CHAOS IN THE N-BODY PROBLEM OF STELLAR DYNAMICS

Douglas C. Heggie

Department of Mathematics,
University of Edinburgh,
Edinburgh, U.K.

1. THE N-BODY PROBLEM OF STELLAR DYNAMICS

Stellar dynamics uses several different models of stellar systems, but in this paper we consider the most fundamental, which is governed by the N-body equations:

$$\ddot{\mathbf{r}}_i = -G \sum_{j=1, j\neq i}^{N} m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3}. \tag{1}$$

Stellar dynamics shares these equations with celestial mechanics (a term which is used here to denote the study of the orbital dynamics of bodies in the solar system), but there are important differences of emphasis. In stellar dynamics all masses are comparable, whereas in celestial mechanics one mass tends to dominate (either the sun or a primary). This has an effect on the methods used and the types of motion which result. Approximate analytical methods are of immense value in celestial mechanics, but not in stellar dynamics, where numerical methods predominate. In celestial mechanics motions tend to be very nearly regular for long intervals of time, whereas in stellar systems motions are highly irregular.

This last point is one that is illustrated by graphical results from numerical integrations (Carnevali & Santangelo 1980), but in the present paper we analyse the irregular nature of motions in stellar systems from a more quantitative point of view; i.e. how sensitively the orbits of the stars depend on the initial conditions. We summarise both numerical data and theoretical investigations on this problem. In addition we consider the implications of these results for the validity and interpretation of numerical simulations.

2. SOME IDEAS OF STELLAR DYNAMICS

2.1 *Some Terminology*

For future reference we list here some useful standard notions of stellar dynamics (cf. Binney & Tremaine 1987), though the applications we have in mind are to star clusters. The kinetic and potential energies of N point masses are, as usual, $T = \frac{1}{2}\sum_1^N m_i \mathbf{v}_i^2$ and $W = -G\sum_{i<j}(m_i m_j/|\mathbf{r}_i - \mathbf{r}_j|)$, respectively. We define the *virial ratio* $q = 2T/|W|$. This is found to be of order unity in stellar systems in dynamic equilibrium, by the virial theorem (which is the extension of Lagrange's identity in the three-body problem to N bodies.) It is convenient to consider a time scale set by the *crossing time* $t_{cr} = 2R/v$, where R is the size of the system (in a sense which may be made precise), and v is the mean stellar speed. Using estimates based on the magnitude of q it is then easy to deduce that $t_{cr} \sim \sqrt{(R^3/(GM))}$, where M is the total mass of the

system.

2.2 A Summary of the Evolution of a Star Cluster

Later we shall present some numerical data whose interpretation depends partly on what is known about the way in which stellar systems evolve, and so here we summarise some ideas about this (cf. Spitzer 1987). If the system is not initially in dynamic equilibrium (e.g. it is collapsing) then it first settles down into dynamic equilibrium, on a time scale of order t_{cr}. The equilibrium is not exact, however, and the system continues to evolve, though much more slowly. Two-body interactions cause stars to be ejected out of the densest part of the cluster, which is called the *core*. The core shrinks (in size and mass); a process referred to as *core collapse*. It takes place on a *relaxation* time scale, which is of order $Nt_{cr}/\ln N$. When the density in the core has become high enough, three-body interactions become sufficiently common to eject substantial numbers of stars from the cluster. This loss of mass eventually reverses the collapse of the core, and causes overall expansion of the entire system.

The eventual outcome of this evolution is somewhat conjectural, but is thought to be an asymptotic state in which escaping single stars, and some escaping binaries, are spreading out to infinity.

3. CHECKING NUMERICAL ERRORS

As already mentioned, many results on the dynamical evolution of star clusters result from numerical computations, using eqs.(1). As always, it is desirable to control the errors in such computations, but this is clearly a difficult task here, as there exist no suitable non-trivial exact solutions against which to test the results. The following partial checks are available.

(i) *Conservation of integrals* In practice the total energy $E = T + W$ is the most useful. Usually it is considered sufficient if the relative numerical error is limited by $|\Delta E/E| < 10^{-4}$ per t_{cr}, or better, (Aarseth 1974). Incidentally, this might seem crude by the standards of celestial mechanics, but several important shortcuts are needed in order to obtain useful results for systems of any size, and accuracy is one of the sacrifices that must be made.

(ii) *Time-reversal* This test is rarely carried out on large systems. An example is the celebrated Burrau problem in the general three-body problem, where Szebehely & Peters (1967) reversed the velocities at the point where the triple system embarked on its final asymptotic motion (a single mass escaping hyperbolically from a binary), and were able to recover the initial coordinates to about three significant figures. This example also shows how unreliable it is to use conservation of energy as a test: even though there were such large errors in the coordinates the total energy was conserved to a relative accuracy better than 10^{-11}!

(iii) *Independent calculations* Numerical stellar dynamics is sometimes referred to as an experimental method, and, like all experimental techniques, should produce reproducible results. But it was shown by Lecar (1968) that different computations with the same initial conditions (using different computers, or even simply different algorithms on the same computer) produced widely divergent results, after the first crossing time or so.

This trial shows that the detailed positions of the stars in an N-body computation cannot, in general, be regarded as being even approximately correct. On the other hand the interpretation of

the results is a little complicated, because we see not only the effects of truncation and rounding errors, but also the propagation of errors made at earlier stages of each calculation. One method of isolating the last mechanism (propagation of errors) is to study the evolution of N-body systems with (slightly) different initial conditions. We now discuss how this can be done in practice, and what the results are.

4. THE GROWTH OF NUMERICAL ERRORS

Consider two N-body systems $I,\ II$ satisfying the same N-body equations but slightly different initial conditions, i.e.

$$\mathbf{r}_i^I(0) = \mathbf{r}_{i0}^I, \quad \dot{\mathbf{r}}_i^I(0) = \dot{\mathbf{r}}_{i0}^I,$$

and

$$\mathbf{r}_i^{II}(0) = \mathbf{r}_{i0}^{II}, \quad \dot{\mathbf{r}}_i^{II}(0) = \dot{\mathbf{r}}_{i0}^{II}.$$

Suppose we now integrate both systems as accurately as possible. How, we ask, does the difference between the solutions evolve?

The "difference" may be measured in several different ways. One possibility is to define

$$\Delta(t) = (\sum(\mathbf{r}_i^I - \mathbf{r}_i^{II})^2 + \sum(\dot{\mathbf{r}}_i^I - \dot{\mathbf{r}}_i^{II})^2)^{1/2}.$$

This choice was made in a classic investigation by Miller (1964), who found, for $4 \leq N \leq 32$, that Δ grows roughly as $\exp(\mu t)$, where μ is constant. There was much scatter, but typical values were given by $\mu t_{cr} \sim 2, 4, 20$ for $N = 8, 12, 32$, respectively. Very comparable results were obtained later by Standish (1968) and by Dejonghe & Hut (1986) for $N = 25$ and $N = 3$, respectively.

These studies immediately explain Lecar's results for $N = 25$; the growth of errors is very rapid on the time scales of interest. And Dejonghe & Hut found that initial errors in Burrau's Problem grow by a factor of order 10^9 up to the time at which Szebehely & Peters carried out time reversal; since errors also grow during the reverse integration by a comparable factor (which is a property of Hamiltonian systems), it is clear that Szebehely & Peters did well to recover the initial conditions even to three significant figures.

It also follows that it is impossible to predict the positions and velocities of the stars in a simulated star cluster for more than a few crossing times. But Miller's results are also puzzling. They suggest that the logarithmic rate at which the errors grow depends on N, and indeed increases roughly linearly with N. If his results could be extrapolated to a rich globular cluster, where $N \sim 10^6$ and $t_{cr} \sim 10^6$yr, they suggest that the time scale for growth of errors is of order 1yr. In this time a star in a globular cluster moves a distance of order 1AU, whereas the mean distance between the stars is of order 10^4AU. What could cause such rapid growth of errors?

For these reasons the author, in collaboration with Goodman & Hut, has carried out further studies, in order to determine better the N-dependence of the growth of errors. In these new studies a standard alternative formulation, using the variational equations, has been employed. Let us abbreviate eqs.(1) as

$$\ddot{\mathbf{r}}_i = \mathbf{a}_i(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N),$$

and let $\Delta\mathbf{r}_i = \mathbf{r}_i^{II} - \mathbf{r}_i^I$ be the difference between two neighbouring solutions. While this difference remains small, it satisfies approximately the linearised equations

$$\Delta\ddot{\mathbf{r}}_i = \sum_{j=1}^{N} \Delta\mathbf{r}_j . \nabla_{\mathbf{r}_j} \mathbf{a}_i, \tag{2}$$

and these equations are sufficient to determine the growing separation of the two solutions within the linear regime.

What follows is a summary of the numerical results reported in Goodman, Heggie & Hut (1990). The systems ranged in size from $N = 4$ to $N = 512$, in steps of a factor of 2. The components of the initial positions and velocities of the stars were chosen from a random distribution on (0,1), then rescaled to make the virial ratio $q = 1$. The components of $\Delta\mathbf{r}_i$ and $\Delta\dot{\mathbf{r}}_i$ were assigned initial values equal to ± 1 at random. At each value of N ten cases were integrated (differing only in the initial conditions, selected as stated), except that only five cases were studied for N = 512. The integrations were continued up to $t = (5/\sqrt{2})t_{cr}$. To measure the growth of the variations $\Delta\mathbf{r}_i$ a different choice from Miller's was made. It turns out that the variations in the velocities $\Delta\dot{\mathbf{r}}_i$ tend to have large 'spikes' during close approaches, whereas $\Delta\mathbf{r}_i$ is better behaved. To smooth the data further, a geometric mean over all the stars was taken, i.e. we define $\ln\Delta = (\sum \ln|\Delta\mathbf{r}_i|)/N$.

Since the purpose of this investigation was to investigate the growth of errors in N-body integrations, it was desirable to use an integration algorithm allowing automatic step-size control based on the local truncation error. The fourth-order Runge-Kutta routines given in Press *et al.* (1986) proved suitable in this respect.

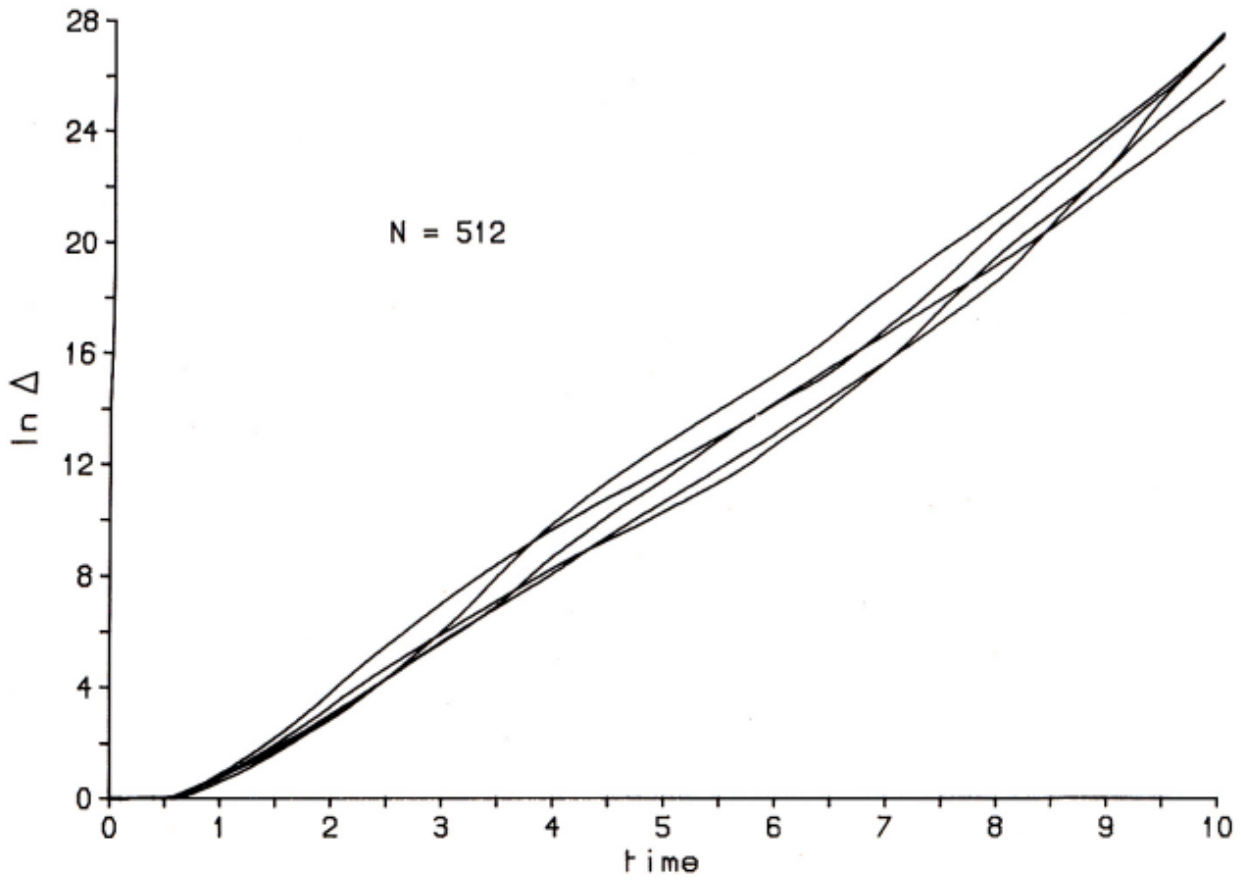

Fig.1 Growth of solutions of the variational equations, for 5 cases with $N = 512$. The ordinate is a logarithmic measure of the size of the solution, as defined in the text, and the abscissa is time.

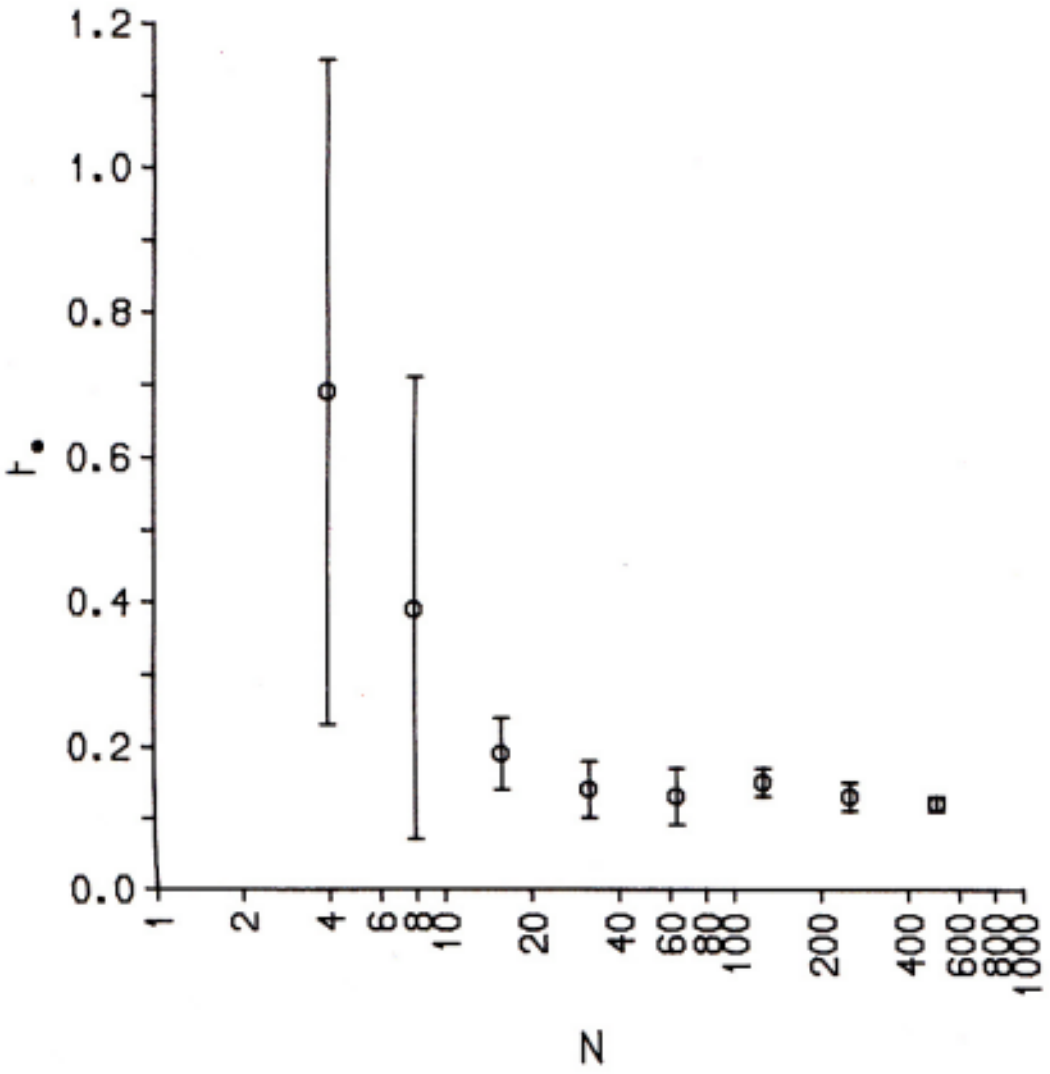

Fig.2 The time scale for the exponential divergence of neighbouring solutions of the N-body problem, as a function of N. The error bars give the standard deviation estimated from the 10 computations carried out for each value of N (except N = 512, for which there are only 5). The ordinate is expressed in units of the crossing time.

The results of these calculations confirmed that the growth of Δ is nearly exponential, especially for large N (Fig.1). Indeed the scatter in the results (across runs made with different random initial conditions) decreases as N increases. The N-dependence of the growth rate was one of the most interesting conclusions (Fig.2). Defining the growth rate as $\mu = d(\ln \Delta)/dt$, it was found that μ decreases as N increases in the range $4 \lesssim N \lesssim 32$, but then levels off up to the largest value we studied. Thus Miller's results, which stopped at $N = 32$, gave a similar N-dependence to what we have found, but clearly cannot be extrapolated to larger N. For large values of N our results are consistent with $\mu \simeq 6/t_{cr}$, with no dependence on N. Since our results are based on a geometric mean over all the stars in each system, it might be thought that the variations $\Delta \mathbf{r}_i$ might be large only for a small number of stars, but Fig.3 shows that this is not so.

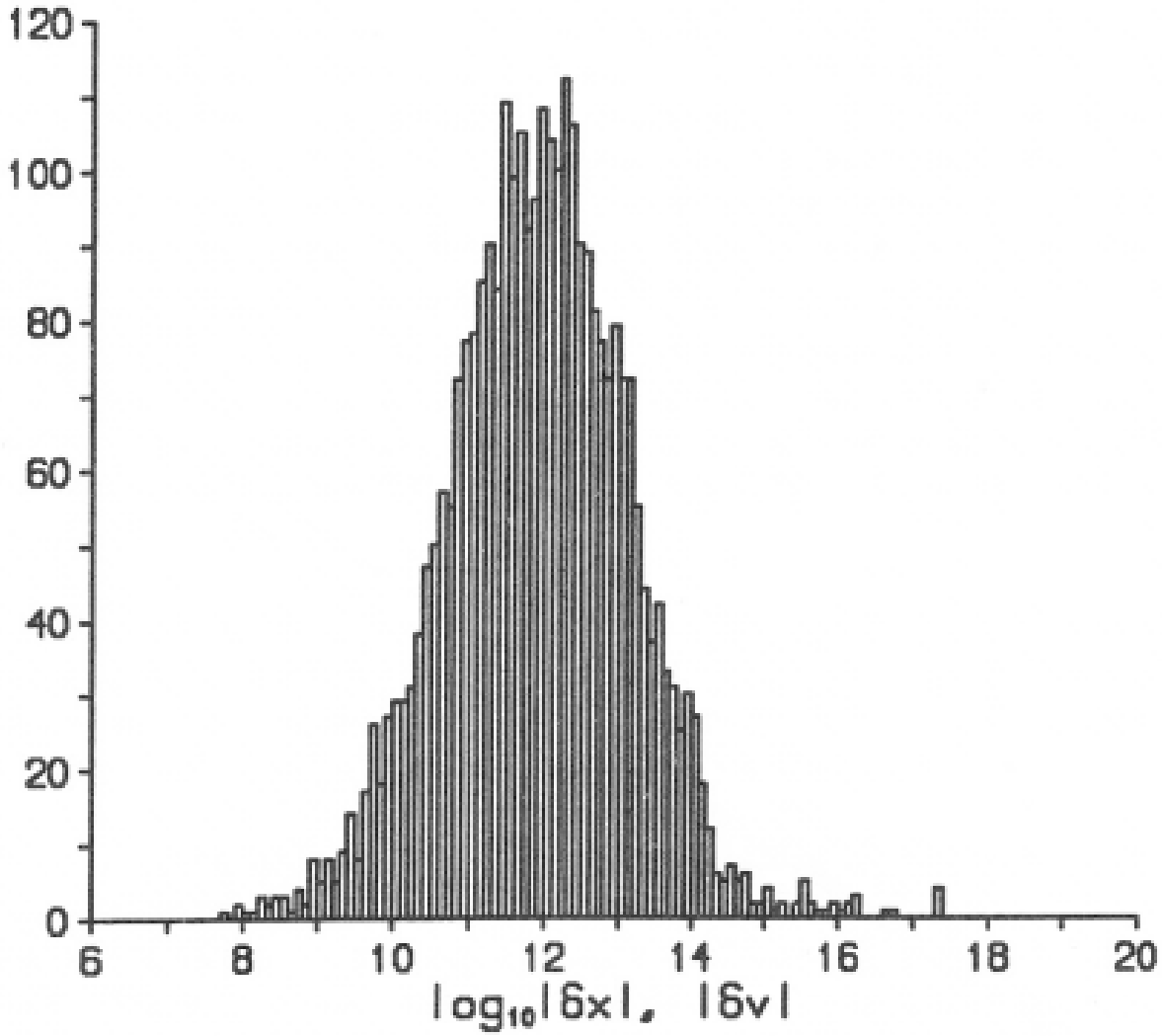

Fig.3 Histogram of values of δx_i and δv_i, i.e. components of the variations in position and velocity, for 512 stars at $t = 10$. Thus 3072 data values are included. The initial values are of order 1.

These new results allow us to sharpen our conclusions about how quickly errors grow in large N-body computations. For example, if such a computation is carried out in double precision then all accuracy is lost at a time such that $\exp(6t/t_{cr}) \sim 10^{18}, \quad \Rightarrow t \simeq 7t_{cr}$. By comparison, the time to the end of core collapse is of order $30t_{cr}$ for $N = 100$, and is approximately proportional to N.

5. LIAPOUNOV EXPONENTS

Before we pass on from these numerical studies to a theoretical analysis of error growth, it is worth attempting to clarify the link between this investigation and the standard language in which such studies are often expressed. Consider a finite dynamical system with state vector $\mathbf{q}$, and let $\Delta\mathbf{q}(t)$ be a solution of the corresponding variational equations. Define

$$\mu = \lim_{t\to\infty} \frac{\ln \|\Delta\mathbf{q}\|}{t}. \tag{3}$$

It can be shown that this limit exists if the system is tolerably smooth, and (for a given initial value $\mathbf{q}(0)$) takes one of a discrete set of values (depending on $\Delta\mathbf{q}(0)$), called *Liapounov characteristic exponents* (see, for example, Lichtenberg & Lieberman 1983). For almost all $\Delta\mathbf{q}(0)$ the value of μ is the *largest* exponent.

This is clearly related to the quantity μ discussed previously, but the present definition is of no value for our purposes, because of the requirement that $t \to \infty$. If the discussion of §2.2 (above) is correct, as $t \to \infty$ the stellar N-body problem is asymptotically integrable, and so $\mu = 0$. (In an integrable system neighbouring orbits deviate linearly with t, not exponentially.) Therefore the largest Liapounov Characteristic Exponent says nothing about loss of predictability, sensitive dependence on initial conditions, etc.

What is needed is a similar concept which measures the divergence of neighbouring orbits over finite intervals of time which are of interest. If we modify eq.(3) and define μ by

$$\mu^{-1} = \frac{t}{\ln(\|\Delta(t)\|/\|\Delta(0)\|)},$$

then we have a definition of a time-dependent quantity, μ, which is sometimes conveniently referred to as a "Liapounov Characteristic Indicator". On the other hand we can emphasise its physical significance if we refer to μ^{-1} as the "time scale of instability", or "e-folding time".

6. THEORY OF THE EXPONENTIAL INSTABILITY

6.1 *Solutions of the N-Body Problem as Geodesics*

We now leave behind the numerical investigation of our problem, and turn to the theoretical approach. There are two lines of attack, one geometrical in nature, the other physical. Both date back to work of Krylov which was published in 1950 (Krylov 1979). In this section we consider the first of these approaches.

We consider a Lagrangian system, like the N-body problem, in which the kinetic energy T is a quadratic form in the generalized velocities $\dot{q}_i$, with coefficients depending on the q_i. Then there is a standard piece of theory (see, for example, Arnold 1978, p.247) which shows that all motions of a fixed energy h give orbits in configuration space which are geodesics for a suitably defined metric. In fact the distance between neighbouring points is given by $ds^2 = 2T(h-U)dt^2$, where U is the potential function. (By our assumption about the form of T, it can be seen that this is a quadratic form in the differentials dq_i, with coefficients which are functions of the generalized coordinates q_i.)

The advantage of casting a dynamical system in this form is that much is known about geodesic flows; and, while the conditions under which such results may be proved may not always apply to the N-body problem, they are strongly suggestive of behaviour to look for. Our interest is in the rate at which orbits diverge, and the rate of divergence of geodesics is described by Jacobi's equation (cf. Arnold 1978, Appendix 1). This is simply the variational equation for geodesic flow, but it is expressed in terms of *covariant* derivatives. The importance of these is that they have a significance independent of the particular coordinates in use; if a covariant derivative vanishes in one system then it vanishes in all.

Let us denote by v the unit tangent vector to a geodesic, and by ξ a vector which takes us from a point on this geodesic to a corresponding point on a neighbouring geodesic. We are interested in the rate of growth of $|\xi|$, and it is easily shown from Jacobi's equation (Gurzadyan & Savvidy 1986) that

$$\frac{d^2|\xi|^2}{ds^2} \geq -2K(\xi, v)|\xi|^2, \tag{4}$$

where K, the Riemannian curvature, can be easily computed in terms of the metric coefficients. If K remained negative we would be able to show that $|\xi|$ eventually grows exponentially with time. Unfortunately, it turns out that K can have either sign in the classical gravitational N-body problem.

The next simplification that can be attempted is to average the right-hand side of eq.(4) over all possible directions of the vectors v and ξ. When this is done it is found that the average value of the Riemann curvature is

$$\langle K \rangle = \frac{2-N}{4N} \frac{\sum_{i=1}^{N} m_i f_i^2}{T^3}, \tag{5}$$

where N is the number of stars, f_i is the acceleration of the ith star, and m_i is its mass. Interestingly, this vanishes when $N = 2$, and it is sometimes said that this corresponds to the integrability of the 2-body problem. Our interest is in larger N, and since then we have $\langle K \rangle < 0$, it is possible to argue (Gurzadyan & Savvidy 1986) that this implies exponential divergence of neighbouring orbits.

Furthermore, a theoretical estimate can be made of the rate of growth, which Gurzadyan & Savvidy attempted to do by estimating the average value of the right-hand side of eq.(5). The average is taken over a sensibly chosen, random distribution of the positions of the stars. Unfortunately, however, the average value of f_i^2 diverges because of the contribution of stars at small distances (on the assumption that the positions of neighbouring stars are uncorrelated). Because of the nature of this difficulty, these authors estimated the average value $\langle f_i^2 \rangle$ essentially by substituting the value of f_i^2 which would be contributed by a typical nearest neighbour. (In a system of radius R, the nearest neighbour is at a distance typically of order $RN^{-1/3}$.) In this way Gurzadyan & Savvidy estimated that the time scale for the divergence of neighbouring orbits in a stellar system is of order $N^{1/3} t_{cr}$.

As far as the author is aware, this was the first attempt to estimate the time scale for the divergence of orbits in the classical gravitational N-body problem. On the other hand the predicted N-dependence is inconsistent with the numerical results summarised in §4, even those published before the time of this theoretical study. It is not hard to see that the problem might lie with the divergence at small distances. While the typical interparticle distance is indeed of order $RN^{-1/3}$, it is not hard to show that, in each crossing time, each star has a close encounter at a distance of order $RN^{-1/2}$, where the rate of divergence of neighbouring orbits is larger than average. Although such encounters are short-lived it is not immediately clear whether a strong, brief effect will be dominated by a weaker, constant effect.

6.2 *The Growth of Errors in "Collisions"*

Perhaps it would be possible to adapt the geometric theory, just described, so as to estimate the effect of occasional close encounters. But there is another type of theory, also apparently due to Krylov, in which this may be attempted. He developed it in the context of the "hard-sphere gas", where the bodies have no effect on each other, except during collisions.

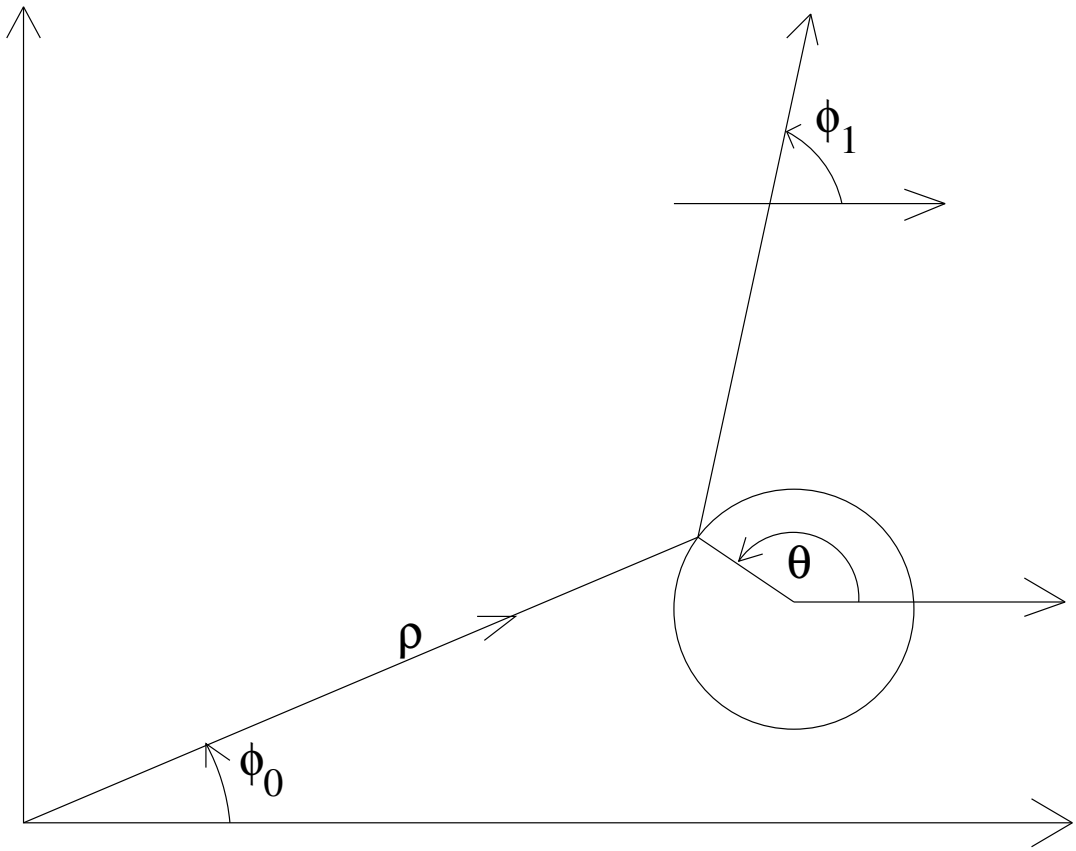

Fig.4 Path of a particle which strikes, and is deflected by, a hard sphere. The notation is explained in the text.

We shall outline the theory for a two-dimensional hard sphere gas first. Suppose a particle sets off from the origin, in a direction making an angle ϕ_0 with the x-axis (Fig.4). It strikes a sphere at a point on the sphere corresponding to the polar angle θ, and then sets off in the direction ϕ_1. Since the angles of incidence and reflection are $\pi - \theta + \phi_0$ and $\theta - \phi_1$, respectively, equality of these angles shows that

$$\phi_1 = \phi_0 + \pi - 2\theta. \tag{6}$$

Now suppose the particle had set off at a slightly different angle $\phi_0 + d\phi_0$. If the surface of the sphere is at a distance ρ from the origin, then the particle now has a transverse spatial displacement

$$dp_1 = \rho d\phi_0, \tag{7}$$

(where we use p to denote the minimum distance at which the particle would pass the centre of the sphere if it were not deflected.) Therefore the particle strikes the surface of the sphere at a point which is displaced by an amount $dp_1/\cos(\pi - \theta + \phi_0)$ clockwise round the surface of the sphere. Equating this to $-r_0 d\theta$, where r_0 is the radius of the sphere, the differential of eq.(6) leads to the result $d\phi_1 = d\phi_0(1 + 2\rho/[r_0 \cos(\pi - \theta + \phi_0)])$. The important point about this result is that errors in the direction of motion are magnified by a factor of order ρ/r_0, and we may imagine this can be large in a dilute gas.

Now we turn to the consideration of the gravitational N-body problem. Here simple estimates (cf. Heggie 1988) show that $\phi_1 \sim \phi_0 - Gm/(pv^2)$, where v is the speed of the particle. Hence

$$d\phi_1 \sim d\phi_0 + Gm(pv)^{-2} dp_1, \;\; \simeq d\phi_0(1 + Gm\rho/[p^2 v^2]). \tag{8}$$

A complication, to which we shall return, is that there is no length-scale in the gravitational problem which corresponds to the radius r_0 of a hard sphere. Instead, there are encounters at a

wide range of distances p, and the distance travelled (ρ) depends on how close the encounters are which we are considering. What can be done (Heggie 1988) is to estimate the effect of those encounters which have the greatest magnifying effect on the errors. It turns out that the most effective encounters are those at a distance of order $RN^{-1/2}$, where, as before, R is a measure of the size of the system. Then the time scale for growth of errors is of order $1t_{cr}$. This is in agreement with the numerical results of Goodman *et al.* which were discussed in §4 above (and were obtained *after* the theoretical estimate!)

We now discuss some refinements of this theory which have been developed recently. An important consideration was advanced by Goodman and is described in more detail in Goodman *et al.* (1990). He pointed out that still closer encounters (at distances much less than $RN^{-1/2}$) affect only a few stars in each crossing time, but have a very large effect on the errors in the positions and velocities of these stars. These strongly affected stars can then "infect" the other stars in the system, because the errors in *all* the stars are coupled together, by the form of the terms on the right-hand side of the variational equations, eq.(2). (The physical reason for this coupling can also be illustrated by Fig.4. In the associated discussion we assumed that the position of the sphere was not subject to error. If, however, the error in its position is large, this will be what mainly determines the error in the direction of the particle after the collision.) It takes several crossing times for this "infection" to spread throughout the entire system, but the net effect is to shorten the time scale for exponential divergence by a factor of order $\ln N$.

It is worth pointing out that there is no numerical evidence for this factor. This may be because such logarithmic variations are rather small in practice, and may be too difficult to detect in the face of the considerable statistical uncertainty in our numerical estimates. On the other hand Goodman's theory makes it clear that the time needed for the "spread of the infection" (its incubation period?) increases with N, whereas our numerical experiments were concluded at a time which was chosen independent of N.

6.3 *The Growth of Errors as a Stochastic Process*

The theoretical estimates we have discussed simply yield orders of magnitude, and teach us how the rate of growth of errors depends on the number of stars, etc. It is much harder to use a result like eq.(8) to derive a numerical estimate of the time scale. The reason for this is that the magnification factor depends on quantities like ρ and p, which vary randomly from one encounter to the next. One solution is to simulate the process numerically, choosing values of these parameters from suitably chosen random distributions, and then multiplying together the effects of successive encounters. This has been done by Goodman (Goodman *et al.* 1990). (Note that this is a different kind of numerical simulation from those discussed in §4. Those were based on the full N-body equations, eq.(1). Here we are using an equation like eq.(8), which is already a great simplification from eq.(1).) In what follows we describe a framework for an *analytical* attack on the same problem. It is mainly due to my colleague A.M. Davie.

First we have to correct eq.(7), which assumes that the particle emerges from the origin with an error only in its initial direction. But if it emerges from its previous encounter, it already has a transverse error in its position, which we may denote by dp_0. Thus eq.(7) is to be replaced by

$$dp_1 = dp_0 + \rho d\phi_0,$$

and eq.(8) by the more fundamental form

$$d\phi_1 = d\phi_0 + Gm(pv)^{-2}dp_1.$$

(A more careful discussion would also be needed to determine the numerical value of the coefficient of the last term.)

These two equations define a map which is symplectic (since its Jacobian is unity), and stochastic (since the quantities $\rho,\ p$ and v will differ from one encounter to another, and may be assumed to be random variables drawn from appropriate distributions.) Introducing a vector $\mathbf{x} = (dp, d\phi)^T$, we may write the iterated map as $\mathbf{x}_n = A_n \mathbf{x}_0$, where the matrices A_n are related by

$$A_{n+1} = \begin{pmatrix} 1 & \rho \\ \dfrac{Gm}{p^2 v^2} & 1 + \rho \dfrac{Gm}{p^2 v^2} \end{pmatrix} A_n.$$

(Note that the quantities $\rho,\ p$ and v should be subscripted to show that they correspond to the nth encounter. These subscripts have been omitted for clarity of notation.) Expressing A_n explicitly by

$$A_n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix},$$

we deduce easily that

$$\begin{aligned} a_{n+1} &= a_n + \rho c_n, (9) \\ &= a_0 \Pi_{i=1}^{n} (1 + \rho_i c_i / a_i). \end{aligned}$$

It is easy to see now that the average logarithmic divergence (per encounter) is given by the average value of $\ln(1 + \rho_i c_i / a_i)$, but it remains to determine how to take the average. In other words, what is the distribution of c_i/a_i? Let us define $z_i = c_i/a_i$. Then, by the kind of argument which led to eq.(9), we immediately obtain the relation $z_{n+1} = Gm(pv)^{-2} + z_n/(1 + \rho z_n)$. This defines a sequence of random variables, since the values of $\rho,\ p$ and v are to be chosen from appropriate distributions. The probability density functions of this sequence can be obtained recursively from

$$\begin{aligned} f_{n+1}(z_{n+1}) &= \int f_n(z_n) f_{\rho,p,v}(\rho, p, v) \delta(z_{n+1} - \frac{Gm}{p^2 v^2} - \frac{z_n}{1 + \rho z_n}) d\rho dp dv dz_n, \\ &= \int f_n(z_n) K(z_n, z_{n+1}) dz_n, \end{aligned}$$

say. If we suppose that the sequence of density functions tends to a limit, i.e. $f_n \to f$, then f must be a solution of the integral equation

$$f(z') = \int f(z) K(z, z') dz. \tag{10}$$

Finally, the rate of exponential divergence of neighbouring orbits can now be computed as $\mu = \langle \ln(1 + \rho z) \rangle / \langle \rho / V \rangle$, where V is the speed of the star. In the computation of the first average, the distribution of z to be used is obtained (in principle) from the solution of the integral equation (10).

7. RELAXATION AND THE GROWTH OF ERRORS

In large stellar systems in dynamic equilibrium, the stars move around on relatively smooth orbits, only occasionally coming close enough to another star to be significantly deflected. Most of the time, they can be thought of as moving in a smooth, almost stationary gravitational field, in which the individual energy of each star is nearly constant. Because of the occasional close approach (and the cumulative effect of encounters which are not so close) its energy varies slightly, but it takes a time of order $N t_{cr} / \ln N$ (the Chandrasekhar relaxation time) to change by a significant amount (cf. Binney & Tremaine 1987).

Any process which changes the energy of individual stars is called a "relaxation" mechanism, and the main reason why Gurzadyan & Savvidy were interested in estimating the rate of divergence of orbits is that they assumed they were also calculating a relaxation rate. Furthermore they found, as we have seen, a time scale for the divergence of orbits which, for large systems at least, must be shorter than Chandrasekhar's time scale. (We saw in §6.1 that they estimated a time scale of order $N^{1/3}t_{cr}$ for the rate of divergence. In §6.2 this has been revised to a time scale of order t_{cr} or even $t_{cr}/\ln N$. This is smaller than Chandrasekhar's relaxation time scale by a factor of N.) Therefore they concluded that they had found a mechanism of relaxation which would cause stellar systems to evolve much more quickly than had been thought hitherto.

In hard-sphere gases it is true that the rate of divergence of orbits is comparable with the rate of relaxation: it takes only a few collisions to randomise the direction of motion of an atom, or to bring its kinetic energy in statistical equilibrium with the energies of the atoms around it. Stellar systems are different, as we now explain.

We have seen that the growth of errors in positions and velocities of stars in stellar systems is exponential, with an e-folding time scale t_e of order $t_{cr}/\ln N$. It is certainly true that errors in energy grow similarly, and on the same time scale. But this *exponential* growth of errors is obtained in a linear approximation (i.e. variational equations, or Jacobi's equation), and ceases to be valid when the linear approximation itself breaks down. It is easy to see that this will certainly happen when the error in the position of a star becomes comparable with the distance to its nearest neighbour.

In order to quantify this statement, let us denote by p and dp the distance of closest approach of two stars, and the error in this distance, respectively. We have seen (in §6.2) that the most effective encounters are those in which $p \sim RN^{-1/2}$, while exponential growth of errors on a time scale of order t_{cr} leads to $dp(t) \sim dp(0)\exp(t/t_{cr})$. Thus the linear approximation breaks down at a time when $dp(t) \sim RN^{-1/2}$, i.e. when $t \sim t_{cr}\ln(R/N^{1/2}dp(0))$. Thereafter the growth of errors in energy slows down, and there is no contradiction with the slow changes in energy predicted by Chandrasekhar's theory.

Actually the predictions of the two theories are quite consistent, in the following sense. The gravitational potential gradient (per unit mass) in a system of radius R, consisting of N stars of mass m each, is of order GNm/R. If the error in position of a star is of order $RN^{-1/2}$, it follows that the resulting error in its potential energy is of order $GmN^{1/2}/R$. Now Chandrasekhar's theory is based on a model in which the energy of a star varies in the fashion of a random walk, and so the change in energy varies with the square root of the time. Knowing the relaxation time given by Chandrasekhar's theory, we can easily estimate the predicted change in energy at the time (of order t_{cr}) when the exponential growth of errors slows down. It turns out to be the same (except for numerical factors) as our estimate of the error in energy at the same time.

There is one respect in which the estimate of the growth of errors in stellar systems teaches us something about galactic dynamics. It is usual to study galactic orbits by assuming, as mentioned above, that the stars move in a smooth potential, and that the graininess of the true potential does not matter over the time scales of interest. This leads to much interesting theory on whether the potentials are integrable, or whether motions are stochastic. In the latter case, it is possible to measure Liapounov Characteristic Exponents, again assuming that the potential is smooth. We now see, however, that the true rate at which neighbouring orbits diverge is of the order of a crossing time, and that encounters with individual stars are fundamental. Thus the approximation of a smooth potential is useful for studying orbits, but not for studying their divergence.

We now return to the issue which motivated this entire study: how reliable are the results of N-body simulations? We now know that the positions and velocities of individual stars are quite unreliable after a few crossing times, which is a period much shorter than that over which we would like to study the dynamics of star clusters. What sense, then, can be made of the results of simulations, which for many purposes seem to be the only way in which some problems of stellar dynamics can be approached? The usual answer given to this question is to assert that *statistical* results are still valid, even though the individual positions and velocities of stars are not known. For example, it is assumed that the average rate at which stars escape can be determined (with only statistical error), even though it is not usually possible to state which stars escape.

Despite the enormous numbers of N-body simulations which have been carried out, it is remarkable that very little has been done to test this assertion. It can also be looked at from another point of view: if all we can be interested in are *statistical* results, why should we bother to compute accurate orbits? How crudely (and cheaply) can we simulate stellar systems and still obtain statistically reliable results? There seems to be no theoretical understanding of these questions, and very little empirical evidence either.

One case which has been studied (Valtonen 1974) is the scattering problem for $N = 3$, i.e. the statistical study of the outcome of encounters between a single star and a binary. Valtonen conducted several series of experiments at different accuracy, and measured the statistical distribution of the final eccentricity of the binary. He found no dependence on the accuracy, judged by conservation of energy E, though the range of accuracies was rather limited:
$0.005 < \langle (\Delta E/E)^2 \rangle^{1/2} < 0.03$.

Larger N-body systems ($N = 16$) have been studied by Smith (1977). The time step for different runs varied over a factor of 7, but he found no discernible differences in his statistical results except when the conservation of energy was grossly violated, i.e. unless $|\Delta E/E|$ exceeded 10 – 100. His conclusion was that it is best to make large numbers of cheap runs, as long as the total energy is roughly conserved in each run.

Against this background, we now present some results of a new study involving somewhat larger systems ($N = 100$). Our aim will be able to test for any dependence of a number of statistical results on integration accuracy. The statistical data we test are chosen to correspond to a number of issues which are of special interest to stellar dynamicists (cf. §2.2).

The particular integration program used was the widely available program known as NBODY1 (Aarseth 1985, Binney & Tremaine 1987). In this program the local time-step Δt depends on a parameter η as $\sqrt{\eta}$, and in this study the chosen values of η were 0.1, 0.03 and 0.01. The value 0.03 is usually recommended, except when close encounters occur. Note that the integration method is of relatively high order (by the standards of stellar dynamics), the local truncation error in position varying as $(\Delta t)^7$. For the initial conditions a so-called Plummer model was selected (cf. Spitzer 1987). The results are scaled to units such that $G = 1, Nm = 1$ and $E = -1/4$; such values are rather standard in the field. For each value of η, 10 runs were carried out, differing in the random numbers used to generate the initial conditions.

It was found, as expected, that the error in the energy increased greatly around the time when the first close binary formed. Such an occurrence is to be expected on the basis of what is known of cluster evolution (cf. §2.2), and there are several refinements in more advanced N-body programs which greatly alleviate this difficulty. They are not present in the basic FORTRAN program NBODY1, and so the simulations were stopped at this point. To be precise, let us define $kT = \langle mv^2 \rangle/3$, i.e. as two-thirds the mean kinetic energy of a single star, in analogy with the

Table 1. Formation of binaries

η	median $\lvert\Delta E/E\rvert$	min(t_{10})	median(t_{10})	max(t_{10})
0.01	0.0003	17	33	55
0.03	0.004	18	33	53
0.1	0.2	10	30	83
0.1*	0.14	10	30	34

* 7 cases with $|\Delta E/E| < 1$

Table 2. Half-mass radius at t_{10}

η	smallest	median	largest
0.01	0.72	0.83	1.18
0.03	0.66	0.92	1.19
0.1	0.80	0.91	1.15

appropriate definition of temperature in kinetic theory. Then we define an energetic binary as one whose energy exceeds $10kT$, and define t_{10} as the time at which the first such binary formed. Table 1 gives some statistical results on this quantity.

Recalling that 10 cases were studied for each value of η, we see that there is very little evidence for a dependence on accuracy. Possibly the distribution of t_{10} is a too wide at low accuracy, but the evidence is not compelling. Incidentally, the theory of core collapse (based on model equations valid for large N, i.e. the Fokker-Planck equation, cf. Spitzer 1987) predicts that this would be complete by about $t \simeq 32$ for $N = 100$. Formation of binaries is thought to be a signature of core collapse (§2.2).

Table 2 presents comparable data on the *half-mass radius* at the same time t_{10}. This is defined as the radius of an imaginary sphere, centred at the densest part of the system, which encloses half of its mass, and its evolution is often used as a simple measure of the evolution of the bulk of the cluster. Fokker-Planck theory predicts that this should increase from 0.77 at $t = 0$ to about 1.10 at the end of core collapse. The values found in these simulations are consistently smaller than this, which perhaps exposes the limitations of the Fokker-Planck model. The important point here is that the values are very consistent among themselves, and show no signs of any dependence on accuracy.

Statistics on the number of stars which escape, and the total energy which they carry off, are shown on Table 3. Here at last is a result which clearly depends very sensitively on integration errors. Perhaps this is not surprising, because escapes tend to follow from relatively close

Table 3. Number of escapers (to time t_{10})

η	Total number	Total energy
0.01	13	0.0088
0.03	8	0.010
0.1	23	12.4
0.1*	14	0.25

* 7 cases with $|\Delta E/E| < 1$; data scaled to 10 cases.

encounters, and these are a major source of error. The last line of the table, in which the very worst of the low-accuracy runs are excluded, shows that the total energy of the escaping stars is even more sensitive to errors than their total number. Incidentally, the problem of escapers is an area where simulations are crucial, as the escape rate is notoriously difficult to predict theoretically (Wielen 1968).

The results also show that the savings achieved by sacrificing accuracy (i.e. setting $\eta = 0.1$ rather than 0.03) are less than a factor of 3 in the median number of integration steps required. Therefore there would be little to gain in carefully tuning η within this range to minimise the number of steps while preserving the statistical reliability of the results. Indeed the conclusion of this study is rather clear: use the recommended value (i.e. $\eta = 0.03$)!

It must be admitted that this brief study leaves many interesting questions untouched. For example, it would be interesting to extend it into the regime of dynamical evolution which follows core collapse, using a more refined N-body program dealing efficiently with close binaries. It would also be desirable to extend it to much larger values of N, values more in line with those which are customary in present-day simulations (e.g. $N \gtrsim 10^3$). The results of the above survey, along with a little theory, are indicative of what to expect, as the following argument shows.

Theory implies that the escape rate (number of escapers per crossing time) should be approximately independent of N, and that the mean energy of one escaper should be $\propto E/N$ approximately. Hence it follows that the rate at which escapers carry off energy is given approximately by $dE/dt \simeq 0.04E/(Nt_{cr})$, where a constant of proportionality has been estimated from the above series of 100-body simulations. Since the energy of escaping stars appears to be a sensitive indicator of integration accuracy, it seems desirable to ensure that the total energy is conserved to an accuracy better than this, which leads to the bound $|\Delta E/E| \ll 0.04t/(Nt_{cr})$ on the cumulative relative error in time t. For $N = 100$ up to the time t_{10} of formation of the first energetic binary, this leads to a limit of order 0.004 on the cumulative relative error (cf. Table 1).

All these considerations do nothing, however, to relieve one possible worry. It may be that numerical simulations of adequate accuracy give consistent results only because they are all equally inaccurate (by the standards which would be necessary to compute reliable positions and velocities for the individual stars.)

9. DESIRABLE PROPERTIES OF N-BODY SIMULATION PROGRAMS

If we are to rely on N-body simulations to produce consistent statistical results, we should consider how they ought to be designed in order to perform this task. A great deal is known about how to control the local truncation error in an integration routine, so that accurate positions and velocities can be ensured (at least for limited intervals of time). Very little seems to be known, by contrast, about how to ensure that the results are statistically valid.

We have already indicated that satisfactory energy conservation is a necessary condition for obtaining reliable statistical results on escaping stars, and we have been able to make this assertion approximately quantitative. Similar conditions can be obtained by requiring that the relaxation of stellar energies (cf. §7 above) should be simulated with sufficient accuracy.

The N-body equations have many other conservation properties, including conservation of the *Poincaré invariants* (cf. Arnold 1978). The first of these is $\omega^2 = \sum dp_i \wedge dq_i$, where each term is a 2×2 determinant, and q_i, p_i are cartesian coordinates and their conjugate momenta. The last invariant gives conservation of volume in the phase space $R^{6N} = \{(q_1, \cdots, q_{3N}, p_1, \cdots, p_{3N})\}$, i.e. Liouville's theorem, which is fundamental to statistical mechanics. Since we are attempting to

ensure the sound statistical behaviour of simulated N-body systems, preservation of the Poincaré invariants would seem to be a desirable property to require. Indeed it is guaranteed if the numerical scheme yields a symplectic map. An example is the familiar time-centred leapfrog, i.e.

$$\begin{aligned} \mathbf{v}_{i+1/2} &= \mathbf{v}_{i-1/2} + \mathbf{a}(\mathbf{r}_i)\Delta t \\ \mathbf{r}_{i+1} &= \mathbf{r}_i + \mathbf{v}_{i+1/2}\Delta t, \end{aligned}$$

where $\mathbf{v}_i$ and $\mathbf{r}_i$ are the velocity and position of a star at the ith step, respectively, $\mathbf{a}$ is the acceleration, and Δt is the time step, which here must be the same for all stars. On the other hand there appears to be no indication that symplectic schemes give better results when a system is chaotic (Channel & Scovel 1988).

10. INSTABILITIES AT THE STATISTICAL LEVEL

Finally, let us speculate on what we shall find when we study the results of N-body simulations, assuming that they are statistically reliable. Our best guide here are results based on simplified models for the evolution of stellar systems. We have already mentioned one of these models: the Fokker-Planck equation, which is a simplified form of the Boltzmann equation; and there are others.

These models show most interesting behaviour in the regime which follows core collapse, when binaries are sufficiently active to cause an overall expansion of the cluster. It turns out that the expansion can be unstable, depending on the number of stars in the cluster. Sometimes the expansion is modified by a regular oscillation, but there are values of N where the expansion looks quite chaotic (Heggie & Ramamani 1989). Indeed it is tempting to summarise the results by stating that, as N increases, the evolution exhibits the familiar period-doubling route to chaos (cf. Guckenheimer & Holmes 1983).

These intriguing observations have been greatly sharpened in recent work by Breeden *et al.* (1990), who have applied some of the standard techniques in chaos theory to study the irregular evolution which occurs for large values of N. For example they have measured positive "Liapounov exponents", and have shown that the solutions occupy an approximately two-dimensional submanifold, rather like the Rössler attractor. In this regard, it would be very interesting to produce a highly simplified model of the evolution of star clusters with three degrees of freedom (which seems to be the minimum needed to exhibit this kind of behaviour), and incorporating the essential physics of N-body systems. The "evaporative model" of stellar systems (cf. Spitzer 1987, §3.1) would seem to be a suitable starting point for such an investigation.

A number of very curious issues are raised by these results. For example they show that stellar systems are still unpredictable, even if we are able to ignore the instability of individual stellar orbits, and concern ourselves only with statistical results. Also they illustrate the old paradoxes by which a Hamiltonian system is able to exhibit behaviour in accordance with the laws of statistical mechanics: it is impossible for a Hamiltonian system to possess a low-dimensional attractor. Indeed these results suggest one other property which N-body simulations ought to possess, namely, correct statistical behaviour in the sense of the laws of thermodynamics.

In conclusion, it can be seen that the stellar N-body problem is an excellent and astrophysically important example of a system which, from every practical point of view, is highly chaotic. Indeed, chaotic systems appear to be as common in stellar dynamics as they are rare in celestial mechanics. Though it is difficult to study rigorously, it provides a concrete illustration of several of the important aspects of more abstract chaotic systems. Finally, it raises important questions about how such systems can be faithfully simulated.

Aarseth, S.J., 1974, Dynamical Evolution of Simulated Star Clusters I. Isolated Models, *Astron. Astrophys.*, 35:237.

Aarseth, S.J., 1985, Direct Methods for N-Body Simulations, in: *Multiple Time Scales*, J.U. Brackbill & B.I. Cohen, eds., Academic Press, New York.

Arnold, V.I., 1978, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York.

Binney, J. & Tremaine, S., 1987, *Galactic Dynamics*, Princeton University Press, Princeton.

Breeden, J.L., Packard, N.H. & Cohn, H., 1990, Chaos in Astrophysical Systems: Core Oscillations in Globular Clusters, preprint, CCSR-90-2 (Center for Complex Systems Research, Dept. of Physics, Beckman Institute, University of Illinois at Urbana-Champaign).

Carnevali, P. & Santangelo, P., 1980, Automated Graphical Plots for the Study of the Gravitational N-body Problem, *Mem.S.A.It.*, 51:529.

Channel, P.J. & Scovel, C., 1988, Symplectic Integration of Hamiltonian Systems, preprint, LA-UR-88-1828 (Los Alamos).

Dejonghe, H. & Hut, P., 1986, Round-Off Sensitivity in the N-Body Problem, in: *The Use of Supercomputers in Stellar Dynamics*, P. Hut & S. McMillan, eds., Springer-Verlag, Berlin.

Goodman, J., Heggie, D.C. & Hut, P., 1990, On the Exponential Instability of N-Body Systems, preprint.

Guckenheimer, J. & Holmes, P., 1983, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York.

Gurzadyan, V.G. & Savvidy, G.K., 1986, Collective Relaxation of Stellar Systems, *Astron. Astrophys.*, 160:203.

Heggie, D.C., 1988, The N-Body Problem in Stellar Dynamics, in: *Long-Term Dynamical Behaviour of Natural and Artificial N-Body Systems*, A.E. Roy, ed., Kluwer, Dordrecht.

Heggie, D.C. & Ramamani, N., 1989, Evolution of Star Clusters After Core Collapse, *M.N.R.A.S.*, 237:757.

Krylov, N.S., 1979, *Works on the Foundations of Statistical Physics*, Princeton University Press, Princeton.

Lecar, M., 1968, A Comparison of Eleven Numerical Integrations of the Same Gravitational 25-Body Problem, *Bull. Astron.*, 3:91.

Lichtenberg, A.J. & Lieberman, M.A., 1983, *Regular and Stochastic Motion*, Springer-Verlag, New York.

Miller, R.H., 1964, Irreversibility in Small Stellar Dynamical Systems, *Ap.J.*, 140:250.

Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T., 1986, *Numerical Recipes*, Cambridge University Press, Cambridge.

Smith, H., Jr., 1977, The Validity of Statistical Results from N-Body Calculations, *Astron. Astrophys.*, 61:305.

Spitzer, L., Jr., 1987, *Dynamical Evolution of Globular Clusters*, Princeton University Press, Princeton.

Standish, E.M., 1968, *Numerical Studies of the Gravitational Problem of N Bodies*, Ph.D. Thesis, Yale University.

Szebeheley, V.G. & Peters, C.F., 1967, Complete Solution of a General Problem of Three Bodies, *A.J.*, 72:876.

Valtonen, M.J., 1974, Statistics of Three-Body Experiments, in: *The Stability of the Solar System and of Small Stellar Systems*, Y. Kozai, ed., Reidel, Dordrecht.

Wielen, R., 1968, On the Escape Rate of Stars from Clusters, *Bull. Astron.*, 3:127.