Is the solar system stable? and Can we use chaos to make measurements?

Jack Wisdom
Massachusetts Institute of Technology, USA

Abstract. This talk addresses two separate questions: "Is the solar system stable?" and "Can we use chaos to make better measurements?" In the first part, a review is presented of the numerical experiments which indicate that the motion of Pluto, and indeed the whole solar system, is chaotic. The time scale for the exponential divergence of nearby trajectories is remarkably short compared to the age of the solar system. In the second part, numerical experiments are presented which indicate that the exponential sensitivity of trajectories to changes in initial conditions and parameters cannot be used to exponentially constrain initial conditions and parameters from trajectory measurements. It does appear though that parameters are better constrained by measurements of chaotic trajectories than might naively be expected.

Introduction
First, it is useful to remind ourselves of the reality of chaos, and just how much fun it is. I have a nice demonstration to show you, of a double pendulum (Fig. 1). The double pendulum is one of the simplest dynamical systems one can build after the pendulum: one pendulum supported at the end of another pendulum, constrained to move in a plane. This simple system exhibits outrageously complicated behavior. How could anyone watch the double pendulum and continue to assume that all solutions of Newton’s equations could be developed in quasiperiodic perturbation expansions? Did hundreds of years really go by without anyone looking at a dynamical system in action? The double pendulum can also be used to illustrate the divided phase-space characteristic of Hamiltonian systems (even though there is friction in the physical pendulum): trajectories with the same energy can be either chaotic or regular depending on the initial condition. Given that such a simple system as the double pendulum exhibits such complicated motion, it is hard to understand why it has taken so long for the importance of chaotic behavior to be realized. Chaos is not an irrelevant mathematical curiosity.
Fig. 1. The double pendulum provides a nice demonstration of chaotic behavior and the divided phase space in simple physical systems.

**Question 1: Is the solar system stable?**

Surely this is one of the oldest questions in modern science. As soon as Newton's equations of motion are written down one has to wonder about the long-term consequences. It has generally been assumed that our solar system is quasiperiodic and consequently stable, and that it is only a matter of time until a mathematical proof of this fact is given. Tremendous progress towards that goal has been made. Arnold (1963) has proven that solar systems are quasiperiodic in large measure provided that the masses, inclinations, and eccentricities of the planets are sufficiently small. On the other hand, we know that dynamical systems generally display chaotic behavior as well as regular behavior, and the solar system is, after all, just another dynamical system. The question of its stability should be approached with an open mind.

What we really would like to know is whether our solar system is on a chaotic or quasiperiodic trajectory. Since the physical experiment runs too slowly for us to decide of the stability of the solar system, numerical experiments have been performed by Sussman and Wisdom.

**1.1. Previous integration**

Numerical integrations require a great deal of computer time. A direct calculation, even with a powerful computer, will not follow each planet around its orbit over a long period of time. The time scale separating the two-body problem and the orbital motion of the planets is much longer. The orbital frequency of a planet around the sun is much longer than the precession time scale for the solar system. The precession time scale is long enough to allow for a couple of integrations by a computer with finite precision.

We have performed calculations (Applegate et al., 1985; 2000) that show that the solar system is chaotic. The computer program designed for this purpose is one of the most sophisticated programs for integrating the equations of motion. It can add, subtract, and multiply numbers with an arbitrary number of decimal places, and it can integrate Newton's equations of motion on a Cray supercomputer.

In our first calculation, the solar system consists of the Sun, Jupiter, Saturn, Uranus, Neptunus, Pluto, and the eight planets. We find that, after a few hundred years, the system is chaotic. This means that the system is sensitive to initial conditions, and that small changes in the initial conditions can lead to large differences in the final state of the system.

One result from this calculation is that the solar system is chaotic. This means that the system is sensitive to initial conditions, and that small changes in the initial conditions can lead to large differences in the final state of the system. This result has important implications for the stability of the solar system, and for the possibility of predicting the future evolution of the system.
too slowly for us to decide the matter, we have approached the question of the stability of the solar system through numerical experiments. Our numerical experiments indicate that, in fact, the solar system is chaotic (Sussman and Wisdom, 1988).

1.1. Previous integration

Numerical integrations of the solar system take an extraordinarily large amount of computer time. This is because there is a tremendous range of time scales. A direct calculation must take steps that are small enough to follow each planet around the sun, yet the motion of the planets is only interesting on time scales of millions of years.

This time scale separation is a consequence of the degeneracy of the two-body problem. The unperturbed two-body problem has only a single frequency, the orbital frequency, even though there are three degrees of freedom after elimination of the center of mass. The degeneracy is broken by planetary perturbations. The largest effect is to make the perihelia and orbital nodes precess. The frequency of these motions is of order the orbital frequency multiplied by the mass ratio of the perturbing planet to that of the sun. The precession time scales range from tens of thousands of years to a couple of million years. Only on a time scale longer than this precession time scale should we expect to find interesting dynamical coupling between the planets.

We have performed our numerical integrations on the Digital Orrery (Applegate et al., 1985). Named after the orreries of the 18th and 19th centuries (Fig. 2), the Digital Orrery (Fig. 3) is a special purpose computer designed specifically for solar system dynamics. It runs at about one-third the speed of a Cray 1, but is smaller than the viewgraph projector. It consists of ten computers which run in parallel. There are ten because there are nine planets plus the sun. Each computer has limited capabilities: it can add, multiply, and take $-\frac{3}{2}$ powers—just enough to integrate Newton’s equations. The mantissa has 56 bits compared to 48 bits on the Cray; it has turned out that this extra precision was crucial.

In our first calculation (Applegate et al., 1985) we integrated the outer planets (Sun, Jupiter, Saturn, Uranus, Neptune, and several massless Plutos) forward and backward in time for about 100 million years. This is longer than the classic million year (±500 000 yr) integration of Cohen, Hubbard, and Oesterwinter (1973) by a factor of more than two orders of magnitude.

One result from that integration was that the best analytic perturbation theory (Bretagnon, 1982) for the solar system was inadequate. Numerically resolving the observed motions of the massive planets into a
Is the solar system stable?

Fig. 2. The mechanical orreries are nice symbols of the apparent clockwork predictability of the motions of the planets.

quasiperiodic series, we found that the spectrum of Jupiter contained terms which were larger than all but seven of the 200 terms listed in

Bretagnon's solution of perturbation theory works.

The motion of Pluto crosses the orbit of Neptune are in an orbital resonance: orbital period of Neptune Pluto. With the system

Fig. 3. The "Digital Orrery" is a computer which was specifically designed to investigate the dynamics of the solar system.

Fig. 4. The power spectrum common logarithm of the p

Fig. 5. The power spectrum perturbation or ries. Lines which are the order in planetary
Fig. 4. The power spectrum of a variable related to the eccentricity of Jupiter's orbit. The common logarithm of the power is plotted vs frequency (in cycles per day).

Bretagnon's solution (see Figs. 4 and 5). The problem was that the perturbation theory was not carried to high enough order.

The motion of Pluto is particularly complicated. The orbit of Pluto crosses the orbit of Neptune. This is only possible because the two planets are in an orbital resonance (Cohen and Hubbard, 1965): three times the orbital period of Neptune is approximately two times the orbital period of Pluto. With the system in this resonance close encounters do not occur.

Fig. 5. The power spectrum of Jupiter was not adequately represented by the best analytic perturbation theories. Lines marked D and E are not recovered in perturbation theories which are third order in the eccentricities and inclinations and second order in the planetary masses.
Is the solar system stable?

Fig. 6. The argument of perihelion of Pluto had a strong long-period modulation, with a period of 34 million years.

Fig. 7. There was a very low to Pluto's eccentricity.

spectra or the present associated with chaotic spectrum is difficult to whether or not near calculation of the Ly displays log₁₀ γ versus

Fig. 8. The inclination of 1 million year integration.

even though the orbits cross. Integrations by Williams and Benson (1973) showed that Pluto was involved in yet another resonance: Pluto's perihelion (the longitude at which the planet is closest to the sun) and its ascending node (the longitude at which the orbit plane crosses the plane perpendicular to the angular momentum of the solar system) are locked together. The regression of the perihelion and the ascending node have precisely the same periods. The difference between the two angles, which is called the argument of perihelion, oscillates about π/2 with a period of about 3.8 million years.

Several new features in the motion of Pluto were revealed by our calculation. We found a surprisingly large number of strong, long-period variations. The argument of perihelion showed a strong modulation with a period of 34 million years (Fig. 6). The variable \( h = e \sin \varpi \), where \( e \) is the orbital eccentricity and \( \varpi \) is the longitude of perihelion, had a strong component with a period near 137 million years (Fig. 7). (This variable plays an important role in analytic theories.) This frequency may be associated with a near resonance between one of the fundamental frequencies associated with Pluto and one of the fundamental frequencies of the system of massive planets. The inclination of Pluto showed some evidence of a period longer than our integration, or perhaps even a secular decline (Fig. 8). The most suspicious bit of evidence was the very noisy power spectrum of the resonance variable associated with the basic 3:2 commensurability of the orbital periods (Fig. 9). The power spectrum of a quasiperiodic trajectory should have no more independent frequencies than the number of degrees of freedom. Noisiness of power
Williams and Benson's other resonance: Pluto's closest to the sun) and its plane crosses the plane solar system) are locked he ascending node have one of the two angles, which put $\pi/2$ with a period of

Fig. 7. There was a very long period (137 million year) component in a variable related to Pluto's eccentricity.

spectra or the presence of a broadband component have been widely associated with chaotic behavior. Unfortunately, the noisiness of a power spectrum is difficult to quantify.

The proper thing to do to determine whether a trajectory is chaotic or quasiperiodic is to compute the Lyapunov exponents, which measure whether or not neighboring trajectories diverge exponentially. Our first calculation of the Lyapunov exponent is shown in Fig. 10. The plot displays $\log_{10} \gamma$ versus $\log_{10}(t - t_0)$, where $\gamma = \ln[d(t)/d(t_0)]/(t - t_0)$.

Fig. 8. The inclination of Pluto had periods longer than could be resolved by our 214 million year integration.
Is the solar system stable?

Fig. 9. Expanded views of power spectra for Pluto were suspiciously noisy.

and \( d(t) \) the phase-space distance between neighboring Plutos. There was no sign that \( \gamma \) was leveling off to a positive Lyapunov exponent.

1.2. New integration

We were compelled by the very long periods and noisy spectra we found in our integrations to carry out a longer integration of the solar system. The motion of Pluto just seemed too complicated.

Our earlier integration was limited to \( \pm 100 \) million years because of the accumulation of roundoff errors. We found a trick that allowed us to

Fig. 10. The computation of the Lyapunov exponent did not show any indication of a nonzero exponent greater than \( 10^{-8.4} \) yr\(^{-1} \).

significantly reduce much longer times. The orbits of some planets.

Fig. 11. Relative error.

in energy corresponds to.

The origin of this which can be estimated is that for 0.0 W. Kahan, we in a system of which there was a small step size which eliminated the length of the test integration. The computations were made over million days versus the measured in days, of the solar mass.

The best step size for our new integ
re suspiciously noisy.

... ring Plutos. There was... exponential.

noisy spectra we found on of the solar system.

illion years because of... that allowed us to

significantly reduce our numerical error and extend our integrations to much longer times. In all direct long-term numerical integrations of planetary orbits the error energy is found to grow linearly with time (see Fig. 11). This energy error dominates all other errors; it leads to a quadratic growth of the error in all positions and longitudes since the linear error in energy corresponds to a linear error in frequency which when integrated yields an error in longitude which grows quadratically in time. The origin of this energy error is not understood; there is as yet no theory which can estimate the rate of the linear growth. The naive expectation would be that the error in energy should behave like a random walk and grow with the square root of time. However, even though we do not understand why the error grows as it does we can make use of it. We found in our numerical studies that the slope of the energy error depended on the step size; in fact, for some step sizes the energy error was positive and for others the error was negative. Following a suggestion of W. Kahan, we investigated whether there might be special step sizes for which there was no linear growth in the energy error. We found that this was indeed the case. More importantly, we found that the special step size which eliminated the linear energy error did not depend on the length of the test integrations, but rather became better defined as the test integrations were extended. Figure 12 shows the energy error after 5 million days versus step size. The units are not important, but time is measured in days, distance is in astronomical units, and mass is in units of the solar mass. Figure 13 shows the energy error after a billion days. The best step size of these test runs is 32.7 days, which is what we chose for our new integrations. We extensively studied the numerical accuracy...
of integrations with various step sizes. These errors were determined by integrating forward in time, then reversing the system to recover the initial state. This is a valid test since our integrator is not microscopically reversible or explicitly symplectic. Figure 14 shows the round trip errors in the position of Jupiter, which display a minimum at the special step size of 32.7 days. Thus removing the energy error actually gives a better long-term trajectory (as indicated by the round trip errors). It is curious that at the special step size the round trip errors in all the planets are comparable; this is not true for other step sizes.

Table I lists for comparison the rate of growth of energy error in

Table I. Energy errors

<table>
<thead>
<tr>
<th>Integration</th>
<th>Energy Error (after 100 million years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHO (1973)</td>
<td>-1.2 × 10^{-5}</td>
</tr>
<tr>
<td>K &amp; N (1984)</td>
<td>-1.5 × 10^{-5}</td>
</tr>
<tr>
<td>200 MYR (1986)</td>
<td>-2.0 × 10^{-5}</td>
</tr>
<tr>
<td>LONGSTOP (1986)</td>
<td>-2.5 × 10^{-5}</td>
</tr>
<tr>
<td>LONGSTOP (1987)</td>
<td>-3.0 × 10^{-5}</td>
</tr>
<tr>
<td>This work</td>
<td>-3.5 × 10^{-5}</td>
</tr>
</tbody>
</table>

Fig. 12. The energy error after 5 million days as a function of step size.

Fig. 13. The energy error after a billion days as a function of step size.

Fig. 14. The error in steps of 32.7 days and then back at the same step size.

several long-term significantly more livations of the solar for nearly a billion at the end of the longitude of Pluton months on the D computer time.
tion of step size.

s were determined by system to recover the is not microscopically the round trip errors are at the special step actually gives a better energy error. It is curious in all the planets are

h of energy error in

Fig. 14. The error in recovering the position of Jupiter after integrating forward a billion days and then backward a billion days vs step size. The round trip errors are a minimum at the same step size that minimizes the energy error.

several long-term integrations of the solar system. Our new integration is significantly more accurate and longer than all other long-term integrations of the solar system. We simulated the motion of the outer planets for nearly a billion years. The estimated error in the longitude of Jupiter at the end of the integration is only a few degrees, and the error in the longitude of Pluto is a few arc minutes. The calculation took about five months on the Digital Orrery; this is about 4 MegaFlop-years of computer time.

Table I. Energy errors in the various long-term integrations of the outer planets. Our first integration is marked "200 MYR (1986)." In our new integration the growth of the energy error is about three orders of magnitude smaller than all previous long-term integrations.

| Integration       | Interval of integration | $\frac{d}{dt} \frac{E - E_0}{|E_0|} [\text{yr}^{-1}]$ |
|-------------------|-------------------------|---------------------------------------------------|
| CHO (1973)        | 1 000 000 yr            | $2.4 \times 10^{-16}$                             |
| K&N (1984)        | 6 000 000 yr            | $5 \times 10^{-16}$                               |
| 200 MYR (1986)    | 214 000 000 yr          | $1.8 \times 10^{-16}$                             |
| LONGSTOP (1986)   | 8 000 000 yr            |                                                   |
| LONGSTOP (1987)   | 100 000 000 yr          | $-2.5 \times 10^{-16}$                           |
| This work         | 845 000 000 yr          | $-3.0 \times 10^{-19}$                           |
Fig. 15. The argument of perihelion of Pluto continues to display a 34 million year modulation.

The argument of perihelion of Pluto is displayed in Fig. 15. The 34 million year oscillation we previously observed (Fig. 6) is the longest period which is noticeable in the time series for the argument of perihelion. The 137 million year modulation of the quantity $h = e \sin \omega$, is also still the longest period noticeable (Fig. 16). The inclination was not secularly declining, but there may be a weak component with a period near 600 million years (Fig. 17).

The most important indicator of chaotic behavior is the largest Lyapunov exponent. We computed the largest Lyapunov exponent in several different ways. The simplest method is to just look at the divergence of nearby trajectories. Figure 18 shows the divergence of two Plutos.

Fig. 16. The eccentricity of Pluto continues to display the 137 million year modulation.

Fig. 18. The logarithm of growth indicates exponential simple two-trajectory evolution.
display a 34 million year

red in Fig. 15. The 34
(Fig. 6) is the longest
he argument of perihe-
olarity \( h = e \sin \alpha \), is also
ne inclination was not
ponent with a period

behavior is the largest Ly-
unov exponent in sev-
look at the divergence of two Plutos.

Already here there is evidence of exponential divergence of trajectories. The slope gives an exponential divergence time scale of about 20 million years. This simple method has the disadvantage that the two trajectories eventually separate so much that the two trajectories do not any longer represent the divergence of “nearby” trajectories, and ultimately the distance between the trajectories saturates. In this case two Plutos cannot separate more than about 45 AU, if the amplitude of the oscillation of the basic 3:2 orbital resonance variable does not significantly change. There are two prescriptions for improving on this simple Lyapunov exponent calculation. One way is to “renormalize” the distance between the two

Fig. 17. The inclination of Pluto varies somewhat irregularly.

Fig. 18. The logarithm of the distance between two “Plutos” is plotted vs time. The linear growth indicates exponential divergence with a time scale of about 20 million years. The simple two-trajectory method saturates at a distance of about 45 AU.
Is the solar system stable?

Fig. 19. The logarithm of the distance between neighboring Plutos computed with the linearized variational equations.

trajectories, i.e., the test trajectory is repeatedly brought back to be near the reference trajectory preserving the direction (in phase space) between them, and keeping track of how much the test trajectory was moved. The problem with this method is knowing how often to renormalize. In a system with a small Lyapunov exponent, we found that too frequent renormalization can contaminate the result. The best method is to integrate the linearized variational equations along with the equations of motion. This method needs renormalization only if the distance grows so large that it cannot be represented in the computer—a problem not encountered here. Figure 19 shows the divergence of nearby Plutos computed using the variational method. The agreement with the two Pluto calculation is striking (Fig. 20). The two calculations differ significantly only when the two trajectory method is close to saturation. Figure 21 shows the conventional (more conservative) plot of the calculation of the Lyapunov exponent. The exponent is clearly leveling off at a positive value, in a manner typical of Lyapunov exponent calculations. The inverse of the Lyapunov exponent is about 20 million years. Thus our calculation indicates that the motion of Pluto is chaotic, with a surprisingly short divergence time scale.

It is interesting to further examine the divergence of nearby trajectories by displaying the distance between Plutos on a log \( d \) versus log \( t \) plot (Fig. 22). Several different experiments are superimposed to show the envelope. The solid curve is an exponential with a 20 million year e-folding time scale. The dashed curve is a power law with distance proportional to the 3/2 power of the time. The plot shows that the initial divergence of the Plutos follows the 3/2 power law. Numerical experi-

Fig. 20. The two \( \tau \) are in excellent ag

ments on a sys

imilar behavio

with the square

that when the:

gence begins to

gence acts as a

illustrates the c

tion during a cr

ization is repea

Fig. 21. The conv

positive exponent i
Fig. 20. The two methods for computing the divergence of trajectories give results which are in excellent agreement up to the point of saturation.

Fig. 21. The conventional representation of a Lyapunov calculation clearly indicates a positive exponent with a divergence time scale of about 20 million years.
Is the solar system stable?

Fig. 22. A log-log plot of the distance between two Plutos vs time emphasizes the early stages of the separation. The initial divergence due to round-off error provides a seed to the exponential divergence associated with chaotic behavior.

vergence, the measured Lyapunov exponent may have no relation to the true exponent. The variational method does not have this problem.

There is a disturbing aspect to this result that the motion of Pluto is chaotic. In other examples of chaotic behavior in the solar system (see Wisdom, 1987 for a review) trajectories always look decidedly irregular when there is a positive Lyapunov exponent. That is not the case here. For example, the argument of perihelion (Fig. 16) looks remarkably regular. The inclination is a little more complicated, but not obviously irregular. The strongest evidence for irregular behavior is in the power spectra. Figure 23 shows a portion of the power spectrum of $h = e \sin \omega$

Fig. 23. This expanded view of the Pluto spectrum still has a broadband character.

Fig. 24. This expanded periodic line spectrum.

for Pluto. Figure shows the power spectrum of the power line spectrum; the that the broadband spectrum of Neptune consistent with our

In planetary mot esting of the orbita closely related to the longitudes of the pl motion on the associated with the age out and produc gitudes are ignorable the average, consta the sampled semim orbital element of I Pluto and Neptune. The semimajor axis 28 show expanded Neptune shows pat! On the other hand, of the semimajor ax
time emphasizes the early off error provides a seed to have no relation to the save this problem.

The motion of Pluto is the solar system (see decidedly irregular is not the case here. '6) looks remarkably ed, but not obviouslyavior is in the power spectrum of $h = e \sin \varpi$ broadband character.

for Pluto. Figure 24 shows, for comparison, a similar portion of the power spectrum of Neptune. The power spectra are qualitatively different. The spectrum of Neptune is consistent with a very complicated line spectrum; the spectrum of Pluto has a broadband character. Note that the broadband component is stronger than most of the lines in the spectrum of Neptune. The broadband component of Pluto's spectrum is consistent with our measurement of a positive Lyapunov exponent.

In planetary motion the semimajor axes are usually the most uninteresting of the orbital elements. The reason is that the semimajor axes are closely related to the momenta conjugate to the fastest moving angles, the longitudes of the planets. Because of the large separation of time scales, the motion on the orbital time scale is not very important. The effects associated with the extremely rapid variations in longitude tend to average out and produce no cumulative effect. The fact that the orbital longitudes are ignorable implies that the variables conjugate to them are, on the average, constant and hence not very interesting. Curiously, plots of the sampled semimajor axis appear more irregular than plots of any other orbital element of Pluto. Figures 25 and 26 show the semimajor axes of Pluto and Neptune (for comparison) over several hundred million years. The semimajor axis of Pluto is markedly more irregular. Figures 27 and 28 show expanded portions of these time series. The semimajor axis of Neptune shows patterns characteristic of a sampled quasiperiodic signal. On the other hand, the word "turbulent" is brought to mind by the plot of the semimajor axis of Pluto. Apparently, sampling the semimajor axis
Is the solar system stable?

Fig. 25. The time series of the semimajor axis of Pluto has an irregular appearance.

has brought out the chaotic character of Pluto's motion. Perhaps there is a clue here to the dynamical resonance mechanism responsible for Pluto's chaotic motion?

Since we performed this numerical experiment which indicated that the motion of Pluto was chaotic an analogous experiment has been performed by J. Laskar (1989) for the whole solar system. With today's computers and today's computer time allotments it is impossible to directly integrate the whole solar system for hundreds of millions of years. Instead, Laskar has studied the whole solar system by first analytically averaging the equations of motion to yield a system of about 150,000 secular equations which he then integrated on a conventional supercomputer. Figure 29 shows the calculation of the Lyapunov exponent from

Fig. 26. The time series of the semimajor axis of Neptune has a quasiperiodic appearance.

Fig. 27. On a shorter time scale irregular appearance.

his paper. Laskar has fit exponential divergence long-term evolution of

One word of caution one of a kind calculation of our results to uncertain initial conditions of the tiny and with the true unlikely to me since tv
as an irregular appearance.

motion. Perhaps there is something responsible for Pluto's

t which indicated that the experiment has been per-

ar system. With today's computers it is impossible to di-

reducions of millions of years. We first analytically sys-

em by first analytically solving a system of about 150,000

conventional supercomputer model. Lyapunov exponent from

his paper. Laskar has found that the whole solar system is chaotic with an

exponential divergence time of only about 5 million years. The detailed

long-term evolution of the Earth is unpredictable!

One word of caution is in order. Both of these calculations are so far

one of a kind calculations. We have not yet evaluated the sensitivity of

our results to uncertainties in our knowledge of the true masses and

initial conditions of the planets. It may be that the chaotic zones are very

tiny and with the true masses the system is not chaotic. This seems

unlikely to me since two very different kinds of calculation have been

Fig. 27. On a shorter time scale the time series of the semimajor axis of Pluto still has an

irregular appearance.

Fig. 28. On a shorter time scale the time series of the semimajor axis of Neptune still

appears quasiperiodic.
Is the solar system stable?

Fig. 29. Calculation of the Lyapunov exponent for the whole solar system indicates exponential divergence with a time scale of only 5 million years.

performed and both gave positive Lyapunov exponents. Still the sensitivity to parameters needs to be addressed.

It is difficult to evaluate the implications of these results. The chaotic character of the solar system opens many possibilities. The high eccentricity and inclination of Pluto's orbit are puzzling since we understand that planets form from a nebula which collapses to a disk; all the planets should more or less be in the same plane. However, I have shown (Wisdom, 1987) that asteroids placed in the chaotic zones associated with orbital resonances with Jupiter often evolve to high inclination and eccentricity. Thus Pluto may have formed in an orbit near the plane of the other planets, and evolved through the chaotic zone to its current high eccentricity and high inclination orbit. This seems a very natural suggestion. Of course, it could also be the case that the chaotic zone in which Pluto moves is very tiny with only exceedingly small holes into other parts of the phase space. In this case, Pluto may always have had much the same orbit. In an extreme view, the chaotic character of Pluto's motion opens the possibility that Pluto was initially in an orbit radically different from the orbit in which it is found today. Chaos enlarges the possibilities rather than constrains them.

The chaotic character of the dynamics of the solar system also precludes any rigorous statement about its future evolution (at the moment). The fact that the solar system is 4.6 billion years old does not demand any conclusion regarding its future stability. Recall my curious results regarding the motion of asteroids in the 3:1 orbital resonance with Jupiter (Wisdom, 1982, 1983). Figure 30 shows the eccentricity of a chaotic asteroid trajectory versus time. For over a hundred thousand years the eccentricity display long periods of stability, but then suddenly it breaks loose, an example of a chaotic zone (Wisdom, 1985). Until we observe Pluto and the rest of the asteroid belt, the possibility of similar behavior cannot be ruled out. We should be asking whether the solar system has been stable over a hundred thousand years from now.

Question 2: Could the nearby planet be stable?

The hallmark of chaotic systems is the exponential
whole solar system indicates

ments. Still the sensitiv-

se results. The chaotic

ities. The high eccen-

since we understand

a disk; all the planets

zones associated with

near the plane of the

ne to its current high

very natural suggest-

ne in which small holes into other

always have had much

character of Pluto's

character of Pluto's

in an orbit radically

Chaos enlarges the

olar system also pre-

olution (at the mo-

n years old does not

y. Recall my curious

orbital resonance with

the eccentricity of a

a hundred thousand

Fig. 30. Eccentricity vs time of a chaotic asteroidlike trajectory. These trajectories often
display long periods of apparently regular motion, followed by bursts of irregular motion.
During the regular periods the trajectory is stuck to an island in the phase space.

years the eccentricity looks quasiperiodic, then it becomes more irregular,
and then suddenly it begins shooting to large values. The reason for this
behavior is that the trajectory is originally stuck to an island, from which
it breaks loose, and eventually it finds its way into a different region of the
chaotic zone in which large jumps in eccentricity take place (Wisdom,
1985). Until we understand the structure of the chaotic zone in which
Pluto and the rest of the solar system move, it is not possible to exclude
the possibility, say, that the orbit of the Earth will suddenly begin to
exhibit similar wild excursions in eccentricity. It is not even possible to
say whether the Earth will still be a member of the solar system a billion
years from now.

Question 2: Can we use chaos to make measurements?
The hallmark of chaos is the sensitive dependence on initial conditions.
The exponential divergence of chaotic trajectories precludes long-term
prediction given limited knowledge of the state of the system. Any error in our knowledge of either the state of the system or the system parameters will exponentially explode into complete uncertainty in the state of the system. Is it possible, though, to turn this sensitivity to our advantage? Rather than trying to predict the future from limited knowledge, can we use the sensitivity of a system undergoing chaotic motion to deduce the properties of the system? Can we measure the behavior of a system and deduce, exponentially well, properties of the system which were required to produce the observed behavior. Several years ago (Wisdom, 1987) I suggested that it might be possible to use noisy observations of the brightness of Hyperion, the chaotically tumbling satellite of Saturn (Wisdom et al., 1984; Klauder, 1989), to determine exponentially well both the state of rotation of Hyperion and the ratios of its principal moments of inertia. Letting the dream get out of hand, it even seemed conceivable that if extensive observations of this otherwise uninteresting 400 km satellite could be made, an arbitrarily large amount of information could be deduced about the Saturn system, without a spacecraft going there. And why restrict oneself to observations of objects which are difficult to observe? Surely the behavior of the double pendulum depends outrageously sensitively on the parameters of the system such as the lengths and masses of the individual pendula, and more importantly on the value of the gravitational acceleration g. Can we make a chaotic double pendulum to measure g, or a chaotic Cavendish balance to measure the gravitational constant G? Unfortunately, it appears there is a sort of "classical uncertainty principle," which prevents us from using the extreme sensitivity of chaotic behavior to make exponentially accurate measurements.

There has recently been considerable interest in the closely related problem of "noise reduction" (Farmer and Sidorowich, 1987, 1988; Kostelich and Yorke, 1989). Given the dynamical system (with its parameters), it certainly is possible to use noisy measurements to exponentially constrain the trajectory of the system, away from the end points of the time series. In the work reported here, I used neither of the techniques for noise reduction mentioned above. Rather, I use the first part of the shadowing algorithm as described by Hammel et al. (1989). By the way, the results of Hammel et al., are especially nice, if anyone is unfamiliar with them. They prove, by "computer proof," that for particular trajectories of particular maps, there are true trajectories of the map which stay close to the numerically computed trajectory for an astonishingly long time. Typical results are that for trajectories computed with 64 bit arithmetic true trajectories shadow computed trajectories within a relative accuracy of $10^{-3}$ for $10^6$ iterations. Who could ask for more of a numerical trajec-

tory? The first trajectory a more careful roundoff error was mentioned I works pretty well. We algorithm does fail. The algorithm result. Despite the essence of the problems, as we sho:

In order to fo

The map is used I make measure each time step, being measured all the cations of embbe the dynamical s is unknown. Thi the dynamical s ation, to what e exponentially do

Figure 31 illu dynamical syste The common lc the true orbit, at the shadow orbi number. We see observations on except at the en machine is not could be recove
tially with the l

Now what h noise with a slig trates the result $10^{-12}$. In each
the system. Any error or the system param-
certainty in the state of
sitivity to our ad
ing chaotic motion to
ure the behavior of a
s of Saturn
exponentially well
ratios of its principal
hand, it even seemed
thereby uninteresting
amount of informa-
without a spacecraft
is of objects which are
pendulum depends
system such as the
more importantly on
we make a chaotic
balance to mean-
appears there is a sort
us from using the
exponentially accurate
in the closely related
vich, 1987, 1988; Ko-
em (with its param-
end points of the
t of the techniques for
first part of the shad-
By the way, the
one is unfamiliar with
rticular trajectories of
which stay close to
ingly long time. Typ-
bit arithmetic true
relative accuracy of
a numerical traject-
tory? The first part of the proof is to compute from the computed tra-
jectory a more consistent pseudotrajectory which is less contaminated by 
the roundoff error. That this algorithm could be used for noise reduction 
was mentioned by Sidorovich and Farmer (1988). I have found that it 
works pretty well, though it is not ideal for a number of reasons. The 
algorithm does not work for regular orbits. The algorithm sometimes 
fails. The algorithm does not give a natural estimate of the error in the 
result. Despite these shortcomings the shadow algorithm captures the 
illusion of the problem with parameter estimation from noisy measure-
ments, as we shall shortly see.

In order to focus on the matters of principle rather than gory details I 
set up a highly idealized set of numerical experiments. I chose to study a 
simple map, the standard map:

\[
\begin{align*}
    y' &= y + K \sin x \\
    x' &= x + y' \mod 2\pi.
\end{align*}
\]  

(1)

The map is used to generate a time series of the trajectory \((x_n, y_n)\). Then 
I make measurements of the trajectory by "observing" both \(x\) and \(y\) at 
each time step. Observations are performed by adding to each variable 
being measured a Gaussian random number with standard deviation \(\sigma\). I 
measure all the state variables in order to avoid the unnecessary 
complications of embedding. In performing the noise reduction I assume that 
the dynamical system, the standard map, is known, but the parameter \(K\) 
is unknown. This avoids problems of having to build an approximation to 
the dynamical system from the data. Given this intentionally ideal situ-
ation, to what extent is it possible to determine the parameter \(K\)? Is it 
exponentially determined by the observations?

Figure 31 illustrates the now familiar result of noise reduction with 
the dynamical system, including the parameter, known. In this case \(K = 1\.

The common logarithm of the difference between the observations and 
the true orbit, as well as the common logarithm of the difference between 
the shadow orbit and the true orbit, are plotted as a function of iteration 
number. We see that the shadow orbit, which is computed from the noisy 
observations only, recovers the true orbit to about machine precision 
except at the end points of the time series. Of course, the precision of the 
machine is not a fundamental limitation; in principle a true trajectory 
could be recovered with an accuracy which continues to grow exponent-
ially with the length of the time series.

Now what happens as we use the shadow algorithm to remove the 
noise with a slightly incorrect value of the parameter \(K\)? Figure 32 illus-
trates the result for three different values of \(\delta K = K - K_{true} = 0, 10^{-8}\), 
and \(10^{-12}\). In each case the shadow algorithm finds a shadow orbit, and the
error in recovering the true orbit is roughly of order the error in the parameter $K$. We can see already that if this is generally true, then it will be difficult to determine parameters. The differences between the shadow orbits and the true orbit are far below the errors in the measurements. The measurements, which are all we know in general, apparently cannot distinguish which one is correct.

Perhaps we should try anyway. An objective measure of goodness of fit is

$X^2 = \frac{\sum}{2}
$

where $(\bar{x}, \bar{y})$ is the orbit. In the sum I time series, exclude number of points in fitting $K$ by minimi

We can estimate the number of obs these shadow orbit shadow orbits seem mine the shadow or then use that shac shadow algorithm a true orbit is still rec form a simple one | Fig. 33, for the $i$th measured position the position of the trajectory varies as: shadow position fro linearity can be see $\delta K = 10^{-12}$ are ve for small $\delta K$ is a si tories form a single
Fig. 33. Illustration of the behavior of the shadow orbits at the \( i \)th point. The shadow orbits form a single parameter family of orbits.

\[
\chi^2 = \frac{1}{2N_0^2} \sum_{i=0}^{N-1} \left[ (\bar{x}_i - \bar{x})^2 + (\bar{y}_i - \bar{y})^2 \right],
\]  

(2)

where \((\bar{x}_i, \bar{y}_i)\) is the observational data, and \((\bar{x}, \bar{y})\) is the shadow orbit. In the sum I only included those points in the interior part of the time series, excluding the exponential transients at the ends. \(N\) is the number of points included in the sum. We can try to determine the best fitting \(K\) by minimizing the \(\chi^2\) of the shadow orbit as a function of \(K\).

We can estimate how the error in the best fitting \(K\) should depend on the number of observations. There are some interesting properties of these shadow orbits which will help us to do this. The first is that the shadow orbits seem to be uniquely determined. For instance, if I determine the shadow orbit with a slightly wrong value of the parameter, and then use that shadow orbit as the initial observational data for the shadow algorithm again, this time with the correct value of \(K\), then the true orbit is still recovered. Away from the end points, the shadow orbits form a simple one parameter family of trajectories. This is illustrated in Fig. 33, for the \(i\)th iteration. The point with the error bars indicates the measured position for this iteration. The dot labeled \((x, y)\) represents the position of the true trajectory. The curve represents how the shadow trajectory varies as a function of \(\delta K\), for this iteration. The distance of the shadow position from the true position is linearly proportional to \(\delta K\). The linearity can be seen in Fig. 32, where the curves for \(\delta K = 10^{-8}\) and \(\delta K = 10^{-12}\) are vertically displaced copies of one another. The linearity for small \(\delta K\) is a simple consequence of the fact that the shadow trajectories form a single parameter family. From the picture of the shadow
Is the solar system stable?

Fig. 34. Parameter estimation errors vs the number of points in the segment of the trajectory inside the exponential transients. The errors decrease as $1/N$ rather than $1/\sqrt{N}$ as was naively expected.

The trajectory given in Fig. 33 we can estimate how the best fitting $\delta K$ will vary with the number of points in the sum. Each measurement provides additional information as to where along the curve the true trajectory lies. The error in $\delta K$ should thus be proportional to $1/\sqrt{N}$, which is not very spectacular.

Curiously, this $1/\sqrt{N}$ expectation is not correct. Figure 34 shows the results of a number of numerical experiments. For each $N$ several runs were made with different random numbers, the best fitting $\delta K$ were determined by the procedure outlined above for each of these runs. The mean and standard deviation of the mean of the common logarithms of these $\delta K$'s are plotted versus the logarithm of $N$. The dotted curve shows the expected $1/\sqrt{N}$ behavior. The dashed curve is simply $\sigma/N$. Apparently, this provides an excellent fit to the data.

Where does the extra factor of $1/\sqrt{N}$ come from? Let us continue the calculation of $\chi^2$. It is convenient to introduce vector notation, $\mathbf{x}_i = (x_p, y_i)$ for the true orbit, $\bar{x}_i$ for the shadow orbit, and $\bar{x}_i$ for the observations. Because of the linear dependence of the shadow orbit on $\delta K$, for small $\delta K$, we can write

$$\bar{x}_i = x_i + \delta KC_i.$$  \hfill (3)

The observational data can be represented:

$$\bar{x}_i = x_i + n_i,$$  \hfill (4)

where $n_i$ are two coefficients generator with standard deviations.

Solving for the $\delta K$

Now if all the $C_i$ magnitude of $\delta K$ scale identically the average trajectory. This curious for each value of $j$ resulting plot of the irregular way. At difficult to shadow the irregular close. The longer the trajectory.

There seems to though the trajectory initial conditions trajectories are all wrong system parameters. The chaos I uncertainty in all exponentially accent orbit. Given $K$, the trajectory, the can deduce from them.

This particular sensitive deponder used to determine though that chaos...
where \( n_i \) are two-dimensional samples from a Gaussian random number generator with standard deviation \( \sigma \). The sum for \( \chi^2 \) becomes

\[
\chi^2 = \frac{1}{2N\sigma^2} \sum_{i=0}^{N-1} \| \bar{x}_1 - \bar{x}_i \|^2
\]

(5)

\[
= \frac{1}{2N\sigma^2} \sum_{i=0}^{N-1} \| n_i - \delta KC_i \|^2.
\]

(6)

Solving for the \( \delta K \) which gives a minimum \( \chi^2 \) we find

\[
\delta K = \frac{\sum_{i=0}^{N-1} n_i C_i}{\sum_{i=0}^{N-1} C_i C_i}.
\]

(7)

Now if all the \( C_i \) have comparable lengths, say, \( \| C_i \| \approx C \), then the magnitude of \( \delta K \) scales as \( \sqrt{N}/NC \), i.e., proportional to \( 1/\sqrt{N} \), again. Evidently the average length of the \( C_i \) grows with the length of the trajectory. This curious suggestion was tested by computing the quantity:

\[
D = \frac{\sum_{i=0}^{N-1} \| C_i \|}{\sum_{i=0}^{N-1} C_i C_i}
\]

(8)

for each value of \( j \) up to the full length of the computed trajectory. The resulting plot of \( D \) versus \( j \) did indeed fall as \( 1/\sqrt{N} \), but in a very irregular way. Apparently, some segments of the trajectory are more difficult to shadow than others, and the \( 1/\sqrt{N} \) behavior of \( D \) results from the irregular close encounters of the trajectory with these difficult spots. The longer the trajectory, the more likely a really difficult spot is encountered.

There seems to be a sort of “classical uncertainty principle.” Even though the trajectories are very sensitive, and almost any variation of initial conditions or parameters lead to exponential divergence, chaotic trajectories are also very flexible, so flexible in fact that even with the wrong system parameters there are chaotic trajectories that can fit the data. The chaos has reduced the dimension of the unknowns from an uncertainty in all of \( x, y, z \), and \( K \), to just an uncertainty in \( K \). There is an exponentially accurate constraint between any given value of \( K \) and the orbit. Given \( K \), the true trajectory can be determined; given any point on the trajectory, the parameter \( K \) can be determined. Unfortunately, all we can deduce from the data is the extremely accurate relationship between them.

This particular set of numerical experiments indicates then that the sensitive dependence of chaotic motion on system parameters cannot be used to determine system parameters exponentially well. It does appear though that chaos, in a somewhat mysterious way, gives an extra im-
Is the solar system stable?

Improvement in the determination of system parameters by one extra factor of \(1/\sqrt{N}\) beyond the naive estimate. Perhaps even this level of improvement could be useful in a physical experiment.

I think it would be worthwhile to carry out similar experiments on continuous systems, and higher dimensional systems.

Acknowledgments

The work reported in this paper was carried out in collaboration with G. J. Sussman, and was supported in part by a NSF Presidential Young Investigator Award and by the NASA Planetary Geology and Geophysics program.

References


C., and Sussman, G. J., "A

n collaboration with G.
SF Presidential Young
Geology and Geophys-

etters by one extra factor n this level of improve-

imilar experiments on ems.

approaches of Pluto to Nep-
ents of the outer planets for Ephemesis XXII, 1 (1973).

ernmental Data Using Time-
viour of the Solar System," at the Motion of Pluto is uene–Pluto System," Astron.

r asteroidal motion near the