The Sitnikov problem

Jan Michálek & Ole Gunnar Røsholt Hovland

December 2023

Abstract

This paper, working as a supplement to the lecture given by Michálek and Hovland on the 19th of December 2023, delves into the intricate dynamics of the three-body problem, with a particular focus on one of its constrained variants known as the Sitnikov problem.

Initially, the three-body problem is introduced in a broad context, establishing its fundamental principles and demonstrating its chaotic nature through a series of detailed plots. Subsequently, the discussion narrows to the Sitnikov problem, elaborating on its unique formulation. The paper explores the solution behavior under specific conditions, gradually transitioning to a more generalized analysis that considers the eccentricity of the primary bodies' orbits. Drawing upon the work "Stable and Random Motions" by Jürgen Moser[2], the paper provides a comprehensive review of the system's behavior, highlighting the utility of the shift map as a tool in understanding the dynamics at play. Concluding the paper, additional plots are presented, which shed light on the Sitnikov problem's distinct characteristics.

1 Three body problem - TBP

When studying celestial mechanics, one of the most famous problems is the *Three-body problem* (more generally *N-body problem)*. The goal is to compute the trajectories of the three bodies in free fall, given their initial position and momentum. For simplicity, we consider point masses, and the trajectories are computed using Newton's laws of motion.

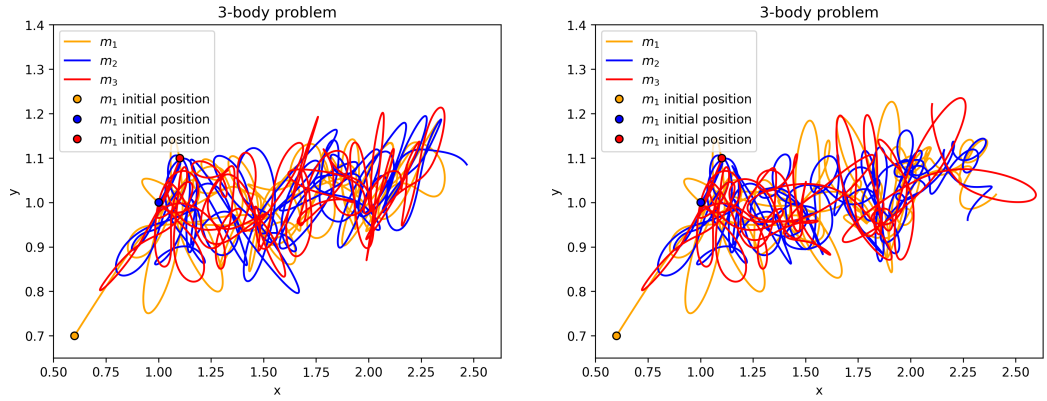
For only two bodies this problem has well-defined solutions. Things get more complicated when we introduce the third body. We can set the problem in the following way:

Problem 1. Consider three bodies with masses m_1, m_2, and m_3, and positions $\mathbf{r}_1, \mathbf{r}_2$, and $\mathbf{r}_3$ respectively. The motion of each body is governed by Newton's law of universal gravitation and Newton's laws of motion. The differential equations for the motions of the three bodies are:

$$
\begin{aligned}
m_1\ddot{\mathbf{r}}_1 &= G\frac{m_1 m_2}{|\mathbf{r}_2-\mathbf{r}_1|^3}(\mathbf{r}_2-\mathbf{r}_1) + G\frac{m_1 m_3}{|\mathbf{r}_3-\mathbf{r}_1|^3}(\mathbf{r}_3-\mathbf{r}_1),\\
m_2\ddot{\mathbf{r}}_2 &= G\frac{m_2 m_1}{|\mathbf{r}_1-\mathbf{r}_2|^3}(\mathbf{r}_1-\mathbf{r}_2) + G\frac{m_2 m_3}{|\mathbf{r}_3-\mathbf{r}_2|^3}(\mathbf{r}_3-\mathbf{r}_2),\\
m_3\ddot{\mathbf{r}}_3 &= G\frac{m_3 m_1}{|\mathbf{r}_1-\mathbf{r}_3|^3}(\mathbf{r}_1-\mathbf{r}_3) + G\frac{m_3 m_2}{|\mathbf{r}_2-\mathbf{r}_3|^3}(\mathbf{r}_2-\mathbf{r}_3),
\end{aligned}
$$

where $\ddot{\mathbf{r}}_i$ represents the acceleration of the i-th body, and G is the gravitational constant.

Posing the problem in three dimensions we obtain a set of 9 second-order differential equations. The problem can be stated equivalently using 18 first-order differential equations. To get some more intuition on the problem, some selected initial conditions were chosen to be computed numerically. For this task, we created a general simulator capable of computing the orbits of the three bodies given any initial conditions (including the amount of mass each body has). Because of the point-like nature of the equations, if two bodies in the simulation got too close to each other, one would fly far away from the other two. To compensate for this, increasing the stability of the integration, we added a padding term to our equations. This means that there is a small repellent force near each body, that pushes them away from each other. Despite having more stable solutions, the system is extremely sensitive to initial conditions. In Figure 1, we see that even a tiny change in initial conditions leads to vastly different trajectories within a relatively short time.



(a) Simulation of the TBP

(b) Modified Simulation

Figure 1: Two plots where the only difference in initial conditions (see Table 1) is an additional speed of 10^{-6} in the z direction of m_3. The problem was simulated in three dimensions, and projected down to two dimensions for plotting purposes.

Body	x	y	z	v_x	v_y	v_z
m_1	0.6	0.7	0.7	0.04	0	0.1
m_2	1	1	1	1	10^{-6}	0
m_3	1.1	1.1	1.1	1	0.2	0.2

Table 1: Initial conditions for the three-body problem simulations in Figure 1 (a)

There exist some special classes of solutions, usually with some special pattern or periodicity. More than 1000 different analytic solutions were discovered in 2017 [1]. Despite this, we cannot give a general closed solution problem.

The TBP is therefore often reduced to the *Restricted three-body problem*, also known as the RTBP. There are many different ways of reducing the TBP, all with different advantages. A reduction we will consider is supposing that one of the bodies has zero mass. From the system of equations above we see that this means that the massless body does not impact the other two bodies. From here there are many ways of simplifying the problem further, but we will focus on the simplifications that lead us to the *Sitnikov problem*.

2 Sitnikov problem formulation - RTBP

The problem was stated by the Russian mathematician Kiril Alexandrovič Sitnikov. He was studying the celestial mechanics and stated the problem in the following way:

Problem 2 (Sitnikov problem)**.** Consider two mass points, primaries, of equal mass $m_1 = m_2 = \frac{1}{2}$ moving under Newton's law on the elliptic orbits, while the center of mass, O, stays in rest. These two mass points are moving in the plane s. Now consider a third mass point $m_3 = 0$ moving on the line L perpendicular to the plane s and going through the center of mass O. The period of two primaries is normalized to 2π and the length unit is normalized in a way, that the gravitational constant is equall to 1. Let z be the coordinate of the line L. Then the differential equation describing the motion of the third body can be written as:

$$\frac{d^2z}{dt^2} = -\frac{z}{(z^2 + r^2(t))^{\frac{3}{2}}} \tag{1}$$

where $r(t + 2\pi) > 0$ is the distance of one of the primaries from the center of mass.

The problem can be stated in more general way, but since we can always normalize and adjust coordinates we will consider this setting from the beginning. The function, that has to be described in more detail is $r(t)$. This function takes into account the eccentricity of the two primaries, denoted ϵ. For small value of $\epsilon > 0$ we can find that:

$$r(t) = \frac{1}{2}(1 - \varepsilon \cos t) + \mathcal{O}(\varepsilon^2).$$

Notice that for the $\epsilon = 0$ the problem reduces to:

$$\frac{d^2z}{dt^2} = -\frac{z}{(\sqrt{1 + z^2})^3}, \tag{2}$$

where we have circular orbits of the primaries. The visualisation of the problem is provided in Figure 2.

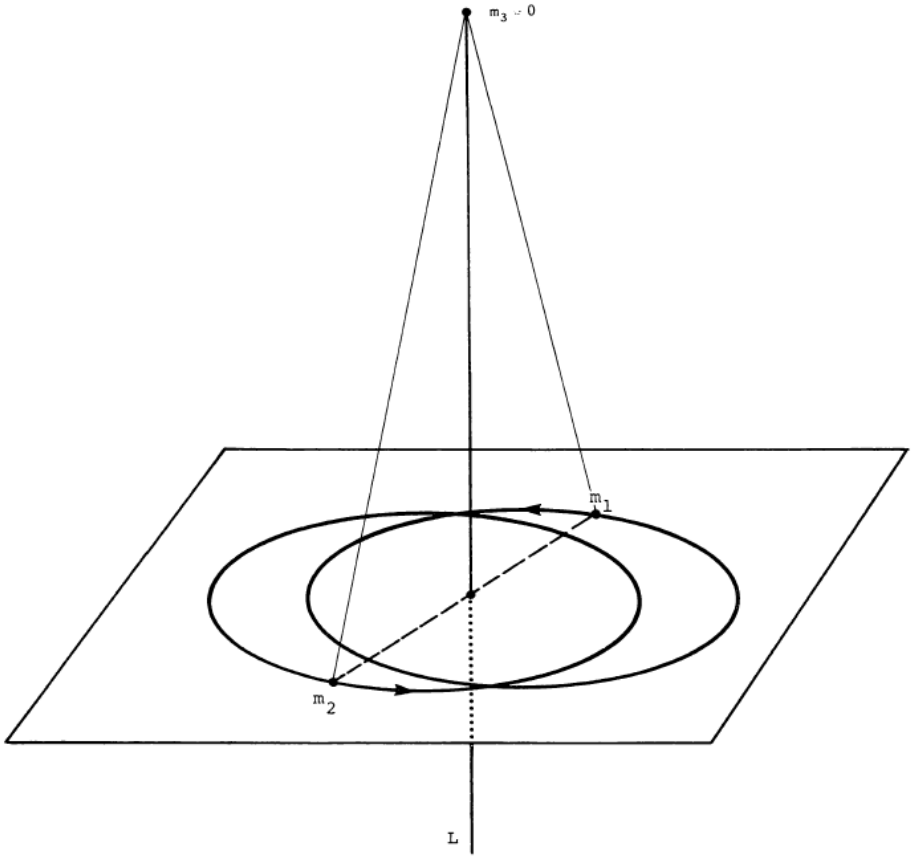


Figure 2: Restricted three body problem

3 Problem solution

To discuss the solution we can start by considering the edge cases. First assume that the initial position of the third body is near zero and same goes for the initial velocity. We can expect that such a solution should be at least bounded for all real t. On the other hand we can assume the initial speed to be large and we can intuitively expect the third body to escape the gravitational influence of the primaries. We can also imagine a third body coming from infinity and then oscillating around the origin.

In either case we need to encapsulate the oscillation in a convenient manner. We will do it in the following way. Let us assume a solution $z(t)$, then we can introduce t_k, where $k \in \mathbb{Z}$, such that $z(t_k) = 0$. Now we can introduce number s_k in the following way:

$$s_k = \left\lfloor \frac{t_{k+1} - t_k}{2\pi} \right\rfloor \tag{3}$$

Notice that 2π is the period of the two primaries and when we assume that t_k is ordered according to its size $t_k < t_{k+}$ then with the use of the floor function we can say that the number s_k says how many rotations the primaries do between the crossing of the origin plane by the third body.

Since we assume $k \in \mathbb{Z}$ we obtain a doubly infinite sequence. When the third body comes from the infinity this sequence is terminated from the left and when the third body escapes to infinity the sequence is terminated from the right. The main result can be generally stated in the following way:

Theorem 1. Given a sufficiently small eccentricity $\varepsilon > 0$ there exists an integer $m = m(\varepsilon)$ such that any sequence s with $s_k \geq m$ corresponds to a solution of the above differential equation.

4 Reduction to mapping

Every trajectory of the third body is entirely described by only two variables. First we need to set the time t_0, when the $z(t_0) = 0$, and then we need to add the velocity at this point $z'(t_0)$. We say that $v_0 = |z(t_0)|$ and also add that by t_0 we mean $t_0 \mod 2\pi$. Notice that equation 1 is invariant under the reflection over the origin plane, so we do not need to distinguish between positive and negative velocities.

Now we will interpret t_0 and v_0 as polar coordinates in the plane, where t_0 is the angle variable and v_0 is the radius. In the next step, we define mapping:

$$\phi : (v_0, t_0) \mapsto (v_1, t_1). \tag{4}$$

This map takes the solution to its next zero - if it exists. To ensure the existence of (v_1, t_1) we need to know where ϕ is defined. It can be shown that there exists an analytic simple closed curve in $\mathbb{R}^2$ in whose interior, D_0 where the mapping ϕ is defined. Outside this interior, the solutions escape the gravitational well of the other two bodies.

Lastly, for the lemma, the reflection mapping can be defined as:

$$\rho : (v, t) \mapsto (v, -t). \tag{5}$$

With these definitions, we can state the following lemma, which can help us relate these mappings.

Lemma 1. ϕ maps D_0 onto a domain $D_1 = \phi(D_0)$ which agrees with the reflected domain $D_1 = \rho(D_0)$. Moreover, ϕ preserves the area element v, dv, dt, and

$$\phi^{-1} = \rho^{-1} \circ \phi \circ \rho. \tag{6}$$

Natural question is how the eccentricity ϵ impacts the relation between D_0 and D_1. If $\epsilon = 0$ then we obtain two identical discs of radius 2.

But for $\epsilon > 0$, $D_0 \neq D_1$ and the boundary curves ∂D_0, ∂D_1 intersect on the symmetry line nontangentially at a point P, see Figure 3.

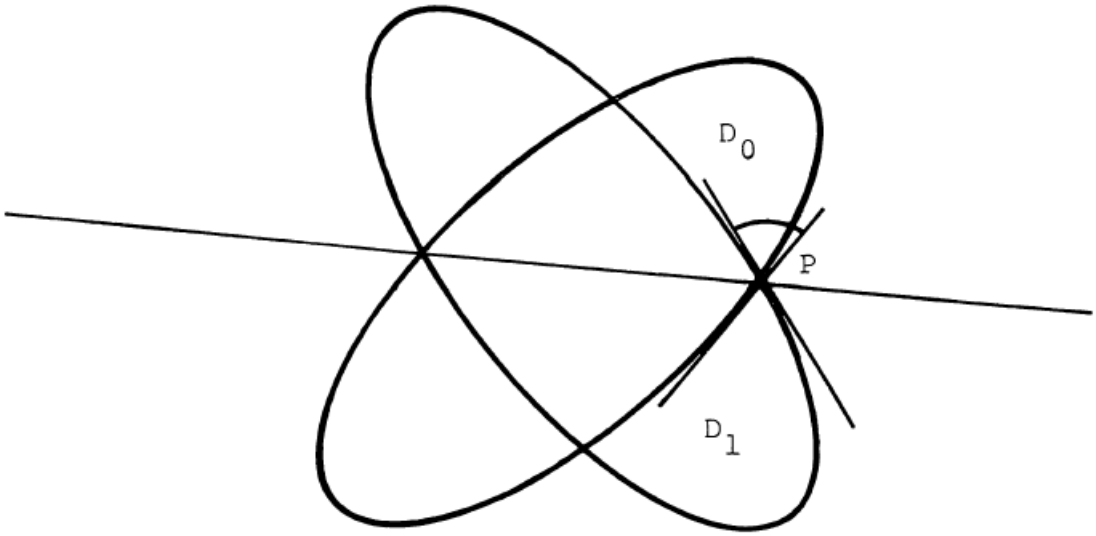


Figure 3: Intersection of the boundaries

It can be shown that we must have $v_0 < 2$ to have a bounded solution. By examining the mapping ψ for the $\epsilon = 0$ we observe that the velocity of the third body in the time of crossing the primaries plane is always the same, $v_0 = v_1$ and the time of the return is just the function of the v_0, $t_1 = t_0 + T(v_0)$. The mapping then takes the following form:

$$\begin{aligned} &\phi : (v_0, t_0) \mapsto (v_1, t_1) \\ &\phi : (v_0, t_0) \mapsto (v_0, t_0 + T(v_0)). \end{aligned} \tag{7}$$

To plot the mapping, we can imagine two concentric circles, if $v_0 < 2$. They are rotated by an angle $T(v_0)$, which tends to ∞ as v_0 approaches 2. The image of any radius is a curve spiraling faster and faster, infinitely around the origin, as it approaches the boundary. This is what is sketched in Figure 4a. The curve meets the boundary non-tangentially, meaning that it crosses the boundary and does not only touch it.

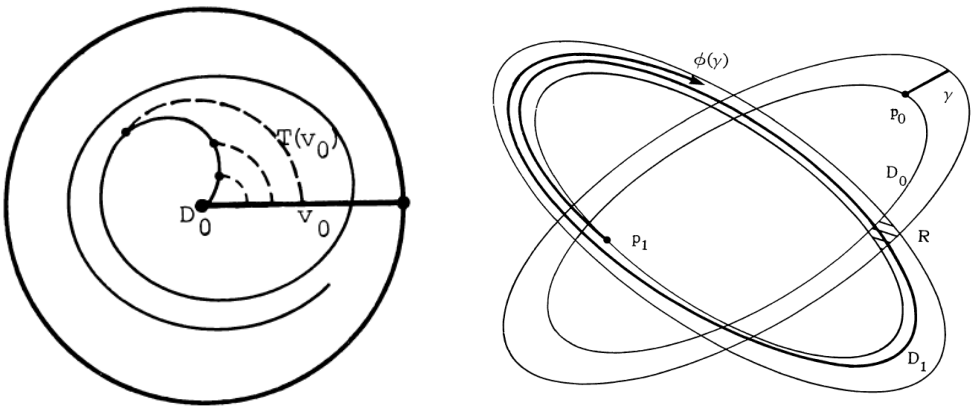
For the case when $\epsilon > 0$ the situation is similar, however now we do not have two identical discs. We parameterize the curve γ as:

$$\gamma : v_0 = v_0(\lambda), t_0 = t_0(\lambda) \qquad 0 \leq \lambda \leq 1. \tag{8}$$

This curve is C^1 and meets ∂D_0 at p, which corresponds to $\lambda = 0$. As in Figure 4a, the curve crosses the boundary non-tangentially. The image curve,

$$\phi(\gamma) : v_1 = v_1(\lambda), t_1 = t_1(\lambda), \tag{9}$$

also displays the spiraling behavior, see Figure 4b. The curve approaches the boundary ∂D_1 as $t_1(\lambda)$ approaches infinity. $t_1(\lambda) \to \infty$ as $\lambda \to 0$.



(a) (b)

Figure 4: Comparison of curves for $\epsilon = 0$ and $\epsilon > 0$

Before continuing to the next section we need to define two more necessary concepts.

Definition 1 (Bundle of Sectors Near the Boundary)**.** Since ∂D_0 is continuously differentiable. Then for a sufficiently small $\delta > 0$ we can define $D_0(\delta)$ as the set of points in D_0 whose distance from ∂D_0 is less than δ. For each point $p \in D_0(\delta)$, there exists a unique closest point $q \in \partial D_0$, provided δ is sufficiently small.

We define two bundles of sectors within $D_0(\delta)$:

- The bundle $\Sigma_0 = \Sigma_0(\delta^{\frac{1}{3}})$ assigns to each point $p \in D_0(\delta)$ a set of lines forming an angle $\leq \delta^{\frac{1}{3}}$ with the line going through p parallel to the tangent of ∂D_0 at q.
- The bundle Σ_0' assigns to each point the set of lines complementary to that of Σ_0.

With these definitions the natural question is what happens with the set D_0 and the bundles Σ_0 under the ϕ mapping. It can be shown, that ϕ transforms $D_0(\delta)$ into $D_0(\delta^\beta)$, where $0 < \beta < 1$. To map bundles we can write that $\Sigma_0 = \Sigma_0\left(\delta^{\frac{1}{3}}\right)$ is under mapping ϕ taken into $\Sigma_1 = \Sigma_1\left(\delta^{\frac{\beta}{3}}\right)$. Moreover, if we let $\zeta_0 \in \Sigma_0'$ and $\zeta_1 = d\phi(\zeta_0)$, with ξ_0 denoting the orthogonal projection of ζ_0 onto the center line of Σ_0 and ξ_1 similarly representing the orthogonal projection of ζ_1 onto the center line of Σ_1, then the following inequality holds:

$$|\xi_1| \leq \delta^{-\frac{1}{3}}|\xi_0|.$$

To get a clearer intuition on these definitions, in Figure 5 we can see a sketch of both a bundle and its complementary bundle. The figure also clearly shows how the mapping works on bundles.

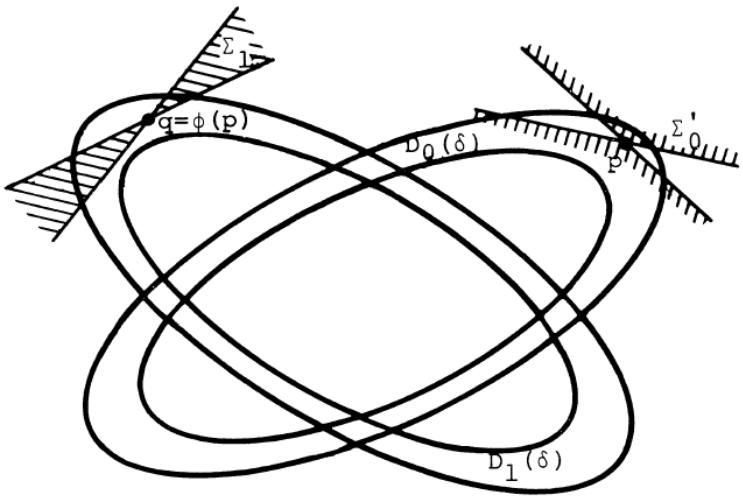


Figure 5: ϕ mapping of the $D_0(\delta)$ and Σ_0

5 The shift as subsystem of ϕ

We will now relate our system to the shift map, more exactly we will say that the mapping ϕ has a shift map as its subsystem. We first define necessary concepts.

Consider an alphabet Δ, which is a set that may be either finite or countably infinite. Using elements from this alphabet, we construct a doubly infinite sequence:

$$s = (\ldots, s_{-2}, s_{-1}, s_0, s_1, s_2, \ldots),$$

where each term s_k is a member of Δ. The collection of all such sequences is denoted by S, which is defined as:

$$S = \{s = (\ldots, s_{-2}, s_{-1}, s_0, s_1, s_2, \ldots) \,|\, s_k \in \Delta, \forall k \in \mathbb{Z}\} \tag{10}$$

and is called a *sequence space*.

We would also like to add a topology structure to this space S. This can be done in the following way.

Definition 2 (Topology on space S)**.** A topology on S is induced by specifying a basis of open neighborhoods. For a given sequence $s^* \in S$, we define the neighborhood basis as:

$$\mathcal{U}_j(s^*) = \{s \in S \,|\, s_k = s_k^*, \forall k \text{ such that } |k| < j\}, \tag{11}$$

for each $j \in \mathbb{N}$. The collection of these neighborhoods for all j and s^* generates the topology on S

This turns S into a topological space. We are adding just a short example to show how this topology works.

Example 1. Let us consider an example with an alphabet $\Sigma = \{a, b\}$. Define a specific sequence s^* in S as:

$$s^* = (\ldots, a, a, b, a, b, a, \ldots).$$

Now consider the neighborhood basis for s^* when $j = 2$. The neighborhood $\mathcal{U}_2(s^*)$ consists of all sequences in S that agree with s^* at positions -1 and 0. So in this neighborhood, all sequences must have b at position -1 and a at position 0, regardless of other positions. Some sequences in $\mathcal{U}_2(s^*)$ could be:

$$(\ldots, a, b, a, b, a, \ldots), \quad (\ldots, b, b, a, a, a, \ldots), \quad (\ldots, a, b, a, b, b, \ldots), \quad \text{etc.}$$

All these sequences share the property that they match s^* at the required positions, demonstrating what it means for two sequences to be "close" in this topological space.

With this topological space, we can introduce a shift map, that will play a key role.

Definition 3 (Shift Map)**.** The shift map σ is a function defined on a sequence space S. For a sequence $(s_k)_{k\in\mathbb{Z}}$, where each s_k is a symbol from the alphabet, the shift map σ is defined by

$$\begin{aligned} \sigma &: S \to S \\ \sigma((s_k)_{k\in\mathbb{Z}}) &= (s_{k+1})_{k\in\mathbb{Z}} \end{aligned} \tag{12}$$

That is, σ maps the sequence to another sequence where each symbol is replaced by its successor in the original sequence.

In the context of symbolic dynamics, the shift map is used to study the dynamical properties of symbolic systems, where the action of σ can be thought of as "shifting" all symbols one position to the left.

We can now define what does it mean to have shift as a subsystem.

Definition 4 (Shift as a Subsystem of a Mapping)**.** Let ϕ be a continuous mapping on a compact metric space X. A shift map σ is said to be a subsystem of ϕ if there exists a subset $Y \subseteq X$ that is invariant under ϕ (i.e., $\phi(Y) \subseteq Y$) and a homeomorphism $h : Y \to \Sigma$, where Σ is a shift space, such that the following diagram commutes:

$$\begin{array}{ccc} Y & \xrightarrow{\phi|_Y} & Y \\ \downarrow h & & \downarrow h \\ \Sigma & \xrightarrow{\sigma} & \Sigma \end{array}$$

Here, $\phi|_Y$ denotes the restriction of ϕ to Y, and the commutativity of the diagram means that $h \circ \phi|_Y = \sigma \circ h$. This implies that the dynamics of ϕ on Y are topologically conjugate to the dynamics of the shift map on Σ, making $(Y, \phi|_Y)$ a subsystem that is dynamically equivalent to the shift system.

All those definitions will now come in handy. Our system also has a shift as its subsystem. We present the following theorem, which for the sake of brevity, we are not going to prove. However, a brief overview of the proof is provided

Theorem 2. The mapping ϕ in D_0 possesses the shift $\overline{\sigma}$ on $D(\overline{\sigma}) \subset \overline{S}$ as a subsystem. Moreover, there exists a hyperbolic invariant set I homeomorphic to S on which ϕ is topologically equivalent to σ.

This theorem allows us to study the dynamics of the original system (equation 1) from a different perspective. A set of orbits can be viewed as a discrete sequence from the sequence space S. Let us now introduce some definitions needed for the actual proof.

We begin by introducing the symmetric component R of the domain $D_0(\delta) \cap D_1(\delta)$. Notice that this element contains in its closure the point P laying on the symmetry line. We consider sufficiently small δ and want to examine the boundaries of R. From Figure 6a wee notice that element R is bounded by four differentiable curves. Two of them are the boundaries ∂D_0 and ∂D_1 and other two depends on particular choice of δ.

It is important to notice that two of these curves are nontangential to ∂D_0 and from the previous results we can derive that the image of R under ϕ spirals towards ∂D_1, see Figure 6a. Thus $\phi(R)$ and R intersect in infinitely many components.

Aside from possibly finitely many of these components, they will connect opposite sides of R. For the components of $\phi(R) \cap R$, we denote them in order by $U_1, U_2, U_3, \ldots$. Similarly, we set

$$V_k = \rho U_k \qquad k = 1, 2, \ldots$$

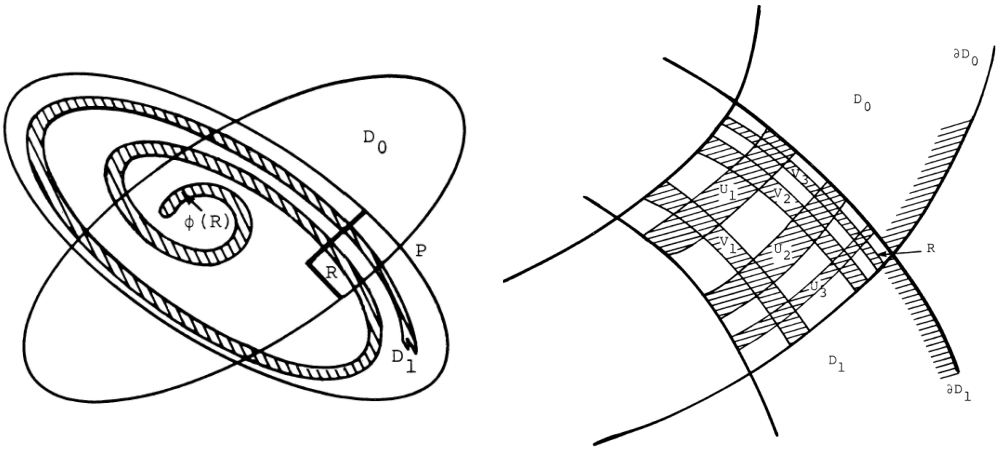
which are components of

$$\phi^{-1}(R) \cap R = \phi^{-1}(R) \cap \rho R = \rho(\phi(R) \cap R).$$

In fact we can connect U_k with V_k by

$$\phi(V_k) = U_k \qquad k = 1, 2, \ldots$$

To imagine these components see Figure 6b, which provides a detailed view on the element R.



(a) (b)

Figure 6: Intersection of R and $\phi(R)$

To prove the Thm. 2 one has to map R to the square Q and then verify that U_k and V_k satisfy the properties of the "horizontal" resp. "vertical stripes". The bundles Σ_0, and Σ_1 are perfect for this task. After showing that R maps to Q, the proof continues by obtaining the shift $\overline{\sigma}$ of the compact sequence space $\overline{S}$ as a subsystem of ϕ.

6 Chaos in the Sitnikov problem

The dynamical complexities of the Sitnikov problem lead to chaotic behavior. In this section, we will show this by connecting Tsheorem 1 and 2, and discussing how this demonstrates chaos in the system.

Theorem 1 emerges as a natural extension of Theorem 2. The domain V_k signifies the starting points of orbits with return times given by $t_1 - t_0 = 2\pi(k+c+\theta)$, with c as a constant and $\theta \in [0, 1)$. This relationship allows us to express the orbit's behavior as a sequence $s \in S$, prescribing the integral part of the return time as an arbitrary sequence $s_k + [c]$.

An orbit's trajectory, whether it escapes or is captured, is encoded in the form of its corresponding sequence. For sequences bounded for $k \leq 0$, the solution $z(t)$ is similarly bounded for $t \leq 0$, meaning it is an escape orbit. The reverse temporal solution $z(-t)$ defines a capture orbit.

An infinite amount of hyperbolic periodic orbits arises from the periodic sequences within S. This is because any periodic sequence $s \in S$ corresponds to a periodic point of $r(s)$ for ϕ. This means that every point of $r(s)$ is an initial point for a periodic orbit. The hyperbolic nature (defined in Tshm. 2) of $r(S)$ ensures that each periodic point is hyperbolic. The existence of periodic orbits can also be shown.

The set $r(S)$ is a hyperbolic invariant set, a property that is indicative of chaos. This means that the set is very sensitive to changes in initial conditions. Moreover, $r(S)$ has by definition an uncountable number of periodic points. Each periodic point is an anchor around which orbits exhibit regular, repeating patterns. However, due to the dense arrangement of these points and their associated orbits, the overall behavior of the system remains unpredictable.

In summary, the existence of the hyperbolic invariant set $r(S)$ embodies the essence of the system's chaotic dynamics through its sensitive initial conditions and the density of periodic points. These features indicate that the Sitnikov problem indeed exhibits chaotic behavior.

7 Conclusion

The investigation of the Sitnikov problem reveals an intricate phase space structure that is a telltale sign of chaotic systems. The correspondence between the elements of the sequence space S and the orbital dynamics of the system is not merely an analogy; it is a rigorous mathematical relationship that encodes the initial conditions and subsequent evolution of the system's trajectories. The presence of a hyperbolic set within the phase space demonstrates the system's sensitivity to initial conditions. This sensitivity is quantified by the exponential divergence of nearby trajectories, a property that is central to the definition of chaos.

Some geometric intuition of these chaotic dynamics is provided through the intersections of stable and unstable manifolds and will be discussed in the epilogue.

8 Epilogue

8.1 Homoclinic points

As an epilogue to our proof of chaos in the RTBP let us look at how the previously defined shift can give us some more insight into the system. We begin by introducing homoclinic and heteroclinic points. It should be noted that these points are also related to the systems' invariant manifolds.

Definition 5 (Heteroclinic point)**.** Let p, q be fixed points of a diffeomorphism ϕ. Then let r be different from these points. If

$$\phi^k(r) \to p \qquad \text{for} \quad k \to +\infty$$
$$\phi^k(r) \to q \qquad \text{for} \quad k \to -\infty$$

then r is a heteroclinic point.

The definition for a homoclinic point is even simpler as one define it in the following way.

Definition 6 (Homoclinic point)**.** Let p be a fixed point of a diffeomorphism ϕ. Then let r be different from p. If

$$\phi^k(r) \to p \qquad \text{for} \quad k \to +\infty$$
$$\phi^k(r) \to p \qquad \text{for} \quad k \to -\infty$$

then r is a homoclinic point.

Points approaching p, after applying the map iteratively, lie on the invariant curve W_p^s. Similarly, points escaping from p, after applying the map iteratively, lie on the invariant curve W_p^u. With this notation, we define heteroclinic points of p, q as an intersection of these curves:

$$W_p^u \cap W_q^s = \{p, q\}\,.$$

Homoclinic points of p can be defined as:

$$W_p^u \cap W_p^s = \{p\}\,.$$

Lets say that these curves intersect at some point r (also assuming that the crossing is transversal), then this point is homoclinic, see Figure 7.

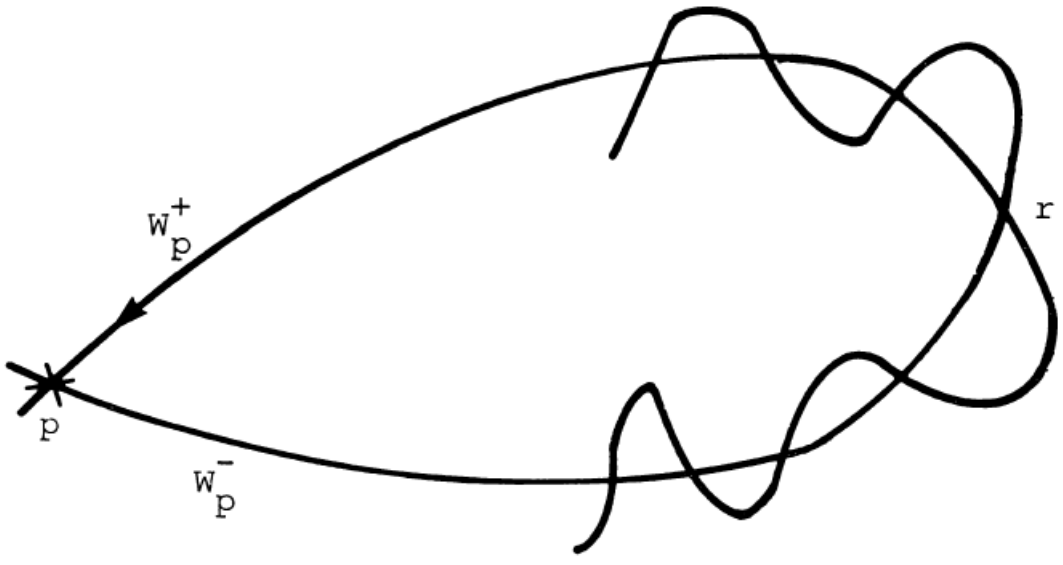


Figure 7: Homoclinic point r belonging to the hyperbolic point p

This particular setting is because the point p of the intersection of the boundaries ∂D_0, ∂D_1 defined in the previous chapters, plays the role of the homoclinic point belonging to an unstable periodic orbit at infinity.

8.2 Transversal map $\tilde{\phi}$

Consider a region R delineated in the proximity of a homoclinic point r. This region is bounded by segments of the stable (W_p^s) and unstable (W_p^u) curves at r, with the remaining boundaries being linear and parallel to the tangents at the manifolds' intersection. For a given point q in R, let $k(q)$ denote the least positive integer for which the iterate $\phi^k(q)$ resides within R, provided such an integer exists. Define the domain $D(\tilde{\phi})$ to consist of all points q in R for which an integer k is found. The map $\tilde{\phi}$ is thus characterized as follows:

$$\tilde{\phi}(q) = \phi^k(q) \quad \text{for all} \quad q \in D(\tilde{\phi}). \tag{13}$$

The obvious assumption is, that $D(\tilde{\phi}) \neq \emptyset$. This and even more is proven by the following theorem.

Theorem 3. Let ϕ be a C^∞-diffeomorphism with a homoclinic point r, and suppose the curves W_p^s and W_p^u of a hyperbolic fixed point p intersect transversally at r. Then, in any neighborhood of r, the transversal map $\tilde{\phi}$ of a quadrilateral has an invariant subset I homeomorphic to the set S (space of sequences on $N = \infty$ symbols) via a homeomorphism $\tau : S \to I$, such that the following conjugacy condition holds:

$$\tilde{\phi} \circ \tau = \tau \circ \sigma.$$

$$\begin{array}{ccc} S & \xrightarrow{\sigma} & S \\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\tau} \\ I & \xrightarrow[\tilde{\phi}]{} & I \end{array}$$

Furthermore, the homeomorphism τ can be extended to a map $\overline{\tau} : \overline{S} \to \overline{I}$ satisfying:

$$\tilde{\phi} \circ \overline{\tau} = \overline{\tau} \circ \overline{\sigma}$$

when both sides are restricted to $D(\tilde{\phi})$.

With the help of Theorem 3 we can conclude our results in the following way.

Theorem 4. The homoclinic points belonging to p are dense in I.

This result implies, that the behaviour near the intersection of W_p^s and W_p^u is very complicated, see Figure 8. Natural question is, how can we get other homoclinic or heteroclinic points. The answer is surprisingly very simple.

Homoclinic points can be obtained by taking a sequences such that their tail ends are periodic with the same block repeating at the right and the left. This means, that in both directions the orbit approaches same behaviour getting closer to the desired homoclinic point.

Heteroclinic points are obtained in the similar manner. We take sequences that posses two different repeating blocks for the right and for the left end.

Taking into account all previous results, we can also show, that the existence of one single homoclinic point imply existence of infinitely many other homoclinic and heteroclinic points.

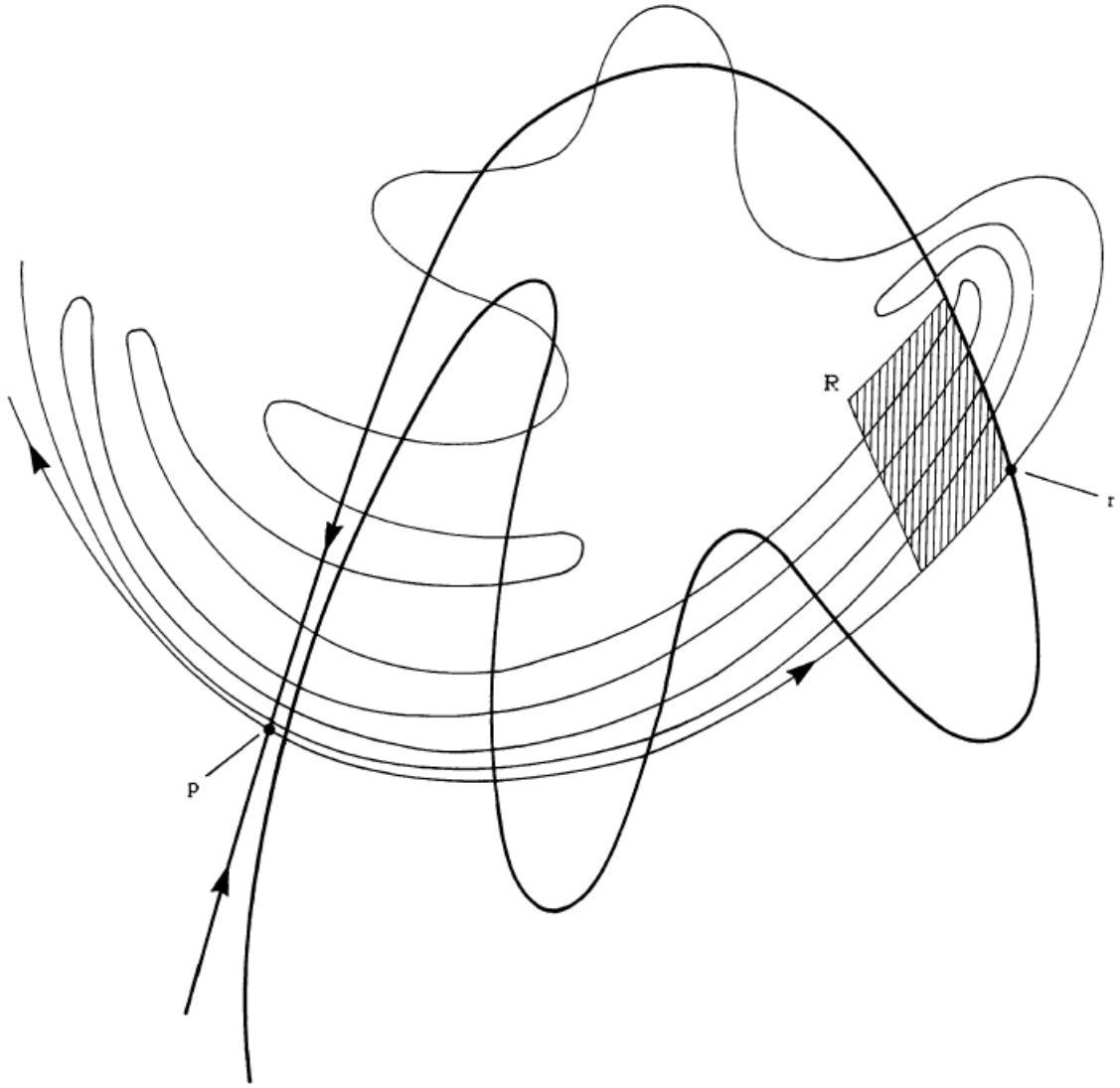


Figure 8: Set of intersections of W_p^s, W_p^u

References

[1] Xiaoming Li, Yipeng Jing, and Shijun Liao. Over a thousand new periodic orbits of a planar three-body system with unequal masses. *Publications of the Astronomical Society of Japan*, 70(4), May 2018.

[2] JÜRGEN MOSER. *Stable and Random Motions in Dynamical Systems: With Special Emphasis on Celestial Mechanics (AM-77)*. Princeton University Press, rev - revised edition, 1973.