

# High-order regularised symplectic integrator for collisional planetary systems

Antoine C. Petit, Jacques Laskar, Gwenaël Boué, and Mickaël Gastineau

ASD/IMCCE, CNRS-UMR8028, Observatoire de Paris, PSL University, Sorbonne Université, 77 avenue Denfert-Rochereau, 75014 Paris, France  
e-mail: [antoine.petit@obspm.fr](mailto:antoine.petit@obspm.fr)

Received 26 April 2019 / Accepted 15 June 2019

## ABSTRACT

We present a new mixed variable symplectic (MVS) integrator for planetary systems that fully resolves close encounters. The method is based on a time regularisation that allows keeping the stability properties of the symplectic integrators while also reducing the effective step size when two planets encounter. We used a high-order MVS scheme so that it was possible to integrate with large time-steps far away from close encounters. We show that this algorithm is able to resolve almost exact collisions (i.e. with a mutual separation of a fraction of the physical radius) while using the same time-step as in a weakly perturbed problem such as the solar system. We demonstrate the long-term behaviour in systems of six super-Earths that experience strong scattering for 50 kyr. We compare our algorithm to hybrid methods such as MERCURY and show that for an equivalent cost, we obtain better energy conservation.

**Key words.** celestial mechanics – planets and satellites: dynamical evolution and stability – methods: numerical

## 1. Introduction

Precise long-term integration of planetary systems is still a challenge today. The numerical simulations must resolve the motion of the planets along their orbits, but the lifetime of a system is typically billions of years, resulting in computationally expensive simulations. In addition, because of the chaotic nature of planetary dynamics, statistical studies are often necessary, which require running multiple simulations with close initial conditions (Laskar & Gastineau 2009). This remark is particularly true for unstable systems that can experience strong planet scattering caused by close encounters.

There is therefore considerable interest in developing fast and accurate numerical integrators, and numerous integrators have been developed over the years to fulfill this task. For long-term integrations, the most commonly used are symplectic integrators. Symplectic schemes incorporate the symmetries of Hamiltonian systems, and as a result, usually conserve the energy and angular momentum better than non-symplectic integrators. In particular, the angular momentum is usually conserved up to a roundoff error in symplectic integrators.

Independently, Kinoshita et al. (1991) and Wisdom & Holman (1991) developed a class of integrators that are often called mixed variable symplectic (MVS) integrators. This method takes advantage of the hierarchy between the Keplerian motion of the planets around the central star and the perturbations induced by planet interactions. It is thus possible to make accurate integrations using relatively large time-steps.

The initial implementation of Wisdom & Holman (1991) is a low-order integration scheme that still necessitates small time-steps to reach machine precision. Improvements to the method have since been implemented. The first category is symplectic correctors (Wisdom et al. 1996; Wisdom 2006). They consist of a modification of the initial conditions to improve the scheme

accuracy. Because it is only necessary to apply them when an output is desired, they do not affect the performance of the integrator. This approach is for example used in WHFAST (Rein & Tamayo 2015). The other approach is to consider higher order schemes (McLachlan 1995b; Laskar & Robutel 2001; Blanes et al. 2013). High-order schemes permit a very good control of the numerical error by fully taking advantage of the hierarchical structure of the problem. This has been used with success to carry out high-precision long-term integrations of the solar system (Farrés et al. 2013).

The principal limitation of symplectic integrators is that they require a fixed time-step (Gladman et al. 1991). If the time-step is modified between each step, the integrator remains symplectic because each step is symplectic. However, the change in time-step introduces a possible secular energy drift that may reduce the interest of the method. As a consequence, classical symplectic integrators are not very adapted to treat the case of systems that experience occasional close encounters where very small time-steps are needed.

To resolve close encounters, Duncan et al. (1998) and Chambers (1999) provided solutions in the form of hybrid symplectic integrators. Duncan et al. (1998) developed a multiple time-step symplectic integrator, SYMBA, where the smallest time-steps are only used when a close encounter occurs. The method is limited to an order two scheme, however. The hybrid integrator MERCURY (Chambers 1999) moves terms from the perturbation step to the Keplerian step when an interaction between planets becomes too large. The Keplerian step is no longer integrable but can be solved at numerical precision using a non-symplectic scheme such as Burlisch–Stoer or Gauss–Radau. However, the switch of numerical method leads to a linear energy drift (Rein & Tamayo 2015). Another solution to the integration of the collisional  $N$ -body problem has been proposed in (Hernandez & Bertschinger 2015; Hernandez 2016;

Dehnen & Hernandez 2017). It consists of a fixed time-step second-order symplectic integrator that treats every interaction between pairs of bodies as Keplerian steps.

Another way to build a symplectic integrator that correctly regularises close encounters is time renormalisation. Up to an extension of the phase space and a modification of the Hamiltonian, it is indeed always possible to modify the time that appears in the equations of motion. As a result, the real time becomes a variable to integrate. Providing some constraints on the renormalisation function, it is possible to integrate the motion with a fixed fictitious time-step using an arbitrary splitting scheme. Here we show that with adapted time renormalisation, it is possible to resolve close encounters accurately.

While time renormalisation has not been applied in the context of close encounters of planets, it has been successful in the case of perturbed highly eccentric problems (Mikkola 1997; Mikkola & Tanikawa 1999; Preto & Tremaine 1999; Blanes & Iserles 2012), see Mikkola (2008) for a general review. We here adapt a time renormalisation proposed independently by Mikkola & Tanikawa (1999) and Preto & Tremaine (1999). We show that it is possible to use the perturbation energy to monitor close encounters in the context of systems with few planets with similar masses. We are able to define a MVS splitting that can be integrated with any high-order scheme.

We start in Sect. 2 by briefly recalling the basics of the symplectic integrator formalism. In Sect. 3 we present the time renormalisation that regularises close encounters, and we then discuss the consequence of the renormalisation on the hierarchical structure of the equations (Sect. 4). In Sect. 5 we numerically demonstrate over short-term integrations the behaviour of the integrator at close encounter. We then explain (Sect. 6) how our time regularisation can be combined with the perihelion regularisation proposed by Mikkola (1997). Finally, we show the results of long-term integration of six planet systems in the context of strong planet scattering (Sect. 7) and compare our method to a recent implementation of MERCURY described in Rein et al. (2019), to SYMBA, and to the non-symplectic high-order integrator IAS15 (Sect. 8).

## 2. Splitting symplectic integrators

We consider a Hamiltonian  $H(\mathbf{p}, \mathbf{q})$  that can be written as a sum of two integrable Hamiltonians,

$$H(\mathbf{p}, \mathbf{q}) = H_0(\mathbf{p}, \mathbf{q}) + H_1(\mathbf{p}, \mathbf{q}). \quad (1)$$

A classical example is given by  $H_0 = T(\mathbf{p})$  and  $H_1 = U(\mathbf{q})$ , where  $T(\mathbf{p})$  is the kinetic energy and  $U(\mathbf{q})$  is the potential energy. In planetary dynamics, we can split the system as  $H_0 = K(\mathbf{p}, \mathbf{q})$ , where  $K$  is the sum of the Kepler problems in Jacobi coordinates (e.g. Laskar 1990) and  $H_1 = H_{\text{inter}}(\mathbf{q})$  is the interaction between the planets.

Using the Lie formalism (e.g. Koseleff 1993; Laskar & Robutel 2001), the equation of motion can be written

$$\frac{dz}{dt} = \{H, z\} = L_H z, \quad (2)$$

where  $z = (\mathbf{p}, \mathbf{q})$ ,  $\{\cdot, \cdot\}$  is the Poisson bracket<sup>1</sup>, and we note  $L_f = \{f, \cdot\}$ , the Lie differential operator. The formal solution of Eq. (2) at time  $t = \tau + t_0$  from the initial condition  $z(t_0)$  is

$$z(\tau + t_0) = \exp(\tau L_H) z(t_0) = \sum_{k=0}^{+\infty} \frac{\tau^k}{k!} L_H^k z(t_0). \quad (3)$$

<sup>1</sup> We use the convention  $\{f, g\} = \sum_i \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i}$ .

In general, the operators  $L_{H_0}$  and  $L_{H_1}$  do not commute, so that  $\exp(\tau L_H) \neq \exp(\tau L_{H_0}) \exp(\tau L_{H_1})$ . However, using the Baker-Campbell-Hausdorff (BCH) formula, we can find the coefficients  $a_i$  and  $b_i$  such that

$$\exp(\tau(L_H + L_{H_{\text{err}}})) = \mathcal{S}(\tau) = \prod_{i=1}^N \exp(a_i \tau L_{H_0}) \exp(b_i \tau L_{H_1}), \quad (4)$$

where  $H_{\text{err}} = O(\tau^r)$  is an error Hamiltonian depending on  $H_0$ ,  $H_1$ ,  $\tau$  and the coefficients  $a_i$  and  $b_i$ .

Because  $H_0$  and  $H_1$  are integrable, we can explicitly compute the evolution of the coordinates  $z$  under the action of the maps  $\exp(\tau L_{H_0})$  and  $\exp(\tau L_{H_1})$ . The map  $\mathcal{S}(\tau)$  is symplectic because it is a composition of symplectic maps. Moreover,  $\mathcal{S}(\tau)$  exactly integrates the Hamiltonian  $H + H_{\text{err}}$ .

If there is a hierarchy in the Hamiltonian  $H$  in the sense that  $|H_1/H_0| \simeq \varepsilon \ll 1$ , we can choose the coefficients such that the error Hamiltonian is of the order of

$$\sum_{i=1}^n O(\tau^r \varepsilon^i), \quad (5)$$

(see McLachlan 1995b; Laskar & Robutel 2001; Blanes et al. 2013; Farrés et al. 2013). For small  $\varepsilon$  and  $\tau$ , the solution of  $H + H_{\text{err}}$  is very close to the solution of  $H$ . In particular, it is thought that the energy error of a symplectic scheme is bounded. Because  $H_{\text{err}}$  depends on  $\tau$ , a composition of steps  $\mathcal{S}(\tau)$  also has this property if the time-step is kept constant. Otherwise, the exact integrated dynamics changes at each step, leading to a secular drift of the energy error.

In planetary dynamics, we can split the Hamiltonian such that  $H_0$  is the sum of the Keplerian motions in Jacobi coordinates and  $H_1$  is the interaction Hamiltonian between planets, which only depends on positions and thus is integrable (e.g. Laskar 1990). This splitting naturally introduces a scale separation  $\varepsilon$  given by

$$\varepsilon = \frac{\sum_{k=1}^N m_k}{m_0} \quad (6)$$

where  $N$  is the number of planets,  $m_k$  is the mass of the  $k$ th planet, and  $m_0$  is the mass of the star. If the planets remain far from each other,  $H_1$  is always  $\varepsilon$  small with respect to  $H_0$ .

The perturbation term is of the order of  $\varepsilon/\Delta$ , where  $\Delta$  is the typical distance between the planets in units of a typical length of the system. During close encounters,  $\Delta$  can become very small, and the step size needs to be adapted to  $\varepsilon/\Delta_{\text{min}}$ . Here  $\Delta_{\text{min}}$  is the smallest expected separation between planets normalised by a typical length of the system.

## 3. Regularised Hamiltonian

### 3.1. General expression

In order to construct an adaptive symplectic scheme that regularises the collisions, we extend the phase space and integrate the system with a fictitious time. Let  $s$  be such that

$$dt = g(\mathbf{p}, p_t, \mathbf{q}) ds, \quad (7)$$

where  $g$  is a function to be determined and  $p_t$  is the conjugated momentum to the real time  $t$  in the extended phase space. In order to have an invertible function  $t(s)$ , we require  $g$  to be positive. We consider the new Hamiltonian  $\Gamma$  defined as

$$\Gamma(\mathbf{p}, p_t, \mathbf{q}, t) = g(\mathbf{p}, p_t, \mathbf{q}) (H(\mathbf{p}, \mathbf{q}) + p_t). \quad (8)$$

$\Gamma$  does not depend on  $t$ , therefore  $p_t$  is a constant of motion. The equations of motion of this Hamiltonian are

$$\frac{dt}{ds} = \{\Gamma, t\} = g(\mathbf{p}, p_t, \mathbf{q}) + \frac{\partial g}{\partial p_t}(\mathbf{p}, p_t, \mathbf{q})(H(\mathbf{p}, \mathbf{q}) + p_t), \quad (9)$$

and for all function  $f(z)$

$$\frac{df(z)}{ds} = \{\Gamma, f(z)\} = g(\mathbf{p}, p_t, \mathbf{q})\{H, f\} + (H + p_t)\{g, f(z)\}. \quad (10)$$

In general,  $H$  is not a constant of motion of  $\Gamma$ . We have

$$\frac{dH}{ds} = \{\Gamma, H\} = (H + p_t)\{g, H\}. \quad (11)$$

If we choose initial conditions  $z_0$  such that  $p_t = -H(z_0)$ , we have  $\Gamma|_{t=0} = 0$ . Because  $\Gamma$  is constant and  $g$  is positive, we deduce from Eq. (3) that we have at all times

$$H + p_t = 0. \quad (12)$$

Because  $p_t$  is also a constant of motion,  $H$  is constant for all times. We can simplify the Eqs. of motion (9) and (10) into

$$\begin{aligned} \frac{dt}{ds} &= g(\mathbf{p}, p_t, \mathbf{q}) \\ \frac{df}{ds}(z) &= g(\mathbf{p}, p_t, \mathbf{q})\{H, f(z)\}. \end{aligned} \quad (13)$$

On the manifold  $p_t = -H(t_0)$ , Eq. (13) describe the same motion as Eq. (2). We call them the regularised equations.

We now wish to write  $\Gamma$  as a sum of two integrable Hamiltonians such as in Sect. 2. Based on previous works (Preto & Tremaine 1999; Mikkola & Tanikawa 1999; Blanes & Iserles 2012), we write

$$H + p_t = (H_0 + p_t) - (-H_1), \quad (14)$$

for  $H = H_0 + H_1$  and we define  $g$  as

$$g(\mathbf{p}, p_t, \mathbf{q}) = \frac{f(H_0 + p_t) - f(-H_1)}{H_0 + p_t + H_1}, \quad (15)$$

where  $f$  is a smooth function to be determined.  $g$  is the difference quotient of  $f$  and is well defined when  $H_0 + p_t + H_1 \rightarrow 0$ . We have

$$g(\mathbf{p}, p_t, \mathbf{q})|_{H_0+p_t=0} = f'(H_0 + p_t) = f'(-H_1). \quad (16)$$

With this choice of  $g$ , the Hamiltonian  $\Gamma$  becomes

$$\Gamma = f(H_0 + p_t) - f(-H_1) = \Gamma_0 + \Gamma_1, \quad (17)$$

where we note  $\Gamma_0 = f(H_0 + p_t)$  and  $\Gamma_1 = -f(-H_1)$ . We remark that  $\Gamma_0$  (resp.  $\Gamma_1$ ) is integrable because it is a function of  $H_0 + p_t$  (resp.  $H_1$ ), which is integrable. Moreover, we have

$$\begin{aligned} L_{\Gamma_0} &= f'(H_0 + p_t)L_{H_0+p_t}, \\ L_{\Gamma_1} &= f'(-H_1)L_{H_1}. \end{aligned} \quad (18)$$

Because  $H_0 + p_t$  (resp.  $H_1$ ) is a first integral of  $\Gamma_0$  (resp.  $\Gamma_1$ ), we have

$$\begin{aligned} \exp(\sigma L_{\Gamma_0}) &= \exp(\sigma f'(H_0 + p_t)L_{H_0+p_t}) = \exp(\tau_0 L_{H_0+p_t}), \\ \exp(\sigma L_{\Gamma_1}) &= \exp(\sigma f'(-H_1)L_{H_1}) = \exp(\tau_1 L_{H_1}), \end{aligned} \quad (19)$$

where

$$\tau_0 = \sigma f'(H_0 + p_t) \text{ and } \tau_1 = \sigma f'(-H_1). \quad (20)$$

The operator  $\exp(\sigma L_{\Gamma_0})$  ( $\exp(\sigma L_{\Gamma_1})$ ) is equivalent to the regular operator  $\exp(\tau_0 L_{H_0+p_t})$  ( $\exp(\tau_1 L_{H_1})$ ) with a modified time step. The operator  $\exp(\sigma L_{\Gamma})$  can be approximated by a composition of operators  $\exp(\sigma L_{\Gamma_k})$

$$\mathcal{S}_{\Gamma}(\sigma) = \prod_{i=1}^N \exp(a_i \sigma L_{\Gamma_0}) \exp(b_i \sigma L_{\Gamma_1}). \quad (21)$$

With the BCH formula,  $\mathcal{S}_{\Gamma}(\sigma) = \exp(\sigma(L_{\Gamma} + L_{\Gamma_{\text{err}}}))$ , where  $\Gamma_{\text{err}}$  is an error Hamiltonian that depends on  $\sigma$ . The symplectic map  $\mathcal{S}_{\Gamma}(\sigma)$  exactly integrates the modified Hamiltonian  $\Gamma + \Gamma_{\text{err}}$ . The iteration of  $\mathcal{S}_{\Gamma}(\sigma)$  with fixed  $\sigma$  is a symplectic integrator algorithm for  $\Gamma$ .

When the timescale  $\sigma$  is small enough, the numerical values of  $H_0$  and  $H_1$  do not change significantly between each step of the composition. We have  $\mathcal{S}_{\Gamma}(\sigma) \simeq \mathcal{S}(\tau)$  with  $\tau \simeq \sigma f'(-H_1)$ . In other words,  $\mathcal{S}_{\Gamma}$  behaves as  $\mathcal{S}$  with an adaptive time-step while keeping the bounded energy properties of a fixed time-step integrator.

### 3.2. Choice of the regularisation function

We wish the step sizes (20) to become smaller when planets experience close encounters. These time-steps are determined by the derivative of  $f$ . For nearly Keplerian systems, Mikkola & Tanikawa (1999) and Preto & Tremaine (1999) studied renormalisation functions such that  $f'(x) \propto x^{-\gamma}$ , where  $\gamma > 0$  (this corresponds to power-law functions and the important case of  $f = \ln$ ).

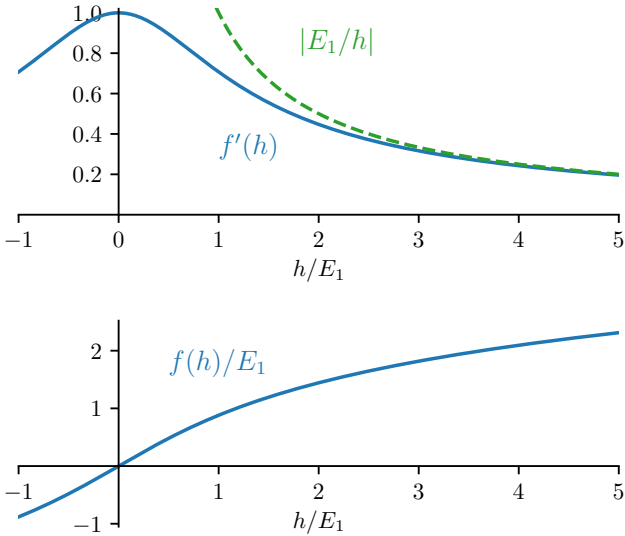
However, these authors considered splitting of the type  $H_0 = T(\mathbf{p})$  and  $H_1 = U(\mathbf{q})$ . As pointed out in Blanes & Iserles (2012), when the Hamiltonian is split as the Keplerian part plus an integrable perturbation, it appears that both terms  $K(\mathbf{p}, \mathbf{q}) + p_t$  and  $-H_1$  can change signs, resulting in large errors in the integration.

We remark that the use of  $f = \ln$  gives the best result when two planets experience a close encounter, but it leads to large energy errors far away for collision when  $H_1$  is nearly 0. Based on these considerations, we require  $f$  to verify several properties to successfully regularise the perturbed Keplerian problem in presence of close encounters:

- $f'$  should only depend on the magnitude of the perturbation  $H_1$ , therefore we require  $f'$  even (and  $f$  odd).
- As already pointed out,  $t$  must be an increasing function of  $s$ , therefore  $f' > 0$ .
- $f'$  should be smooth, therefore we exclude piecewise renormalisation functions.
- We require that the regularisation vanishes in absence of perturbation i.e.  $f'(0) = 1$ . It results that for vanishing perturbation,  $\sigma = \tau$ .
- Let  $E_1$  be a typical value of  $H_1$  far from close encounters that we determine an expression for below.  $f'$  should only depend on  $H_1/E_1$  in order to only track relative changes in the magnitude of the perturbation.
- For high values of the perturbation (e.g. during close encounters), we require  $f'(H_1) \sim |E_1/H_1|$  such that we preserve the good properties pointed out by previous studies (Preto & Tremaine 1999; Mikkola & Tanikawa 1999).
- To reduce the computational cost of the integrator,  $f'$  has to be numerically inexpensive to evaluate.

These properties lead to a very natural choice for  $f'$  (see in Fig. 1). We choose

$$f'(h) = \frac{1}{\sqrt{1 + \left(\frac{h}{E_1}\right)^2}}. \quad (22)$$



**Fig. 1.** Upper panel:  $f'$ , Eq. (22) as a function of  $h/E_1$ . The asymptotic value for  $h/E_1 \rightarrow +\infty$  is given in green. Lower panel:  $f/E_1$ , Eq. (23) as a function of  $h/E_1$ .

Taking the odd primitive of Eq. (22), we find

$$f(h) = E_1 \operatorname{arcsinh}\left(\frac{h}{E_1}\right). \quad (23)$$

For now on,  $f$  always refers to definition (23). With this choice of the function  $f$ , the Hamiltonian  $\Gamma$  takes the form

$$\Gamma = E_1 \operatorname{arcsinh}\left(\frac{H_0 + p_t}{E_1}\right) + E_1 \operatorname{arcsinh}\left(\frac{H_1}{E_1}\right), \quad (24)$$

where we used the oddity of  $f$ .

We need to define  $E_1$  more explicitly. When planets are far from each other, their mutual distance is of the same order as the typical distance between the planets and the star. Using the same idea as in Marchal & Bozis (1982); Petit et al. (2018), we define a typical length unit of the system based on the initial system energy  $E_0$ . We have

$$a_{\text{typical}} = -\frac{\mathcal{G}M^*}{2E_0}, \quad (25)$$

where  $M^* = \sum_{0 \leq i < j} m_i m_j$ . The typical value for the perturbation Hamiltonian far away from collision can be defined as

$$E_1 = \frac{\mathcal{G}m^*}{a_{\text{typical}}} = \frac{2|E_0|m^*}{M^*}, \quad (26)$$

where  $m^* = \sum_{1 \leq i < j} m_i m_j$ . We note that we have  $E_1/E_0 = \mathcal{O}(\varepsilon)$ .

The behaviour of the higher order derivative of  $f$  is useful for the error analysis, and in particular, their dependence in  $\varepsilon$ . The  $k$ th derivative of  $f$  has for expression

$$f^{(k)}(h) = E_1^{1-k} \operatorname{arcsinh}^{(k)}\left(\frac{h}{E_1}\right) = \mathcal{O}(\varepsilon^{1-k}). \quad (27)$$

## 4. Order of the scheme

### 4.1. Analytical error estimates

As explained in Sect. 2, most of the planet dynamics simulations are made with a simple a second-order scheme such as the Wisdom–Holman leapfrog integrator (Wisdom & Holman 1991). It is indeed possible to take advantage of the hierarchy

between  $H_0$  and  $H_1$ . This can be done by adding symplectic correctors (Wisdom et al. 1996; Rein & Tamayo 2015) or by cancelling the term of the form  $\varepsilon \tau^k$  up to a certain order (Laskar & Robutel 2001).

Hierarchical order schemes such as  $\mathcal{SAB}\mathcal{A}(10, 6, 4)$  (Blanes et al. 2013; Farrés et al. 2013) behave effectively as a tenth-order integrator if  $\varepsilon$  is small enough. Cancelling only selected terms reduces the number of necessary steps of the scheme, which reduces the numerical error and improves the performances.

Unfortunately, this property cannot be used for the regularised Hamiltonian because  $\Gamma_0$  and  $\Gamma_1$  are almost equal in magnitude. Nevertheless, it should be noted that the equations of motion (13) and the Lie derivatives (18) keep their hierarchical structure. The Poisson bracket of  $\Gamma_0$  and  $\Gamma_1$  gives

$$\{\Gamma_0, \Gamma_1\} = f'(H_0 + p_t) f'(H_1) \{H_0, H_1\}. \quad (28)$$

Because  $f'$  does not depend directly on  $\varepsilon$  (by choice,  $f'$  only tracks relative variations of  $H_1$ ),  $\{\Gamma_0, \Gamma_1\}$  is of the order of  $\varepsilon$ . However, for higher order terms in  $\sigma$  in  $H_{\text{err}}$ , it is not possible to exploit the hierarchical structure in  $\varepsilon$ . We consider the terms of the order of  $\sigma^2$  in the error Hamiltonian for the integration of  $\Gamma$  using the leapfrog scheme. We have (e.g. Laskar & Robutel 2001)

$$\Gamma_{\text{err}} = \frac{\sigma^2}{12} \{\{\Gamma_0, \Gamma_1\}, \Gamma_0\} + \frac{\sigma^2}{24} \{\{\Gamma_0, \Gamma_1\}, \Gamma_1\} + \mathcal{O}(\sigma^4). \quad (29)$$

In order to show the dependence in  $\varepsilon$ , we develop the Poisson brackets in Eq. (29)

$$\begin{aligned} \Gamma_{\text{err}} = & \frac{\sigma^2}{12} f'(H_0 + p_t)^2 f'(H_1) \{H_0, H_1\} \{H_0, H_1\} \\ & - \frac{\sigma^2}{12} f'(H_0 + p_t)^2 f''(H_1) (\{H_0, H_1\})^2 \\ & + \frac{\sigma^2}{24} f'(H_1)^2 f'(H_0 + p_t) \{H_0, H_1\} \{H_1, H_1\} \\ & + \frac{\sigma^2}{24} f'(H_1)^2 f''(H_0 + p_t) (\{H_0, H_1\})^2 + \mathcal{O}(\sigma^4). \end{aligned} \quad (30)$$

The first and third terms only depend on  $f'$  and nested Poisson brackets of  $H_0$  and  $H_1$ . Their dependency on  $\varepsilon$  is thus determined by the Poisson brackets as in the fixed time-step case. The first term is of the order of  $\varepsilon \sigma^2$  and the third of the order of  $\varepsilon^2 \sigma^2$ . On the other hand, the second and last terms introduce the second derivative of  $f$  as well as a product of Poisson bracket of  $H_0$  and  $H_1$ . Equation (27) shows that they are of the order of  $\varepsilon \sigma^2$ .

In order to cancel all terms of the order of  $\varepsilon \sigma^2$ , it is necessary to cancel both terms by  $\sigma^2$  in Eq. (29). Thus, the strategy used in Blanes et al. (2013) does not provide a scheme with a hierarchical order because every Poisson bracket contributes with terms of the order of  $\varepsilon$  to the error Hamiltonian.

It is easy to extend the previous result to all orders in  $\sigma$ . We consider a generic error term of the form

$$\Gamma_{\text{gen}} = \sigma^{n-1} \{\{\Gamma_{k_0}, \Gamma_{k_1}\}, \dots, \Gamma_{k_n}\}, \quad (31)$$

where  $k_j$  is either 0 or 1 and  $k_0 = 0$  and  $k_1 = 1$ . The development of  $\Gamma_{\text{gen}}$  into Poisson brackets of  $H_0$  and  $H_1$  contains a term that has for expression

$$\sigma^{n-1} f^{(n_0)}(H_0 + p_t) f^{(n_1)}(H_0 + p_t) f_1^{n_0} f_0^{n_1} (\{H_0, H_1\})^{n-1}, \quad (32)$$

where  $n_j$  is the number of  $\Gamma_j$  in  $\Gamma_{\text{gen}}$ , and  $f'_1 = f'(H_1)$ . Because  $n = n_0 + n_1$ , we deduce from Eq. (27) that the term Eq. (32) is of

the order of  $\varepsilon\sigma^{n-1}$ . Thus, the effective order of the Hamiltonian is always  $\varepsilon\sigma^{r_{\min}}$ , where  $r_{\min}$  is the smallest exponent  $r_k$  in Eq. (5). The error is still linear in  $\varepsilon$ . Hence, it is still worth using the Keplerian splitting.

We use schemes that are not dependent on the hierarchy between  $H_0$  and  $H_1$ . McLachlan (1995a) provided an exhaustive list of the optimal methods for fourth-, sixth-, and eighth-order integrators. Among the schemes he presented, we select a sixth-order method consisting of a composition of  $n = 7$  leapfrog steps introduced by Yoshida (1990) and an eighth-order method that is a combination of  $n = 15$  leapfrog steps. A complete review of these schemes can be found in Hairer et al. (2006). The coefficients of the schemes are given in Appendix B.

In order to solve the Kepler step, we adopt the same approach as Mikkola (1997); Rein & Tamayo (2015). The details on this particular solution as well as other technical details are given in Appendix A.

#### 4.2. Non-integrable perturbation Hamiltonian

When the classical splitting of the Hamiltonian written in canonical heliocentric coordinates is used,  $H_1$  depends on both positions and momenta. We can write  $H_1$  as a sum of two integrable Hamiltonian  $H_1 = T_1 + U_1$ , where  $T_1$  is the indirect part that only depends on the momenta and  $U_1$  is the planet interaction potential, only depending on positions (Farrés et al. 2013). We thus approximate the evolution operator (19) by

$$\exp(\sigma L_{\Gamma_1}) = \exp\left(\frac{\tau_1}{2} L_{T_1}\right) \exp(\tau_1 L_{U_1}) \exp\left(\frac{\tau_1}{2} L_{T_1}\right) + O(\varepsilon^3 \tau_1^3). \quad (33)$$

The numerical results suggest that heliocentric coordinates give slightly more accurate results at constant cost. It is possible to use an alternative splitting that is often called democratic heliocentric splitting (Laskar 1990; Duncan et al. 1998). With this partition of the Hamiltonian, the kinetic and the potential part of the perturbation Hamiltonian commute. Therefore, the step  $\exp(\sigma L_{\Gamma_1})$  is directly integrable using the effective step size  $\tau_1$  Eq. (20). In the following numerical tests, we always use the classical definition when we refer to heliocentric coordinates.

### 5. Error analysis near the close encounter

#### 5.1. Time-step and scheme comparison

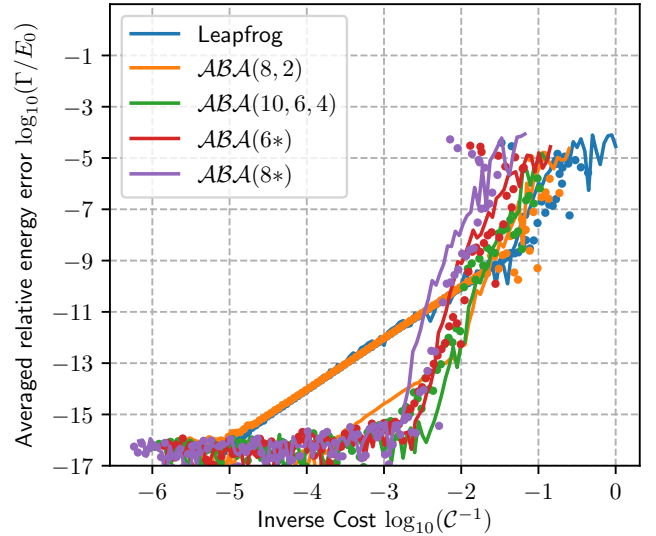
In this section, we test how a single close encounter affects the energy conservation. To do so, we compare different integration schemes for a two-planet system, initially on circular orbits, during an initial synodic period

$$T_{\text{syn}} = \frac{2\pi}{n_1 - n_2} = \frac{2\pi}{n_2(\alpha^{-3/2} - 1)}, \quad (34)$$

where  $n_i$  is the mean motion of planet  $i$  and  $\alpha$  is the ratio of the semi-major axis.

Because the time-step is renormalised, we need to introduce a cost function that depends on the fictitious time  $s$  as well as on the number of stages involved in the scheme of the integrator. We define the cost of an integrator as the number of evaluations of  $\exp(a\sigma L_{\Gamma_0}) \exp(b\sigma L_{\Gamma_1})$  that are required to integrate for a given real time period  $T_{\text{syn}}$

$$C = \frac{s_{\text{syn}} n}{T_{\text{syn}} \sigma}, \quad (35)$$



**Fig. 2.** Comparison between various schemes detailed in the body of the text for the integration of a single synodic period for a system of equal-mass planets initially on planar circular orbits with  $\varepsilon = 2m/m_0 = 10^{-5}$  and  $\alpha = 0.8$ . The solid lines represent the integrations using a fixed real time-step and the dots the adaptive time-steps. The closest planet approach is 0.19992 AU.

where  $s_{\text{syn}}$  is the fictitious time after  $T_{\text{syn}}$ ,  $\sigma$  is the fixed fictitious time-step, and  $n$  is the number of stages of the integrator. We also compare the renormalised integrators to the same scheme with fixed time-step. For a fixed time-step, the cost function is simply given by  $C_{\text{fixed}} = n/\tau$ , where  $\tau$  is the time-step (Farrés et al. 2013). We present different configurations in Figs. 2–4.

In the two first sets, we integrate the motion of two equal-mass planets on circular orbits, starting in opposition with respect to the star. In both simulations, we have  $\varepsilon = (m_1 + m_2)/m_0 = 10^{-5}$ , the stellar mass is  $1 M_{\odot}$ , and the outer planet semi-major axis is 1 AU. In both figures, we represent the averaged relative energy variation

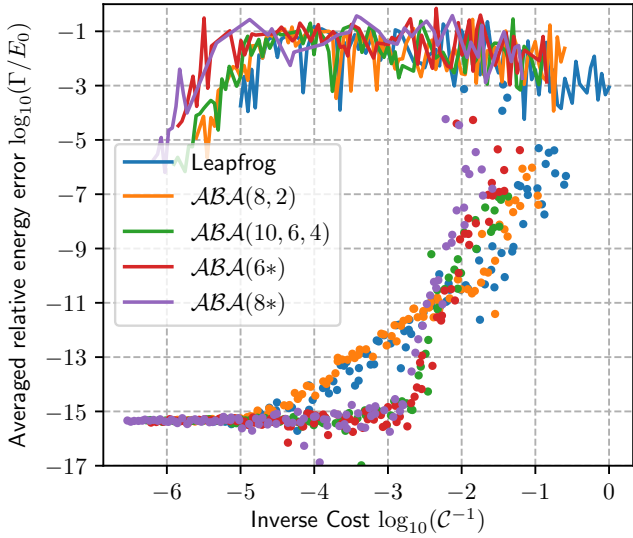
$$\frac{\Gamma}{E_0} \simeq f'(H_1) \frac{\Delta E}{E_0}. \quad (36)$$

as a function of the inverse cost  $C^{-1}$  of the integration for various schemes:

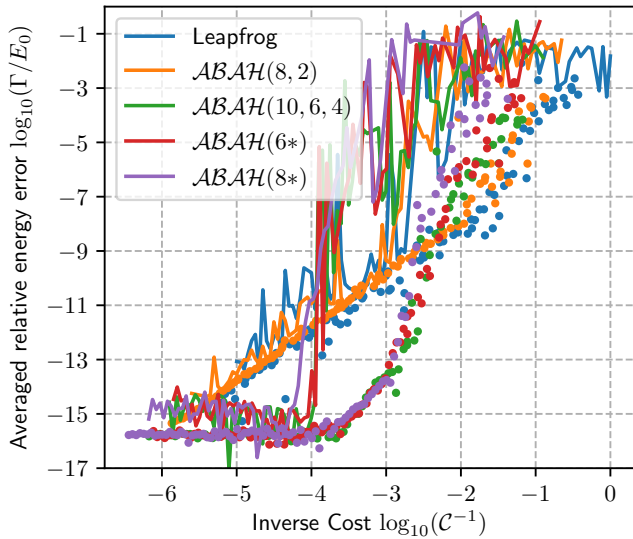
- the classical order 2 leapfrog  $\mathcal{ABA}(2, 2)$ ,
- the scheme  $\mathcal{ABA}(8, 2)$  from Laskar & Robutel (2001),
- the scheme  $\mathcal{ABA}(10, 6, 4)$  from Blanes et al. (2013),
- the sixth- and eighth-order schemes from McLachlan (1995a) that we introduced in the previous section.

For each scheme, we plot with a solid line the result of the fixed time-step algorithm and with dots the results of the adaptive time-step integrator. In the results presented in Figs. 2 and 3, the systems are integrated in Jacobi coordinates. In Fig. 4 we integrate in heliocentric coordinates. For the three cases, the integrations were also carried out with the other set of coordinates with almost no differences.

In Fig. 2 the initial semi-major axis ratio  $\alpha$  is 0.8 and the closest approach of the two planets is 0.19992 AU. In this case, a fixed time-step algorithm provides accurate results, and when a generalized order scheme such as  $\mathcal{ABA}(8, 2)$  or  $\mathcal{ABA}(10, 6, 4)$  is used, the performances are better with a fixed time-step than with an adaptive one. However, we see no sensitive differences in accuracy between the fixed and adaptive versions of the  $\mathcal{ABA}(6^*)$  and  $\mathcal{ABA}(8^*)$  schemes. For the most efficient



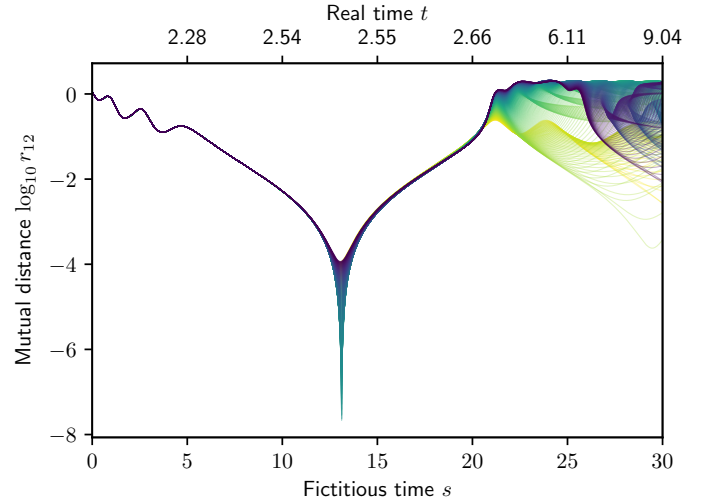
**Fig. 3.** Same as Fig. 2 for a system of two equal-mass planets initially on planar circular orbits with  $\varepsilon = 2m/m_0 = 10^{-5}$  and  $\alpha = 0.97$ . The closest planet approach is  $3.68 \times 10^{-5}$  AU.



**Fig. 4.** Same as Fig. 2 for a system of two equal-mass planets initially on planar circular orbits with  $\varepsilon = 2m/m_0 = 10^{-3}$  and  $\alpha = 0.9$ . The closest planet approach is  $1.7 \times 10^{-2}$  AU. This system is integrated in heliocentric coordinates to demonstrate that we obtain similar results to the case of Jacobi coordinates.

schemes, machine precision is reached for an inverse cost of a few  $10^{-3}$ . This case illustrates the point made in Sect. 4 that there is no advantage in using the hierarchical structure of the original equation in the choice of the scheme.

In Fig. 3 the initial semi-major axis ratio  $\alpha$  is 0.97 and the closest approach of the two planets is  $3.68 \times 10^{-5}$  AU. The radius of a planet of mass  $5 \times 10^{-6} M_\odot$  with a density of  $5 \text{ g cm}^{-3}$ , denser than Earth, is  $5.21 \times 10^{-4}$  AU, that is, ten times greater. We can compute the density that corresponds to a body of mass  $5 \times 10^{-6} M_\odot$  and radius half the closest approach distance. We obtain a density of  $1.1 \times 10^5 \text{ g cm}^{-3}$ , which is close to the white dwarf density. In this case, the fixed time-step algorithm is widely inaccurate while the performance of the adaptive time-step algorithm remains largely unchanged. This demonstrates how powerful this new approach is for the integration of systems with a few planets.



**Fig. 5.** Mutual distance of the planets as a function of the fictitious time for 1000 initial conditions chosen as described in the body of the text. On the upper axis, we give the real time for the initial condition that reached the closest approach.

We also give an example with higher mass planets in Fig. 4. We plot the relative energy error as a function of the inverse cost for a system of two planets of mass  $5 \times 10^{-4} M_\odot$  that correspond to the case of  $\varepsilon = 10^{-3}$ . The initial semi-major axis ratio is 0.9 to ensure a closest-approach distance of  $1.7 \times 10^{-2}$  AU.

For this particular close encounter, it is still possible to reach machine precision with a fixed time-step at the price of a very small time-step. On the other hand, the close encounter is perfectly resolved for the adaptive time-step, even if  $\varepsilon$  is larger by two orders of magnitude. Because of the higher masses, it is necessary to take smaller time-steps to ensure that the relative energy error remains at machine precision.

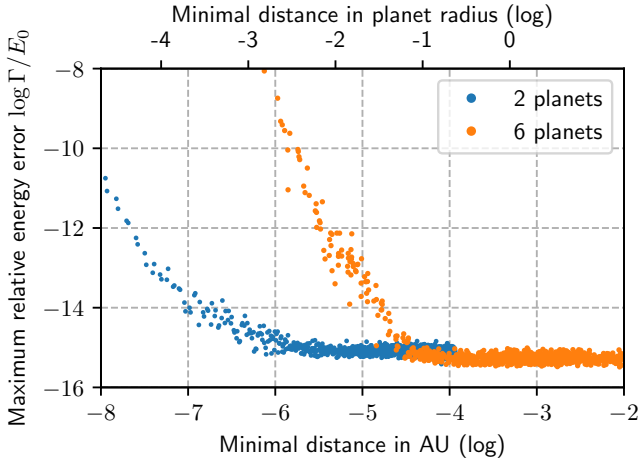
## 5.2. Behaviour at exact collision

To demonstrate the power of our algorithm in resolving close encounters, we considered the case of two planets that almost exactly collide (as if the bodies were material points). We take two planets of mass  $m = 10^{-5} M_\odot$ , the first on an orbit of eccentricity 0.1 and semi-major axis 0.95 AU, and the second on a circular orbit of semi-major axis 1 AU. We then fine-tuned the mean longitudes of the two planets to ensure that we approach the exact collision. We integrated 1000 initial conditions for which we linearly varied  $\lambda_1$  between  $-0.589$  and  $-0.587$ , and we took  $\lambda_2 = -1.582$ .

In Fig. 5 we represent for each initial condition the mutual distance of the planets as a function of the fictitious time  $s$ . In the upper axis, we give the real time of the particular initial condition that gave the closest approach. In this simulation, all initial conditions reached a mutual distance smaller than  $1.1 \times 10^{-4}$  AU and the closest approach is  $5.5 \times 10^{-9}$  AU. For the chosen masses, we can compute the physical radius of such bodies assuming a particular bulk density. We have

$$R_p = 5.21 \times 10^{-2} \left( \frac{\rho}{1 \text{ g cm}^{-3}} \right)^{-1/3} \left( \frac{m_p}{1 M_\odot} \right)^{1/3}, \quad (37)$$

where  $\rho$  is the bulk density of the planet. The radius is only a weak function of the density over the range of planet bulk densities. When we assume a very conservative value of  $\rho = 6 \text{ g cm}^{-3}$ , we obtain a planet radius of  $6.17 \times 10^{-4}$  AU, that is, well above the close encounters considered here.



**Fig. 6.** Relative energy error as a function of the closest mutual approach reached during a close encounter of super-Earths in a two-planet system and in a six-planet system. The particular initial conditions are described in the text. In the upper part, we give the mutual distance in units of planet radii for a planet of density  $\rho = 6 \text{ g cm}^{-3}$ .

We plot the energy error as a function of the closest approach distance in Fig. 6. We also considered a more practical case where four other planets were added to the system. The additional planets were located on circular orbits, had the same masses of  $10^{-5} M_{\odot}$  and were situated outside of the orbit of the second planet, with an equal spacing of 0.15 AU. Their initial angles are randomly drawn but were similar in every run.

In both set-ups, we used the heliocentric  $\mathcal{AB}\mathcal{A}(8^*)$  scheme, and the fictitious time-step was  $\sigma = 10^{-2}$ . In the case of the two-planet system, machine precision is reached even for an approach of approximately  $10^{-6}$  AU, which corresponds to a planet radius of about  $10^{-3}$ . When planets are added, it is more difficult to reach very close encounters because the other planets are perturbed. Nonetheless, approaches up to  $6 \times 10^{-7}$  AU are reached, and the close encounter is resolved at machine precision up to a closest distance of  $3 \times 10^{-5}$  AU or a planet radius of  $5 \times 10^{-2}$ .

## 6. Pericenter regularisation

When planet–planet scattering is studied, the closest distance between the planet and the star may be reduced. As a result, the fixed time-step becomes too large and the passage at pericenter is insufficiently resolved. In order to address this problem, we can combine the close-encounter regularisation of Sect. 3 with a regularisation method introduced by Mikkola (1997).

We first detail Mikkola’s idea. When integrating a few-body problem Hamiltonian, a time renormalisation  $d\tilde{t} = g(\mathbf{q})du$  can be introduced, where

$$g(\mathbf{q}) = \left( \sum_{i=1}^N \frac{A_i}{q_i} \right)^{-1}. \quad (38)$$

Here  $A_i$  are coefficients that can be chosen arbitrarily. In this article, we define them as

$$A_i = \frac{m_i a_{\text{typical}}}{\sum_{j=1}^N m_j}, \quad (39)$$

where  $a_{\text{typical}}$ , defined in Eq. (25), is the typical length scale of the system. The new Hamiltonian  $\Upsilon$  in the extended phase space  $(\mathbf{q}, t, \mathbf{p}, p_t)$  is

$$\Upsilon = \Upsilon_0 + \Upsilon_1 = g(\mathbf{q})(H_0 + p_t) + g(\mathbf{q})H_1. \quad (40)$$

As in Sect. 3, in the sub-manifold  $\{p_t = H(0)\}$ ,  $\Upsilon$  and  $H$  have the same equations of motion up to the time transformation. If  $H_1$  only depends on  $\mathbf{q}$  (using Jacobi coordinates, e.g.),  $\Upsilon_1$  is trivial to integrate.

### 6.1. Kepler step

It is also possible to integrate  $\Upsilon_0$  as a modified Kepler motion through the expression of  $g$  (Mikkola 1997). We denote  $v_0 = \Upsilon_0(0)$  and  $\tilde{H}_0 = g^{-1}(\mathbf{q})(\Upsilon_0 - v_0)$ .  $\tilde{H}_0$  has the same equations of motion as  $\Upsilon_0$  up to a time transformation  $d\tilde{t} = g^{-1}(\mathbf{q})du$ . We have

$$\begin{aligned} \tilde{H}_0 &= H_0 + p_t - v_0 \sum_{i=1}^N \frac{A_i}{q_i} \\ &= p_t + \sum_{i=1}^N \frac{\|\mathbf{p}_i\|^2}{2m_i} - \frac{\mu_i m_i + v_0 A_i}{q_i} = p_t + \sum_{i=1}^N \tilde{K}_i, \end{aligned} \quad (41)$$

where  $\tilde{K}_i$  is the Hamiltonian of a Keplerian motion of planet  $i$  with a modified central mass

$$\tilde{\mu}_i = \mu_i \left( 1 + \frac{v_0 A_i}{\mu_i m_i} \right) = \mu_i \left( 1 + \frac{v_0 a_{\text{typical}}}{\mu_i \sum_{j=1}^N m_j} \right). \quad (42)$$

However, the time equation must be solved as well. We integrate with a fixed fictitious time-step  $\Delta u$ . The time  $\Delta\tilde{t}(u)$  is related to  $u$  through the relation

$$\Delta u = \int_0^{\Delta\tilde{t}} g^{-1}(\mathbf{q}(\tilde{t}))d\tilde{t} = \sum_{i=1}^N A_i \int_0^{\Delta\tilde{t}} \frac{1}{r_i(\tilde{t})} d\tilde{t}, \quad (43)$$

where  $r_i(\tilde{t})$  follows a Keplerian motion. We can rewrite Eq. (43) with the Stumpff formulation of the Kepler equation (Mikkola 1997; Rein & Tamayo 2015, see Appendix A.1). Because  $\int_0^{\Delta\tilde{t}} \frac{1}{r_i(\tilde{t})} d\tilde{t} = X_i$ , we have

$$\Delta u = \sum_{i=1}^N A_i X_i. \quad (44)$$

As a consequence, the  $N$  Stumpff–Kepler equations

$$\Delta\tilde{t} = r_{0i} X_i + \eta_{0i} G_2(\beta_{0i}, X_i) + \zeta_{0i} G_3(\beta_{0i}, X_i) = \kappa_i(X_i), \quad (45)$$

must be solved simultaneously with Eq. (44). To do so, we used a multidimensional Newton–Raphson method on the system of  $N + 1$  equations consisting of the  $N$  Kepler Eqs. (45) and (44) of unknowns  $Y = (X_1, \dots, X_N, \Delta\tilde{t})$ . The algorithm is almost as efficient as the fixed-time Kepler evolution because it does not add up computation of Stumpff’s series. At step  $k$ , we can obtain  $Y^{(k+1)}$  through the equation

$$Y^{(k+1)} = Y^{(k)} - dF^{-1}(Y^{(k)})(F(Y^{(k)})), \quad (46)$$

where

$$F = \begin{pmatrix} \kappa_1(X_1) - \Delta\tilde{t} \\ \vdots \\ \kappa_N(X_N) - \Delta\tilde{t} \\ \sum_{i=1}^N A_i X_i - \Delta u \end{pmatrix}, \quad (47)$$

and

$$dF = \begin{pmatrix} \kappa'_1(X_1) & 0 & \cdots & -1 \\ 0 & \ddots & 0 & -1 \\ \cdots & 0 & \kappa'_N(X_N) & -1 \\ A_1 & \cdots & A_N & 0 \end{pmatrix}, \quad (48)$$

with  $\kappa'_i$  being the derivative with respect to  $X_i$  of  $\kappa_i$  (Eq. (45)).

Equation (46) can be rewritten as a two-step process where a new estimate for the time  $\Delta\tilde{t}^{(k+1)}$  is computed and is then used to estimate  $X_i^{(k+1)}$ . We have

$$\Delta\tilde{t}^{(k+1)} = \left( \Delta u + \sum_{i=1}^N \frac{A_i (\kappa_i(X_i^{(k)}) - \kappa'_i(X_i^{(k)})X_i^{(k)})}{\kappa'_i(X_i^{(k)})} \right) / \left( \sum_{i=1}^N \frac{A_i}{\kappa'_i} \right), \quad (49)$$

and

$$X_i^{(k+1)} = \frac{\Delta\tilde{t}^{(k+1)} + \kappa'_i(X_i^{(k)})X_i^{(k)} - \kappa_i(X_i^{(k)})}{\kappa'_i(X_i^{(k)})}. \quad (50)$$

## 6.2. Heliocentric coordinates

In heliocentric coordinates,  $H_1^H$  depends on  $\mathbf{p}$  as well. As a result,  $gH_1^H$  cannot be easily integrated, and it is not even possible to write it as a sum of integrable Hamiltonians. To circumvent this problem, we can split  $gH_1^H$  into  $gU_1^H$  and  $gT_1^H$ . The potential part  $gU_1^H$  is integrable, but a priori,  $gT_1^H$  is not integrable.

The integration of  $gT_1^H$  can be integrated using a logarithmic method as proposed in Blanes & Iserles (2012). The evolution of  $gT_1^H$  during a step  $\Delta t$  is the same as the evolution of  $\log gT_1^H = \log g + \log T_1^H$ , which is separable, for a step  $\mathcal{T}_1 \Delta t$  where  $\mathcal{T}_1 = gT_1^H|_{t=0}$ . Therefore, we can approximate  $\log gT_1^H$  using a leapfrog scheme, and the error is of order  $\Delta u^2 \varepsilon^3$  as well.

Then, we can approximate the heliocentric step using the same method as in Sect. 4.2. In this case, it is necessary to approximate the step even when democratic heliocentric coordinates are used because  $g(\mathbf{q})$  does not commute with  $T_1^H$ .

## 6.3. Combining the two regularisations

We showed that the Hamiltonian  $\Upsilon$  is separable into two parts that are integrable (or nearly integrable for heliocentric coordinates). Therefore we can simply regularise the close encounters by integrating the Hamiltonian

$$\tilde{\Gamma} = f(\Upsilon_0 + p_u) + f(\Upsilon_1) = f(g(\mathbf{q})(H_0 + p_t) + p_u) + f(g(\mathbf{q})H_1), \quad (51)$$

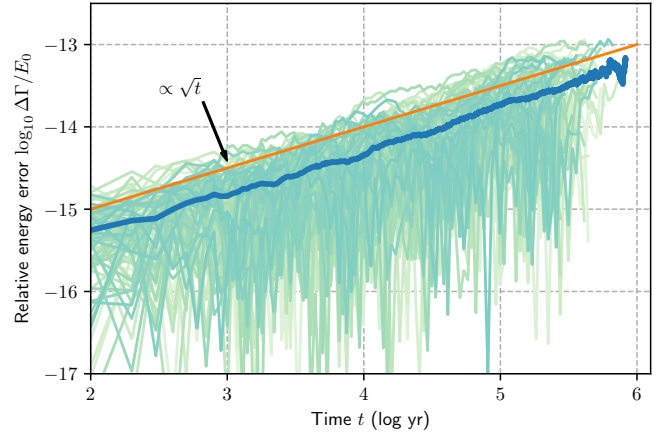
where  $p_u$  is the momentum associated with the intermediate time  $u$  used to integrate  $\Upsilon$  alone. The time equation is

$$\frac{dt}{ds} = \frac{\partial \tilde{\Gamma}}{\partial p_t} = f'(\Upsilon_0)g(\mathbf{q}). \quad (52)$$

We need to place ourselves on the sub-manifold such that  $\Upsilon_0 + p_u + \Upsilon_1 = 0$  and  $H_0 + p_t + H_1 = 0$ , in order to have the same equations of motion for  $\tilde{\Gamma}$ ,  $\Upsilon$  and  $H$ . Both of these conditions are fulfilled by choosing  $p_t = -E_0$  and  $p_u = 0$ .

## 7. Long-term integration performance

So far, we only presented the performance of the algorithm for very short integrations. In this section, we present the long-term behaviour of the integrator for systems with a very chaotic nature. We consider two different configurations, a system composed of equal planet masses on initially circular, coplanar and equally spaced orbits, used as a test model for stability analysis since the work of Chambers et al. (1996). The second is a similar system but with initially moderate eccentricities and inclinations.



**Fig. 7.** Evolution of the relative energy error for 100 systems composed of six planets on initially circular orbits. The light curves represent individual systems, and the thicker blue curves show the average over the 100 systems. The orange straight line is proportional to  $\sqrt{t}$  in order to show that the integrator follows Brouwer's law. The orange line is the same in all energy error figures for comparison.

### 7.1. Initially circular and coplanar systems

We integrated 100 systems of six initially coplanar and circular planets. The planet masses were taken to be equal to  $10^{-5} M_{\oplus}$ , and the semi-major axis of the outermost planet was fixed to 1 AU, while the semi-major axis ratios of the adjacent planets were all equal to 0.88. These values were chosen to ensure that the lifetime of the system was of the order of 300 kyr before the first collision. The fixed fictitious time-step was  $\sigma = 10^{-2}$  yr. Because of the renormalisation, this approximately corresponds to a fixed time-step of  $6.3 \times 10^{-3}$  yr in terms of computational cost. The initial period of the inner planet was of the order of 0.38 yr, that is, we have about 50 steps per orbit.

The simulations were stopped when two planet centres approached each other by less than half the planet radii, assuming a density of  $6 \text{ g cm}^{-3}$ . This stopping criterion is voluntarily nonphysical as it allows for a longer chaotic phase that leads to more close encounters. We also monitored any encounters with an approach closer than 2 Hill radii at 1 AU (0.054 AU) and recorded its time, the planets involved, and the closest distance between the two planets.

For the majority of the integrations, we observe moderate semi-major axis diffusion without close encounters. About 1 kyr before the final collision, the system enters a true scattering phase with numerous close encounters. The integrations lasted on average 353 kyr, the shortest was 129 kyr long and the longest 824 kyr. On average, we recorded 557 close encounters, and 68% occurred during the last 10 000 yr. On average, 6.4 approaches below 0.01 AU per system were recorded. Of these very close encounters, 95% occurred during the last 1000 yr.

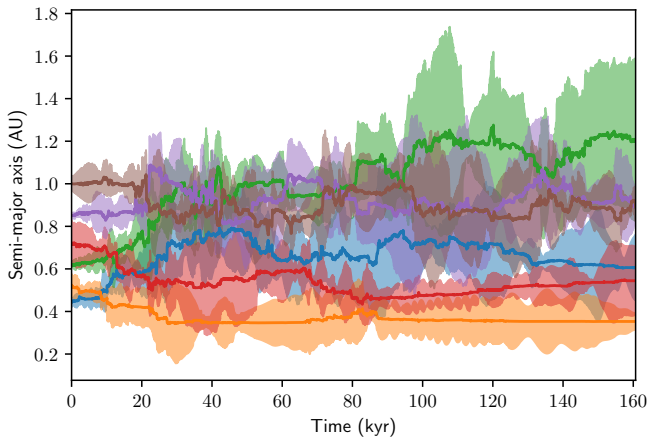
The relative energy error evolution is shown in Fig. 7. We observe that the integrator follows the Brouwer (1937) law because the energy error behaves as a random walk (the error grows as  $\sqrt{t}$ ). The scattering phase does not affect the energy conservation: no spikes of higher error values occur towards the end of the integrations.

### 7.2. Planet scattering on inclined and eccentric orbits

Our second long-term test problem was a set of planetary systems with moderate eccentricities and inclinations. The goal is here to test the integrator in a regime where strong scattering

**Table 1.** Summary of the main statistics and specifications of the spatial and eccentric runs.

Run	J1	J2	H1	H2
Coordinates	Jacobi	Jacobi	Heliocentric	Heliocentric
Pericenter	No	Yes	No	Yes
Equivalent fixed time-step	$2.6 \times 10^{-3} \text{ yr}^{-1}$	$2.6 \times 10^{-3} \text{ yr}^{-1}$	$2.9 \times 10^{-3} \text{ yr}^{-1}$	$2.9 \times 10^{-3} \text{ yr}^{-1}$
Time-step per shortest orbit	112	112	102	102
Average lifetime	53.3 kyr	53.0 kyr	52.6 kyr	46.3 kyr
Average number of close encounters	$1.26 \times 10^4$	$1.17 \times 10^4$	$1.20 \times 10^4$	$1.07 \times 10^4$
Average number of very close encounters	458	415	426	384


**Fig. 8.** Evolution of the semi-major axis, periapsis, and apoapsis for a typical initial condition. The bold curve is the semi-major axis, and the filled region represents the extent of the orbit. The mutual inclinations are not represented.

occurs on long timescales with a much higher frequency in very close encounters (shorter than 0.01 AU). In addition, this case helps to demonstrate the advantage of the pericenter regularisation. The initial conditions were chosen as follows:

- Similarly to the previous case, we considered systems of six planets with an equal mass of  $10^{-5} M_{\odot}$ .
- The semi-major axis ratio of adjacent pairs was 0.85.
- The Cartesian components of the eccentricities ( $e \cos(\varpi)$ ,  $e \sin(\varpi)$ ) and inclinations ( $i \cos(\Omega)$ ,  $i \sin(\Omega)$ ) were drawn from a centred Gaussian distribution with a standard deviation of 0.08. The average eccentricities and inclinations were about 0.11.
- The mean longitudes were chosen randomly.

We also kept the same stopping criterion and recorded the close encounters (smaller than 2 Hill radii).

The integrations were performed with two sets of coordinates, Jacobi and heliocentric; with and without pericenter regularisation. The main statistics from the runs are summarised in Table 1. For all integrations, we used the eighth-order scheme and a fictitious time-step  $\sigma = 4 \times 10^{-3}$ . The systems on average experience a close encounter every 5 yr. A typical evolution is presented in Fig. 8. In this particular run, we see numerous planet exchanges, almost 35 000 close encounters occur, and the system experiences more than 1200 very close encounters. Nevertheless, the final relative energy error is  $6.8 \times 10^{-15}$ .

We plot in Figs. 9a–d the relative energy errors of runs J1, J2, H1, and H2, respectively. We plot in light colours the relative energy error of individual systems and in thicker blue the median of the relative energy error. The choice to use the median was made because some initial conditions lead to a worse energy

conservation, which makes the average less informative. For the longer times, the median is no longer reliable because the majority of the simulations had already stopped.

Despite the extreme scattering that occurs, the energy is conserved in most systems within the numerical round-off error prescription. We also remark that the heliocentric coordinates seem to provide a more stable integrator even though the interaction step is approximated. In the two cases without the pericenter regularisation, J1 and H1, it seems that the energy error is no longer proportional to  $\sqrt{t}$  but linear in  $t$ .

As explained above, because of the high eccentricities, the innermost pericenter passage is not very well resolved, and therefore the step-to-step energy variation is no longer close to machine precision. This effect has been confirmed by the detailed analysis of the time-series close to energy spikes by restarts of the integration from a binary file. We observed that the energy did not change during the close encounter, but shortly after, when the inner planet passed closer to the star because of the scattering.

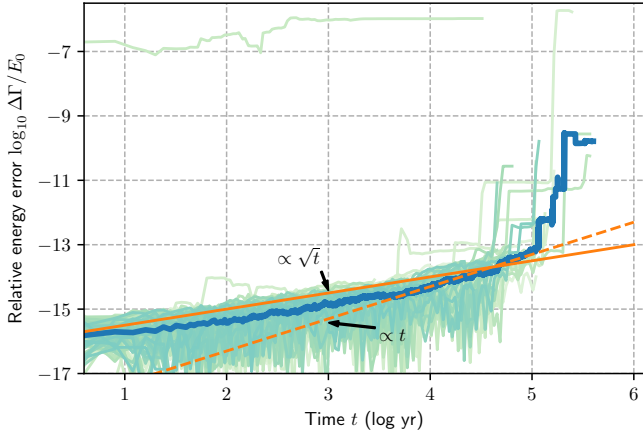
We also observe an initial condition well above the machine-precision level in all four runs. In this particular system, we have  $e_5(t=0) = 0.81$ , which gives an initial pericenter at 0.15 AU. In this extreme case, the pericenter regularisation leads to an improvement by two orders of magnitude of the error in Jacobi coordinates. The heliocentric coordinates remain more efficient, however, and we do not see an improvement in this particular case.

In all runs, the time distribution of the close encounters appears to be rather uniform. Moreover, the probability distribution function (PDF) of the closest approach during a close encounter is linear in the distance, as shown in Fig. 10. This proves that the systems are in a scattering regime where the impact parameters for the close encounters are largely random (see also Laskar et al. 2011, Fig. 5).

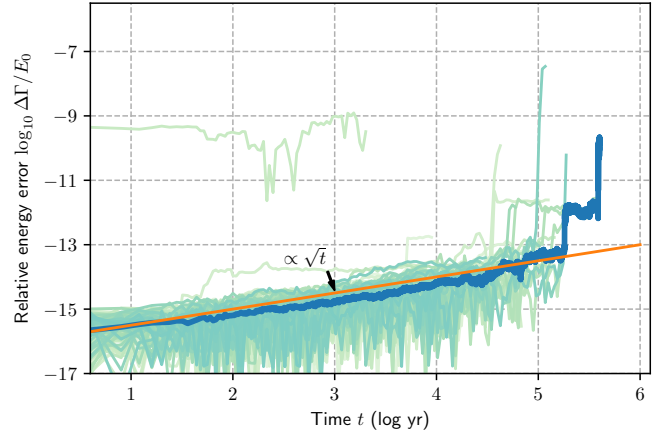
## 8. Comparison with existing integrators

We ran the same eccentric and inclined initial conditions using the REBOUND (Rein & Liu 2012) package integrators IAS15 (Rein & Spiegel 2015) and MERCURIUS (Rein et al. 2019). IAS15 is an adaptive high-order Gauss-Radau integrator, and MERCURIUS is a hybrid symplectic integrator similar to MERCURY (Chambers 1999). This choice is motivated because the REBOUND team made recent and precise comparisons with other available integrators (Rein & Spiegel 2015; Rein & Tamayo 2015). Moreover, we used a very similar implementation of the Kepler step. We also ran the same tests with the multi-time-step symplectic integrator SYMBA (Duncan et al. 1998).

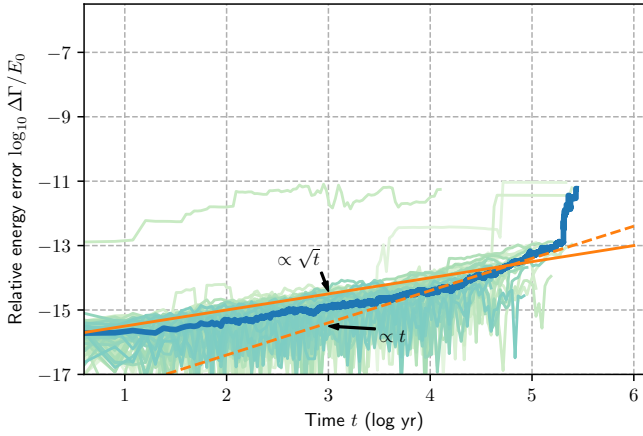
We plot in Fig. 11 the energy error median that results from the integration. Our integrator performs similarly to the



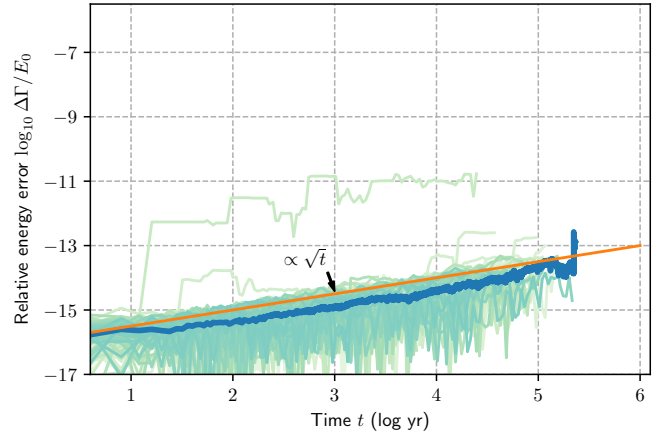
(a) J1: Jacobi coordinates, no pericenter regularisation.



(b) J2: Jacobi coordinates, pericenter regularisation.

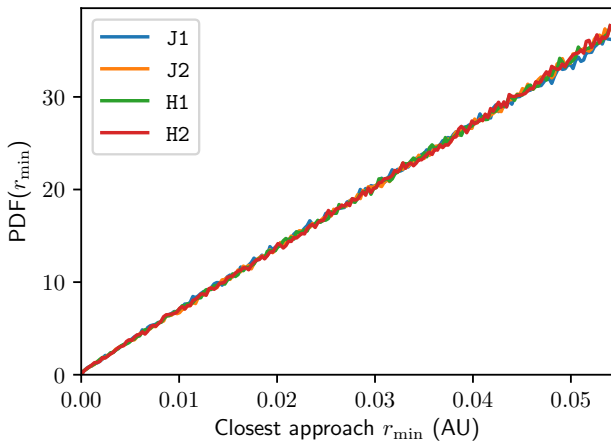


(c) H1: heliocentric coordinates, no pericenter regularisation.



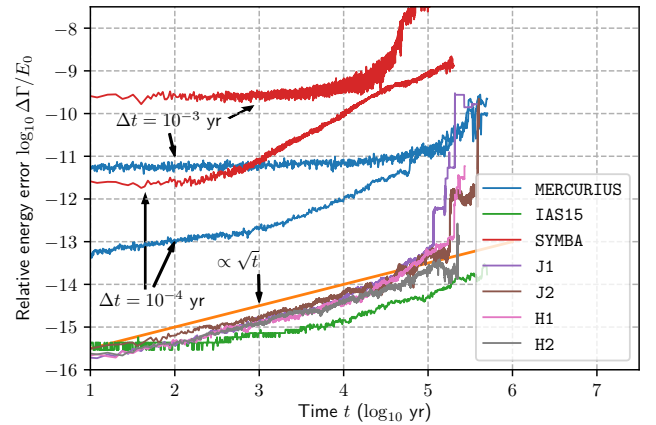
(d) H2: heliocentric coordinates, pericenter regularisation.

**Fig. 9.** Evolution of the energy error for the four runs detailed in the text and Table 1. The light curves represent individual systems, and the thicker blue curves show the medians over the 100 systems. The orange filled lines are proportional to  $\sqrt{t}$  and are the same in all energy error figures for comparison. The dashed orange lines represents a linear error in time to show the deviation from the Brouwer law. We discuss the outliers in the main text.



**Fig. 10.** PDF of the closest approach during close encounters. The PDFs of the four runs are almost identical.

non-symplectic integrator IAS15 and better than MERCURIUS and SYMBA. Our integrator runs at the same speed as MERCURIUS and SYMBA with a time-step of  $\Delta t = 10^{-4}$ . However, IAS15 is much faster on this particular problem (by a factor 10 on



**Fig. 11.** Comparison of the different adaptive integrator runs presented in Table 1 to other codes: IAS15, MERCURIUS, and SYMBA. The same initial conditions were integrated. We plot the median of the energy error.

average). MERCURIUS and SYMBA are not designed to work efficiently at such low-energy errors. However, our motivation for this work was to find a precise symplectic integrator for close

encounters. We thus tried to obtain the best precision from MERCURIUS and SYMBA to make the comparison relevant.

## 9. Discussion

We showed that time renormalisation can lead to very good results for the integration of highly unstable planetary systems using a symplectic integrator (Sect. 3). The algorithm can use a scheme of arbitrary order, however, we have not been able to take advantage of the hierarchical structure beyond the first order in  $\varepsilon$ . This is due to the non linearity of the time renormalisation as explained in Sect. 4. It is possible, however, to cancel every term up to arbitrary order in the time-step using non-hierarchical schemes.

Our time renormalisation uses the perturbation energy to monitor when two planets encounter each other. The algorithm is thus efficient if the two-planet interaction contributes to a significant part of the perturbation energy. As a result, this integrator is particularly well adapted for systems of a few planets (up to a few tens) with similar masses (within an order of magnitude). Moreover, it behaves extremely well in the two-planet problem even when one planet mass is significantly lower than the other.

However, the algorithm is not yet able to spot close encounters between terrestrial planets in a system that contains giant planets, such as the solar system. In the case of the solar system, the perturbation energy is dominated by the interaction between Jupiter and Saturn. We plan to address this particular problem in the future.

## References

- Blanes, S., & Iserles, A. 2012, *Celest. Mech. Dyn. Astron.*, 114, 297  
 Blanes, S., Casas, F., Farrés, A., et al. 2013, *Appl. Num. Math.*, 68, 58  
 Brouwer, D. 1937, *AJ*, 46, 149  
 Chambers, J. E. 1999, *MNRAS*, 304, 793  
 Chambers, J., Wetherill, G., & Boss, A. 1996, *Icarus*, 119, 261  
 Dehnen, W., & Hernandez, D. M. 2017, *MNRAS*, 465, 1201  
 Duncan, M. J., Levison, H. F., & Lee, M. H. 1998, *AJ*, 116, 2067  
 Farrés, A., Laskar, J., Blanes, S., et al. 2013, *Celest. Mech. Dyn. Astron.*, 116, 141  
 Gladman, B., Duncan, M., & Candy, J. 1991, *Celest. Mech. Dyn. Astron.*, 52, 221  
 Hairer, E., Lubich, C., & Wanner, G. 2006, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations* (Berlin: Springer Science & Business Media)  
 Hernandez, D. M. 2016, *MNRAS*, 458, 4285  
 Hernandez, D. M., & Bertschinger, E. 2015, *MNRAS*, 452, 1934  
 Kahan, W. 1965, *Commun. ACM*, 8, 40  
 Kinoshita, H., Yoshida, H., & Nakai, H. 1991, *Celest. Mech. Dyn. Astron.*, 50, 59  
 Koseleff, P. V. 1993, in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, eds. G. Cohen, T. Mora, & O. Moreno, Lecture Notes in Computer Science (Berlin: Springer Berlin Heidelberg), 213  
 Laskar, J. 1990, in *Les Méthodes Modernes de La Mécanique Céleste. Modern Methods in Celestial Mechanics*, 89  
 Laskar, J., & Gastineau, M. 2009, *Nature*, 459, 817  
 Laskar, J., & Robutel, P. 2001, *Celest. Mech. Dyn. Astron.*, 80, 39  
 Laskar, J., Gastineau, M., Delisle, J.-B., Farrés, A., & Fienga, A. 2011, *A&A*, 532, L4  
 Marchal, C., & Bozis, G. 1982, *Celest. Mech.*, 26, 311  
 McLachlan, R. 1995a, *SIAM J. Sci. Comput.*, 16, 151  
 McLachlan, R. I. 1995b, *BIT Num. Math.*, 35, 258  
 Mikkola, S. 1997, *Celest. Mech. Dyn. Astron.*, 67, 145  
 Mikkola, S. 2008, in *Dynamical Evolution of Dense Stellar Systems* (Cambridge: University Press), 246, 218  
 Mikkola, S., & Innanen, K. 1999, *Celest. Mech. Dyn. Astron.*, 74, 59  
 Mikkola, S., & Tanikawa, K. 1999, *Celest. Mech. Dyn. Astron.*, 74, 287  
 Petit, A. C., Laskar, J., & Boué, G. 2018, *A&A*, 617, A93  
 Preto, M., & Tremaine, S. 1999, *AJ*, 118, 2532  
 Rein, H., & Liu, S.-F. 2012, *A&A*, 537, A128  
 Rein, H., & Spiegel, D. S. 2015, *MNRAS*, 446, 1424  
 Rein, H., & Tamayo, D. 2015, *MNRAS*, 452, 376  
 Rein, H., Hernandez, D. M., Tamayo, D., et al. 2019, *MNRAS*, 485, 5490  
 Stiefel, E. L., & Scheifele, G. 1971, *Linear and Regular Celestial Mechanics*. (Berlin: Springer)  
 Stumpff, K. 1962, *Himmelsmechanik, Band I* (Berlin: VEB Deutscher Verlag der Wissenschaften)  
 Wisdom, J. 2006, *AJ*, 131, 2294  
 Wisdom, J., & Holman, M. 1991, *AJ*, 102, 1528  
 Wisdom, J., Holman, M., & Touma, J. 1996, *Fields Inst. Commun.*, 10, 217  
 Yoshida, H. 1990, *Phys. Lett. A*, 150, 262

## Appendix A: Implementation

We give technical details on our implementation choices in this appendix.

### A.1. Kepler equation

The key step in any Wisdom–Holman algorithm is the numerical resolution of the Kepler problem,

$$H_{\text{Kepler}} = \frac{\mathbf{v}^2}{2} - \frac{\mu}{r}, \quad (\text{A.1})$$

where  $\mu = \mathcal{G}M$  and  $M$  is the central mass in the set of coordinates used in the integration. Because this is the most expensive step from a computational point of view, it is particularly important to optimise it. Here, we closely followed the works by Mikkola (1997) and Mikkola & Innanen (1999) and refer to them for more details. Rein & Tamayo (2015) presented an unbiased numerical implementation that can be found in the package REBOUND<sup>2</sup>.

We consider a planet with initial conditions  $\mathbf{r}_0$  and  $\mathbf{v}_0$ . The goal is to determine the position of the planet  $\mathbf{r}$  and its velocity  $\mathbf{v}$  along the Keplerian orbit after a time  $t$ . In order to avoid conversions from Cartesian coordinates into elliptical elements, we use the Gauss  $f$ - and  $g$ -function<sup>3</sup> formalism (e.g. Wisdom & Holman 1991). We have

$$\begin{aligned} \mathbf{r} &= f\mathbf{r}_0 + g\mathbf{v}_0, \\ \mathbf{v} &= \dot{f}\mathbf{r}_0 + \dot{g}\mathbf{v}_0, \end{aligned} \quad (\text{A.2})$$

where the values of  $f$ ,  $g$ ,  $\dot{f}$  and  $\dot{g}$  depend on  $t$ ,  $\mathbf{r}_0$ , and  $\mathbf{v}_0$  and are given in Eq. (A.7). When two planets encounter each other, their orbits may become hyperbolic. To be able to resolve these events as well as ejection trajectories, we use a formulation of the Kepler problem that allows hyperbolic orbits. In order to do so, Stumpff (1962) developed a general formalism that contains the hyperbolic and elliptical case in the same equations. Moreover, this approach avoids the singularity for an eccentricity close to 1. Stumpff introduced special functions

$$c_n(z) = \sum_{j=0}^{+\infty} \frac{(-z)^j}{(n+2j)!}. \quad (\text{A.3})$$

The  $c$ -functions allows us to compute the so-called  $G$ -functions (Stiefel & Scheifele 1971), which are defined as

$$G_n(\beta, X) = X^n c_n(\beta X^2). \quad (\text{A.4})$$

In this formalism, the Kepler equation takes the form (Stumpff 1962)

$$t = r_0 X + \eta_0 G_2(\beta, X) + \zeta_0 G_3(\beta, X), \quad (\text{A.5})$$

of unknown  $X = \int_0^t dt/r$  and where

$$\begin{aligned} \beta &= \frac{2\mu}{r_0} - \mathbf{v}_0^2, \\ \eta_0 &= \mathbf{r}_0 \cdot \mathbf{v}_0, \\ \zeta_0 &= \mu - \beta r_0. \end{aligned} \quad (\text{A.6})$$

In Eq. (A.5),  $X$  plays a similar role to the eccentric anomaly in the classical form. Equation (A.5) can be solved by the Newton method (Rein & Tamayo 2015). The new position and velocity are then obtained with Eq. (A.2) and

$$\begin{aligned} f &= 1 - \mu \frac{G_2(\beta, X)}{r_0}, & \dot{f} &= -\frac{\mu G_1(\beta, X)}{r_0 r}, \\ g &= t - \mu G_3(\beta, X), & \dot{g} &= 1 - \mu \frac{G_2(\beta, X)}{r}, \end{aligned} \quad (\text{A.7})$$

where  $r = r_0 + \eta_0 G_1 + \zeta_0 G_2$ .

### A.2. Computing the effective time-step

In both the Kepler and the perturbation step, it is important to precisely compute the effective time-step (20). For the Kepler step in particular,  $\tau_0$  depends on the difference  $H_0 - E_0$ , where  $H_0$  and  $E_0$  are similar. To avoid numerical errors, we compute the initial energy with compensated summation (Kahan 1965). We save the value and the associated error. We then evaluate the Keplerian energy  $H_0$  using compensated summation and then derive the difference. Compensated summation is also used to update the positions and velocities, and to integrate the real time equation.

During the numerical tests, we realised that in Jacobi coordinates, the perturbation energy  $H_1$  is most of the time lower by almost an order of magnitude than the typical planet energy interaction  $E_1$ . On the other hand,  $H_1$  and  $E_1$  are roughly of the same order for heliocentric coordinates. This shows that the algorithm is less efficient in the detection of an increase in interaction energy (that monitors the close encounters). To circumvent this problem, we slightly modify Eq. (20) for Jacobi coordinates. The effective step sizes are computed with

$$\tau_0 = \sigma f'(H_0 - E_0 + E_1) \text{ and } \tau_1 = \sigma f'(-(H_1 - E_1)). \quad (\text{A.8})$$

The total energy sum is still zero,

$$(H_0 - E_0 + E_1) + (H_1 - E_1) = 0, \quad (\text{A.9})$$

which preserves the equation of motion according to Eq. (13). With this modification, the results between Jacobi and heliocentric coordinates are comparable.

<sup>2</sup> <https://rebound.readthedocs.io/en/latest/>

<sup>3</sup> These functions are different from the time-renormalisation functions used in this article.

## Appendix B: MacLachlan high-order schemes

In [McLachlan \(1995a\)](#), the scheme coefficients  $w_i$  are the coefficients for a symmetric composition of second-order steps. We list in [Table B.1](#) the corresponding coefficients  $a_i$  and  $b_i$  for the

schemes that we used,  $\mathcal{ABA}(6^*)$  and  $\mathcal{ABA}(8^*)$ . McLachlan provided 20 significant digits, therefore we made the computation in quadruple precision and truncated to the appropriate precision.

**Table B.1.** Coefficients of the schemes used in this article.

Scheme	Order	Stages	$a_i$	$b_i$
$\mathcal{ABA}(6^*)$	6	7	$a_1 = 0.39225680523877863191$ $a_2 = 0.51004341191845769875$ $a_3 = -0.471053385409756436635$ $a_4 = 0.068753168252520105975$	$b_1 = 0.78451361047755726382$ $b_2 = 0.23557321335935813368$ $b_3 = -1.17767998417887100695$ $b_4 = 1.3151863206839112189$
$\mathcal{ABA}(8^*)$	8	15	$a_1 = 0.370835182175306476725$ $a_2 = 0.166284769275290679725$ $a_3 = -0.109173057751896607025$ $a_4 = -0.191553880409921943355$ $a_5 = -0.13739914490621317141$ $a_6 = 0.31684454977447705381$ $a_7 = 0.324959005321032390205$ $a_8 = -0.240797423478074878675$	$b_1 = 0.74167036435061295345$ $b_2 = -0.409100825800031594$ $b_3 = 0.19075471029623837995$ $b_4 = -0.57386247111608226666$ $b_5 = 0.29906418130365592384$ $b_6 = 0.33462491824529818378$ $b_7 = 0.31529309239676659663$ $b_8 = -0.79688793935291635398$

**References.** The values are computed from [McLachlan \(1995a\)](#).