Eur. J. Phys. 17 (1996) 295–302. Printed in the UK


# Chaos in the three-body problem: the Sitnikov case

**Daniel Fernández Hevia† and Antonio F Rañada‡**

Departamento de Física Teórica, Universidad Complutense, 28040 Madrid, Spain

Received 28 December 1995, in final form 1 January 1996

**Abstract.** We consider the Sitnikov problem, an especially simple case of the restricted three-body problem, which becomes highly chaotic as the value of a certain parameter is increased, the signs of chaos appearing neatly. We argue that it provides a very clear example for introducing students to non-integrability and chaos.

**Resumen.** Consideramos el problema de Sitnikov, un caso especialmente simple del problema de los tres cuerpos, que se hace altamente caótico al aumentar el valor de un cierto parámetro, apareciendo nítidamente los signos del caos. Creemos que ofrece un ejemplo muy claro para introducir a los estudiantes en la no integrabilidad y el caos.

## 1. Introduction

In standard courses on analytical dynamics, the Liouville theorem plays an important role: it states that an $N$ degrees of freedom Hamiltonian system is completely integrable if it has $N$ functionally independent constants of motion in involution (i.e. such that the Poisson bracket of any two of them vanishes). Moreover, the celebrated Noether theorem gives a correspondence between such constants and the symmetries of the system, provided that the latter are known. In this way, the absence of enough symmetry is known to be the key fact that breaks the integrability and leads to irregular behaviours in dynamical systems. It is customary to mention, as an example of a system without enough symmetry, and hence non-integrable, the very famous three-body problem. However, at undergraduate level, explicit examples of the irregular dynamics due to its non-integrability are too complex to be considered.

We will discuss here an especially simple case of this problem, called the Sitnikov problem, explicitly showing the chaotic behaviour it displays through the so-called 'indicators of chaos' (phase space orbits, power spectrum, autocorrelation function, Liapunov exponent), and interpreting the results in the light of the Kolmogorov–Arnold–Moser (KAM) theorem which we briefly outline in section 2. In section 3 we introduce the Sitnikov system and its main features. In section 4 we present the results of the numerical computations and, finally, we conclude in section 5 that the system is suitable for undergraduate courses in theoretical mechanics.

Here lies the interest of this problem: it allows one to get the taste of the general three-body case with a simple one-dimensional dynamical system, its study being clearly accesible to beginners in numerical methods.

† Now at Departamento de Física, Universidad de Oviedo, Oviedo, Spain.

‡ Author to whom correspondence should be addressed.

## 2. The KAM theorem

One of the most celebrated dynamical results of the last decades is the KAM theorem but unfortunately it is too difficult for undergraduate students. We will sketch it here in a very simple fashion, leaving aside the technical details. This is enough for our purposes; a more rigourous formulation can be found in [1, 2].

Let $H_0$ be a completely integrable $N$-dimensional Hamiltonian, $H_1$ a perturbation, and $\lambda$ a small parameter. Consider the Hamiltonian $H = H_0 + \lambda H_1$. As $H_0$ is integrable, action-angle variables do exist and the unperturbed motion in phase space takes place in $N$-dimensional embedded tori. The perturbation usually breaks some symmetry and renders the problem non-integrable.

The KAM theorem gives us detailed information about what happens: some of the original tori still exist, still containing regular phase space orbits, but some others are destroyed by the perturbation, the corresponding orbits being free to explore higher-dimensional regions of phase space (always bounded, if $N \leq 2$, by the conserved tori). The number of conserved tori is smaller, the larger the perturbation.

As most of the time the systems are assumed to be autonomous, it seems convenient to explain briefly how to treat the non-autonomous case. Let $H_0$ be

0143-0807/96/050295+08\$19.50 

a completely integrable one-dimensional Hamiltonian with action-angle variables $(I, \varphi)$, for which the phase space orbits are closed and periodic, each one with its frequency $\omega$. In the $(I, \varphi, t)$ space, the orbits are contained in embedded invariant cylinders parallel to the $t$ axis. Now, suppose we add a small perturbation $H_1$, i.e. that the Hamiltonian changes to $H = H_0 + \lambda H_1$. We are interested in the case in which $H_1$ is periodic in time. If the system is non-degenerate, that is if the condition $\partial^2 H_0/\partial I^2 \neq 0$ is fulfilled, the KAM theorem guarantees the conservation of all those cylinders the frequency of which is irrational enough, in the sense that for all integers $n$ and $m$

$$\left|\omega - \frac{m}{n}\right| > \frac{k(\lambda)}{n^{2.5}}, \qquad \lim_{\lambda\to 0} k(\lambda) = 0.$$

These cylinders may be, however, distorted.

We will see that our numerical computations display a very clear image of this structure.

## 3. The Sitnikov problem

The system consists of two equal-mass primary stars ($m_1 = m_2 = m/2$) moving in Keplerian ellipses of eccentricity $e$ ($0 \leq e < 1$) in the $xy$ plane, while a test body of negligible mass (called 'the planetoid') moves along the $z$ axis [3,4]. The origin of coordinates has been taken as the centre of mass of the primaries. It is easy to see [5] that the motion of the planetoid remains forever in the $z$ axis.

We will use units in which the total mass of the primaries is $m = 1$, its period is $2\pi$, and the gravitational constant $G = 1$. The motion of the test body is one-dimensional, and constitutes a one-dimensional dynamical system.

Let $r(t)$ be the distance of the primaries to the origin. Then it is a simple exercise to see that the Hamiltonian function by unit mass of the planetoid is

$$H(z, v, t) = \frac{v^2}{2} - \frac{1}{\sqrt{z^2 + r^2(t)}}. \qquad (1)$$

Note that, from well known results in the two-body problem [12], and taking into account our choice of units, we have

$$2r(\varphi) = \frac{1 - e^2}{1 + e\cos\varphi} = 1 - e\cos\varphi + \mathrm{O}(e^2).$$

We can eliminate the polar angle $\varphi$ as a function of $t$ from

$$t = \int_0^{\varphi} \frac{4r^2(\varphi')}{\sqrt{1 - e^2}}\, \mathrm{d}\varphi'$$

and, hence, performing the integral (discarding terms of higher orders in $e$) we arrive finally at

$$r(t) = \tfrac{1}{2}(1 - e\cos t) + \mathrm{O}(e^2), \qquad (2)$$

so, taking (2) into account, we may write (1) up to first order in the eccentricity as

$$\begin{aligned} H(z, v, t) &= H_0(z, v) + eH_1(z, v, t) \\ &= \frac{v^2}{2} - \frac{1}{\sqrt{z^2 + 1/4}} - e\frac{\cos t}{4(z^2 + 1/4)^{3/2}}. \end{aligned} \qquad (3)$$

Note that $H(z, v, t) = H(z, v, t + 2\pi)$, i.e. the Hamiltonian has period $2\pi$ in time, which is an important technical detail in the rigourous analytical studies of the problem [1,2] as it allows one to apply the KAM theorem. In the numerical calculations we have restricted ourselves to small values of the eccentricity (that, as we see, acts as perturbation parameter $\lambda = e$), $e < 0.08$, so that (3) will be our Hamiltonian henceforth.

The Hamiltonian equations of motion are

$$\dot{z} = v, \qquad \dot{v} = -\frac{8z}{(4z^2 + 1)^{3/2}} - e\frac{24z}{(4z^2 + 1)^{5/2}}\cos t. \qquad (4)$$

For $e = 0$, we see that $H$ is time-independent (hence invariant under time translation), the energy $H = E$ being so conserved (which is just a consequence of the aforementioned symmetry through Noether's theorem). It is thus a one-dimensional autonomous system with one constant of motion and it is then completely integrable. In fact its complete solution is expressable in terms of elliptic integrals [5,6], but for $e > 0$ this symmetry is destroyed, because $H$ takes an explicit dependence on time. Neither the energy, nor any other quantity is now conserved thus rendering the problem non-integrable.

Now, by means of the KAM theorem, we can make predictions about the expected behaviour of this system. In order to visualize the situation better, we take time as a third coordinate of phase space, so that the solutions of Hamilton's equations for the unperturbed motion are contained in invariant embedded cylinders. When we go to the perturbed motion, i.e. to small but non-vanishing values of $e$, some of these cylinders will remain, but others will be destroyed giving place to erratic solutions bounded by the conserved ones. This is exactly the picture we find in the next section when looking at phase space orbits.

Numerical studies of certain aspects of the Sitnikov problem appear in [7,8]. For a modern formulation and some new results on this problem see [9,10].

## 4. The chaos indicators

Here we assume some acquaintance with the most commonly used indicators, as they are described in standard texts, [2,11,14] so we do not repeat here a detailed description.

The equations of motion were integrated using three different methods, varying the integration steps in order to be sure of the accuracy of the results. The methods employed were a typical fourth-order Runge–Kutta approach, the Adams–Moulton predictor–corrector [15] and the very accurate method of Bulirsch and Stoer [16].

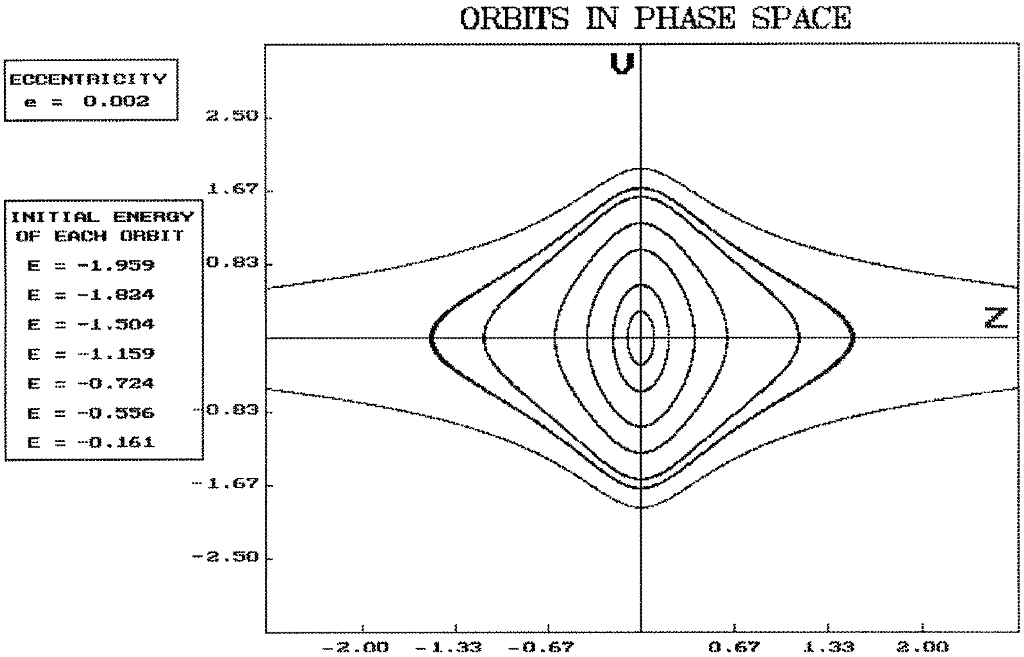


**Figure 1.** Phase portrait for $e = 0.002$.

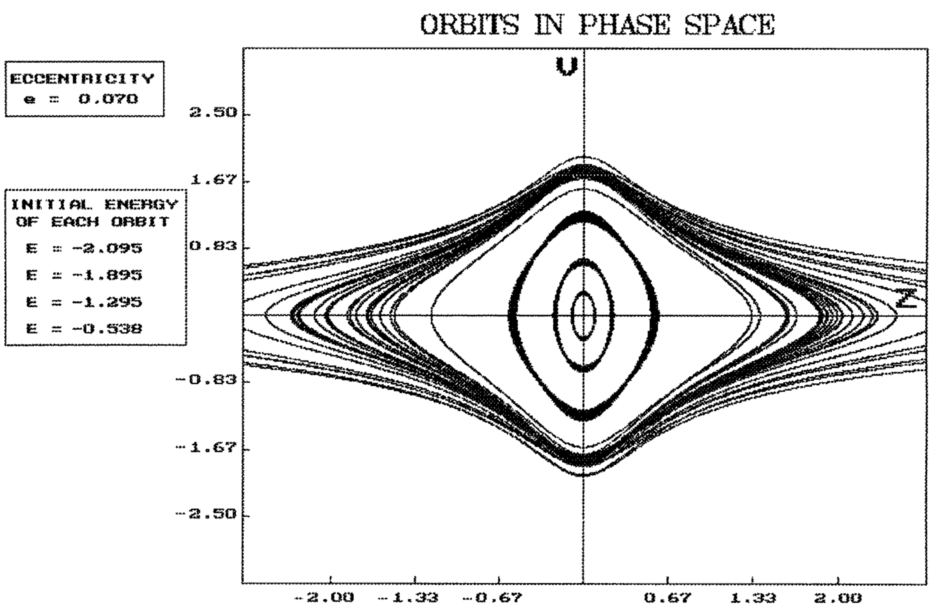


**Figure 2.** Phase portrait for $e = 0.07$.

### 4.1. Orbits in phase space

Remember that we are allowing time to act as a third coordinate of phase space. In autonomous systems with more than one degree of freedom, it is customary to study the Poincaré map, defined by means of a suitable section of the phase space (see [11–14]). However, it is easier and equally clear in this one-degree-of-freedom non-autonomous case to use instead the stroboscopic

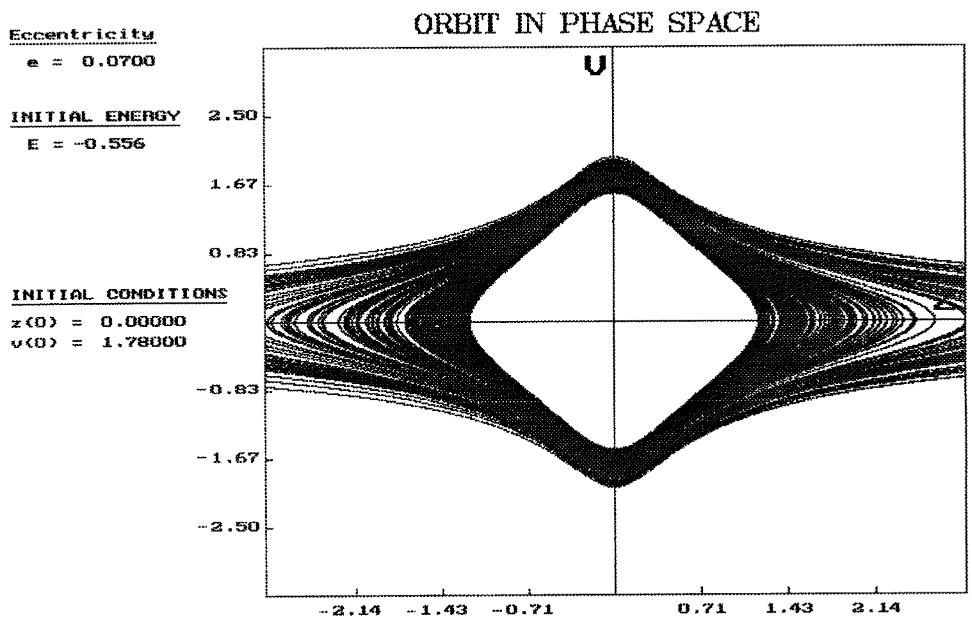


**Figure 3.** Chaotic trajectory in phase space.

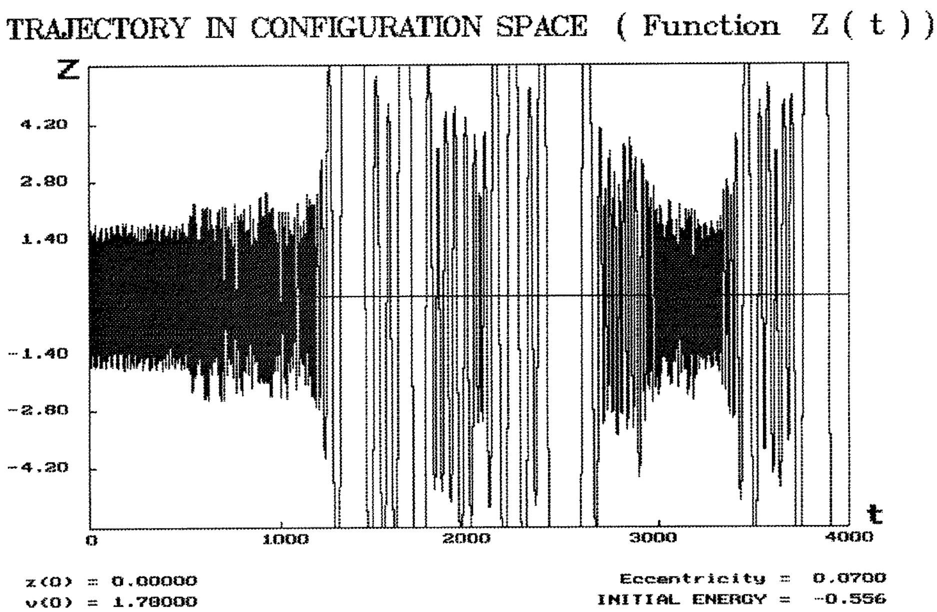


**Figure 4.** Representative erratic trajectory in the chaotic regime.

projection, obtained by plotting the position $z$ and the velocity $v$ of the planetoid (corresponding to a certain solution) for time instants $t_n = 2n\pi$.

For $e = 0$, the motion occurs in invariant cylinders that project over closed curves of the plane $(z, v)$. The origin corresponds to a stable equilibrium point with a minimum energy $E = -2$, and it is surrounded by periodic and bounded orbits of negative increasing energy until we arrive at the orbit corresponding to $E = 0$ which is the separatrix between bounded and unbounded motions of the planetoid.

For $e > 0$ this structure is clearly broken. This is barely visible in figure 1, for the very small value of $e = 0.002$, but it is clearly seen in figure 2, which represents only four solutions with the higher value of the eccentricity $e = 0.07$: just a few invariant cylinders survive and between them the orbits exibit a characteristic area filling property. Now look at figure 3: it shows only one solution, with initial energy close to zero (where even the unperturbed ($e = 0$) motion is unstable, because, in the zone in phase space close to the $E = 0$ separatrix, even the smallest perturbation may change drastically the nature of the motion, so rendering unbounded an originally bounded trajectory), in greater detail. No conserved tori bound the orbit now and it is even possible for the planetoid to escape the system and arrive at infinity, as originally proved by Sitnikov [3]. The erratic nature of the orbit is clear from the figure.

### 4.2. Trajectories in configuration space

As we are dealing with a one-degree-of-freedom system, it is possible to explicitly display the motion $z(t)$ of the planetoid. In figure 4 we see instead a solution computed for $e > 0$; even when no definitive conclusions could be made just by looking at this figure, no regularity is observed here and the motion seems highly erratic.

### 4.3. Power spectrum

Apparently irregular trajectories in configuration or phase space may still hide an underlying regular multiperiodic behaviour. The power spectrum of this solution is a well known way of clarifying this point discerning between complicated multiperiodic motions and the truly chaotic ones.

Let $\{z_j\} = \{z(j\Delta t), j = 1, \ldots, n\}$ be a discretization of the trajectory (called 'the signal'), obtained by numerical integration over a total time lapse $t_{\max} = n\Delta t$, $\Delta t$ being a fixed time interval, the integration step for instance. Define the discrete Fourier transform of the signal as [13]

$$\tilde{z}_k = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} z_j \, \mathrm{e}^{-2\pi \mathrm{i} kj/n},$$

which can be interpreted as a discretization of a function of the frequency, $\tilde{z}_k = \tilde{z}(k\Delta f)$, with $\Delta f = 2\pi/t_{\max}$. The power spectrum is defined then as $E_k = |\tilde{z}_k|^2$. It is, therefore, a function of $k$ (i.e. of the frequency), its peaks giving the frequencies of the motion.

In figure 5, we see the power spectrum (scaled logarithmically) of a solution when $e = 0$, which clearly shows only one fundamental frequency and its second harmonic. Comparing it with figure 6, corresponding to the same solution as figure 6, the aperiodic nature of the motion is quite clear: it displays a continuous spectrum of frequencies.

### 4.4. Autocorrelation functions

Intimately related to the power spectrum, these functions most clearly exhibit the loss of information along the trajectory. Let $\langle z \rangle = \sum_1^n z_k/n$ be the mean value of the signal. The autocorrelation function of the signal $C_m$ is defined as

$$C(m\Delta t) \equiv C_m = \frac{1}{n} \sum_{i=1}^{n} z'_i z'_{i+m}, \qquad z'_i = z_i - \langle z \rangle.$$

The function $C_m$ gives a measure of how much $z_i - \langle z \rangle$ (the difference between the signal and its mean value) keeps a memory of its value after an interval $m\Delta t$ of time. If the signal changes only slightly over $m$ steps, all the terms in the sum will have the same sign and the correlation will be high, $C_m$ being large. If, on the other hand, the signal changes erratically, the summands will cancel and $C_m$ will be small. This explains why the autocorrelation function of chaotic systems decreases very fast to zero with increasing $m$, while in the regular case the behaviour is very different, without going to zero or being even periodic.

The autocorrelation function of a typical periodic $e = 0$ solution is itself periodic and no loss of information occurs. However in figure 7, corresponding to the same chaotic solution already considered, the function tends clearly to zero after some time, revealing that information is lost.

### 4.5. Lyapunov exponents

For regular systems, the rate of separation of initially nearby trajectories (with close initial conditions) is linear, so that any error in the initial conditions is not greatly magnified along the evolution. But for chaotic systems this rate may be exponential, generating the so-called strong sensitivity to the initial conditions (SSIC), which is often taken [17] as the very definition of chaotic dynamics. When SSIC appears, the lack of infinite accuracy in the determination of the initial conditions leads to the impossibility of making accurate predictions about the behaviour of the system.

If we consider two close initial conditions, separated by an initial distance $D_0$, and let the corresponding solutions evolve, then, after a certain time, the distance between them behaves, for chaotic systems in which SSIC is present, as

$$d = d_0 \, \mathrm{e}^{Kt},$$

where $K$ is the leading Lyapunov exponent. In this way, a numerically computed positive Lyapunov exponent is taken to be a very good indicator of the existence of SSIC and, hence, of the chaotic nature of the motion [18, 19].

To compute the Lyapunov exponent, we have used here the algorithm of Benettin *et al* [18]: take an initial condition $(z_0, v_0)$ and a certain time lapse $\tau = k\Delta t$ (for some $k$) and form the sequence $\{(z_0, v_0), (z_\tau, v_\tau), (z_{2\tau}, v_{2\tau}), \ldots\}$ (taken from the output of the numerical integration). Another initial condition $(z'_0, v'_0)$ is then chosen (at a small distance $d_0$ in phase

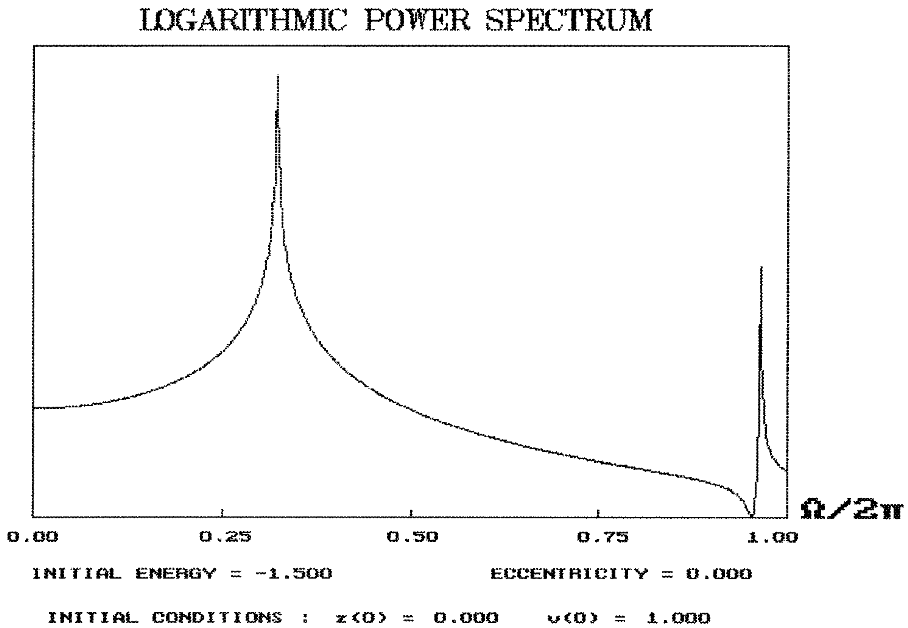


**Figure 5.** Power spectrum of a periodic trajectory for $e = 0$.

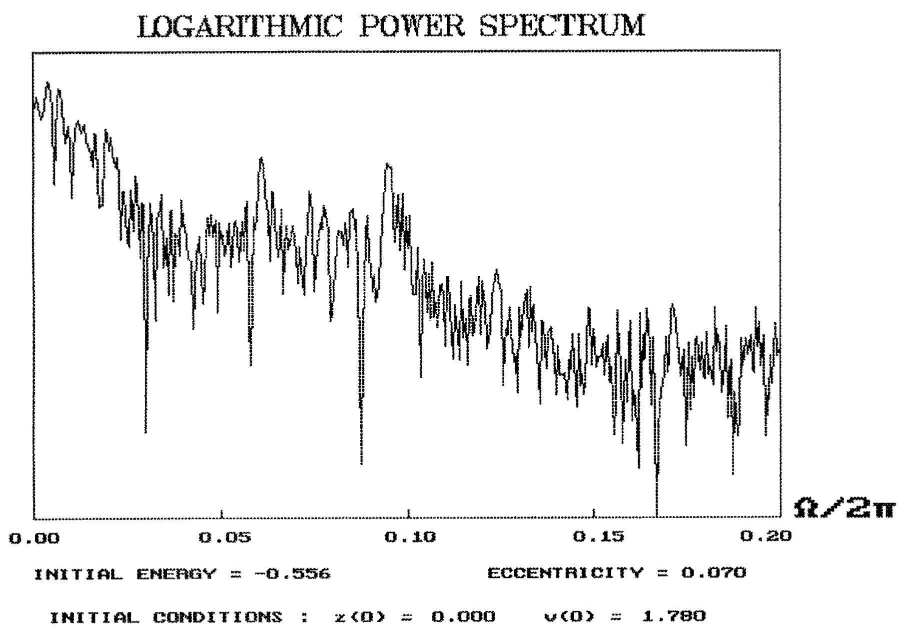


**Figure 6.** Power spectrum for the solution of figure 4.

space from the first one). The corresponding solution is integrated during the interval $\tau$ until the value $(z'_\tau, v'_\tau)$, computing then the distance $d_1 = \mathrm{dist}[(z_\tau, v_\tau), (z'_\tau, v'_\tau)]$. The second trajectory then approaches to the first one, by taking a point in the line joining $(z_\tau, v_\tau)$ and $(z'_\tau, v'_\tau)$, such that its distance to $(z_\tau, v_\tau)$ is again $d_0$. This point is taken as the initial condition of another trajectory which is integrated during the time $\tau$ to obtain the point $(z''_\tau, v''_\tau)$

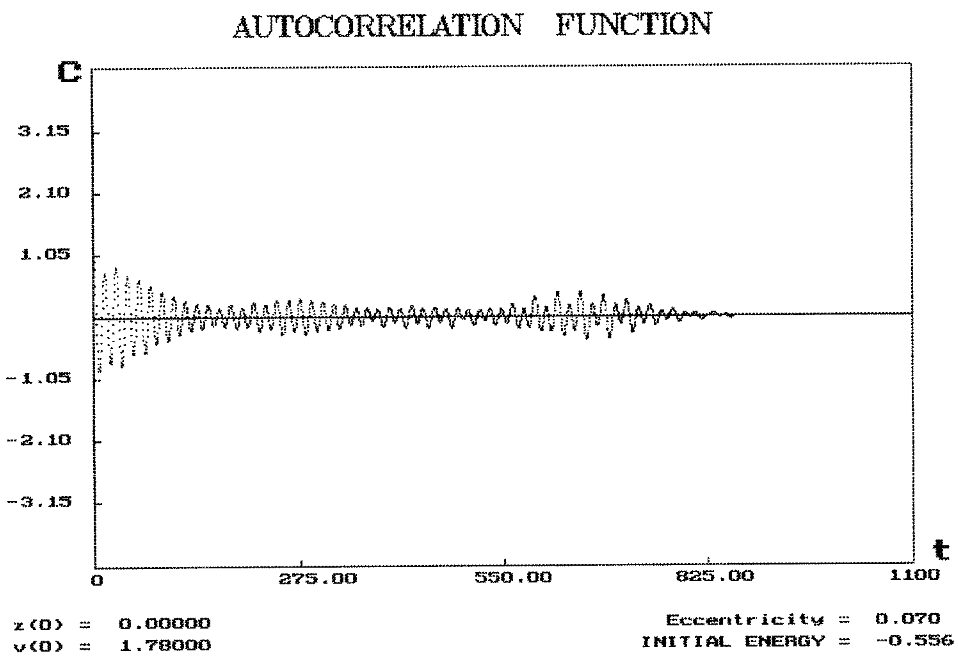


**Figure 7.** Autocorrelation function of the solution of figure 4.

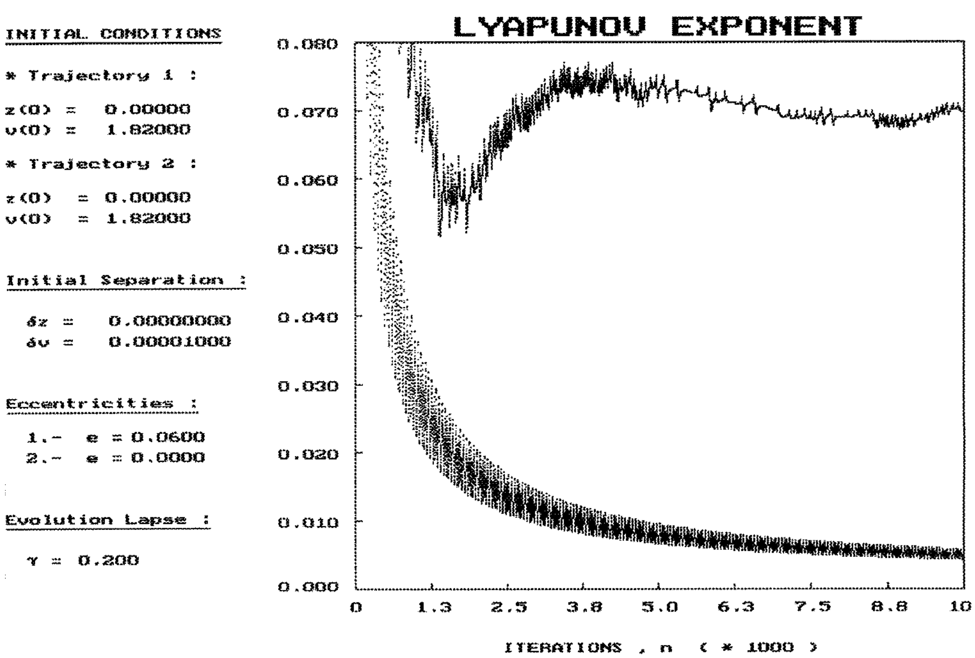


**Figure 8.** Liapunov numbers for two solutions with the same initial conditions, one with $e = 0$, the other with $e = 0.06$.

and define $d_2 = \mathrm{dist}[(z_2\tau, v_2\tau), (z''_\tau, v''_\tau)]$. The process of integrating during time $\tau$ and approaching to the first trajectory is iterated, obtaining therefore, a sequence of distances $\{d_i\}$, after which the numbers

$$k_p = \frac{1}{p\tau}\sum_{i=1}^{p} \ln \frac{d_i}{d_0},$$

are computed. The important result is that this sequence converges and that

$$K = \lim_{p\to\infty} k_p,$$

where $K$ is the Lyapunov exponent. It is clear from the explanation of the algorithm that it measures how fast a trajectory separates from the neighbouring ones. If $K > 0$, the separation grows exponentially, which is a sign of chaos. We show in figure 8 the evolution of $k_p$, which clearly converges to values of $K$.

We have made the calculation for the same initial conditions but with $e = 0$ the first time and $e > 0$ the second. The difference between the two cases is apparent: when $e = 0$ the exponent tends clearly to zero indicating that no chaos is present, but when $e > 0$ this parameter converges to a positive non-null value, hence confirming the existence of chaotic dynamics.

## 5. Conclusions

We have performed a study of the onset of chaotic motion in the Sitnikov problem and of the general features of this system as predicted by the KAM theorem.

At undergraduate level, it is clearly impossible to present detailed and rigourous analytical studies of nonlinear systems. This applies, in particular, to the three-body system, which is unfortunate in view of its enormous importance in the development of dynamical ideas as the first case of chaos ever considered. It is, therefore, desirable to find simple and accesible ways to teach chaotic dynamics to these students, and this is where numerical experiments with the indicators of chaos and some applications of the KAM theorem may be very useful.

The Sitnikov problem seems an attractive example for exploring and understanding the world of chaos.

Moreover, it has the additional interest of presenting a one-degree-of-freedom non-autonomous system, when most of the examples of chaoticity appearing in the literature are higher dimensional but autonomous. We think that it may be useful for undergraduate courses in classical mechanics.

## References


[1] Berry M 1983 Regular and irregular motion *Topics in Nonlinear Dynamics (Conf. Proc. 46)* ed S Jorna (New York: American Institute for Physics) pp 16–120

[2] Lichtenberg A J and Lieberman M A 1983 *Regular and Stochastic Motion* (Berlin: Springer)

[3] Sitnikov K 1960 *Dokl. Akad. Nauk. USSR* **133** 303–6

[4] Moser J 1973 *Stable and Random Motions in Dynamical Systems (Ann. Math. Studies 77)* (Princeton, NJ: Princeton University Press)

[5] Loarte A, Serrano J M, Mena G A and Morón C 1986 *Movimiento Caótico. Problema de Sitnikov* (Madrid: Trabajos del Seminario de Mecánica, Universidad Complutense)

[6] Tabor M 1988 *Chaos and Integrability in Nonlinear Dynamics* (New York: Wiley–Interscience).

[7] Dvorak K 1993 *Celestial Mech. Dyn. Astron.* **56** 71–80

[8] Hagel J and Trenkler T 1993 *Celestial Mech. Dyn. Astron.* **56** 81–98

[9] Alfaro J M and Chiralt C 1993 *Celestial Mech. Dyn. Astron.* **55** 351–67

[10] Hagel J 1992 *Celestial Mech. Dyn. Astron.* **53** 267–92

[11] Rañada A F 1988 Phenomenology of chaotic motion *Methods and Applications of Nonlinear Dynamics* ed A W Saenz (Singapore: World Scientific) pp 1–93

[12] Rañada A F 1990 *Dinámica Clásica* (Madrid: Alianza)

[13] Bergé P, Pomeau Y and Vidal Ch 1990 *L'Ordre dans le Chaos* (Paris: Hermann)

[14] Schuster H G 1989 *Deterministic Chaos* (Weinheim: VCH)

[15] Ralston A 1965 *A First Course in Numerical Analysis* (New York: McGraw-Hill)

[16] Bulirsch R and Stoer J 1966 Numerical treatment of ordinary differential equations by extrapolation methods *Num. Math.* **8** 1–13

[17] Wiggins S 1988 *Global Bifurcations and Chaos* (Berlin: Springer)

[18] Benettin G, Galgani L and Strelcyn J M 1976 *Phys. Rev.* A **14** 2338–45

[19] Benettin G and Strelcyn J M 1978 *Phys. Rev.* A **17** 773–84