 5–20 Myr; since the outer planets do not exhibit chaotic behaviour in Laskar's calculations, we may infer that this chaos is driven by mean-motion rather than secular resonances.

Despite the chaos, the semi-major axes, eccentricities, and inclinations of the eight major planets appear to vary regularly over timescales from at least 100 Myr (inner planets) to 1 Gyr (outer planets). On longer timescales, Laskar (1994) finds large and irregular variations in the eccentricities and inclinations of the terrestrial planets, especially Mercury and Mars.

The presence of chaos in the orbits of the planets has profound implications for the structure and evolution of the solar system:

(i) The precise positions of the planets and the shapes of their orbits are unpredictable on timescales $\gtrsim 10^8$ yr; thus, for example, the impulse from the launch of a single interplanetary spacecraft changes the position of Earth by ~ 1 radian after ~ 200 Myr.

(ii) In general, chaotic dynamical systems with many degrees of freedom are unstable: the chaotic regions in phase space are connected so that the trajectory can wander over large distances ('Arnold diffusion'), although the diffusion time is generally large and uncertain. Thus it is highly probable that the solar system is unstable, although the timescale for macroscopic changes is likely to be many times its lifetime.

(iii) Our present understanding of the solar system is already sufficient to follow planetary orbits accurately for several times the Liapunov time. Thus more accurate measurements of the properties of the present solar system will not improve our understanding of its fate. The most we can do is to assign probabilities to various fates by integrating an ensemble of solar systems. A first effort of this kind has been made by Laskar (1994), who found a small but non-zero chance that Mercury would be ejected or collide with Venus in less than 5 Gyr.

(iv) If planets can be ejected in the future, then perhaps they were also ejected in the past. Low-mass planets are more likely to be ejected than high-mass ones, and a number of ejections of low-mass planets could have occurred without exciting the eccentricities or inclinations of the survivors to values larger than observed. Thus the planetary system may have looked quite different just after planet formation was complete, perhaps having many more planets than the current nine. Support for this view is provided by integrations that follow thousands of test particles on orbits between the giant planets for up to 10^9 yr (Levison and Duncan 1993, Holman and Wisdom 1993); these simulations show that almost all test particles between the orbits of Jupiter and Neptune are lost by ejection or collision over the lifetime of the solar system.

(v) Dynamical chaos may have played a role in solar system formation. Experiments with artificial solar systems prepared by placing the giant planets in randomly chosen circular orbits show that many such systems are unstable, and most exhibit chaos with Liapunov times far shorter than those observed in the solar system (Quinlan 1992). Thus some aspect of the planet formation process must have favoured the rare configurations that are only weakly chaotic—or perhaps only weakly chaotic planetary systems are hospitable to astronomers.

The chaotic behaviour of planetary orbits is ironic, since the solar system has long been the prototypical example of the philosophical view—for which Laplace was one of the early spokesmen—that the universe is deterministic and predictable.

Summary

A thorough understanding of the long-term evolution of planetary orbits would require integrating an ensemble of planetary systems for 5–10 Gyr. This task is still more than two orders of magnitude beyond our present capability for N-body integrations. Nevertheless, from shorter integrations and other tools, we are beginning to understand the principal features of the dynamical evolution of the planetary system over its lifetime.

This review was prepared during a research leave at the Institute of Astronomy, which was made possible by a Killam Research Fellowship and a grant from the Raymond and Beverly Sackler Foundation.

References

Applegate, J. H., Douglas, M. R., Gürsel, Y., Sussman, G. J., and Wisdom, J. 1986. *Astron. J.* **92**, 176.

Arnold, V. I. 1978. *Mathematical Methods of Celestial Mechanics* (New York: Springer).

Channell, P. J., and Scovel, C. 1990. *Nonlinearity* **3**, 231.

Cohen, C. J., and Hubbard, E. C. 1965. *Astron. J.* **70**, 10.

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

Duncan, M. J., and Quinn, T. 1993. *Ann. Rev. Astron. Astrophys.* **31**, 265.

Eckert, W. J., Brouwer, D., and Clemence, G. M. 1951. *Astron. Papers Amer. Ephemeris* **12**, 1.

Ferraz-Mello, S. (ed.) 1992. *Chaos, Resonances, and Collective Dynamical Phenomena in the Solar System* (Dordrecht: Kluwer).

Forest, E. 1987. 'Canonical integrators as tracking codes'. SSC Central Design Group Technical Report SSC-138.

Holman, M. J., and Wisdom, J. 1993. *Astron. J.* **105**, 1987.

Kinoshita, H., and Nakai, H. 1984. *Cel. Mech.* **34**, 203.

Kinoshita, H., Yoshida, H., and Nakai, H. 1991. *Cel. Mech.* **50**, 59.

Lambert, J. D., and Watson, I. A. 1976. *J. Inst. Maths Applics.* **18**, 189.

Laskar, J. 1985. *Astron. Astrophys.* **144**, 133.

Laskar, J. 1989. *Nature* **338**, 237.

Laskar, J. 1990. *Icarus* **88**, 266.

Laskar, J. 1994. *Astron. Astrophys.* **287**, L9.

Laskar, J., Quinn, T., and Tremaine, S. 1992. *Icarus* **95**, 148.

Levison, H. F., and Duncan, M. J. 1993. *Astrophys. J.* **406**, L35.

Newhall, X X, Standish, E. M., and Williams, J. G. 1983. *Astron. Astrophys.* **125**, 150.

Quinlan, G. D. 1992, in Ferraz-Mello (1992).

Quinlan, G. D., and Toomre, A. 1994. In preparation.

Quinlan, G. D., and Tremaine, S. 1990. *Astron. J.* **100**, 1694.

Quinn, T., and Tremaine, S. 1990. *Astron. J.* **99**, 1016.

Quinn, T. R., Tremaine, S., and Duncan, M. 1991. *Astron. J.* **101**, 2287.

Richardson, D. L., and Walker, C. F. 1989. *J. Astronaut. Sci.* **37**, 159.

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

Sackmann, I.-J., Boothroyd, A., and Kraemer, K. E. 1993. *Astrophys. J.* **418**, 457.

Saha, P., and Tremaine, S. 1992. *Astron. J.* **104**, 1633.

Saha, P., and Tremaine, S. 1994. *Astron. J.* **108**, 1962.

Sussman, G. J., and Wisdom, J. 1988. *Science* **241**, 433.

Sussman, G. J., and Wisdom, J. 1992. *Science* **257**, 56.

Touma, J., and Wisdom, J. 1994. *Astron. J.* **108**, 1943.

Will, C. M. 1981, *Theory and Experiment in Gravitational Physics*, (Cambridge: Cambridge University Press).

Wisdom, J., and Holman, M. 1991. *Astron. J.* **102**, 1528.

Wisdom, J., and Holman, M. 1992. *Astron. J.* **104**, 2022.

Wisdom, J., Holman, M., and Touma, J. 1994, in *Integration Algorithms for Classical Mechanics*, proceedings of a workshop at the Fields Institute.

Yoshida, H. 1990. *Phys. Lett. A* **150**, 262.
Yoshida, H. 1993. *Cel. Mech.* **56**, 27.