

Accurate orbit propagation in the presence of planetary close encounters

Davide Amato,¹★ Giulio Baù² and Claudio Bombardelli¹

¹Space Dynamics Group, School of Aerospace Engineering, Technical University of Madrid, Plaza del Cardenal Cisneros 3, E-28040 Madrid, Spain

²Mathematics Department, University of Pisa, Largo Bruno Pontecorvo 5, I-56127 Pisa, Italy

Accepted 2017 May 18. Received 2017 May 10; in original form 2017 March 6

ABSTRACT

We present an efficient strategy for the numerical propagation of small Solar system objects undergoing close encounters with massive bodies. The trajectory is split into several phases, each of them being the solution of a perturbed two-body problem. Formulations regularized with respect to different primaries are employed in two subsequent phases. In particular, we consider the Kustaanheimo–Stiefel regularization and a novel set of non-singular orbital elements pertaining to the *Dromo* family. In order to test the proposed strategy, we perform ensemble propagations in the Earth–Sun Circular Restricted 3-Body Problem (CR3BP) using a variable step size and order multistep integrator and an improved version of Everhart’s RADAU solver of 15th order. By combining the trajectory splitting with regularized equations of motion in short-term propagations (1 year), we gain up to six orders of magnitude in accuracy with respect to the classical Cowell’s method for the same computational cost. Moreover, in the propagation of asteroid (99942) Apophis through its 2029 Earth encounter, the position error stays within 100 metres after 100 years. In general, as to improve the performance of regularized formulations, the trajectory must be split between 1.2 and 3 Hill radii from the Earth. We also devise a robust iterative algorithm to stop the integration of regularized equations of motion at a prescribed physical time. The results rigorously hold in the CR3BP, and similar considerations may apply when considering more complex models. The methods and algorithms are implemented in the NAPLES FORTRAN 2003 code, which is available online as a GitHub repository.

Key words: methods: numerical – celestial mechanics – minor planets, asteroids: general.

1 INTRODUCTION

The reliable and accurate prediction of planetary close encounters is paramount in Space Situational Awareness (SSA), dynamical astronomy and planetary science.

Interest on this topic was first spurred in the 19th century by the search for the origin of short-periodic comets. According to the capture hypothesis, originally attributed to Laplace and further developed by Le Verrier (Everhart 1969), they hail from long-periodic comets experiencing close encounters with Jupiter that decrease their orbital energy. It was one of these encounters – ended up with the disintegration of Comet Shoemaker–Levy 9 in Jupiter’s atmosphere in 1994 July – that drew attention on the possibility that such a fierce event could happen on the Earth. The episode built momentum for modern large-scale asteroid surveys, which spectacularly increased the number of catalogued main belt and near-Earth asteroids (NEAs) and posed the basis for impact monitoring activities.

Close encounters have been found to induce exponential growth of the distance between initially close trajectories, leading to the decrease of Lyapunov times (Tancredi 1998). In fact, initially close trajectories diverge linearly after one encounter due to different

post-encounter orbital periods. However, the divergence accumulates multiplicatively after each encounter, eventually resulting in an exponential behaviour (Valsecchi et al. 2003). The latter is a necessary condition for chaotic trajectories to arise (Devaney et al. 2003, p. 50).

Discovering NEAs with highly non-linear dynamics requires to routinely propagate thousands of initial conditions, following orbital updates coming from observations. The required computational load might become critical when new-generation NEA surveys will settle into place and greatly increase the frequency of observations. For NEAs in which the orbit determination process is non-linear, some of the currently operating impact monitoring systems propagate the orbital probability density function by using a Monte Carlo approach in a unidimensional subspace, the Line of Variations (LOV). The LOV is sampled, generating several thousands of virtual asteroids (VAs) that need to be numerically propagated with a sophisticated dynamical model for subsequent analysis (Milani et al. 2005). Using efficient numerical techniques might bring about significant savings in computational cost. Increasing the accuracy of numerical propagation can also improve the estimation of impact probabilities.

When integrating numerically the equations of motion, planetary close encounters introduce a sudden spike in the magnitude on the

* E-mail: d.amato@upm.es

right-hand side (i.e. the vector field describing the evolution of the state vector), which may disrupt the integration. Considerable effort has been dedicated to improving the accuracy of the solutions in such cases.

Everhart (1974, 1985), while working on cometary debris, developed a very efficient single-step integrator, `RADAU`, which has been extensively applied in the integration of planetary close encounters and multiple star systems. Its strength relies in the choice of Gauss–Radau spacings for the substeps, which allow to build a 15th-order numerical scheme with only seven substeps. This integrator has been recently improved by Rein & Spiegel (2015) in the `IAS15` code, which achieves outstanding performances in the integration of cometary orbits affected by close encounters with Jupiter.

Dynamical astronomers have classically used symplectic numerical schemes in long-term integrations covering time spans comparable to the age of the Solar System. Due to their Hamiltonian-preserving nature they keep the energy error bounded in many situations; however, the symplectic character is lost if the time step of the integration is changed. Nevertheless, several symplectic algorithms capable of efficiently handling close encounters have been developed. In the method of Chambers (1999), implemented in the `MERCURY` package, a switch from a completely symplectic to a hybrid regime is applied when two bodies get close enough. A conventional integrator is used for the close encounter terms of the Hamiltonian together with those corresponding to the two-body problems with the Sun, while the other terms are solved symplectically. The method by Wisdom (2017), which only treats massless *test particles*, improves on the latter by employing Jacobi coordinates rather than democratic heliocentric ones, and by using a smoother switching function. Duncan, Levison & Lee (1998) devised a variant of mixed-variable symplectic methods that applies different time steps to the individual force components during close encounters. Using universal Kepler solvers (Wisdom & Hernandez 2015), Hernandez & Bertschinger (2015) proposed a second-order symplectic and time-reversible integrator, whose ability to reproduce close encounters efficiently was demonstrated by Hernandez (2016).

Even with non-symplectic methods, collisional problems can still be solved efficiently in the Hamiltonian formalism, especially if other properties such as time reversibility are exploited. The method by Kvaerno & Leimkuhler (2000) is based on the time-reversible integration of Hamiltonian equations. It makes use of a continuous switching function to handle close encounters; the latter are numerically integrated by a high-order scheme using Gauss–Legendre spacings. While the integration is performed with a fixed step size, the method makes use of a first-order Sundman transformation to achieve automatic step size regulation. Gonçalves Ferrari, Boekholt & Portegies Zwart (2014) also presented a method for the integration of the N -body problem through splitting of the Hamiltonian into separate two-body problems. The method is not time reversible (Hernandez & Bertschinger 2015), and second-order convergence is achieved by a composition of self-adjoint first-order maps.

Hamiltonian methods are ideal for integrating N -body systems in which most of the bodies are mutually interacting. Symplectic methods enable reliable integrations on time spans comparable to the age of the Solar system and are particularly suited for qualitative studies. The mathematical formulation allows for continuous splitting strategies, in which the terms of the Hamiltonian relative to particles undergoing encounters are interchanged continuously between the main and perturbing parts. However, including dissipative forces requires special care (Mikkola 1997; Tsang et al. 2015), and poses challenges to the accuracy of these methods. Moreover,

using variable step sizes is not straightforward, and high orders of integration are cumbersome to achieve.

Here, we are concerned with formulations of the *perturbed two-body problem*, which is a special case of the N -body. We assume the existence of a primary body that exerts an acceleration significantly larger than those of the remaining $N - 1$, which can be considered as perturbing bodies. This approach accurately models the motion of an asteroid with negligible mass, and is the one commonly used in near-Earth asteroid monitoring activities. The positions and velocities of the $N - 1$ remaining bodies are assumed to be retrieved from tabulated high-precision ephemerides such as the Jet Propulsion Laboratory (JPL) Development Ephemerides (Folkner et al. 2014). In contrast to Hamiltonian methods, the implementation of both conservative and dissipative perturbations is straightforward, and high-order adaptive solvers can be used without any particular precaution.

In this work, regularized formulations of the perturbed two-body problem are used to propagate close encounters with the Earth. Regularization is an analytical procedure that removes the $1/r^2$ singularity from equation (1). As it will be briefly explained in Section 2, it is usually achieved by a transformation of time and of the spatial coordinates. These formulations of dynamics have shown excellent numerical performances due to several characteristics: the solution is well behaved also close to the primary body; the independent variable is an angle-like quantity (also called *fictional time*) that produces an analytical step size regulation in physical time; the instability of Keplerian motion¹ is mitigated or suppressed and the equations of motion are linear in the Keplerian case (Stiefel & Scheifele 1971; Deprit, Elpe & Ferrer 1994). Linearization of equation (1) is also possible without eliminating the singularity at collision. In Burdet’s *focal* method (Burdet 1969), for example, one of the state variables is the inverse of the orbital radius. Here, we also consider these formulations as ‘regularized’ since the vector field does not explicitly present denominators with powers of r . Due to these characteristics, regularized formulations considerably reduce the global truncation error. The accumulation of round-off error can also be mitigated by adopting a suitable numerical scheme (Hairer, McLachlan & Razakrivony 2008; Rein & Spiegel 2015).

Regularization has proved highly beneficial in integrating the perturbed two-body problem, and regularizations for multiple bodies have been devised (Szebehely 1967, chap. 3). We employ two regularized methods: the Kustaanheimo–Stiefel (K–S) regularization (Stiefel & Scheifele 1971), and a set of non-singular orbital elements (Baù et al. 2015), which stems from the *Dromo* formulation originally developed by Peláez, Hedo & Rodríguez de Andrés (2007) for tether bead models. The *Dromo* elements are strictly related to Burdet’s focal variables. The numerical power and mathematical elegance of this method stimulated further work in recent years, which resulted in the inclusion of a perturbing potential (Baù, Bombardelli & Peláez 2013) and of time elements (Baù & Bombardelli 2014). Furthermore, Baù et al. (2015) particularized the original *Dromo* variables to the case of negative total energy, while Roa & Peláez (2015) and Baù et al. (2016) did the same for the positive case in Minkowskian and Euclidean geometry, respectively.

¹ Consider at a certain epoch a particle following a Keplerian ellipse and another particle having a slightly different position and velocity. The two orbital energies will differ by some amount, and since the orbital period depends on the energy, for arbitrary small variations of the position and velocity, after some time, the distance between the two particles will become larger than a given bound (Stiefel & Scheifele 1971, p. 75).

Formulations of dynamics based on orbital elements are particularly suited to the integration of weakly perturbed problems. However, close encounters generate quasi-impulsive perturbations on a heliocentric scale, which irremediably deteriorate the performance of this class of methods. Therefore, we propose to split the trajectory into heliocentric and geocentric phases, which unfold on different time and distance scales. This approach underlies analytical methods that decompose the full three-body problem into two-body problems with distinct primaries, such as patched conics (Kemble 2006) and Öpik’s (Öpik 1976; Valsecchi et al. 2003). In a similar manner, we try to exploit the extremely favourable numerical characteristics of element methods for weakly perturbed problems by switching reference systems and primary bodies during the integration of a planetary close encounter.

The question of *where* (or equivalently, when) to switch during the integration arises quite naturally, and elicits a non-trivial answer. Intuitively, we would switch systems at the surface separating regions of gravitational influence in a restricted three-body problem. Thus, the question ultimately shifts to the definition of such regions. Laplace’s sphere of influence and the Hill sphere (Battin 1999, p. 395) are suitable candidates; recently also definitions in terms of the variation of the two-body energy have been attempted (Araujo et al. 2008). However, none of these explicitly guarantees minimization of the numerical error, in particular when regularized formulations are involved. While this objective requires a rigorous analysis taking into account the particular numerical scheme and formulation of dynamics, we have embarked on a preliminary study of the matter.

The performance of the regularized formulations is compared to that of the equations of motion in Cartesian coordinates, both with and without splitting the trajectory. This is accomplished by conducting large-scale simulations in the planar circular, restricted Earth–Sun three-Body problem (CR3BP). In these simulations, we propagate an ensemble of trajectories that are affected by a single close encounter with the Earth. The initial conditions are derived from three parameters that fix the encounter geometry. For propagations in which the trajectory is split, the test particles enter and exit the geocentric phase at an assigned distance from the Earth. The accuracy is measured by computing position and energy errors with respect to a reference propagation obtained in quadruple precision, while the number of right-hand side evaluations is considered as a computational cost metric. Moreover, we also investigate the impact of the distance at which the primary bodies are switched on the accuracy and computational cost of integrations with regularized formulations. We extend our results to the long-term propagation (on a centenary scale) of the asteroid (99942) Apophis. We propagate the trajectory for 100 yr with the presented methods and analyse the growth of the numerical error in more detail. For the Apophis test case, we also compare our methods to the Bulirsch–Stoer (BS) integrator and the hybrid method by Chambers (1999), which are implemented in the MERCURY package. Thus we assess the feasibility of using the latter methods, which are widely employed in the planetary systems studies, to asteroid impact monitoring.

In all of the simulations, the regularized formulations and the Cowell equations are integrated by employing adaptive solvers. We use a multistep solver with variable step size and order, and a single-step solver with variable step size based on the Everhart–Radau numerical scheme (Rein & Spiegel 2015). They are implemented in the LSODAR and XRA15 packages, respectively.

The paper is organized as follows. We illustrate the prominent features of the regularized formulations used in this work in Section 2. The propagation strategy based on splitting the close

encounter into two weakly perturbed two-body problems is detailed in Section 3. Relevant aspects of the solvers used in the work are shown in Section 4. We propose a simple parametrization of planar close encounters in the CR3BP in Section 5, which we use in the large-scale simulations performed in Section 6. Along with Section 7, in which we analyse the long-term propagation of Apophis, Section 6 contains the key results of the work. These include the comparison of regularized formulations to the equations of motion in Cartesian coordinates, and the study on the choice of switch distances in the context of trajectory splitting. Section 8 gives details on the software implementation, while conclusive remarks are in Section 9.

2 REGULARIZED FORMULATIONS OF DYNAMICS

In special perturbations, the most straightforward way of solving the perturbed two-body problem is by numerically integrating the equations of motion in Cartesian coordinates,

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3}\mathbf{r} + \mathbf{F}, \quad (1)$$

where \mathbf{r} is the position with respect to the primary body, μ is its gravitational parameter and \mathbf{F} is the perturbing acceleration. This approach, known as *Cowell’s method* or *Cowell’s formulation* (Battin 1999, p. 447), is simple and robust, especially in situations where the magnitude of the perturbation \mathbf{F} is comparable to that of the main gravitational acceleration. However, the direct integration of equation (1) can be disadvantageous from the computational point of view. In fact, its solutions are unstable even in the unperturbed case, i.e. for Keplerian motion. This implies that the propagation of numerical error is quite fast. It also exhibits a singularity for $r = 0$ that poses limits on the step size when close to the primary body, and makes the integration of collisional orbits impossible. For these reasons, sophisticated and adaptive numerical solvers are required to reach satisfactory levels of accuracy, in particular when close encounters are involved.

All of these issues can be ameliorated by *regularizing* the equations of motion, i.e. by analytically removing the singularity in equation (1). Regularization is achieved in two steps. The first one is to change the independent variable from the physical time t to a fictitious time s by means of the generalized Sundman transformation

$$\frac{dt}{ds} = f(\mathbf{y}, s)r^\alpha, \quad f > 0, \quad (2)$$

where f , in general, is a function of the state vector \mathbf{y} and is constant if the motion is unperturbed, and α is a positive constant that has to be smaller than $3/2$. Depending on the value of α , a uniform step-size distribution in s results in step sizes in t that are smaller when the particle is closer to one or to both of the apses; this is a favourable situation for numerical integration. The second step is to represent the two-body problem with linear differential equations that do not contain the singularity $r = 0$. This result is usually achieved by embedding some Keplerian integrals in the equations of motion (Bond & Allman 1996, chap. 9) and introducing new state variables, rather than position and velocity.

Linearity is a desirable feature for improving the numerical performance and can be obtained also without regularization. In Burdet’s *focal* method (Burdet 1969) the Keplerian motion is decomposed into the radial displacement with respect to the primary and the free rotation of the radial direction in space. Then, if the true anomaly is taken as an independent variable ($\alpha = 2$ in equation

1), the inverse of the orbital radius and the radial unit vector satisfy linear differential equations with constant coefficients.

By applying the variation of parameters (VOP) technique to the solution of the linearized equations, we can define new orbital elements, and use them as state variables. These quantities, being integrals of the motion, exhibit a smooth evolution for weakly perturbed problems. Due to this characteristic, element formulations are highly efficient when the magnitude of the perturbing acceleration \mathbf{F} is small. Moreover, the same beneficial properties of the parent variables are inherited by the derived elements: they may be well defined for circular and equatorial orbits, and at collision.

All these schemes transforming the two-body problem into a set of linear differential equations are here referred to as regularized methods. Note that they suffer from some little drawbacks: an iterative procedure is necessary to know the state vector \mathbf{y} at specific epochs (see Appendix A); position and velocity must be obtained from \mathbf{y} through algebraic formulas and the vector \mathbf{y} is composed of redundant variables.

Regularization has a long history in celestial mechanics, and several sets of regularized equations exist in the literature. We consider two of them in this work: the Kustaanheimo–Stiefel (K-S) regularization and *EDromo*: a set of orbital elements developed by Baù et al. (2015), which is related to Burdet’s method.

2.1 Kustaanheimo–Stiefel regularization

The K–S regularization has seen extensive use in dynamical astronomy, and it is described in detail in the text by Stiefel & Scheifele (1971). It is based on the classical Sundman transformation, i.e. $f = \alpha = 1$ in equation (1), and on a mapping from $\mathbf{u} \in \mathbb{R}^4$ to $\mathbf{r} \in \mathbb{R}^3$ given by

$$\mathbf{x} = \mathbf{L}(\mathbf{u})\mathbf{u}, \quad \mathbf{x} = (\mathbf{r}; 0), \quad (3)$$

where the matrix $\mathbf{L}(\mathbf{u})$ is made by the four parameters $(u_1; u_2; u_3; u_4) = \mathbf{u}$ as follows:

$$\mathbf{L}(\mathbf{u}) = \begin{pmatrix} u_1 & -u_2 & -u_3 & u_4 \\ u_2 & u_1 & -u_4 & -u_3 \\ u_3 & u_4 & u_1 & u_2 \\ u_4 & -u_3 & u_2 & -u_1 \end{pmatrix}. \quad (4)$$

Let the perturbation \mathbf{F} stem from a potential V and a non-conservative acceleration \mathbf{P} :

$$\mathbf{F} = -\frac{\partial V}{\partial \mathbf{r}} + \mathbf{P}. \quad (5)$$

The total energy $\tilde{\varepsilon}$ is the sum of the Keplerian energy ε and the perturbing potential V ,

$$\tilde{\varepsilon} = \varepsilon + V. \quad (6)$$

Denoting with a prime differentiation with respect to s , the differential equation for the K–S parameters is written as

$$\mathbf{u}'' = \frac{\tilde{\varepsilon}}{2}\mathbf{u} - \frac{1}{4}\frac{\partial}{\partial \mathbf{u}}(\|\mathbf{u}\|^2 V) + \frac{\|\mathbf{u}\|^2}{2}(\mathbf{L}^T(\mathbf{u}) \cdot \mathbf{P}). \quad (7)$$

Note that equation (7) becomes linear for Keplerian motion, furthermore if $\tilde{\varepsilon} < 0$ it represents four scalar harmonic oscillators of the same frequency. Time must be obtained by the Sundman transformation *sono stati aggiunti, ma non ci vanno*:

$$t' = \|\mathbf{u}\|^2, \quad (8)$$

or by introducing a time element (Stiefel & Scheifele 1971). A key operation for improving the efficiency of K–S regularization is to

regard the total energy $\tilde{\varepsilon}$ as a state variable instead of computing it from \mathbf{u} , \mathbf{u}' and t . The reason becomes clear if we consider the case in which $\mathbf{P} = \mathbf{0}$ and V does not depend on time. Then, from the differential equation

$$\tilde{\varepsilon}' = \|\mathbf{u}'\|^2 \frac{\partial V}{\partial t} + 2[\mathbf{u}' \cdot (\mathbf{L}^T(\mathbf{u}) \cdot \mathbf{P})], \quad (9)$$

we have $\tilde{\varepsilon}(s) = \tilde{\varepsilon}(0)$, while on the other hand the function $\tilde{\varepsilon}(\mathbf{u}, \mathbf{u}', t)$ will deviate from $\tilde{\varepsilon}(0)$ due to errors affecting time and the K–S parameters.

The state vector is $\mathbf{y} = (\mathbf{u}; \mathbf{u}'; t; \tilde{\varepsilon})$, and its dimension is 10. The set of differential equations to be integrated is given by equations (7) (rewritten as 8 first-order equations), (8) and (9). The transformations between the Cartesian and K–S state vectors are possible through explicit algebraic relations that are given in Stiefel & Scheifele (1971, p. 33).

2.2 EDromo

The method Dromo (Peláez et al. 2007; Baù et al. 2013) consists of seven orbital elements: three of them allow us to recover the motion along the radial direction and the remaining four fix the orientation of the orbital plane and a departure point on it. The angular displacement between the radial unit vector and this point is provided by the independent variable, which is the true anomaly when the motion is unperturbed. Dromo elements can be derived from the focal method developed by Burdet (1969) through the VOP technique (more details can be found in the introduction of Baù et al. 2015).

EDromo was born from the idea of applying the same decomposition of the dynamics as in Dromo but with the time transformation²

$$t' = \frac{r}{\sqrt{-2\tilde{\varepsilon}}}, \quad \tilde{\varepsilon} < 0, \quad (10)$$

where $\tilde{\varepsilon}$ is the total energy (equation 6). This choice was made to improve the performance of Dromo for highly eccentric motion under third-body perturbations. As in Dromo, the main role is played by an intermediate reference frame $\{x, y, z\}$ with the axis z oriented as the angular momentum vector. The evolution of this frame is given by the four components $(\lambda_4, \lambda_5, \lambda_6, \lambda_7)$ of a unit quaternion. From the differential equation of the orbital radius, which becomes linear in the two-body problem, two orbital elements (λ_1, λ_2) can be defined by applying the VOP method. These turn out to be the projections of the eccentricity vector (or its generalized version, see Baù et al. 2015) on the axes x, y of the intermediate frame. The quantities λ_1, λ_2 and $\lambda_3 = -1/(2\tilde{\varepsilon})$ allow us to compute the radial solution and the angle ν between the position vector \mathbf{r} and the departure axis x , as follows:

$$r = \lambda_3 \rho, \quad r' = \lambda_3 \zeta, \quad (11)$$

$$\nu = s + 2 \arctan\left(\frac{\zeta}{m + \rho}\right), \quad (12)$$

where

$$\rho = 1 - \lambda_1 \cos s - \lambda_2 \sin s, \quad \zeta = \lambda_1 \sin s - \lambda_2 \cos s, \quad (13)$$

$$m = \sqrt{1 - \lambda_1^2 - \lambda_2^2} \quad (14)$$

and s is the fictitious time.

² The following equations for EDromo involve non-dimensional quantities. Reference mass, length and time are chosen so that the gravitational parameter of the primary is equal to 1.

The perturbations are split into conservative and non-conservative terms according to equation (5), and projected on the radial, local-horizontal and out-of-plane axes as $\mathbf{F} = (R, T, N)$ and $\mathbf{P} = (R_p, T_p, N_p)$. The differential equations for the spatial elements are written as

$$\lambda'_1 = (Rr - 2V)r \sin s + \Lambda_3 [(1 + \rho) \cos s - \lambda_1], \quad (15)$$

$$\lambda'_2 = (2V - Rr)r \cos s + \Lambda_3 [(1 + \rho) \sin s - \lambda_2], \quad (16)$$

$$\lambda'_3 = 2\lambda_3^3 \left(R_p \zeta + T_p n + \frac{\partial V}{\partial t} \sqrt{\lambda_3 \rho} \right), \quad (17)$$

$$\lambda'_4 = N \frac{r^2}{2n} (\lambda_7 \cos v - \lambda_6 \sin v) + \frac{\omega_z}{2} \lambda_5, \quad (18)$$

$$\lambda'_5 = N \frac{r^2}{2n} (\lambda_6 \cos v + \lambda_7 \sin v) - \frac{\omega_z}{2} \lambda_4, \quad (19)$$

$$\lambda'_6 = N \frac{r^2}{2n} (-\lambda_5 \cos v + \lambda_4 \sin v) + \frac{\omega_z}{2} \lambda_7, \quad (20)$$

$$\lambda'_7 = N \frac{r^2}{2n} (-\lambda_4 \cos v - \lambda_5 \sin v) - \frac{\omega_z}{2} \lambda_6, \quad (21)$$

where

$$\Lambda_3 = \frac{1}{2\lambda_3} \lambda'_3, \quad n = \sqrt{m^2 - 2\lambda_3 \rho^2 V} \quad (22)$$

and

$$\omega_z = \frac{n - m}{\rho} + \frac{1}{m(1 + m)} [(2V - Rr) \times (2 - \rho + m)r + \Lambda_3 \zeta (\rho - m)]. \quad (23)$$

Physical time is computed by either a constant or a linear time element. We have, respectively

$$\lambda'_{0,c} = \lambda_3^{3/2} [(Rr - 2V)r + \Lambda_3 (2\zeta - 3s)], \quad (24)$$

$$\lambda'_{0,1} = \lambda'_{0,c} + \lambda_3^{3/2} (1 + 3\Lambda_3 s). \quad (25)$$

In unperturbed motion the elements $\lambda_{0,c}, \lambda_1, \dots, \lambda_7$ are constants, while $\lambda_{0,1}$ is a linear function of the independent variable. As for the K–S formulation, the conversion between the EDromo elements and the Cartesian state vector is managed by explicit algebraic relations (Baù et al. 2015).

Note that the EDromo formulation is only defined for $\tilde{\varepsilon} < 0$. Analogous formulations for $\tilde{\varepsilon} > 0$ have been developed by Roa & Peláez (2015) and Baù et al. (2016). Also, equations (18)–(21) are singular for $n = 0$. This condition, when $V = 0$, is satisfied if the angular momentum vanishes.

3 PROPAGATION STRATEGY

In planetary close encounters, the fundamental character of the perturbed two-body problem changes on time-scales that can be orders of magnitude smaller than the period of the heliocentric orbit. Almost suddenly, the gravitational acceleration from the perturbing body may become larger than that of the primary. Thus, heliocentric orbits experience strong, quasi-impulsive perturbations.

Uncertainties that are small on a Solar system scale have a significant impact on the close encounter, since the latter takes place on much smaller (planetary) time and distance scales. Moreover, trajectories that are close before the encounter diverge afterwards; this divergence can be exponential after subsequent encounters, thereby spawning chaotic dynamics. Therefore, some kind of adaptation of

the integration process to the close encounter is mandatory in order to mitigate the propagation error. This can be provided by numerical solvers with variable step size or order. The sudden growth of the magnitude of \mathbf{F} in equation (1) is detected by an increase in the local truncation error, which causes the integrator to choose a shorter step size and a higher order. Adaptive solvers, which are commonly used in Cowell's method, are applied here to regularized formulations. We show that they can successfully accommodate the change in the dynamics brought about by the encounter if the propagation is split into three phases. Far from the perturbing body, the orbit is computed in a heliocentric phase, where the primary body and the origin of the reference system is the Sun. When a certain *switch distance* R_{sw} from a major Solar System body (i.e. a planet) is reached, the integration is stopped, and the position and velocity are computed with respect to the planet. This becomes the new primary body, while the Sun is regarded as a perturber. The new perturbed two-body problem is numerically solved until the distance R_{sw} is reached again. Then, another switch of reference systems and primary bodies is applied, so that the Sun and the planet come back to play the roles of main attractor and perturber, respectively.

The procedure described above is here called *trajectory splitting*, and it can be automatized and repeated indefinitely for multiple close encounters. Conceptually, it is similar to the analytical technique of patched conics for designing interplanetary trajectories, with the difference that any perturbation can be added in the physical model. The important effect of the proposed approach is to transform a strongly perturbed two-body problem into a series of less perturbed two-body problems. When employing the Cowell formulation, this strategy does not bring about evident advantages. In the right-hand side of equation (1), the primary and the perturber are treated equally. Splitting the trajectory only switches their roles, but the equation formally remains the same. On the other hand, the performance of regularized formulations may be substantially improved. In the K–S formulation, a weaker perturbation produces a smaller rate of variation of the total energy, which results in more stable differential equations for the K–S parameters. Also methods based on orbital elements, like EDromo, become faster and more accurate.

There are some caveats to trajectory splitting. If the independent variable is not the physical time, stopping the propagation requires an iterative algorithm (see Appendix A). Moreover, particular care must be taken for the selection of the optimal value of the switch distance.

4 NUMERICAL SOLVERS

As to extensively study the performances of regularized sets of equations, we choose two solvers with different numerical schemes. Both of them are adaptive, in the sense that either just the step size or both the step size and the order can be changed along the integration. We adjust the accuracy by setting the value of the dimensionless parameter ϵ , which is a threshold on the magnitude of the local truncation error.

The solvers are equipped with event location algorithms, which allow us to stop the integration when a function of the independent variable reaches a desired value. They are necessary to detect close encounters and to assess the value of the physical time, if this is not the independent variable (see Section 4.3).

The length of the first step of a propagation can be passed to the solver. We discuss this aspect and propose a way to get a reasonable initial step size in Appendix B.

The numerical schemes used by the solvers in this work are not time reversible. Although time reversibility is a desirable property for an integrator, its lack does not affect the relative strengths of the methods presented in any way.

4.1 LSODAR multistep solver

The Livermore Solver for Ordinary Differential Equations with Automatic Root-finding (LSODAR), implemented in the ODEPACK FORTRAN 77 package,³ is a multistep solver with variable step size and order. LSODAR includes automated root-finding and scheme switching capabilities (Radhakrishnan & Hindmarsh 1993). It uses the Adams–Bashforth–Moulton (ABM) method, and if the equations become stiff (Petzold 1983) it switches to a backward differentiation formula (BDF) implicit method. The step size is regulated as to comply with absolute and relative tolerances on the local truncation error. In our simulations, these quantities are set to be both equal to the accuracy parameter ϵ . As to achieve larger step sizes, the solver is endowed with an order-changing algorithm that selects orders of up to 12 and 5 for the ABM and the BDF methods, respectively.

4.2 XRA15 single-step solver

We first implemented the variable step size 15th-order RADAU solver (Everhart 1985) in the XRA15 FORTRAN 2003 subroutine. The core of this code is the implicit, single-step numerical scheme with seven stages developed by Everhart (1974). Then, in order to improve the accuracy and robustness of the original solver, we modified the step size control and predictor-corrector algorithms as proposed by Rein & Spiegel (2015).

Everhart’s RADAU can integrate either first- or second-order equations. However, Everhart (1974) recognizes that ‘results are up to 10^6 times more accurate at the same number of function evaluations’ when the position is obtained from the acceleration instead of the velocity. Therefore, in Cowell’s method we directly apply XRA15 to the second-order equation (1). On the other hand, for the K–S regularization we write a system of 10 differential equations of the first-order.⁴

While LSODAR has been used especially in research areas different from celestial mechanics, like analytical chemistry and computational biology, RADAU was conceived as a powerful tool for orbit propagation in the presence of close encounters.

4.3 Event location

An *event* is said to occur at the value $s = s^*$ of the independent variable, if s^* is a root of the *event function* $g(s, \mathbf{y}(s))$, where $\mathbf{y}(s)$ is the state vector. The event location problem consists in finding the roots of g during the propagation.

In our numerical simulations we have to solve this problem for two cases. Time is a function of s in regularized methods, thus stopping the propagation at a prescribed time t_f requires to find the root of the equation

$$t(s) - t_f = 0. \quad (26)$$

³ <http://www.netlib.org/odepack/>, last visited 2016 November 28.

⁴ We did not investigate the possibility of directly integrating equation (7).

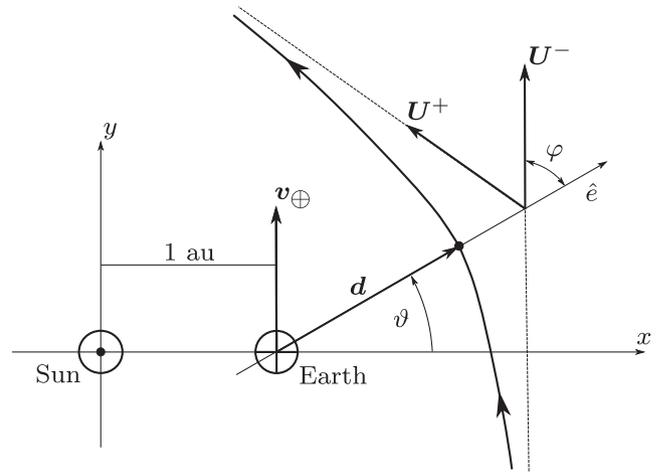


Figure 1. Parametrization of the close encounter through (d, e, ϑ) (Section 5). The geocentric position vector at the closest approach distance d is in the same direction of the eccentricity vector \hat{e} of the geocentric hyperbola.

Also, if the trajectory splitting is activated (see Section 3), we want to know for which value of s the distance r from the encountered body is equal to the switch radius R_{sw} , that is

$$r(s) - R_{sw} = 0. \quad (27)$$

LSODAR has a very efficient algorithm to solve the event location problem. When a change of sign in the event function g is detected between two integration steps, LSODAR finds all the roots inside the step in sequence, starting from the leftmost (Hiebert & Shampine 1980). The value of the root is refined iteratively using a modified *regula falsi* algorithm (Dahlquist & Björck 1974, chap. 6). This method proved extremely reliable, and we did not find any cases in which the detection of a close encounter failed. The strict tolerances used in our tests imply that the steps taken by the solver are small enough for the event location algorithm to easily locate and converge to the roots of equations (26) and (27). For the XRA15 solver, we developed a reliable event location algorithm for equation (26), which is described in detail in Appendix A.

5 PARAMETRIZATION OF PLANAR CLOSE ENCOUNTERS

Most of the tests performed in this work rely on massive numerical simulations in the planar, Sun–Earth CR3BP. We introduce a simple parametrization of planar close encounters to determine initial conditions for the integrations. Also, we use the patched conics approach to derive the heliocentric orbital elements from the encounter parameters. These elements allow us to analyse the impact of the dynamics on the amplification of the numerical error.

5.1 Close encounter parameters

Let us consider the epoch in which the propagated object reaches the closest point to the Earth’s centre of mass. A close encounter can be characterized by these three parameters: the minimum approach distance (d); the eccentricity of the geocentric hyperbola ($e > 1$) and the angle ϑ between the position vector \mathbf{d} and the Sun–Earth direction (see Fig. 1). Two more parameters describing the hyperbolic encounter are of particular interest: the asymptotic velocity U and the angle φ ($\in [0, \pi/2]$) between the incoming (or outgoing) asymptote and the apse line of the hyperbola. By the vis-viva and

conic equations, we obtain U and φ from d and e as

$$U = \sqrt{\frac{\mu_{\oplus}}{d}} (e - 1), \quad (28)$$

$$\cos \varphi = \frac{1}{e}, \quad (29)$$

where μ_{\oplus} is the Earth's gravitational parameter. Equation (28) shows that level curves of the Keplerian energy ε are straight lines on the plane (d, e) , with steeper slopes corresponding to higher values of ε . Since $\varepsilon = U^2/2$ we speak equivalently of *high (low)* energy and *fast (slow)* encounters. Note that smaller values of d correspond to higher velocities, while the directions of the asymptotes are not affected. Moreover, higher values of e correspond to faster motion along hyperbolic trajectories that tend to be rectilinear.

5.2 From encounter parameters to heliocentric elements

We can gain valuable knowledge about the effect of close encounters in the planar CR3BP by looking at the variations in the orbital elements of the heliocentric trajectory. These can be estimated with the patched conics approximation, which is based on two hypotheses: the motion is Keplerian, and the close encounter happens instantaneously.

Consider a heliocentric reference system $\{S, x, y\}$ such that the Earth lies on the x -axis at 1 au from the Sun, and let \mathbf{i}, \mathbf{j} be the unit vectors of the axes x, y , respectively. We are concerned with computing the energy and eccentricity before and after the encounter, because they strongly affect the numerical performance (see Section 6.3.1). In the following, we use the superscripts ‘ $-$ ’ for the pre-encounter quantities, and ‘ $+$ ’ for the post-encounter ones.

As shown in Fig. 1, the orientation of the asymptotic velocity vectors $\mathbf{U}^-, \mathbf{U}^+$ with respect to the Earth's velocity \mathbf{v}_{\oplus} depends on ϑ and φ . Adding \mathbf{U}^{\mp} to $\mathbf{v}_{\oplus} = v_{\oplus} \mathbf{j}$, we obtain the expressions for the heliocentric velocities,

$$\mathbf{v}^{\mp} = \mathbf{U}^{\mp} + \mathbf{v}_{\oplus} = \pm U \cos(\varphi \pm \vartheta) \mathbf{i} + [U \sin(\varphi \pm \vartheta) + v_{\oplus}] \mathbf{j}. \quad (30)$$

By employing equation (30) along with the vis-viva equation we can write the Keplerian energies before and after the encounter as

$$\varepsilon^{\mp} = \frac{U^2 - v_{\oplus}^2}{2} + U v_{\oplus} \sin(\varphi \pm \vartheta). \quad (31)$$

The corresponding eccentricities are readily obtained as

$$e^{\mp} = \sqrt{1 + 2 \frac{\varepsilon^{\mp}}{v_{\oplus}^2} \left(1 + \frac{U}{v_{\oplus}} \sin(\varphi \pm \vartheta)\right)^2}. \quad (32)$$

From equation (30) we have that the change in velocity $\Delta \mathbf{v}$ due to the encounter is

$$\Delta \mathbf{v} = \mathbf{v}^+ - \mathbf{v}^- = -2U \cos \varphi (\cos \vartheta \mathbf{i} + \sin \vartheta \mathbf{j}). \quad (33)$$

Its magnitude $\Delta v = 2U \cos \varphi$ depends only on the shape of the hyperbola. Fig. 2 displays level curves of Δv over the set of pairs (d, e) that we considered in our simulations. While ϑ does not influence the rotation of the velocity vector from \mathbf{U}^- to \mathbf{U}^+ , it has instead a significant impact on the heliocentric dynamics. Indeed, from equation (31), the energy variation results

$$\Delta \varepsilon = \varepsilon^+ - \varepsilon^- = -2U v_{\oplus} \sin \vartheta \cos \varphi = \mathbf{v}_{\oplus} \cdot \Delta \mathbf{v}, \quad (34)$$

and both open and closed orbits may be obtained by changing ϑ for fixed values of d and e .

The analytical variations of the orbital elements are in good agreement with numerical tests, and adequately describe the evolution of the particles in the limits of the patched-conics approach.

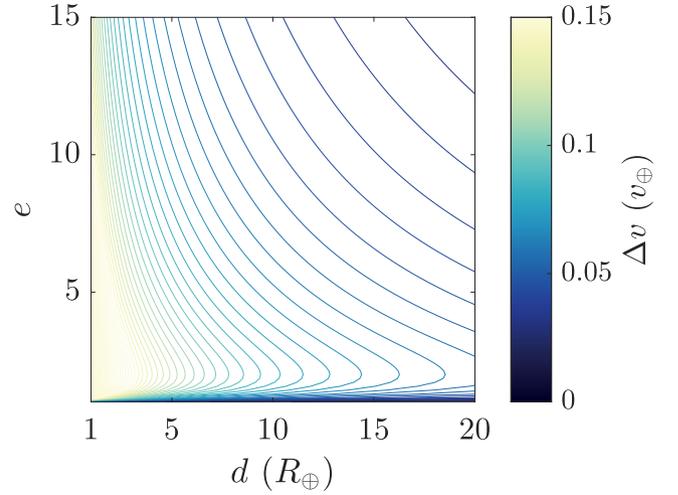


Figure 2. Level curves of $\Delta v(d, e)$ computed through equation (33). It is expressed in units of Earth orbital velocity v_{\oplus} ; the distance d is in Earth radii.

6 LARGE-SCALE NUMERICAL EXPERIMENTS

The performances of the formulations presented in Section 2 are analysed in the planar, Sun–Earth CR3BP by means of massive ensemble integrations, which are carried out with both `LSODAR` and `XRA15` solvers. The main focus of the numerical experiments is to investigate the efficiency of the trajectory splitting strategy (Section 3) when it is applied with regularized methods. We also study the effect of the switch distance R_{sw} : its choice must be taken with some care to avoid the build-up of truncation error and increases in the computational effort.

In the following the *propagation method* will denote the strategy, i.e. with or without trajectory splitting, the adopted formulation(s) and the solver. First, we expound on the implementation of the numerical experiments and on the quantification of the performance, then we present the results.

6.1 Test description

Given d, e, ϑ (see Fig. 1) we compute the geocentric position and velocity of the particle at the epoch t_{ca} of its closest approach with the Earth. Starting from them, we obtain a *reference* trajectory by integrating forwards and backwards until $t_i = t_{\text{ca}} - \Delta t/2$ and $t_f = t_{\text{ca}} + \Delta t/2$, respectively. The propagation is done in quadruple precision by applying Cowell's formulation and the `LSODAR` solver with a very strict accuracy parameter ϵ . The position and velocity relative to the Sun at time t_i are then set as initial conditions for the *test* simulations. These span the interval $[t_i, t_f]$, and are executed in double precision using different propagation methods. Test trajectories are compared to the reference to evaluate their accuracy, as explained in Section 6.1.1. If trajectory splitting is applied, several switch distances R_{sw} are considered in order to investigate the impact of this parameter on the numerical performance of a propagation method. Our tests are built up by taking thousands of different combinations of (d, e, ϑ) .

The time interval Δt has been set equal to 1 yr, so that we can thoroughly study *only one* close encounter. Much longer propagations would increase the probability of experiencing multiple encounters, and would require prohibitively high computational times. On the other hand, a shorter Δt would not allow an appreciable accumulation of the truncation error.

6.1.1 Performance metrics

We quantify the accuracy by means of the relative position and energy error metrics. The former is computed as

$$\delta r = \frac{\|\mathbf{r}_{f,\text{test}} - \mathbf{r}_{f,\text{ref}}\|}{\|\mathbf{r}_{f,\text{ref}}\|}, \quad (35)$$

where $\mathbf{r}_{f,\text{test}} = \mathbf{r}_{\text{test}}(t_f)$ and $\mathbf{r}_{f,\text{ref}} = \mathbf{r}_{\text{ref}}(t_f)$ are the heliocentric positions for the test and reference propagations, respectively, at the final time t_f . The latter is obtained as

$$\delta \varepsilon = \left| \frac{\varepsilon_{f,\text{test}} - \varepsilon_{f,\text{ref}}}{\varepsilon_{f,\text{ref}}} \right|, \quad (36)$$

where ε is the two-body energy of the asteroid with respect to the Sun, and the subscripts have the same meaning as before. The quantity $\delta \varepsilon$ can be understood as a secular drift of the asteroid position in the along-track direction, which in the long run may cause its loss. Thus, in short-term propagations the energy error is more important than the position error.

As to capture the overall behaviour of the error of a propagation method for encounter hyperbolas of different shapes we average the quantities δr , $\delta \varepsilon$, over the pairs (d, e) for a fixed ϑ , as follows:

$$\bar{x}(\vartheta) = \left(\prod_{i=1}^n \prod_{j=1}^m x(d_i, e_j, \vartheta) \right)^{\frac{1}{nm}}, \quad (37)$$

where x can be either δr or $\delta \varepsilon$. As it will be shown later, the errors can vary by up to three orders of magnitude depending on the values of (d, e) . By taking the geometric average the influence of large outliers is mitigated, giving a more effective information.

As a metric for the computational cost of each propagation, we use the number of evaluations FE of the right-hand side of the equations of motion. This metric has the advantage of being machine independent; furthermore, it is representative of the computational time especially when the evaluation of the perturbing forces is time-consuming. We also average FE according to equation (37).

6.2 Choice of the accuracy parameter

The growth of truncation error during the integration can be contained by using small values of the accuracy parameter ϵ (Section 4), which lead to shorter integration steps. However, decreasing ϵ also results in more floating-point operations, which generate round-off error. If ϵ is small enough, the global (i.e. accumulated) round-off error will surpass the global truncation error. This situation is to be avoided, since the accuracy is degraded and the numerical scheme might go into instability.

In our simulations, ϵ is chosen according to the following criterion. For each propagation method, we computed the lowest possible ϵ such that the numerical error is dominated by the truncation error. This value, which is here called *round-off limit*, is found by an iterative procedure. Each simulation is repeated using decreasing values of ϵ ,

$$\epsilon_1 > \epsilon_2 > \dots > \epsilon_R > \epsilon_{R+1}. \quad (38)$$

The global truncation error, which is measured through the average position error, also decreases until reaching a small enough ϵ ,

$$\overline{\delta r}(\vartheta; \epsilon_1) > \overline{\delta r}(\vartheta; \epsilon_2) > \dots > \overline{\delta r}(\vartheta; \epsilon_R), \quad (39)$$

$$\overline{\delta r}(\vartheta; \epsilon_R) < \overline{\delta r}(\vartheta; \epsilon_{R+1}), \quad (40)$$

for a given ϑ . Condition (40) signals the entry into a round-off-dominated regime. We define ϵ_R as the round-off limit, that is the

value of the accuracy parameter below which the numerical error is dominated by round-off. Different propagation methods are then compared at the round-off limit. We elect the *best method* as the one that reaches the highest accuracy and with the lowest computational cost.

The round-off limit mainly depends on the solver. For LSODAR, it is around or even below machine zero in double precision. For XRA15 it is more difficult to find this value. The very high order of the solver implies that the convergence of the integration is very sensitive to the parameter ϵ : values that are just a few orders of magnitude greater than the round-off limit might produce a very large global truncation error. Therefore, the XRA15 solver requires some calibration to find the range of ϵ that lead to stable propagations. Also, the round-off limit in XRA15 is noticeably dependent on the type of propagated trajectory, and on the perturbations. While this phenomenon does not strongly affect the numerical performances in our simulations, it should be taken into account to avoid numerical instabilities in practical applications. Algorithms that mitigate the round-off error, as those described in Rein & Spiegel (2015), would lower the round-off limit and thereby allow us to choose conservative values for ϵ without the need for calibration. We found a good estimate for the round-off limit to be 10^{-3} for element methods, and between 10^{-6} and 10^{-7} for coordinate-based formulations. These values are quite close to the estimates derived in Rein & Spiegel (2015) to reach machine precision in the integration of Newtonian equations. The specific value of ϵ that guarantees the best performance must be chosen on a case-by-case scenario, taking into account the considered perturbations and the propagation span.

The threshold on the local truncation error for most applications will be higher than the round-off limit. However, as long as the round-off error does not dominate the integration we do not expect the relative strengths of the presented methods to change significantly. Regularized formulations ultimately work by reducing the local truncation error, therefore a gain in performance at the round-off limit ϵ_R will be reflected in propagations at any other $\epsilon > \epsilon_R$.

6.3 Simulations without trajectory splitting

We analyse results from large-scale propagations carried out by keeping the Sun as the primary body for the whole integration span. More specifically, we consider one million different combinations of the encounter parameters d, e, ϑ (see Fig. 1) uniformly distributed over $[\mathbf{R}_\oplus, 20 \mathbf{R}_\oplus] \times [1.01, 15] \times [0^\circ, 360^\circ]$. Note that most of the values of d and e correspond to deep and slow encounters, which are very challenging to integrate due to larger perturbations.

Our purpose is to assess the performances of regularized formulations compared to Cowell's. This analysis also allows us to gain insight on the dependence of the numerical error on the parameters (d, e, ϑ) , which affect the orbital elements of the heliocentric trajectory and the magnitude of perturbations. Understanding such dependence sheds light on the characteristics of the formulations of dynamics involved in the study, and lets us identify particularly critical close encounters.

We will only show plots of the energy error in this section, since it is the most critical for short-term propagations (Section 6.1.1). The position error exhibits a similar behaviour.

6.3.1 Impact of the dynamics on the performance

The relative energy error $\delta \varepsilon$ is plotted in Figs 3 and 4 as a function of d and e for propagations carried out with LSODAR and XRA15, respectively. The upper panels refer to the Cowell's formulation and the lower ones to the K-S. The number of function evaluations FE

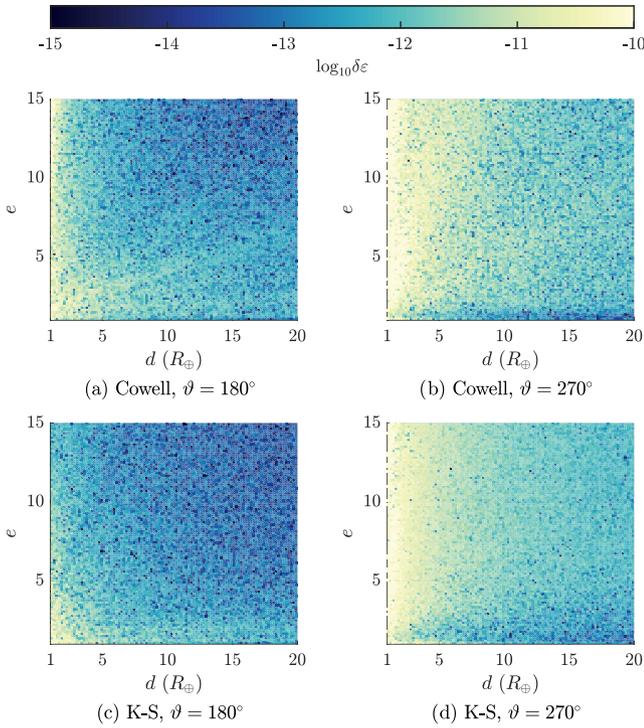


Figure 3. Relative energy error $\delta\epsilon$ (equation 36) as a function of the close encounter parameters d , e , for $\vartheta = 180^\circ, 270^\circ$ (see Fig. 1). The minimum approach distance d is measured in Earth radii. LSODAR was used with Cowell’s (upper panels) and K–S (lower panels) formulations. The accuracy parameter is $\epsilon = 10^{-16}$.

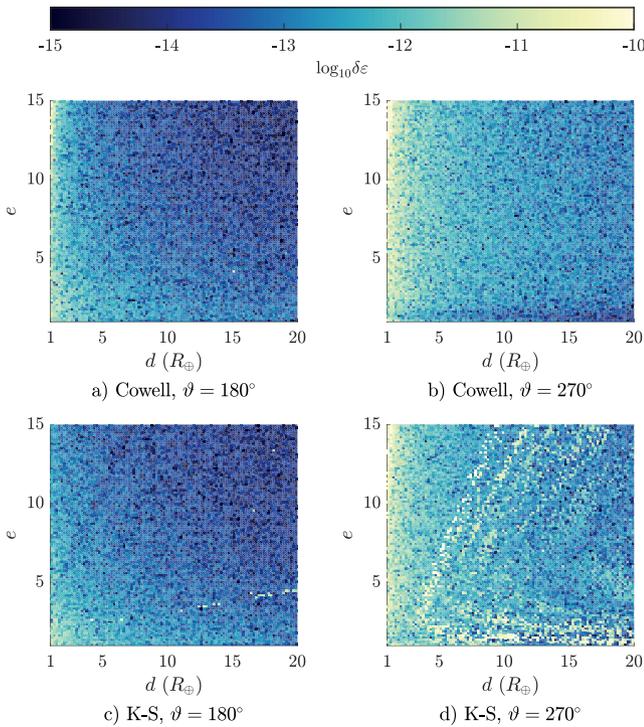


Figure 4. Relative energy error $\delta\epsilon$ (equation 36) as a function of the close encounter parameters d , e , for $\vartheta = 180^\circ, 270^\circ$ (see Fig. 1). The minimum approach distance d is measured in Earth radii. XRA15 was used with Cowell’s (upper panels) and K–S (lower panels) formulations. The accuracy parameters are $\epsilon = 10^{-6}$ and 10^{-7} , respectively.

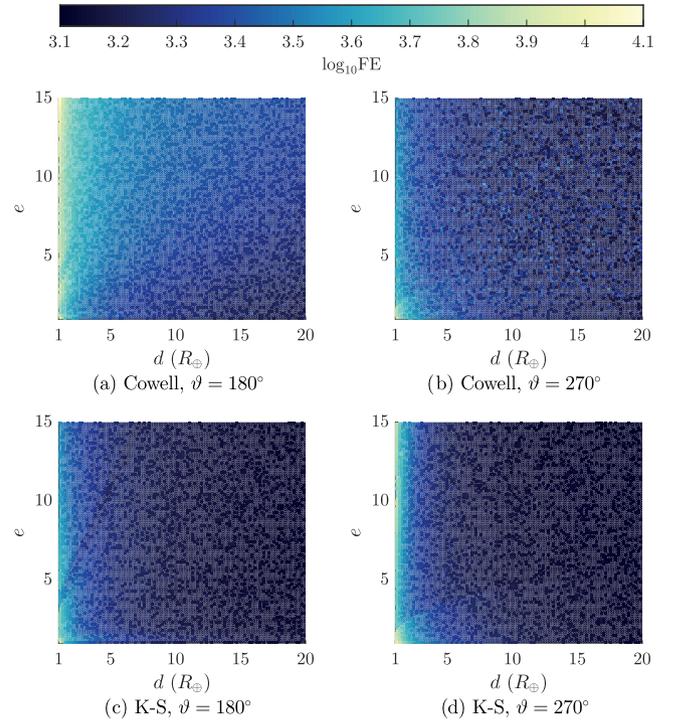


Figure 5. Number of function evaluations FE for the same simulations of Fig. 3.

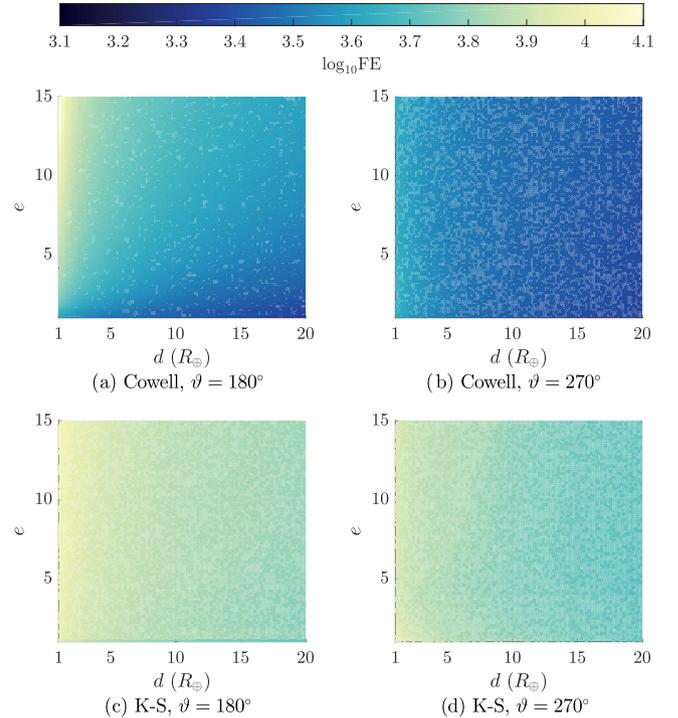


Figure 6. Number of function evaluations FE for the same simulations of Fig. 4.

of all these simulations is shown in Figs 5 and 6. We show the plots for the values $\vartheta = 180^\circ, 270^\circ$, which approximately correspond to extrema of $\delta\epsilon$ with respect to ϑ . The energy error depends on the heliocentric eccentricities before and after the encounter (e^\pm , equation 32) and on the change in velocity imposed to the particle (Δv , equation 33). The latter is strictly related to the magnitude of

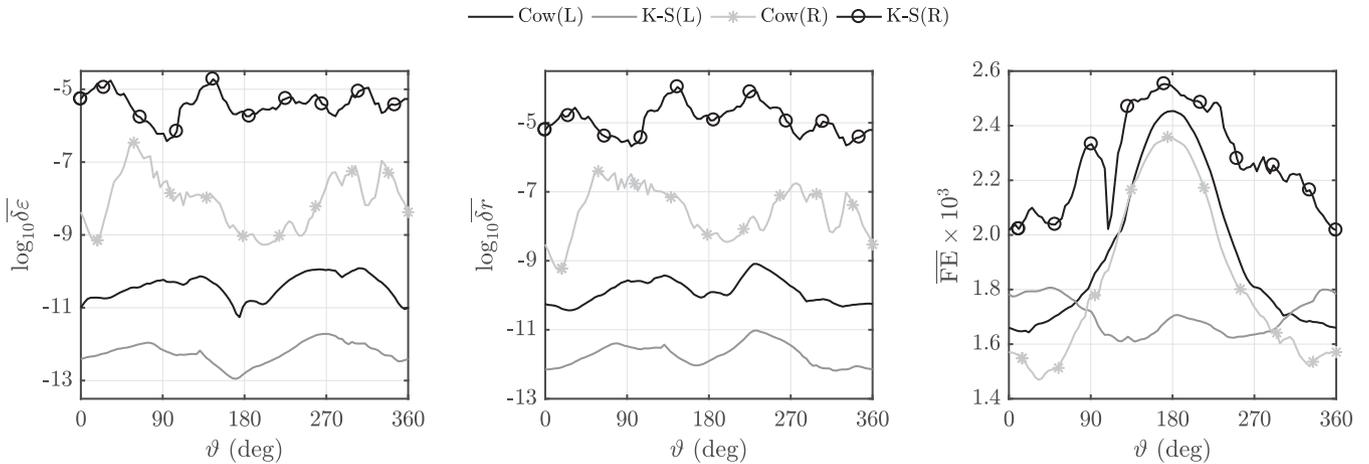


Figure 7. Average relative energy error $\overline{\delta \varepsilon}$, relative position error $\overline{\delta r}$, and number of function evaluations \overline{FE} as functions of ϑ (equation 37). The labels ‘Cow’ and ‘K–S’ refer to the Cowell’s and Kustaanheimo–Stiefel formulations. The letters ‘L’ and ‘R’ stand for the LSODAR and XRA15 numerical solvers. The accuracy parameters ϵ for K–S(R), Cow(L) and Cow(R) are chosen so that the computational cost is comparable to the one of K–S(L). We show for this method the best possible performance working in double precision.

the perturbation generated by the close encounter. Comparing the error plots with Fig. 2 shows that accuracy degrades where the Δv is larger. Moreover, observe that quasi-parabolic encounters, taking place at $e \approx 1$, imply $\Delta v \approx 0$ in the patched conics approximation, and in fact they are propagated very accurately (this is particularly evident by looking at the blue regions in the lowest part of the plots for $\vartheta = 270^\circ$). The left half of the (d, e) region corresponds to high values of e^\pm and Δv . Both are detrimental to the accuracy of the propagations and increase the computational cost.

Fig. 7 displays the behaviour of the average relative errors in energy and position, and of the average number of function evaluations, with respect to ϑ , computed by equation (37). These plots reveal how the performance is influenced through the angle ϑ by the heliocentric Keplerian energy (ε^\pm , equation 31) and its variation $\Delta \varepsilon$ (equation 34). The former affects the computational cost: more negative values of the energy correspond to orbits with faster time-scales and shorter periods. Because the time interval of propagation is the same for all the simulations, the number of steps will be higher for these orbits. A faster time-scale also implies sharper changes in the perturbing accelerations, which increase the cost. Since for most of close encounters φ takes values close to 90° ($\varphi > 70^\circ$ in 86% of the simulations), we have that variations in ε^\pm are mainly driven by variations of $\cos \vartheta$. In fact, the number of function evaluations in Cowell’s formulation rises markedly around $\vartheta = 180^\circ$. The quantity $\Delta \varepsilon$ affects instead the accuracy. As we could expect from equation (34) the error exhibits smaller values when $\sin \vartheta$ is close to 0.

6.3.2 Comparison of the propagation methods

The combination of K–S and LSODAR allows one to obtain the best performance. Fig. 7 shows that this method is two and four orders of magnitude more accurate than integrating Newtonian equations with LSODAR and XRA15, respectively, for a comparable computational cost.

The remarkable efficiency displayed by the K–S regularization vanishes when the XRA15 solver is used. Figs 4 and 6 show that K–S is computationally more expensive than Cowell for the same level of accuracy. Therefore, quite a large value of ϵ had to be chosen in

Fig. 7 with energy and position errors that are roughly three orders of magnitude larger. This unusual behaviour of K–S arises because in the initialization phase the XRA15 corrector loop requires many more iterations to converge with respect to Cowell. The resulting number of function evaluations becomes dominant when the time interval of propagation is short, as in this case. We are not aware of any previous works in which the Everhart–Radau numerical scheme has been employed in the integration of regularized formulations, hence this issue might have come unnoticed until now.

The propagation method that, in double precision, reaches the highest accuracy as measured in terms of $\overline{\delta \varepsilon}$, $\overline{\delta r}$ is Cowell’s formulation and XRA15. It obtains the lowest average energy and position errors, on the order of 10^{-13} and 10^{-12} , respectively. In fact, on the scale of 1 au a relative position error of 10^{-11} corresponds to sub-metric accuracy. However, the computational cost required by the method is very high and makes it uninteresting for this comparison. The combination of K–S and LSODAR with $\epsilon = 10^{-16}$ is only slightly less accurate, but the number of function evaluations is 2.5 times smaller.

6.3.3 Impact of LSODAR order changes on accuracy

We highlight a phenomenon that arises with the LSODAR solver when it is used with the Cowell formulation. For values of the accuracy parameter bigger than about 10^{-15} , the numerical error shows an erratic behaviour in the (d, e) region (see Fig. 8a). This is due to how LSODAR handles very rapid variations on the right-hand side of the differential equations. If the predicted step size for a given integration order is substantially reduced to satisfy the bound on the local truncation error, then the order is increased so that a larger step size can be applied. For each simulation of Fig. 8(a), we computed the average order \bar{q} over the number of integration steps. By juxtaposing the patterns of the energy error and of the average order as in Figs 8(a) and (b), we find out that order changes have a dramatic impact on the accuracy. Apart from these considerations, choosing higher values of ϵ in LSODAR does not bring a qualitative change of the numerical error behaviour as a function of (d, e) . Finally, we carried out the same type of analysis using K–S. For this

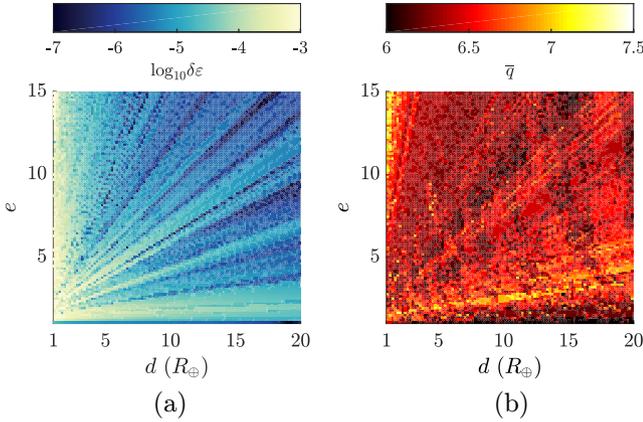


Figure 8. Relative energy error $\delta\epsilon$ (equation 36) and average integration order \bar{q} as a function of the close encounter parameters d , e for $\vartheta = 180^\circ$ (see Fig. 1). The minimum approach distance d is measured in Earth radii. LSODAR was used with Cowell’s formulation and the accuracy parameter is $\epsilon = 10^{-8}$.

formulation the numerical error has a much smoother behaviour, and it is less dependent on \bar{q} .

6.3.4 Element formulations

We do not show here results obtained with the EDromo formulation; its performance is much less sensitive to the heliocentric eccentricity, but it is heavily affected by the magnitude of Δv . Using EDromo without performing the trajectory splitting is disadvantageous with both solvers as high values of the numerical error and function evaluations are obtained. We also noticed that LSODAR often switches to the BDF method to prevent the arising of stability issues. This signals that the right-hand side of the EDromo equations is not well behaved if the Sun is kept as the primary body throughout all the integration. The reason is that a fast and strong perturbation on a heliocentric time-scale considerably deforms the trajectory, which cannot be well described by orbital elements. By introducing trajectory splitting it is possible to reduce the magnitude of the perturbations, thus substantially improving the performance of element formulations, as it will be shown in the next section.

6.4 Simulations with trajectory splitting

We describe the numerical experiments for the case in which the trajectory is split during the integration of a close encounter according to the procedure detailed in Section 3.

6.4.1 Detection of close encounters

For *test* simulations that apply trajectory splitting we want to know when the distance r between the particle and the Earth is equal to some fixed value R_{sw} (Section 3). First, in the *reference* integration we solve the event location problem $r(t) - R_{sw} = 0$. Since we only consider one close encounter, there exist two solutions t_1 , t_2 , representing the times of beginning and end of the geocentric phase. Then, for each test simulation we solve the event location problem $t(s) - t^* = 0$, where t^* will be t_1 , t_2 in sequence. In practice, the switch of the primary is done when the geocentric distance

is equal to R_{sw} along the *reference* propagation.⁵ The reason for using an event function based on time rather than on distance is that the former, being a strictly increasing function of s , is more convenient for the application of an algorithm for root-finding as the one described in Appendix A.

6.4.2 Impact of the switch distance on the performance

The primary body is switched at a *suitable* geocentric distance R_{sw} along the trajectory. Ideally, we would like R_{sw} to be optimal in the sense of minimizing the numerical error and the computational cost. Finding such a switch distance is not trivial: it should be neither too large nor too small, since strong third-body perturbations would arise in the geocentric and heliocentric propagation phases, respectively. This is the rationale that is often used to justify the change of the origin of the coordinate system at the sphere of influence, or at the Hill sphere, whose radii for the Earth are approximated by R_{SoI} and R_{Hill} , respectively,

$$R_{SoI} = a_{\oplus} \left(\frac{\mu_{\oplus}}{\mu_{\odot}} \right)^{2/5} \approx 0.0062 \text{ au}, \quad (41)$$

$$R_{Hill} = a_{\oplus} \left(\frac{\mu_{\oplus}}{3\mu_{\odot}} \right)^{1/3} \approx 0.01 \text{ au}. \quad (42)$$

However, these criteria do not take into account the characteristics of the numerical scheme and of the formulation, hence they might not guarantee the best efficiency. For instance, in regularized schemes the perturbing force does not enter the right-hand side of the differential equations alone, because it is multiplied by functions of the state vector.

The optimal switch distance is searched for numerically as follows. We consider the same pairs (d, e) as for the simulations without trajectory splitting. The average errors, $\delta\epsilon$, δr , and function evaluations FE are computed from equation (37) for $\vartheta = 0^\circ, 90^\circ, 180^\circ, 270^\circ$ and 100 values of R_{sw} in the range $[0.2, 8] \times R_{SoI} \approx [0.0012, 0.049] \text{ au}$. Many propagation methods are included in the comparison, with the possibility of adopting distinct formulations for the heliocentric and geocentric phases. Only the performance of the best methods is shown here.

Fig. 9 displays results for propagations done with LSODAR and the K–S formulation, while in Fig. 10 EDromo is employed (with a constant time element) instead of K–S when the origin of the coordinate system is the Sun. The energy and the position errors are not affected by R_{sw} if this is greater than 0.01 au, while for shorter distances the errors may rise by up to one order of magnitude. On the other hand, with $\vartheta = 90^\circ, 180^\circ$ and 270° the computational cost shows pronounced minima for values of R_{sw} between 0.01 and 0.025 au. The case $\vartheta = 0^\circ$ deserves further attention: in this geometry, most of the particles approach the Earth’s orbit nearly tangentially. A larger period of the heliocentric orbit is spent in a highly perturbed region close to the Earth, so the change of primary body is more convenient at a larger distance. In general, switch distances smaller than $1 R_{SoI}$ must be avoided since they cause the computational cost to increase of up to 60 per cent, without bringing any improvement of the accuracy.

The best performance with the XRA15 solver is obtained by employing EDromo in the heliocentric phase and Cowell in the geocentric phase. The behaviours of the errors and computational cost

⁵ For a generic test simulation the geocentric distance at $t = t_1$, $t = t_2$ will differ by some small amount from R_{sw} . This does not affect considerations on the performance of the methods as a function of R_{sw} .

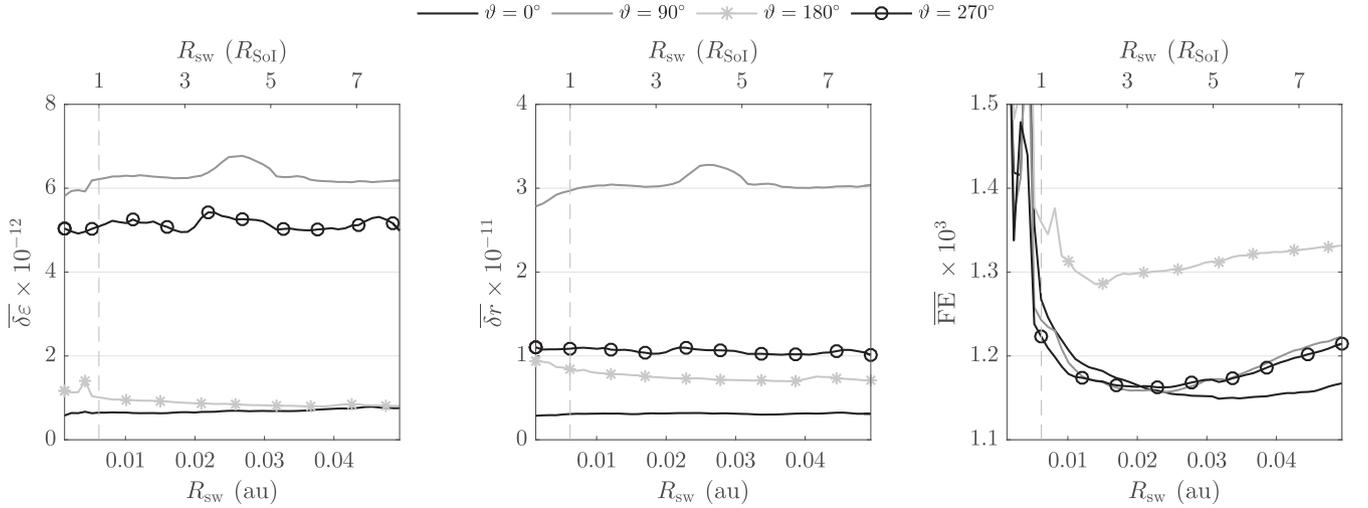


Figure 9. Average relative energy error $\overline{\delta\varepsilon}$, relative position error $\overline{\delta r}$ and number of function evaluations \overline{FE} as functions of the switch distance R_{sw} for four values of ϑ (equation 37). The K–S formulation is used with LSODAR and the accuracy parameter is $\epsilon = 10^{-15}$. The dashed grey line corresponds to one radius of the sphere of influence.

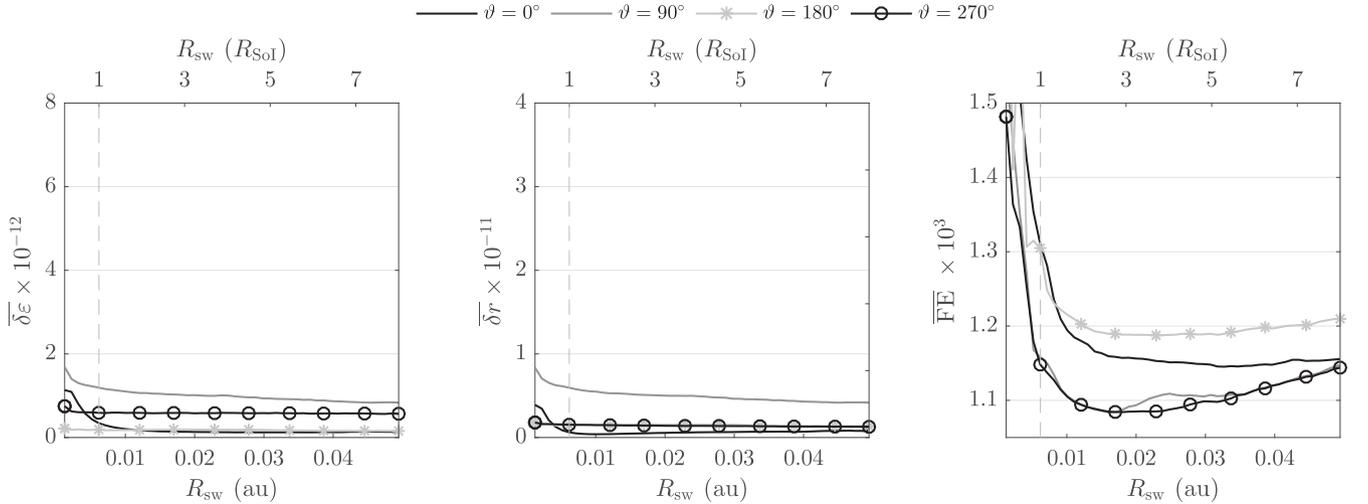


Figure 10. Same as Fig. 9. Here EDromo is used in the heliocentric propagation and K–S in the geocentric one. The numerical solver is LSODAR and the accuracy parameter is $\epsilon = 10^{-15}$.

shown in Fig. 11 depart from those obtained with LSODAR. The position and energy errors exhibit an evident ‘plateau’ for R_{sw} approximately between 1 and 3 Hill radii; outside this region, they may increase by up to two orders of magnitude. The number of function evaluations decreases steadily with R_{sw} , only showing signs of growth near the highest values of the considered range. We note that while $\epsilon = 10^{-3}$ is the round-off limit only for EDromo, it does not allow us to get the best accuracy with Cowell’s formulation. This is the reason why for R_{sw} bigger than $3 R_{Hill}$ the error accumulated in the geocentric propagation starts to deteriorate the accuracy of the whole integration. If a stricter accuracy parameter is used, as $\epsilon = 10^{-5}$, the error plateau is extended but the number of function evaluations, for the same accuracy, becomes comparable to that of Cowell’s method without trajectory splitting.

As a simple rule of thumb, splitting the propagation between 1.2 and 3 Hill radii yields the best numerical performances. On the other hand, by choosing one radius of the sphere of influence we do not achieve, in general, the best performance with regularized methods. Note that although the main features of close encounters

are well replicated in the CR3BP, the value of R_{sw} maximizing the performance must be chosen depending on the particular problem.

6.4.3 Comparison of the propagation methods

We analysed the performances of the propagation methods that apply the trajectory splitting in terms of accuracy and computational cost for several values of ϑ in the range $[0^\circ, 360^\circ]$, as previously done in Fig. 7. Taking advantage of the results shown in Section 6.4.2 we selected the value 0.015 au for the switch distance because it guarantees an almost optimal efficiency in all the simulations. The best methods at the round-off limit are the pairs EDromo, K–S with LSODAR ($\epsilon = 10^{-15}$), and EDromo, Cowell with XRA15 ($\epsilon = 10^{-3}$). In particular, while the latter is more accurate, the former is computationally faster. In Figs 12 and 13 we separately show their performance together with other propagation methods, which do not split the trajectory and adopt the same solver, for a comparable average number of function evaluations. Concerning the integration of K–S equations with LSODAR, changing the primary

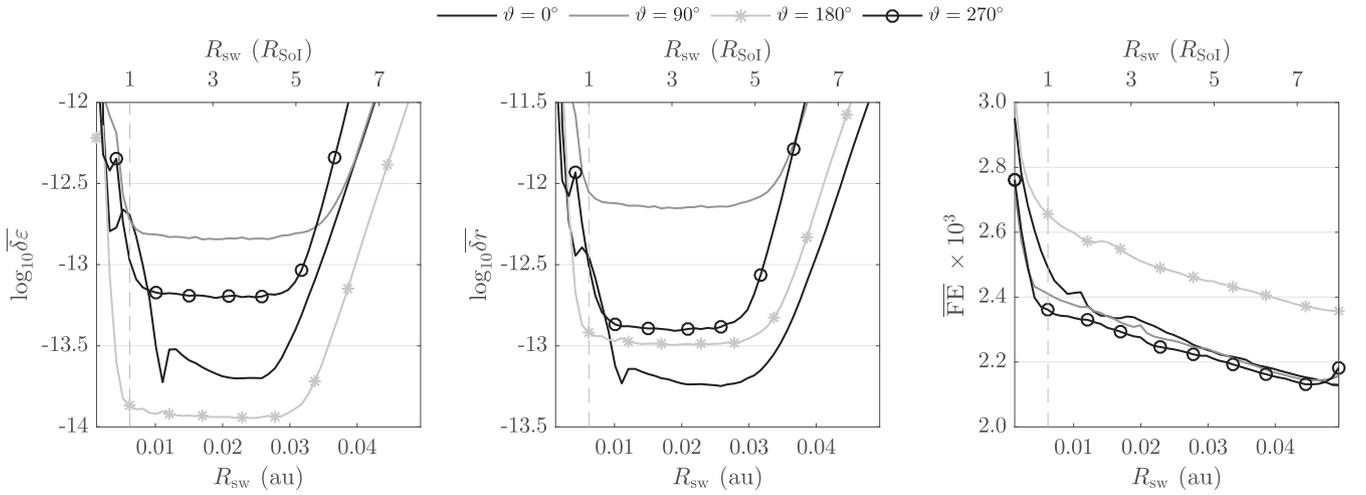


Figure 11. Same as Fig. 9. Here EDromo is used in the heliocentric propagation and Cowell in the geocentric one. The numerical solver is XRA15 and the accuracy parameter is $\epsilon = 10^{-3}$.

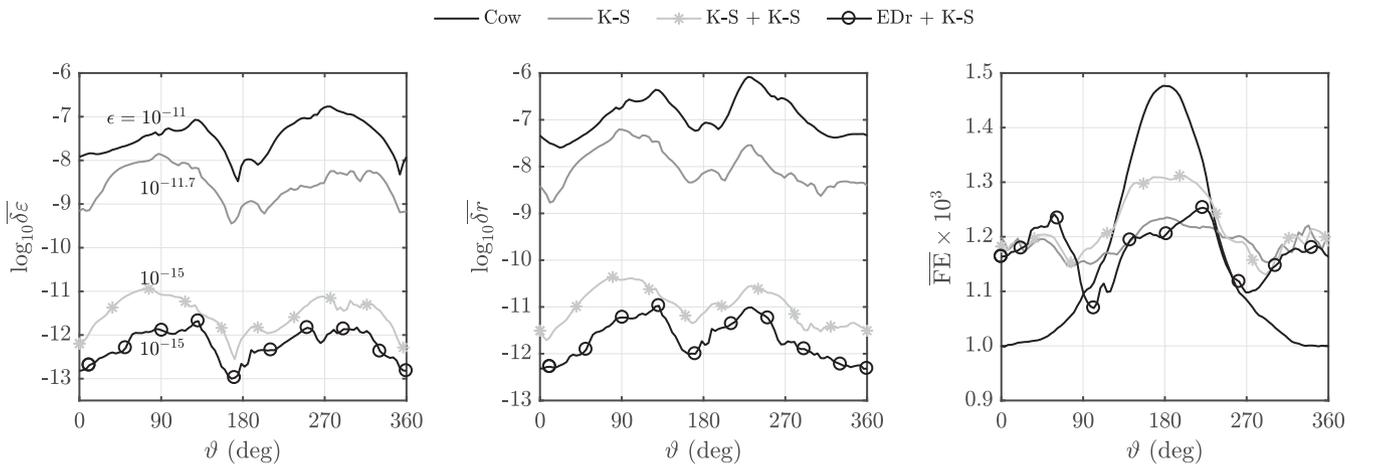


Figure 12. Performance comparison of several propagation methods with the LSODAR solver in terms of the average relative energy and position errors ($\overline{\delta\epsilon}$, $\overline{\delta r}$) and number of function evaluations (FE). The curves labelled ‘Cow’ and ‘K-S’ refer to propagations with Cowell’s and Kustaanheimo–Stiefel formulations without performing any change of the primary. In the other two methods that we consider the trajectory is split at $R_{sw} = 0.015$ au with K-S employed in the geocentric propagation and either K-S or EDromo in the heliocentric phase. The accuracy parameters ϵ for Cow, K-S, K-S + K-S are chosen so that the computational cost is comparable to the one of EDr + K-S. We show for this method the best possible performance working in double precision.

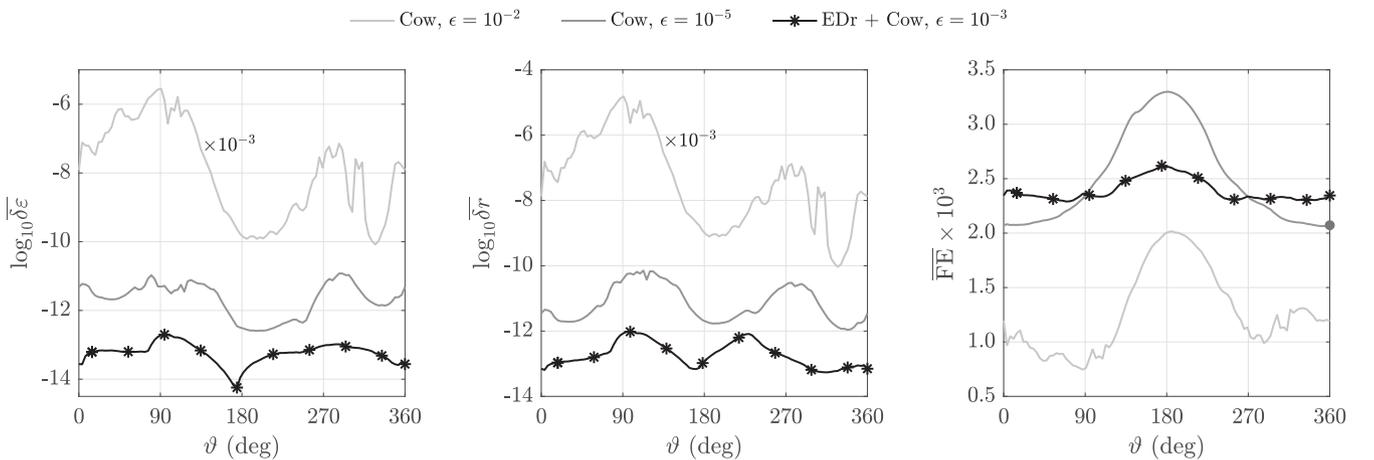


Figure 13. Same as Fig. 12 but here we consider the XRA15 solver. Cowell’s method with the Sun at the origin of the coordinate system is compared to EDromo and Cowell’s formulations employed in the heliocentric and geocentric phases, respectively, with $R_{sw} = 0.015$ au. We show for EDr + Cow the best possible performance working in double precision.

improves the accuracy by three orders of magnitude. When K–S is employed only in the geocentric propagation and EDromo in the heliocentric one, the errors are reduced by four orders of magnitude instead. The mitigation of the error in the pre-encounter phase afforded by this element formulation is highly beneficial to numerical efficiency. For the XRA15 solver, using EDromo and Cowell when the Sun and the Earth are, respectively, the primary body enhances the accuracy by two orders of magnitude with respect to Cowell’s formulation without trajectory splitting. The use of K–S instead of Cowell in the geocentric phase does not lead to a further improvement for this solver. It is remarkable that although regularized methods have an overhead due to the event location algorithm implemented in XRA15, EDromo exhibits an excellent numerical performance. Finally, note that for a similar computational effort of the combination EDromo, K–S in Fig. 12, Cowell’s method loses up to six more digits in the correct final position (see Fig. 13). In fact, the energy error oscillates violently between 10^{-5} and 10^{-10} , indicating that the exact evolution of some trajectories is not even coarsely reproduced by the numerical integration.

6.4.4 Regularized element methods for hyperbolic motion

A regularized element formulation conceptually analogous to EDromo has been recently developed for orbits with strictly positive total energy (Roa & Peláez 2015; Baù et al. 2016). We carried out many simulations adopting this scheme in the geocentric phase. However, some trajectories exhibit peculiar dynamical phenomena that complicate its use with trajectory splitting. Therefore, we do not show any results, and we only explain the issues that arise. For slow encounters the Keplerian energy along the geocentric motion may attain a negative value at some point, leading to a temporary capture. In this case the formulation fails. Another problematic event, which may occur in the highly perturbed region far from the Earth, is when the angular momentum vanishes. The method incurs in a singularity and the whole propagation fails. It is always possible to avoid these scenarios by setting a sufficiently small switch radius. However, by proceeding in this way we are not sure to minimize the numerical error, therefore the performance analysis would get more difficult.

7 LONG-TERM PROPAGATION OF (99942) APOPHIS

In the previous section we dealt with short-term, bidimensional propagations in the planar CR3BP. We extend our analysis to the long-term, three-dimensional propagation of a NEA. As a test case of particular interest for impact monitoring activities, we choose the asteroid (99942) Apophis. This NEA will experience a close encounter with the Earth in April 2029 at a minimum distance d of $6 R_{\oplus}$. Such a small value of d corresponds to a very deep encounter, which may be followed by possible resonant returns. We use the same physical model as in the previous section, that is the Sun–Earth CR3BP. Note that an actual impact risk analysis would require a more refined physical model, as in Farnocchia et al. (2013).

7.1 Reference trajectory

The *reference* trajectory is obtained with the same procedure described in Section 6.1. The position and velocity at the epoch of the closest approach are given by the nominal solution available

in JPL’s HORIZONS system.⁶ The total duration of the propagation is 100 yr, which is the time interval habitually considered for impact monitoring activities. We checked that the main features of the 2029 encounter are replicated in our propagation. In particular, we reproduce Apophis’ jump from the Apollo to the Atens group.

7.2 Amplification of round-off error

The amplification of an initial round-off error after the close encounter imposes a limit to the maximum accuracy achievable by the numerical integration. As to quantify this lower bound, we carry out another propagation in which the magnitude of the initial velocity v_i , relative to the Sun, is perturbed as follows:

$$\bar{v}_i = v_i (1 + p), \quad p = 3\epsilon_{\text{dp}}, \quad (43)$$

where $\epsilon_{\text{dp}} = 2.22 \times 10^{-16}$ is the double-precision machine zero. The propagation is performed in quadruple precision and with the same formulation, solver and accuracy parameter of the reference one. Hence, the difference between the two trajectories is due to the effect of the encounter on the initial deviation in velocity.

By never performing propagations under the round-off limit (Section 6.2), we ensure that the measured errors are not affected by round-off error. However, understanding its behaviour for both solvers may be especially important for highly accurate propagations on time spans longer than a century. We estimated the rate of growth of round-off error by propagating (99942) Apophis in the unperturbed case for up to 10^6 integration steps, and by monitoring the error in the two-body orbital energy. When LSODAR is employed with the coordinate-based formulations Cowell and K–S, the round-off error accumulates linearly in time. If XRA15 is used instead, round-off grows as the square root of time with the Cowell formulation, thus following Brouwer’s law (Brouwer 1937). This result is in agreement with that obtained by Rein & Spiegel (2015) for the IAS15 solver. Thus, we discover that the achievement of Brouwer’s law in IAS15 is not exclusively due to the suppression of round-off error by compensated summation, since this device is not implemented in XRA15. When the K–S formulation is integrated with XRA15, the round-off error again grows linearly.

If the EDromo formulation is employed with a constant time element, there is no round-off accumulation in the unperturbed problem since the right-hand side of equations (15)–(21) and of equation (25) is identically zero. Integrating equation (24) for the linear time element with LSODAR again gives rise to a linear drift in the orbital energy, but with the XRA15 solver the error is kept within machine precision. The reason for this quite interesting result might be connected to the very high order of the solver, but it is not yet completely understood.

7.2.1 Considerations on chaos

In the same way as in equation (43), we have also perturbed the single components of the initial velocity and position for several values of p , with $3\epsilon_{\text{dp}} < p < 10\epsilon_{\text{dp}}$. In all of the tests, the distance δr_f between the perturbed and the reference positions at the end of the propagation does not exceed 10^{-5} au. In both the pre-encounter and post-encounter phases, the distance δr between the perturbed and reference trajectories grows linearly. Moreover, the close encounter amplifies the δr by the same factor for all values of p . Therefore, the

⁶ <http://ssd.jpl.nasa.gov/?horizons>, last visited 2016 November 22. Site manager: Ryan S. Park.

Table 1. Numerical error and computational cost for propagations of 100 yr of the asteroid (99942) Apophis. The errors δr , δv and δJ are, respectively, in position, velocity and on the Jacobi constant (the latter is a relative error). The computational cost is shown by the number of function evaluations and the CPU time. The accuracy parameter of EDr + K–S and Cow corresponds to the round-off limit, while for K–S+K–S and Cow* it is chosen to have a computational cost similar to that of EDr + K–S. The last two rows refer to the conservative Bulirsch-Stoer (BS) integrator and the hybrid method of the MERCURY software, which are ran at the round-off limit. For these solvers, ϵ is the tolerance of the BS integrator. The time step of the mixed-variable symplectic integrator is 0.1 d.

Solver	Formulation	$\log_{10}\epsilon$	δr (au)	δv (au d ⁻¹)	δJ	FE	CPU time (s)
LSODAR	EDr + K–S	– 16	4.00×10^{-8}	3.15×10^{-11}	1.01×10^{-15}	16 980	0.134
	Cow	– 15	1.65×10^{-6}	1.29×10^{-9}	1.11×10^{-13}	60 881	0.178
	Cow*	– 8	0.85	6.10×10^{-4}	1.55×10^{-6}	17 364	0.0660
	K–S + K–S	– 13	3.21×10^{-5}	2.53×10^{-8}	2.16×10^{-12}	17 391	0.0934
XRA15	EDr + K–S	– 3	9.61×10^{-10}	7.57×10^{-13}	4.81×10^{-16}	49 077	0.159
	Cow	– 7	3.53×10^{-8}	2.78×10^{-11}	9.54×10^{-16}	110 013	0.251
	Cow*	– 1.4	1.01	6.83×10^{-4}	2.19×10^{-5}	50 854	0.0856
	K–S + K–S	– 1.5	7.47×10^{-4}	5.93×10^{-7}	9.32×10^{-11}	48 260	0.128
BS2	Cow	– 14	8.29×10^{-9}	1.11×10^{-10}	3.82×10^{-13}	66 875	0.333
Hyb	–	– 14	1.51×10^{-5}	2.04×10^{-7}	5.63×10^{-12}	–	0.658

magnitude of δr_f is proportional to p . In fact, while the pre-encounter distance is strongly amplified due to the encounter, the regular evolution of the perturbed trajectories indicates that the problem is not chaotic on the considered time span. While a single close encounter results in a linear divergence of initially close trajectories, chaotic dynamics only arise due to *sequences* of close encounters through the mechanism mentioned in the introduction. Thus, we do not expect the dynamics to be chaotic on this time span.

7.3 Performance analysis

We run in double precision *test* simulations starting 50 yr before the 2029 close encounter and covering a time interval of 100 yr. Several propagation methods are considered with LSODAR and XRA15. We include in our comparison Cowell’s formulation, keeping the Sun always at the origin of the coordinate system and regularized schemes that are used with trajectory splitting at $R_{sw} = 0.016$ au. More specifically, the Kustaanheimo–Stiefel regularization is employed in the geocentric propagation, and either K–S or EDromo in the heliocentric phases. For EDromo the linear time element is implemented instead of the constant one, because it is more efficient for long-term propagations (Baù & Bombardelli 2014).

We measure the accuracy of the propagations by computing the errors in position and velocity with respect to the reference. Moreover, we take into account the relative error on the Jacobi constant J (Battin 1999, p. 373),

$$\delta J(t) = \left| \frac{J(t) - J_i}{J_i} \right|, \quad (44)$$

where $J_i = J(t_i)$ is the initial value. We estimate the computational cost by measuring the number of function evaluations and the computational time. The latter is computed as the average of 10 identical runs on an Intel i7 quad-core machine at 3.10 GHz, using Linux Mint 18 and the GFORTRAN 5.4.0 compiler without optimizations.

The numerical error and computational cost for each of the tests are displayed in Table 1. At the round-off limit (Section 6.2) the propagation method that reaches the highest accuracy is the pair EDromo, K–S for both numerical solvers. When compared to Cowell’s method the position and velocity errors at the end of the propagation are two orders of magnitude smaller. We also report the performance of Cowell’s formulation and the pair K–S, K–S for

a comparable number of function evaluations to that of the best method. Cowell’s final position error is in the order of 1 au, while for the combination EDromo, K–S it can be as low as 100 m. This value is even smaller than the difference between the reference trajectory and the one starting with the perturbed initial velocity. It is also notable that by employing EDromo instead of K–S in the heliocentric propagation the position accuracy is improved by three and six orders of magnitude when using LSODAR and XRA15, respectively.

The errors on position, velocity and Jacobi constant of the propagation methods as a function of time are displayed in Figs 14–16. Since the 2029 close encounter amplifies the error in position and velocity by a factor of 10 000, the best way to reduce the final error is by improving the accuracy in the computation of the pre-encounter arc of the trajectory. EDromo achieves the lowest pre-encounter errors in position, velocity and on the Jacobi constant. In particular, the position error is kept around 10^{-12} or 10^{-13} au depending on the numerical solver. It is difficult to find a reliable fit to the errors with a power law in the pre-encounter part of the propagation. This is due to the fact that the asteroid only completes 63 revolutions from 1979 to 2030, thus the fit is not performed on a very large time span.

Contrary to the errors in position and velocity, the error on the Jacobi constant does not show a significant increase after the close encounter. However, in impact monitoring activities it is the position of the asteroid the quantity that must be known most accurately. Therefore, the quality of numerical integrations should be measured by using errors on the fast variables, i.e. position and velocity.

In all the simulations in which trajectory splitting is applied the number of function evaluations relative to the geocentric propagation accounts for less than 3% of the total. As long as the error accumulated in this portion does not compromise the accuracy of the post-encounter heliocentric phase, the choice of the formulation to be used when the Earth is the primary body is of secondary importance.

7.3.1 Propagation efficiency

For all the methods and for both solvers considered, we repeated the propagation of (99942) Apophis with decreasing values of the accuracy parameter ϵ . The final errors and number of function evaluations are displayed in Figs 17–19.

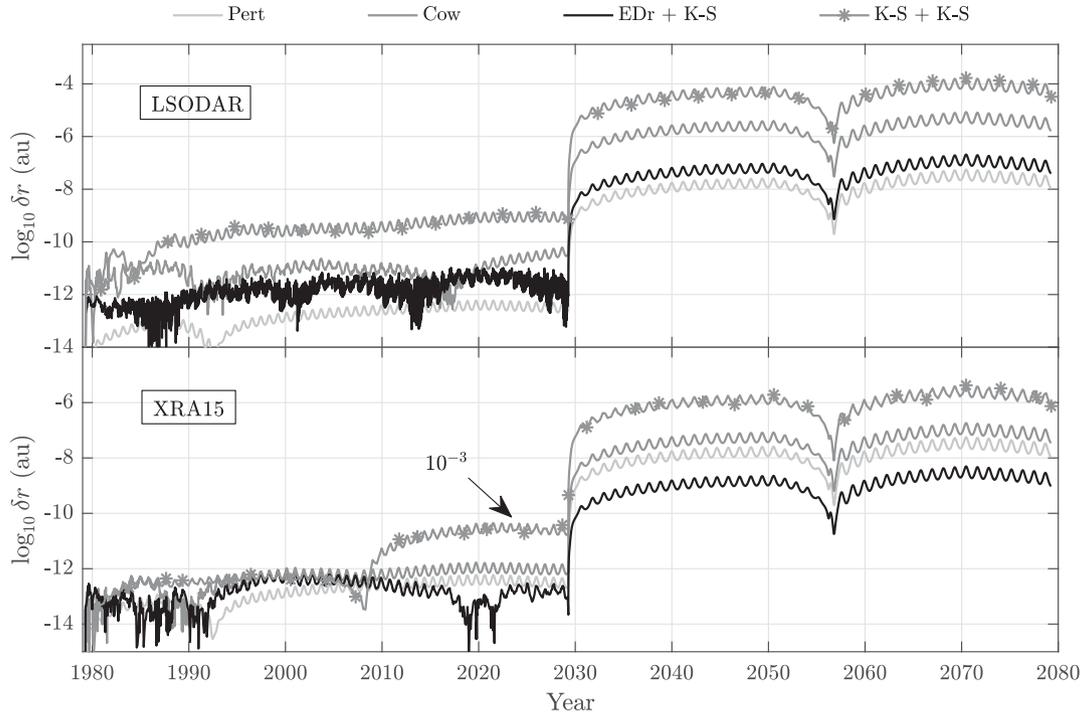


Figure 14. Time evolution of the position error of several methods for the long-term propagation of the asteroid (99942) Apophis. The curve ‘Pert’ is obtained in quadruple precision and with a tiny variation of the initial velocity (more details are given in Section 7.2). The other curves refer to the same simulations of Table 1. Cow and EDr + K–S represent the highest accuracy that one can get from Cowell’s formulation (keeping the Sun as primary body) and the combination EDromo, K–S with the latter employed in the geocentric propagation. If we use K–S also in the heliocentric phase instead of EDromo, the resulting error is shown by K–S + K–S for a computational cost similar to that of EDr + K–S. The errors of the K–S + K–S method for the xRA15 solver are multiplied by the factor of 10^{-3} .

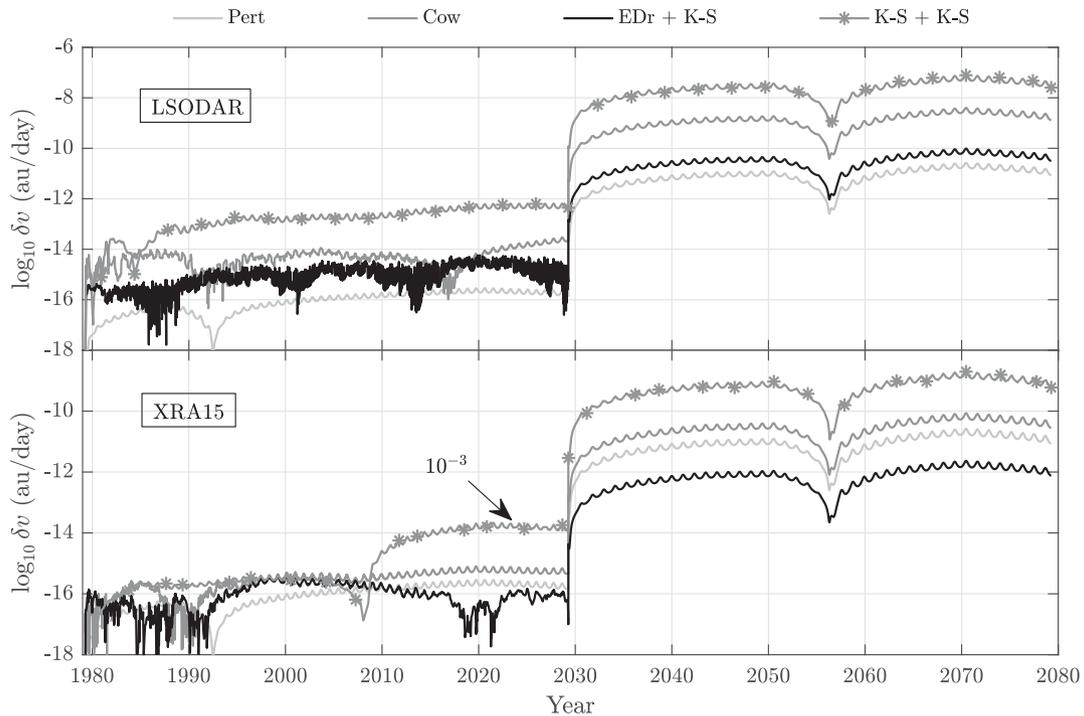


Figure 15. Time evolution of the velocity error of several methods for the long-term propagation of the asteroid (99942) Apophis. The description of the curves is the same as in Fig. 14.

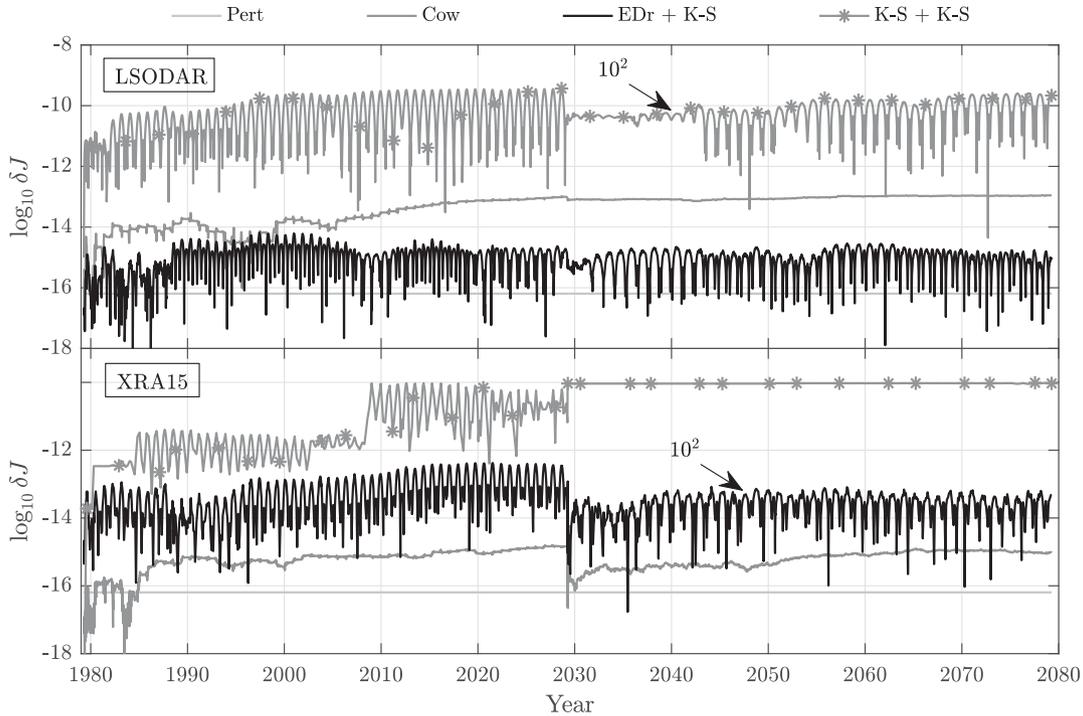


Figure 16. Time evolution of the relative error on the Jacobi constant (equation 44) of several methods for the long-term propagation of the asteroid (99942) Apophis. The description of the curves is the same as in Fig. 14. Some of the curves are multiplied by the factors displayed in the plots.

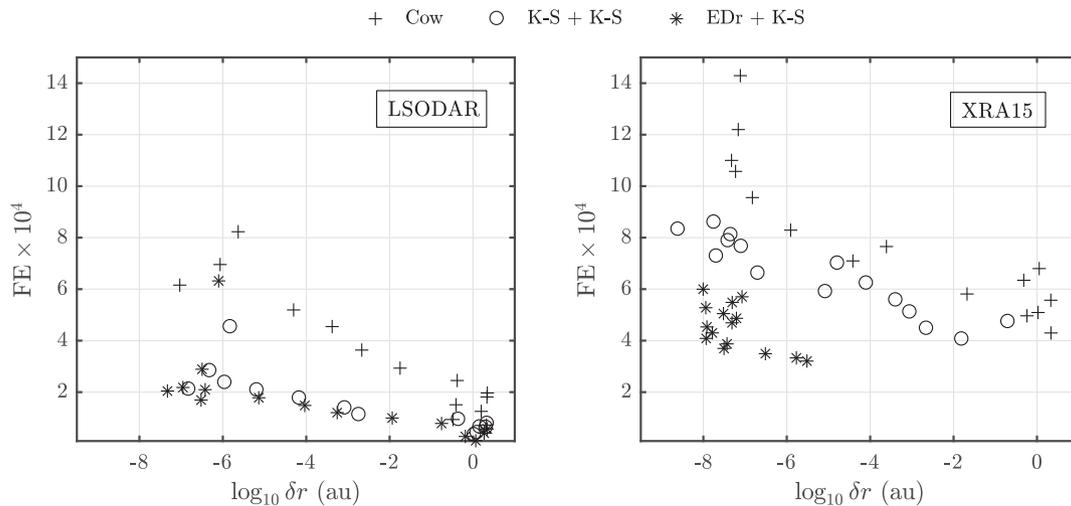


Figure 17. Function evaluations versus final position error of several methods for the long-term propagation of the asteroid (99942) Apophis. The names of the curves refer to simulations performed with Cowell’s formulation (curve ‘Cow’), K–S in the geocentric phase and either K–S or EDromo in the heliocentric phase (curves ‘K–S + K–S’ and ‘EDr + K–S’). For the LSODAR solver, the curves were obtained by decreasing ϵ from 10^{-3} to 10^{-19} . For the XRA15 solver, ϵ was decreased from 0.01 to 10^{-8} when using Cowell’s formulation and to 10^{-4} for the others.

Trajectory splitting with EDromo in the heliocentric phase and K–S in the geocentric achieves the highest efficiency. It always exhibits a number of function evaluations from two to three times smaller than the Cowell formulation for both solvers. In particular, for the propagations performed at the round-off limit, the computational cost decreases by the factors of 3.6 and 2.2 with LSODAR and XRA15, respectively (Table 1). The advantage of using EDromo is particularly relevant for a solver of very high order, such as XRA15. In fact, the local truncation error is proportional to the $(k + 1)$ -th derivative of the right-hand side of the equations of motion. The smoothing of the right-hand side achieved by using

EDromo equations results in a much smaller value of this derivative. The higher the order of the solver k , the more the local truncation error is consequently reduced.

Using the K–S formulation in the heliocentric phase is also advantageous in combination with the LSODAR solver. For coarse accuracies the corresponding method shows similar performances to EDromo; however, the latter allows one to reach higher accuracies. The K–S formulation is less efficient when used in combination with XRA15, having a number of function evaluations that is intermediate between that of Cowell’s method and EDromo.

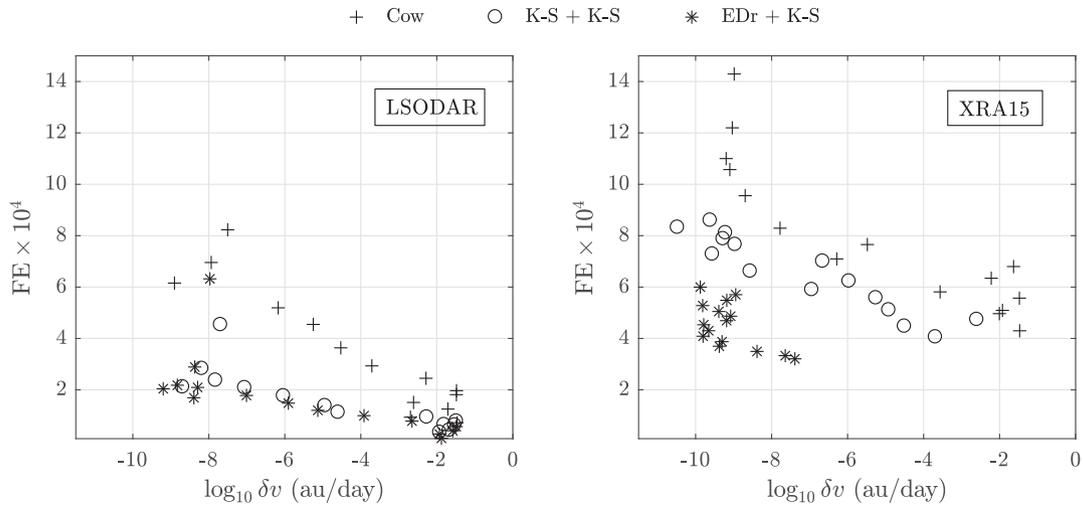


Figure 18. Function evaluations versus final velocity error of several methods for the long-term propagation of the asteroid (99942) Apophis. The description of the curves is the same as in Fig. 17. The points are obtained by varying ϵ .

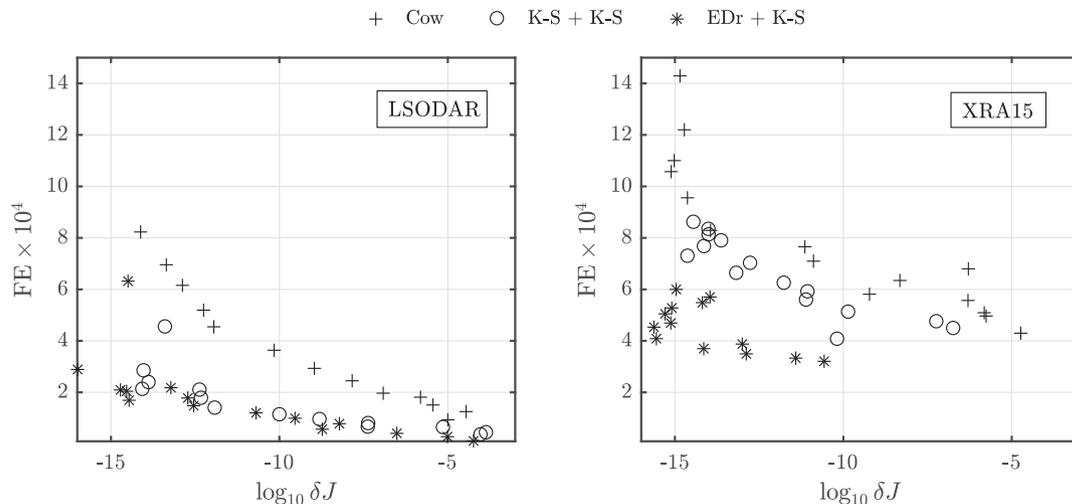


Figure 19. Function evaluations versus final relative error on the Jacobi constant (equation 44) of several methods for the long-term propagation of the asteroid (99942) Apophis. The description of the curves is the same as in Fig. 17. The points are obtained by varying ϵ .

Note that the smallest values of ϵ that we considered are smaller than the round-off limit, as to illustrate the entry into a round-off-dominated regime. As outlined in Section 6.2, round-off error starts dominating when the computational cost increases while the accuracy worsens. Using trajectory splitting with the regularized formulations also makes the integration more robust with respect to the accumulation of round-off error. The number of function evaluations after the round-off limit is passed (leftmost points in Figs 17–19) is not as high as in the case of the Cowell formulation.

7.3.2 Comparison with the *MERCURY* code

Having assessed that the method using EDromo is the most efficient among those examined, we compare it to the Bulirsch–Stoer method for conservative equations (Press et al. 1992, chap. 16) and the hybrid symplectic method of Chambers (1999), which are implemented in the *MERCURY* package.⁷ The first method integrates

second-order equations in heliocentric Cartesian coordinates using an approach based on Richardson extrapolation. The latter uses a second-order, mixed-variable symplectic scheme to integrate the dynamics with the Sun as a primary body. A close encounter starts when the distance between two bodies decreases below a *changeover radius* R_{crit} , fixed at the beginning of the propagation. During a close encounter, the dynamics are integrated at high precision using the BS method. The switch of integrators is handled through a continuous function, which ensures the formal symplecticity of the switching algorithm (Chambers 1999; Wisdom 2017). We propagate at the maximum available accuracy by decreasing the time step and the BS tolerance until the errors in position and velocity start increasing, with a procedure analogous to that explained in Section 6.2. We find that the round-off limit is reached with a time step of 0.1 d for the mixed-variable symplectic method and a tolerance of 10^{-14} for the BS method. We use the default value of 3 Hill radii for the changeover radius.

The numerical error and computational cost for the tests are reported in Table 1. We do not count the number of function evaluations for the hybrid method, and only report its computational

⁷ <http://star.arm.ac.uk/jec/mercury/mercury6.tar>, last visited 2017 April 5.

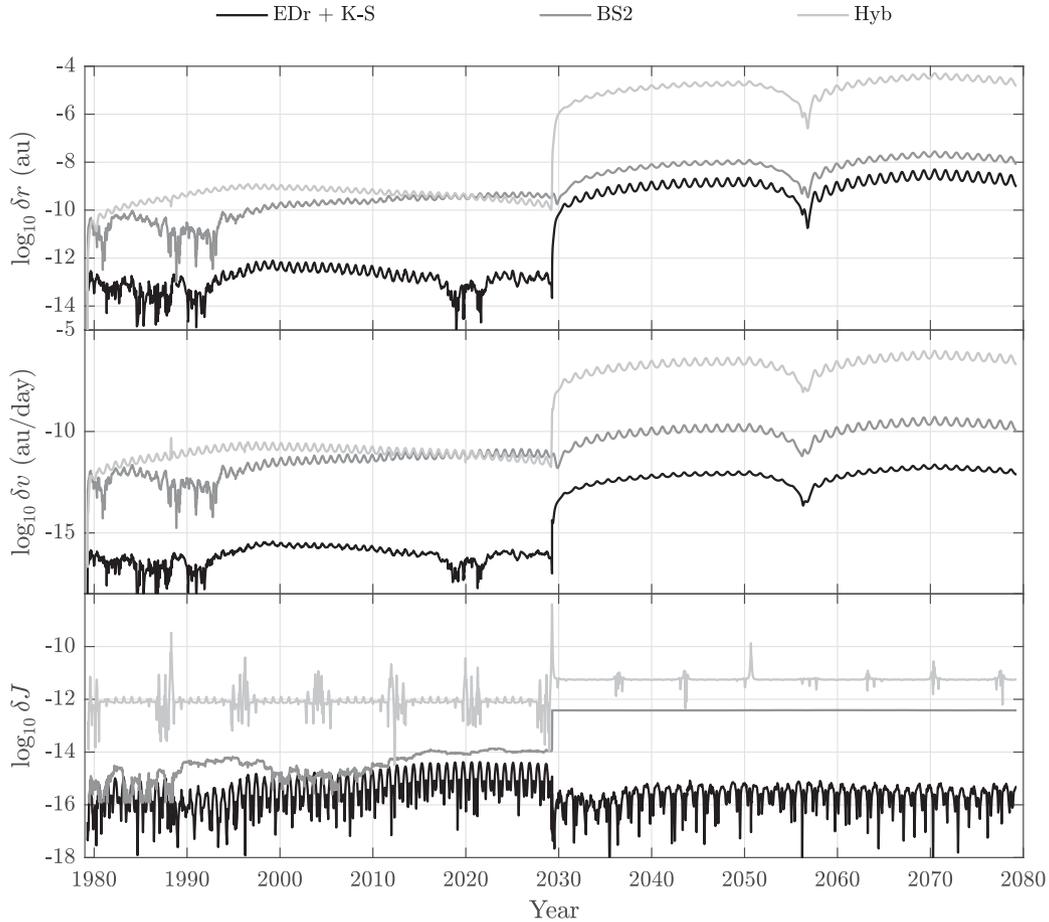


Figure 20. Time evolution of the errors on position, velocity and Jacobi constant (equation 44) for the long-term propagation of the asteroid (99942) Apophis. The curves ‘EDr+K-S’ are the same as those relative to the `XRA15` solver in Figs 14–16. The curves ‘BS2’ and ‘Hyb’ are obtained at the round-off limit with the conservative Bulirsch–Stoer integrator and the hybrid method implemented in the `MERCURY` code. Spikes in the error on the Jacobi constant take place when the distance from the Earth is less than 0.3 au.

time. In fact there are two choices to consider for the count of the function evaluations in the mixed-variable symplectic scheme, i.e. the evaluation of the Keplerian term and of the interaction term of the Hamiltonian, but choosing either of these does not result in an accurate metric of computational cost.

The BS method is four orders of magnitude more accurate in position and two times less computationally expensive than the hybrid, and appears to be competitive for this problem. However, it does not achieve the performance of the EDromo method, which has a similar accuracy in position with a computational cost that is 30% lower when used with the `XRA15` solver. The latter also requires a computational cost four times lower than the hybrid method while being four orders of magnitude more accurate, and keeps the final error on the Jacobi constant close to machine zero.

Fig. 20 shows the time history of the errors. Note that the time span of the integration is probably short enough for the global truncation error of conventional methods (such as BS) to be lower than the one of the symplectic, even if the truncation error on the Hamiltonian is bounded in the latter method. The BS method exhibits the smallest error jumps due to the close encounter, although it does not realize the lower final errors. This suggests that it might be applied efficiently to the integration of other formulations; however, a full characterization of the BS method is outside the scope of this paper.

Finally, we emphasize that the integrators in the `MERCURY` package have been recently shown to be not numerically symplectic

(Hernandez 2016). In this case, the lack of symplecticity does not appear to be concerning due to the satisfying accuracy in conserving the Jacobi constant.

8 SOFTWARE IMPLEMENTATION

The large-scale numerical experiments were performed with the Numerical Analysis of Planetary Encounters (`NAPLES`) FORTRAN 2003 code. To facilitate the reproducibility of the results, we make it available online as a GitHub repository, <https://github.com/dkamato/NAPLES>. The `XRA15` solver and its auxiliary files are also contained in the code.

9 CONCLUSIONS

The aim of this work is to improve numerical performances in the orbit computation of objects that experience planetary close encounters. We propose a strategy that consists of changing the origin of the coordinate system from the Sun to the encountered planet when the distance between the particle and the planet is shorter than a given value R_{sw} . Therefore we have a sequence of perturbed two-body problems, and each of them is solved by integrating regularized equations of motion with an adaptive solver. We consider the Kustaanheimo–Stiefel regularization (K–S) and EDromo, a formulation based on non-singular orbital elements. The numerical

solvers that we employ are a multistep and a single-step scheme, denominated `LSODAR` and `XRA15`, respectively. The latter is an improvement on Everhart's `RADAU` integrator of 15th-order (Rein & Spiegel 2015), and it includes an event location algorithm that we have developed for propagations with regularized formulations.

A systematic study is carried out through ensembles of short-term propagations in the planar Earth–Sun CR3BP. Each propagation is parametrized by three quantities, which univocally define a trajectory. In this way we simulate many different kinds of close encounters. Concerning propagations carried out without splitting the orbit, K–S shows an excellent performance when the numerical solver is `LSODAR`. For the same computational cost, this method increases the accuracy by up to four orders of magnitude with respect to the integration in Cartesian coordinates by `XRA15`. We also note that it is not convenient to combine this solver with the K–S formulation in short-term simulations due to the high number of function evaluations required by the initialization phase. The splitting strategy allows us to obtain the best performances. Use of `EDromo` in the heliocentric propagation and K–S in the geocentric one, both with `LSODAR`, is much more accurate than Cowell's method for the same computational effort.

We also studied the impact of the switch radius R_{sw} on the numerical efficiency. With an empirical analysis, we find out that both accuracy and cost of computation are improved if the primary body is changed between 1.2 and 3 Hill radii, rather than at the sphere of influence. The exact value of the switch radius maximizing the performance must be chosen depending on the specific physical problem.

We extended our investigation to the long-term propagation of the asteroid (99942) Apophis, which undergoes a very deep close encounter with the Earth in 2029. The propagation method that in double precision achieves the highest accuracy is given by the combination of `EDromo` and K–S with the `XRA15` solver. The former is employed to propagate the motion when the primary body is the Sun, the latter when it is the Earth. In our simplified model, this method leads to a position error of 100 m after 100 yr. It is notable that the best propagation with Cowell's method is two orders of magnitude less accurate with a computational cost that is 3.6 and 2.2 times higher, depending on the solver. Also, the combination of `EDromo` and K–S with `XRA15` is four orders of magnitude more accurate than the hybrid method implemented in the `MERCURY` package, with a computational cost four times smaller. We assess that the latter method is better suited for qualitative studies, in which the time spans of integration may be several orders of magnitude larger than the ones considered.

We stress that adaptive solvers represent the best option to integrate Newtonian equations in strongly perturbed problems. The advantages of regularizations with respect to Cowell's formulation are much greater if fixed step size, fixed order numerical solvers are used.

Applications beyond the problem of asteroid propagations in the Solar System (e.g. planetary formation) are possible. However, additional work will be needed to address possible limitations of the method under extreme conditions. For instance, in the case of very fast encounters due to retrograde objects, one may need to consider a more sophisticated switch criterion. For a forthcoming work, we are investigating the possibility of adopting a smoother criterion based on the estimated local truncation error, rather than on the position. Finally, we expect that integrating regularized formulations with the Bulirsch–Stoer solver would achieve an outstanding numerical performance due to its efficiency. This assertion has to be verified through additional numerical tests.

ACKNOWLEDGEMENTS

This work is funded by the European Commission's Framework Programme 7, through the Stardust Marie Curie Initial Training Network, FP7-PEOPLE-2012-ITN, Grant Agreement 317185.

The authors are grateful to the anonymous reviewer for the constructive comments and suggestions. D.A. thanks Aaron J. Rosenbren for the inspiring discussions about encounters and his practical support, and Francesca Guerra for providing valuable information on long-term asteroid test cases. D.A. also extends his sincere thanks to Angelo Graziosi for the permission to use his original code as a basis for `XRA15`, and for his useful assistance on Everhart–Radau solvers.

REFERENCES

- Araujo R. A. N., Winter O. C., Prado A. F. B. A., Vieira Martins R., 2008, *MNRAS*, 391, 675
- Battin R., 1999, *An Introduction to the Mathematics and Methods of Astrodynamics*. American Institute of Aeronautics and Astronautics, Reston, VA
- Baù G., Bombardelli C., 2014, *AJ*, 148, 43
- Baù G., Bombardelli C., Peláez J., 2013, *Celest. Mech. Dyn. Astron.*, 116, 53
- Baù G., Bombardelli C., Peláez J., Lorenzini E., 2015, *MNRAS*, 454, 2890
- Baù G., Amato D., Bombardelli C., Milani A., 2016, in *Proceedings of the 6th International Conference on Astrodynamics Tools and Techniques (ICATT)*. Darmstadt, Germany
- Bond V. R., Allman M. C., 1996, *Modern Astrodynamics: Fundamentals and Perturbation Methods*. Princeton Univ. Press, Princeton, NJ
- Brouwer D., 1937, *AJ*, 46, 149
- Burdet C. A., 1969, *J. Reine Angew. Math.*, 238, 71
- Chambers J. E., 1999, *MNRAS*, 304, 793
- Dahlquist G., Björck Å., 1974, *Numerical Methods*. Prentice-Hall, Englewood Cliffs, NJ
- Deprit A., Elipe A., Ferrer S., 1994, *Celest. Mech. Dyn. Astron.*, 58, 151
- Devaney R. L., 2003, *An Introduction to Chaotic Dynamical Systems*. Westview Press, Boulder, CO
- Duncan M., Levison H., Lee M., 1998, *AJ*, 116, 2067
- Everhart E., 1969, *AJ*, 74, 735
- Everhart E., 1974, *Celest. Mech.*, 10, 35
- Everhart E., 1985, in Carusi A., Valsecchi G. B., eds, *Proc. IAU Colloq. Vol. 83, Dynamics of Comets: Their Origin and Evolution*. Reidel, Dordrecht, p.185
- Farnocchia D., Chesley S. R., Chodas P. W., Micheli M., Tholen D. J., Milani A., Elliott G. T., Bernardi F., 2013, *Icarus*, 224, 192
- Folkner W. M., Williams J. G., Boggs D. H., Park R. S., Kuchynka P., 2014, *Interplanet. Network Progress Rep.*, 196, 1
- Gonçalves Ferrari G., Boekholt T., Portegies Zwart S. F., 2014, *MNRAS*, 440, 719
- Hairer E., McLachlan R., Razakarivony A., 2008, *BIT Numer. Math.*, 48, 231
- Hernandez D. M., 2016, *MNRAS*, 458, 4285
- Hernandez D. M., Bertschinger E., 2015, *MNRAS*, 452, 1934
- Hiebert K., Shampine L., 1980, *Technical Report SAND80-0180*. Sandia Laboratories, Albuquerque, NM
- Kemle S., 2006, *Interplanetary Mission Analysis and Design*. Springer, Berlin
- Kvaerno A., Leimkuhler B., 2000, *SIAM J. Sci. Comput.*, 22, 1016
- Mikkola S., 1997, *Celest. Mech. Dyn. Astron.*, 68, 249
- Milani A., Chesley S. R., Sansaturio M. E., Tommei G., Valsecchi G. B., 2005, *Icarus*, 173, 362
- Öpik E. J., 1976, *Interplanetary Encounters: Close-Range Gravitational Interactions*. Elsevier, Amsterdam

- Peláez J., Hedo J. M., Rodríguez de Andrés P., 2007, *Celest. Mech. Dyn. Astron.*, 97, 131
- Petzold L., 1983, *SIAM J. Sci. Stat. Comput.*, 4, 136
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, *Numerical Recipes in FORTRAN. The Art of Scientific Computing*. Cambridge Univ. Press, Cambridge
- Radhakrishnan K., Hindmarsh A. C., 1993, Technical Report UCRL-ID-113855. Lawrence Livermore National Laboratory
- Rein H., Spiegel D. S., 2015, *MNRAS*, 446, 1424
- Roa J., Peláez J., 2015, *Celest. Mech. Dyn. Astron.*, 123, 13
- Stiefel E., Scheifele G., 1971, *Linear and Regular Celestial Mechanics*. Springer-Verlag, Berlin
- Szebehely V., 1967, *Theory of Orbits. The Restricted Problem of Three Bodies*. Academic Press, New York
- Tancredi G., 1998, *Celest. Mech. Dyn. Astron.*, 70, 181
- Tsang D., Galley C. R., Stein L. C., Turner A., 2015, *ApJ*, 809, L9
- Valsecchi G. B., Milani A., Gronchi G. F., Chesley S. R., 2003, *A&A*, 408, 1179
- Wisdom J., 2017, *MNRAS*, 464, 2350
- Wisdom J., Hernandez D. M., 2015, *MNRAS*, 453, 3015

APPENDIX A: OUTPUT AT PRESCRIBED VALUES OF PHYSICAL TIME

We present an algorithm to find the value of the independent variable that corresponds to a given physical time t^* . Since t increases with s (see equation 2), the function

$$g(s) = t(s) - t^* \quad (\text{A1})$$

has only one root s^* . Let the j -th ($j \geq 1$) integration step be the range of values of $s \in [s_{j-1}, s_j]$. During the propagation we check the condition

$$g(s_j)g(s_{j-1}) \leq 0. \quad (\text{A2})$$

If A2 is verified for $j = i$ and $g(s_i), g(s_{i-1}) \neq 0$, we search for a root of $g(s)$ in the open interval $\Sigma = (s_{i-1}, s_i)$. For this purpose, we find successively better approximations τ_n to s^* by Newton's method:

$$\tau_{n+1} = \tau_n - \frac{t(\tau_n) - t^*}{t'(\tau_n)}, \quad (\text{A3})$$

until the value of $|g(\tau_{n+1})|$ is sufficiently small. Note that t' is the derivative of t with respect to s and its expression is given in equation (2). The initial guess τ_0 is chosen through the following procedure. Since the integration step is much shorter than the orbital period, we generically have that the second derivative $t''(s)$ changes sign at most once in Σ . If $t''(s)$ is strictly monotonic, we directly set

$$\begin{aligned} \tau_0 &= s_i \text{ for } t''(s) > 0, \\ \tau_0 &= s_{i-1} \text{ for } t''(s) < 0. \end{aligned} \quad (\text{A4})$$

Assume that the function $t(s)$ has an inflection point at $s_f \in \Sigma$ (i.e. $t''(s_f)t''(s_{i-1}) < 0$). We divide Σ in the two subintervals $\Sigma_\ell = (s_{i-1}, s_1]$, $\Sigma_r = [s_1, s_i)$, where

$$s_1 = \frac{s_i + s_{i-1}}{2}. \quad (\text{A5})$$

After computing $g(s_1), t''(s_1)$, we know in which subinterval s^* and s_f are located. If it is the same one for both of them, the bisection is applied again to such subinterval by introducing s_2 . This operation is carried on until s_k ($k \geq 1$) is between s^* and s_f . Then, we choose

Table A1. Expressions of $t'(s)$ and $t''(s)$ as functions of the state vector and the independent variable s for the formulations used in this work (see Section 2). Prime denotes differentiation with respect to s .

	$t'(s)$	$t''(s)$
K-S	$\ \mathbf{u}\ ^2$	$2(\mathbf{u} \cdot \mathbf{u}')$
EDromo	$\lambda_3^{3/2} \rho$	$\frac{3}{2} \sqrt{\lambda_3} \rho \lambda_3' + \lambda_3^{3/2} (\zeta - \lambda_1' \cos s - \lambda_2' \sin s)$

τ_0 as follows:

$$\begin{aligned} \tau_0 &= s_i \text{ if } s^* > s_k \text{ and } t''(s_i) > 0, \\ \tau_0 &= s_{i-1} \text{ if } s^* < s_k \text{ and } t''(s_{i-1}) < 0, \\ \tau_0 &= s_k \text{ if } s^* > s_k \text{ and } t''(s_i) < 0, \text{ or } s^* < s_k \text{ and } t''(s_{i-1}) > 0. \end{aligned} \quad (\text{A6})$$

If $s_f = s^*$, the bisection method itself allows us to find s^* after 52 iterations working in double precision. Finally, the case in which $t''(s_i)t''(s_{i-1}) = 0$ is handled as in (A4).

The algorithm described in this section has been used when integrating the K-S and EDromo equations with Everhart's `RADAU` numerical scheme. Note that we need to evaluate the state variable $t(s)$, and its derivatives $t'(s)$ and $t''(s)$, which take the expressions reported in Table A1, at several points $\tilde{s} \in (s_{i-1}, s_i)$. Since Everhart's scheme is single step this can be accomplished in a straightforward and accurate way by propagating the state vector from s_{i-1} to \tilde{s} , with just a small overhead in terms of the total number of function evaluations.

APPENDIX B: INITIAL STEP SIZE SELECTION

In both solvers, the size of the first step of a propagation can be assigned by the user. In general, it will be either reduced or enlarged according to internal checks on the local truncation error and on the numerical stability. Note that when trajectory splitting is applied there are three propagations in the same simulation, each one with its own initial step size. Its value must be chosen by balancing two factors. On one hand, since the ratio between the lengths of two subsequent steps is bounded, it cannot be too small, as to avoid performing a large number of steps until it stabilizes around a suitable value. On the other hand, if it is too large the solvers will reject many integration steps until it becomes small enough to satisfy the local truncation error check.

We estimate the initial step size h_0 by means of an empirical formula that takes into account the characteristics of the motion and the accuracy requested by the integration. It is computed as

$$h_0 = -\frac{S}{n \log_{10} \epsilon}, \quad (\text{B1})$$

where ϵ is the accuracy parameter and n, S are found as follows. If the osculating orbit at the initial time of the propagation is closed we set $n = 100$, and S equal to the change in the independent variable s that corresponds to one revolution. If it is open, $n = 600$ and S is the increase in s that we have after covering the arc of the orbit contained in a sphere of radius 0.1 au and centred in the primary body.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.