A Brief Survey of Issues Relating to the Reliability of Simulation of the Large Gravitational N-body Problem

Ph.D. Qualifying Depth Paper, Dept. of Computer Science, University of Toronto
Wayne B. Hayes

Note to readers: Those unfamiliar with common terms from stellar dynamics like *softening, collision, regularization, etc.*, should probably read the glossary first. (Page 23.)

1 Introduction

A gravitational N-body system consists of a set of N particles of mass m_i, and position $\mathbf{r}_i$, $i = 1, ..N$, all moving under the force of their mutual gravitational attraction. The force particle j exerts on particle i is

$$\mathbf{F}_{ij} = \frac{Gm_i m_j}{r_{ij}^2}\bar{\mathbf{r}}_{ij} = \frac{Gm_i m_j}{|\mathbf{r}_{ij}|^3}\mathbf{r}_{ij} = -\mathbf{F}_{ji} \tag{1}$$

where G is Newton's gravitational constant, $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$, $r_{ij} = |\mathbf{r}_{ij}|$, and $\bar{\mathbf{r}}_{ij} = \mathbf{r}_{ij}/r_{ij}$. The total force on particle i is the sum of the forces exerted on it by all the other particles.

We study the reliability of direct numerical simulation of the large N gravitational N-body problem. A partial list of errors introduced in such simulations includes [33]: finite-N sampling, or *discreteness noise*, because the N used in simulations usually is much smaller than the N of the system being modelled; *force softening*, which eliminates the singularity at $r = 0$ and lessens spurious collisional effects introduced by finite-N sampling [9]; using fast but *approximate force algorithms* such as tree codes [6, 45] or particle-mesh methods (*eg.,* [61]); numerical ordinary differential equation (ODE) integration *truncation* error; and *machine roundoff* error. The goal of the study of reliability is to ascertain what effect these *input* or *local* errors have on the *output* or *global* error. Less stringently, we can be interested in the effect of these errors on statistical measurements taken from the simulations.

What is meant by "reliability" will certainly depend on what we are trying to measure or produce with our simulations. For example, the goal of a simulation of a cold, collisionless disk may be simply to reproduce global effects like spiral arms, or to test the effect of various distributions of dark matter. In such simulations, the microscopic dynamics of individual particles are of little interest as long as the global structure is valid. In contrast, when simulating a highly collisional globular cluster, it is widely known (*eg.,* [75]) that microscopic effects, like close encounters and formation of binary stars in the core, have significant impact on the global evolution of the cluster. Thus, to reliably reproduce global properties of globular clusters, these microscopic details must be taken into account, even though the particular happenings of any individual simulation may be of little interest.

We are interested both in the theory and practice of error control. Although we realize that a practical error measure must be devised in order to control error, we are also interested in the theoretical question of *what is the strongest statement that we can make about how a numerical simulation reflects the real system?* A very strong statement, for example, might have something to do with shadowing of phase-space trajectories, whereas for practical error control we may attempt to conserve energy to some given accuracy. We would be particularly interested in a statement that would link, for example, the quality of energy conservation to the shadowing property.

The present paper is a brief tour of some of the literature concerning the topic of the title. The survey is by no means complete[1]. We will look at some simulations done by astronomers, and the kinds of measurements taken. We will then connect N-body simulation to the study of chaos, because the N-body problem has *sensitive dependence on initial conditions*, and to the study of *shadowing*, which is one of many kinds of

[1] *I.e.,* the author is under no illusions about the remaining extent of his ignorance.

global error analysis techniques for ODE integration. Finally, we will briefly mention the connection between shadowing and Boundary Value Problems for ODEs.

2 N-body simulation

2.1 Softening and discreteness noise

One of the fundamental limitations of current N-body simulations of stellar systems is that they usually contain several orders of magnitude fewer particles ("stars") than real stellar systems contain. Thus, each simulated particle feels a much more "lumpy" or *granular* gravitational field than a star would feel in a real system. For collisional systems this may mean that timescales and statistics are distorted by some approximately computable amount [38], but for collisionless systems this spurious collisionality is simply a source of error [39]. The effect of close encounters can be moderated via *force softening*, but the granular effect of the gravitational potential is strong even for particles that are separated by large distances ([75, section 1.3a], [68]). A common force computation method for collisionless systems is the *particle-mesh* method (*eg.,* [61]), in which softening is implicit and equal to the distance between mesh points. For particle-particle or tree methods for force computation, softening is added explicitly by modifying the gravitational force as

$$|F_{ij}| = G\frac{m_i m_j}{r_{ij}^2 + \epsilon_{soft}^2}, \tag{2}$$

where ϵ_{soft} is the softening parameter. It is usually chosen to approximate the average inter-particle distance. However, Sellwood [68] notes that if a disk is being simulated in two dimensions, then the softening parameter ϵ_{soft} can instead be used to approximate a disk of thickness ϵ_{soft}. He also argues that a simple smoothing of the gravitational potential is probably an oversimplification, since real spiral galaxies contain giant molecular clouds and star clusters that make the potential significantly less smooth than the theorist's ideal. He includes a diagram clearly showing that spurious spiral structure can develop in a *cold disk* as a result of discreteness noise. This structure is clearly visible in a disk of 2,000 particles after only 1 revolution, and is visible even in a disk with 20,000 particles after a few revolutions. Discreteness noise can also introduce spurious heating of a collisionless *cold disk*, so that artificial disk-cooling mechanisms must be introduced [68, 57].

Dyer and Ip [22] note that equation (2) is physically inconsistent: particles act like Plummer distributions when they are contributing to the field, but are points when the field acts upon them, leading to a violation of Newton's Third Law and thus a violation of conservation of momentum. They fix this with a model in which softened particles are treated as uniform density spheres. Non-overlapping spheres interact as points, while a polynomial is used to compute the force on spheres that overlap. (The spheres are not solid in any physical sense. They can pass through each other like "ghosts".)

When approximate force computation algorithms are used, they are generally the next most dominant source of error. Barnes and Hut [7] analyzed their tree code [6] for errors and found that it works well for collisionless systems, but their results were inconclusive about collisional systems. (Jernigan and Porter [45] introduce a tree method that is completely different from the Barnes and Hut algorithm, and works very well for collisional systems.) Barnes and Hut [7] found that errors in global conservation of energy were dominated by force errors, and concluded that a modest error tolerance in force of about one part in 10^2 per particle was sufficient to conserve global energy to one part in 10^{3-4}, and that "further improvements in tree methods for a fixed N is much less desirable than simply increasing N." Hernquist, Hut, and Makino [39] came to a similar conclusion after doing many experiments with various accuracies of force computation and values of N from 10^2 to 10^5. They measured individual particle energies (which are conserved in collisionless systems), and found that, for a given N, the energy deviations were almost identical regardless of the force error, and that energy errors decrease uniformly as N is increased. In other words, the effects of discreteness noise dominate those of force errors, so that simply increasing N is the most desirable improvement that can be made to current N-body simulations. They state that they believe this result is "*independent of N*" (their emphasis). This author believes that they are probably correct for values of N likely to be used in the foreseeable future, but the statement is clearly false in general, because we could, in principle, use the actual N of a real system.

A bizarre alternative is offered by Bryant [11, 12], who introduces a model in which velocities are always bounded and stellar encounters are interpreted as mutual absorption and emission of the simulated bodies.

2.2 Statistics of collisional N-body simulations

Most simulations of stellar N-body systems are done for the purpose of observing global evolution or for taking statistics of global measurements. The microscopic details of which star goes where are rarely of interest. On this basis, a sufficient condition for simulation reliability is that all global measurements of interest from the real system should be reproduced by the simulation. In fact, the assumption that macroscopic measurements are valid (the *MMV* assumption), even though the microscopic details are possibly wrong, is a necessary working assumption for practicing N-body researchers. It is difficult to see how this assumption could be verified conclusively without an understanding of how the simulation treats microscopic details (for example, a set of self-consistent simulations may contain systematic errors such that they are all wrong in the same way), but most practitioners of stellar N-body simulation take the MMV hypothesis as a working assumption anyway. Heggie [36] states that this assumption "is little better than an article of faith." Nonetheless, we are led to the question of what accuracy level is necessary in order for us to collect undistorted statistics.

When measuring statistical quantities, we must keep in mind that even if we had a perfect N-body simulator, we could only run it a finite number of times, so there would still be the fundamental noise that is present in any finite-sized statistical sample. Thus, there is little use in attempting to measure statistics any more accurately than this limit dictates [74, 73, 35, 38]. Also, a disadvantage of using direct N-body simulation to collect statistics is that only a small amount of the wealth of information in a full N-body simulation is ever extracted.[2] Thus direct N-body simulation is not a particularly information-efficient method of gathering statistics [35].

Smith [74] used a 2nd order variable timestep predictor/corrector with no regularization and double precision arithmetic to test the hypothesis that some statistics are relatively insensitive to integration accuracy. The statistics he measured in globular cluster simulations were the first and third quartic mean stellar density, energy, and radius, as well as the median radius. He compared simulations with varying degrees of accuracy using various statistical tests and found the statistics to disagree at only the 10% level, *i.e.,* there was little evidence to conclude that the simulations gave different results according to integration accuracy. He concluded that the statistics taken are reliable unless energy conservation is grossly violated. However, Lecar [53] chose to measure different statistics, and found them to vary widely depending on which machine or algorithm was used to integrate the problem. This will be further discussed below.

2.2.1 Efficiency of many small *vs.* one large simulation

Smith [73] and Heggie [38] have noted that n simulations each with N/n particles is much less expensive than 1 simulation with N particles, and is also completely trivial to parallelize. Such sets of simulations can also give better statistics based on the larger sample size, as long as it is known how to scale the statistics from smaller simulations to larger ones. Unfortunately, this scaling is non-trivial because different processes scale differently with N, so it is possible that they interact differently in current simulations than they do in real clusters [38]. Again, we are led to the conclusion that increasing N is the most desirable way of increasing the reliability of N-body simulations.

[2] That is, no finite list of statistics completely describes a full phase-space trajectory.

Smith [73] continued his previous study ([74]) to ascertain the effects of different numbers of particles on statistical measurements. He measured density profiles and binary star formation rates for clusters with $N = 8, 16, 32, 64$ and found that, although the scaled statistics for $N = 8$ differed substantially from those with higher N, the statistics for $N = 16$ and higher agreed quite well. Furthermore, the number of binaries formed scaled as $N_b \sim N^{0.45 \pm 0.05}$, consistent with a theoretical estimate of $\sqrt{N}$. He didn't measure information on escapers, since escapes happen sufficiently rarely that $N = 64$ doesn't provide enough of a sample. He concluded that better statistics can be taken with many small simulations than with 1 large one.

Although this method is less effective at modelling some processes that really do need a large N, such as gravothermal effects, it may be reliable enough to study processes such as core collapse rates, escape rates, binary star formation rates, density profiles, and gradients of velocity dispersion with radius [38]. However, this does not mean that we can learn everything there is to learn about these statistics by running millions of simulations with $N = 64$, because discreteness noise is still a factor. What these results say is that at any given time, if N_{big} is the largest simulation one is currently capable of running, it may be more productive to run 100 simulations with $N = N_{big}/10$ than a single simulation with $N = N_{big}$.

2.3 Exponential divergence in N-body systems

In this author's opinion, there is currently no conclusive evidence that the MMV assumption is correct. That it is generally accepted seems mostly to stem from the heuristic observation that N-body simulation results often *look* quite reasonable, and in fact sometimes reproduce results predicted by theory. For example, theoretically predicted phenomena in globular clusters such as mass segregation, core collapse, binary star formation rates, and relaxation times all seem to be reproduced reasonably faithfully by good N-body simulations [38]. Astronomers have also been successful in producing realistic-looking spiral structure in collisionless disks, but it is not clear that the simulations produce this structure for the same reason that real spiral galaxies do [68].

One of the strongest reasons for suggesting caution in the interpretation of results of N-body simulations is that N-body systems display *sensitive dependence on initial conditions*. This means that, at the particle level, two solutions whose phase-space initial conditions differ by an arbitrarily small amount will diverge away from each other exponentially with time. Since a numerical simulation constantly introduces small errors (cf. section 1) to the solution, we are guaranteed that our numerical solution diverges exponentially away from the true solution with the same initial conditions. This is the case even if quantities such as energy are conserved to arbitrary accuracy, since there is an infinite number of qualitatively different phase-space trajectories occupying the same energy hyper-surface. Indeed, this exponential magnification of small errors led one astronomer to exclaim, "This view calls into question the very need for accurate simulation." [36]

Miller [58] seems to have been the first to demonstrate that N-body systems display exponential divergence of nearby trajectories, even though all first integrals were well conserved and the evolution of the systems appeared plausible. The e-folding time is defined to be the interval of time it takes for a quantity to increase by a factor of e, Euler's constant (2.7182...). The e-folding time t_e of the exponential divergence seems to be of order the crossing time, t_{cr}, although it increases (*i.e.,* the rate of exponential divergence decreases) as the system becomes less collisional. This has been both derived analytically [36, 37, 27] and measured experimentally [36, 46, 47, 27]. Dejonghe and Hut [21] showed that this exponential divergence can happen even in the smallest possible case of $N = 3$, where they demonstrated that nearby initial conditions can be pushed apart by as much as a factor of 10^{150} during a complex 3-body encounter, although $10^{8} - 10^{15}$ was more the norm. (Their method was to keep the displacement in the linear regime by letting the trajectories diverge, then "pulling them together" along the shortest path, then letting them diverge some more, *etc.* Thus they were able to measure divergences larger than could be represented by the floating point system.) Lecar [53] co-ordinated a study between many researchers, each of whom independently computed the solution to a 25-body problem with identical initial conditions. They measured the radius of the N^{th}

body from the centre of mass, the moment of inertia, and the average squared radial & tangential velocities. They found that different algorithms and computers gave results in which these measurements differed by as much as 100%. Although none of these statistics are the same as the ones mentioned above as being reliable, it is disturbing to note that some statistical measures appear to be reliable while others are not.

Kandrup [46] apparently confuses (according to [27]) exponential divergence of nearby trajectories with relaxation, and provides an unconvincing argument that an inadequately defined phenomenon he calls "mixing" occurs on a crossing timescale, concluding that this mixing is a hitherto unknown form of relaxation.

Kandrup, Smith, and Willmes [47, 48, 49] performed hundreds of numerical experiments to study the details of how exponential divergence is distributed amongst individual particles, and how it behaves with respect to various parameters. They found that the distribution of e-folding times amongst particles had a long tail, indicating that the global divergence was dominated by a small number of particles that, on average, had a divergence rate about 4 times higher than the average, and this distribution changed little with changing initial conditions. (Interestingly, they did not mention *which* particles had the highest divergence rate; this author presumes that they would be the particles that underwent more or harder collisions than average.) They found that the distribution of perturbations amongst particles Δr_i after 4 crossing times was approximately Gaussian in $\log(\Delta r)$, and that the softening parameter had little effect on these results until it was greater than the average inter-particle separation, at which point divergence of nearby trajectories became weaker. They found that when they had particles of unequal mass, the e-folding time was independent of mass. This seems surprising, since it would be plausible to expect the e-folding times to increase with increasing mass, because heavier particles move more slowly and thus suffer fewer encounters, and the encounters are mostly with less massive particles. Perhaps this is offset by the encounters lasting longer with low velocity, and the fact that massive particles sink to the centre of the cluster where they can encounter other massive particles more often. Finally, Kandrup *et al.* postulate that the exponential divergence occurs not only due to the magnifying effect of collisions, but also due to the "pulling apart" force exerted by the $1/r^3$ tidal force.

In contrast, Goodman, Heggie, and Hut [27] argue that the exponential divergence is *solely* a result of encounters, and is greatest in particles that undergo the strongest encounters. They postulate that a completely smooth gravitational potential would display no exponential divergence of nearby trajectories, *i.e.*, $t_e = \infty$. They build a very elegant model that explains the exponential divergence based on magnification of perturbations during encounters, and derive that t_e is of order $t_{cr}/\ln\ln N$, which is nearly independent of N for large N if softening is maintained at the average inter-particle separation, or if softening is not used at all. However, if N is increased with *constant* softening, then $t_e \sim (N\varepsilon^2/R^2)^{1/3}$ for large N. This is partly the basis on which they base their conclusion that $t_e \to \infty$ for a smooth potential. Based on individual trajectory divergence, they derive a bound on the average energy error of a single star per crossing time of $E_{tot}/N^{3/2}$, and postulate that a collisionless simulation is "accurate enough" if individual particle energies are conserved to this tolerance.

Heggie [37] studied many aspects of chaos in stellar N-body systems. He provides yet another derivation of $t_e \sim Kt_{cr}$, similar to Goodman *et al.* [27], and uses numerical experiments to estimate $K \approx \frac{1}{6}$ for large N. Since the derivation is based on the magnification of perturbations during encounters, he concludes that smooth potentials are good for modelling orbits, but not their divergence. Finally, he experiments with integraton accuracy and finds that the only important statistic that seems highly sensitive to integration accuracy is the escape energy and number of escapers: with lax integration tolerances, stars tend to escape the cluster more often, and with higher energy, whereas, with high integration accuracy, fewer stars escape and with less energy. This is reasonable, since escapes usually happen as a result of close approaches, which can tax the integrator and lead to unreal 2-body encounters if too low an accuracy is used. This seems to be a strong argument for the use of regularization in simulations of collisional systems. Furthermore, since escape rates and escape energies are notoriously difficult to predict theoretically [37], it seems important that collisions be handled correctly in order to get a reliable measures of this phenomenon.

One surprising observation is that, to this author's knowledge, nobody has tried doing numerical experiments of exponential divergence in a cold disk, which could potentially help settle the question of whether exponential divergence exists in collisionless systems.

3 Integrators

Unless otherwise stated, the term "system" in this section will mean an autonomous Hamiltonian system. The discussion of symplectic integration for Hamiltonian systems in this section draws heavily from Sanz-Serna [66], which is a very readable introduction to the subject. Another seminal paper is Channel and Scovel [13].

The question of what kind of integrator to use to integrate N-body systems is one that has had plenty of attention. No discussion of stellar system N-body integrators would be complete without mention of Aarseth's classic integration scheme for collisional systems [1], which includes regularization. Makino [55] derived formulae for optimal order and timestep criteria for Aarseth-type integrators. For collisionless N-body systems, and indeed many non-astronomical Hamiltonian systems, symplectic integrators have enjoyed widespread popularity. Recent years have also seen a growing interest in exactly conserving integrators.

The popularity of symplectic integrators stems from the fact that the time-t flow of any Hamiltonian system is a *symplectic* or *canonical* mapping. One consequence of this is that the mapping is volume preserving in phase space. This is a consequence of Liouville's Theorem (see, for example, [56]), which states that the phase-space density of trajectories inside an ensemble remains constant with time. A symplectic integrator also preserves phase space volume to within the machine precision. As a result, symplectic integrators tend to preserve qualitative properties of phase space trajectories: trajectories do not cross, and although energy is not exactly conserved, energy fluctuations are bounded. In fact, it can be shown that the phase space of a linear Hamiltonian problem simply undergoes slight stretches and/or contractions along various axes, so for example a circular trajectory in 2 dimensions becomes an ellipse. Furthermore, if the Hamiltonian function is sufficiently smooth, then the discrete solution produced by a symplectic integrator lies on the exact solution of a problem that has a slightly perturbed Hamiltonian. In general, even for non-smooth Hamiltonians, the discrete solution lies exponentially close in ε to an exact solution to a nearby Hamiltonian problem, where ε is the local error [66].

Non-symplectic integrators generally turn a conservative system into a dissipative one, while a symplectic integrator, on average over time, preserves the conservative nature of the system. Earn and Tremaine [23] point out that roundoff error actually destroys the exact symplecticness of a time- and/or space-continuous system. They demonstrate how one can build a precisely symplectic discrete map over a lattice discretization of space using *integer* arithmetic, so that if the modeled problem is Hamiltonian, then the discrete model is also *exactly* Hamiltonian, but with a slightly different Hamiltonian function. This builds confidence in the qualitative nature of the solution being similar to that of the continuous system.

Since symplectic integrators preserve the topological structure of trajectories in phase space, they are probably more reliable for long-term integrations than non-symplectic integrators, since a non-symplectic integrator causes the system to become dissipative, which in turn can badly distort the topological structure of the phase flow over long periods. However, symplectic integrators are often implicit, and require more function evaluations and smaller timesteps to produce the same computational accuracy as compared to standard non-symplectic integrators [66]. Thus, for short-term integrations where accuracy is paramount, a standard non-symplectic integrator may be more efficient.

Another drawback of symplectic integration is that any integration scheme which is symplectic for constant timestep does not remain symplectic if the timestep is changed based solely and explicitly on the state vector [71]. Recall that the time-Δt map produced by a symplectic integrator lies exactly (or exponentially close to) the true time-Δt flow of a slightly perturbed Hamiltonian problem. However, if Δt changes, then the problem we are close to also changes. Thus, if we vary the timestep, it appears that there is no single Hamiltonian problem our symplecticly-generated solution remains close to.

Nonetheless, several researchers have begun to experiment with variable timestep symplectic integrators. Hut, Makino and McMillan [40] introduce an implicit method of symmetrization which is intended to preserve approximate time symmetry. For the Leapfrog integrator, their method reduces to

$$y_1 = f(y_0, \delta t), \quad \delta t = \frac{1}{2}(h(y_0) + h(y_1)) \tag{3}$$

where $h(y)$ is a function used to compute the timestep, based solely on the state vector. They tested their symmetrized leapfrog on a Kepler problem with eccentricity $e = 0.9$, using 1 implicit iteration per timestep and 10^3 timesteps per revolution. The errors in energy, semi-major axis, eccentricity, time of perihelion passage and longitude of pericenter all have moderately sized periodic errors, but secular drift is nil or linear with a small slope, compared to the non-symmetrized error of linear or quadratic, respectively. They also demonstrate how a 4th-order generalization of Leapfrog can use their time-symmetrization technique, and use it to perform an impressive integration of an orbit with eccentricity $e = 1 - 10^{-6} = 0.999999$. With 3 iterations per timestep of the implicit time-stepping symmetrization equation, they were able to integrate this trajectory using 2×10^3 timesteps per orbit, with a drift in the above quantities of only one part in 10^7, over 10^3 revolutions. This seems very impressive. They do not mention any theoretical results on the symplecticness of their method, however.

Skeel and Biesiadecki [72] introduce another method for using different timesteps while preserving the symplectic nature of the integrator. They show that, if it is possible to split the Hamiltonian into independent additive terms, then each term may be integrated with a different, but constant, timestep, and the integration remains exactly symplectic. Each timestep must be an integer multiple of the smallest timestep. For example, if there are only two step sizes and the large timestep is a multiple M of the smaller, then the force at timestep n is computed as

$$F_n = \begin{cases} F_n^{\text{hard}} + M F_n^{\text{soft}}, & n \text{ a multiple of } M, \\ F_n^{\text{hard}}, & \text{otherwise.} \end{cases} \tag{4}$$

They show that an *artificial* splitting of the Hamiltonian into "quickly" and "slowly" varying components can also utilize their method. They test their method on the Kepler problem, setting the "quickly" varying component to be the half of the orbit that is closest to the central mass, and the slowly varying component to be the other half of the orbit. Their results are similar to the second-order results of Hut *et al.* [40]. It would be interesting to see if a 4th order version of their method would also give comparable results to Hut *et al.* [40]. The error of this method is analyzed more fully in Littell, Skeel, and Zhang [54], where they come to the unsurprising conclusion that the smallest stepsize times the highest frequency of the "quickly" changing components must be small. In other words, their method still needs an adequate sample of the highest frequency changes in order to integrate them. They prove the also unsurprising result that, if ω is the maximum frequency in the system, integrated with timestep h, and $\hat{\omega}$ is the maximum frequency of one of the "slower" components, then the most efficient timstep for that component is $\hat{h} = h\omega/\hat{\omega}$.

In addition to symplectic integration, there is an active and quickly blooming body of literature on exactly conserving algorithms. In general, there does not exist an integrator that is both symplectic and exactly energy conserving for Hamiltonian systems [69, 71]. However, since energy is not the only quantity conserved in most systems, there is still the possibility of building a symplectic integrator that conserves other quantities, although there are limits to what the integrator can conserve in comparison to the real system [25]. Furthermore, if the system is not Hamiltonian, then a symplectic integrator provides no added benefit, whereas a conserving integrator may still be of great value. One must be careful, however, how one enforces energy conservation. It is plausible that a method which naïvely enforces global energy conservation may introduce spurious mechanisms of energy *transport* within the system [69].

Shadwick, Bowman, and Morrison [69] introduce a non-traditional integration scheme which is explicit, and exactly conserving of all quantities one is willing to program it to conserve. The drawback is that the method is not general: one needs to derive a complicated set of explicit formulae for each new integrator-problem pair. One applies the integration method to one's problem *on paper*, and then derives exactly how the conserved quantities of interest are not conserved. Then one simply adds terms to the integration routine that compensate for the non-conservation. If the standard version of the integrator is explicit, then so is the modified version. Although it is a very tedious process, it seems to have impressive results. The authors present 3 examples: the "three-wave" problem, the Lotka-Volterra predator-prey problem, and the Kepler problem, each demonstrated with both Euler's method and a 2nd order predictor-corrector. In the Kepler problem they force conservation of energy, angular momentum, and an obscure quantity known as the Runge-Lenz vector. They note that non-conservation of phase-space volume in Hamiltonian systems can be used as a measure of error, since one can no longer use conservation of other quantities as a measure of error. The authors, however, do not clearly state the limitations of their method. For example, it is not clear precisely why one couldn't use their method to force conservation of phase space volume as well as energy, which we know to be impossible [71].

We will close our discussion of integrators with a remark by Sanz-Serna [66, p.278] (d is the dimension of configuration space):

> Conservation of energy restricts the dynamics of the numerical solution by forcing the computed points to be on the correct $(2d-1)$-dimensional manifold $H =$ constant, but otherwise poses no restriction to the dynamics: within the manifold the points are free to move anywhere and only motions orthogonal to the manifold are forbidden. When d is large this is clearly a rather weak restriction. On the other hand, symplecticness restricts the dynamics in a more global way: all directions in phase space are taken into account.

Athough I think he makes an extremely interesting point, I note that his remark does not address the possibility that Shadwick *et al.* [69] have demonstrated: that energy is not the only quantity that can be conserved. Thus, it is this author's opinion that much interesting work remains in the area of integration of conservative systems.

4 Miscellaneous comments concerning dynamical systems

Szebehely [79] quotes Poincaré's definition of a non-integrable system as one in which "no analytical and globally valid integrals (invariant relations between variables) exist besides energy, momentum, and center of mass." He notes several formal ways of recognizing a chaotic system. Using frequency analysis, for example, chaotic motion is associated with a broad spectrum while periodic motion is associated with peaks. Poincaré's surface-of-section method turns periodic orbits into points, quasi-periodic orbits into lines, and chaotic motion into an irregular distribution of points.

Skeel [70] notes that a probabilistic formulation of chaos in the scope of molecular dynamics spreads an initially small ensemble into a uniform distribution via mixing. He cautions that the backwards error idea of interpreting numerical trajectories as being exact solutions to a nearby problem is not foolproof: first, for backward error to be a reasonable interpretation, it is necessary to ensure that numerical errors are much smaller than modelling errors. Second, backward error analysis should never be offered as more than a partial justification: it is still necessary to show, for example, that errors in the force field do not seriously effect the quantities of interest.

A very interesting, but overly pessimistic, paper is one by Adams *et al.* [2]. They note that a numerical solution to an ODE *always* remains close to a true solution of the problem for *some* length of time. The interesting quantity is the typical length scale of each of these sections of the numerical solution. This is similar to the idea of *shadow segments* discussed by Hayes [33]. The points at which the numerical solution switches which true solution it is following they define as a point of *diversion*. This is similar to the idea of *glitches* in the shadowing literature, and an area where a diversion/glitch is likely to occur they call a *perturbation sensitive* neighborhood. The phenomenon of numerical solutions encountering glitches/diversions is termed *computational chaos*. They argue that diversions may be detectable if some first integral is violated at that point; however, as was shown in the errata to Hayes [33], glitches in N-body solutions, for example, do not appear to be detectable via violations of energy conservation alone.

One very interesting observation made by Adams *et al.* concerns integration of the Lorenz equations using a "classical Runge-Kutta method" (they do not specify exactly which one). Using enclosure techniques, they prove the existence of a periodic orbit, and then study the occurrence of diversions when using the RK method to integrate the orbit. They find that the orbit is periodic to graphical accuracy for all time when using a timestep of $h = 1/256$, and also for $h = 1/64$. However, for $h = 1/128$ they find that the orbit diverges drastically from the periodic orbit. "This unexpected non-monotonic dependence of the accuracy of the solution on h," they emphasize, "is *unpredictable*". They discuss several other examples of unpredictable behaviour of a numerical solution in comparison with the solution produced by enclosure methods. Based on these results, they question whether true solutions of the Lorenz equations, or even of celestial mechanical systems that are generally believed to be chaotic [81, 78, 52], are in fact chaotic. They further use this result to claim that shadowing results for ODEs are practically useless since, they (erroneously) claim, there is no iron-clad guarantee of a bound on local error. This is simply not true for systems for which the right hand side of the ODE is sufficiently smooth [30]. For example, Grebogi *et al.* [28] build a shadow for a numerical solution of the forced, damped pendulum equation, which has rigorous, explicit bounds on truncation error. Finally, Adams *et al.* come to the astonishingly pessimistic (and wrong) conclusions that

(I) Concerning quantitatively reliable information on individual true solutions y^* in a "chaotic set", enclosure methods are the only practically available approach, unless a first integral of the ODEs is known; and

(II) An unknown but presumably large portion of the published results on "chaotic sets of solutions to ODEs" is more concerned with computational chaos than with ODE-chaos.

Although their results with the Lorenz attractor are interesting and worthy of further study, their conclusions are misguided.

Corless [19] argues that perhaps too much emphasis is put on exponential magnification of numerical errors. He bases this conclusion on the observation that even real systems have perturbations that, for practical and even fundamental purposes, we necessarily ignore[19, p.32]:

> Mathematical modelling of real phenomena *always* requires approximation and neglect of small effects. One neglects, for example, the effect of the gravitational attraction of Jupiter on one's earthbound experiment.... Similarly, one ignores 'small' stochastic terms in ordinary differential equation models of many phenomena, or 'small' non-autonomous perturbations of physics experiments (such as the effect of passing trucks). So a numerical analysis of methods of solving ODEs which puts truncation and roundoff errors on the same basis as modelling, measurement, and data errors would be a completely successful analysis.... We [have to] study the effects of perturbations, of course, but we have to do this even if we know the exact solution of the specified problem.

He does note, however, that there are sometimes fundamental problems when trying to imbed a discrete solution into a continuous system. For example, Euler's method applied to $y' = y^2$ produces $u \to u + hu^2$, which has derivative 0 at $u = -1/2h$. Thus the map is not a diffeomorphism, whereas the h-flow of any continuous dynamical system is a diffeomorphism.

5 Shadowing

5.1 Introduction

Dynamical systems often display sensitive dependence on initial conditions: small changes in the conditions at any point in the orbit produce solutions that exponentially diverge from each other, leading to a vastly different solution a short time later. Since a numerical method introduces small perturbations like roundoff and truncation error, we must naturally ask what effect these errors have on the validity of numerical solutions.

A true trajectory of a dynamical system lying near an approximate trajectory is called a *shadow*. A pleasant introduction to the shadowing idea is provided by Sanz-Serna and Larsson [67]. Although there exist other methods of global error analysis, shadowing is unique in that it demands that the equation governing the ODE remain fixed, while the initial conditions are perturbed in an effort to find a nearby true solution that remains close to the computed solution. The measures of error for shadowing are the distance between the numerical trajectory and the shadow, and the length of time the two remain near to each other. This is different from more common forms of backward error analysis for ODEs (see, *eg.,* [19]) and also the argument used in favour of symplectic integrators, in which it is shown that the numerical solution lies along the exact solution of an ODE system whose governing equation has been slightly perturbed. Although the latter approach is useful in some contexts, it is not clear that changing the governing equation is always the best way to measure error, especially if the governing equation satisfies special conditions that, if violated, could produce qualitatively different solutions. For example, a Hamiltonian system may not remain Hamiltonian if small, arbitrary, time-varying perturbations are allowed in its governing equation. Most practitioners of physical simulation are optimistic that their results are at least statistically valid, but "even optimists need criteria to decide how accurate the integration has to be to yield valid statistical results." [62]

We start with some definitions. The terms *orbit*, *trajectory*, and *solution* are used interchangeably. We assume a well-scaled problem where all macroscopic quantities of interest are of order unity. Throughout this section, $|\cdot|$ denotes the Euclidian norm, and $||\cdot||$ denotes the max norm, although any other norm should give similar results.

Definition. A *true trajectory* $\{\mathbf{x}_i\}_{i=a}^{b}$ of f satisfies $\mathbf{x}_{i+1} = f(\mathbf{x}_i)$ for $a \leq i < b$.

We are interested in the case that a and b are finite integers. For a discrete map, f may be a simple equation, such as the logistic equation $f(x) = 1 - 2x^2$, which maps the interval $[-1, 1]$ onto itself. For an ODE system like the N-body problem, $f(\mathbf{x})$ represents the one-timestep flow through $\mathbf{x}$.

Definition. $\{\mathbf{p}_i\}_{i=a}^{b}$ is a δ_f*-pseudo-trajectory*, also called a *noisy orbit*, for f if $||\mathbf{p}_{i+1} - f(\mathbf{p}_i)|| < \delta_f$ for $a \leq i < b$, where δ_f is the noise amplitude.

For a discrete map, δ_f is often the machine epsilon; for an ODE system that has position $\mathbf{p}_i$ at time t_i, it is the error per step.

Definition of shadowing. The true trajectory $\{\mathbf{x}_i\}_{i=a}^{b}$ $\delta_{\mathbf{x}}$*-shadows* $\{\mathbf{p}_i\}_{i=a}^{b}$ if $||\mathbf{x}_i - \mathbf{p}_i|| < \delta_{\mathbf{x}}$ for $a \leq i \leq b$.

Definition. For $a \leq i < b$, the *1-step error* made between step i and step $i+1$ of the pseudo-trajectory $\{\mathbf{p}_i\}_{i=a}^{b}$ is $\mathbf{e}_{i+1} = \mathbf{p}_{i+1} - f(\mathbf{p}_i)$.

Thus, a true trajectory is one whose 1-step errors are identically zero, and a δ_f-pseudo-trajectory is one whose 1-step errors satisfy $||\mathbf{e}_i|| < \delta_f$ for $a < i \leq b$.

Definition. The pseudo-trajectory $\{\mathbf{p}_i\}_{i=a}^{b}$ has a *glitch* at point $i = G_0 < b$ if for some relevant $\delta_{\mathbf{x}}$ there exists a true trajectory that $\delta_{\mathbf{x}}$-shadows $\{\mathbf{p}_i\}_{i=a}^{G_0}$, but no true trajectory that $\delta_{\mathbf{x}}$-shadows $\{\mathbf{p}_i\}_{i=a}^{G}$ for $G > G_0$.

As noted by Quinlan and Tremaine [62], "The failure [to find a shadow] does not [seem to] arise from a gradual increase in distance over time. Instead, there are isolated regions of a few iterations ("glitches") over which no true orbit can be found that shadows the noisy orbit."

Definition. Refinement is a process, similar to Newton's Method, that iteratively takes a noisy orbit and tries to produce a nearby orbit with less noise, *i.e.,* one with smaller 1-step errors. A refinement iteration is *successful* if before the iteration the trajectory has noise δ_f^0, after the iteration it has noise δ_f^1, and

$$\delta_f^1 < \mu\delta_f^0 \text{ for some reasonable } \mu \in (0,1). \tag{5}$$

Otherwise the refinement iteration is *unsuccessful.*

Shadowing was first discussed by Anosov [3] and Bowen [10], in relation to *hyperbolic* systems. In a 2-dimensional hyperbolic system, there are two special directions called the *unstable* (or *expanding*) and the *stable* (or *contracting*) directions, which may vary with time and generally are not orthogonal. Small perturbations along the stable direction decay exponentially in forward time, while small perturbations in the unstable direction grow exponentially in forward time. The two directions reverse rôles in backward time. In addition, the angle between the stable and unstable directions is uniformly bounded away from 0. In such a system, Anosov and Bowen proved that $\forall\varepsilon > 0\ \exists\ \delta > 0$: if $\mathbf{X}$ is a δ-pseudo orbit of arbitrary length, then there exists a true orbit that stays within ε of $\mathbf{X}$.

5.2 Tutorial

Before going into detail about shadowing algorithms and lemmas, it will be instructive to highlight how shadowing differs from other methods of error analysis, and to give some simple examples.

We look at error analysis techniques for the numerical solution of

$$y'(t) = f(t,y).$$

Let $y(t)$ be the true solution, and $\tilde{y}(t)$ be a suitably smooth and accurate interpolant of the (discrete) computed solution.

Forward error analysis is the most stringent. It states that the numerical solution remains close to the true solution with the same initial conditions:

$$\tilde{y}(0) = y(0), \quad |\tilde{y}(t) - y(t)| < \varepsilon.$$

Backward error analysis states that the numerical solution lies on the solution to a slightly perturbed problem. For example, *Defect control* is a method of backward error analysis demanding that the derivative of the numerical solution remain close to the right hand side computed along the numerical solution:

$$\tilde{y}(0) = y(0), \quad |\tilde{y}'(t) - f(t,\tilde{y})| < \varepsilon.$$

This is also called the method of *modified equations*, because it tries to find a slightly different right hand side whose true solution stays closer to the numerical solution than the true solution of the original equation. Finally, shadowing is another form of backward error analysis that says that we remain close to some true solution $\hat{y}$ of the true ODE, but that $\hat{y}$ has slightly different initial conditions:

$$\tilde{y}(0) = y(0) \neq \hat{y}(0), \quad \hat{y}'(t) = f(t,\hat{y}), \quad |\tilde{y}(t) - \hat{y}(t)| < \varepsilon, \quad 0 \le t \le T_s,$$

where T_s is the shadow length.

A simple example of a shadowable system is provided by Quinlan and Tremaine [63]. Let $x'' = x$, *i.e.,* $x' = v$, $v' = x$ in phase space, and let $x(-\infty) = v(-\infty) = 0$. The true solution is $x = v = 0\ \forall t$. Now introduce "noise" $v = \varepsilon$ at $t = 0$. The noisy solution becomes

$$x(t) = \begin{cases} 0, & t < 0 \\ \varepsilon\frac{e^t - e^{-t}}{2}, & t \ge 0 \end{cases}$$

A shadow of this noisy solution is

$$y(t) = \varepsilon e^t/2,$$

which remains within $\varepsilon/\sqrt{2}$ (in phase space) from $x(t)$ for all t.

Next we offer a taste of the proof of an almost "trivial" theorem: if a 1D map is contracting, then noisy orbits are shadowed. First we prove a lemma.

LEMMA [1D Fixed Point Lemma]: Let $f : [0,1] \to [0,1]$ be a continuous map. Then there exists a fixed point $x_* \in [0,1]$ of f, *i.e.,* $f(x_*) = x_*$.

PROOF [44]: If $f(0) = 0$ or $f(1) = 1$ we are done. Thus, assume $0 < f(0)$ and $f(1) < 1$. Consider $g(x) = x - f(x)$. Clearly $g(0) = 0 - f(0) < 0$ and $g(1) = 1 - f(1) > 0$. Since $g(x)$ is continuous, it must cross 0 as x increases from 0 to 1. Hence, there is some x_* in $[0,1]$ for which $g(x_*) = 0 = x_* - f(x_*)$, so $f(x_*) = x_*$. $\square$

A slightly more complex argument is required to prove the existence of a fixed point for an arbitrary interval X. For our purposes we will assume the above lemma holds for an arbitrary closed finite interval, not just $[0,1]$.

THEOREM [1D Contracting Map Shadowing Theorem]: Let X be a continuous 1-dimensional interval and let $f : X \to X$ be a continuous, uniformly contracting scalar map[3], *i.e.,* $\forall\, x,\ x+\delta \in X,\ |f(x+\delta) - f(x)| \le \rho|\delta|,\ \rho < 1$. Note that if f' exists and is continuous then we can take $\rho = \sup_{x\in X} |f'(x)|$. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that every δ-pseudo orbit is ε-shadowed.

OUTLINE OF PROOF: We demonstrate an asymptotically necessary condition for $\delta(\varepsilon)$, but do not prove sufficiency. By the 1D fixed point lemma, there exists a fixed point $x_* = f(x_*) \in X$. Furthermore, since f is a stable (*i.e.,* contracting) map, x_* is a *stable* fixed point, which all true orbits asymptotically approach. To devise the necessary condition on δ, assume pessimistically that after every iteration, the noisy orbit jumps by δ in the direction away from x_*. Assume, without loss of generality, that f is increasing in an ε-neighborhood of x_*, so that $f(x_* + \varepsilon) > f(x_*)$. Then, in order to stay within ε of x_*, we require that a point on the noisy orbit starting ε away from x_* is mapped to a new point which is also no more than ε away from x_*:

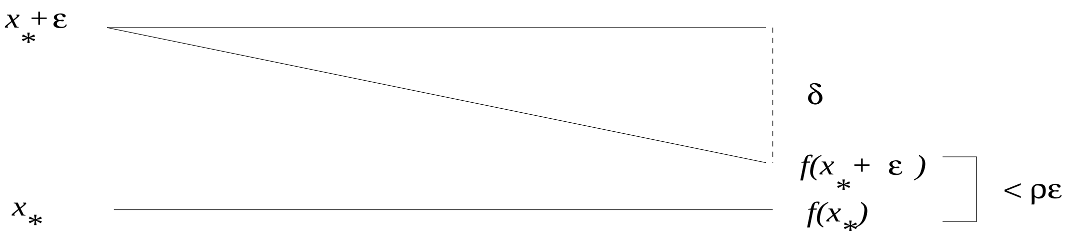

Figure 1: FIGURE FOR 1D CONTRACTING MAP SHADOWING THEOREM

Thus,

$$(f(x_* + \varepsilon) - f(x_*)) + \delta \le \rho\varepsilon + \delta < \varepsilon \implies \delta < \varepsilon(1-\rho). \quad \square$$

Notice that the closer $\rho \sim f'$ is to zero, the more stable f is, so that it can accomodate a larger noise amplitude δ.

If f were instead uniformly expanding, then we would expect our numerical solution to be exponentially pushed away from the true solution. This is true, but note that if f is expanding, then f^{-1} is contracting. In this case, we can apply a theorem similar to the above in *reverse* time.

A simple example of a system which is *not* shadowable, at least by the definitions we've seen so far, was provided by Yorke [personal communication]. Let $x'' = 0$, the solution of which is straight-line motion

[3] *Remark*: In one dimension, this is synonymous with uniformly hyperbolic, except there are no expanding directions.

at constant velocity v_0 (assume $v_0 = x'(0) \neq 0$). This motion is about as regular, and non-hyperbolic, as one could possibly imagine. Now, add noise of size ε in x' at $t = 0$. Since all true solutions have constant velocity, but our noisy solution has velocity $x' = v_0,\ \ t < 0$, and a *different* velocity $x' = v_0 + \varepsilon,\ \ t \geq 0$, it is easy to see that any true solution will diverge linearly away from the noisy solution inside at least one of the intervals $(-\infty, 0],\ \ [0, \infty)$. Thus, no true solution exists that remains close to the noisy solution for both $t \leq 0$ and $t \geq 0$.

5.3 Shadowing non-hyperbolic systems

For non-hyperbolic systems, we may have to be satisfied with finite-length shadows. The first studies of shadows for non-hyperbolic systems appear to be Beyn [8] and Hammel *et al.* [31]. Grebogi, Hammel, Yorke, and Sauer [32, 28] provide the first rigorous proof of the existence of a shadow for a non-hyperbolic system over a non-trivial length of time. For their method to work, the system does not need to be uniformly hyperbolic, but only "strongly" hyperbolic. Their method consists of two parts. First, they *refine* a noisy trajectory using an iterative method that produces a nearby trajectory with less noise. This procedure will be discussed in detail below. When refinement converges to the point that the noise is of order the machine precision, they invoke *containment*, which can rigorously prove the existence of a nearby true trajectory.

The containment process in two dimensions consists of building a parallelogram M_i around each point $\mathbf{p}_i$ of the refined numerical trajectory such that two sides $E_i^{1,2}$ are parallel to the expanding direction, while the other two sides $C_i^{1,2}$ are parallel to the contracting direction. In order to prove the existence of a shadow, the image under the map f of M_i must map over M_{i+1} such that $f(M_i)$ makes a "plus sign" with M_{i+1} (Figure 2):

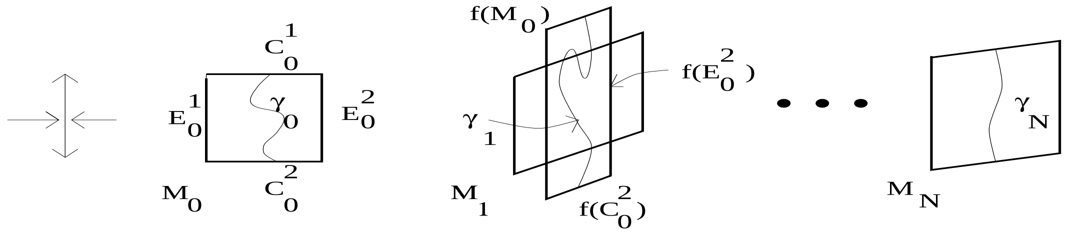

Figure 2: CONTAINMENT.

$$f(E_i^k) \cap C_{i+1}^j \neq \emptyset,\ \ j,k = 1,2, \quad \text{and} \quad f(C_i^j) \cap E_{i+1}^k = \emptyset,\ \ j,k = 1,2. \tag{6}$$

To ensure this occurs, a bound on the second derivative of f is needed, and the expansion and contraction amounts need to be resolvable by the machine precision. The proof of the existence of a true orbit then relies on the following argument. Draw a curve γ_0 in M_0 from one contracting edge to the other, *i.e.,* roughly parallel with the expanding direction. Its image $f(\gamma_0)$ is then stretched such that there is a section of $f(\gamma_0)$ lying wholly within M_1, and in particular $f(\gamma_0)$ leaves M_1 through the contracting sides at both ends. Let γ_1 be the section of $f(\gamma_0)$ lying wholly within M_1. Now look at $f(\gamma_1)$ in M_2. Repeat this process along the orbit, producing γ_N lying wholly within the final parallelogram M_N. Then any point lying along γ_N, traced backwards, represents a true trajectory that stays within $M_i,\ \ i = N, N-1, \ldots, 1, 0$, and we are done.

With this picture, there is a nice geometric interpretation of the requirement that the angle between the stable and unstable directions be bounded away from 0: if the angle gets too small, then the parallelogram essentially loses a dimension, and $f(M_i)$ doesn't make a "plus sign" with M_{i+1}. Practically speaking, this occurs when the angle becomes comparable with the noise amplitude of the refined orbit. Hence, the more accurate the orbit, the longer it can be shadowed.

5.4 Refinement of a numerical trajectory

The refinement procedure of Grebogi *et al.* ([28], hereinafter referred to as GHYS) is analagous to Newton's method for finding a zero of a function. GHYS presented their algorithm for the two dimensional case. We will also be looking closely at a generalization by Quinlan and Tremaine ([62, 63], with [63] hereinafter referred to as QT) to arbitrary-dimensional Hamiltonian systems.

The basic idea is as follows. Let $\mathbf{P} = \{\mathbf{p}_i\}_{i=0}^{S}$ be a trajectory with S steps that has noise $\delta > \eta\varepsilon_{mach}$, where ε_{mach} is the machine precision, and η is some constant significantly greater than 1 that allows room for improvement towards the machine precision. Let $\mathbf{e}_{i+1} = \mathbf{p}_{i+1} - f(\mathbf{p}_i)$ be the 1-step error at step $i+1$, where $|\mathbf{e}_{i+1}| < \delta$ for all i. The set of all 1-step errors is represented by $\mathbf{E} = \{\mathbf{e}_i\}_{i=1}^{S}$, and is estimated by a numerical integration technique that has higher accuracy than that used to compute $\mathbf{P}$. This describes a function, which we call g, which takes as input the entire orbit $\mathbf{P}$ and whose output is the set of 1-step errors $\mathbf{E}$, *i.e.,* $g(\mathbf{P}) = \mathbf{E}$. Since the 1-step errors are assumed to be small, $||\mathbf{E}||$ is small. That is, $\mathbf{P}$ is close to a zero of g, if one exists. A zero of the function would represent an orbit in which the 1-step errors are identically zero, *i.e.,* a true orbit. This is an ideal situation in which to run Newton's method. If the method converges, then a true orbit lies nearby.

Assume we have a noisy D-dimensional orbit $\mathbf{P} = \{\mathbf{p}_i\}_{i=0}^{S}$, and it has a shadow $\{\mathbf{x}_i\}_{i=0}^{S}$. Then

$$\mathbf{x}_{i+1} = f(\mathbf{x}_i) \text{ and } \mathbf{p}_{i+1} = \tilde{f}(\mathbf{p}_i),$$

where $\tilde{f}$ is an approximation to f with a typical accuracy or *noise* of δ_f. Here, δ_f should be the accuracy typically used by practitioners of the problem being studied. Now suppose we compute the 1-step errors

$$\mathbf{e}_{i+1} = \mathbf{p}_{i+1} - f(\mathbf{p}_i),$$

although numerically we use $\hat{f}$, a better approximation to f than $\tilde{f}$. Then $\hat{\mathbf{c}}_i \equiv \mathbf{x}_i - \mathbf{p}_i$ represents a correction term that perturbs point $\mathbf{p}_i$ towards the shadow. So,

$$\hat{\mathbf{c}}_{i+1} = \mathbf{x}_{i+1} - \mathbf{p}_{i+1} = f(\mathbf{x}_i) - f(\mathbf{p}_i) - \mathbf{e}_{i+1} = \frac{\partial f(\mathbf{p}_i)}{\partial \mathbf{p}_i}\hat{\mathbf{c}}_i - \mathbf{e}_{i+1} + O(|\hat{\mathbf{c}}_i|^2) \tag{7}$$

In the spirit of Newton's method, we ignore the $O(|\hat{\mathbf{c}}_i|^2)$ in (7), and so one refinement iteration defines the corrections along the entire orbit:

$$\mathbf{c}_{i+1} = L_i\mathbf{c}_i - \mathbf{e}_{i+1}, \tag{8}$$

where $L_i \equiv \frac{\partial f(\mathbf{p}_i)}{\partial \mathbf{p}_i}$ is the linearized perturbation map. For a discrete map, L_i is just the Jacobian of the map at step i. For a system of ODEs, L_i is the Jacobian of the integral of the ODE from step i to step $i+1$.[4] The computation of the linear maps L_i, called *resolvents* for an ODE system, takes most of the CPU time during a refinement. If we let D be the dimension of the problem, then the resolvent has $O(D^2)$ terms in it, and it takes $O(D^3)$ time to compute. Hayes and Jackson [33, 34] list several optimizations to the procedure

[4] In other words, let

$$\dot{\mathbf{p}} = \mathbf{h}(t, \mathbf{p}) \tag{9}$$

be the first-order ODE. Note that $\mathbf{p}_{i+1} = f(\mathbf{p}_i)$ is the solution of (9) using $\mathbf{p}_i$ as the initial condition and integrating $\mathbf{h}$ to time $i+1$. Then $J = \frac{\partial \mathbf{h}}{\partial \mathbf{p}}$ is the Jacobian of $\mathbf{h}(t, \mathbf{p})$. The Jacobian measures how $\dot{\mathbf{p}}$ changes if $\mathbf{p}$ is changed by a small amount. The *resolvent* $R(t_{i+1}, t_i)$ is the integral of $J(t)$ along the path $\mathbf{p}(t)$, and describes how a small perturbation $\delta\mathbf{p}$ from $\mathbf{p}_i$ at time t_i gets mapped to a perturbation from $\mathbf{p}_{i+1}$ at time t_{i+1}. That is, $R(t_{i+1}, t_i)$ is the solution of the *variational equation*

$$\frac{\partial R}{\partial t} = J(t, \mathbf{p}(t))R(t, t_i), \quad R(t_i, t_i) = I,$$

where I is the identity matrix. The reason the arguments to R seem reversed is for notational convenience: they satisfy the identity $R(t_2, t_0) = R(t_2, t_1)R(t_1, t_0)$, and so a perturbation $\delta\mathbf{p}$ at time t_0 gets mapped to a perturbation at time t_2 by the matrix-matrix and matrix-vector multiplication $R_2\delta\mathbf{p} = R_1R_0\delta\mathbf{p}$ [30]. Finally, the linear map in the GHYS refinement procedure is $L_i = R(t_{i+1}, t_i)$.

that speed the process considerably, and should be considered if one is interested in using the algorithm in higher dimensions. Presumably, if one is interested in studying simpler high-dimensional systems, a chaotic map would be a better choice than an ODE system, because no variational equation integration is needed.

For simplicity of explanation, we assume $D=2$ for the remainder of this subsection, deferring discussion of the higher D case to the next subsection. If the problem were not chaotic, the correction terms $\mathbf{c}_i$ could be computed directly from (8). But since L_i will amplify any errors in $\mathbf{c}_i$ not lying in the stable direction, computing the $\mathbf{c}_i$'s by iterating (8) forward will amplify errors and typically produce nothing but noise. Therefore, we split the error and correction terms into components in the stable ($\mathbf{s}_i$) and unstable ($\mathbf{u}_i$) directions at each timestep:

$$\mathbf{e}_i = e_{u_i}\mathbf{u}_i + e_{s_i}\mathbf{s}_i, \qquad \mathbf{c}_i = c_{u_i}\mathbf{u}_i + c_{s_i}\mathbf{s}_i. \tag{10}$$

Since it is not known *a priori* which direction is unstable at each timestep, the unstable vector $\mathbf{u}_0$ at time t_0 is initialized to an arbitrary unit vector. The linearized map is then iterated forward with

$$\bar{\mathbf{u}}_{i+1} = L_i\mathbf{u}_i, \qquad \mathbf{u}_{i+1} = \bar{\mathbf{u}}_{i+1}/|\bar{\mathbf{u}}_{i+1}|. \tag{11}$$

Since L_i magnifies any component that lies in the unstable direction, and assuming we are not so unlucky to choose a $\mathbf{u}_0$ that lies too close to the stable direction, then after a few e-folding times $\mathbf{u}_i$ will point roughly in the unstable direction at t_i. Similarly, the stable unit direction vectors $\mathbf{s}_i$ are computed by initializing $\mathbf{s}_S$ to an arbitrary unit vector and iterating backward,

$$\bar{\mathbf{s}}_i = L_i^{-1}\mathbf{s}_{i+1}, \qquad \mathbf{s}_i = \bar{\mathbf{s}}_i/|\bar{\mathbf{s}}_i|. \tag{12}$$

Substituting (10) into (8) yields

$$c_{u_{i+1}}\mathbf{u}_{i+1} + c_{s_{i+1}}\mathbf{s}_{i+1} = L_i(c_{u_i}\mathbf{u}_i + c_{s_i}\mathbf{s}_i) - (e_{u_{i+1}}\mathbf{u}_{i+1} + e_{s_{i+1}}\mathbf{s}_{i+1}). \tag{13}$$

While L_i magnifies errors in the unstable direction, it damps them in the stable direction. Likewise, L_i^{-1} damps errors in the unstable direction and magnifies errors in the stable direction. Thus the c_u terms should be computed backward, and the c_s terms forward. Taking components of (13) in the unstable direction at step $i+1$, iterate backward on

$$c_{u_i} = (c_{u_{i+1}} + e_{u_{i+1}})/|\bar{\mathbf{u}}_{i+1}|, \tag{14}$$

and taking components in the stable direction, iterate forward on

$$c_{s_{i+1}} = |L_i\mathbf{s}_i|c_{s_i} - e_{s_{i+1}}. \tag{15}$$

The initial choices for c_{s_0} and c_{u_S} are arbitrary as long as they are small — smaller than the maximum shadowing distance — because (15) damps initial conditions and (14) damps final conditions. GHYS and QT choose them both as 0. This choice is probably as good as any, but it can be seen here that if one shadow exists, there are infinitely many of them.[5] Another way of looking at these initial choices for c_{s_0} and c_{u_S} is that they "pinch" the growing components at the end point, and the backward-growing components at the initial point, to be small. That is, *boundary conditions* are being forced on the problem so that the exponential divergence is forcibly masked, if possible, making the solution of (8) numerically stable.

Note that, counter-intuitively, these boundary conditions demand that the initial condition for the shadow

[5] For any system, *even a chaotic one*, given any true orbit of fixed length, a small enough perturbation in the initial condition in any direction produces a small change in the final condition, although for chaotic systems this perturbation must be exponentially small in the length of the orbit. (If the perturbation is restricted to the stable subspace, then obviously a similar solution will be obtained.) Thus given any true orbit that δ-shadows a noisy orbit, there exist infinitely many true orbits nearby that also shadow it. However, it may be that all the true orbits are packed into a space unresolved by the machine precision.

and noisy orbits differ along the unstable direction. In fact, this *must* occur if the change in initial conditions is to have any effect. That is, when looking for a shadow, if perturbations were only allowed in the stable direction, those perturbations would die out, leading the "shadow" to follow the true orbit that passes through our initial conditions — the one that is already known to diverge exponentially from the computed orbit.

5.4.1 QT's generalization to arbitrary Hamiltonian systems

If the configuration space is D dimensional, then there are $2D$ dimensions in the phase space. It can be shown that in a Hamiltonian system, the number B of stable and unstable directions is equal, although $B < D$ is possible.[6] At time t_i, let $\{\mathbf{u}_i^j\}_{j=1}^D$ represent D unstable unit vectors, and let $\{\mathbf{s}_i^j\}_{j=1}^D$ represent D stable unit vectors. For any particular timestep, it will be convenient if the unstable vectors are orthonormal to each other, and the stable vectors are orthonormal to each other. However, the stable and unstable vectors together will not in general form an orthogonal system.

The vectors are evolved as in the two dimensional case, except using Gram-Schmidt orthonormalization to produce two sets of D-orthonormal vectors at each timestep. Arbitrary orthonormal bases are chosen, $\mathbf{u}_0^j$ at time t_0, and $\mathbf{s}_S^j$ at time t_S, and then evolved according to

$$\bar{\mathbf{u}}_{i+1}^j = L_i \mathbf{u}_i^j, \qquad \bar{\mathbf{s}}_i^j = L_i^{-1} \mathbf{s}_{i+1}^j.$$

At each step i, two Gram-Schmidt orthonormalizations are done: one on $\{\bar{\mathbf{u}}_{i+1}^j\}_{j=1}^D$ to produce $\{\mathbf{u}_{i+1}^j\}_{j=1}^D$, and another on $\{\bar{\mathbf{s}}_i^j\}_{j=1}^D$ to produce $\{\mathbf{s}_i^j\}_{j=1}^D$. After a few e-folding times, $\mathbf{u}_i^1$ points in the most unstable direction at step i, $\mathbf{u}_i^2$ points in the second most unstable direction, *etc.* Likewise, $\mathbf{s}_i^1$ points in the most stable direction at time i, $\mathbf{s}_i^2$ points in the second most stable direction, *etc.*

The multidimensional generalization of (10) is the obvious

$$\mathbf{e}_i = \sum_{j=1}^{D} (e_{u_i^j} \mathbf{u}_i^j + e_{s_i^j} \mathbf{s}_i^j), \qquad \mathbf{c}_i = \sum_{j=1}^{D} (c_{u_i^j} \mathbf{u}_i^j + c_{s_i^j} \mathbf{s}_i^j). \tag{16}$$

To convert the 1-step error at step i from phase-space co-ordinates $\mathbf{e}'_i = \{e_i^k\}_{k=1}^{2D}$ to the stable and unstable basis $\mathbf{e}_i = \left\{ \{e_{u_i^j}\}_{j=1}^D, \quad \{e_{s_i^j}\}_{j=1}^D \right\}$, one constructs the matrix V_i whose columns are the unstable and stable unit vectors, $V_i = (\mathbf{u}_i^1 \mathbf{u}_i^2 \cdots \mathbf{u}_i^D \mathbf{s}_i^1 \mathbf{s}_i^2 \cdots \mathbf{s}_i^D)$, and solves the system $V_i \mathbf{e}_i = \mathbf{e}'_i$.

From (8) and (16) the equation for the correction co-efficients in the unstable subspace at step $i+1$ is

$$\sum_{j=1}^{D} (c_{u_{i+1}^j} + e_{u_{i+1}^j}) \mathbf{u}_{i+1}^j = \sum_{j=1}^{D} c_{u_i^j} L_i \mathbf{u}_i^j$$

[6] It is not hard to show that the Jacobian of the N-body problem has the form

$$J = \begin{pmatrix} 0 & I \\ J_a & 0 \end{pmatrix},$$

where $J_a = \frac{\partial \mathbf{a}}{\partial \mathbf{r}}$ is symmetrical, and has real but not necessarily positive eigenvalues. If μ is an eigenvalue of J_a, then $\pm\sqrt{\mu}$ are eigenvalues of J. If μ is positive, this gives one expanding and one contracting direction; if μ is negative, it gives directions that neither expand nor contract, but "rotate" in some sense. These are the directions that are non-hyperbolic.

which are projected out along $\mathbf{u}_{i+1}^k$ producing

$$c_{u_{i+1}^k} + e_{u_{i+1}^k} = \sum_{j=k}^{D} c_{u_i^j} U_i^{kj}$$

where the scalar $U_i^{kj} = \mathbf{u}_{i+1}^k \cdot L_i \mathbf{u}_i^j = \mathbf{u}_{i+1}^k \cdot \bar{\mathbf{u}}_{i+1}^j$, and the Gram-Schmidt process ensures $U_i^{kj} = 0$ if $j < k$. As stated previously, the boundary condition requires the unstable component of the corrections to be small at timestep S, and their stable components to be small at timestep 0. For simplicity we take $\{c_{u_S^j} = 0\}_{j=1}^D$, as did QT, and the co-efficients are computed backward using

$$c_{u_i^k} = \frac{1}{U_i^{kk}} \left(c_{u_{i+1}^k} + e_{u_{i+1}^k} - \sum_{j=k+1}^{D} c_{u_i^j} U_i^{kj} \right).$$

We first solve for $\{c_{u_i^D}\}_{i=0}^S$, which does not require knowledge of the other c_u co-efficients, then solve for $\{c_{u_i^{D-1}}\}_{i=0}^S$, *etc.*

Again from (8) and (16), the equation for the correction co-efficients in the stable subspace at timestep i is

$$\sum_{j=1}^{D} (c_{s_{i+1}^j} + e_{s_{i+1}^j}) L_i^{-1} \mathbf{s}_{i+1}^j = \sum_{j=1}^{D} c_{s_i^j} \mathbf{s}_i^j,$$

which are projected out along $\mathbf{s}_i^k$ producing

$$c_{s_i^k} = \sum_{j=k}^{D} (c_{s_{i+1}^j} + e_{s_{i+1}^j}) S_i^{kj},$$

where $S_i^{kj} = \mathbf{s}_i^k \cdot L_i^{-1} \mathbf{s}_{i+1}^j = \mathbf{s}_i^k \cdot \bar{\mathbf{s}}_i^j$, and the Gram-Schmidt process ensures $S_i^{kj} = 0$ if $j < k$. The boundary condition at step 0 is $\{c_{s_0^j} = 0\}_{j=1}^D$, and the co-efficients are computed forward using

$$c_{s_{i+1}^k} = \frac{1}{S_i^{kk}} \left(c_{s_i^k} - \sum_{j=k+1}^{D} (c_{s_{i+1}^j} + e_{s_{i+1}^j}) S_i^{kj} \right) - e_{s_{i+1}^k}.$$

As with the unstable corrections, we first compute $\{c_{s_i^D}\}_{i=0}^S$ which does not require knowledge of the other c_s co-efficients, then we compute $\{c_{s_i^{D-1}}\}$, *etc.*

5.4.2 Discussion and Results of the GHYS/QT algorithm

There is no guarantee that refinement converges towards a true orbit; if there was, then all noisy orbits would be shadowable. In fact, even if some refinements are successful, numerical refinement alone does not prove rigorously that a *true* shadow exists; it only proves the existence of a numerical shadow, *i.e.,* a trajectory with less noise than the original.

QT [63] argue that if the refinement algorithm fails, there is good reason to believe that no shadow exists. They apply two arguments. First, from the more rigourous study of simpler systems, glitches are known to exist and are not just a failure of any particular refinement algorithm. Second, QT's results are consistent with a conjecture by GHYS on the frequency of glitches. However, Hayes [33] frequently saw cases in which the algorithm failed to find a shadow for noisy orbits of length S, but succeeded in finding a shadow for the superset of length $2S$. Hence, the algorithm failed to find a shadow of length S, even though one exists.

Hayes [33] has preliminary results that seem to indicate that shadow lengths scale inversely with the number of dimensions of the system, at least for the gravitational N-body problem. This would imply that shadows get shorter as the number of degrees of freedom increases, which is a very pessimistic conclusion for simulations of high-dimensional systems.

We note that the GHYS/QT refinement algorithm is trivially parallelizable, since the computation of each L_i is completely independent of all the others. For the same reason, it also has excellent locality of reference in a serial implementation, so virtual memory paging is minimized. Once the L_i's are computed, it may also be worth parallelizing the recurrence (8) [43].

5.5 Shadowing Lemmas, and theoretical results

The GHYS method for rigorously proving that a shadow exists is *constructive.* That is, one explicitly constructs a subspace around the numerical orbit which must contain a true orbit. This is essentially an interval arithmetic, or *enclosure* method (see, *eg.,* [59]). In contrast, *shadowing lemmas* are capable of proving the existence of a shadow non-constructively [14, 15, 16, 17, 18, 60, 80]. However, they always involve computing the Jacobian of the map, or solving the variational equations of the ODE, and then estimating hyperbolicity, so they are all roughly equal in computational expense.

5.5.1 Shadowing lemmas for discrete maps

Chow and Palmer [14] provide a proof of the shadowing lemma in 1 dimension when f is a (non-uniformly) expanding map. (Recall that if f is expansive in forward time, then it is contractive in reverse time. Thus we want f to be as expansive as possible when iterated forward, *i.e.,* we want f'^{-1} to be as small as possible, to facilitate shadowing.) Let $\{y_n\}_{n=0}^{N+1}$ be a pseudo-orbit of a C^2 continuous map $f:[0,1]\to[0,1]$. They define the "expansiveness" as

$$\sigma = \sup_{n=0}^{N} \sum_{m=n}^{N} |f_n'^{-1} f_{n+1}'^{-1} \cdots f_m'^{-1}|, \quad f_i = f(y_i).$$

Each $f_i'^{-1}$ measures how *contractive* the map is when applied backwards at step i, so that smaller is "better". So, each term in the sum measures how contractive the map is when applied backwards from step m down to n. Thus, the larger the sum of absolute values is, the less contractive the map is, when applied backwards from any point beyond n, down to n. Finally, the supremum measures how non-contractive the map can be anywhere. The larger σ is, the less contractive the inverse map f^{-1} is, and the more difficult it is to shadow the orbit.

Then they define a term I would call the "error diminishing factor":

$$\tau = \sup_{n=0}^{N} \left| \sum_{m=n}^{N} f_n'^{-1} f_{n+1}'^{-1} \cdots f_m'^{-1}(y_{m+1} - f(y_m)) \right|, \quad f_i = f(y_i),$$

where $y_{m+1} - f(y_m)$ is the 1-step error of step $m+1$. The description is similar to that of σ. For example, the $n=0$ term measures by how much all the 1-step errors are contracted when they are propagated backwards to step 0.

THEOREM[14]: Let $f:[0,1]\to[0,1]$ be a C^2 function with

$$M = \sup\{|f''(x)| : 0 \le x \le 1\}.$$

Let $\mathbf{y} = \{y_n\}_{n=0}^{N+1}$ be a pseudo-orbit of f such that

$$2M\sigma\tau \le 1.$$

Then there is an exact orbit $\mathbf{x} = \{x_n\}_{n=0}^N$ with

$$\frac{\tau}{1+\frac{1}{2(1+\sqrt{1-2M\sigma\tau}}} \le \sup_{n=0}^{N} |x_n - y_n| \le \varepsilon \equiv \frac{2\tau}{1+\sqrt{1-2M\sigma\tau}}.$$

OUTLINE OF PROOF: We will not discuss the lower bound except to note the interesting fact that it says that the shadow can not be arbitrarily close to the noisy orbit. It is the upper bound that concerns us.

Let the 1-step error at each step be $e_{n+1} = y_{n+1} - f(y_n)$ and let z_n be the offset from the noisy orbit to the true orbit we will find, $x_n = y_n + z_n$. The $\mathbf{z} = \{z_n\}_{n=0}^N$ are analogous to the correction terms c_i of the GHYS/QT refinement algorithm defined in equation (7) on page 14. Then we find that the z_n need to satisfy

$$z_{n+1} = f'(y_n)z_n - e_{i+1} + g_n(z_n),$$

where

$$g_n(z) = f(y_n + z) - f(y_n) - f'(y_n)z.$$

g_n measures the quality of the 1st order Taylor approximation to $f(y_n + z)$, so $g_n(z) \in O(z^2)$ for small z. It comprises the terms we ignored in equation (7) that go to zero as y_n approaches a true orbit.

Denote by S the set of sequences $\mathbf{z} = \{z_n\}_{n=0}^N$ with $|z_n| \le \varepsilon$ for all n. The following mapping T is analogous to 1 refinement of the GHYS/QT algorithm, with the boundary condition $z_{N+1} = 0$. However, it includes, via $\{g_n\}_{n=0}^N$, the second-and-higher order terms ignored by GHYS/QT, so that if one could actually construct T, it would find the exact shadow in a single "refinement" iteration:

$$(T\mathbf{z})_n = -\sum_{m=n}^{N} f_n'^{-1} f_{n+1}'^{-1} \cdots f_m'^{-1}(g_m(z_m) - e_{m+1}).$$

Chow and Palmer go on to show that T maps S into itself by showing that $|g_n(z)| \le \frac{1}{2}M|z|^2$ and hence that $|(T\mathbf{z})_n|$ is bounded by $\tau + \frac{1}{2}M\sigma\varepsilon^2 = \varepsilon$. Finally, since T maps S into itself, they invoke Brouwer's Fixed Point Theorem to prove that there is a fixed point $\mathbf{z}^T \in S$ of T, proving that a shadow exists within S. □

Since computing σ and τ in full would be very expensive ($O(N^3)$ time), they instead compute

$$\sigma_p = \sup_{n=0}^{N} \sum_{m=n}^{\min(n+p,N)} |f_n'^{-1} \cdots f_m'^{-1}|$$

where p is a positive integer chosen as follows. For $p = 0, 1, \ldots, N$, we define

$$\mu_p = \sup_{n=0}^{N-p} |f_n'^{-1} \cdots f_{n+p}'^{-1}|.$$

Then we choose the smallest p so that $\mu_p < 1$. What μ_p says in words is, "ensure that any sequence of p iterations of f^{-1}, taken as a unit, perform a contraction". Said another way, if $\mu_p < 1$, then there does not exist an expansive sequence of backwards iterations of f lasting longer than p. Thus, all backwards sequences of iterations of f longer than p will be contracting, so that, as a whole, the trajectory is "shadowable". The larger p is, however, the less shadowable the trajectory becomes until, if $p \sim N$, there is too little hyperbolicity. Using a similar argument they define a τ_p, using the same p as μ_p above.

Finally, Chow and Palmer apply this algorithm to the quadratic map

$$f(x) = ax(1-x), \quad 1 \le a \le 4.$$

They show that with $a = 3.8$ and $y_0 = 0.6$, they can choose $p = 30$, and can shadow the orbit for $N = 426,000$ iterates, with a shadowing distance of 9.76×10^{-6}. They also perform a detailed analysis of roundoff errors to ensure that their bound is completely *rigorous*, but we do not detail their analysis here.

Hadeler [29] relates Chow & Palmer's one dimensional Shadowing Lemma to the older and more well understood work related to Newton's method and Kantorovich's theorem. Kantorovich's theorem is a statement relating G, G', G'^{-1} and an initial guess $\mathbf{x}$ to the convergence of Newton's method. Hadeler shows how to translate between Chow & Palmer's Shadowing Lemma and Kantorovich's theorem, showing the relation between the former's σ and τ quantities to quantities needed in order for Kantorovich's theorem to hold. This means that we are closer to being able to apply the vast amount of literature related to Newton's method to shadowing.

Chow and Palmer extended their analysis to maps in higher dimensions [15]. Although the details are different for the multidimensional case, the basic idea is the same as the one dimensional case. Letting $x_n = y_n + z_n$ be a true orbit offset by z_n from the noisy orbit y_n, one defines an operator L that computes 1-step errors from offsets:

$$L : Z \to E,$$

where Z is the set of possible offsets and E is the set of possible 1-step errors. Note that if the space is D dimensional and there are N points in the orbit, then $Z \subseteq \Re^{ND}$, but $E \subseteq \Re^{(N-1)D}$, so L is many-to-one. (This is not surprising: there are many different noisy orbits that can possess the same set of 1-step errors.) Finally, our task is to solve $L\mathbf{z} = \mathbf{0}$. However, since L is many-to-one, L^{-1} is not unique and our task is to choose the L^{-1} with smallest norm, so that $||\mathbf{z}||$ is minimized. Chow and Palmer show that if the map has strong hyperbolicity, then the L^{-1} with smallest norm is the one that we would expect if we understand the GHYS/QT algorithm: after decomposing space into the stable and unstable subspaces, z_0 has zero component in the stable subspace, while z_N has zero component in the unstable subspace. They demonstrate their method on the Hénon map, and show that for a noisy orbit in double precision with 1-step errors of $2^{-54} \approx 10^{-15.5}$, they can use $p = 40$ and shadow $N = 333,000$ iterates with an $||L^{-1}|| = 113277 \approx 10^5$. This means that the shadowing distance is about 10^5 times the size of the 1-step errors, giving a shadow distance of about $10^{-10.5}$.

This "magnification factor", the ratio between shadow distance and local error, is termed the "brittleness" of an orbit by Dawson *et al.* [20]. They show that if the number of positive and negative Lyapunov exponents changes, or if a particular one fluctuates about zero, then this magnification factor can blow up. If the brittleness is of order the inverse of the machine epsilon, then all accuracy is lost as the shadowing error is comparable to the size of the variables themselves. The reason that a fluctuating exponent is bad is related in the following diagram, where δ is the local error, and the vertical direction is initially contracting, but then becomes expanding (Figure 3). As we see, (a) an ensemble of trajectories that starts off in an ε-ball is first

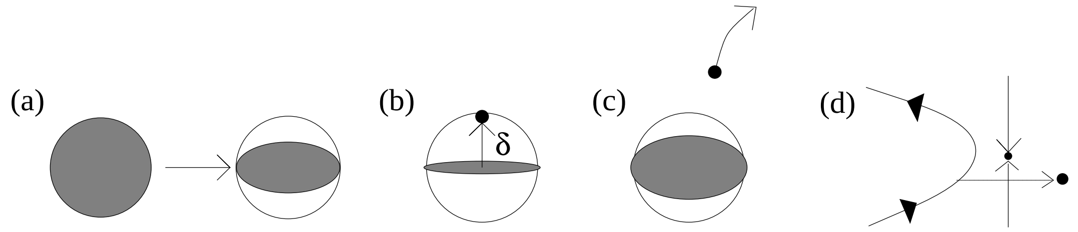

Figure 3: Fluctuating Lyapunov exponent in the "vertical" direction.

compressed into a sheet. (b) If the local error steps outside this sheet, and then the direction becomes an expanding direction, then (c) the numerical trajectory diverges away from all true trajectories that started in the original ε-ball, (d) possibly entering regions of phase space with qualitatively different behaviour than the true trajectories.

However, Dawson *et al.* make the strong claim that they believe this kind of fluctuation of Lyapunov exponents is "common" in high dimensional systems, with the only justification apparently being that there are so many dimensions, there must be a fluctuating exponent *somewhere*. It is unclear if this argument applies to all systems. For example, Hamiltonian systems always have equal numbers of positive and negative exponents.

5.5.2 Shadowing lemmas for ODE systems

There is a fundamental difference between a discrete map and a discrete solution to an ODE. Local errors of the former are restricted to being "space like" — there is no concept of the passage of time between iterations of the map. The latter, however, can have errors in space *as well as time*. That is, the numerical error in the length of each timestep can accumulate, leading the numerical solution to have a slightly different time scale than the real system.

In the integration of periodic or almost periodic systems, like the solar system, this is also known as phase error, because the numerical solution may have a slightly different period than the true solution. Thus, although the "orbit" may be reproduced correctly by the numerical solution, the time at which the real particle and simulated particle pass a certain point may differ. This is the case even if the integrator is symplectic [76, 26].

Thus, when attempting to shadow a numerical solution, it may be necessary to "rescale" time [16, 17, 18, 80]. The definition of a shadow of an ODE system must then be modified:

Definition of ODE shadowing: An approximate trajectory $\mathbf{y} = \{y_n\}_0^N$ with timesteps $\{h_n\}_0^{N-1}$ is ε-shadowed by a true solution if there exists a sequence of points $\mathbf{x} = \{x_n\}_0^N$ with timesteps $\{\tau_n\}_0^{N-1}$ such that $x_{n+1} = \phi_{\tau_n}(x_n)$ where $\phi_{\Delta t}$ is the Δt-flow of the system, and $|y_n - x_n| < \varepsilon$ and $|\tau_n - h_n| < \varepsilon$.[7]

In other words, the numerical solution is shadowed if it closely follows the *path* of a true solution, but at time t it is allowed to be up to εt time units ahead of, or behind, the true solution. This linear growth of errors is due to a lack of hyperbolicity in the direction of the flow in phase space [80]. For large t, this can be a significant difference, so a shadowing algorithm which does not take the re-scaling of time into account is likely to grossly underestimate the length of the shadow. Coomes, Koçak, and Palmer [17, 18] dramatically demonstrate this when they show that the "map method" is capable of shadowing the Lorenz equations for only 10 time units, while their method, which rescales time, can find shadows lasting 10^5 time units — ten *thousand* times longer!

Finally, note that the non-shadowable example given in the tutorial ($x'' = 0$, page 12) *is* obviously shadowable if time is rescaled. This matches what our intuition would say: as long as we only care about qualitative properties of the solution, why should it matter that the numerical solution traverses the path at a slightly different velocity than the true solution, as long as the trajectories, as a whole, remain near to each other?

6 Two-point Boundary Value Problems for ODEs

There is an intimate link between shadowing and two-point boundary value problems (2PBVPs) for ODEs. In contrast to the initial value problem (IVP)

$$\mathbf{y}'(t) = f(t, \mathbf{y}), \ \ \mathbf{y}(a) = \alpha,$$

which specifies all conditions at one point, a 2PBVP specifies conditions at two points, generally endpoints:[8]

$$\mathbf{y}'(t) = f(t, \mathbf{y}), \ \ g(\mathbf{y}(a), \mathbf{y}(b)) = \mathbf{0}.$$

[7] We may need to normalize the latter as $|\tau_n - h_n| / \max(h_n, \tau_n)$ in case h_n or τ_n gets too small.
[8] We will ignore the rescaling of time for this section, because I don't understand enough about it yet.

There are also multi-point BVPs, which place requirements at multiple points along the interval. These can be transformed into 2PBVPs by treating each sub-interval as an independent 2PBVP and then adding continuity conditions between sub-intervals.

Let $\mathbf{x}(t)$, $t \in [a,b]$ be an approximate trajectory and let $\mathbf{y}(t)$ be a true trajectory. Let $U(t)$ and $S(t)$ be the unstable and stable subspaces along $\mathbf{x}(t)$, respectively. The boundary conditions that we apply for $\mathbf{y}$ to be a shadow of $\mathbf{x}$ (cf. eqns. (14) and (15) on page 15 and ensuing discussion) are

$$(\mathbf{y}(a) - \mathbf{x}(a)) \cap S(a) = \emptyset, \quad (\mathbf{y}(b) - \mathbf{x}(b)) \cap U(b) = \emptyset, \tag{17}$$

where '$\cap$' denotes projection into a subspace. In practice, we could relax these boundary conditions to allow $\mathbf{x}$ and $\mathbf{y}$ into the "forbidden" subspaces by an amount less than the shadowing distance. So, there exists a shadow *iff* (17) has a solution.

The concept of hyperbolicity used in shadowing is identical to the concept of *dichotomy* in BVPs [60, 16]:

Definition: For positive constants K, λ, the system

$$\partial_t \Phi(t,\tau) = f_z(z(t),t)\Phi(t,\tau), \quad \Phi(\tau,\tau) = I$$

is said to be (K,λ)-hyperbolic [16] or to have *exponential dichotomy* [5, p.116] on the interval $[a,b]$ if for a given $s \in [a,b]$ there exist linear subspaces $S(s)$ and $U(s)$ such that, if $y_0 \in S(s)$,

$$|\Phi(t,a)\Phi^{-1}(s,a)y_0| \leq Ke^{-\lambda(t-s)}|y_0|$$

for $t \geq s$ and $s,t \in [a,b]$, and, if $y_0 \in U(s)$,

$$|\Phi(t,a)\Phi^{-1}(s,a)y_0| \leq Ke^{-\lambda(s-t)}|y_0|$$

for $s \geq t$ and $s,t \in [a,b]$. Here, Φ is a solution to the variational equation of the problem, *i.e.,* a resolvent (cf. footnote on page 14).

Dichotomy is closely related to well-conditioning of 2PBVPs, in the sense that if there is a dichotomy, then it is easy to put bounds on the conditioning constants. It is related to Green's functions, in the sense that if there are bounds, exponential in time, on the Green's functions, then there is a dichotomy [5, section 3.4.3]. The splitting of the solution components into the stable and unstable subspaces is called *decoupling* in the BVP literature, and standard algorithms exist to handle decoupling.

The GHYS/QT refinement algorithm, it turns out, is an implementation of a form of *multiple shooting*, which is simply a method of solving 2PBVPs by splitting the interval into many sub-intervals, computing an approximate solution S_i on each sub-interval, and then requiring that the solutions S_i match at the boundaries. Many topics of potential interest to the study of shadowing, including multiple shooting, are discussed in Ascher *et al.* [5]. Enright [24] suggests ways to improve the performance of multiple shooting, based on clever ways of choosing the sub-interval endpoints. Krogh, Keener, and Enright [51] show how it is possible to reduce the size of the resolvent matrix since there are really only k unknowns at one end and $N-k$ at the other, instead of integrating the entire variational equation. So, in the GHYS/QT algorithm, we could reduce the size of the resolvents by half. Chow and Van Vleck [16] show explicitly how shadowing is related to multiple shooting.

The field of 2PBVPs is comparatively mature in comparison to the study of shadowing. For example, methods for *a posteriori*, rigorous error bounds on 2PBVPs existed in 1981 [50], six years before GHYS. Clearly, the study of shadowing may benefit from interaction with the study of 2PBVPs.

7 Discussion and future directions

In this author's opinion, the most exciting recent prospect is the application of rescaling of time discussed above, particularly by Coomes *et al.* [18]. The fact that they could shadow the Lorenz equations 10^4 times longer using a rescaling of time is remarkable. Miller [58] has a diagram showing the distance between two nearby solutions with 8 particles, in which 2 of the particles form a temporary binary star with a highly elliptical orbit. There are very large spikes in the separation of the two solutions, representing the times of closest approach of the two stars, because at closest approach the stars speed up immensely. Thus, the phase space separation of the two nearby solutions temporarily increases dramatically. Without a re-scaling of time, it is unlikely that a shadowing algorithm would be able to shadow such an orbit, when in fact it is not unreasonable to say that a true solution that temporarily moves away from a numerical solution in this fashion can still be labelled a shadow. It is already known that glitches in N-body simulations occur most frequently near close approaches [62, 63, 33]. It will be interesting to see how often the glitches can be eliminated by a rescaling of time. It may be related to Chow & Palmer's μ_p — perhaps p measures the number of timesteps in an encounter.

Another interesting point in regard to non-hyperbolicity is the observation that N-body systems have some purely imaginary eigenvalues (cf. footnote on page 16). This is not surprising, since there are components of regular motion in most N-body systems. It may be that these "rotating" directions can not be "pinned to zero" at the endpoints of the shadow. Perhaps it may be possible to handle this non-hyperbolic structure by allowing these "rotating" directions to have non-zero boundary conditions at the endpoints, unlike the exponential directions.

Quinlan and Tremaine [62, 63] ask the question of whether shadows, even if they exist, are typical of true orbits chosen at random. It would seem odd if they were not. Perhaps bi-shadowing may be applicable. Bi-shadowing studies not only the existence of a true solution close to a numerical solution, but the existence of possible numerical solutions close to arbitrary true solutions.

Many authors have demonstrated that the rate of exponential divergence decreases as a system becomes less collisional [47, 48, 49, 27, 63]. Hayes [33] demonstrates that a simple scaling of the N-body problem gives a shadow length inversely proportional to N, which seems rather disappointing. However, Quinlan and Tremaine, and Hayes, discuss the possibility that, even if shadowing becomes more difficult in higher dimensions, perhaps the fact that N-body systems become smoother and less collisional with higher N may compensate for this, allowing smooth systems with large numbers of particles to be shadowable.

There are many, many more interesting areas of further research. The author's Master's Thesis [33] discusses many of them. It also discusses other measures of error that may be used if shadowing turns out to be too stringent a measure of error.

Finally, there may be the possibility of proving rigorously that a shadow does *not* exist. The literature on this subject seems rather sparse; it is mentioned briefly in Adams *et al.* [2] and Iserles *et al.* [42].

Glossary

Like most groups of technical people, stellar dynamicists have a language all their own. This glossary is informal, and constitutes an exceedingly brief and fleeting crash course in stellar dynamics for those members of my committee not familiar with the terms. It is designed to be read sequentially, not alphabetically.

collision A *collision* occurs when two particles pass close enough to each other to "substantially" alter at least one of their pre-encounter paths. They don't actually have to "touch" in any physical sense. If they are part of a larger parent body (*eg.,* a galaxy) then their orbits will also be altered. The definition of "substantial" depends on context, but most real gravitational systems are either clearly *collisional* (*eg.,* globular clusters) or *collisionless* (most galaxies). No system is completely collisionless, but galactic disks, for example, have relaxation times that are orders of magnitude longer than the age of the universe.

encounter An *encounter* is a term that applies to any 2-body interaction. Although *collision* and *encounter* are often used interchangeably, a collision usually refers explicitly to a close or strong encounter.

crossing time The average time it takes a body to travel from one side of the system to the other, sometimes defined as

$$t_{cr} = \frac{v_{rms}}{R_f}$$

where R_f is the radius of a sphere that contains a fraction f (usually 1/2) of the mass of the system, and v_{rms} is the root mean square particle speed. In disk systems, it is half the average orbital period.

relaxation In a real system, each particle has a birth date (*eg.,* the formation of a star from a condensing gas cloud), at which time it enters an initial orbit. As the system ages, particles are perturbed from their initial orbits both by the slowly changing global potential, and by collisions. The process acting on individual particles is similar to a random walk; the cumulative effect on the system is called *relaxation.* A system is *relaxed* when no information remains about the initial orbits of most particles. More formally, if we look at the collisions experienced by a single particle and ϕ_c is the deflection in radians caused by each collision c, then if most particles satisfy

$$\sum_c |\phi_c| \approx\geq 1,$$

then the system is relaxed. Similar arguments can be made for the relative change in orbital energy or orbital angular momentum per collision. A slightly more detailed analysis [9, chapter 4] shows that the relaxation time is about $\frac{N}{8 \ln N}$ crossing times.

equipartition Each particle has kinetic energy $\frac{1}{2}mv^2$ (v measured relative to the centre of mass of the entire cluster of objects). When two stars of unequal mass collide, they tend to exchange energy so as to equalize their kinetic energies. (This is the same principle used by interplanetary probes for gravity assists: they gain speed, relative to the sun, by passing close to planets.) This results in lighter particles gaining energy and entering higher orbits while larger ones lose energy and sink towards the centre of the system. If we consider particles to be members of a gas, then their kinetic energy can also be interpreted as their *temperature* (nothing to do with the surface temperature of an actual star). When the temperature of each particle has become statistically independent of its mass, then the system has achieved equipartition of energy. It generally takes about a relaxation time.

cold disk A *cold* system is characterized by low RMS collision velocities (*i.e.,* low velocity dispersion), and higher temperatures mean higher velocity dispersion. A spherical system cannot be cold because every point in the system has orbits with vastly different velocity vectors passing through it. Only disks can be cold, where nearby orbits have similar orbital vectors and thus low relative speeds. The disks of spiral galaxies are generally cold and collisionless, but they slowly heat up as the system relaxes (over time periods much longer than the current age of the universe).

globular cluster A small, dense, roughly spherically symmetric group of $10^5 \sim 10^6$ old stars. There are about 150 known globular clusters orbiting our Milky Way Galaxy. Since a globular cluster is small and dense, it is highly collisional. The relaxation time of a globular cluster is roughly 3-5 billion years, and they are believed to be composed of the very first stars to form during the formation of our Galaxy about 10 billion years ago; thus existing globular clusters are relaxed (and they are observed to be quite equipartitioned).

spiral galaxy A group of roughly 10^{8-10} stars, arranged with an ellipsoidal centre of old stars and a sparser thin disk of younger stars. The core can be collisional, the disk is generally not.[9]

elliptical galaxy A voluminous group of roughly $10^{10} \sim 10^{11}$ stars. They look a lot like globular clusters but are many orders of magnitude larger, sometimes having diameters larger than spiral galaxies. They can be collisional in the core, but generally not outside of it.

Plummer model A popular model used to describe the gravitational potential (and thereby the density of particles) in spherical stellar systems. The Plummer Potential is defined by

$$\Phi_P(b) = -\frac{GM}{\sqrt{r^2 + b^2}}$$

where b is the single parameter and $\Phi_P(0)$ gives the potential of a point mass. The gravitational force between two non-overlapping "Plummer spheres" is

$$F = \frac{Gm_1 m_2}{r^2 + b^2}$$

force softening and discreteness noise Finally, we are usually restricted to simulate with an N which is many orders of magnitude smaller than the actual number of particles in the system. (For example, even a small globular cluster has 10^5 members, and it's just possible that we may be able to fully simulate such a system without softening in the next few years.) Thus, each particle in our simulation can be representative of thousands or millions of particles in the real system. Each simulated particle can be interpreted as either A) the total mass of the group it represents, at the centre of mass of the group, or B) a randomly chosen member of that group ("Monte carlo sampling"). In either case the potential energy function of the system is much more "lumpy" than it should be, and two-body interactions will be much more important than in the real system. For example, using interpretation (A), two simulated particles passing by each other with zero distance encounter infinite force; whereas in the real system, two groups of particles thrown at each other generally pass through each other almost unscathed. (Because the diameter of a star is about 10^{-9} of the average distance between them.) This results in the simulated system being much more collisional than the real system, and having a much higher relaxation rate.

A simple method to attempt to partially alleviate this problem is to "soften" the force between two simulated particles. The simplest method is to use model (A) and treat the particles as Plummer spheres, replacing Newton's gravitational force law

$$F = \frac{Gm_1 m_2}{r^2} \tag{18}$$

with

$$F = \frac{Gm_1 m_2}{r^2 + \varepsilon^2} \tag{19}$$

where ε is a constant chosen to approximate the average separation between simulated particles. Other methods (like the particle mesh methods, *eg.,* [61]) contain implicit softening. However, softening only partially alleviates this problem of *discreteness noise.* The only known way to substantially improve it is simply to increase N [39].

[9] Galactic clusters of stars and areas in the spiral arms perturbed by giant molecular clouds can have substantially higher local collision rates. Both these effects are negligible in elliptical galaxies, because ellipticals are hotter (and are thus less likely to contain clusters) and are generally cloudless.

regularization If a system is being simulated without softening, then close encounters between particles can occur, and sometimes 3- or more-body interactions can result in the formation of a *closed* binary or multiple star system. To accurately integrate these tight orbits takes time-steps many orders of magnitude smaller than is typical for non-interacting stars. Multiple star systems form an integral part of the evolution of globular clusters, and thus cannot be ignored. Without special treatment the integration of tight binary star systems can soak up the majority of your CPU time and slow the simulation immensely [1]. *Regularization* is the process of speeding up the simulation of these orbits by solving them analytically as perturbed two-body systems.

Here are some more terms that are not to do with stellar dynamics, but astronomy in general.

Dark Matter matter that we can't see, but we know must exist for various reasons, like spiral galaxies rotating so fast that, if the light matter was all that was there, the galaxy would fly apart. *Something* must be holding it together, assuming gravity is still $1/r^2$ at galaxy-sized distances.

Particle-Particle method the obvious $O(N^2)$ method for computing all the forces between all the particles of an N-body system.

Particle-Mesh Method compute a finite-term expansion of the gravitational potential over a grid (mesh) in space, then interpolate the potential at each particle. The gradient of the potential is the force on that particle.

Tree method Split the universe hierarchically up into boxes, such that any box containing more than one particle is split into 2^d equal-sized boxes, where d is the dimension, 2 or 3. Then, when computing the force on a particle, far away particles can be lumped into boxes and an approximate force from all of them can be computed by assuming all the particles are at their mutual centre of mass in the box. The time is $O(N \log N)$ to compute all the forces.

Mass segregation a result of equipartition: massive bodies sink to the bottom of a cluster over roughly a relaxation time.

Core collapse a phenomenon, independent of mass segregation (because it happens even if all the masses are the same), whereby energy is transfered out of the core of a collisional system, like a globular cluster. The core gets more and more dense, until, in the limit, it becomes infinitely dense. In reality, core collapse is halted by a mechanism I'm not familiar with, but involves the formation of binary stars in the core.

Perihelion point of closest approach of a solar system object to the Sun. Also peri-astron, closest approach to any star.

References

[1] Sverre J. Aarseth. Direct Methods for N-Body Simulations. In *Multiple Time Scales*, pages 377–418. Academic Press, Inc., 1985.

[2] E. Adams, W. F. Ames, W. Kühn, W. Rufeger, and H. Spreuer. Computational chaos may be due to a single local error. *Journal of Computational Physics*, 104:241–250, 1993.

[3] D. V. Anosov. Geodesic Flows and Closed Riemannian Manifolds with Negative Curvature. *Proc. Steklov Inst. Math*, 90:1, 1967.

[4] U. Ascher and R. D. Russel, editors. *Numerical Boundary Value ODEs*, volume 5 of *Progress in Scientific Computing*. Birkhäuser Boston, 1985.

[5] Uri M. Ascher, Robert M. M. Mattheij, and Robert D. Russel. *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations.* Prentice-Hall Series in Computational Mathematics. Prentice-Hall, 1988.

[6] Josh Barnes and Piet Hut. A hierarchical $O(N \log N)$ force-calculation algorithm. *Nature*, 324:446–449, 4 December 1986.

[7] Joshua E. Barnes and Piet Hut. Error analysis of a tree code. *Astrophysical Journal Supplement Series*, 70:389–417, 1989.

[8] W.-J. Beyn. On Invariant Closed Curves for One-Step Methods. *Numer. Math.*, 51:103–122, 1987.

[9] James Binney and Scott Tremaine. *Galactic Dynamics.* Princeton Series in Astrophysics. Princeton University Press, 1987.

[10] R. Bowen. ω-Limit Sets for Axiom A Diffeomorphisms. *Journal of Differential Equations*, 18:333, 1975.

[11] John G. Bryant. A formulation of the N-body problem where the velocities are bounded. In Roy [64], pages 363–369.

[12] John G. Bryant. A new interpretation of collisions in the N-body problem. In Roy [65], pages 501–507.

[13] P. J. Channell and C. Scovel. Symplectic integration of Hamiltonian systems. *Nonlinearity*, 3:231–259, 1990.

[14] Shui-Nee Chow and Kenneth J. Palmer. On the numerical computation of orbits of dynamical systems: the one-dimensional case. *Dynamics and Differential Equations*, 3:361–380, 1991.

[15] Shui-Nee Chow and Kenneth J. Palmer. On the numerical computation of orbits of dynamical systems: the higher dimensional case. *Journal of Complexity*, 8:398–423, 1992.

[16] Shui-Nee Chow and Erik S. Van Vleck. A shadowing lemma approach to global error analysis for initial value ODEs. *SIAM Journal on Scientific Computing*, 15(4):959–976, July 1994.

[17] Brian A. Coomes, Hüseyin Koçak, and Kenneth J. Palmer. Shadowing orbits of ordinary differential equations. *Journal of Computational and Applied Mathematics*, 52:35–43, 1994.

[18] Brian A. Coomes, Hüseyin Ko cak, and Kenneth J. Palmer. Rigorous computational shadowing of orbits of ordinary differential equations. *Numerische Mathematique*, 69:401–421, 1995.

[19] Robert M. Corless. Error backward. *Contemporary Mathematics*, 172:31–62, 1994.

[20] Silvina Dawson, Celso Grebogi, Tim Sauer, and James A. Yorke. Obstructions to Shadowing When a Lyapunov Exponent Fluctuates about Zero. *Physical Review Letters*, 73(14):1927–1930, 3 Oct 1994.

[21] Herwig Dejonghe and Piet Hut. Round-off sensitivity in the N-body problem. In Hut and McMillan [41], pages 212–218.

[22] Charles C. Dyer and Peter S. S. Ip. Softening in N-body simulations of collisionless systems. *The Astrophysical Journal*, 409:60–67, May 1993.

[23] David J. D. Earn and Scott Tremaine. Exact solutions for Hamiltonian systems using integer arithmetic. In Sundelius [77], pages 143–148.

[24] Wayne H. Enright. Improving the performance of numerical methods for two point boundary value problems. In Ascher and Russel [4], pages 107–119.

[25] Z. Ge and J. E. Marsden. Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators. *Phys. Lett. A.*, 133:134–139, 1988.

[26] O. Gonzalez and A. M. Stuart. Remarks on the qualitative properties of modified equations. Preprint, 1996.

[27] Jeremy Goodman, Douglas C. Heggie, and Piet Hut. On the exponential instability of N-body systems. *The Astrophysical Journal*, 415:715–733, 1993.

[28] Celso Grebogi, Stephen M. Hammel, James A. Yorke, and Tim Sauer. Shadowing of Physical Trajectories in Chaotic Dynamics: Containment and Refinement. *Physical Review Letters*, 65(13):1527–1530, 24 September 1990.

[29] K. P. Hadeler. Shadowing orbits and Kantorovich's theorem. *Numerische Mathematique*, 73:65–73, 1996.

[30] Ernst Hairer, Syvert Paul Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations I – Nonstiff Problems.* Springer-Verlag, 1980.

[31] Stephen M. Hammel, James A. Yorke, and Celso Grebogi. Do Numerical Orbits of Chaotic Dynamical Processes Represent True Orbits? *Journal of Complexity*, 3:136–145, 1987.

[32] Stephen M. Hammel, James A. Yorke, and Celso Grebogi. Numerical Orbits of Chaotic Dynamical Processes Represent True Orbits. *Bull. Am. Math. Soc.*, 19:465–470, 1988.

[33] Wayne Hayes. Efficient Shadowing of High Dimensional Chaotic Systems with the Large Astrophysical N-body Problem as an Example. Master's thesis, Dept. of Computer Science, University of Toronto, January 1995. Available via anonymous ftp from `ftp.cs.utoronto.ca/pub/reports/na/hayes-95-msc`.

[34] Wayne Hayes and Kenneth R. Jackson. A fast shadowing algorithm for high dimensional ODE systems. Unpublished.

[35] D. C. Heggie. Star cluster dynamics: Mathematical models. In Hut and McMillan [41], pages 13–22.

[36] Douglas C. Heggie. The n-body problem in stellar dynamics. In Roy [64], pages 329–347.

[37] Douglas C. Heggie. Chaos in the n-body problem of stellar dynamics. In Roy [65], pages 47–62.

[38] Douglas C. Heggie. Statistics of Small-N Simulations. Preprint astro-ph/9511124, November 1995.

[39] Lars Hernquist, Piet Hut, and Jun Makino. Discreteness noise versus force errors in N-body simulations. *Astrophysical Journal Letters*, 402:L85–L88, 1993.

[40] Piet Hut, Jun Makino, and Steve McMillan. Building a better leapfrog. *Astrophysical Journal Letters*, 443:L93–L96, 1995.

[41] Piet Hut and S. L. W. McMillan, editors. *The Use of Supercomputers in Stellar Dynamics.* Springer-Verlag, 1986.

[42] A. Iserles, A. T. Peplow, and A. M. Stuart. *SIAM J. Numer. Anal.*, 28:1723, 1991.

[43] K. R. Jackson and R. N. Pancer. The parallel soluton of ABD systems arising in numerical methods for BVPs for ODEs. Technical Report 255/91, Dept. of Computer Science, University of Toronto, May 1992.

[44] Ken Jackson. Personal communication.

[45] J. Garrett Jernigan and David H. Porter. A tree code with logarithmic reduction of force terms, hierarchical regularization of all variables, and explicit accuracy controls. *Astrophysical Journal Supplement Series*, 71:871–893, 1989.

[46] Henry E. Kandrup. The time scale for "mixing" in a stellar dynamical system. *Physics Letters A*, 140:97–100, September 1989.

[47] Henry E. Kandrup and Haywood Smith. On the sensitivity of the N-body problem to small changes in initial conditions. *The Astrophysical Journal*, 374:255–265, 1991.

[48] Henry E. Kandrup and Haywood Smith. On the sensitivity of the N-body problem to small changes in initial conditions. II. *The Astrophysical Journal*, 386:635–645, 1992.

[49] Henry E. Kandrup, Haywood Smith, and David Willmes. On the sensitivity of the N-body problem to small changes in initial conditions. III. *The Astrophysical Journal*, 399:627–633, 1992.

[50] Gershon Kedem. A posteriori error bounds for two-point boundary value problems. *SIAM J. Numer. Anal.*, 18(3), June 1981.

[51] Fred T. Krogh, J. P. Keener, and Wayne H. Enright. Reducing the number of variational equations in the implementation of multiple shooting. In Ascher and Russel [4], pages 121–135.

[52] Jaques Laskar. Large-scale chaos in the solar system. *Astronomy and Astrophysics*, 287L:9–12, 1994.

[53] M. Lecar. A comparison of eleven numerical integrations of the same gravitational 25-body problem. *Bull. Astron.*, 3:91, 1968.

[54] Todd R. Littell, Robert D. Skeel, and Meiqing Zhang. Error analysis of symplectic multiple time stepping. Preprint, December 1995.

[55] Junichiro Makino. Optimal order and time-step criterion for Aarseth-type N-body integrators. *The Astrophysical Journal*, 369:200–212, 1991 March 1.

[56] Jerry B. Marion and Stephen T. Thornton. *Classical Dynamics of Particles and Systems.* Harcourt Brace Jovanovich, third edition, 1988.

[57] Dimitri Mihalas and James Binney. *Galactic Astronomy*, pages 570–571. Freeman, second edition, 1981. A short note at the back of the book, "A Model of Galactic Evolution", simulations by Frank Hohl of NASA Langly Research Centre.

[58] R. H. Miller. Irreversibility in small stellar dynamical systems. *The Astrophysical Journal*, 140:250, 1964.

[59] Ned Nedialkov. Validated solutions of initial value problems for ordinary differential equations. Ph.D. Qualifying paper, University of Toronto Computer Science, June 1996.

[60] Kenneth J. Palmer. Exponential dichotomies, the shadowing lemma and transversal homoclinic points. In U. Kirchgraber and H. O. Walther, editors, *Dynamics Reported*, volume 1. Wiley and Teubner, 1988.

[61] D. Pfenniger and D. Friedli. Computational issues connected with 3D N-body simulations. *Astronomy and Astrophysics*, 270:561–572, 1993.

[62] Gerald D Quinlan and Scott Tremaine. Shadow orbits and the gravitational N-body problem. In Sundelius [77], pages 143–148.

[63] Gerald D. Quinlan and Scott Tremaine. On the reliability of gravitational N-body integrations. *Monthly Notices of the Royal Astronomical Society*, 259:505–518, 1992.

[64] A. E. Roy, editor. *Long-Term Dynamical Behaviour of Natural and Artificial N-body Systems.* Klewer Academic, 1988.

[65] A. E. Roy, editor. *Predictability, Stability, and Chaos in N-body Dynamical Systems.* Plenum, 1991.

[66] J. M. Sanz-Serna. Symplectic integrators for Hamiltonian problems: an overview. In A. Iserles, editor, *Acta Numerica 1992*, pages 243–286. Cambridge University Press, 1992.

[67] J. M. Sanz-Serna and S. Larsson. Shadows, Chaos, and Saddles. *Appld. Numer. Math.*, 13:181–190, 1993.

[68] J. A. Sellwood. Disc galaxy dynamics on the computer. In Hut and McMillan [41], pages 5–12.

[69] B. A. Shadwick, John C. Bowman, and P. J. Morrison. Exactly conservative integrators. Preprint, Department of Physics and Institute for Fusion Studies, University of Texas at Austin, August 1995.

[70] R. Skeel. Meaning of molecular dynamics. Unpublished, e-mail skeel@cs.uiuc.edu.

[71] R. D. Skeel and C. W. Gear. *Physica D*, 60:311–313, 1992.

[72] Robert D. Skeel and Jeffrey J. Biesiadecki. Symplectic integration with variable stepsize. *Annals of Numerical Mathematics*, 1:1–9, 1994.

[73] H. Smith Jr. The dependence of statistical results from N-body calculations on N. *Astronomy and Astrophysics*, 76:192–199, 1979.

[74] Haywood Smith Jr. The validity of statistical results from N-body calculations. *Astronomy and Astrophysics*, 61:305–312, 1977.

[75] Lyman Spitzer Jr. *Dynamical Evolution of Globular Clusters.* Princeton series in astrophysics. Princeton University Press, 1987.

[76] Andrew M. Stuart and O. Gonzalez. Comments on qualitative properties of integrators, 1996. Talk given at SIAM Conference in Dynamical Systems, Snowbird, Utah, May 1995. See also Gonzalez and Stuart (1996).

[77] B. Sundelius, editor. *Dynamics of Disc Galaxies.* Göteborg, Sweden, 1991.

[78] G. J. Sussman and Jack Wisdom. Chaotic evolution of the solar-system. *Science*, 257:56–62, 1992.

[79] Victor Szebehely. Chaos, stability, and predictability in Newtonian dynamics. In Roy [65], pages 63–71.

[80] Eric S. Van Vleck. Numerical shadowing near hyperbolic trajectories. *SIAM Journal on Scientific Computing*, 16(5), 1995.

[81] Jack Wisdom. Long-term evolution of the solar-system. *IAU Symposia*, pages 17–24, 1992.