ON THE ADVANTAGES OF USING A STRICT HIERARCHY TO MODEL ASTRODYNAMICAL PROBLEMS

Diogene Alessandro Dei Tos
Ph.D. candidate, Department of Aerospace Science and Technology, Politecnico di Milano, via la Masa 34, 20156, Milan, Italy; diogenealessandro.deitos@polimi.it

Francesco Topputo
Assistant Professor, Department of Aerospace Science and Technology, Politecnico di Milano, via la Masa 34, 20156, Milan, Italy; francesco.topputo@polimi.it

In this paper an algorithm is developed that combines the capabilities and advantages of several different astrodynamical models of increasing complexity. Splitting these models in a strict hierarchical order yields a clearer grasp on what is available.

With the effort of developing a comprehensive model overhead, the equations for the spacecraft motion in simpler models can be readily obtained as particular cases. The proposed algorithm embeds the circular and elliptic restricted three-body problems, the four-body bicircular and concentric models, an averaged $n$-body model, and, at the top hierarchic ladder, the full ephemeris space-based restricted $n$-body problem. The equations of motion are reduced to the assignment of 13 time-varying coefficients, which multiply the states and the gravitational potential to reproduce the proper vector field. This approach is powerful because it allows, for instance, an efficient and quick way to check solutions for different dynamics and parameters. It is shown how a gradual increase of the dynamics complexity greatly improves accuracy, the chances of success and the convergence rate of a continuation algorithm, applied to low-energy transfers.

I. INTRODUCTION

The circular restricted three-body problem, addressed CRTBP later in this work, is the ideal model to design unique solutions, ranging from Lagrange point orbits to low energy transfers [22]. These trajectories embed the effect of two gravitational attractions in a natural way, and therefore they are more accurate than the conics, solutions of the classic two-body problem. The phase space portrait of the CRTBP has been thoroughly studied in the past, with special attention to the dynamics in the neighbourhood of the collinear libration points [15, 11, 8]. This is because most of the dynamics of the restricted problem can be related to the equilibrium points and their invariant stable and unstable manifolds. From a practical perspective, these points possess properties that make them valuable candidates for space missions. Herschel, Planck and Gaia about the Sun–Earth $L_2$, and SOHO and Genesis about Sun–Earth $L_1$, are just few examples of typical libration points missions.

Several other missions are planned that make explicit use of three-body dynamics, such as Plato and Euclid. Finally, the concepts of weak stability boundaries and ballonic capture [13], proved reliable from the rescue of the Japanese spacecraft Hiten [3], will be applied to the ESA’s cornerstone mission BepiColombo [14].

However, when the three-body orbits are reproduced in more comprehensive gravitational models, large errors are found. That is, as the three-body orbits are defined in the regions of phase space where the sensitivity is high, the additional terms of solar system produce large effects along the orbits. An automatic algorithm able to correct orbits in the real solar system model is in these circumstances of great aid to space mission design.

Several works exist in literature that present a way to account for the full gravitational dynamics of the solar system. Dynamical substitutes are found by continuation in [10] and [11]; through a reduction to the centre manifold and by numerically imposing the solution to be quasi-periodic in [8]; or selecting a finite number of frequencies that reasonably represents the major contribution of the celestial bodies [9]. Corrections have also been obtained retaining a very large number of frequencies in an analytical power series expansion of the gravitational potential [12].

The purpose of this paper is to establish a clear hierarchy in the gravitational models available to the designer, and to explicitly exploit it for the continuation of typical three-body orbits in the $n$-body problem, modelled through precise ephemeris data. To achieve the objective, an automatic algorithm has been implemented. A tool is developed that combines the capabilities and advantages of several different astrodynamical models of increasing complexity. Splitting these models in a strict hierarchical order allows a clearer grasp on what is available. The differential equations governing the dynamics of a massless particle within the vector field generated by the $n$ celestial bodies in the solar system are written as perturbation of the CRTBP in a non-uniformly rotating and pulsating frame. In this way, the equations for the spacecraft motion in simpler models can be readily obtained as particular cases of those in the general model. The equations of motion are reduced to the assignment of 13 time-varying coefficients, which multiply the states and the gravitational potential to reproduce the proper vector field.

The refinement is carried out by means of a modified multiple shooting technique, and the problem is solved for a finite set of variables. This approach is powerful because it allows an efficient and quick way to check solutions for different dynamics and parameters. It is shown how a gradual increase of the dynamics complexity greatly improves accuracy, the
chances of success and the convergence rate of a continuation algorithm, applied to periodic orbits.

The approach used in this work possesses similar traits compared to the one developed in [16] and [20]. Nonetheless, the explicit exploitation of a gravitational hierarchy in the models represents a new approach. The results of the Earth–Moon system serve as solid benchmark to validate this procedure compared to others and prove this method correct and reliable. The results obtained in this work further improve the algorithm developed in [6] and [11]. The explicit exploitation of increasingly accurate gravitational models is a new way to tackle the search of trajectories within the $n$-body problem, aimed at efficiently finding trajectories with prescribed features.

This paper is organised as follows. In Section II the dynamical models that are used for the numerical computations are described, paying careful attention to their hierarchical order. Section III is the core of this work and details the algorithm that refines trajectories in the real solar system model. The methodology and numerical procedure are explained. The results are illustrated and discussed in Section IV, where families of Halo orbits are corrected with the proposed algorithm. Lastly, the conclusions are drawn in Section V.

II. DYNAMICAL MODELS

A great variety of astrodynamical models are available to the designer. As the complexity of such models increases new solutions appear due to the richer content of the vector field. The drawback is that no analytical solution is available, and consequently it’s very difficult to have a general insight on the solutions. Extensive computational searches are usually required in order to hit the desired optimal trajectory.

In this section the main astrodynamical models are shown to be a particular case of the roto-pulsating $n$-body problem. With this approach, a single set of equations can be used to represent the whole domain of possible gravitational models, simply varying the coefficients and the potential function of the model on the top of the hierarchic ladder.

II.1 The problem of $n$ bodies

The most general model for the description of the motion of a massless particle subjected to the gravitational field of other $n-1$ celestial bodies is the $n$-body problem, whose geometry is shown in Figure 1. The dynamics of the particle $P_k$ of mass $m_k$, $k = 1, \ldots, n$, whose Cartesian coordinates are $\mathbf{R}_k = (X_k, Y_k, Z_k)^T$ is governed by Newton’s universal law of gravitation. Applying Newton’s second law of motion and assuming constant masses, the equation may be written as:

$$m_k \ddot{\mathbf{R}}_k = \sum_{j=1, j \neq k}^{n} \frac{G m_j m_k}{R_{jk}^3} (\mathbf{R}_j - \mathbf{R}_k) \quad k = 1, \ldots, n \quad (1)$$

Eq. (1) is written in an inertial reference frame and represents a set of 6n first order ordinary differential equations.

However, Astrodynamics is mainly concerned with the study of artificial objects, in this contest the hypothesis of restricted dynamics has been applied with accurate results. The artificial object moves in the vectorial field created by the $n$ celestial bodies, without affecting their motion. Another significant simplification can be obtained if the trajectories of the primaries are proper time-dependent functions. Let $S$ be the set of celestial bodies, the motion of an artificial satellite is represented by

$$\ddot{\mathbf{R}} = \sum_{j \in S} \mu_j \frac{\mathbf{R}_j - \mathbf{R}}{||\mathbf{R}_j - \mathbf{R}||^3} \quad (2)$$

where $\mu_j = G m_j$ is the mass parameter, $G$ is the universal constant of gravitation, and $\mathbf{R} = (X, Y, Z)^T$ is the position of the artificial satellite.

The solar system model used throughout this work consists in Newton equations for the restricted $n$-body problem. We avail ourselves of the JPL ephemeris data DE430 [7] to determine in a precise way the states of the Sun, the planets and the Moon at given epochs with respect to an inertial reference frame whose origin is located at the solar system barycentre. More precisely, the Spice toolkit has been used to determine the states of the celestial bodies [17].

The equations of the solar system restricted $n$-body problem, or simply SSRnBP, are written as perturbation of the CRTBP, by means of a time-dependent coordinates transformation [9], in this way a better insight of each term can be attained. Moreover, this model fits well in the hierarchical approach followed in this paper. Let $\mathbf{R}$ and $\mathbf{V}$ be the dimensional position and velocity, respectively, of a massless body $P_1$ in the inertial solar system barycentric frame, and let $\rho$ be its adimensional position in the new rotating and pulsating reference frame, defined by a pair of primaries $P_1$ and $P_2$ (see Fig. 2). The transformation between the solar system barycentric reference frame, and the new non-inertial reference is then

$$\mathbf{R} = \mathbf{b} + k \mathbf{C} \rho$$

where

$$\mathbf{b}(t) = \frac{m_1 \mathbf{R}_1 + m_2 \mathbf{R}_2}{m_1 + m_2}$$

$$k(t) = ||\mathbf{R}_2 - \mathbf{R}_1||$$

$$\mathbf{C}(t) = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$$
and

\[
\begin{align*}
\mathbf{e}_1 &= \frac{R_2 - R_1}{k} \\
\mathbf{e}_2 &= \mathbf{e}_3 \times \mathbf{e}_1 \\
\mathbf{e}_3 &= \frac{(V_2 - V_1) \times (R_2 - R_1)}{|| (V_2 - V_1) \times (R_2 - R_1) ||}
\end{align*}
\]

In Eq. (4) and Eq. (5), \(m_i, R_i, \) and \(V_i\) are the mass, position, and velocity of \(P_i\), respectively, \(i = 1, 2\). The transformation is hence composed by two parts. The first is a translation of the frame centre from the solar system barycentre to the primaries centre of mass, \(b\). The second is a rotation by means of the orthogonal cosine angle matrix \(C\), and a scaling by means of the time-dependent factor \(k\). The rotation is such that the primaries are always aligned with the \(x\)-axis of the new frame. The scaling factor \(k\), which is the actual distance between the primaries, adjusts their positions so as to be fixed in time with respect to the new frame of reference. As a result, the new frame rotates in a non-uniform fashion and pulsates in order to guarantee some convenient features, primarily suggested from the CRTBP. In this paper the new gravitational model will be addressed as *roto-pulsating n-body problem*, or simply \(\text{RPnBP}\).

The Lagrangian of the complete gravitational model is:

\[
\mathcal{L}(\mathbf{R}, \dot{\mathbf{R}}, t) = T - V = \frac{1}{2} \dot{\mathbf{R}} \cdot \mathbf{R} + \sum_{j \in S} \frac{Gm_j}{|| \mathbf{R} - \mathbf{R}_j ||}
\]

where \(S\) is the set of all celestial bodies included in the solar system model, except for the spacecraft itself, and \(m_j\) their masses. The dots indicate derivation with respect to dimensional time, \(t\).

Without loss of generality, a constant time reference is chosen for the adimensionalisation equal to the mean motion of the primaries, \(n\). With this choice the average primaries revolution period is \(2\pi\), that is

\[
\tau = n(t - t_0) = \frac{2\pi}{T} = \sqrt{\frac{G(m_1 + m_2)}{a^3}}
\]

where \(a\) is the mean distance between the primaries for a long time interval (a period of 400 years will suffice). Note that the mean distance between primaries is not the semi-major axis of the ellipse described by their orbits, the eccentricity modifies it by a factor \(\sqrt{1 - e^2}\). The initial epoch \(t_0\) is used to shift the non autonomous problem to have null initial adimensional time.

Applying the transformation [9] to the Lagrangian, and then carrying out the Lagrangian mechanics operations, we obtain the equations of motion in the roto-pulsating frame of reference. After some manipulation [10]:

\[
\rho'' = \frac{1}{n} \left( \frac{k}{k} \frac{\rho - \rho_b}{\rho - \rho_\ast} + \frac{1}{n^2} \left( \frac{k}{k} \frac{\rho - \rho_\ast}{\rho - \rho_b} + \frac{1}{n^2} \rho \right) \Omega \right) + \frac{\rho_b}{n^2 k^2} \Omega
\]

in which primes indicate derivatives with respect to the adimensional time, \(\tau\), and \(\Omega\) is the potential function of the \(\text{RPnBP}\) defined as

\[
\Omega = (1 - \mu) \frac{\rho - \rho_1}{|| \rho - \rho_1 ||^3} + \mu \frac{\rho - \rho_2}{|| \rho - \rho_2 ||^3} + \sum_{j \in S^*} \hat{\rho}_j \frac{\rho - \rho_j}{|| \rho - \rho_j ||^3}
\]

where \(S^*\) is the collection of celestial bodies except for the primaries, and \(\hat{\rho}_j = m_j/(m_1 + m_2)\) is a convenient form to express their mass parameters. Mixed derivative notation stems acknowledging that ephemeris data is numeric, discrete, and provided for regular dimensional time. The vector Eq. [5] might be written per components,

\[
\begin{align*}
x'' &= b_1 x' + b_2 y' + b_3 y + b_4 z + b_5 z + b_{13} \Omega_{/z} \\
y'' &= b_2 x' + b_3 y' + b_4 y - b_5 y - b_{10} y + b_{12} \Omega_{/y} \\
z'' &= b_3 x' + b_4 z' + b_5 z - b_{11} z - b_{12} z + b_{13} \Omega_{/z}
\end{align*}
\]

with coefficients

\[
\begin{align*}
b_1 &= \frac{\mathbf{b} \cdot \mathbf{e}_1}{kn^2} & b_7 &= -\frac{1}{n^2} \left( \frac{k}{k} \mathbf{e}_1 \cdot \mathbf{e}_1 \right) \\
b_2 &= \frac{\mathbf{b} \cdot \mathbf{e}_2}{kn^2} & b_8 &= \frac{1}{n^2} \mathbf{e}_1 \cdot \mathbf{e}_3 \\
b_3 &= \frac{\mathbf{b} \cdot \mathbf{e}_3}{kn^2} & b_9 &= \frac{1}{n^2} \left( \frac{k}{k} \mathbf{e}_2 \cdot \mathbf{e}_1 + \mathbf{e}_2 \cdot \mathbf{e}_1 \right) \\
b_4 &= \frac{1}{n^2} \frac{k}{k} & b_{10} &= -\frac{1}{n^2} \left( \frac{k}{k} \mathbf{e}_2 \cdot \mathbf{e}_2 \right) \\
b_5 &= \frac{2}{n^2} \mathbf{e}_1 \cdot \mathbf{e}_1 & b_{11} &= \frac{1}{n^2} \left( \frac{k}{k} \mathbf{e}_3 \cdot \mathbf{e}_2 + \mathbf{e}_3 \cdot \mathbf{e}_2 \right) \\
b_6 &= \frac{2}{n^2} \mathbf{e}_3 \cdot \mathbf{e}_2 & b_{12} &= -\frac{1}{n^2} \left( \frac{k}{k} \mathbf{e}_3 \cdot \mathbf{e}_3 \right) \\
b_{13} &= \frac{\mu \mathbf{S} + \hat{\mu} \rho}{kn^2} 
\end{align*}
\]

These are the equations used throughout this work. Note that a straightforward balance between centrifugal and gravitational forces would require the coefficient \(b_{13}\) to be unity when considering just two main bodies (i. e., the CRTBP). In the general case this coefficient will oscillate about this value due to the variability of \(k\).

II.II  Circular restricted three-body problem

The lower level in the hierarchic ladder is represented by the circular restricted three-body problem, or simply CRTBP. Let us consider a body \(P_3\) of mass \(m_3\) in the vector field of
two primaries \( P_1, P_2 \) of masses \( m_1 \) and \( m_2 \), respectively, such that the condition \( m_3 \ll m_2 < m_1 \) is satisfied. In the CRTBP the primaries revolve in planar configuration at constant angular speed. The motion of the third body is studied in a rotating reference frame, named synodic reference frame, whose origin is located at the primaries centre of mass, the \( x \) axis is always aligned with the \( P_1P_2 \) direction, the \( z \) axis is orthogonal to the primaries plane of motion, and the \( y \) axis forms a right-hand term. By means of a proper adimensionalisation \([19]\) the equations of motion depend only on the mass parameter, defined as \( \mu = m_2/(m_1 + m_2) \). The adimensionalisation is such that the distance between the primaries, their angular speed and the sum of their masses are set to a unity value. In this system the positions of \( P_1 \) and \( P_2 \) are fixed, being \( P_1 \) located at \((-\mu, 0, 0)\) and \( P_2 \) at \((1 - \mu, 0, 0)\). The equations of motion read

\[
\ddot{x} - 2\dot{y} = \Omega_y^{(3)} \quad \ddot{y} + 2\dot{x} = \Omega_x^{(3)} \quad \ddot{z} = \Omega_z^{(3)} \tag{12}
\]

where the three-body potential function can be expressed as

\[
\Omega^{(3)} = \frac{1}{2}(x^2 + y^2) + \frac{1 - \mu}{r_1} + \frac{\mu}{r_2} + \frac{1}{2}\mu(1 - \mu) \tag{13}
\]

and terms \( r_1 = \sqrt{(x + \mu)^2 + y^2 + z^2} \) and \( r_2 = \sqrt{(x - 1 + \mu)^2 + y^2 + z^2} \) are the scalar distances between the third mass and the primaries.

The dynamic equations describing this model can hence be seen as a particular case of the more general RP\(n\)BP, and are indeed obtained by simply assigning proper values to the coefficients, i.e., \( b_i = 0 \) for \( i \neq 7, 10, 13 \), \( b_7 = b_{10} = b_{13} = 1 \), and \( b_5 = 2 \).

II.III Elliptic restricted three-body problem

The next step in the hierarchy is the elliptic restricted three-body problem, or ERTBP. This model studies the motion of a massless particle, \( P \), under the gravitational field generated by the mutual elliptic motion of two primaries, \( P_1 \) and \( P_2 \), of masses \( m_1, m_2 \), respectively. The equations of motion for \( P \) are \([19]\) \([13]\)

\[
x'' - 2y' = \omega_x \quad y'' + 2x' = \omega_y \quad z'' = \omega_z \tag{14}
\]

The subscripts in Eq. \(14\) mean the partial derivative of

\[
\omega(x, y, z, \theta) = \frac{\Omega^{(3)}}{1 + e_p \cos \theta} \tag{15}
\]

where the potential function is the same defined in the CRTBP, Eq. \(13\). Primes denote derivatives with respect to the new independent variable: the true anomaly, \( \theta \).

Eqs. \(14\) are written in a non uniformly rotating, barycentric, adimensional coordinate frame where \( P_1 \) and \( P_2 \) have fixed positions \((-\mu, 0, 0)\) and \((1 - \mu, 0, 0)\), respectively, and \( \mu \) is the mass parameter of the system as per the CRTBP. This coordinate frame isotropically pulsates as the \( P_1P_2 \) distance, assumed to be the unit length. It varies according to the mutual position of the two primaries with respect to \( \theta \), the true anomaly of the system. This is the independent variable, and plays the role of time: \( \theta \) is equal to zero when \( P_1P_2 \) are at their periastrides, as both primaries orbit their barycentre in similarly oriented ellipses having common eccentricity \( e_p \).

Normalising the period of the primaries to \( 2\pi \), the dependence of true anomaly on time is

\[
\frac{d\theta}{dt} = \frac{(1 + e_p \cos \theta)^2}{(1 - e_p)^3/2} \tag{16}
\]

Unlike the CRTBP, the true anomaly in Eq. \(15\) makes the elliptic problem nonautonomous. Thus, any qualitative feature of this problem strictly depends on the true anomaly, \( \theta \).

The coefficients of the elliptic three-body problem are not constant, but depend on the true anomaly.

\[
b_5 = 2 \quad b_{12} = -\frac{e_p \sin \theta}{1 + e_p \cos \theta} \tag{17}
\]

Note that these coefficients reduce to those of the CRTBP if the eccentricity of the primaries is zero, \( e_p = 0 \).

II.IV Restricted four-body problem

Including an additional massive body into the bargain can potentially lead to completely different dynamical behaviour and solutions. As the geometry becomes more complex, so does the hierarchical ladder. There is a bifurcation of models at this point: 1) a system of two bodies revolves about a massive celestial body, leading to the bicircular behaviour; or 2) two bodies revolve in similar fashion around a massive primary, that is the circular concentric hypothesis.

II.IV.1 Bicircular four-body model

The bicircular problem, addressed BCP hereafter, is a restricted non-coherent model that considers two primaries and a third gravitational perturbation:

1. two primaries \( P_1 \) and \( P_2 \) (e.g., the Earth and the Moon) revolve in circular orbits around their barycentre, \( B \);
2. at the same time, \( B \) and the third body \( P_3 \) (e.g., the Sun) are moving in circular orbits around the centre of masses of the whole system (e.g., Earth–Moon–Sun), \( B' \);
3. the primaries and the third body moves in the same plane.

For the sake of clarity the system Earth–Moon–Sun is analysed here. The results can easily be generalised and extended to other selection of primaries with the proper adjustments in the model parameters. Fig. 6 shows the geometry of the BCP. Beginning with an inertial frame, we perform a change of variables to write the equations in synodical Earth–Moon coordinates. However, in this case attention must be exerted in the transformation since the Earth–Moon barycentre ceases to be an inertial point, due to the perturbation of the third body. From Newton’s law the derivation of the equations of motion for a massless particle \( P \) is straightforward. The adimensionalisation paradigm closely follows the one used for the CRTBP. In particular, the adimensionalisation is such that the distance between the primaries, their angular speed and the sum of their masses are set to a unity value. In this system the positions of \( P_1 \) and \( P_2 \) are fixed, being \( P_1 \) located at \((-\mu, 0, 0)\) and \( P_2 \) at \((1 - \mu, 0, 0)\). To properly define the state
of the fourth body in adimensional coordinates, new mean quantities have to be introduced: $m_3$, the Sun mass, $a_3$, the distance between the Earth–Moon barycentre and the Sun, $m_3$ and $\omega_3$ the mean angular velocity of the Sun in inertial and synodic coordinates, respectively. All these quantities are not independent. The following equalities hold:

\begin{align}
\omega_3 &= 1 - n_3 \\
a_3^3 m_3^2 &= 1 + m_3 \\
\alpha &= \omega_3 t
\end{align}

The first one is easy to understand if we remind of the mean angular velocity of the Moon is 1, in synodic coordinates. The second is a consequence of the third Kepler’s law in adimensional coordinates. In the third equality, $\alpha$ is the phase angle between the Earth–Moon line and the Sun. The values used for these quantities are [1]

\begin{align}
a_3 &= 388.81114 \\
\omega_3 &= 0.9291959855 \\
m_3 &= 328900.54
\end{align}

With these parameters the position of the Sun in the Earth–Moon synodic frame can be written as

$$\rho_S = a_S \begin{bmatrix} \cos(\alpha + \alpha_0) & \sin(\alpha + \alpha_0) & 0 \end{bmatrix}^T$$

where $\alpha_0$ is the initial phase angle of the Sun, which depends on the initial epoch and makes the BCP nonautonomous.

The equations of motion read

$$\ddot{\rho} + C^T (2\nu \dot{\rho} + C^T \dot{\rho}) = (\mu_2 - 1) \left[ \frac{\rho - \rho_S}{||\rho - \rho_S||^3} - (1 - \mu_1) \frac{\mu_1 \rho + \rho_S}{||\mu_1 \rho + \rho_S||^3} + \mu_1 \left( \frac{1 - \mu_1}{||1 - \mu_1||} \right) \frac{r - \rho_S}{||r - \rho_S||^3} \right]$$

where $\rho$ and $\rho_S$ are the adimensional positions of the massless particle and the Sun, respectively, $r = (1, 0, 0)$ is the vector from the Earth to the Moon, $C$ is the rotation matrix of the Earth–Moon synodic frame. Let $m_E, m_M, m_S$ be the masses of the Earth, Moon, and Sun, respectively, then

$$\mu_1 = \frac{m_M}{m_E + m_M}$$

is the mass parameter of the Earth–Moon system. Primes denote here derivatives with respect to the adimensional time.

We obtain a set of equations that are similar to the equations of the CRTBP, and that surprisingly possesses the very same set of coefficients of the CRTBP. The BCP is hence considered as a not so small perturbation of the CRTBP, next step of the hierarchy. The major difference between the two models lays in the definition of the potential function. Indeed, the BCP potential embeds three new terms, between square brackets in Eq. (23). The first is simply due to the Sun attraction, whilst the other two stem from the nonzero acceleration of the Earth–Moon centre of mass that depends on the actual distances between the primaries and the Sun.

The BCP is suitable to every gravitational system in which one small body is orbiting a larger one, e.g., a natural satellite system, which is on turn orbiting another very massive celestial body, e.g., a star. The gravity of the more massive body can be treated as a perturbative action. Examples can be readily found in the solar system: Sun plus Earth–Moon, Mars–Phobos, Jupiter–Europa, Saturn–Titan, Neptune–Triton, and so on.

II.IV.2 Concentric circular four-body problem

The concentric circular four-body problem, or simply CCP, is the second horizontal hierarchical part of the restricted four-body problem. It is a restricted non-coherent model that, unlike the BCP, considers one primary and two secondary bodies. It is assumed that the motion of a massless object is governed by three primaries $P_1, P_2, P_3$, of masses $m_1, m_2, m_3$, respectively. One of the primaries is much more massive than the other two, $m_1 \gg m_2, m_3$. In a quasi-inertial reference frame centred at $P_1$, the bodies $P_2$ and $P_3$ rotate about $m_1$ in circles of radii $r_2$ and $r_3$ and with angular velocities $\omega_2$ and $\omega_3$, respectively. The circular orbits are coplanar. The geometry of the CCP is shown in Fig. [4].

The equations of motion for $P$ are first written as perturbation of a simple Kepler problem [2], where the main attractor is $P_1$.

$$\ddot{r} + \mu_1 \frac{r}{||r||^3} = - \sum_{j=2}^N \mu_j \left( \frac{d_j}{||d_j||^3} + \frac{r_{j1}}{||r_{j1}||^3} \right)$$

where $d_j = (x_j - x_1)/||x_j - x_1||$ is the position vector of the secondary $j$ relative to $P_1$.
where \( \mu_j = Gm_j \), \( d_j = R - R_j \) is the distance between the massless particle and the \( j \)th small attractor, and \( r_j = R_j - R_1 \) is the distance between the massive central body and the small attractors \( j = 2, 3 \), respectively. Eq. (24) is transformed into a \( P_1 P_2 \) synodic frame by means of a proper rotation \( \mathbf{r} = r_2 \dot{C} \rho \). The new position vector is hence \( \rho = (x, y, z)^T \). Dimensionless time is again obtained through the mean motion of \( P_2 \) about \( P_1 \). The equations of motion read
\[
\begin{align*}
x'' - 2y' &= \Omega^{(CCP)}_{f}\frac{x}{\rho} \\
y'' + 2x' &= \Omega^{(CCP)}_{f}\frac{y}{\rho} \\
z'' &= \Omega^{(CCP)}_{f}\frac{z}{\rho}
\end{align*}
\]
(25)
where the potential for the CCP is defined as
\[
\Omega^{(CCP)} = \frac{1}{2} (x^2 + y^2) + \frac{1 - \mu_1}{\|\rho\|} + \mu_1 \left( \frac{1}{\|\rho - \rho_2\|} \right)
\]

(26)
and \( \mu_1 = \frac{m_2}{m_1 + m_2} \) is the synodic mass parameter, and \( \mu_3 = \frac{m_2}{m_1} \) is the perturbation mass ratio. The position vectors of the two small attractors in the new rotating frame are
\[
\rho_2 = (1, 0, 0)^T \quad \rho_3 = \left( \frac{r_3}{\|r_3\|} \cos \varphi, \sin \varphi, 0 \right)^T
\]
(27)
where the phase angle of \( P_2 \) with respect to the line \( P_1 - P_2 \) in synodic coordinates is \( \varphi = \dot{\omega} t + \varphi_0 \), and the apparent motion of \( P_3 \) about the synodic frame is \( \dot{\Omega} = \frac{d_j}{2} - 1 \). The CCP is nonautonomous due to the initial phase angle \( \varphi_0 \), that depends on the relative position of the celestial bodies at the initial epoch.

The coefficients of the circular concentric four-body model are the very same of the CRTBP. This should not surprise because the baseline dynamics has not changed, being the well-known synodic three-body coordinates. The perturbation of the fourth body varies however the potential function.

Extreme care should be exerted when using the circular concentric model. Indeed, due to the derivation of the motion equations, the centre of the reference frame is not the barycentre of the synodic coordinates, but it is the massive body directly, assumed to be quasi-inertial.

II.V Averaged coefficients

The step before using ephemeris data to calculate the states of the celestial bodies is the creation of a database that contains the information and data on several choices of synodic reference frames. In particular the equations of motion are the same as the RPNBP, Eqs. (10). However, the coefficients are not variable functions of time, they are instead constant values that represent the average value of that coefficient for a selected pair of primaries. The average process is a time-average mathematical operator applied to the coefficients, long time spans are required to include the dynamics of all the solar system bodies. This operation avoids computing the coefficients in Eqs. (11), which is expensive. The average is applied to all the possible selection of primaries.

### Tab. 1: Coefficients the Earth–Moon restricted hierarchic models

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>CRTBP</th>
<th>BCP</th>
<th>CCP</th>
<th>ERTBP</th>
<th>Mean model</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b_1 )</td>
<td>0</td>
<td>0</td>
<td>-3.209e-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_2 )</td>
<td>0</td>
<td>0</td>
<td>3.223e-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_3 )</td>
<td>0</td>
<td>0</td>
<td>5.753e-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_4 )</td>
<td>0</td>
<td>0</td>
<td>-2.993e-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_5 )</td>
<td>2</td>
<td>2</td>
<td>2.0000271</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_6 )</td>
<td>0</td>
<td>0</td>
<td>-1.727e-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_7 )</td>
<td>1</td>
<td>1</td>
<td>1.00478</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_8 )</td>
<td>0</td>
<td>0</td>
<td>5.399e-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_9 )</td>
<td>0</td>
<td>0</td>
<td>-1.253e-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_{10} )</td>
<td>1</td>
<td>1</td>
<td>1.00478</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_{11} )</td>
<td>0</td>
<td>0</td>
<td>8.551e-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_{12} )</td>
<td>0</td>
<td>0</td>
<td>-0.001611</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b_{13} )</td>
<td>1</td>
<td>1</td>
<td>1.00747</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As an example Tab. 1 lists the 13 coefficients of the Earth–Moon mean problem, calculated with a time average procedure on a 359.1-year span. For comparison the tables also displays the coefficients of the other gravitational models outlined so far. The mean values of the coefficients do not detach very much from the CRTBP. Perhaps, the only exception being isolated to coefficient \( b_2 \) that relates the second time derivative of \( z \) to its linear part, \( z \). Fig. 5 displays the absolute value maximum and minimum percent differences between some of the coefficients of the CRTBP and ERTBP, defined as \( \Delta_{\text{CRTBP}} b_j(f, e_p) = 100 \cdot (b_j^{(\text{CRTBP})} - b_j^{(\text{ERTBP})}) \), \( j = 7, 10, 12, 13 \). It can be seen that for typical primaries within the solar system (roughly 80% of primaries has \( e_p < 0.1 \)) the eccentricity contribution does not significantly vary.
the coefficients of the model. However, as a critical case, the coefficients of the Sun–Mercury ERTBP exhibits variations of up to 25% with respect to those of the CRTBP.

The harmonic content of the dynamics is clearly visible in Fig. 6, displaying the Fourier transform of $b_9$. The peaks in the Fourier transform correspond firstly to the rotopulsating main frequency, and secondly to the largest perturbations which affects the synodic dynamics. As for any Fourier procedure, the most relevant parameters to be specified are the size, $T_0$, of the sampling interval, and the number, $N$, of equally spaced sampling points within such interval. These parameters define the Nyquist critical frequency, $f_N = \frac{N}{2T_0}$, that fixes the window within which the frequencies (true or aliased) will be found. The number of samples has been chosen such that the Fourier analysis can detect frequencies at least equal to \( \frac{1}{T_0} \), and at the same time to let the solar system completely exhaust its dynamics. A period larger than 250 years suffice the purpose. What is more, the number of samples chosen has been selected as a power of two; due to requirements of the Fast Fourier Transform (commonly known as FFT) algorithm, used to carry out the transforms in a fast and efficient way.

As far as the Earth–Moon system is concerned, the coefficients span an interval of 359.1 years with $N = 2^{18}$ samples, which results in a time rate of 12 hours; on the other hand, the Sun–Mercury systems coefficients are sampled with $N = 2^{19}$ with a time rate of 12 hours, providing a totality of roughly 718.2 years.

Finally, in order to reduce the leakage, the functions to be transformed are truncated by means of a Hanning window function of order 2:

$$H_{T_0}(t) = \frac{2}{3} \left(1 - \cos \frac{2\pi t}{T_0}\right)^2$$

As expected, the majority of the perturbative contributions appear in the Fourier transform of some coefficients. Note, however, that not all the coefficients feel the same perturbative effects. For example, considering the Earth–Moon case, coefficients $b_1$, $b_3$ and $b_5$ show a completely different harmonic trend. As far as the Earth–Moon system is concerned, the perturbation with highest Fourier transform magnitude has period of 32 days roughly; this contribution is conjectured to be caused by the Sun.

To sum up, Fig. 7 shows the flow diagram associated to the hierarchy of the gravitational models. It’s interesting to note the horizontal behaviour of the four-body models, indispensable to consider different kind of relative position for the perturbations.

---

\(1\) Note that Pluto has an orbital period of roughly 247 years.
II. VI Potential functions

Increasing the complexity of the gravitational model in hierarchical order has lead to new dynamical equations whose coefficients do not vary significantly from the CRTBP. In the BCP and CCP cases the coefficients remain even unmodified. This is because the nonlinear terms are mostly contained by a proper redefinition of the potential function, $\Omega$. Fig. 8 shows coloured parametric surface of section of the potential function of the BCP, on the left, and its percentage difference with the CRTBP, on the right. In order to produce a bidimensional gradient result, the height $z$ has been set to 0. Moreover, being the bicircular problem non autonomous, the epoch has been chosen at the date March 11th 2015, 12h 56m 00s TDB. With this choice the initial phase angle of the Sun with respect to the Earth–Moon line is roughly 115°. Several tests have been run with different epochs within a lunar period. The areas of minimum and maximum percentage variation follow the Sun direction (black arrow) and the whole picture rotates counterclockwise. Except this rotating variation, the main shape of the potential seems not to vary significantly within roughly 28 days from the result presented on the right figure.

The typical regions of motion associated to the Jacobi constant are maintained. The surfaces of zero-velocity, even though now are variable, have retained the typical progression shape: oval, dumbbell, horseshoe, and tadpole. The fundamental difference is that the perturbation of the Sun creates regions at higher and lower levels of potential. The main consequence is that there exist special configurations of the spacecraft at certain epoch where the energy required to escape the Earth–Moon attraction and open the passage at $L_1$ and $L_2$ is lower. Conversely, the opposite behaviour exists as well; the Sun can indeed act upon the spacecraft to favour the stability of a capture orbit in the synodic system. It is very peculiar to see how the action of the Sun is symmetric. The central region has very low values of variation, and bifurcates along the direction normal to the Sun resulting in bigger regions with negative variation of potential. The remaining regions, clustered in the Sun and anti-Sun directions have positive variation of potential. The small boundary regions of the saddle that have zero variation of potential are the regions where the perturbation of the Sun does not play any role in the vector field of the massless particle, and presumably are the best conditions to propagate Halo-type orbits.

III. METHODOLOGY

In this section the attention is focused on the logical and mathematical procedure to continue trajectories calculated in the CRTBP toward the RPNBP. The selection of the orbit to refine is based on two factors: the stability of the orbit within the three-body frame, and the sensibility to the increased chaotic content associated to the larger number of retained celestial bodies.

Halo-like orbits are deemed to satisfy these requirements, and, on top of that, they are being widely used for current space missions. The computation of these orbits must account for the non-linear terms that arise in the linearised CRTBP when large amplitude orbits are considered. These solutions are obtained through a numerical approach, based on perturbation techniques, in order to correct the analytic initial estimates, and on continuation techniques, in order to expand the infinitesimal orbits. Halo orbits are periodic orbits which bifurcate from the planar Lyapunov orbits when the in-plane and out-of-plane frequencies of the linearised vector field are equal. This is a 1:1 resonance that appears as a consequence of the nonlinear terms of the equations and, hence, these 1-D invariant tori have to be searched as series expansion with a single frequency. In details, once the out-of-plane $A_z$ amplitude overcomes a limit value, the frequency of the in-plane oscillatory motion achieves the value of the frequency of the one out of the plane, and three-dimensional halo orbits emerge. Fig. 9 represents a family of Halo orbits of the $L_1$ Sun-Jupiter system, calculated via high-order differential corrections [21].

![Fig. 8: Potential function, $\Omega(x, y, 0)$, colour-gradient visualisation](image-url)
In this work, a *Lindstedt–Poincaré method* is used to compute accurately some of the solutions of the centre manifold. This process is based on finding a parametric family of trigonometric expansions that satisfy the equations of motion, up to a sufficiently high order. The potential function is expanded by means of Taylor series and terms up to the second order have been retained. According to [13], Legendre polynomial are used to retain high-order terms. The solution of the linear periodic part of these equations remains unvaried. These linear solutions are already Lissajous trajectories. Following the procedure in [6], when the nonlinear terms (in Legendre polynomials form) are considered, the complete solution is sought as formal series in powers of the amplitudes \( A_i \) and \( \Delta A_i \).

The main objective is to continue these Halo orbits in the RPnBP exploiting the hierarchy of the gravitational models to smooth the nonlinear terms gradients and to ensure the success of the refinement procedure. These orbits have already been demonstrated to exist in the complete ephemeris model [6], even though with small quasi-periodic modification in the baseline shape and oscillation frequency. However, the computation of these trajectories is computationally very extensive and require one ad-hoc procedure that is very cumbersome from the mathematical perspective and that can fail according to the precision of the initial condition and to the time scale used. We argue that this is because the passage from CRTBP to RPnBP is too sharp, and the algorithm might encounter difficulties in the minimisation of the objective function at hand. Applying the same algorithm sequentially from the CRTBP to the RPnBP and following the hierarchic order of the gravitational models established in this paper, convergence is attained and the algorithm is prone to easily find local optima that satisfy the requirements.

The refinement procedure, inspired by [16], is achieved by means of an iterative algorithm that consists of two steps: 

1. **Evaluation of a compliant initial seed orbit**
2. **Modified multiple shooting**

The two-point boundary value problem (TPBVP) is formulated slightly different from usual, leading to the name modified multiple shooting. In particular, the technique has to cope with the fact that no boundary conditions are actually known, and the sole requirement is to produce a piecewise continuous trajectory which stays as ‘close’ as possible in phase space to the initial seed. In order to attain this, the multiple shooting is coupled with an optimisation. In the first place, the classical optimal problem is translated into a *non-linear programming* (often termed NLP) method by means of direct transcription of the dynamics and the problem is then solved for a finite set of variables when a proper objective function is specified [4, 5]. As opposite to the optimal control problem, no dynamics is involved into a NLP problem, because in this case the dynamics is merely seen as a constraint that the NLP must satisfy. Let \( T_0 \) be the initial epoch, which should be specified due to the non-autonomous nature of the \( n \)-body problem, and \( \Delta T \) the time-span covered by a certain set of nodes. The basic procedure for trajectories refinement consists basically of the following steps:

**Step 1** Using a simplified gravitational model, generate a sequence of nodes as initial guess for Step 2;

**Step 2** Fix the initial epoch, \( T_0 \), and for a given time-span \( \Delta T \), perform the modified multiple shooting with the initial guess;

### III.I The modified multiple shooting

The fundamental principles in the classical multiple shooting technique are preserved, and eventually the zero of a non-linear multi-variable function has to be found by Newton method. However, the classic multiple shooting has been modified so as to deal with free boundary conditions. In particular, the equations that represent the boundary conditions are erased. Only the collection of defect vectors, here termed \( c \), is maintained.

\[
\begin{bmatrix}
  c_1(s_1, s_2) \\
  c_2(s_2, s_3) \\
  \vdots \\
  c_{m-1}(s_{m-1}, s_m)
\end{bmatrix}
= 
\begin{bmatrix}
  \zeta_1 \\
  \zeta_2 \\
  \vdots \\
  \zeta_{m-1}
\end{bmatrix}
= 0 \quad (29)
\]

where the defect vector is the difference between the flow of the dynamical equations, \( \varphi \), and the control point, as shown in Fig. 10

\[
\zeta_k = \varphi(t_{k+1}; t_k, s_k) - s_{k+1} \quad k = 1, \ldots, m-1 \quad (30)
\]

Note that, for the \( n \)-body problem, \( c(s) : \mathbb{R}^{6m} \rightarrow \mathbb{R}^{6(m-1)} \) and \( s \in \mathbb{R}^{6m} \). The problem is to seek the zero of the function \( c(s) = 0 \). An optimisation procedure is implemented in order to shatter the underdetermination and provide for the 6 missing equations. Let \( f(s) : \mathbb{R}^{6m} \rightarrow \mathbb{R}^3 \) be the objective function.
function, the minimisation problem is stated as:

$$\min_s \frac{1}{2} c^T c \quad \text{subject to} \quad c(s) = 0 \quad (31)$$

The problem is solved numerically. The Jacobian of the function $c$, $J_c$, is calculated by means of an approximated forward finite difference scheme.

$$J_c(s) = \begin{bmatrix} \Phi_1 -I_{nXn} & 0 & 0 \\ 0 & \Phi_2 -I_{nXn} & 0 \\ 0 & \ldots & \Phi_{m-1} -I_{nXn} \end{bmatrix} \quad (32)$$

where the state transition matrices, $\Phi_k$, are given by

$$\Phi_k = D_{s_k} F_k(s) = D_{s_k} \varphi(t_{k+1}; t_k, s_k) \quad (33)$$

where $D_{s_k}(\cdot)$ is the gradient with respect to $s_k$.

Basically the method consists of three steps:

1. calculate the Halo orbit and store the initial conditions;
2. apply the continuation method described above to make a step further in the hierarchic ladder and update the initial conditions;
3. if the model is RPNBP stop, otherwise reiterate from step 1.

### IV. RESULTS

The initial condition for a Earth–Moon $L_1$ Halo orbit are flown under the vector gravitational fields in the hierarchy. As expected, the CRTBP reproduces again the original Halo orbit. Within the other dynamics, the orbit deviates from the original path and follows different trajectories, which might in some cases be substantially different from the expected quasi-periodic behaviour. This is because the other celestial bodies produces perturbation along the orbit that are not negligible. Fig. 11 shows the propagation of the Halo orbits in the different dynamical models. The BCM stays very close to the CRTBP, making more than a complete revolution along the Halo orbit before escaping towards the Earth. This happens because at the selected initial the perturbation of the Sun is minimum along the trajectory. However, as time passes the Earth–Moon system revolves about the Sun, whose action eventually pulls away the orbit. The RPNBP and the ERTBP behave in similar fashion near the libration point, but detaches very quickly from the Halo trajectory and are temporarily captured in a selenocentric orbit. These two models produce similar results because at such distances the effect of the eccentricity of the lunar orbit produces greater effects than the solar gravity. Lastly, the averaged coefficients model is not able at all to simulate the nonlinearity of the gravitational field, and results in a very fast escape from the desired configuration. This model shall be used with extreme care.

Fig. 11: $x - y$ (top), $x - z$ (middle), $y - z$ (bottom) projections of a Earth–Moon Halo orbit of amplitude equal to $A_z = 385$ [Km], propagated in different gravitational models
Halo orbits, associated to different amplitude parameters, of the northern family (first coefficient of the Fourier expansion of the z-coordinate in normalised units, in which the unit length is taken as the distance between the libration point and the small primary). Different values of $A_z$ univocally correspond to different energy levels, and thus to different values of the Jacobi constant, $C_J$. In other words, given any of these two parameters, one particular Halo orbit is specified. Results are presented in Fig. 12 which displays the refinement of three halo orbits about the Earth–Moon first libration point, associated to three different values of the amplitude, $A_z = \{0.01, 0.03, 0.06\}$. Note that larger amplitudes mean larger orbital path, and therefore smaller Jacobi energy, $C_J$. The refinement has been done for 150 primaries revolutions, but for the sake of presentation clarity, only the first year is graphed. It is clear that the numerical refinement procedure of a Halo orbit produces a quasi-periodic one, whereas the baseline shape and size do not change for most of the cases.

The advantage of using a hierarchic modelling appears in the long period continuation. The normal algorithm can indeed handle periods equal to a few Earth–Moon rotations with proper selection of the Fourier extrapolation. On the other hand, the hierarchic ladder algorithm can handle time spans up to several hundreds Earth–Moon revolutions without the need of extrapolating based on frequency content. This is because the transition among models is smoother, the optimisation is well-behaved, and convergence properties more ideal.

A Fourier analysis of the refined halo-type orbit shows that the main single Halo frequency is maintained. In addition, since the refined orbit is now quasi-periodic, other frequencies appear. In particular, the frequency corresponding to main gravitational perturbers appear. Note that the perturbative effects of other massive planets are present, but cannot be resolved by the Fourier transform due to the limited total period.

V. CONCLUSION

In this paper an algorithm has been developed, that continues and refines trajectories calculated in the CRTBP towards the more complex ephemeris-based roto-pulsating n-body problem. The intrinsic hierarchic behaviour of the gravitational models at hand has been deeply employed to attain convergence and to improve the efficiency of the algorithm. The difference in the vectorial field had been made soother thanks to the progressive increase of included perturbation, explicitly using a hierarchic ladder. In particular several Earth–Moon $L_1$ Halo orbits have been refined with the proposed method. These test cases demonstrate the validity and the efficiency of the algorithm when compared to the same method without the hierarchy.

Acknowledgements

The first author would like to thank the Italian Space Agency, ASI, and the Space Generation Advisory Council, SGAC, for the grant to attend IAC 2015 and Space Generation Congress 2015.

Bibliography


