

Dynamics, Geometry and Solar Sails[☆]

Ariadna Farrés^a, Àngel Jorba^b

^a*Institut de Matemàtica, Universitat de Barcelona,
Gran Via 585, 08007 Barcelona, Spain*

^b*Departament de Matemàtiques i Informàtica, Universitat de Barcelona,
Gran Via 585, 08007 Barcelona, Spain*

Abstract

This note focuses on some dynamical aspects of a solar sail. The first part of the paper is a survey of the use of dynamical systems tools to control a solar sail near an unstable equilibrium point of the Earth-Sun system. The second part focuses on new results on the dynamics near an equilibrium point of a sail near an asteroid. The main tool is a reduction to the centre manifold to focus on the bounded motions. In both cases, the role of the geometrical structures of the phase space is highlighted.

Keywords: Periodic orbits, centre manifolds, Halo orbits

2010 MSC: 37J15, 65P30, 70F07

1. Introduction

Dynamical systems have proven to be a useful tool for the design of space missions. For instance, the use of invariant manifolds is now common to derive control and transfer strategies. In this note we focus on a specific kind of low thrust propulsion, known as solar sailing. Solar sailing is based on the use of large membrane mirrors to take advantage of the solar radiation pressure to propel the spacecraft. Although the acceleration produced is smaller than the one achieved by a traditional chemical thruster, solar radiation pressure acts continuously and it is unlimited in time. This makes some long term missions more accessible, and opens a wide new range of possible applications that cannot be achieved by a traditional spacecraft.

Up to now, three solar sails have been successfully deployed in space: IKAROS, NanoSail-D2 and LightSail-A. IKAROS (Interplanetary Kite-craft Accelerated by Radiation Of the Sun) is a Japan Aerospace Exploration Agency experimental spacecraft with a 14×14 m² sail. The spacecraft was launched on May 21st 2010, together with Akatsuki (Venus Climate Orbiter). On December 8th 2010, IKAROS passed by Venus at about 80.800 km. NanoSail-D2 is a small solar

[☆]Work supported by the grants MTM2015-67724-P and 2014 SGR 1145.

Email addresses: ari@maia.ub.es (Ariadna Farrés), angel@maia.ub.es (Àngel Jorba)

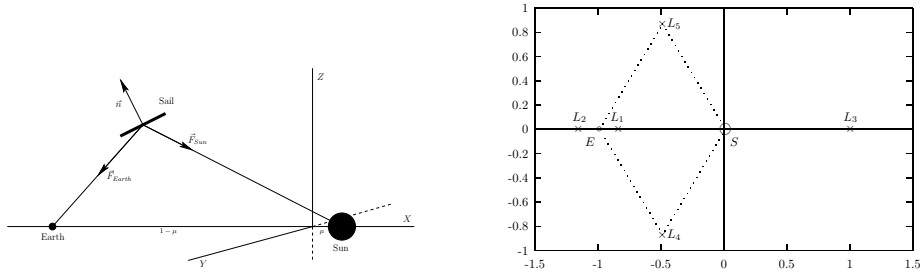


Figure 1: Left: Scheme of the forces acting on the sail. Right: The five equilibrium points of the Restricted Three-Body Problem.

sail (10 m^2 , 4 kg) deployed by NASA on January 2011 in a low Earth orbit, that reentered the atmosphere on September 17th 2011. LightSail-A is a small test spacecraft (32 m^2) of the Planetary Society, that was launched on May 20th 2015 and deployed its solar sail on June 7th 2015. It reentered the atmosphere on June 14th 2015.

In this paper we will focus on the dynamics of a solar sail in a couple of situations. We will introduce this problem focusing on a solar sail in the Earth-Sun system. In this case, the model used will be the Restricted Three Body Problem (RTBP for short) plus solar radiation pressure (see Figure 1, left). The effect of the solar radiation pressure on the RTBP produces a 2D family of “artificial” equilibria, coming from the well known equilibria of the RTBP (see Figure 1, right or [25] for more details). This new equilibria can be parametrised by the orientation of the sail. We will describe the dynamics around some of these “artificial” equilibrium points. We note that, due to the solar radiation pressure, the system is Hamiltonian only for two cases: when the sail is perpendicular to the Sun - sail line; and when the sail is aligned with the Sun - sail line (i.e., no sail effect). The main tool used to understand the dynamics is the computation of centre manifolds, for both the Hamiltonian and non-Hamiltonian cases.

The second example is the dynamics of a solar sail close to an asteroid. Note that, in this case, the effect of the sail becomes very relevant due to the low mass of the asteroid. We will use, as a model, a modified Hill problem that includes the effect of the solar radiation pressure, to describe some aspects of the natural dynamics of the sail.

The paper is organised as follows. Section 2 is a short introduction to the dynamics of a solar sail and its applications. Section 3 is a summary of some previous work of the authors on the use of the geometry of the phase space to control a sail. Section 4 introduces a modification of the classical Hill problem to model a solar sail close to an asteroid. Section 5 explains the so-called reduction to the centre manifold for the previous model, and finally Section 6 uses the centre manifold to describe the phase space.

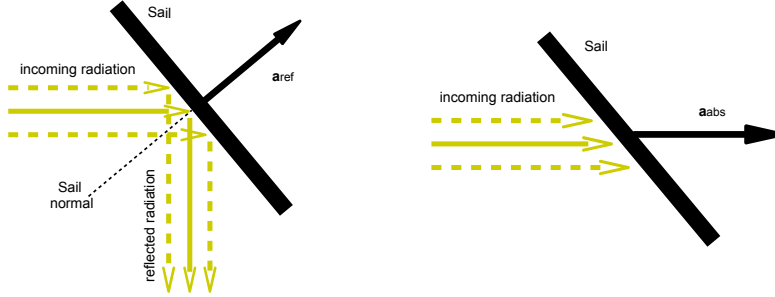


Figure 2: The effect of light on a solar sail. Left: Force produced by the reflected light. Right: Force produced by the absorbed light.

2. Solar sail models

Here, a solar sail is modelled as a flat surface that reflects a large portion of the sunlight, while a small portion is absorbed. The reflected photons produce an impulse in the normal direction of the sail, while the absorbed photons produce an impulse in the opposite direction of the Sun, see Figure 2. We note that the impulse produced by a reflected photon is the sum of two impulses: the impulse produced by its absorption and the impulse produced by its emission. The sail orientation is given by the normal vector to the surface of the sail, \vec{n} and it is parametrised by two angles, α and δ . The acceleration can be written as

$$\vec{a} = \bar{\beta} \langle \vec{r}_s, \vec{n} \rangle \left(\rho \langle \vec{r}_s, \vec{n} \rangle \vec{n} + \frac{1}{2}(1 - \rho)\vec{r}_s \right), \quad (1)$$

where ρ denotes the reflectivity coefficient ($\rho = 0$ corresponds to a perfect solar panel that absorbs all the photons, and $\rho = 1$ corresponds to a perfectly reflecting solar sail). The scalar factor $\bar{\beta}$ depends on the size of the sail, its distance to the Sun, the total mass of the satellite and the units used (see [5] for more details).

In this section, to simplify the discussion, we will assume that the sail is perfectly reflecting, that is, $\rho = 1$ (the model with $\rho < 1$ will be used in the case described in Section 4).

2.1. A dynamical model

Here we use the Restricted Three Body Problem (RTBP) taking the Sun and Earth as primaries and including the solar radiation pressure. In this case, writing $\bar{\beta} = \beta \frac{m_s}{r_{ps}^2}$, we have that the acceleration of the sail is given by

$$\vec{a} = \beta \frac{m_s}{r_{ps}^2} \langle \vec{r}_s, \vec{n} \rangle^2 \vec{n},$$

where now β is a constant, that can be seen as the ratio of the solar radiation pressure in terms of the solar gravitational attraction ($\beta = 1$ means that, if the

sail is perpendicular to Sun direction, the effect of the solar radiation pressure on the sail equals the gravitational attraction of the Sun). With current technology, it is considered reasonable to take $\beta \approx 0.05$ ([20]). This means that a spacecraft of 100 kg has a sail of $58 \times 58 \text{ m}^2$.

The equations of motion are:

$$\begin{aligned}\ddot{x} &= 2\dot{y} + x - (1 - \mu) \frac{x - \mu}{r_{ps}^3} - \mu \frac{x + 1 - \mu}{r_{pe}^3} + \beta \frac{1 - \mu}{r_{ps}^2} \langle \vec{r}_s, \vec{n} \rangle^2 n_x, \\ \ddot{y} &= -2\dot{x} + y - \left(\frac{1 - \mu}{r_{ps}^3} + \frac{\mu}{r_{pe}^3} \right) y + \beta \frac{1 - \mu}{r_{ps}^2} \langle \vec{r}_s, \vec{n} \rangle^2 n_y, \\ \ddot{z} &= - \left(\frac{1 - \mu}{r_{ps}^3} + \frac{\mu}{r_{pe}^3} \right) z + \beta \frac{1 - \mu}{r_{ps}^2} \langle \vec{r}_s, \vec{n} \rangle^2 n_z,\end{aligned}\tag{2}$$

where $\vec{n} = (n_x, n_y, n_z)$ is the normal to the surface of the sail with

$$\begin{aligned}n_x &= \cos(\phi(x, y) + \alpha) \cos(\psi(x, y, z) + \delta), \\ n_y &= \sin(\phi(x, y, z) + \alpha) \cos(\psi(x, y, z) + \delta), \\ n_z &= \sin(\psi(x, y, z) + \delta).\end{aligned}$$

Here, $\vec{r}_s = (x - \mu, y, z)/r_{ps}$ is the Sun - sail direction and the angles ϕ and ψ refer to the position of the probe w.r.t. the Sun, in spherical coordinates (see [7] for details).

The properties of this system depend on the values of the parameters. If $\alpha = \delta = 0$ the system is Hamiltonian for any value of β . If $\alpha = 0$ and $\delta \neq 0$, the system is not Hamiltonian but it is still reversible, under the symmetry $R : (x, y, z, \dot{x}, \dot{y}, \dot{z}, t) \rightarrow (x, -y, z, -\dot{x}, \dot{y}, -\dot{z}, -t)$. For general values of the parameters, the flow of (2) preserves volume but is neither Hamiltonian nor reversible. Of course, if β is small, it is close to a Hamiltonian system.

It is well known that the RTBP has 5 equilibrium points ($L_i, i = 1, \dots, 5$, see Figure 1, right). For small β , these 5 points are replaced by 5 continuous families of equilibria, parametrised by α and δ . For a small value of β , we have 5 disconnected families of equilibria near the classical L_i . For larger values of β , these families merge into each other. We end up having two disconnected surfaces, S_1 and S_2 , where S_1 is like a sphere and S_2 is like a torus around the Sun ([21, 7, 9]).

2.2. Interesting missions applications

There are some proposed missions where the capabilities of solar sails are fundamental ([20]). One of the first is the so-called Geostorm Warning Mission, whose goal is the continuous observation of the Sun to provide information and early warnings of geomagnetic storms. The proposed location is near the point L_1 of the Earth-Sun system. In this case, due to the effect of the sail, the equilibrium point L_1 is closer to the Sun, which is better for this mission. Moreover, if the sail is not orthogonal to the Sun direction, the equilibrium point is displaced away from the Earth-Sun line (see Figure 3) which makes radio communications possible: if the probe were sitting on the Earth-Sun line

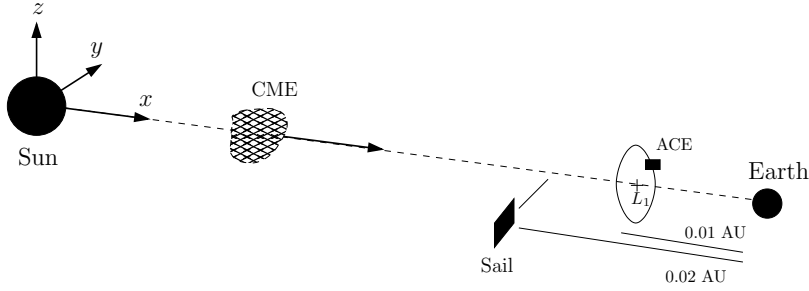


Figure 3: The Geostorm Warning Mission.

then, as seen from Earth, it would be on the middle of the solar disk and its radio signals would be completely masked by the noise coming from the Sun.

Another interesting mission is the Polar Observer mission, to provide a continuous monitoring of Earth poles. Note that the use of conventional satellites on circumpolar orbits would require several of them to have a continuous coverage. Using the same idea as before, we can use the orientation of the sail to displace the equilibrium point upwards as shown in Figure 4.

Finally, there is an increasing interest on sending manned missions to Mars. There is a long list of problems to be solved before making such a journey possible. Here we note that it would be highly interesting to be able to keep the communications with Earth when the two planets are lined up at opposite sides of the Sun (which happens every two years, roughly). One option is to use the same idea as with the Polar Observer Mission, and to displace the equilibrium point upwards so that the solar sail has a clean view of both planets. Of course, this can be done either putting the sail over the Earth, or over Mars (see Figure 5).

We must observe that, in these three missions, the equilibrium point is always unstable. Therefore, some kind of strategy is needed to keep the sail there.

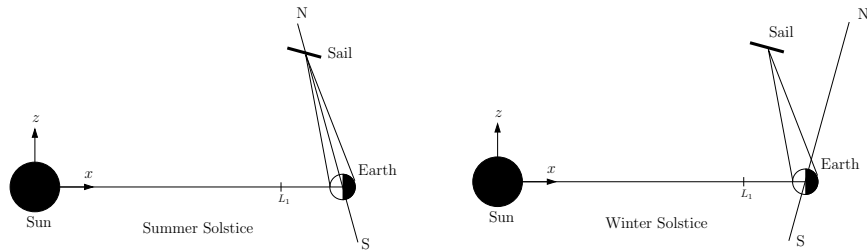


Figure 4: The Polar Observer Mission.

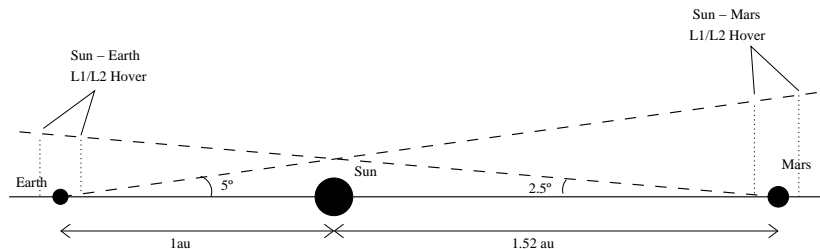


Figure 5: Configurations to maintain continuous communication between Earth and Mars.

3. Controlling a solar sail

In this section we focus on the station keeping of a solar sail near an unstable equilibrium point. The main idea is to look at the linear dynamics around the equilibrium point, study how it varies when the sail orientation changes, and then use this information to change the sail orientation (i.e. the phase space) to make the natural dynamics act in our favour, keeping the trajectory close to a given equilibrium point.

As an example, we focus on the mission Geostorm, where the equilibrium points are unstable with two real eigenvalues, $\lambda_1 > 0$, $\lambda_2 < 0$, and two pair of complex eigenvalues, $\nu_{1,2} \pm i\omega_{1,2}$, with $|\nu_{1,2}| \ll |\lambda_{1,2}|$. As the real parts $\nu_{1,2}$ are very small, we will assume that they are zero. Of course, we will see later (in Section 3.4) that the control strategy still works when this assumption is dropped. If $\nu_{1,2}$ are zero, the linear dynamics at the equilibrium point is of the type saddle \times centre \times centre. We describe the trajectory of the sail in three reference planes defined by the eigendirections, see Figure 6.

For small variations of the sail orientation, the equilibrium point, eigenvalues and eigendirections have a small variation. We will describe the effects of the changes on the sail orientation on each of these three reference planes.

3.1. The hyperbolic direction

Let us start by discussing the control for the hyperbolic directions. When we are close to the equilibrium point, p_0 , the trajectory escapes along the unstable

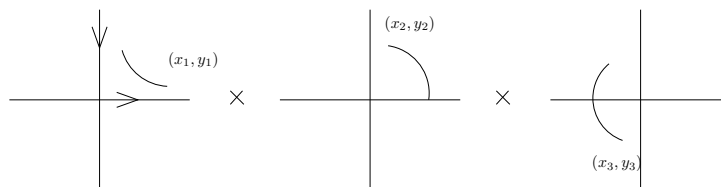


Figure 6: Motion near the equilibrium point, as seen in the reference planes given by the saddle and the two centres.

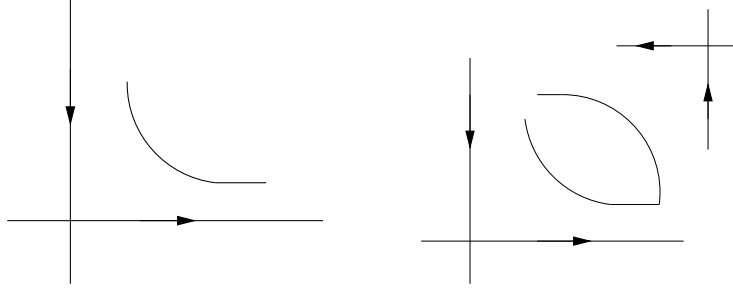


Figure 7: Motion in the hyperbolic direction. Left: The trajectory escapes along the unstable direction. Right: If, using the parameters of the system, we can move the equilibrium point to a suitable place, the unstable manifold of the new point will push the trajectory back.

direction (see Figure 7, left). If we change the sail orientation the equilibrium point is shifted, and the trajectory will escape along the new unstable direction. A first goal is to find a new sail orientation (α, δ) so that the trajectory will come close to the stable direction of p_0 so that, when the initial orientation is restored, the trajectory approaches p_0 (see Figure 7, right). As the trajectory escapes again, the process is repeated.

3.2. The centre directions

In these planes, the trajectory is a rotation around the equilibrium point p_0 . Note that, if the sail orientation is changed, the equilibrium point is shifted and the trajectory will rotate around a new equilibrium point p_1 . Therefore, a sequence of changes on the sail orientation results in a sequence of rotations around different equilibrium points. At this point, it is important to note that a sequence of rotations around different points can result in an unbounded motion (see Figure 8, right). It can be shown that a sufficient condition for the motion to be bounded is that (the centre projection of) p_1 is chosen on the segment that goes from p_0 to the actual position of the probe (see [7] for the proof). In this way, the centre part of the trajectory will not spiral out.

3.3. The control algorithm

The control algorithm is based on looking for a sequence of orientations of the sail such that the natural dynamics of the saddle and centres projection of the dynamics is as described in the previous sections. In this way, the natural dynamics of the sail will keep it close to the selected equilibrium point.

To implement this process, note that we do not know explicitly the position of the equilibrium points $p(\alpha, \delta)$, but we can easily compute the linear approximation of this function, $p(\alpha, \delta) = p(\alpha_0, \delta_0) + Dp(\alpha_0, \delta_0) \cdot (\alpha - \alpha_0, \delta - \delta_0)^T$. We could also compute a second order expansion, but this linear approximation is enough. It is also enough to use the eigenspaces of p_0 for all these points (adding the first order corrections for the eigenvalues when moving α and δ does not

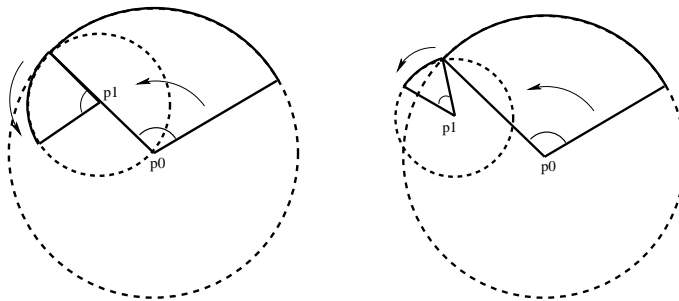


Figure 8: Motion in the centre directions. Left: The second point p_1 makes the trajectory to approach the initial point p_0 . Right: The second point p_0 makes the trajectory to move away from the origin. See the text for more details.

improve the results, because the variation of the angles is small). Hence, we have a two parameter family of equilibria, and we have several conditions to attain, namely, on the relative positions of the saddle and the two centres (that makes up to 6 conditions). This is solved by means of a least squares method.

Of course, it may happen that the solution of this overdetermined system is not good enough to control the sail. For instance, imagine that the saddle shown in Figure 6 (left) moves up and down with α , and orthogonal to the paper page with δ . Then, it would be impossible to control the sail since no matter what we do the saddle is always pushing the trajectory away. On the other hand, if for some angles the (projection of the) saddle moves to the “other side” of the sail, then this can be used to push the trajectory back (at least, to the stable direction of the previous point). To have this second situation a suitable non-degeneracy (or transversality) condition is needed. This condition is satisfied in the Geostorm mission considered here (see [7] for more details).

3.4. Simulations

The next step is to test this control strategy using the full set of equation. That is, we use the linearized dynamics (with the assumption that $\nu_{1,2}$ are zero) to compute the change on the sail orientation, but use the complete set of equations for its motion.

We have done a Monte Carlo simulation taking 1000 random initial conditions near the point and trying to control each one for a time span of 30 years (this is a very long time for a space mission). We have tested the robustness of our strategy by including random errors on the position and velocity determination (this means that we do not know exactly the position of the sail in phase space), as well as on the orientation of the sail at each manoeuvre (this means that the sail is not exactly at the required orientation). These errors try to imitate the real situation found in a space mission.

We have taken $\beta = 0.051689$ which refers, for instance, to a satellite of 130 kg mass with a 67×67 m square sail. The results of the simulation are

	Success	Max. Time	Min. Time	Ang. Vari.
No Error	100 %	45.87 days	24.13 days	1.43°
Error Pos.	100 %	45.85 days	24.13 days	1.43°
Error Pos. & Ori. [*]	100 %	53.90 days	21.59 days	1.42°
Error Pos. & Ori. [†]	97 %	216.47 days	15.54 days	1.67°

Table 1: Statistics for the Geostorm mission taking 1000 simulations. Considering errors on the sail orientation of order 0.5° (^{*}) and 2.2° ([†]). See the text for more details.

summarised in Table 1. The meaning of the entries is the following: The row “No Error” refers to the simulations without any errors in the phase space position of the sail, or in its orientation. The row “Error Pos.” refers to simulations with errors in the phase space position, the row “Error Pos. & Ori” adds errors in the desired orientation for the sail. The first of these two rows has used orientation errors of half a degree, while the second one of 2.2 degrees (when the control strategy starts to fail). The columns “Max. Time” and “Min. Time” refer to the time between manoeuvres (it is convenient to have a reasonable time between manoeuvres, for instance a month is a good number) and the last column is the angular variation of the sail as seen from the Earth. The conclusion is that the proposed method works perfectly, provided that the errors in the orientation of the sail are below 2°. If the orientation errors were larger than this value, then we should allow the sail to move far away from the point (to make the errors in the orientation relatively smaller) and, hence, use second (or higher) order terms for the manifolds.

4. The Augmented Hill problem

The classical Hill problem is a simplification of the RTBP introduced by G.W. Hill ([15]) to study the motion of the Moon. In short, the Hill problem aims to describe the motion of an infinitesimal particle (Moon) attracted by two massive bodies: the first (Earth) is fixed at the origin, while the second body (Sun) is assumed to be far away, in such a way that it produces a uniform and constant vector field around the mass sitting at the origin. The reference frame rotates uniformly such that the x axis points in the direction of the Sun. The motion of the infinitesimal particle is driven by this constant vectorfield plus the gravitational attraction of the mass at the origin. For more details see, for instance, [25, 22].

In this section we are interested in the motion of a solar sail near an asteroid. Due to the large difference between the masses of Sun, asteroid and probe, and the difference between the asteroid-Sun and the asteroid-probe distances, it seems natural to use a Hill model for this situation. As we also want to account for the effect of the solar radiation pressure on the sail we add this force as an extra term. The equations of motion are then,

$$\ddot{x} - 2\dot{y} = -\frac{x}{r^3} + 3x + a_x,$$

$$\begin{aligned}\ddot{y} + 2\dot{x} &= -\frac{y}{r^3} + a_y, \\ \ddot{z} &= -\frac{z}{r^3} - z + a_z,\end{aligned}$$

where (x, y, z) denotes the position of the solar sail, $r = \sqrt{x^2 + y^2 + z^2}$, and $\vec{a} = (a_x, a_y, a_z)$ is the acceleration given by the solar sail. In this case, as the direction of the Sun is constant, $r_s = (1, 0, 0)$, expression (1) takes the form

$$\begin{aligned}a_x &= \bar{\beta}(\rho \cos^3 \alpha \cos^3 \delta + 0.5(1 - \rho) \cos \alpha \cos \delta), \\ a_y &= \bar{\beta}(\rho \cos^2 \alpha \cos^3 \delta \sin \alpha), \\ a_z &= \bar{\beta}(\rho \cos^2 \alpha \cos^2 \delta \sin \delta),\end{aligned}\tag{3}$$

where α and δ are the angles defining the orientation of the sail. The normalised units of distance and time are $L = (\mu_{sb}/\mu_{sun})^{1/3}R$ and $T = 1/\omega$, where μ_{sb} and μ_{sun} are the gravitational parameters for the small body (asteroid) and the Sun, R is the Sun-asteroid mean distance, and $\omega = \sqrt{\mu_{sun}/R^3}$ is its frequency.

Due to the nature of the Hill model, the normalisation of units is different from the one of the RTBP. In the normalised units for the Hill problem, we have that $\bar{\beta} = K_1(A/m)\mu_{sb}^{-1/3}$, where $K_1 \approx 7.8502$ if the area A is given in m^2 and the mass m is measured in kg. Defining momenta as $p_x = \dot{x} - y$, $p_y = \dot{y} + x$, $p_z = \dot{z}$, this system is described by the Hamiltonian function

$$\begin{aligned}H &= \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) + yp_x - xp_y - \frac{1}{2}(2x^2 - y^2 - z^2) - \frac{1}{r} \\ &\quad - a_x x - a_y y - a_z z.\end{aligned}\tag{4}$$

4.1. Equilibrium points

It is well-known that the Hill problem (that is, the case $\bar{\beta} = 0$ in (3)) has two equilibrium points, $L_{1,2}$, symmetrically located around the asteroid, with coordinates $(\pm 3^{-1/3}, 0, 0)$ (see [25]). The effect of the solar radiation pressure is to displace these points. For instance, if the sail is perpendicular to the Sun direction ($\alpha = \delta = 0$), the position of $L_{1,2}$ move towards the Sun (on the x axis) as β increases. If the sail is not perpendicular to the Sun direction, the equilibrium points are displaced outside the x axis, see [10, 11] for more details.

4.2. Families of periodic and quasi-periodic orbits

For realistic values of β , these equilibrium points are unstable: the linearised dynamics at the point is the product of two linear oscillations and one linear saddle. Skipping the hyperbolic directions, the linear dynamics around the point can be described as the product of two harmonic oscillators. If the frequencies of these two oscillators are linearly independent over rationals, their product produces a two parametric family¹ of quasi-periodic orbits that sometimes are

¹The two parameters of the family are the amplitudes of the two linear oscillations.

referred as Lissajous orbits. When the nonlinear part of the system is added, the well-known Lyapunov centre theorem (see, for instance, [22]) ensures that for each harmonic oscillator there exists a one-parametric family of periodic orbits that can be seen as the extension of these oscillations to the full nonlinear system. When the amplitude of these periodic orbits goes to zero (i.e., when the periodic orbits tend to the equilibrium point) their period goes to the period of the corresponding harmonic oscillations.

If the frequencies of the two centre directions are not only linearly independent over rationals but also Diophantine, the two parametric family of Lissajous orbit can be extended into the nonlinear system as a two-parametric Cantorian family of quasi-periodic motions with two basic frequencies. As before, when the amplitudes of these quasi-periodic orbits go to zero (i.e., when the quasi-periodic orbits tend to the equilibrium point) their frequencies go to the frequencies of the harmonic oscillators. The normal directions to these orbits contain a saddle (it comes from the saddle of the equilibrium point) so they are unstable. For more details on these topics, see [18, 19, 4].

The previous discussion shows the existence of a two-parametric (Cantorian) family of unstable quasi-periodic orbits close to the equilibrium points. We want to know if this local behaviour extends to a larger region around the point, and what kind of orbits exist in this large region. Note that this question is very difficult to answer by means of numerical simulations: the numerical integration of orbits starting near these points produces trajectories that escape quite fast due to the unstable character of the region. To solve this problem, we will perform the so-called reduction to the centre manifold.

5. Reduction to the centre manifold

It is based on performing a sequence of normalising transformations on the Hamiltonian function (4) with the only purpose of decoupling the centre directions from the hyperbolic ones. These transformations are applied on a power expansion of the Hamiltonian at the equilibrium point and, hence, the final Hamiltonian (which is a power expansion) is only valid on a neighbourhood of the point. As we will see, this neighbourhood is quite large. At the end of the process, we obtain the power expansion of a two-degrees of freedom Hamiltonian system, which can be seen as the initial three-degrees of freedom Hamiltonian without the directions of the saddle. Choosing suitable Poincaré sections, the dynamics reduces to a family of two dimensional maps (being the parameter of the maps the energy level of the Hamiltonian), and can be easily displayed. This technique has already been used to study the neighbourhood of the collinear points of the RTBP; see, for instance, [12] (reprinted as [13]) and [16, 17, 14]. For the use of this technique in a non-Hamiltonian situation, see also [8, 9]. Now let us discuss how it works in the Hamiltonian setting.

Let us start by assuming that the equilibrium point has been translated to the origin and the Hamiltonian has been expanded in power series. Moreover, we have chosen suitable (complex) coordinates such that the linear part of the

system is in diagonal form. This means that the Hamilton function can be written as

$$H(q, p) = H_2(q, p) + \sum_{n \geq 3} H_n(q, p), \quad (5)$$

where $H_2 = \lambda_1 q_1 p_1 + \sqrt{-1} \omega_1 q_2 p_2 + \sqrt{-1} \omega_2 q_3 p_3$, and H_n denotes an homogeneous polynomial of degree n .

5.1. The Lie series method

Let us recall that, if $F(q, p)$ and $G(q, p)$ are two real functions (where, as usual, q denotes the positions and p the momenta), their Poisson bracket is defined as

$$\{F, G\} = \frac{\partial F}{\partial q} \frac{\partial G}{\partial p} - \frac{\partial F}{\partial p} \frac{\partial G}{\partial q}.$$

In what follows, we will use the following notation. If $z = (z_1, \dots, z_n)$ is a vector of complex numbers and $k = (k_1, \dots, k_n)$ is an integer vector, we denote by z^k the value $z_1^{k_1} \dots z_n^{k_n}$ (in this context we define 0^0 as 1). Moreover, we define $|k|$ as $|k_1| + \dots + |k_n|$.

A change of variables is called canonical when it preserves the Hamiltonian form (for any Hamiltonian function) of the equations of motion. It is not difficult to show a transformation is canonical if and only if the differential of the change (on any point) is a symplectic matrix. Canonical transformations are very useful both from the theoretical and practical point of view, since they allow to work on a single function (the Hamiltonian) instead of a system of differential equations.

It is not easy to obtain canonical changes of variables in a explicit form, since it is very difficult to impose that its differential at any point is a symplectic matrix. Fortunately, there exist several techniques to produce such transformations. The one used here is based on the following properties of Hamiltonian flows:

1. Let $\Phi_t(x, y)$ be the time t flow of a Hamiltonian system. Then, $(q, p) = \Phi_t(x, y)$ is a canonical transformation.
2. Let $G(q, p)$ a Hamiltonian system (with ℓ degrees of freedom) and let $(q_0(t), p_0(t))$ be an orbit of this system. Then, for any smooth function f ,

$$\frac{d}{dt} f(q_0(t), p_0(t)) = \{f, G\}(q_0(t), p_0(t)). \quad (6)$$

Then, it is not difficult to see that to transform a Hamiltonian H by means of the time 1 flow of a Hamiltonian G , we can apply the formula

$$\hat{H} \equiv H + \{H, G\} + \frac{1}{2!} \{\{H, G\}, G\} + \frac{1}{3!} \{\{\{H, G\}, G\}, G\} + \dots, \quad (7)$$

where \hat{H} denotes the transformed Hamiltonian. This formula follows from the application of the Taylor formula for the transformation and using (6) for the derivatives involved. The Hamiltonian G is usually called the generating function of the change of variables.

If the Hamiltonian H and the generating function G are truncated power expansions, the expression (7) is very suitable for effective computations, since it can be easily implemented on a computer. Some authors use the so-called Lie triangle for this implementation. We note that it is possible to organise the computations to avoid computing this triangle and saving memory, see [16]. One can argue that a practical problem for this kind of transformation is that it is defined by an infinite series. This is not a problem since we usually work with a finite truncation of these series. This will produce a high order approximation to the results wanted that, in many cases, are good enough for practical purposes. On the other hand, it is possible to derive rigorous estimates on the size of this remainder so one can obtain bounds on the error of the results obtained with the truncated series.

5.2. Practical implementation

It is easy to check that, if P and Q are homogeneous polynomials of degree r and s respectively, then $\{P, Q\}$ is a homogeneous polynomial of degree $r + s - 2$. This property is very useful for the computer implementation of a transformation given by a generating transformation G . For instance, let us assume that we want to eliminate the monomials of degree 3 of the Hamiltonian, as it is usually done in a normal form scheme. Let us select as a generating function a homogeneous polynomial of degree 3, G_3 . Then, it is immediate to check that the terms of the transformed Hamiltonian \hat{H} satisfy

- degree 2: $\hat{H}_2 = H_2$,
- degree 3: $\hat{H}_3 = H_3 + \{H_2, G_3\}$,
- degree 4: $\hat{H}_4 = H_4 + \{H_3, G_3\} + \frac{1}{2!} \{\{H_2, G_3\}, G_3\}$,
- ...

Hence, to kill the monomials of degree 3 one has to look for a G_3 such that $\{H_2, G_3\} = -H_3$. Let us denote

$$\begin{aligned} H_3(q, p) &= \sum_{|k_q|+|k_p|=3} h_{k_q, k_p} q^{k_q} p^{k_p}, \\ G_3(q, p) &= \sum_{|k_q|+|k_p|=3} g_{k_q, k_p} q^{k_q} p^{k_p}, \end{aligned}$$

where $\eta_1 = \lambda_1$, $\eta_2 = \sqrt{-1}\omega_1$ and $\eta_3 = \sqrt{-1}\omega_2$. As

$$\{H_2, G_3\} = \sum_{|k_q|+|k_p|=3} \langle k_p - k_q, \eta \rangle g_{k_q, k_p} q^{k_q} p^{k_p}, \quad \eta = (\eta_1, \eta_2, \eta_3),$$

we immediately obtain

$$G_3(q, p) = \sum_{|k_q|+|k_p|=3} \frac{-h_{k_q, k_p}}{\langle k_p - k_q, \eta \rangle} q^{k_q} p^{k_p}.$$

Observe that $|k_q| + |k_p| = 3$ implies $\langle k_p - k_q, \eta \rangle \neq 0$. Note that G_3 is so easily obtained because of the “diagonal” form of H_2 .

We are not interested in a complete normal form, but only in uncoupling the central directions from the hyperbolic one. Hence, it is not necessary to cancel all the monomials in H_3 but only some of them. Moreover, as we want the radius of convergence of the transformed Hamiltonian to be as big as possible, we will try to choose the change of variables as close to the identity as possible. This means that we will kill the least possible number of monomials in the Hamiltonian. Hence, we kill the monomials $q^{k_q} p^{k_p}$ such that the first factor of k_q is different from the first factor of k_p (we will see in a moment that this is enough to produce the centre manifold). This implies that the generating function G_3 is

$$G_3(q, p) = \sum_{(k_q, k_p) \in \mathcal{S}_3} \frac{-h_{k_q, k_p}}{\langle k_p - k_q, \eta \rangle} q^{k_q} p^{k_p},$$

where \mathcal{S}_n , $n \geq 3$, is the set of indices (k_q, k_p) such that $|k_q| + |k_p| = n$ and the first component of k_q is different from the first component of k_p . Then, the transformed Hamiltonian \hat{H} takes the form

$$\hat{H}(q, p) = H_2(q, p) + \hat{H}_3(q, p) + \hat{H}_4(q, p) + \dots, \quad (8)$$

where $\hat{H}_3(q, p) \equiv \hat{H}_3(q_1 p_1, q_2, p_2, q_3, p_3)$ (note that \hat{H}_3 depends on the product $q_1 p_1$, not on each variable separately²). This process can be carried out up to a finite order N , to obtain a Hamiltonian of the form

$$\bar{H}(q, p) = \bar{H}_N(q, p) + R_N(q, p), \quad (9)$$

where $H_N(q, p) \equiv H_N(q_1 p_1, q_2, p_2, q_3, p_3)$ is a polynomial of degree N and R_N is a remainder of order $N + 1$ (note that H_N depends on the product $q_1 p_1$).

Neglecting the remainder and applying the canonical change given by $I_1 = q_1 p_1$, we obtain the Hamiltonian $\bar{H}_N(I_1, q_2, p_2, q_3, p_3)$ that has I_1 as a first integral. Setting $I_1 = 0$ we obtain a 2DOF Hamiltonian, $\bar{H}_N(0, \bar{q}, \bar{p})$, $\bar{q} = (q_2, q_3)$, $\bar{p} = (p_2, p_3)$, that represents (up to some finite order N) the dynamics inside the centre manifold.

It is important to note the absence of small divisors during this process. The denominators that appear in the generating functions, $\langle k_p - k_q, \eta \rangle$, can be bounded from below when $(k_q, k_p) \in \mathcal{S}_N$: using that η_1 is real and that $\eta_{2,3}$ are purely imaginary, we have

$$|\langle k_p - k_q, \eta \rangle| \geq |\lambda_1|, \quad \text{for all } (k_q, k_p) \in \mathcal{S}_N, \quad N \geq 3.$$

For this reason, the divergence of this process is very mild. The divergence of normalising transformations in the absence of small divisors has been considered in other contexts, see, for instance, [1] or [2]; for a more general discussion

²Because we have killed the monomials $q^{k_q} p^{k_p}$ of H such that the first component of k_q is different from the first component of k_p

see [23]. This is clearly observed when this process is stopped at some degree N , since the remainder is very small in a quite big neighbourhood of the equilibrium point.

5.3. Changes of variables

It is also possible to compute an explicit expression for the nonlinear change of variables between the coordinates of (5) and (9), by applying each generating function G_3, \dots, G_N to each coordinate: for instance,

$$\begin{aligned}\tilde{q}_j &= q_j + \{q_j, G_3\} + \frac{1}{2!} \{\{q_j, G_3\}, G_3\} + \frac{1}{3!} \{\{\{q_j, G_3\}, G_3\}, G_3\} + \dots, \\ \tilde{p}_j &= p_j + \{p_j, G_3\} + \frac{1}{2!} \{\{p_j, G_3\}, G_3\} + \frac{1}{3!} \{\{\{p_j, G_3\}, G_3\}, G_3\} + \dots,\end{aligned}$$

is the transformation that links the coordinates of (8) given by the (q, p) variables to the coordinates of (5) given by the coordinates (\tilde{q}, \tilde{p}) . The successive application of the generating functions G_4, \dots, G_N produces the final transformation. See [16] for details.

6. Dynamics in the centre manifold for the Augmented Hill Problem

Next step is to display the dynamics of the 2DOF Hamiltonian system $H_{cm} = \bar{H}_N(0, \bar{q}, \bar{p})$. Let us call (q_h, p_h) the variables in the normalised coordinates related to the horizontal oscillations, and (q_v, p_v) the variables related to the vertical oscillations. We consider the Poincaré section $q_v = 0$ (in other words, we are “slicing” the vertical motions).

In what follows, we fix $\rho = 0.85$ and $\beta = 5$. Note that, as β is an adimensional value that compares the solar radiation pressure with the gravitation of the asteroid, the smaller the asteroid mass is, the larger the β . The value $\beta = 5$ corresponds to a small sail near a large asteroid like Ceres ([10, 11]). Moreover, we will focus on the dynamics near the point L_2 . Hence, we perform the reduction to the centre manifold for this case. As this process is very fast, the reduction is done for each set of considered parameters.

Let us first consider the case $\alpha = \delta = 0$ and select the energy level $H_{cm} = 0.4$ (corresponding to $H = -4.51907174$ in (4)). The corresponding Poincaré section is shown in Figure 9 (left), in coordinates (q_h, p_h) . This map is obtained as follows: for each initial data $(q_h^{(0)}, p_h^{(0)})$, we compute the only positive value $p_v^{(0)}$ such that the point $(q_h^{(0)}, p_h^{(0)}, q_v^{(0)} = 0, p_v^{(0)})$ is in the energy level $H_{cm} = 0.4$. This point is used as initial data for the flow of the Hamiltonian H_{cm} and Figure 9 (left) shows the intersections of the orbit with the section $q_v = 0$, for several orbits. Figure 9, right, displays these same points but in the initial (synodic) coordinates (x, y, z) of the Hill problem. Note that most of the trajectories in this map are invariant curves. In the initial flow, they correspond to quasi-periodic motions with two basic frequencies: a frequency related to a vertical oscillation (by vertical we mean outside of the (x, y) coordinates of (4)), and a frequency related to the rotation on the invariant curve shown in the

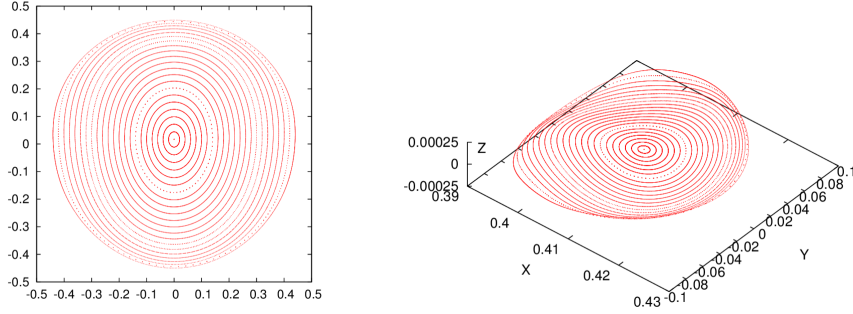


Figure 9: $H_{cm} = 0.4$, $\alpha = \delta = 0$. Left: Poincaré section in centre manifold coordinates. Right: The left plot but in synodical Hill coordinates.

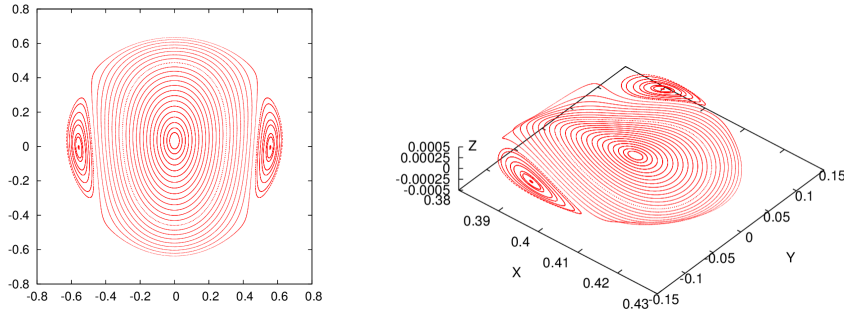


Figure 10: $H_{cm} = 0.8$, $\alpha = \delta = 0$. Left: Poincaré section in centre manifold coordinates. Right: The left plot but in synodical Hill coordinates.

Poincaré section. The fixed (elliptic) point in the centre of the plot corresponds to a vertical periodic Lyapunov orbit, and the outside boundary of the plot corresponds to a planar periodic Lyapunov orbit.

We note that, as $H_{cm}(0) = 0$ and the Hessian of H_{cm} at the origin is positive definite, the set $H_{cm} = h$ is diffeomorphic to \mathbb{S}^3 if h is small.

Figure 10 corresponds to the energy level $H_{cm} = 0.8$ ($H = -4.45085751$ in (4)). The main difference with the previous figure is that two new elliptic points have appeared: the so called Halo orbits.

6.1. Halo orbits

Halo orbits are a very well known class of periodic orbits of the Restricted Three-Body Problem, located near the collinear equilibrium points $L_{1,2,3}$. They are periodic orbits which bifurcate from the planar Lyapunov periodic orbits when the in plane (or intrinsic) and out of plane (or normal) frequencies are equal. In other words, they are born at a 1:1 resonance involving the frequency of a periodic orbit with one of its normal frequencies.

The importance of these orbits became clear with the mission ISEE 3 (International Sun-Earth Explorer) that was launched in 1978. The purpose of this

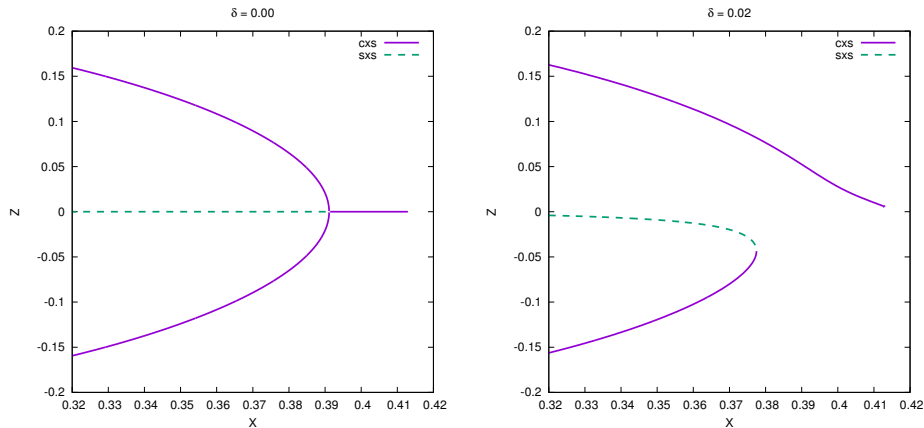


Figure 11: Left: Continuation of the planar Lyapunov family of periodic orbits when $\alpha = 0$, $\delta = 0$, showing the pitchfork bifurcation where the family of Halo orbits is born. Right: Periodic orbits when $\alpha = 0$, $\delta = 0.02$. Both plots show the (x, z) coordinates of the periodic orbits in the Poincaré section $y = 0$.

probe was to study the Sun so it was desirable to place it in between Earth and Sun, to have a continuous monitoring of the activity of the Sun. Note that to place it at the L_1 point satisfies this requirement but with the following drawback: as seen from the Earth, the spacecraft would be in the middle of the solar disk. This means that an antenna pointing to the probe is also pointing to the Sun, and the noise coming from the Sun would make it impossible to receive any data from the probe. Halo orbits provide a very good alternative to place the spacecraft: As seen from the Earth, a Halo orbit with low vertical amplitude is seen as moving East and West through the solar disk, but if the orbit is selected with a large enough vertical amplitude, the orbit is seen moving around the solar disk, without crossing it. This is the reason for using the word “Halo” to name these orbits: they somehow remind one of a halo around a saint. Note that a probe on one of these orbits can keep permanent communications with the Earth while it has a continuous coverage of the Sun. Pioneer works in this direction are [6, 24] (Sun-Earth), [26] (Sun-Jupiter), and [3] (Earth-Moon).

The Augmented Hill model can be seen as a perturbation of the RTBP: on the one hand, the Hill problem is a simplification of the RTBP that is very natural in the present situation and, on the other hand, we have added the effect of the radiation pressure. Hence, it is not a surprise to still find Halo orbits in this augmented Hill problem. If we fix $\alpha = \delta = 0$ (we recall that β is already fixed at 5), we can continue the planar Lyapunov family of periodic orbits by varying the energy level (we could also use an arc-length continuation) and then the Halo orbits appear as a pitchfork bifurcation of the planar Lyapunov periodic orbits.

It is interesting to note that one of the effects of the extra parameters δ is to break the symmetry of the pitchfork, as shown in Figure 11. As an illustration, Figure 12 shows the pitchfork bifurcation in two different Poincaré sections of

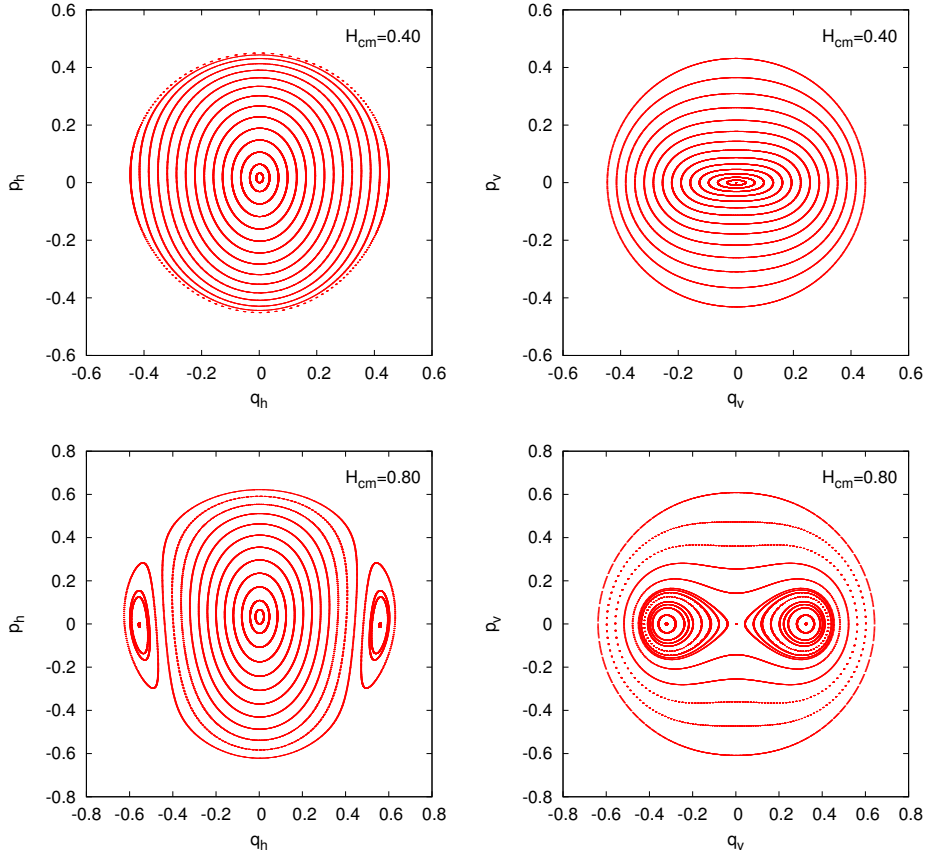


Figure 12: Poincaré section in the centre manifold, for $\alpha = 0$, $\delta = 0$ and different energy levels. Left plots: (q_h, p_h) (horizontal) section. Right plots (q_v, p_v) (vertical) section.

the centre manifold: in the left column of Figure 12 we show what we call the *horizontal* section: the slice is given by $q_v = 0$ and we display the (q_h, p_h) plane (the value of p_v is obtained from the energy level). The outer boundary of this section corresponds to the planar Lyapunov orbit and the vertical Lyapunov orbits appears as a single dot at the middle. Note that this orbit is completely included in the section and, hence, its pitchfork bifurcation is not well seen, so that the Halo orbits seem to appear from nowhere. On the other hand, in the right column we show what we call the *vertical* section: the slice is $q_h = 0$ and we display the (q_v, p_v) plane (as before, the value of p_h is obtained from the energy level). Now the outer boundary of the section is given by the vertical Lyapunov orbit, while the planar Lyapunov orbit appears as a dot at the middle. This section is transversal to the bifurcating planar Lyapunov orbit and then the pitchfork is displayed properly. Figure 13 contains the same plots as Figure 12 but for $\delta = 0.02$, where you can see the effect of the symmetry breaking on the Poincaré section.

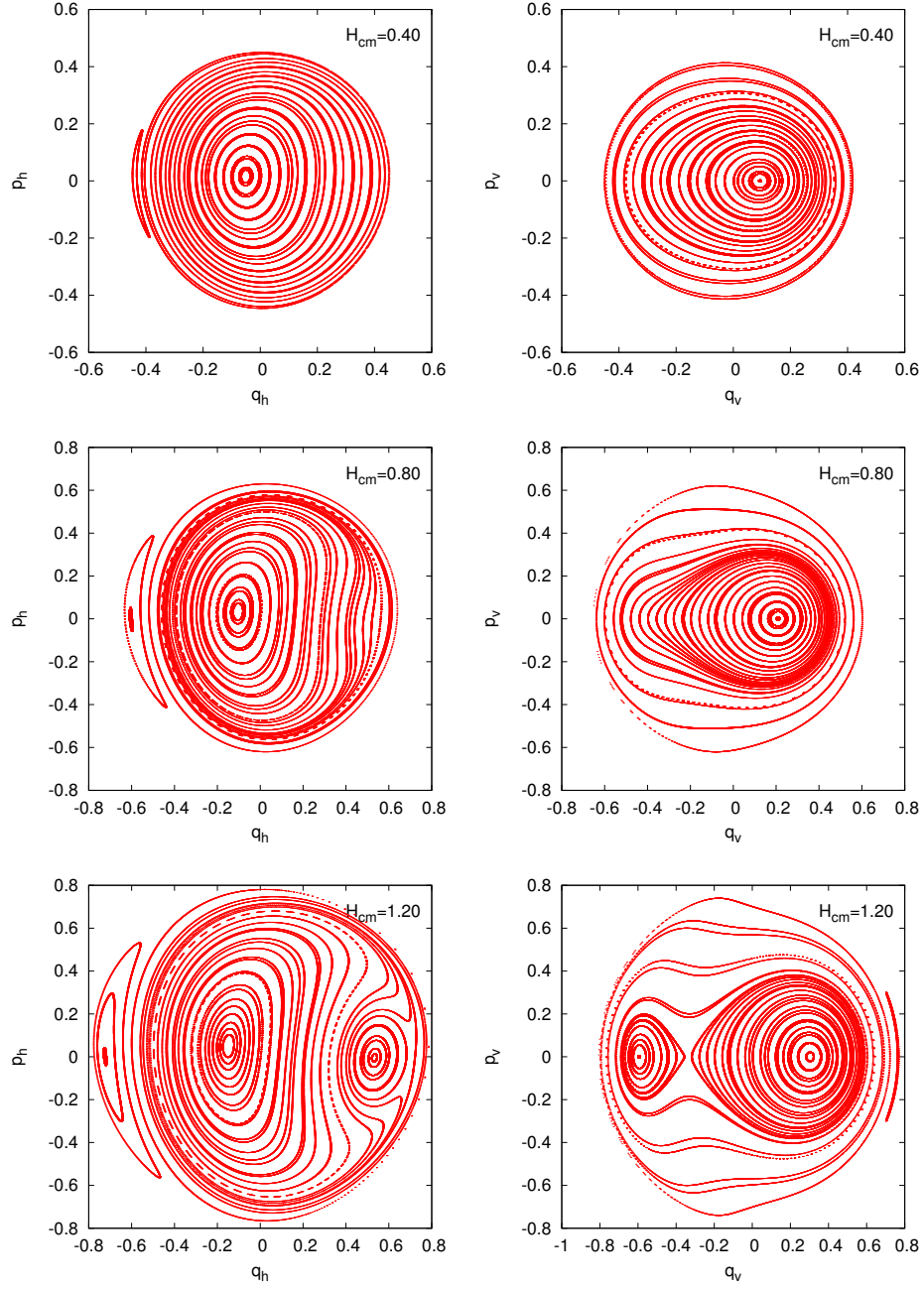


Figure 13: Poincaré section in the centre manifold, for $\alpha = 0$, $\delta = 0.02$ and different energy levels. Left plots: (q_h, p_h) (horizontal) section. Right plots (q_v, p_v) (vertical) section. Compare with Figure 12

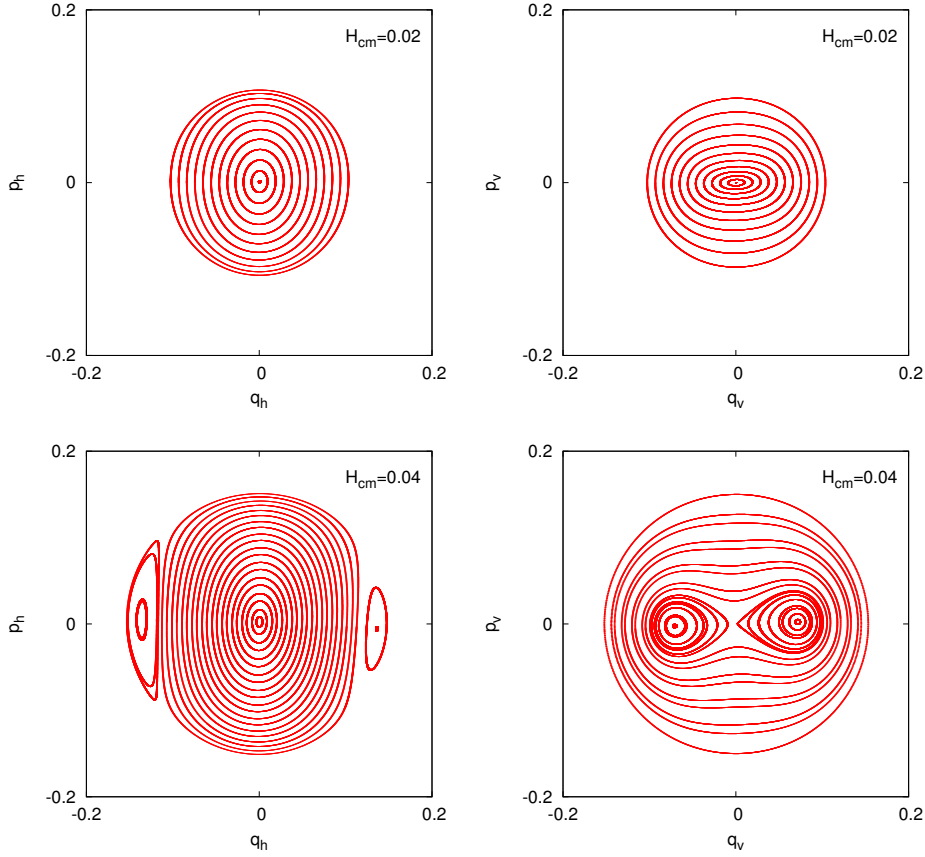


Figure 14: Same as Figure 12 but for $\alpha = 0.49$, $\delta = 0$.

Now let us fix $\delta = 0$ and move α (this keeps the symmetry responsible for having the pitchfork bifurcation that created the Halo orbits). When α starts increasing from zero, this bifurcation point starts moving towards the equilibrium point and, when $\alpha = \alpha_{crit} \approx 0.507819585540$, the branching of Halo orbits takes place from the equilibrium point, which is in a 1:1 resonance. Figures 14, 15 and 16 show the phase space before, at and after this resonance. It is clear that the pitchfork bifurcation point crosses over the equilibrium point and goes from the planar to the vertical family, so the horizontal and vertical Poincaré sections interchange roles. A complete analysis of the dynamics near $\alpha = \alpha_{crit}$ by using the necessary number of parameters to unfold this resonance is actually work in progress. In particular, we are interested in the (partial) unfolding given when moving δ , and also in the effect of the parameter β . This would describe the geometry of the phase space that a solar sail would find when navigating near this point.

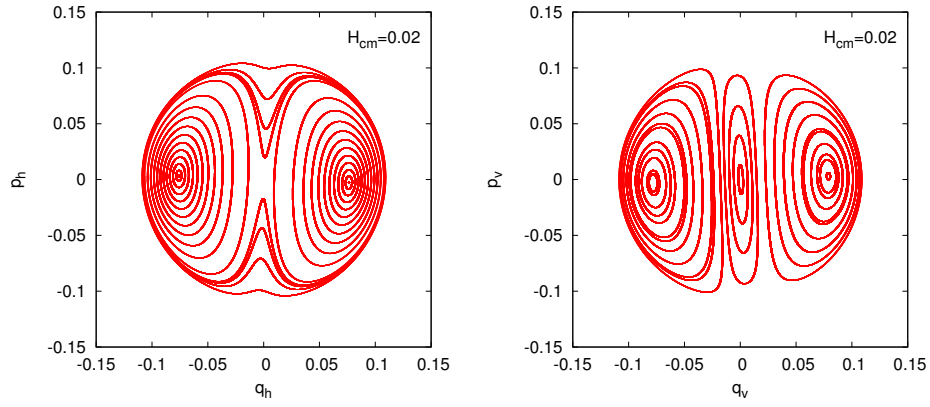


Figure 15: Same as Figure 12 but for $\alpha = 0.5078195855399$, $\delta = 0$.

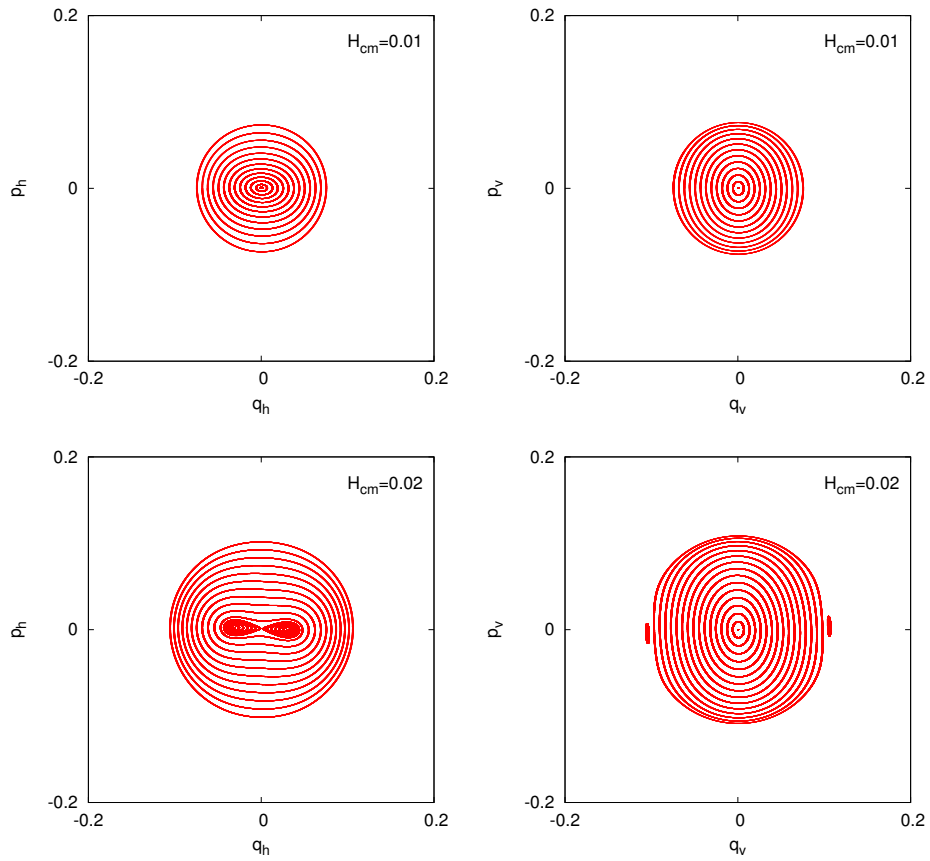


Figure 16: Same as Figure 12 but for $\alpha = 0.52$, $\delta = 0$.

7. Conclusions

We have discussed the use of dynamical systems tools to understand the geometry of the phase space of a solar sail. As introduction, we have surveyed some previous results of the authors and we have shown how to use the knowledge of the phase space to control a sail in the Earth-Sun system. Then, we have used the Augmented Hill problem as a model for a sail near the point L_2 of the asteroid-Sun system. As this is a 3 degree of freedom Hamiltonian system, we have used the reduction to the centre manifold to describe the geometry of the phase space near the point L_2 , showing the occurrence of periodic orbits, bifurcations and quasi-periodic motions.

References

- [1] Bazzani, A., Giovannozzi, M., Servizi, G., Todesco, E., Turchetti, G., 1993. Resonant normal forms, interpolating Hamiltonians and stability analysis of area preserving maps. *Phys. D* 64, 66–97.
- [2] Benettin, G., Giorgilli, A., 1994. On the Hamiltonian interpolation of near-to-the-identity symplectic mappings with application to symplectic integration algorithms. *J. Statist. Phys.* 74, 1117–1143.
- [3] Breakwell, J., Brown, J., 1979. The ‘Halo’ family of 3-dimensional periodic orbits in the Earth-Moon restricted 3-body problem. *Celestial Mech.* 20, 389–404.
- [4] Broer, H., Hanßmann, H., Jorba, A., Villanueva, J., Wagener, F., 2003. Normal-internal resonances in quasiperiodically forced oscillators: a conservative approach. *Nonlinearity* 16, 1751–1791.
- [5] Dachwald, B., Seboldt, W., Macdonald, M., Mengali, G., Quarta, A., McInnes, C., Rios-Reyes, L., Scheeres, D., Wie, B., Görlich, M., et al., 2005. Potential Solar sail degradation effects on trajectory and attitude control, in: *AIAA Guidance, Navigation, and Control Conference and Exhibit*. doi:10.2514/6.2005-6172.
- [6] Farquhar, R., Kamel, A., 1973. Quasi-periodic orbits about the translunar libration point. *Celestial Mech.* 7, 458–473.
- [7] Farrés, A., Jorba, A., 2008. A dynamical system approach for the station keeping of a solar sail. *J. Astronaut. Sci.* 56, 199 – 230.
- [8] Farrés, A., Jorba, À., 2010a. On the high order approximation of the centre manifold for ODEs. *Discrete Contin. Dyn. Syst. Ser. B* 14, 977–1000.
- [9] Farrés, A., Jorba, À., 2010b. Periodic and quasi-periodic motions of a solar sail close to SL_1 in the Earth-Sun system. *Celestial Mech.* 107, 233–253.

- [10] Farrés, A., Jorba, À., Mondelo, J., 2014a. Orbital dynamics for a non-perfectly reflecting solar sail close to an asteroid, in: Proceedings of the 2nd IAA Conference on Dynamics and Control of Space Systems, Rome, Italy.
- [11] Farrés, A., Jorba, À., Mondelo, J., Villac, B., 2014b. Periodic motion for an imperfect solar sail near an asteroid, in: Macdonald, M. (Ed.), *Advances in Solar Sailing*. Springer, Berlin, Heidelberg, pp. 885–898.
- [12] Gómez, G., Jorba, À., Masdemont, J., Simó, C., 1991. Study refinement of semi-analytical Halo orbit theory. ESOC Contract 8625/89/D/MD(SC), final report. European Space Agency.
- [13] Gómez, G., Jorba, À., Masdemont, J., Simó, C., 2001. Dynamics and mission design near libration points. Vol. III, *Advanced methods for collinear points*. volume 4 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc.
- [14] Haro, A., Canadell, M., Luque, A., Mondelo, J.M., Figueras, J.L., 2016. The Parameterization Method for Invariant Manifolds. From Rigorous Results to Effective Computations. volume 195 of *Applied Mathematical Sciences*. Springer-Verlag.
- [15] Hill, G., 1878. Researches in the Lunar Theory. *Amer. J. Math.* 1, 5–26, 129–147, 245–260.
- [16] Jorba, À., 1999. A methodology for the numerical computation of normal forms, centre manifolds and first integrals of Hamiltonian systems. *Exp. Math.* 8, 155–195.
- [17] Jorba, À., Masdemont, J., 1999. Dynamics in the centre manifold of the collinear points of the Restricted Three Body Problem. *Phys. D* 132, 189–213.
- [18] Jorba, À., Villanueva, J., 1997. On the normal behaviour of partially elliptic lower dimensional tori of Hamiltonian systems. *Nonlinearity* 10, 783–822.
- [19] Jorba, À., Villanueva, J., 2001. The fine geometry of the Cantor families of invariant tori in Hamiltonian systems, in: European Congress of Mathematics, Vol. II (Barcelona, 2000). Birkhäuser, Basel. volume 202 of *Progr. Math.*, pp. 557–564.
- [20] McInnes, C., 1999. *Solar Sailing: Technology, Dynamics and Mission Applications*. Springer-Praxis.
- [21] McInnes, C., McDonald, A., Simmons, J., MacDonald, E., 1994. Solar sail parking in restricted three-body system. *Journal of Guidance, Control and Dynamics* 17, 399–406.

- [22] Meyer, K., Hall, G., Offin, D., 2009. Introduction to Hamiltonian dynamical systems and the N -body problem. volume 90 of *Applied Mathematical Sciences*. Second ed., Springer, New York.
- [23] Pérez-Marco, R., 2003. Convergence or generic divergence of the Birkhoff normal form. *Ann. of Math. (2)* 157, 557–574.
- [24] Richardson, D., 1980. Analytic construction of periodic orbits about the collinear points. *Celestial Mech.* 22, 241–253.
- [25] Szebehely, V., 1967. *Theory of Orbits*. Academic Press.
- [26] Zagouras, C., Kazantzis, P., 1979. Three-dimensional periodic oscillations generating from plane periodic ones around the collinear Lagrangian points. *Astrophys. Space Sci.* 61, 389–409.