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.

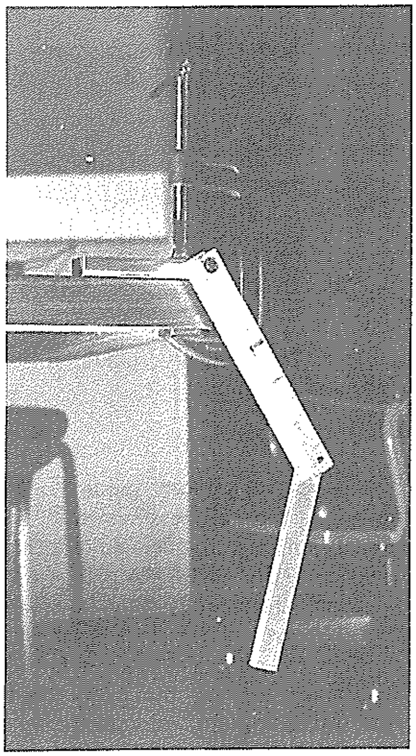


Fig. 1. The double pendulum provides a nice demonstration of chaotic behavior and the divided phase space in simple physical systems.

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 deci
of the stability of the so
numerical experiments i
(Sussman and Wisdom,

1.1. Previous integration

Numerical integrations
amount of computer tim
time scales. A direct cal
follow each planet arou
interesting on time scal
This time scale separ
two-body problem. The
frequency, the orbital fr
freedom after eliminatio
by planetary perturbatio
orbital nodes precess.
orbital frequency multip
that of the sun. The pre
of years to a couple of m
precession time scale s
coupling between the pl
We have performed
(Applegate *et al.*, 1985
centuries (Fig. 2), the
puter designed specifica
one-third the speed of a
jector. It consists of ten
because there are nine
capabilities: it can add,
integrate Newton's equ
bits on the Cray; it has
In our first calculatio
planets (Sun, Jupiter,
Plutos) forward and ba
is longer than the clas
Cohen, Hubbard, and C
orders of magnitude.
One result from that
tion theory (Bretagnon,
merically resolving the

?

n of chaotic behavior and the

lern science. As soon as
ɔne has to wonder about
ı assumed that our solar
ɜ, and that it is only a
s fact is given. Tremen-
de. Arnold (1963) has
large measure provided
of the planets are suffi-
dynamical systems gen-
behavior, and the solar
em. The question of its
ıd.
r our solar system is on
hysical experiment runs

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 –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

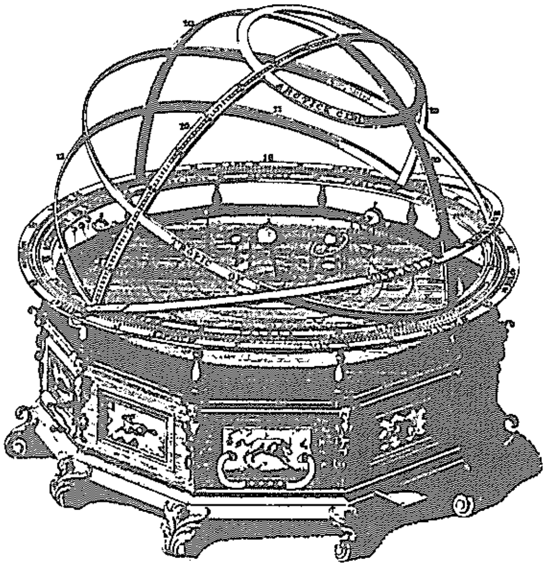


Fig. 2. The mechanical orreries are nice symbols of the apparent clockwork predictability of the motions of the planets.

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

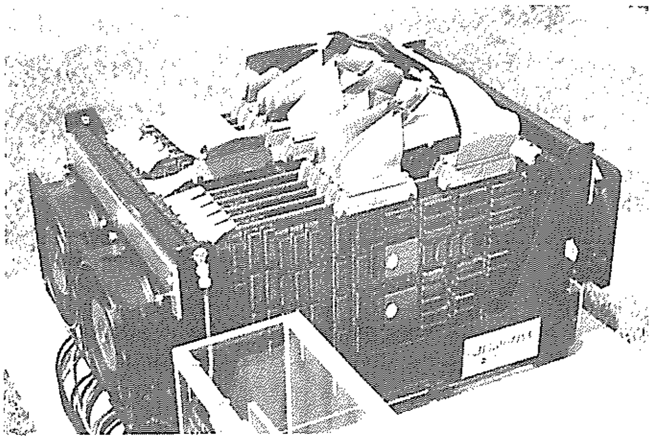


Fig. 3. The "Digital Orrery" is a computer which was specifically designed to investigate the dynamics of the solar system.

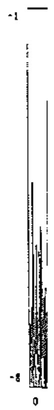


Fig. 4. The power spectrum
common logarithm of the p

Bretagnon's solution
perturbation theory w
The motion of Plut
crosses the orbit of Ne
are in an orbital reson
orbital period of Neptu
Pluto. With the syster

Fig. 5. The power spectrum
perturbation theories. Lines
which are th [illegible] order in t
planetary m

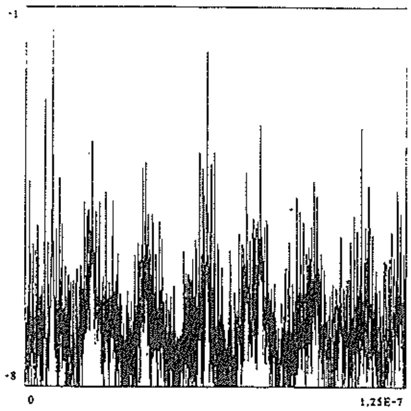


Fig. 4. The power spectrum of a variable related to the eccentricity of Jupiter's orbit. The common logarithm of the power is plotted vs frequency (in cycles per day).

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

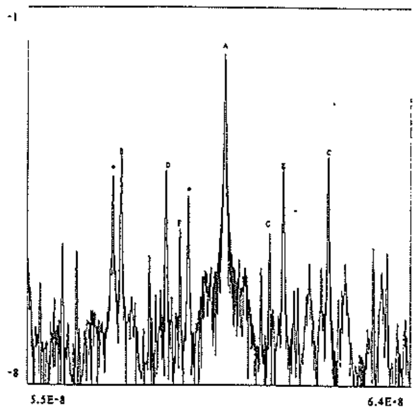


Fig. 5. The power spectrum of Jupiter was not adequately represented by the best analytic perturbation theories. Lines marked D and E are not recovered in perturbation theories which are third order in the eccentricities and inclinations and second order in the planetary masses.

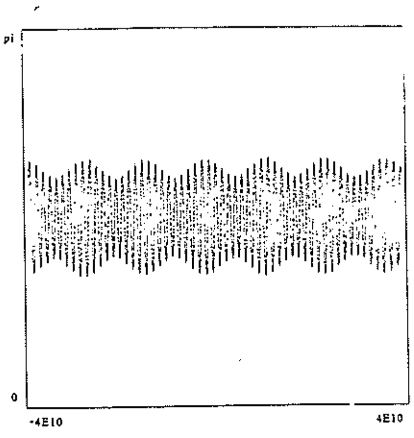


Fig. 6. The argument of perihelion of Pluto had a strong long-period modulation, with a period of 34 million years.

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 $\pi/2$ 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 ϖ 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

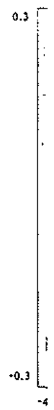


Fig. 7. There was a very lo
to Pluto's eccentricity.

spectra or the presen
associated with chaotic
spectrum is difficult to
The proper thing to
quasiperiodic is to co
whether or not neighb
calculation of the Ly
displays $\log_{10} \gamma$ versus

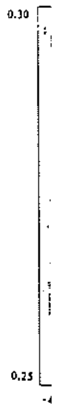
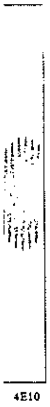


Fig. 8. The inclination of
million year integration.



ng-period modulation, with a

Williams and Benson
other resonance: Pluto's
losest to the sun) and its
plane crosses the plane
olar system) are locked
he ascending node have
n the two angles, which
out $\pi/2$ with a period of

were revealed by our
er of strong, long-period
strong modulation with
le $h = e \sin \varpi$, where e
de of perihelion, had a
n years (Fig. 7). (This
eories.) This frequency
one of the fundamental
fundamental frequencies
of Pluto showed some
, or perhaps even a sec-
f evidence was the very
ssociated with the basic
g. 9). The power spec-
no more independent
om. Noisiness of power

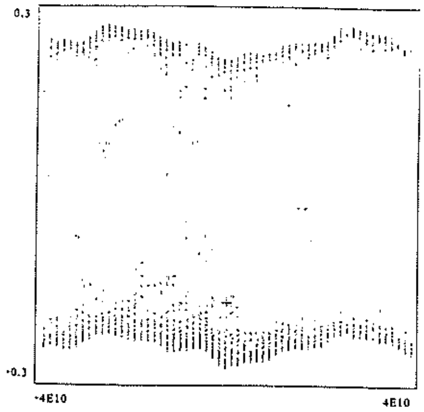


Fig. 7. There was a very long period (137 million year) component in a variable related to Pluto's eccentricity.

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 $\gamma = \ln[d(t)/d(t_0)]/(t - t_0)$,

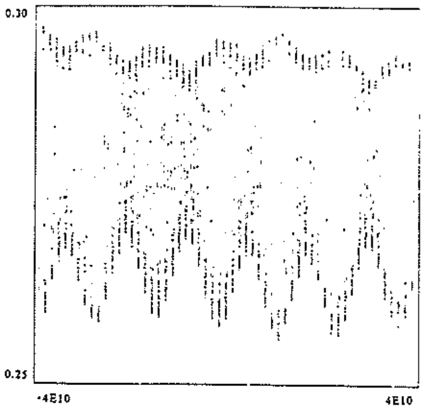


Fig. 8. The inclination of Pluto had periods longer than could be resolved by our 214 million year integration.

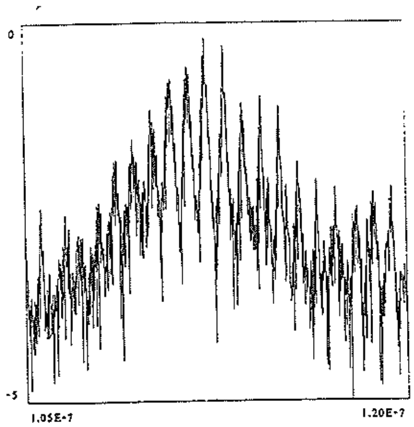


Fig. 9. Expanded views of power spectra for Pluto were suspiciously noisy.

and $d(t)$ the phase-space distance between neighboring Plutos. There was no sign that γ 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 ±100 million years because of the accumulation of roundoff errors. We found a trick that allowed us to

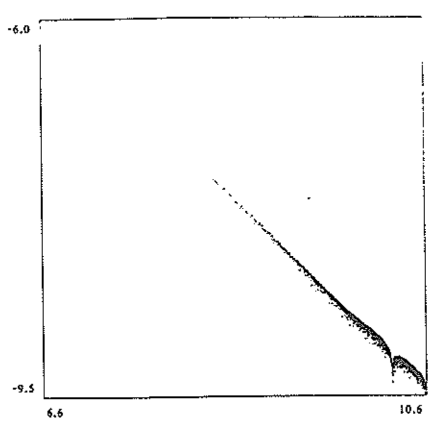


Fig. 10. The computation of the Lyapunov exponent did not show any indication of a nonzero exponent greater than $10^{-6.8}$ yr^{-1}.

Fig. 11. Relative er

significantly redu
much longer time
etary orbits the er
11). This energy
growth of the err
in energy corresp
grated yields an
The origin of this
which can estima
would be that the
grow with the sq
understand why
found in our nur
pended on the ste
positive and for o
W. Kahan, we in
which there was n
was indeed the ca
size which elimir
length of the test i
integrations were
million days versi
measured in days,
of the solar mass.
The best step size
for our new integr

re suspiciously noisy.

ɪring Plutos. There was
ɪnov exponent.

ɪoisy spectra we found
ɔn of the solar system.

illion years because of
rick that allowed us to

t show any indication of a

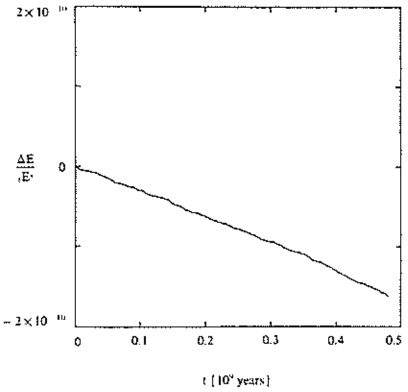


Fig. 11. Relative energy error in our new long-term integration of the outer planets.

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

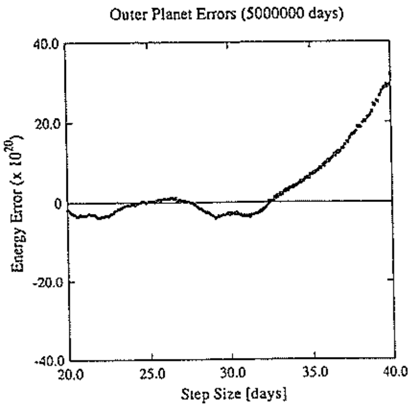


Fig. 12. The energy error after 5 million days as a function of step size.

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

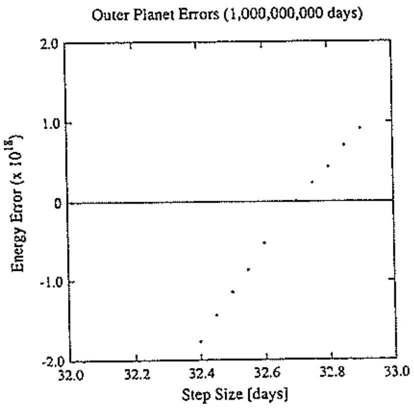


Fig. 13. The energy error after a billion days as a function of step size.

Fig. 14. The error in
days and then backwa
at the same step size

several long-term
significantly more
tions of the solar
for nearly a billio
at the end of the
longitude of Plut
months on the D
puter time.

Table I. Energy error
integration is marked
energy error is abou
integrations.

Integration
CHO (1973)
K&N (1984)
200 MYR (1986)
LONGSTOP (1986)
LONGSTOP (1987)
This work

ıction of step size.

s were determined by
ystem to recover the
is not microscopically
the round trip errors
ım at the special step
ıctually gives a better
) errors). It is curious
in all the planets are

:h of energy error in

ıction of step size.

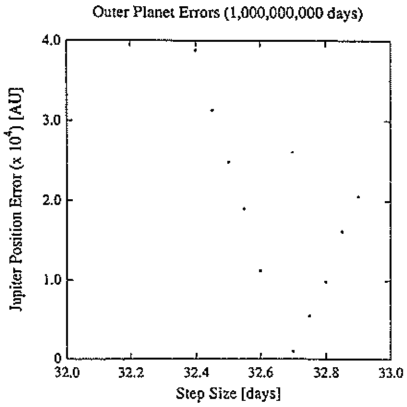


Fig. 14. The error in recovering the position of Jupiter after integrating forward a billion days and then backward a billion days vs step size. The round trip errors are a minimum at the same step size that minimizes the energy error.

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	$\frac{d}{dt}\frac{E-E_0}{\lvert E_0\rvert}$ [yr^{-1}]
CHO (1973)	1 000 000 yr	2.4×10^{-16}
K&N (1984)	6 000 000 yr	5×10^{-16}
200 MYR (1986)	214 000 000 yr	1.8×10^{-16}
LONGSTOP (1986)	8 000 000 yr	
LONGSTOP (1987)	100 000 000 yr	-2.5×10^{-16}
This work	845 000 000 yr	-3.0×10^{-19}

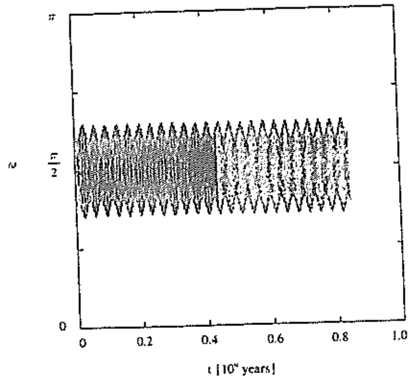


Fig. 15. The argument of perihelion of Pluto continues to display a 34 million year modulation.

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 $h = e \sin \varpi$, 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.

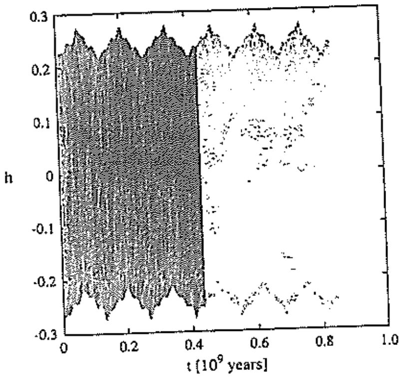


Fig. 16. The eccentricity of Pluto continues to display the 137 million year modulation.

Fig. 17. T

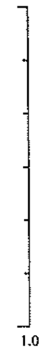
Already here there
The slope gives an
years. This simple
eventually separate
represent the diverg
tance between the
separate more than
basic 3:2 orbital re
are two prescriptio
calculation. One w

Fig. 18. The logarithm
growth indicates expon
simple two-trajectory

› display a 34 million year

ʼed in Fig. 15. The 34
[Fig. 6) is the longest
he argument of perihe-
ntity $h = e \sin \varpi$, is also
ne inclination was not
nponent with a period

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



37 million year modulation.

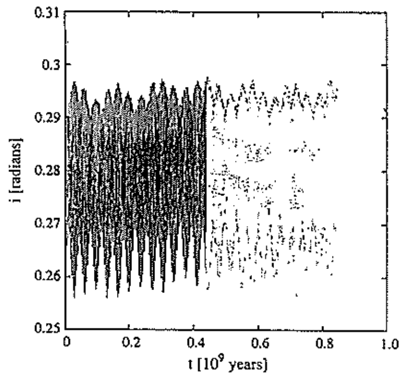


Fig. 17. The inclination of Pluto varies somewhat irregularly.

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

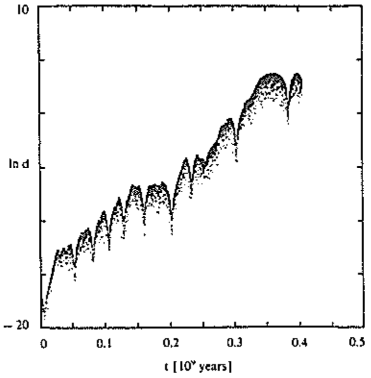


Fig. 18. The logarithm of the distance between two "Plutos" is plotted vs time. The linear growth indicates exponential divergence with a time scale of about 20 million years. The simple two-trajectory method saturates at a distance of about 45 AU.

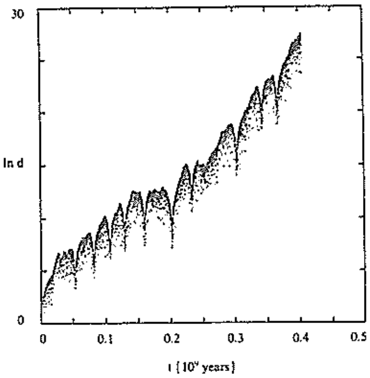


Fig. 19. The logarithm of the distance between neighboring Plutos computed with the linearized variational equations.

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 n
are in excellent ag

ments on a sys
similar behavio
with the square
that when the s
gence begins to
gence acts as a
illustrates the c
tion during a ca
ization is repea

Fig. 21. The conv
positive exponent

tos computed with the

;ht back to be near
ase space) between
ry was moved. The
renormalize. In a
that too frequent
method is to inte-
the equations of
distance grows so
a problem not en-
earby Plutos com-
vith the two Pluto
differ significantly
ıration. Figure 21
calculation of the
off at a positive
culations. The in-
years. Thus our
ic, with a surpris-

of nearby trajecto-
d versus log *t* plot
osed to show the
0 million year e-
vith distance pro-
ws that the initial
Numerical experi-

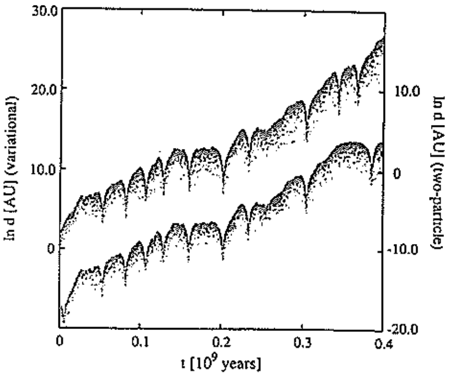


Fig. 20. The two methods for computing the divergence of trajectories give results which are in excellent agreement up to the point of saturation.

ments on a system of Plutos without planetary perturbations showed a similar behavior for the divergence. (This divergence can be contrasted with the square-law behavior observed for the round trip errors.) We see that when the slope of the exponential exceeds the power-law the divergence begins to follow the exponential. Thus the initial power-law divergence acts as a "seed" for the later exponential divergence. This plot illustrates the disaster that could result from too frequent renormalization during a calculation of a small Lyapunov exponent. If the renormalization is repeatedly carried out during the power-law phase of the di-

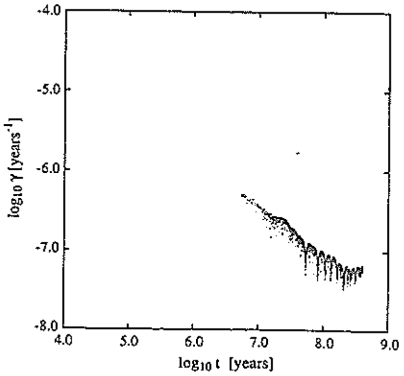


Fig. 21. The conventional representation of a Lyapunov calculation clearly indicates a positive exponent with a divergence time scale of about 20 million years.