SHADOW ORBITS AND THE GRAVITATIONAL N-BODY PROBLEM

GERALD D QUINLAN AND SCOTT TREMAINE
Canadian Institute for Theoretical Astrophysics, McLennan Labs, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 1A1, Canada.
E-mail: quinlan@cita.utoronto.ca
E-mail: tremaine@cita.utoronto.ca

1. Chaos in the gravitational N-body problem

The gravitational N-body problem (GNP) is to find the solution of the equations of motion

$$\frac{d^2\mathbf{r}_i}{dt^2} = -G\sum_{j\neq i} m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3}, \tag{1}$$

where m_i and $\mathbf{r}_i$ are the mass and position of particle i, $i = 1, \ldots, N$. In the usual systems examined in stellar dynamics (such as galaxies and star clusters) the motion is chaotic, in the sense that small perturbations to the position of an individual particle grow exponentially with time, $|\Delta\mathbf{r}| \sim \exp(\mu t)$, where μ is the (maximum) Lyapunov exponent. For hot systems such as star clusters or elliptical galaxies, the e-folding time is $\mu^{-1} \sim t_{\rm cr}$, or possibly $\mu^{-1} \sim t_{\rm cr}/\ln N$, where $t_{\rm cr}$ is the crossing time and N is the number of stars (Goodman, Heggie, and Hut 1988; Heggie 1991; Kandrup and Smith 1991).

Because of this exponential divergence the orbits found by numerical integration of the equations of motion (1) — we call these "noisy" orbits to distinguish them from the "true" orbits that are exact solutions of the equations — are extremely sensitive to numerical errors (Miller 1964). The exponential growth of errors means that there is no obvious way to estimate how accurately a numerical integration should be done. Comparing the results after more than a few crossing times with those of independent calculations (using different timesteps, programs, computers, etc.) almost always shows large differences (Lecar 1968). The most common check is the conservation of integrals such as the total energy (or angular momentum), but this is a weak test because the errors in the positions can be large even though the energy is conserved to high accuracy.

The absence of convincing accuracy checks means that numerical solutions of the GNP are sometimes viewed with suspicion (Miller 1964). Most dynamicists believe that the overall statistical properties of the final state are likely to be correct, despite the possibly large errors in the individual phase-space positions, but even optimists need criteria to decide how accurate the integration has to be to yield reliable statistical results.

Here we briefly describe a novel measure of the accuracy of N-body integrations. We shall say that a noisy orbit is accurate if there is a true orbit that remains close

B. Sundelius (ed.) Dynamics of Disc Galaxies (1991) Göteborg, Sweden, pp. 143–148.

to the noisy orbit throughout the integration (the true orbit is said to "shadow" the noisy orbit). Thus we are not worried if the noisy orbit diverges exponentially from the true orbit with the same initial conditions, so long as there is some other true orbit, with slightly different initial conditions, that remains close to the noisy orbit. The maximum phase-space distance between the noisy orbit and the shadow orbit (the "shadow distance") then parametrizes the accuracy of the integration. If a shadow orbit exists, the noisy orbit represents the true dynamics in the sense that it remains close to some true orbit of the system.

2. Constructing shadow orbits

The idea of a shadow orbit can be illustrated by a simple example. Consider the equation $x'' = x$, or, in position-velocity phase-space, $x' = v$, $v' = x$, with initial conditions $x = v = 0$ at time $t = -10$. Imagine that we integrate this equation exactly except for noise consisting of a discontinuous jump ϵ in velocity at $t = 0$. Thus the noisy orbit is

$$x_{\rm n}(t) = \begin{cases} 0 & \text{if } t < 0, \\ \frac{1}{2}\epsilon(e^t - e^{-t}) & \text{if } t \geq 0. \end{cases} \tag{2}$$

After the jump the noisy orbit diverges exponentially from the true orbit $x(t) = 0$. However, there is a shadow orbit $x_{\rm sh}(t) = \epsilon e^t/2$, which is a true orbit that remains within $\epsilon/\sqrt{2}$ (in phase space) of the noisy orbit at all times.

A more realistic example of shadowing is described by Grebogi et al. (1990; hereafter GHYS). They examined a two-dimensional Hamiltonian map of the unit square onto itself (known as the standard map):

$$J_{n+1} = J_n + (K/2\pi)\sin(2\pi\theta_n) \pmod 1, \tag{3}$$

$$\theta_{n+1} = \theta_n + J_{n+1} \pmod 1. \tag{4}$$

For the choice $K = 3$ examined by GHYS, most of phase space is filled with chaotic orbits, and the separation between a noisy orbit and a true orbit typically increases by a factor of 3 with each iteration. Thus even in double precision arithmetic (with 16 decimal places accuracy) the noisy orbit will be completely different from the true orbit starting from the same initial conditions after about 30 iterations. Yet the noisy orbit can be shadowed for much longer: in one example GHYS constructed a noisy orbit with one-step errors (i.e., the errors made in going from one iteration to the next) of order $\delta_f = 10^{-14}$, and proved that this orbit was shadowed by a true orbit for $N = 10^7$ iterations, with shadow distance $\delta_x = 10^{-8}$. The long shadowing time is impressive considering the rapid rate at which nearby orbits diverge.

For most dynamical systems, including the standard map, noisy orbits cannot be shadowed by true orbits for arbitrarily long times. The failure does not arise through a gradual increase in the shadow distance as the orbit is shadowed for longer and longer times. Instead, there are isolated regions of a few iterations ("glitches") over which no true orbit can be found that shadows the noisy orbit. GHYS conjecture that for a chaotic two-dimensional Hamiltonian map in which the typical noise per step is $\delta_f > 0$, there should be true orbits that shadow noisy orbits within $\delta_x \sim \sqrt{\delta_f}$ for $N \sim 1/\sqrt{\delta_f}$ iterations before encountering a glitch.

To prove the existence of shadow orbits GHYS used two algorithms called containment and refinement (see also Hammel, Yorke, and Grebogi 1988). Refinement is an iterative procedure analogous to Newton's method for root finding; it takes as input a noisy orbit and after iterating to convergence returns a true orbit that shadows the original noisy orbit (thus it "refines" the noisy orbit by reducing the noise at each iteration). Like Newton's method, refinement is superconvergent: the number of significant digits in the approximation to the true orbit typically doubles on each iteration. In some cases the iteration does not converge, which is a sign that the noisy orbit contains a glitch. In practice without exact arithmetic we cannot iterate the refinement to convergence, and we stop the iteration when the one-step errors in the refined orbit are comparable with the precision of the arithmetic. The existence of a shadow orbit can then be proved with the containment procedure, which involves the construction of a sequence of small boxes in phase space around the points of the noisy orbit in such a way that a true orbit must pass through them.

Our experience suggests that in practice the refinement procedure alone is sufficient to test for the existence of shadow orbits (although this test is not rigorous without the containment procedure). This conclusion is based on experiments done on the standard map with the symbolic manipulation language Maple using 100 significant decimal places. When iterating the refinement procedure on a noisy orbit we always found one of two outcomes: either the refinement would converge geometrically, with the one-step errors in the refined orbit becoming $\leq 10^{-100}$ after a small number of iterations, or the refinement would stop converging after just one or two iterations, with the errors in the refined orbit being little better (and often worse) than those in the noisy orbit. This result strongly suggests (but does not prove) that if a noisy orbit with errors of, say, $\delta_f \simeq 10^{-6}$ can be refined until the errors are reduced to $\simeq 10^{-12}$ then further iterations of the refinement procedure with exact arithmetic would yield a true orbit to arbitrarily high accuracy. If we allow this assumption then we can carry out refinement calculations using double-precision floating-point arithmetic rather than Maple, which makes the calculations go much faster. This is the approach that we take in the N-body experiments described below.

We have generalized the refinement procedure of GHYS in a straightforward way to work for Hamiltonian systems with more than two phase-space coordinates, and have tested it on four-dimensional standard maps (two standard maps coupled together) and on two- and four-dimensional maps describing the motion of particles through gravitating systems. We found that these maps obeyed the conjecture of GHYS for two-dimensional maps.

3. Results from Aarseth's NBODY1 program

We present some preliminary results on shadowing in gravitational N-body systems using noisy orbits produced by Aarseth's popular and widely available NBODY1 program (Aarseth 1985; Binney and Tremaine 1987; see Sellwood 1987 for a review of techniques for integrating the GNP). This program contains an accuracy parameter η that determines the time step δt for a given particle through the relation $\delta t = (\eta F/\ddot{F})^{1/2}$, where F and $\ddot{F}$ are the force and the second time derivative of the force acting on the particle. The recommended value is $\eta = 0.03$, which usually conserves the total energy to better than one part in 10^4 over one crossing time. The truncation errors in position and velocity vary as $(\delta t)^7$ and $(\delta t)^6$ respectively, and hence the one-step phase-space error (δ_f) varies as $(\delta t)^6$. It is customary in N-body work to soften the force law by replacing the denominator in equation (1) by $(|\mathbf{r}_i - \mathbf{r}_j|^2 + h^2)^{3/2}$, with h being the softening length, but we used $h = 0$ to have a more challenging problem.

It is not computationally feasible to search for shadow orbits in the GNP when N is large. We have therefore examined a simpler problem in which $N-1$ of the particles are fixed and only one particle is free to move. For the fixed particles we chose 100 particles of mass $m = 0.01$ from Plummer's density distribution, $\rho(r) \propto (1 + r^2/r_p^2)^{-5/2}$, with $G = 1$ and r_p (the Plummer radius) $= 3\pi/16$ (the units are chosen so that the total mass and gravitational energy are 1 and $-\frac{1}{2}$ respectively). If the Plummer model were in virial equilibrium with an isotropic velocity distribution, the central velocity dispersion would be $v_{\rm rms}(0) = \sqrt{8/3\pi}$. We define the crossing time by $t_{\rm cr} = r_p/v_{\rm rms}(0)$. With these units the position and velocity of a typical orbit are both of order unity.

We construct a noisy orbit by starting a particle at the centre of the cluster with a random velocity direction and integrating for a given number of time steps using NBODY1. The noisy orbit consists of the phase-space coordinates returned by the integrator at each time step; the errors are typically $\delta_f \simeq 10^{-5}(\eta/0.03)^3$. We then apply the refinement procedure to the noisy orbit, iterating until the errors in the refined orbit stop shrinking. The integrations required in the refinement procedure are done by the Bulirsch-Stoer method and are much more accurate (and time-consuming) than the original noisy integration. If we can reduce the one-step errors in the refined orbit to 10^{-12} or better we assume that a shadow orbit has been found; we then increase the number of time steps by a factor of 1.3 and repeat the process, continuing until we find the maximum time $T_{\rm max}$ for which a shadow orbit can be found.

An example is shown in Figure 1. The solid line is the noisy orbit of a particle starting at the origin with velocity $1.5v_{\rm rms}(0)$, integrated for about $39\,t_{\rm cr}$. The long-dashed line is an accurate Bulirsch-Stoer integration starting from the same initial conditions; this orbit soon deviates significantly from the noisy orbit and we stop plotting it after $21\,t_{\rm cr}$. The short-dashed line is the shadow orbit, which starts from a position displaced from the origin by about 10^{-4} and remains within a distance 0.021 (in phase space) of the noisy orbit throughout the integration. At most times steps the distance between the noisy and shadow orbits is much smaller than 0.021, the typical distance being $\leq 10^{-3}$. Efforts to shadow the noisy orbit for longer than shown in Figure 1 were not successful;

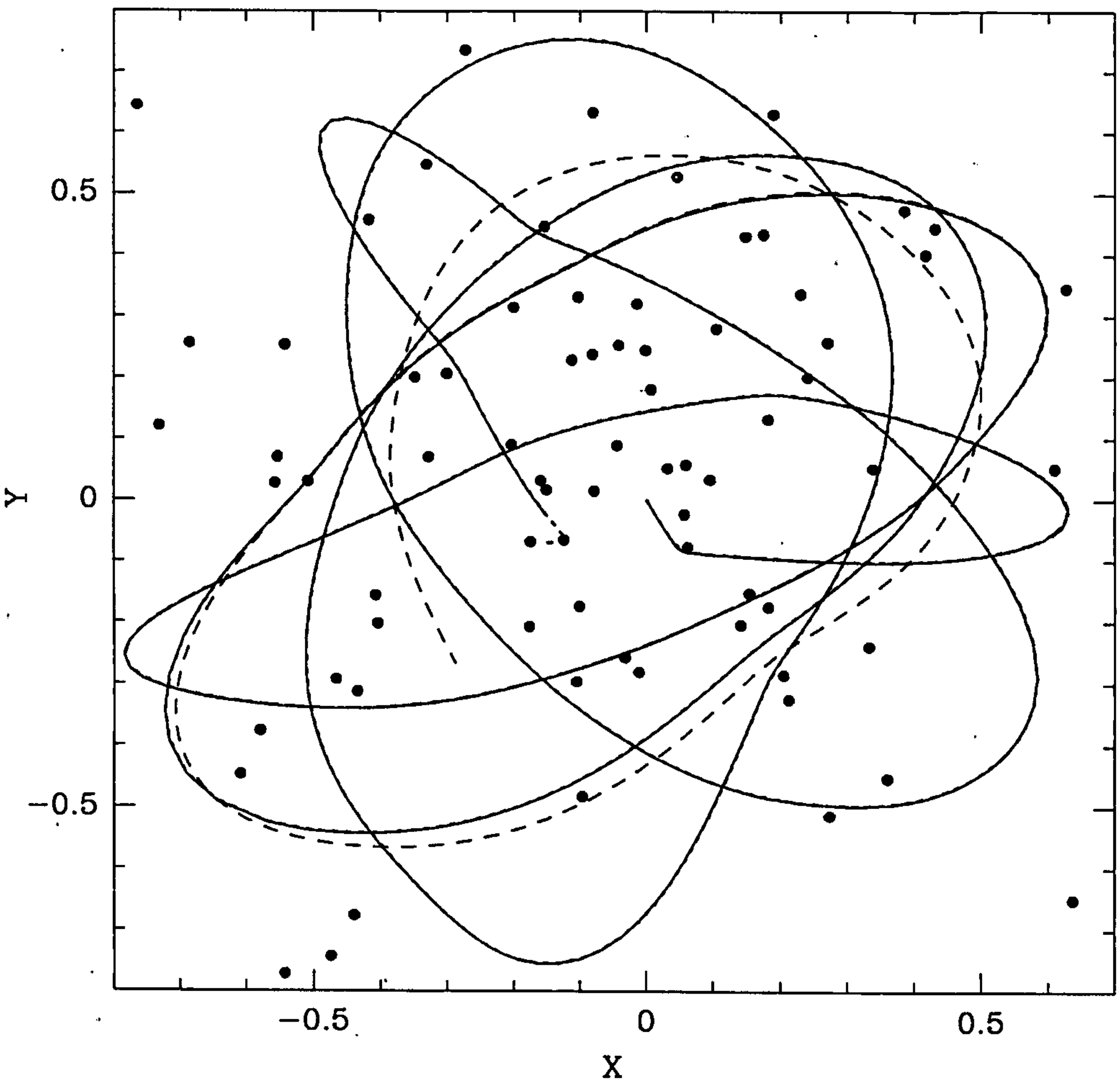

Figure 1. The solid line is the projection onto the x-y plane of the (noisy) orbit of a particle starting at the origin with velocity $(v_x,v_y,v_z)\simeq(0.41,-0.73,1.08)$. The filled circles are the projected positions of the fixed particles in the cluster. The orbit is integrated for a time $t \simeq 39\, t_{\rm cr}$ using accuracy parameter η =0.02. The long-dashed line is an accurate Bulirsch-Stoer integration starting from the same initial conditions, integrated for about $21\, t_{\rm cr}$. The short-dashed line is the shadow orbit. The phase-space separation between the shadow and noisy orbits reaches a maximum of 0.021 (resulting mostly from the separation in velocity, not position) near the point x =0.188, $y = -0.281$. The shadow orbit can be seen extending from the end of the noisy orbit (because the shadow orbit has been plotted for a slightly longer time interval).

the glitch occurred at a close encounter with the fixed particle at the end of the shadow orbit in the figure.

We have integrated a variety of noisy orbits with various values of the accuracy parameter η. In general $T_{\max}$ increases and the shadow distance δ_x decreases as η decreases (i.e. as the integrations become more accurate). We have routinely shadowed orbits for up to $20 - 40\,t_{cr}$. In most cases the glitch that prevented the noisy orbit from being shadowed for a longer time occurred at a close encounter with one of the fixed particles. Some preliminary experiments with softened force laws suggest that the shadow times can be much longer if the effects of close encounters are suppressed.

4. Summary

The use of shadow distance as an accuracy measure has two principal limitations. First, the computation of a refined orbit is much more expensive than the computation of the original noisy orbit, and therefore it is not practical to use shadowing to provide direct accuracy estimates for all numerical solutions of the GNP, especially not for problems with a large number of particles moving simultaneously. The best that can be done is to estimate the errors in a few representative systems. Second, there is no guarantee that the shadow orbits are "typical" (i.e., that their statistical properties are the same as those of true orbits chosen at random), although there is no reason to suppose otherwise. In spite of these limitations, the existence of shadow orbits shows that numerical integrations of N-body systems can yield orbits that closely resemble true orbits, and, we believe, puts the validity of these integrations on a sounder footing.

References

Aarseth, S. J. (1985) in *Multiple Time Scales* (Academic Press) pp. 377–418.
Binney, J. and Tremaine, S. (1987) *Galactic Dynamics* (Princeton University Press).
Grebogi, C., Hammel, S. M., Yorke, J. A. and Sauer, T. (1990) *Phys. Rev. Lett.*, **65**, 1527–1530.
Hammel, S. M., Yorke, J. A. and Grebogi, C. (1988) *Bull. Am. Math. Soc.*, **19**, 465–469.
Goodman, J., Heggie, D. C. and Hut, P., (1988) preprint.
Heggie, D. C. (1991) in *Predictability, Stability, and Chaos in N-body Dynamical Systems*, ed. A. E. Roy, in press.
Kandrup, H. E. and Smith H. (1991,) *Astrophys. Journal*, **374**, 255–265.
Lecar, M. (1968) *Bull. Astronomique*, **3**, 91–104.
Miller, R. H. (1964) *Astrophys. Journal*, **140**, 250–257.
Sellwood J. A. (1987) *Ann. Rev. Astron. Astrophys.*, **25**, 151–186.