



**HAL**  
open science

# A differentiable N-body code for transit timing and dynamical modelling - I. Algorithm and derivatives

Eric Agol, David M. Hernandez, Zachary Langford

## ► To cite this version:

Eric Agol, David M. Hernandez, Zachary Langford. A differentiable N-body code for transit timing and dynamical modelling - I. Algorithm and derivatives. Monthly Notices of the Royal Astronomical Society, 2021, 507, pp.1582-1605. 10.1093/mnras/stab2044 . insu-03748190

**HAL Id: insu-03748190**

**<https://insu.hal.science/insu-03748190>**

Submitted on 9 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A differentiable $N$ -body code for transit timing and dynamical modelling – I. Algorithm and derivatives

Eric Agol <sup>1,2★†</sup>, David M. Hernandez <sup>3,4</sup> and Zachary Langford <sup>1</sup>

<sup>1</sup>*Astronomy Department, University of Washington, Seattle, WA 98195, USA*

<sup>2</sup>*Institut d'Astrophysique de Paris, 98bis Boulevard Arago, Paris F-75014, France*

<sup>3</sup>*Harvard-Smithsonian Center for Astrophysics, 60 Garden St, MS 51, Cambridge, MA 02138, USA*

<sup>4</sup>*Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA*

Accepted 2021 June 30. Received 2021 June 3; in original form 2021 March 22

## ABSTRACT

When fitting  $N$ -body models to astronomical data – such as transit times, radial velocity, and astrometric positions at observed times – the derivatives of the model outputs with respect to the initial conditions can help with model optimization and posterior sampling. Here, we describe a general purpose symplectic integrator for arbitrary orbital architectures, including those with close encounters, which we have recast to maintain numerical stability and precision for small step sizes. We compute the derivatives of the  $N$ -body coordinates and velocities as a function of time with respect to the initial conditions and masses by propagating the Jacobian along with the  $N$ -body integration. For the first time, we obtain the derivatives of the transit times with respect to the initial conditions and masses using the chain rule, which is quicker and more accurate than using finite differences or automatic differentiation. We implement this algorithm in an open source package, `NbodyGradient.jl`, written in the JULIA language, which has been used in the optimization and error analysis of transit-timing variations in the TRAPPIST-1 system. We present tests of the accuracy and precision of the code, and show that it compares favourably in speed to other integrators that are written in C.

**Key words:** planetary systems – planets and satellites: dynamical evolution and stability.

## 1 INTRODUCTION

The gravitational  $N$ -body problem refers to the integration of the positions and velocities of a set of  $N$  point-particles forward or backward in time using Newton's equations, after specifying their masses and initial phase-space coordinates. The solution of the  $N$ -body problem can be put to many uses, for example, matching observational data on a set of astronomical bodies, estimating the long-term stability or sensitivity to initial conditions of a model system, or determining the outcome of interactions between bodies.

For each of these applications, it can be beneficial to be able to compute the derivatives of the state of the system at a given time with respect to the initial conditions and masses. As the  $N$ -body problem is highly non-linear, non-linear optimizers are needed to find the model parameters with the maximum a posteriori probability (MAP), or maximum-likelihood estimate. Derivative-free optimization, such as Nelder–Mead, can be slow to converge, and typically becomes less efficient as the number of dimensions grows. Hence, derivatives can significantly speed up the process of optimization. Once the MAP is found, the Hessian can be computed with derivatives to estimate the uncertainties on parameters. Then, each parameter can be varied along a fixed grid, and the non-linear optimization can be re-run to trace out the likelihood profile. Finally, the posterior can be

sampled using Bayesian techniques that take advantage of derivatives to improve the efficiency of the sampling in high dimensions.

Finite-difference derivatives can be easy to compute numerically; however, finite differences are computationally expensive and limited by numerical precision. The computation of derivatives along with  $N$ -body integration can yield higher precision with less computation time, enabling more effective application of non-linear optimization and parameter uncertainty estimation. This calculation can be laborious, involving propagating derivatives through each time-step of an integration, but the result can be much more computationally efficient and accurate relative to computing derivatives with finite differences.

The calculation of derivatives of the  $N$ -body problem has been investigated in prior work. Mikkola & Innanen (1999) and Rein & Tamayo (2015) derive the variational equations of the symplectic integrator of Wisdom & Holman (1991), the ‘Wisdom–Holman (WH)’ method, to obtain the tangent map of an  $N$ -body system as a function of time, from which the positional variations may be derived as a function of variations in the initial phase-space coordinates. This is implemented in the `WHFast` integrator in the `REBOUND` package based on the WH method (Rein & Tamayo 2015). Second-order variational equations were derived by Rein & Tamayo (2016) for a high-order integrator (IAS15) that is assumed to exactly solve the  $N$ -body equations. Pál (2010) used a Lie-integration scheme, including derivatives with respect to the initial orbital elements and masses, to fit for planet–planet perturbations in radial-velocity-detected systems.

\* E-mail: [agol@uw.edu](mailto:agol@uw.edu)

† Guggenheim Fellow.

## 1.1 Algorithm

The purpose of this paper is to implement first-order derivatives in a symplectic integrator, including the mass derivatives, and allowing for a system hierarchy that is more general than standard symplectic integrators, and which includes derivatives of the transit times with respect to the initial conditions, which is currently absent in the literature. Instead of using the variational equations, which assume exact solution of the  $N$ -body problem for obtaining the derivatives, we compute the derivatives of the  $N$ -body symplectic map, with the goal of yielding a more precise result for the Jacobian of the state of the system at a given time with respect to the specified initial conditions. Although the Rein & Tamayo (2016) integrator could have been put to use for this problem, we are interested in developing a complementary code that trades generality and precision for potentially more speed.

The basic integrator we use has been described in two prior papers: Hernandez & Bertschinger (2015) and Dehnen & Hernandez (2017). The novel aspect underlying the integrator is to allow all bodies to be treated on equal footing. A universal Kepler solver (Wisdom & Hernandez 2015) is used to integrate pairs of bodies with Keplerian drifts forward in time, interspersed with constant-velocity corrections that are negative in time, while using operator splitting to create a symplectic and time-symmetric integrator out of the original concept proposed by Gonçalves Ferrari, Boekholt & Portegies Zwart (2014). A potential advantage of this approach is the adaptability to different problems with various geometries, such as hierarchical triples, pairs of binaries, or other more complex hierarchies (Hamers & Portegies Zwart 2016). The popular WH method, and its variants, that uses different coordinates and Hamiltonian splittings (Hernandez & Dehnen 2017) assume that there is a dominant mass and widely separated planets. For general applications, these assumptions can be too constraining.

A drawback of this integrator is the potential for numerical cancellation errors to accumulate due to alternating negative and positive time-steps that are a necessary part of the algorithm. In developing this code, we found that these cancellations caused numerical errors that accrue in proportion to the number of time-steps. This becomes more significant when the time-steps are short, as more steps are required for a given integration time. We have rectified this problem by combining the negative and positive time-steps into a single step, and cancelling the terms analytically, which we find reduces the numerical errors significantly. Thus, another goal of this paper is to describe this improved integrator.

## 1.2 Motivation: TTVs and photodynamics

The particular application we have in mind is the detection and characterization of exoplanet systems. Planetary interactions become important when data are of high precision, or if integrations are carried out on long time-scales to study system stability. The first example of non-Keplerian interactions being important was the pulsar exoplanet system PSR 1257 + 12 (Wolszczan & Frail 1992). As had been predicted, the interactions of the planets were detected in the pulsar timing, and then used to confirm the planetary nature of the system, as well as measure the inclinations and masses of the planets by breaking the mass-inclination degeneracy that accompanies Doppler shifts (Malhotra et al. 1992; Rasio et al. 1992; Peale 1993; Wolszczan 1994). Secondly, high-precision radial velocity measurements of exoplanet systems also require accounting for planet–planet interactions. An early example of this is GJ 876, which required an  $N$ -body integration to match the observed stellar

radial velocity instead of treating the radial velocity signal as a sum of unperturbed Keplerians (Laughlin & Chambers 2001). Thirdly, the *Kepler* spacecraft yielded sufficient precision of the times of transit of exoplanets to produce a novel means of detecting and characterizing exoplanets: transit-timing variations (TTVs; Holman & Murray 2005; Agol et al. 2005). The Kepler-9 planet system showed strong anti-correlated variations in the times of transit relative to a fixed ephemeris, which allows for measurement of the planet masses (Holman et al. 2010; Freudenthal et al. 2018; Borsato et al. 2019). Currently, several planets have been detected with TTVs, while hundreds have been characterized (see Agol & Fabrycky 2017; Jontof-Hutter 2019, and references therein).

TTVs are entirely due to non-Keplerian motion of the planetary orbits. In the Newtonian two-body problem, transits occur at regular intervals, and so the transit times are uniformly spaced in time with the orbital period of the system. When three or more bodies interact, each pair of bodies no longer follows a Keplerian orbit, but is perturbed by the other bodies in the system. In the planetary case, the perturbations of the times of transit by other planets are typically small compared with the orbital period of the planet. TTVs are defined as the residuals of a linear fit to the times of transit (Agol et al. 2005), and so by definition TTVs are imparted by non-Keplerian motion. Consequently, the presence of TTVs typically requires an  $N$ -body model for the computation of the times of transit.

The advent of the detection of TTVs spurred theoretical models for short-term planetary dynamics. Analytic prescriptions exist for TTVs (e.g. Agol et al. 2005; Nesvorný & Beaugé 2010; Lithwick, Xie & Wu 2012; Nesvorný & Vokrouhlický 2014; Deck & Agol 2015; Agol & Deck 2016; Deck & Agol 2016; Hadden & Lithwick 2016; Nesvorný & Vokrouhlický 2016). However, the dynamics of multiplanet systems is sufficiently complex that any analytic prescription is only accurate in a confined region of parameter space and/or limited time-scales, and generally needs to be checked against numerical integration since it is unknown beforehand whether these restrictions apply to the masses and orbital elements of a particular system (e.g. Deck & Agol 2015; Jontof-Hutter et al. 2016; Hadden & Lithwick 2017; Linial, Gilbaum & Sari 2018; Yoffe, Ofir & Aharonson 2021). On the other hand, numerical integration can be much more computationally expensive and can accrue numerical errors.

When optimizing a TTV model, the gradient of the likelihood is often required to find the direction in which the variation of the initial conditions will improve the likelihood. The likelihood gradient in turn requires the gradient of each transit time with respect to the initial conditions. Often finite differences are used to estimate this gradient; however, finite-difference derivatives are limited by the numerical accuracy of the integration; see Rein & Tamayo (2016) regarding the drawbacks of finite-difference derivatives. This can cause difficulty in optimizing numerical TTV model fits. In addition, finite differences are expensive to compute as at least two integrations are required for each parameter, and truncation or round-off errors can limit the precision. For planetary systems, this requires  $14N$  integrations for  $N$  planets.

Computing the posterior parameter distributions for observed systems requires numerous evaluations of the likelihood, which becomes difficult to explore for high-dimensional planetary systems due to the ‘curse of dimensionality’ making grid-based and Markov-chain-based integrations prohibitive. This can be ameliorated by Hamiltonian Markov chains, which require computation of the derivatives of the likelihood function (Girolami & Calderhead 2011).

Dynamical interactions have also been measured in systems where multiple stars are present. The triple-star KOI-126 was characterized

with a ‘photodynamical’ model (Carter et al. 2011), which requires coupling an  $N$ -body code to a photometric model. The architecture of this system prohibits the use of a standard WH-type symplectic integrator to describe the dynamics as there is a binary star in orbit about a more massive star. Similarly, circumbinary planets, such as Kepler-16b, have been found which require a photodynamical model of the stars and planets (Doyle et al. 2011). One can imagine more complex geometries, such as a planet–moon pair orbiting a binary star, which would also require a photodynamical code to model. The computational expense of each of these models is significant, so that obtaining converged posterior parameters is a challenge. An advantage of photodynamical models is that the covariance between transit parameters and orbital parameters may be computed directly from the light curve.

Given these observational modelling problems, the motivation for this code is to provide derivatives of a general  $N$ -body integrator for short-term integrations to model stellar and planetary systems with arbitrary hierarchy, and to compute the derivatives of the model with respect to the initial conditions to allow for better optimization of the model likelihood and to explore parameter space more efficiently with Hamiltonian Markov chain Monte Carlo.

In Section 2, we describe the original symplectic integrator algorithm, and discuss its numerical instability. In Section 3, we give the modifications to the algorithm we have made to prevent numerical cancellation. In Section 4, we introduce the derivatives of the algorithm. In Section 5, we describe the implementation and precision of the algorithm. In Section 6, we compare the algorithm with other  $N$ -body integrators. Finally, in Section 7 we conclude. Throughout the paper, we provide hyperlinks via check marks (✓) after equations that point to the lines of code in GITHUB that implement these equations. As the code has been tested with unit tests, we interpret this as a validation of these equations.

## 2 OVERVIEW OF SYMPLECTIC INTEGRATOR

We carry out the  $N$ -body integration with a symplectic integrator (Channell & Scovel 1991) that uses Kepler steps to integrate pairs of bodies, interspersed with constant velocity corrections, thus treating each and every body in an identical manner (Hernandez & Bertschinger 2015). The advantage of this approach relative to the WH method is that the integrator can be used as an all-purpose integrator for studying systems with a range of architectures. The integrator is especially powerful when binaries at any scale are present. A fourth-order corrector gives higher precision to this integrator without much additional computational cost (Dehnen & Hernandez 2017); hence, we refer to the algorithm as DH17<sup>1</sup> in what follows. DH17 is mathematically written by equation (29) in Dehnen & Hernandez (2017), using  $\alpha = 0$ . In Section 4.9 and Algorithm 2, we present a generalization of the method, described mathematically by equation (39) in Dehnen & Hernandez (2017). This generalization is also referred to as DH17 as Dehnen & Hernandez (2017) also called both methods the same name. The methods described here are all time-reversible and time-symmetric (Hairer, Lubich & Wanner 2006; Hernandez & Bertschinger 2018). We give an overview of DH17, along with transit-time finding, in algorithm 1, which uses a fixed time-step,  $h$ , from initial time  $t_0$  over a duration  $t_{\max}$ .

<sup>1</sup>In Dehnen & Hernandez (2017), this algorithm was called ‘DH16’. We have used ‘DH17’ to reflect the publication year.

Unfortunately, the DH17 algorithm is numerically unstable. Consequently, we have modified the DH17 algorithm, and present a modified algorithm, which we will refer to as AHL21, in which we combine pairs of steps of the DH17 algorithm into a single step. The AHL21 algorithm is mathematically identical to the DH17 algorithm; however, because of carrying out the cancellation analytically rather than numerically, it is more numerically stable, as we describe in Section 3. But first we start by outlining the DH17 algorithm and its drawbacks.

### 2.1 DH17 algorithm summary

The original DH17 algorithm is given in Algorithm 1. The algorithm is derived from splitting the Hamiltonian into pairwise Keplerian terms:

$$\begin{aligned} H &= T + V, \\ &= T + \sum_i \sum_{j>i} V_{ij}, \\ &= T + \sum_i \sum_{j>i} (K_{ij} - T_{ij}), \end{aligned} \quad (1)$$

where  $T$  is the kinetic energy,  $V$  is the total potential energy, while  $T_{ij}$ ,  $V_{ij}$ , and  $K_{ij}$  are the kinetic, the potential, and the total energy of a pair of bodies  $i$  and  $j$ . The  $K_{ij}$  term is the two-body Hamiltonian, whose solution amounts to a Keplerian orbit whose centre of mass moves at a constant velocity, hence the notation ‘ $K$ ’ (Hernandez & Bertschinger 2015). Note that the minus sign in front of the kinetic energy term indicates a *backward* drift in time.

A symplectic integration is achieved by successively solving each component of the Hamiltonian, using a Kepler solver and a simple drift, to give the positions and velocities at the start of the next component (Dehnen & Hernandez 2017). The creation of a second-order map from this splitting of the Hamiltonian involves division of each time-step of duration  $h$  into two sub-steps of duration  $h/2$ . In each of the substeps, the order of application of the terms is reversed to cancel first-order error terms. In addition, a fourth-order velocity corrector is added in the middle of the time-step, which amounts to applying tidal accelerations to the velocities that are neglected in the two-body elements of the Hamiltonian, yielding much higher precision without much additional computational effort; this results in Algorithm 1.

As mentioned, there is an unfortunate drawback to the DH17 algorithm, which is the *negative* time-step. In cases in which the potential energy term is small, the  $K_{ij}$  and  $-T_{ij}$  terms nearly cancel. What this means is that the motion induced by these terms in the mapping can be nearly equal and opposite, causing numerical cancellation which leads to round-off errors that accumulate with time. This has two different causes. First, the centre-of-mass portion of these Hamiltonians is identical, and thus cancels exactly (Dehnen & Hernandez 2017). Secondly, if the acceleration is weak, or the time-step is short enough that the acceleration does not result in a significant change in velocity, then the Keplerian step is nearly inertial, and so the entire Kepler step is very nearly equal and opposite to the negative drift. These two sources of cancellation can lead to numerical errors when implementing the algorithm. We found these errors to be severe for long integrations, for weakly interacting bodies (e.g. pairs of planets), or for very short time-steps, which compounds the error more rapidly. We present a solution to this issue in the next section, which is our first main result.

As a side note, the DH17 algorithm differs in its accuracy from symplectic integrators typically used for planetary systems. Planetary

**Data:** Initial Cartesian coordinates and masses at time  $t = t_0$ .

**Result:** Integration of  $N$ -body system over time  $t_{\max}$ , and resulting times of transit and derivatives.

```

for  $t - t_0 < t_{\max}$  do
    Drift all particles for time  $h/2$ ;
    for all pairs of particles  $(i, j)$  do
        Drift particles  $i$  and  $j$  for time  $-h/2$ ;
        Apply a Kepler solver to advance the relative
        coordinates of  $i$  and  $j$  by  $h/2$ ;
        Advance center of mass coordinates of  $i$  and  $j$  by  $h/2$ ;
    end
    Apply fourth-order velocity correction to all particles over
    time step  $h$ ;
    for reversed pairs of particles  $(i, j)$  do
        Apply a Kepler solver to advance the relative
        coordinates of  $i$  and  $j$  by  $h/2$ ;
        Advance center of mass coordinates of  $i$  and  $j$  by  $h/2$ ;
        Drift particles  $i$  and  $j$  for time  $-h/2$ ;
    end
    Drift all particles for time  $h/2$ ;
    if transit has occurred for particles  $i$  and  $j$  then
        Refine transit time, and save.
    end
    Increment time  $t$  by  $h$ .
end
    
```

**Algorithm 1:** Transit times with DH17 symplectic integration

systems with a dominant mass are described by a Hamiltonian:

$$H = H_A + \epsilon H_B, \quad (2)$$

where  $\epsilon \ll 1$  in which the terms  $H_A$  and  $H_B$  are typically chosen to be integrable. For this Hamiltonian, the WH method (Wisdom & Holman 1991) has been developed which carries an energy error of  $\mathcal{O}(\epsilon h^2)$ ,<sup>2</sup> where  $h$  is the step size. In contrast, the DH17 algorithm is fourth order and its error is  $\mathcal{O}(\epsilon h^4)$ . However, while WH assumes  $\epsilon \ll 1$ , DH17 does not require this assumption.

### 3 THE MORE ACCURATE AHL21 ALGORITHM

In this paper, we present a modified version of the DH17 algorithm in which the negative drifts ( $-T_{ij}$ ) and Keplerian steps ( $K_{ij}$ ) are combined algebraically, so that leading-order terms are cancelled by hand. This exact cancellation prevents the accumulation of round-off and truncation errors that occur when implementing the DH17 algorithm. We find that this approach gives a higher precision numerical algorithm yielding results that obey the expected  $h^4$  scaling of the algorithm down to machine precision on the time-scales we have tested.

To describe this new approach, we first need to summarize the application of these two sub-steps.

#### 3.1 Kinetic-energy drift

The drift term is the most straightforward: each particle (for  $T$ ) or pair of particles (for  $T_{ij}$ ) simply drifts inertially,

$$\mathbf{x}_i(t+h) = \mathbf{x}_i(t) + h\mathbf{v}_i(t), \quad (3)$$

<sup>2</sup>Different conventions are used for the scalings with  $\epsilon$ . In the convention of Hernandez & Dehnen (2017), the error scales as  $\mathcal{O}(\epsilon^2 h^2)$ . In this convention, all scalings get an extra factor of  $\epsilon$ .

for a time-step  $h$ , where  $\mathbf{x}_i(t)$  and  $\mathbf{v}_i(t)$  are, respectively, the position and velocity vectors of the  $i$ th body at time  $t$ . Again, note that  $-T_{ij}$  in equation (1) indicates that  $h$  is negative.

#### 3.2 Universal Kepler step

To carry out the  $K_{ij}$  mapping, we use a universal Kepler solver to compute the change in the relative position between the bodies (Wisdom & Hernandez 2015). The Kepler solver uses a universal Kepler equation based upon the initial positions and velocities of a pair of bodies at the start of a step. The solution of Kepler's equation enables a mapping of the initial phase-space coordinates to the final phase-space coordinates after a time  $h$  assuming pairwise Keplerian motion (i.e. neglecting every other body in the system).

The equation of motion in Cartesian coordinates derived from the  $K_{ij}$  Hamiltonian for each Kepler step is given by

$$\ddot{\mathbf{x}}_{ij} = -\frac{k\mathbf{x}_{ij}}{r_{ij}^3}, \quad (4)$$

where  $k = Gm_{ij}$ ,  $m_{ij} = m_i + m_j$  is the sum of the masses of the  $i$ th and  $j$ th pair of bodies,  $\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$  and  $r_{ij} = |\mathbf{x}_{ij}|$ . The universal solver transforms the time dependence to an independent variable,  $s$ , defined by  $\dot{s} = ds/dt = r^{-1}$ , where  $r$  is the distance between the bodies, which simplifies the equations of motion. In the rest of this section, we drop the subscript  $ij$  from the mass, coordinate, and velocity vectors, i.e.  $m = m_{ij}$ ,  $r = r_{ij}$ ,  $\mathbf{x} \equiv \mathbf{x}_{ij}$ , and  $\mathbf{v} \equiv \mathbf{v}_{ij}$ . We will refer to the Cartesian coordinates for the Keplerian as  $(\mathbf{x}_0, \mathbf{v}_0)$  at the start of a step and  $(\mathbf{x}, \mathbf{v})$ , a time  $h$  later.

A Kepler step uses the fact that in the two-body problem angular momentum is conserved; thus, the final relative positions and velocities of the two bodies,  $\mathbf{x}$  and  $\mathbf{v}$ , are in the same plane as the initial relative positions and velocities,  $\mathbf{x}_0$  and  $\mathbf{v}_0$ , while the centre-of-mass velocity remains constant and the centre-of-mass position drifts at a constant rate as there are no external perturbers. This means that the final relative positions and velocities can be expressed as a linear combination of the initial relative velocities and positions:

$$\begin{aligned} \mathbf{x} &= f\mathbf{x}_0 + g\mathbf{v}_0 \\ \mathbf{v} &= \dot{f}\mathbf{x}_0 + \dot{g}\mathbf{v}_0, \end{aligned} \quad (5)$$

where  $f$  and  $g$  are Gauss's functions, which we define in more detail below as a function of  $\mathbf{x}_0$ ,  $\mathbf{v}_0$ ,  $h$ , and  $k$ , where  $k = Gm$  is the central force constant.

Then, the equations describing the initial and final states are given by Wisdom & Hernandez (2015), based on Mikkola & Innanen (1999). We define  $r_0 = |\mathbf{x}_0|$  is the initial separation,  $v_0 = |\mathbf{v}_0|$  is the initial relative speed, and  $r = |\mathbf{x}|$ ,  $v = |\mathbf{v}|$  are the separation and relative speed at the end of the step. We define two additional quantities:

$$\eta_0 = \mathbf{x}_0 \cdot \mathbf{v}_0, \quad (6)$$

$$\begin{aligned} \beta &= \frac{2k}{r_0} - v_0^2, \\ &= \frac{2k}{r} - v^2. \end{aligned} \quad (7)$$

Expressing the final positions and velocities in terms of the initial values requires the Gauss  $f$  and  $g$  functions, which are given by

$$f = 1 - \frac{k}{r_0} G_2, \quad (8)$$

$$g = r_0 G_1 + \eta_0 G_2, \quad (9)$$

and their derivatives

$$\dot{f} = -\frac{k}{rr_0}G_1, \quad (10)$$

$$\dot{g} = \frac{1}{r}(r_0G_0 + \eta_0G_1), \quad (11)$$

where  $G_i(\beta, s)$  are four functions whose definitions depend on the sign of  $\beta$  for  $i = 0, \dots, 3$  (Table 1). With these definitions, Wisdom & Hernandez (2015) show that

$$r(s) = r_0G_0 + \eta_0G_1 + kG_2. \quad (12)$$

This equation may also be derived from conservation of angular momentum, requiring  $\mathbf{x}_0 \times \mathbf{v}_0 = \mathbf{x} \times \mathbf{v}$ , which yields the condition  $f\dot{g} - g\dot{f} = 1$ ; this equation is equivalent to equation (12). We transform these equations from  $s$  to  $\gamma = |\beta|^{1/2}s$ , as  $\gamma$  is dimensionless. Equation (12) can be integrated over a time-step,  $h$ , to give an implicit Kepler's equation for  $\gamma$ :

$$h = r_0G_1 + \eta_0G_2 + kG_3, \quad (13)$$

which can be solved using Newton's method to find  $\gamma$  as a function of  $h, r_0, \eta_0, k$ , and  $\beta$ . The functions  $G_0, G_1, G_2$ , and  $G_3$  are defined in Table 1 in terms of trigonometric and hyperbolic functions (Wisdom & Hernandez 2015), which differ based upon whether the bodies are bound (elliptic) or unbound (hyperbolic).<sup>3</sup> As the  $G_i$  functions only depend upon  $\gamma$  and  $\beta$ , once  $\gamma$  is found numerically, the remainder of the Kepler step simply involves algebraic computation.

### 3.3 Combined Kepler and drift step

In the AHL21 algorithm, these two steps,  $-T_{ij}$  and  $K_{ij}$ , are combined in different orders: either a negative drift followed by a Kepler step,  $-T_{ij} + K_{ij}$ , or a Kepler step followed by a negative drift,  $K_{ij} - T_{ij}$  (Algorithm 2). As these operations do not commute, we need to handle each one separately. A diagram showing the order of these mappings is given in Fig. 1. In each case, the position coordinate takes on an intermediate value, which changes the nature of the combined steps.

We describe these two options in the following subsections.

#### 3.3.1 Drift then Kepler (DK)

In the first case, the negative drift is taken first, yielding an intermediate position

$$\hat{\mathbf{x}}_0 = \mathbf{x}_0 - h\mathbf{v}_0. \quad (14)$$

With this modified value of the initial position, the Gauss  $f, g, \dot{f}$ , and  $\dot{g}$  functions need to be computed from  $(\hat{\mathbf{x}}_0, \mathbf{v}_0, k, h)$  after solution of Kepler's equation for  $\gamma$ , so we indicate these functions with a hat, e.g.  $\hat{f} \equiv f(\hat{\mathbf{x}}_0, \mathbf{v}_0, k, h)$ . In addition, we would like to find the difference between the final and initial coordinates,  $\Delta\mathbf{x}_{DK} = \hat{\mathbf{x}} - \mathbf{x}_0$  and  $\Delta\mathbf{v}_{DK} = \hat{\mathbf{v}} - \mathbf{v}_0$ ; this allows for a more accurate computation of these quantities when the step sizes are small. In the combined step, -Drift + Kepler (which we indicate with 'DK'), the resulting term is

$$\begin{aligned} \Delta\mathbf{x}_{DK} &= (\hat{f} - 1)\mathbf{x}_0 + (\hat{g} - h\hat{f})\mathbf{v}_0, \\ \Delta\mathbf{v}_{DK} &= \hat{f}\mathbf{x}_0 + (\hat{g} - h\hat{f} - 1)\mathbf{v}_0, \end{aligned} \quad (15)$$

<sup>3</sup>This equations simplifies to a cubic in the parabolic case when  $\beta = 0$ . The solution of this cubic is also used as the starting guess for the Newton's solver in the elliptic and hyperbolic cases.

where, again,  $\hat{f}, \hat{g}, \hat{f},$  and  $\hat{g}$  are all computed in terms of  $(\hat{\mathbf{x}}_0, \mathbf{v}_0, k, h)$ . The scalar functions in equation (15) are given in Appendix A.

Note that in these equations the 1's are cancelled analytically; this yields more stable computation of the changes in the positions and velocities when these are small.

#### 3.3.2 Kepler then Drift (KD)

In the other case, a Kepler step is applied first, followed by a negative drift. The Kepler step can be computed in terms of the initial coordinates,  $\mathbf{x}_0$  and  $\mathbf{v}_0$ , yielding intermediate coordinates  $(\check{\mathbf{x}}, \check{\mathbf{v}})$ , and then the negative drift is applied resulting in  $\mathbf{x} = \check{\mathbf{x}} - h\check{\mathbf{v}}$  (Fig. 1).

We combine these and take the difference with the initial coordinates,  $\Delta\mathbf{x}_{KD} = \check{\mathbf{x}} - h\check{\mathbf{v}} - \mathbf{x}_0$  and  $\Delta\mathbf{v}_{KD} = \check{\mathbf{v}} - \mathbf{v}_0$ , to give the resulting difference vectors

$$\begin{aligned} \Delta\mathbf{x}_{KD} &= (f - h\dot{f} - 1)\mathbf{x}_0 + (g - h\dot{g})\mathbf{v}_0, \\ \Delta\mathbf{v}_{KD} &= \dot{f}\mathbf{x}_0 + (\dot{g} - 1)\mathbf{v}_0, \end{aligned} \quad (16)$$

where  $f, g, \dot{f}$ , and  $\dot{g}$  are all computed in terms of  $(\mathbf{x}_0, \mathbf{v}_0, k, h)$ , and the 'KD' indicates that the Kepler step precedes the negative drift, Kepler-Drift. The scalar functions in equation (16) are given in Appendix A.

We take care that these functions are evaluated in a numerically stable manner to avoid round-off error due to cancellations between terms at small time-steps. With the combination of the drift and Kepler steps, it turns out that we no longer need the drift of the centre-of-mass coordinates in each Kepler step as these cancel exactly.

Thus, the DH17 algorithm simplifies significantly at the expense of making the substeps slightly more complicated. This new combined algorithm we dub 'AHL21', which is given in Algorithm 2. The fourth-order correction is the same as that given in Dehnen & Hernandez (2017), and is summarized in Section 4.8. The transit-time finding is described below in Section 4.11. An alternate version of the algorithm in which the combined drift and Kepler steps are replaced by a kick for some pairs of bodies is given in Section 4.7.

The primary goal of this paper is to describe the implementation and differentiation of the AHL21 algorithm, yielding the derivatives of the transit times with respect to the initial conditions. Along the way, we compute the derivatives of the state of the system at each time-step with respect to the initial conditions, which may be used for other applications such as photodynamics, radial velocity, astrometry, or computation of Lyapunov exponents. Next we describe the derivative computation.

## 4 DIFFERENTIATION OF SYMPLECTIC INTEGRATOR

We divide the differentiation of algorithm 2 into a series of steps:

(i) Derivative of coordinates at end of a symplectic step with respect to coordinates at the beginning. This includes the Kepler step, drifts, and kicks (Sections 4.4–4.7).

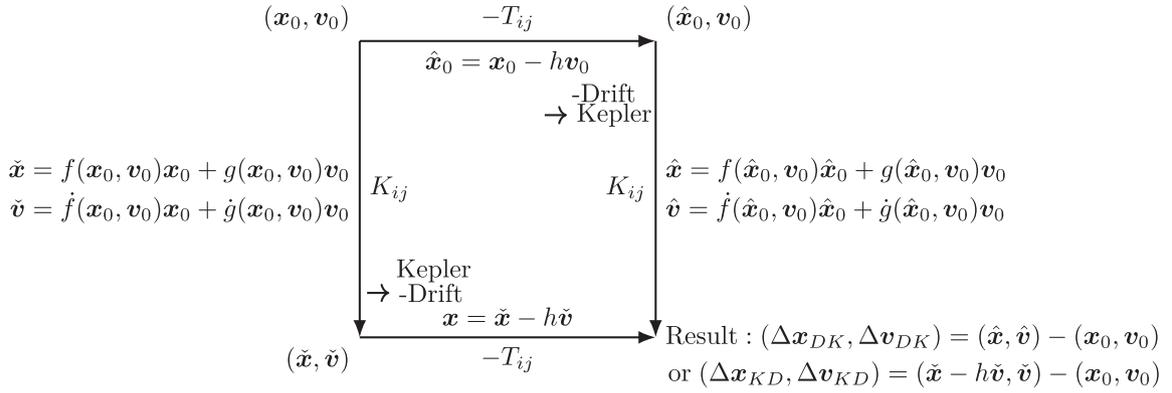
(ii) Derivatives of fourth-order velocity correction (Section 4.8).

(iii) Propagation of Jacobians through each of these steps (Section 4.9).

(iv) Derivative of parameters output at specified times with respect to the coordinates at the symplectic time grid. Here, we give the example of the derivatives of the transit times with respect to the initial conditions (Section 4.11), but this could also include eclipse

**Table 1.** Functions  $G_i(s)$  used in solving Universal Kepler equation, and functions  $H_i$  used later in the combined drift + Kepler step and its derivatives.

Variable	Elliptic	Parabolic	Hyperbolic
$\beta$	$>0$	$=0$	$<0$
$\gamma$	$\sqrt{\beta}s$	$-$	$\sqrt{-\beta}s$
$G_0(\beta, \gamma)$ ✓	$\cos \gamma = 1 - \beta G_2$	1	$\cosh \gamma = 1 - \beta G_2$
$G_1(\beta, \gamma)$ ✓	$\beta^{-1/2} \sin \gamma = 2\beta^{-1/2} \sin \frac{1}{2}\gamma \cos \frac{1}{2}\gamma$	$s$	$(-\beta)^{-1/2} \sinh \gamma = 2(-\beta)^{-1/2} \sinh \frac{1}{2}\gamma \cosh \frac{1}{2}\gamma$
$G_2(\beta, \gamma)$ ✓	$\beta^{-1}(1 - \cos \gamma) = 2\beta^{-1} \sin^2 \frac{1}{2}\gamma$	$\frac{1}{2}s^2$	$\beta^{-1}(1 - \cosh \gamma) = -2\beta^{-1} \sinh^2 \frac{1}{2}\gamma$
$G_3(\beta, \gamma)$ ✓	$\beta^{-1}(\gamma - \sin \gamma)/\sqrt{\beta}$	$\frac{1}{6}s^3$	$\beta^{-1}(\gamma - \sinh \gamma)/\sqrt{-\beta}$
$H_1(\beta, \gamma)$ ✓	$\beta^{-2}(2 - 2\cos \gamma - \gamma \sin \gamma)$	$\frac{1}{12}s^4$	$\beta^{-2}(2 - 2\cosh \gamma + \gamma \sinh \gamma)$
$H_2(\beta, \gamma)$ ✓	$\beta^{-3/2}(\sin \gamma - \gamma \cos \gamma)$	$\frac{1}{3}s^3$	$(-\beta)^{-3/2}(-\sinh \gamma + \gamma \cosh \gamma)$
$H_3(\beta, \gamma)$ ✓	$\beta^{-1}(4 \sin \gamma - \sin \gamma \cos \gamma - 3\gamma)/\sqrt{\beta}$	0	$\beta^{-1}(4 \sinh \gamma - \sinh \gamma \cosh \gamma - 3\gamma)/\sqrt{-\beta}$
$H_5(\beta, \gamma)$ ✓	$\beta^{-1}(3 \sin \gamma - \gamma \cos \gamma - 2\gamma)/\sqrt{\beta}$	0	$\beta^{-1}(3 \sinh \gamma - \gamma \cosh \gamma - 2\gamma)/\sqrt{-\beta}$
$H_6(\beta, \gamma)$ ✓	$\frac{1}{2}\beta^{-2}(9 - 8 \cos \gamma - \cos 2\gamma - 6\gamma \sin \gamma)$	0	$\frac{1}{2}\beta^{-2}(9 - 8 \cosh \gamma - \cosh 2\gamma + 6\gamma \sinh \gamma)$


**Figure 1.** The order of the combined substeps (from upper left corner to lower right corner) has two sequences: first a negative drift followed by a Kepler step, then a Kepler step followed by a negative drift. These two options need to be handled separately, and notation for the intermediate steps is summarized in this diagram.

**Data:** Initial Cartesian coordinates and masses at time  $t = t_0$ .

**Result:** Integration of  $N$ -body system over time  $t_{\max}$ , and resulting times of transit and derivatives.

```

for  $t - t_0 < t_{\max}$  do
    Kick particles in  $A^C$  for time  $h/6$ ;
    Drift all particles for time  $h/2$ ;
    for pairs of particles  $(i, j)$  in  $A$  do
        Apply a combined -Drift+Kepler step for bodies  $i$  and  $j$ 
        over a time  $h/2$  to give the changes in position and
        velocity of  $\Delta x_{DK}$  and  $\Delta v_{DK}$  and update  $x_{ij}$  and  $v_{ij}$ ;
    end
    Apply velocity correction and a kick, both multiplied by
     $2/3$ , to particles in  $A^C$ ;
    Apply velocity correction to particles in  $A$ ;
    for reversed pairs of particles  $(i, j)$  in  $A$  do
        Apply a combined Kepler-Drift step for bodies  $i$  and  $j$ 
        over a time  $h/2$  to give the changes in position and
        velocity of  $\Delta x_{KD}$  and  $\Delta v_{KD}$  and update  $x_{ij}$  and  $v_{ij}$ ;
    end
    Drift all particles for time  $h/2$ ;
    Kick particles in  $A^C$  for time  $h/6$ ;
    if transit has occurred for particles  $i$  and  $j$  then
        Refine transit time, and save.
    end
    Increment time  $t$  by  $h$ .
end
    
```

**Algorithm 2:** Transit times with AHL21 symplectic integration.

times, radial velocity at pre-specified time, or relative positions of the bodies at times of observation. This step involves an AHL21 step with a fractional time duration.

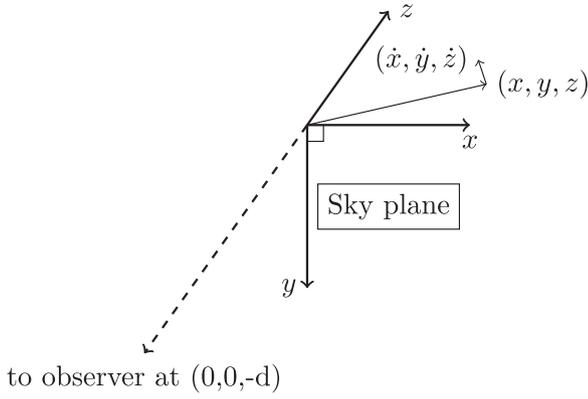
We describe each of these steps in turn, after some preliminaries.

#### 4.1 Notation conventions

The integration is carried out in inertial Cartesian coordinates (Hernandez & Bertschinger 2015), while the initial conditions of the  $N$  bodies are specified in either of two forms: Cartesian coordinates of  $N$  bodies, or orbital elements in a hierarchy of  $N - 1$  Keplerians (which we leave to a future paper). The initial time of the start of the integration,  $t_0$ , requires a snapshot of the phase-space coordinates or orbital elements, which fully specify the problem with the addition of the masses of the bodies,  $\mathbf{m} = \{m_1, \dots, m_N\}$ , which are constant in time. In this section, we describe the phase-space coordinates.

#### 4.2 Code units

We utilize units for masses in  $M_\odot$ , positions in au, and time in days. Our gravitational constant is given by  $GM_\odot = 0.00029598 \text{ au}^3 \text{d}^{-2} M_\odot^{-1}$ . The initial conditions, then, simply need to be specified in terms of masses in  $M_\odot$ , positions in au, and velocities in  $\text{au d}^{-1}$ .



**Figure 2.** Cartesian coordinate system. Body  $i$  is at position  $\mathbf{x}_i = (x_i, y_i, z_i)$  with velocity  $\mathbf{v}_i = (\dot{x}_i, \dot{y}_i, \dot{z}_i)$ .

### 4.3 Cartesian coordinates

The Cartesian coordinates utilize a right-handed coordinate system for which the sky plane is the  $x$ - $y$  plane, while the  $z$ -axis is along the line of sight, increasing away from the observer (Fig. 2). Positions for each body are denoted with a vector  $\mathbf{x}_i(t) = (x_i(t), y_i(t), z_i(t))$ , while velocities are denoted with  $\mathbf{v}_i(t) = (\dot{x}_i(t), \dot{y}_i(t), \dot{z}_i(t))$ , with subscript  $i = 1, \dots, N$  labelling each body, and  $\dot{c} = \frac{dc}{dt}$  indicates time derivative of variable  $c$ . The observer is located at  $\mathbf{x}_{\text{obs}} = (0, 0, -d)$ , where  $d$  is the distance of the observer to the centre of mass of the system.

The initial conditions are completely specified via  $\mathbf{q}(t_0)$ , where  $\mathbf{q}(t) = \{\mathbf{x}_i(t), \mathbf{v}_i(t), m_i; i = 1, \dots, N\}$ . The vector  $\mathbf{q}(t)$  has  $7N$  elements, where the  $7(i-1) + j$ th element refers to planet  $i$  and the  $j$ th element of the vector

$$\mathbf{q}_i(t) = \{x_i(t), y_i(t), z_i(t), \dot{x}_i(t), \dot{y}_i(t), \dot{z}_i(t), m_i\}, \quad (17)$$

where  $j = 1, \dots, 7$ . Note that we take the origin of the coordinates to be the centre of mass of the system, so that a constraint on the initial conditions is  $\sum_i m_i q_{i,j}(t_0) = 0$  for  $j = 1, \dots, 6$ , where  $q_{i,j}$  denotes the  $j$ th element of  $\mathbf{q}_i(t)$ .<sup>4</sup>

The coordinate system is right-handed, with the  $x$ -axis pointing to the right on the sky, then the  $y$ -axis points downwards, so that  $\hat{x} \times \hat{y} = \hat{z}$  points away from the observer, for unit vectors  $\{\hat{x}, \hat{y}, \hat{z}\}$  (Fig. 2).

### 4.4 Derivative of a combined Drift and Kepler step

The building block of this integrator is the universal Kepler solver for integrating pairs of bodies (Wisdom & Hernandez 2015), which we combined with a negative drift, before or after, described in Section 3.3. Standard WH  $N$ -body symplectic integrators (Wisdom & Holman 1991) use an elliptic (bound) Kepler solver for the ‘unperturbed’ motion, while the weaker interactions between low-mass or distant bodies are treated as impulses or kicks alternating with the Kepler-drifts. In the case of the DH17 integrator, a hyperbolic step is needed for the pairwise Keplerian integration of bodies that are unbound; this is used as an alternative to kicks. In this section we summarize the computation of the the Jacobian of the final relative coordinates with respect to the initial coordinates over a time-step with duration  $h$ , and

<sup>4</sup>In general, the centre-of-mass is allowed to move at a constant velocity, which is not implemented in our initial conditions, but could be if required.

then the transformation to the coordinates of the individual bodies.

The variational equations for the Cartesian coordinates depends on the ordering of the negative drift and Kepler step. For the combined negative drift followed by a Kepler step, the change in relative position and velocity is

$$\begin{aligned} \delta \Delta \mathbf{x}_{\text{DK}} &= (\hat{f} - 1) \delta \mathbf{x}_0 + (\hat{g} - h \hat{f}) \delta \mathbf{v}_0 \\ &\quad + \delta(\hat{f} - 1) \mathbf{x}_0 + \delta(\hat{g} - h \hat{f}) \mathbf{v}_0, \checkmark \end{aligned} \quad (18)$$

$$\begin{aligned} \delta \Delta \mathbf{v}_{\text{DK}} &= \hat{f} \delta \mathbf{x}_0 + (\hat{g} - h \hat{f} - 1) \delta \mathbf{v}_0 \\ &\quad + \delta \hat{f} \mathbf{x}_0 + \delta(\hat{g} - h \hat{f} - 1) \mathbf{v}_0, \checkmark \end{aligned} \quad (19)$$

while for a combined Kepler step followed by a negative drift, the variation of the change in relative coordinates is

$$\begin{aligned} \delta \Delta \mathbf{x}_{\text{KD}} &= (f - h \dot{f} - 1) \delta \mathbf{x}_0 + (g - h \dot{g}) \delta \mathbf{v}_0 \\ &\quad + \delta(f - h \dot{f} - 1) \mathbf{x}_0 + \delta(g - h \dot{g}) \mathbf{v}_0, \checkmark \end{aligned} \quad (20)$$

$$\begin{aligned} \delta \Delta \mathbf{v}_{\text{KD}} &= \dot{f} \delta \mathbf{x}_0 + (\dot{g} - 1) \delta \mathbf{v}_0 \\ &\quad + \delta \dot{f} \mathbf{x}_0 + \delta(\dot{g} - 1) \mathbf{v}_0, \checkmark \end{aligned} \quad (21)$$

where we have taken the differential of equations (15) and (16). The first line of each equation we have already computed, while the differentials of the Gauss functions in the second lines remain to be computed.

Each of the differential Gauss function terms involve the basis  $(\mathbf{x}_0, \mathbf{v}_0, k, h)$ , while each of these functions is defined in terms of  $G_i(\beta, \gamma)$ ,  $\beta$ ,  $\gamma$ ,  $r_0$ ,  $r$  and  $\eta_0$  (or  $\hat{G}_i(\hat{\beta}, \hat{\gamma})$ ,  $\hat{\beta}$ ,  $\hat{\gamma}$ ,  $\hat{r}_0$ ,  $\hat{r}$  and  $\hat{\eta}_0$ ). Note that if  $h$  is varied as a function of phase space, symplecticity is lost; however we accept a small symplecticity violation at a single time-step when searching for transit times. A similar choice was made in Deck et al. (2014). Thus, we first need to compute the differential of these Gauss function terms with respect to these intermediate quantities, and then propagate through these differentials using the chain rule to obtain the derivatives with respect to the basis. There is an extra step involved in the drift-first case (DK): since the functions on the right-hand side are defined in terms of  $\hat{\mathbf{x}}_0 = \mathbf{x}_0 - h \mathbf{v}_0$ , we also need to apply the chain rule to  $\hat{\mathbf{x}}_0$  to transform the derivatives to the basis.

The differentials of intermediate quantities are given in Appendix B1.

#### 4.4.1 Differential of drift-then-Kepler step

The differential of the scalar quantities  $\delta(\hat{f} - 1)$ ,  $\delta(\hat{g} - h \hat{f})$ ,  $\delta(\hat{f})$ , and  $\delta(\hat{g} - h \hat{f} - 1)$  should also be scalars, and can be expressed in terms similar to the  $\delta\gamma$  and  $\delta r$  terms given in Appendix B1. Note, however, that as the Kepler step takes place *after* the negative drift, all of these functions are to be computed in terms of  $\hat{\mathbf{x}}_0$  substituted for  $\mathbf{x}_0$ , and so we need to add an extra step in the derivation to find the differentials in terms of  $\mathbf{x}_0$  in lieu of  $\hat{\mathbf{x}}_0$ . The differential of these functions in terms of intermediate scalar quantities is given in appendix B2.

Substituting these differentials into equations (18) and (19), we arrive at expressions for  $\partial \Delta \mathbf{x}_{\text{DK}} / \partial \mathbf{x}_0$ ,  $\partial \Delta \mathbf{x}_{\text{DK}} / \partial \mathbf{v}_0$ ,  $\partial \Delta \mathbf{x}_{\text{DK}} / \partial k$ ,  $\partial \Delta \mathbf{v}_{\text{DK}} / \partial \mathbf{x}_0$ ,  $\partial \Delta \mathbf{v}_{\text{DK}} / \partial \mathbf{v}_0$ , and  $\partial \Delta \mathbf{v}_{\text{DK}} / \partial k$ . We also know  $\partial k / \partial \mathbf{x}_0 = 0$ ,  $\partial k / \partial \mathbf{v}_0 = 0$ , and  $\partial k / \partial k = 1$ , which we insert into a Jacobian,  $\mathbf{J}_{\text{kep}}$ , which is a  $7 \times 7$  matrix. In addition, we need the time derivatives of the coordinates with respect to the time-step,  $h$ ,  $\partial \Delta \mathbf{x}_{\text{DK}} / \partial h$ ,

$\partial \Delta \mathbf{v}_{\text{DK}} / \partial h$ , to obtain the time derivatives of the transit times with respect to the initial coordinates.

This completes the summary of the Jacobian of the combined negative drift then Kepler step for the variation of the relative coordinates of bodies  $i$  and  $j$  (recall that we dropped the  $ij$  subscripts in this section). In the next subsection, we discuss the results for the Kepler step followed by a negative drift.

#### 4.4.2 Differential of Kepler then drift step

The Kepler step followed by a negative drift is slightly simpler as the Gauss functions can be expressed in terms of  $\mathbf{x}_0$  rather than  $\hat{\mathbf{x}}_0$ . The differentials of the scalar functions are given in Appendix B3.

As with the prior combined step, the terms in these differentials may be inserted into a Jacobian,  $\mathbf{J}_{\text{kep}}$ , as well as give the derivatives with respect to  $h$ . This completes the description of the Jacobians computed for the combined drift and Kepler steps for the change in the relative coordinates between bodies  $i$  and  $j$ . This needs to be translated into the variations of the positions and velocities of the individual bodies  $i$  and  $j$ , which we describe next.

### 4.5 Jacobian of combined Kepler + drift step

The foregoing computation gives the variation in the *relative* difference between the positions and velocities of bodies  $i$  and  $j$ . This translates into variations in the positions of bodies  $i$  and  $j$  given by

$$\begin{aligned} \Delta \mathbf{x}_{i,\text{DK}} &= \frac{m_j}{m_i + m_j} \Delta \mathbf{x}_{\text{DK}}, \checkmark \\ \Delta \mathbf{x}_{j,\text{DK}} &= -\frac{m_i}{m_i + m_j} \Delta \mathbf{x}_{\text{DK}}, \checkmark \\ \Delta \mathbf{v}_{i,\text{DK}} &= \frac{m_j}{m_i + m_j} \Delta \mathbf{v}_{\text{DK}}, \checkmark \\ \Delta \mathbf{v}_{j,\text{DK}} &= -\frac{m_i}{m_i + m_j} \Delta \mathbf{v}_{\text{DK}}, \checkmark \end{aligned} \quad (22)$$

and likewise for  $\text{DK} \rightarrow \text{KD}$ , where  $\mathbf{x}_i(t+h) = \mathbf{x}_i(t) + \Delta \mathbf{x}_{i,\text{DK}}$  is carried out with compensated summation.

The Jacobian may be found by differentiating these equations with respect to the initial conditions of the Kepler-drift step, which is straightforward  $\checkmark$  for the position, velocity, and time-step derivatives. However, since this equation involves the masses  $m_i$  and  $m_j$ , the mass derivative of a combined Kepler/drift step involves an additional term, where

$$\begin{aligned} \frac{\partial \Delta \mathbf{x}_{i,\text{DK}}}{\partial m_i} &= -\frac{m_j}{(m_i + m_j)^2} \Delta \mathbf{x}_{\text{DK}} + \frac{Gm_j}{m_i + m_j} \frac{\partial \Delta \mathbf{x}_{\text{DK}}}{\partial k}, \checkmark \\ \frac{\partial \Delta \mathbf{x}_{i,\text{DK}}}{\partial m_j} &= \frac{m_i}{(m_i + m_j)^2} \Delta \mathbf{x}_{\text{DK}} + \frac{Gm_j}{m_i + m_j} \frac{\partial \Delta \mathbf{x}_{\text{DK}}}{\partial k}, \checkmark \end{aligned} \quad (23)$$

and the same equations apply for  $\mathbf{x} \rightarrow \mathbf{v}$ ,  $i \leftrightarrow j$ , and  $\text{DK} \rightarrow \text{KD}$ . The first of these two equations has a cancellation due to the difference in sign between the two terms on the right-hand side. Specifically,  $\Delta \mathbf{x}_{\text{DK}} \propto k$ , so there is a term in the derivative,  $\partial \Delta \mathbf{x}_{\text{DK}} / \partial k$ , which equals  $\Delta \mathbf{x}_{\text{DK}} / k$ , which exactly cancels the first term in the equation. We carry out this cancellation algebraically, thus avoiding roundoff errors which can occur when this term is much larger than the others in  $\partial \Delta \mathbf{x}_{\text{DK}} / \partial k$ . The second equation has both terms with the same sign, so this cancellation does not occur when the derivatives are with respect to the mass of the other body. Here we give the resulting derivatives:

$$\frac{\partial \Delta \mathbf{x}_{i,\text{DK}}}{\partial m_i} = \frac{G^2 m_j}{\hat{\beta} \hat{r}^3 \hat{r}_0^2} [J_1 \mathbf{x}_0 - J_2 \mathbf{v}_0], \checkmark \quad (24)$$

$$\frac{\partial \Delta \mathbf{v}_{i,\text{DK}}}{\partial m_i} = \frac{G^2 m_j}{\hat{\beta} \hat{r}^3 \hat{r}_0^2} [J_3 \mathbf{x}_0 + J_4 \mathbf{v}_0], \checkmark \quad (25)$$

where  $J_1 - J_4$  are functions given in Appendix C. Note that the derivatives of  $\Delta \mathbf{x}_{j,\text{DK}}$  and  $\Delta \mathbf{v}_{j,\text{DK}}$  with respect to  $m_j$  look identical save for replacing  $m_j$  with  $-m_j$ .

Similarly, the mass derivatives in the Kepler followed by drift step are given as

$$\frac{\partial \Delta \mathbf{x}_{i,\text{KD}}}{\partial m_i} = \frac{G^2 m_j}{\beta r^3 r_0^2} [J_5 \mathbf{x}_0 + J_6 \mathbf{v}_0], \checkmark \quad (26)$$

$$\frac{\partial \Delta \mathbf{v}_{i,\text{KD}}}{\partial m_i} = \frac{G^2 m_j}{\beta r^3 r_0^2} [J_7 \mathbf{x}_0 + r_0 J_8 \mathbf{v}_0], \checkmark \quad (27)$$

where  $J_5 - J_8$  are given in Appendix C. As above, the derivatives of  $\Delta \mathbf{x}_{j,\text{KD}}$  and  $\Delta \mathbf{v}_{j,\text{KD}}$  with respect to  $m_j$  look identical save for replacing  $m_j$  with  $-m_j$ . We place all of these derivatives into a Jacobian matrix for each drift + Kepler substep,  $\Delta \mathbf{J}_{\text{DK},ij}$  or  $\Delta \mathbf{J}_{\text{KD},ij}$ .

This completes the computation of the Jacobian of the drift plus Keplerian evolution of bodies  $i$  and  $j$  with respect to one another. Next, we describe the derivatives of the drift step applied at the start and end of each time-step.

### 4.6 Drift

The drift of an individual body is given by

$$\mathbf{x}_i(t+h) = \mathbf{x}_i(t) + h \mathbf{v}_i(t), \quad (28)$$

$$\mathbf{v}_i(t+h) = \mathbf{v}_i(t). \quad (29)$$

This has the straightforward differential of

$$\delta \mathbf{x}_i(t+h) = \delta \mathbf{x}_i(t) + h \delta \mathbf{v}_i(t) + \delta h \mathbf{v}_i(t), \quad (30)$$

$$\delta \mathbf{v}_i(t+h) = \delta \mathbf{v}_i(t). \quad (31)$$

There are two stages at which the drifts are applied: all particles drift at the start and end of each AHL21 step with a duration  $h/2$  (see algorithm 2). We refer to this as  $\mathbf{I} + \Delta \mathbf{J}_{\text{D}}(h)$  for drifting all of the planets. In some cases, it proves to be faster and sufficiently accurate to use instantaneous kicks between pairs of bodies rather than solving the Kepler problem; we now turn to describing this option.

### 4.7 Derivative of kicks

Hernandez & Bertschinger (2015) show that for some pairs of particles (typically distant or unbound), sufficient accuracy may be obtained by applying a gravitational kick between particles, rather than a Keplerian step and negative drift. Letting  $A$  be the set of pairs  $(i, j)$  advanced with drift + Kepler steps, then  $A^C$  is the complementary set which receives pairwise kicks such that  $A \cap A^C = \emptyset$ . Note that if all pairs are in  $A^C$ , the integrator becomes leapfrog.

Algorithm 2 implements this method by applying the pairwise kicks (to set  $A^C$ ) for a time-step  $h/6$  before the initial drifts, then after the combined drift-Kepler is applied to set  $A$ , a second set of kicks is applied for a time-step  $2h/3$  along with separate correction terms for the pairs in  $A$  and  $A^C$ , and then after the second Kepler-drift step is applied to  $A$  in reverse order, there is a final set of pairwise kicks applied to  $A^C$  for a time  $h/6$  after the final drifts (on set  $A^C$ ).

For a pair of particles  $i$  and  $j$ , the kicks applied over a time-step  $h$  are given by

$$\begin{aligned}\Delta \mathbf{v}_i &= -h \frac{Gm_j \mathbf{x}_{ij}}{r_{ij}^3}, \checkmark \\ \Delta \mathbf{v}_j &= h \frac{Gm_i \mathbf{x}_{ij}}{r_{ij}^3}, \checkmark\end{aligned}\quad (32)$$

where, as above,  $\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$ ,  $r_{ij} = |\mathbf{x}_{ij}|$ , which has derivatives given by

$$\delta \Delta \mathbf{v}_i = -h \frac{Gm_j}{r_{ij}^5} \left[ \mathbf{x}_{ij} \frac{\delta m_j}{m_j} r_{ij}^2 + \mathbf{w}_{ij} \right], \checkmark \quad (33)$$

$$\delta \Delta \mathbf{v}_j = h \frac{Gm_i}{r_{ij}^5} \left[ \mathbf{x}_{ij} \frac{\delta m_i}{m_i} r_{ij}^2 + \mathbf{w}_{ij} \right], \checkmark \quad (34)$$

$$\mathbf{w}_{ij} = \delta \mathbf{x}_{ij} r_{ij}^2 - 3 \mathbf{x}_{ij} \mathbf{x}_{ij} \cdot \delta \mathbf{x}_{ij}. \checkmark \quad (35)$$

These differentials yield a Jacobian for the kicks between all pairs of bodies  $(i, j) \in A^C$ . We now move on to describing the derivatives of the fourth-order correction that is used to improve the order of the algorithm.

#### 4.8 Derivative of correction

Dehnen & Hernandez (2017) reduced the error in the Hernandez & Bertschinger (2015) mapping, obtaining a symplectic integrator accurate to  $h^4$ . We incorporate this correction into our integrator, using  $\alpha = 0$  (equation 40 of Dehnen & Hernandez 2017) that only requires one call of the corrector in between the sequences of binary drift-Kepler and Kepler-drift steps (in the middle of algorithm 2). Two corrections need to be computed: one for the pairs in  $A$ , and one for those in  $A^C$ . We describe these in the next two subsections.

##### 4.8.1 Drift + Kepler pairs correction ( $A$ )

The first correction is applied to the velocities of the bodies that are treated with the Kepler + drift splitting, with an impulse term for the  $i$ th body in  $A$  of

$$\Delta \mathbf{v}_i = \frac{h^3}{24} \sum_{i,j \in A} \frac{Gm_j}{r_{ij}^5} \mathbf{T}_{ij}, \checkmark \quad (36)$$

$$\mathbf{T}_{ij} = \mathbf{x}_{ij} \left( \frac{2Gm}{r_{ij}} + 3\mathbf{a}_{ij} \cdot \mathbf{x}_{ij} \right) - r_{ij}^2 \mathbf{a}_{ij}, \quad (37)$$

where  $m = m_i + m_j$ ,  $\mathbf{a}_{ij} = \mathbf{a}_i - \mathbf{a}_j$ , and

$$\mathbf{a}_i = - \sum_{i,j \in A} \frac{Gm_j}{r_{ij}^3} \mathbf{x}_{ij}. \checkmark \quad (38)$$

Note that the sum is only taken over the particles in  $A$ , and no correction is required for the positions. We will define a constant  $C = (Gh^3)/24$  in what follows.

The derivative of this correction term can be computed in two steps, first computing the derivative of  $\mathbf{a}_i$ , and then the derivative of  $\Delta \mathbf{v}_i$ ,

$$\delta \mathbf{a}_i = - \sum_{i,j \in A} \frac{Gm_j}{r_{ij}^5} \left[ \mathbf{x}_{ij} \frac{\delta m_j}{m_j} r_{ij}^2 + \delta \mathbf{x}_{ij} r_{ij}^2 - 3 \mathbf{x}_{ij} \mathbf{x}_{ij} \cdot \delta \mathbf{x}_{ij} \right], \checkmark \quad (39)$$

$$\begin{aligned}\delta \Delta \mathbf{v}_i &= C \sum_{i,j \in A} \left[ \frac{\delta m_j}{r_{ij}^5} - \frac{m_j}{r_{ij}^6} \frac{5 \mathbf{x}_{ij} \cdot \delta \mathbf{x}_{ij}}{r_{ij}} \right] \mathbf{T}_{ij} \\ &+ C \sum_{i,j \in A} \frac{m_j}{r_{ij}^5} \delta \mathbf{T}_{ij}, \checkmark\end{aligned}\quad (40)$$

with

$$\begin{aligned}\delta \mathbf{T}_{ij} &= \delta \mathbf{x}_{ij} \left( \frac{2Gm}{r_{ij}} + 3\mathbf{a}_{ij} \cdot \mathbf{x}_{ij} \right) \\ &+ \frac{2Gm \mathbf{x}_{ij}}{r_{ij}} \left( \frac{\delta m_i + \delta m_j}{m} - \frac{\mathbf{x}_{ij} \cdot \delta \mathbf{x}_{ij}}{r_{ij}^2} \right) \\ &+ \mathbf{x}_{ij} (3 \mathbf{x}_{ij} \cdot \delta \mathbf{a}_{ij} + 3 \mathbf{a}_{ij} \cdot \delta \mathbf{x}_{ij}) \\ &- 2(\mathbf{x}_{ij} \cdot \delta \mathbf{x}_{ij}) \mathbf{a}_{ij} - r_{ij}^2 \delta \mathbf{a}_{ij}. \checkmark\end{aligned}\quad (41)$$

When implementing these equations as computer code, we pre-compute and store the dot products to save computational time.

##### 4.8.2 Correction for fast-kick pairs ( $A^C$ )

The pairs in  $A^C$  also require a correction, but with a slightly simpler relation:

$$\Delta \mathbf{v}_i = \frac{h^3}{36} \sum_{i,j \in A^C} \frac{Gm_j}{r_{ij}^5} [3 \mathbf{x}_{ij} (\mathbf{a}_{ij} \cdot \mathbf{x}_{ij}) - \mathbf{a}_{ij} r_{ij}^2], \quad (42)$$

where the sum is taken only over pairs in  $A^C$ . The derivatives are computed in a manner similar to that described in the prior subsection.

The overall Jacobian for this step is given by  $\mathbf{I} + \Delta \mathbf{J}_{4th}(h)$ , which is the identity matrix for the position and mass component, and is given by  $\partial \Delta \mathbf{v}_i / \partial \mathbf{x}_j$  for the offdiagonal components relating bodies  $i$  and  $j$ . The time derivatives are straightforward as they involve derivatives with respect to  $h^3$ , and so involve the same formulae multiplied by  $3/h$ . With the Jacobians now defined for each component of a time-step, we next describe how we combine these into the Jacobian of a full time-step.

#### 4.9 Jacobian of a time-step

With the Jacobian transformations computed at each step of algorithm 2, we can now compute the complete derivative of each transit time with respect to the initial conditions, keeping track of the product of Jacobians throughout algorithm 2. Now, in each case we compute the change in the coordinates over a time-step, and so the Jacobian of each substep has the form:

$$\mathbf{J}_{\text{substep}} = \mathbf{I} + \Delta \mathbf{J}_{\text{substep}}, \quad (43)$$

where  $\Delta \mathbf{J}_{\text{substep}}$  is the Jacobian of the change in coordinates at the end of the substep with respect to the coordinates at the beginning of the substep. Consequently, the Jacobian can be written as

$$\begin{aligned}\mathbf{J}_{\text{current}} &= (\mathbf{I} + \Delta \mathbf{J}_{\text{substep}}) \mathbf{J}_{\text{prior}} \\ &= \mathbf{J}_{\text{prior}} + \Delta \mathbf{J}_{\text{substep}} \mathbf{J}_{\text{prior}}. \checkmark\end{aligned}\quad (44)$$

The propagation of the Jacobian involves adding terms to the prior Jacobian as a function of each substep. Now, when the time-step is small, this involves very small additions to the Jacobian which can increase the impact of round-off error during the propagation of the derivatives. To mitigate the impact of this, we use compensated summation (Kahan 1965).

As an example, in the 3-body case, an individual AHL21 step looks like

$$\begin{aligned} \mathbf{J}_{\text{AHL21}}(h) &= \left( \mathbf{I} + \Delta \mathbf{J}_{\text{D}} \left( \frac{h}{2} \right) \right) \left( \mathbf{I} + \Delta \mathbf{J}_{\text{DK},12} \left( \frac{h}{2} \right) \right) \\ &\times \left( \mathbf{I} + \Delta \mathbf{J}_{\text{DK},13} \left( \frac{h}{2} \right) \right) \left( \mathbf{I} + \Delta \mathbf{J}_{\text{DK},23} \left( \frac{h}{2} \right) \right) \\ &\times \left( \mathbf{I} + \Delta \mathbf{J}_{4\text{th}}(h) \right) \\ &\times \left( \mathbf{I} + \Delta \mathbf{J}_{\text{KD},23} \left( \frac{h}{2} \right) \right) \left( \mathbf{I} + \Delta \mathbf{J}_{\text{KD},13} \left( \frac{h}{2} \right) \right) \\ &\times \left( \mathbf{I} + \Delta \mathbf{J}_{\text{KD},12} \left( \frac{h}{2} \right) \right) \left( \mathbf{I} + \Delta \mathbf{J}_{\text{D}} \left( \frac{h}{2} \right) \right). \end{aligned} \quad (45)$$

In this example, we do not use fast kicks for any pair of bodies. Note that each Jacobian in this product is computed with the updated state of the system from the prior substep.

We have also implemented a version of the integrator which allows the drift + Kepler interactions to be replaced with kicks for some subset of pairs bodies (Section 4.7). For this version of the integrator an additional three Jacobians must be multiplied.

After taking  $n + 1$  steps whereby the first transit occurs in between steps  $n$  and  $n + 1$ , a substep is taken to find the intermediate time,  $\Delta t = t - nh - t_0$ , which minimizes the sky separation between the two bodies at time  $t$ . To compute the derivatives of the times of transit requires computing the Jacobian of a step with intermediate time  $\Delta t$ , which is used to compute the derivatives of the time of transit with respect to the initial conditions, which we describe next.

#### 4.10 Time derivative of a step

The derivative of the positions and coordinates as a function of the time-step duration,  $\Delta t$ , requires a propagation of the time-step derivative through all of the substeps of a single time-step. The involves applying the chain rule through each of the sub-steps in algorithm 2. Let  $\mathbf{q}_{\text{current}} = \Delta \mathbf{q} + \mathbf{q}_{\text{prior}}$  be the coordinates and velocities of all bodies after applying one component of a substep. Then,

$$\frac{d\mathbf{q}_{\text{current}}}{d\Delta t} = \left( \mathbf{I} + \Delta \mathbf{J}_{\text{substep}} \right) \frac{d\mathbf{q}_{\text{prior}}}{d\Delta t} + \left. \frac{\partial \Delta \mathbf{q}}{\partial \Delta t} \right|_{\mathbf{q}_{\text{prior}}}, \quad (46)$$

where  $\frac{\partial \Delta \mathbf{q}}{\partial \Delta t}$  is the partial derivative of a particular sub-step with respect to the time-step  $\Delta t$ . Note that in the AHL21 algorithm (2), the combined drift and Kepler steps take place over a time  $\Delta t = h/2$ , which introduces a factor of 1/2 in the partial derivatives with respect to  $h$ .

At the end of the step we refer to the derivative over the time-step with respect to  $\Delta t$  as

$$\frac{d\mathbf{q}(t)}{d\Delta t}, \quad (47)$$

where  $t$  is the total simulation time upon completion of the time-step of duration  $\Delta t$ . This may then be used to compute the transit times and their derivatives, as described next.

#### 4.11 Derivative of transit times

We define the times of transit between bodies  $i$  and  $j$  as the point in time where the sky-projected separation is at a minimum, and body  $i$  is in front of body  $j$  (Fabrycky 2010). Since multiple transits can occur between two bodies, we count these with a third index,  $k$ , so that the set of transit times during the time integration is given by  $\{t_{ijk}; \forall i, j, k\}$ . At a transit time, the sky-velocity between the two bodies

must be perpendicular to their sky separation, where the ‘sky plane’ is the  $x$ – $y$  plane; this guarantees an extremum of the sky separation between the bodies. The dot product of the relative sky separation and sky velocity of the two bodies equals zero at the time of transit, and is negative/positive just before/after transit. So, transit times are computed from the constraints

$$\begin{aligned} g_{\text{sky},ij}(t_{ijk}) &= (x_i - x_j)(v_{x,i} - v_{x,j}) + (y_i - y_j)(v_{y,i} - v_{y,j}) = 0, \checkmark \\ z_i &< z_j \\ \frac{dg_{\text{sky},ij}}{dt} &> 0, \end{aligned} \quad (48)$$

where  $i$  is the index of the planet, and  $j$  is the index of the star (Fabrycky 2010), and  $k$  is an index for the number of transits between the bodies.

Throughout the time-integration of a system, transits between a planet and star (or any pair of bodies) are checked for by identifying when  $g_{\text{sky},ij}(t)$  changes sign from negative to positive between two time-steps, and the planet (or occulter) is nearer to the observer than the star. Once a transit time has been identified as occurring between time-steps  $n$  and  $n + 1$ , where  $t_n = t_0 + nh$ , by the condition  $g_{\text{sky},ij}(t_n) < 0$  and  $g_{\text{sky},ij}(t_{n+1}) > 0$  and  $z_i(t_n) < z_j(t_n)$ , then the time of transit is solved for with Newton’s method, which makes use of our Jacobian calculation. Newton’s method is applied to obtain the time  $t_{ijk} = t_n + \Delta t$ , where  $\Delta t$  is the time after  $t_n$  at which  $g_{\text{sky},ij} = 0$ , which is taken as the time of transit. The initial guess for the time of transit,  $\Delta t_{\text{init}}$ , is found by linear interpolation:

$$\Delta t_{\text{init}} = - \frac{g_{\text{sky},ij}(t_n)h}{g_{\text{sky},ij}(t_{n+1}) - g_{\text{sky},ij}(t_n)}. \quad (49)$$

To implement Newton’s method, the system is integrated in between these time-steps with a single AHL21 step, but with a time  $\Delta t < h$ , instead of  $h$ , giving  $\mathbf{q}(t_n + \Delta t)$ . From these coordinates,  $g_{\text{sky}}(t_n + \Delta t)$  is computed between the two bodies, and refined using

$$\delta \Delta t = -g_{\text{sky}} \left( \frac{dg_{\text{sky}}}{dt_{ijk}} \right)^{-1}, \quad (50)$$

where

$$\begin{aligned} \frac{dg_{\text{sky}}}{dt_{ijk}} &= x_{ij} \left( \frac{dv_{x,i}}{d\Delta t} - \frac{dv_{x,j}}{d\Delta t} \right) + y_{ij} \left( \frac{dv_{y,i}}{d\Delta t} - \frac{dv_{y,j}}{d\Delta t} \right) \\ &+ v_{x,ij} \left( \frac{dx_i}{d\Delta t} - \frac{dx_j}{d\Delta t} \right) + v_{y,ij} \left( \frac{dy_i}{d\Delta t} - \frac{dy_j}{d\Delta t} \right), \end{aligned} \quad (51)$$

and  $x_{ij} = x_i - x_j$ , etc. and the time derivatives with respect to  $\Delta t$  are computed with equation (47). Note that in practice since the integration time-step,  $h$ , is fixed, for the transit time derivatives  $\delta t_{ijk} = \delta(\Delta t)$ .

Once a transit time is found, how does it vary with the initial conditions? We focus on the initial conditions just before transit at time  $t_n$ ,  $\mathbf{q}_n = \mathbf{q}(t_n)$ . If  $\mathbf{q}_n$  is perturbed slightly, then the time of transit will change, but the new time of transit must still satisfy  $g_{\text{sky},ij}(t_{ijk} + \delta \Delta t) = 0$ , where the new time of transit is at  $t_n + \Delta t + \delta(\Delta t)$ . So,

$$\frac{\partial g_{\text{sky}}}{\partial \mathbf{q}_n} \delta(\mathbf{q}_n) + \frac{dg_{\text{sky}}}{dt} \delta(\Delta t) = 0, \quad (52)$$

where we have dropped the  $i, j, k$  subscripts from  $t$  in this equation.

Thus, the gradient of each transit time with respect to the state,  $\mathbf{q}_n = \mathbf{q}(t_n)$ , at the beginning of the  $n$ th time-step just preceding the

transit is given by

$$\begin{aligned} \frac{d(\Delta t)}{d\mathbf{q}_n} = & - \left( \frac{d\mathbf{g}_{\text{sky}}}{dt_{ijk}} \right)^{-1} \left[ \left( \frac{\partial x_i}{\partial \mathbf{q}_n} - \frac{\partial x_j}{\partial \mathbf{q}_n} \right) (v_{x,i} - v_{x,j}) \right. \\ & + \left( \frac{\partial v_{x,i}}{\partial \mathbf{q}_n} - \frac{\partial v_{x,j}}{\partial \mathbf{q}_n} \right) (x_i - x_j) \\ & + \left( \frac{\partial y_i}{\partial \mathbf{q}_n} - \frac{\partial y_j}{\partial \mathbf{q}_n} \right) (v_{y,i} - v_{y,j}) \\ & \left. + \left( \frac{\partial v_{y,i}}{\partial \mathbf{q}_n} - \frac{\partial v_{y,j}}{\partial \mathbf{q}_n} \right) (y_i - y_j) \right], \checkmark \end{aligned} \quad (53)$$

where the gradients are computed over the partial time-step, so that, for example,  $\frac{\partial x_j}{\partial \mathbf{q}_n}$  is the component of  $\mathbf{J}_{\text{AHL21}}(\Delta t)$  associated with the  $x$ -component of body  $j$ . Note again that the transit time is  $t_{ijk} = t_0 + nh + \Delta t$ , but since  $h$  and  $t_0$  are fixed,  $dt_{ijk}/d\mathbf{q}_n = d\Delta t/d\mathbf{q}_n$ .

In addition to the transit time,  $t_{ijk}$ , it is also useful to compute the sky velocity,  $v_{\text{sky},ijk}$ , and the impact parameter squared,  $b_{\text{sky},ijk}^2$ , at the time of transit. These can be used to compute transit light curves, as well as measure the variation of the impact parameter and duration as a function of time as additional dynamical constraints on a system. These two quantities are given by

$$v_{\text{sky},ijk}(\mathbf{q}_n, t_{ijk}) = \sqrt{(v_{x,i} - v_{x,j})^2 + (v_{y,i} - v_{y,j})^2}, \checkmark \quad (54)$$

$$b_{\text{sky},ijk}^2(\mathbf{q}_n, t_{ijk}) = (x_i - x_j)^2 + (y_i - y_j)^2, \checkmark \quad (55)$$

where there is a direct dependence upon  $\mathbf{q}_n$  which is propagated to the time of transit within the time-step,  $\Delta t$ , and there is an indirect dependence upon  $\mathbf{q}_n$  through the fact that these are evaluated at the time of transit,  $t_{ijk}(\mathbf{q}_n)$ .

Taking the derivative of these with respect to  $\mathbf{q}_n$  gives

$$\frac{dv_{\text{sky},ijk}}{d\mathbf{q}_n} = \frac{\partial v_{\text{sky},ijk}}{\partial \mathbf{q}_n} + \frac{dv_{\text{sky},ijk}}{d\Delta t} \frac{\partial \Delta t}{\partial \mathbf{q}_n}, \checkmark \quad (56)$$

$$\begin{aligned} \frac{\partial v_{\text{sky},ijk}}{\partial \mathbf{q}_n} = & v_{\text{sky},ijk}^{-1} \left[ (v_{x,i} - v_{x,j}) \left( \frac{\partial v_{x,i}}{\partial \mathbf{q}_n} - \frac{\partial v_{x,j}}{\partial \mathbf{q}_n} \right) \right. \\ & \left. + (v_{y,i} - v_{y,j}) \left( \frac{\partial v_{y,i}}{\partial \mathbf{q}_n} - \frac{\partial v_{y,j}}{\partial \mathbf{q}_n} \right) \right], \checkmark \end{aligned} \quad (57)$$

$$\begin{aligned} \frac{dv_{\text{sky},ijk}}{d\Delta t} = & v_{\text{sky},ijk}^{-1} \left[ (v_{x,i} - v_{x,j}) \left( \frac{dv_{x,i}}{d\Delta t} - \frac{dv_{x,j}}{d\Delta t} \right) \right. \\ & \left. + (v_{y,i} - v_{y,j}) \left( \frac{dv_{y,i}}{d\Delta t} - \frac{dv_{y,j}}{d\Delta t} \right) \right], \checkmark \end{aligned} \quad (58)$$

$$\begin{aligned} \frac{db_{\text{sky},ijk}^2}{d\mathbf{q}_n} = & 2 \left[ (x_i - x_j) \left( \frac{\partial x_i}{\partial \mathbf{q}_n} - \frac{\partial x_j}{\partial \mathbf{q}_n} \right) \right. \\ & \left. + (y_i - y_j) \left( \frac{\partial y_i}{\partial \mathbf{q}_n} - \frac{\partial y_j}{\partial \mathbf{q}_n} \right) \right], \checkmark \end{aligned} \quad (59)$$

Note that we compute  $b_{\text{sky}}^2$  rather than  $b_{\text{sky}}$  to avoid the problem that when the orbits are edge-on, the impact parameter is zero at mid-transit, causing the derivative of  $b_{\text{sky}}$  to be divided by  $b_{\text{sky}} = 0$ , which results in a NaN.

This completes the computation of all of the Jacobians needed to propagate the derivatives of the transit times, and sky velocity/impact parameter, through to the initial conditions, which we describe next.

#### 4.12 Jacobians of positions, velocities, transit times

With the Jacobians computed at each of the steps, we can recursively compute the Jacobian at step  $n$  with  $t = t_0 + nh$  as

$$\mathbf{J}_n = \mathbf{J}_{\text{AHL21}}(h) \mathbf{J}_{n-1}. \quad (60)$$

Starting with the initial state  $\mathbf{q}_0 = \mathbf{q}(t_0)$  and initial Jacobian  $\mathbf{J}_0 = \frac{\partial \mathbf{q}_0}{\partial \mathbf{q}_0} = \mathbf{I}$  (the identity matrix), we iteratively compute the Jacobian at step  $n$  with respect to the state at initial time,  $t_0$  ( $n = 0$ ), giving the Jacobian transformation from  $\mathbf{q}_0$  to  $\mathbf{q}_n$ ,

$$\mathbf{J}_n = \frac{d\mathbf{q}_n}{d\mathbf{q}_0}. \quad (61)$$

Then, the gradient of the transit times is given by

$$\frac{d\Delta t}{d\mathbf{q}(t_0)} = \frac{d\Delta t}{d\mathbf{q}(t_n)} \mathbf{J}_n. \quad (62)$$

We save this gradient for each transit time in an array that is pre-allocated when calling the routine.

In our implementation, we do not compute  $\mathbf{J}_{\text{AHL21}}(h)$  for each step, directly; instead, we iteratively multiply the current Jacobian by the Jacobian for each sub-step.

This completes the description of the algorithm and its derivatives. We now turn to the implementation and testing of the code.

## 5 IMPLEMENTATION AND TESTING

We have developed `NbodyGradient.jl`<sup>5</sup> in the JULIA language for carrying out the foregoing computations. This involves the initialization of the algorithm, the  $N$ -body integration, the finding of transit times, and the Jacobian propagation. Given the complicated nature of the calculations, we have written unit tests for each of the steps in the algorithm; these were critical in developing the code for computing the derivatives, and helped to pinpoint inaccuracies in the DH17 algorithm which led to developing the AHL21 algorithm. We have also created tests of the code as a whole, and carried out comparisons with other codes for both speed and accuracy, which are summarized here.

In this section, we describe some aspects of the implementation of the code (Section 5.1) and the tests we have carried out. We test the  $N$ -body algorithm for accuracy by varying the step size and checking for conservation of energy and angular momentum (Section 5.2), while we check the transit-time algorithm for accuracy by measuring the variation in transit times with step size (Section 5.4). We compare the  $N$ -body integrator with a C implementation to check for speed (Section 5.3). We check the numerical precision of the code by carrying out comparisons with extended precision (Section 5.5), and we check the accuracy of the derivatives by comparing with finite-differences carried out in extended precision (Section 5.5). Most of our tests are carried out with integrations of the outer Solar system and of the TRAPPIST-1 system.

We start by describing the implementation of the algorithm in JULIA.

### 5.1 JULIA implementation

We chose the JULIA language to develop this code (Bezanson et al. 2017), given its several advantages. The high-level, interactive (REPL) capability can make debugging code more straightforward. The just-in-time compiler can make the code execution competitive with compiled C, if attention is paid to memory allocation and type stability. An advantage of JULIA for testing code accuracy is that different numerical types can easily be changed which allows for straightforward computation at different precisions. JULIA also uses multiple dispatch that allows us to automatically select versions

<sup>5</sup><http://github.com/ericagol/NbodyGradient.jl>

of functions that match in precision, and gives us control over computing gradients, transit times, and other outputs. Finally, JULIA is open-source, and thus amenable to distribution and usage amongst scientists.

We have optimized the code keeping in mind several unique aspects of JULIA. Memory allocation and garbage collection were minimized by defining arrays at higher levels that were then passed to subroutine functions to avoid repeated allocation of large arrays. Matrix multiplication can be sped up by utilizing the BLAS linear algebra routines (Blackford et al. 2002), specifically `gemmv`, which gave significant reduction in run-time for the multiplication of Jacobians at each step. For multiplication of the  $J_{DK/KD,ij}$  Jacobian, we found more efficiency by copying the portion of  $J_n$  (times the prior substeps) relevant to bodies  $i$  and  $j$  to a new  $14 \times 7N$  matrix, then using the BLAS routine to carry out the multiplication, and then copying the result back into  $J_n$ . For loops in which we access elements of arrays successively, we try to step through elements which are adjacent in memory, and we also avoid index-checking to save time. Finally, we try to avoid changing the types of variables, and we define the types up front to make this explicit. Thanks to these details, we find a favourable run-time comparison with C (Section 5.3).

Another aspect of our implementation is that the code is simple to use and extendable. Here is an example of running the integrator and computing transit times interactively from the JULIA prompt (REPL), or from within a Jupyter or Pluto notebook<sup>6</sup> (a slightly modified version of the script used in the comparisons with other codes in Section 6):

---

```
using NbodyGradient

# Set up initial conditions from file of orbital elements
# (Initial time, # of bodies, orbital elements file)
ic = ElementsIC(0.0, 8, "elements.txt")
# Time-step. Period of planet b / 100
h = ic.elements[2,2]/100
# Set up integrator (time-step, initial time, elapsed time)
intr = Integrator(h, 0.0, 4533.0)
# Compute and store initial Cartesian coordinates
s = State(ic)
# Allocate arrays for transit times and derivatives
tt = TransitTiming(intr.tmax, ic)
# Run integration & compute transit times w/ derivatives.
intr(s, tt)
```

---

Here, the initial conditions are specified by a file containing rows of orbital elements, `elements.txt`.<sup>7</sup> An integration is triggered by

<sup>6</sup>Further details on running the code can be found in the documentation at <http://github.com/ericagol/NbodyGradient.jl>.

<sup>7</sup>We leave discussion of initial conditions to future work, but for sake of completeness the elements file is set up as follows. Each row is given by the mass, period, time of initial transit, eccentricity vector components, Inclination, and Long. of Ascending Node for each body. The columns are delimited by a comma (`,`). The eccentricity vector is defined by  $(\cos(\varpi), \sin(\varpi))$  where  $e$  is the eccentricity and  $\varpi$  is the long. of periastron. The orbital elements are given in Jacobi coordinates which are converted to Cartesian coordinates to start the integration.

passing a `State` type (`s` in the example) to an `Integrator` (`intr` in the example), along with any ‘output’ type (`tt` in the example): `intr(s, tt)`. The `State` holds the Cartesian coordinates and Jacobian which are updated at each step. Passing the output structure `tt` of type `TransitTiming` tells the integrator to compute transit times of the system and store the results in the `tt` structure. The transit times can be accessed within the structure as `tt.tt`, which is a two dimensional array of size  $N$  by  $N_t$ , which holds the transit times for each planet, each of which have a count in the vector `tt.count` with a maximum allowed value of  $N_t$ . The derivatives with respect to the initial Cartesian coordinates and masses are stored as `tt.dtdq0` which is a 4-dimensional array with the same first two dimensions as `tt.tt`, and the last two dimensions of sizes 7 and  $N$  which hold the derivatives with respect to  $x$ ,  $v$ , and  $m$  for each body ( $\mathbf{q}_0$ ). By utilizing multiple dispatch, adding functionality to the code consists of simply making a new `Integrator` method and a structure to hold the output.

Next, we describe the accuracy of the  $N$ -body algorithm by checking conservation of energy and angular momentum.

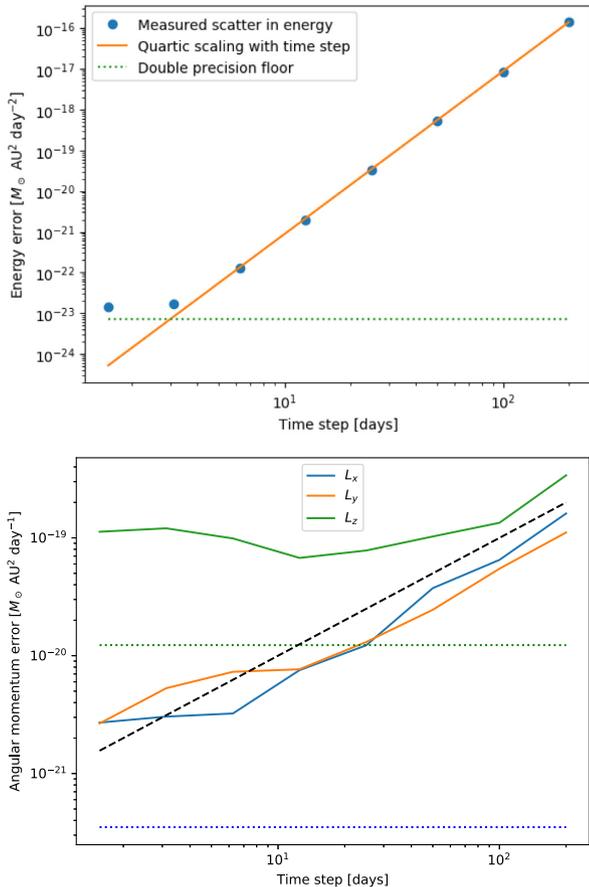
## 5.2 Energy and angular momentum conservation

To test the accuracy of the algorithm, we have carried out integrations of the outer Solar system. We start with positions, velocities, and masses given in Hairer et al. (2006), only including the giant planets (Jupiter, Saturn, Uranus, and Neptune). The mass of the Sun is added to the sum of the masses of the terrestrial planets for a fifth inner body.

We compute the total energy and angular momentum of the system as a function of time. We measured the RMS value for the energy and all three components of the total angular momentum, and varied the time-steps by factors of 2. We expect the energy precision to scale with time-step to the fourth power,  $\alpha h^4$ . Fig. 3 shows that this scaling holds over a range of two orders of magnitude in the time-step. We used time-steps from 1.5625 to 200 d, and the RMS energy and angular momentum was measured over  $\approx 10^6$  time-steps in each case. The upper end of the time-steps was set by the requirement that the time-step be smaller than 1/20 of the shortest orbital period, in this case Jupiter. At the lower end of this range (1.5625 and 3.125 d) we see a deviation from the  $h^4$  scaling thanks to the limit of double-precision representation of the energy. This occurs at approximately  $2^{-52} = 2.2 \times 10^{-16}$  of the absolute error value of each conserved quantity, which is plotted as dotted lines in each panel in Fig. 2 (the energy is  $\approx -3 \times 10^{-8} M_\odot \text{au}^2 \text{d}^{-2}$ , while the total angular momentum is  $\approx 6 \times 10^{-5} M_\odot \text{au}^2 \text{d}^{-1}$ ). The three components of angular momentum show a better conservation precision that is close to double precision for all values of the time-step. Numerical errors accumulate with time-step, and the expectation is that these scale as  $\approx \epsilon_{\text{err}} h^{1/2} t^{1/2} = \epsilon_{\text{err}} h N_S^{1/2}$ , where  $N_S$  is the number of time-steps and  $\epsilon_{\text{err}}$  is a random numerical error (Hairer, McLachlan & Razakarivony 2008). This is also borne out in Fig. 3 that shows a scaling of the error with time-step  $h$  ( $N_S$  is held fixed in these integrations).

Our conclusion is that the numerical integration is behaving as expected: energy is conserved with an accuracy  $\alpha h^4$  above the double-precision limit, and angular momentum is conserved close to double precision, but grows according to Brouwer’s law. Note that the RMS relative error (defined as  $\Delta E/E_0$ , with  $E_0$  the initial energy, and  $\Delta E$  the change in this energy) measures an oscillation that can be orders of magnitude larger than the mean relative error over time.

Given this evidence of accurate behaviour of the  $N$ -body algorithm, we next ask: how does the  $N$ -body implementation fare in computation time?



**Figure 3.** (Top) Conservation of energy. Standard deviation of energy versus step size (the blue dots). The orange line shows  $h^4$  scaling. The green-dotted line shows double-precision limit. (Bottom) Angular momentum conservation. The standard deviation of each component is plotted versus step size. The dotted lines show the double-precision limit for angular momentum ( $L_y$  is zero, so this is not shown). The black-dashed curve shows that angular momentum error scales  $\propto h$  according to Brouwer's law.

### 5.3 Comparing with C implementation

To check that we have optimized the computational speed of the  $N$ -body integrator, we carried out a comparison of the JULIA version of our code with a C implementation without derivatives (Hernandez 2016). First, we compared the Kepler solver (Wisdom & Hernandez 2015) and found that our JULIA implementation matches the C version. Both versions take  $0.15 \mu\text{s}$  per Kepler step for a bound orbit with  $e = 0.5$ .<sup>8</sup> Note that in this comparison we used a version of the Kepler step which is not combined with a backward drift.

Next, we carried out an integration of the outer Solar system (with five bodies, as described in the prior section) with the C implementation. We found that the C implementation runs at the same speed as our JULIA implementation of AHL21. With 50-d time-steps, both versions take about  $4.7 \mu\text{s}$  per time-step.<sup>9</sup> In this comparison, we use the same convergence criterion for the Kepler solver; when we include the fourth-order corrector in AHL21 it increases the run time by  $\approx 10$  percent for the outer Solar

<sup>8</sup>These comparisons were made on a Macbook with a 2.8 GHz Intel Core i7 processor with JULIA v1.6. The C code was compiled with `cc -O3`.

<sup>9</sup>See footnote 8.

system problem. Thus, we conclude that the speed of the JULIA implementation is comparable to compiled C.

When using the convergence of a fractional tolerance of  $10^{-8}$  is reached for the solution to Kepler's equation – we find a bias in the long-term energy conservation that causes it to drift with time. If instead we use the convergence criterion that the eccentric-anomaly,  $\gamma$ , remains unchanged relative to one of the prior two iterations of Newton's solver – then we find that this bias is significantly reduced. This adds iterations to the Kepler-solver, typically 1-2, and thus causes the code to take about 10 percent longer to run, but with the trade-off of better energy conservation. Thus, using the fourth-order corrector and this criterion adds about 20 percent to the overall runtime, amounting to  $5.7 \mu\text{s}$  per time-step compared with the example above.

Now that we have verified the speed and accuracy of the  $N$ -body algorithm, we next examine the accuracy and precision of the transit times as a function of step size.

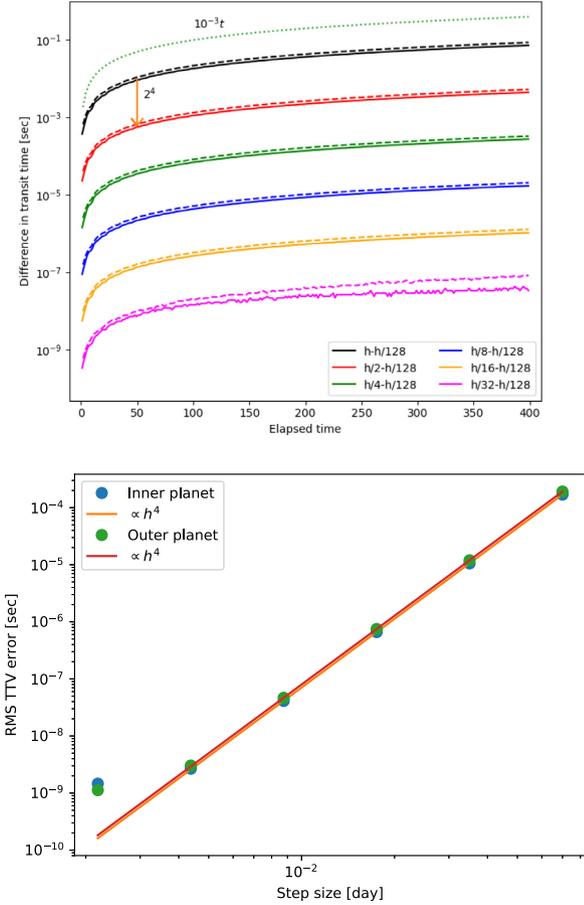
### 5.4 Transit-timing accuracy

Since the numerical accuracy of the integration depends upon the step size parameter,  $h$ , as the AHL21 integrator is fourth order in  $h$  (Fig. 3), we also expect that the accuracy of transit times should scale with  $h^4$ . Further precision could be obtained if we were to use a corrector at the start of the integration; however, such a corrector is left for future work.

Fig. 4 shows the change in the transit times with stepsize for a simulation of TRAPPIST-1 b and c over 400 d. Compared with the times computed with a very small step-size, the transit times drift with time. This behaviour is expected due to the difference between the symplectic Hamiltonian and the full Hamiltonian that contains high-frequency terms that cause the coordinates of the symplectic integrator to be offset from the real coordinates. This offset causes the longitudes to drift with time, and due to the slight difference in orbital frequency, the drift grows linearly as shown in Fig. 4. Since the AHL21 algorithm has order  $h^4$ , these offsets decrease with step-size as  $h^4$  (Fig. 4), and so at small stepsize the symplectic integrator better approximates the real system. In practice, these coordinate offsets are not expected to be important for transit-timing analyses as they will lead to very small differences between the inferred parameters and real parameters, even for large  $h$ . We recommend that the user determine which step size is appropriate by checking the difference in transit times as a function of step size.

For the purposes of TTVs, we are primarily interested in the precision of the non-linear portion of the transit times versus epoch. So, to assess the TTV precision, we subtract a linear fit from the difference between an integration with large  $h$  with an integration with small  $h$ , and then compute the RMS of the residuals. We expect the RMS to scale as  $h^4$ , and Fig. 4 indeed shows that this is the case for large step size for a system with two planets of periods 1.5 and 2.4 d, masses  $3 \times 10^{-5}$  of the star, low eccentricity, and integrated over 400 d (this approximates the inner two planets of the TRAPPIST-1 system). In both cases the TTV precision reaches a value that is  $\approx 10^{-14}$  of each planet's orbital period. We also find that this precision scales in proportion to the ratio of the masses of the planets to the star, as expected (see discussion under equation 2).

Having demonstrated that the accuracy of the transit times scales as expected, we next examine the numerical precision of the computed times, as well as their derivatives.

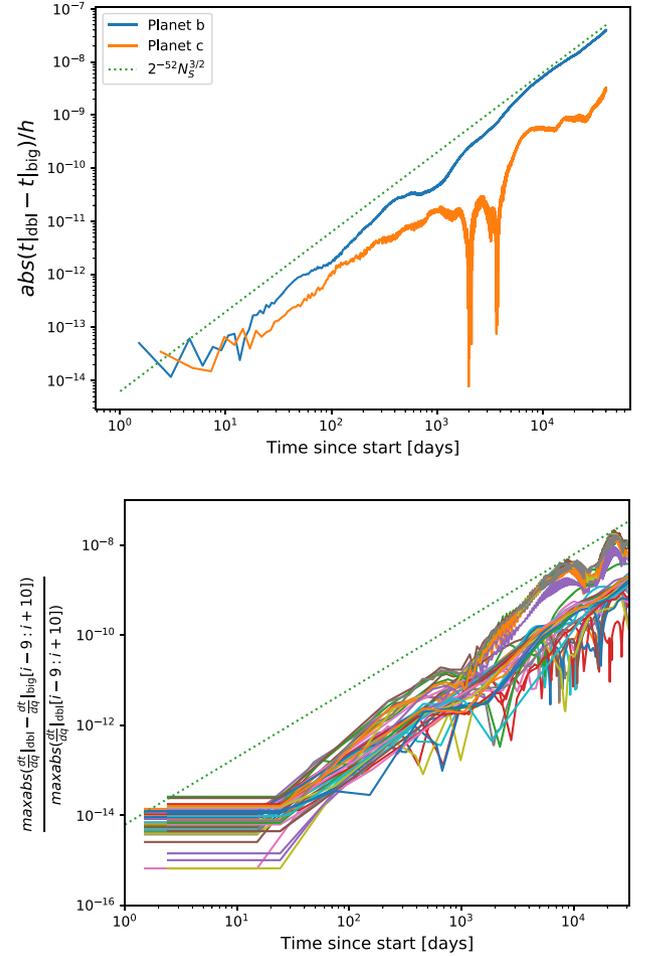


**Figure 4.** (Top) Variation in the transit times with step size. The absolute value of the difference in times between the indicated step size and a step size of  $h/128$  is plotted for the inner two planets of TRAPPIST-1 (solid and dashed, respectively) simulated over 400 d with a nominal step size of  $h = 0.06$  d. The dotted curve shows a linear scaling, while the arrow indicates a  $1/24$  decrease in timing difference when the step-size is halved. (Bottom) RMS precision of TTVs for the TRAPPIST-1/b/c two-planet system with six stepsizes, compared with an integration which is  $1/2$  of the shortest stepsize, after removing a linear fit from the difference to isolate the comparison to the transit-timing variations. For both planets the TTV errors scale as  $h^4$  until double-precision is reached at  $\approx 10^{-9}$  s.

### 5.5 Precision of transit times and their derivatives

Given a fixed step size for the algorithm, we next ask the question: how numerically precise are the transit times computed for that step size? And, how precise are the derivatives computed as a function of the initial conditions? These questions involve the control of truncation and round-off errors in the algorithm, which motivated the development of the AHL21 algorithm.

We check the numerical precision of the algorithm by comparing the transit times and their derivatives computed at both double precision and extended precision (using the double-precision `Float64` type with 64 bits, and the extended-precision `BigFloat` type with 256 bits in JULIA). Fig. 5 shows the difference in the times of transit in the TRAPPIST-1 b and c case computed in double-precision relative to `BigFloat` precision. We find that the computational errors grow at a rate that is bounded by  $\approx 2^{-52} h N_S^{3/2}$ , where  $N_S$  is the number of elapsed time-steps (Fig. 5), as expected for phase-errors based on



**Figure 5.** (Top) Fractional numerical error of transit times computed over 40000 d for TRAPPIST-1 b and c computed from double and `BigFloat` integrations. The error is plotted relative to the time step,  $h$ . (Bottom) Fractional numerical error on the transit time derivatives with respect to the initial Cartesian coordinates and masses. The green-dotted lines in both panels scale as  $2^{-52} N_S^{3/2}$ , where  $N_S$  is the number of time-steps. The maximum absolute derivative differences have been taken over 20 steps, and divided by the maximum absolute derivatives to give the fractional differences.

Brouwer’s Law (Brouwer 1937). The computation was carried out for 400,000 d for an inner orbital period of 1.5 d, for a total of  $\approx 10^7$  time-steps.

Next, we carry out tests of the numerical precision of the Jacobians at each substep in the calculation, as well as for the entire integration interval, and for the transit time derivatives. We do this in two ways: (1) by computing finite-difference derivatives in extended precision and (2) by comparing the derivatives in double precision with derivatives computed with extended precision arithmetic. The finite-difference test checks that the formulas derived in Section 4 are valid, while the extended precision test checks that the numerical implementation is precise.

To compute the finite-difference derivatives, we carry out integrations for each parameter using `BigFloat`, and compute a finite difference approximation of the partial derivative with parameters perturbed just above and below the nominal value:

$$\frac{\partial t}{\partial q_{i,j}} \approx \frac{\Delta t}{\Delta q_{i,j}} = \frac{t(q_{i,j}(1 + \varepsilon_{\text{diff}})) - t(q_{i,j}(1 - \varepsilon_{\text{diff}}))}{2\varepsilon_{\text{diff}} q_{i,j}}, \quad (63)$$

where  $t(q_{i,j})$  indicates the transit time evaluated at initial conditions  $\mathbf{q}_0$  with the  $i$ ,  $j$ th initial condition given by  $q_{i,j}$ . Typically we use  $\varepsilon_{\text{diff}} = 10^{-18}$ , and we find that the finite-difference derivative is insensitive to this value when rounded back to double-precision. We used these finite differences in writing and debugging each substep of the code, and we have created a suite of tests which can be used when further modifying or developing the code. We found that the finite difference derivatives agree with the derivatives computed from propagating the Jacobian, at a level close to the double-precision limit, which validates our implementation of the algorithm based on the formulae in Section 4.

Next, we estimate the fractional numerical errors on the derivatives of the transit times with respect to the initial conditions and masses propagated through the numerical integration (Fig. 5, bottom panel) by comparing the derivatives computed at double precision with those computed at extended precision. We find that these double-precision numerical errors also shows a growth that is bounded by  $2^{-52} N_S^{3/2}$ , according to Brouwer’s Law. In this case, we filtered the derivatives before computing the fractional error taking the maximum absolute value of each derivative over 20 transit times normalized by the maximum absolute derivative over the same 20 times to avoid the case in which the values of the derivatives approach zero. We did not find that the Brouwer’s law limit applied to be the DH17 algorithm – the errors significantly exceeded the Brouwer’s law limit for long integrations – that motivated the development of the AHL21 algorithm. We did not carry out longer integrations due to the high computational expense for `BigFloat` precision.

We conclude that based on these tests the code is performing as expected: the  $N$ -body code is fast and as accurate as the algorithm allows; the transit times are precise; the derivative formulae are correct; and the derivatives are precise. With these validations of the code completed, we now turn to compare our code with other publicly available  $N$ -body and transit-timing codes.

## 6 COMPARISON WITH OTHER CODES

In this section, we compare with two existing open-source  $N$ -body integrators that have been used for transit-timing and  $N$ -body integration: `TTVFast` and `REBOUND`. Although other codes are available, such as `SYSTEMIC` (Meschiari & Laughlin 2010) and `TRADES` (Borsato et al. 2014), as well as numerous proprietary codes for modelling transit timing, `TTVFast` and `REBOUND` are both widely used and open source. These comparisons provide further validation of the accuracy of our code, as well as timing benchmarks of the relative speeds.

### 6.1 Comparison of transit times with `TTVFast` and `REBOUND`

The `TTVFast` approach uses a WH integrator (Wisdom & Holman 1991) with a central dominant body, appropriate for planetary systems orbiting a single star (or planets orbiting a single star in a wide binary). A third-order corrector is used at the start of each simulation to transform from real coordinates to symplectic coordinates. Two versions of `TTVFast` have been developed in FORTRAN and C; here we describe comparisons with the latter.<sup>10</sup> The initial conditions may be specified in either Jacobi or heliocentric orbital elements, or heliocentric Cartesian coordinates. We use the initial Cartesian coordinates from `NbodyGradient` transformed to heliocentric coordinates, and then rotated by  $180^\circ$  about the  $y$ -axis

<sup>10</sup><https://github.com/kdeck/TTVFast>

so that the observer is located along the  $+z$ -axis, the convention adopted in `TTVFast`.

The `TTVFast` algorithm uses an approximate method to find times of transit. When a transit time is found to occur for one of the planets between two time-steps, then two Keplerian integrations between the planet and star are integrated forwards and backwards from the prior and subsequent time-steps, and weighted to approximate the position of the planet relative to the star. Newton’s method is then used to find the time of transit in the same manner described above (Section 4.11).

We have made a comparison of the transit times from `NbodyGradient` with `TTVFast` using the best-fitting initial conditions for the seven-planet TRAPPIST-1 system (Agol et al. 2021). For this comparison, we use a time-step for `TTVFast` which is 0.05 per cent<sup>11</sup> of the orbital period of planet b (Fig. 6) to reduce the difference between the symplectic and real coordinates (we use a larger step of 0.1 per cent for `NbodyGradient` as this integrator is higher precision). We find that over a time-scale of  $\sim 4532$  d (an estimate of the total time between initial and final TRAPPIST-1 observations over the lifetime of JWST), the difference between `TTVFast` and `NbodyGradient` is better than a few milliseconds for all seven planets, with better agreement for the inner planets than for the outer. This agreement is quite good, and we attribute the remaining differences, which grow with time, as being due to differences between the initial mapping and real coordinates which cause phase errors to grow with time.

Although `REBOUND` is not primarily designed for transit-timing, there is a PYTHON notebook in the `REBOUND` repository that gives an example of transit-time computation.<sup>12</sup> We used the same initial Cartesian coordinates as the `NbodyGradient` computation for TRAPPIST-1, and computed the transit times with a tolerance of  $10^{-12}$  d for `REBOUND`. We transform  $z \rightarrow -x$ ,  $x \rightarrow y$ , and  $y \rightarrow -z$  to allow for the fact that the `REBOUND` computation places the observer along the  $x$ -axis rather than along the  $-z$ -axis (as assumed in `NbodyGradient`, Fig. 2). Fig. 6 shows that over 4000 d for TRAPPIST-1, the times agree between `NbodyGradient` and `REBOUND` at the  $< 4 \mu\text{s}$  level. This was computed with a time-step of 0.0015 d for `NbodyGradient`, about 1/1000 of the orbital period of the inner planet, TRAPPIST-1b, to reduce the difference between the symplectic and real coordinates.

Unfortunately, `TTVFast` does not include derivatives, which was part of the motivation for developing the `NbodyGradient` code. However, given that the transit times compare well, and that we have compared the `NbodyGradient` derivatives with finite-differences computed at high precision (Section 5.5), this gives us confidence that the `NbodyGradient` derivatives are also being computed accurately. We have made scripts available for reproducing this comparison in the `NbodyGradient` repository.

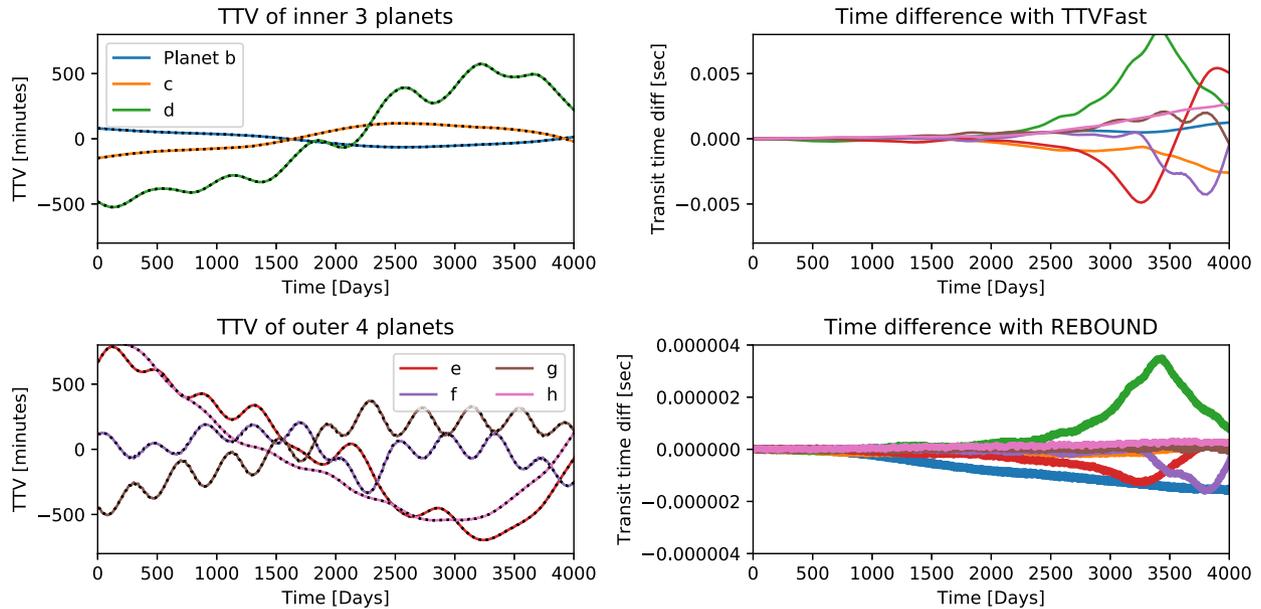
Next, we compare the run-time of `NbodyGradient` with `REBOUND`, with and without gradients.

### 6.2 Run-time comparison with `REBOUND`

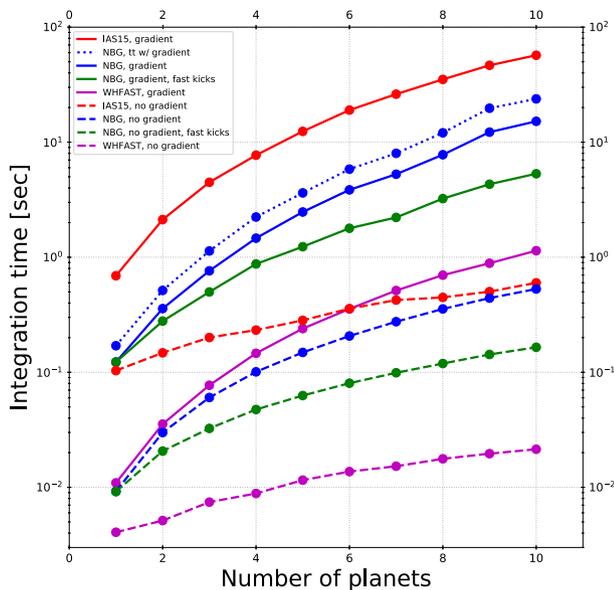
The `REBOUND` integrator `IAS15` (Rein & Tamayo 2015) allows for the computation of the variational equations, and may be used to

<sup>11</sup>We also had to modify `TTVFast` to avoid accumulation of numerical errors that occur for such a small time-step. Rather than adding the time-step to the elapsed time every time-step, we multiply the current number of steps by the time-step to obtain the elapsed time.

<sup>12</sup>See [https://rebound.readthedocs.io/en/latest/ipython\\_examples/TransitTimingVariations](https://rebound.readthedocs.io/en/latest/ipython_examples/TransitTimingVariations) for a description.



**Figure 6.** Transit-timing comparison for TRAPPIST-1. (Left) TTVs from *NbodyGradient* (in colour), and *TTVFast* and *REBOUND* (the black dots) for the seven TRAPPIST-1 planets over 4000 d. (Right) Timing differences in seconds between *NbodyGradient* and *TTVFast* (top) and *REBOUND* (bottom); the colours are the same as in left-hand panel. Note that the panel in the upper right has a range that has been expanded by  $6 \times 10^6$  relative to the left-hand panels, while the lower right-hand panel is expanded by a factor of  $1.2 \times 10^{10}$ .



**Figure 7.** Computational time of the *NbodyGradient* code without (blue), with fast kicks for planet pairs (green), and with transit-time computation (magenta) versus *REBOUND* IAS15 (red), with (solid) and without (dashed) computation of the gradient. We also compare with the *REBOUND* *WHFast* algorithm (magenta), both with gradients (solid) and without (dashed).

model systems with close encounters and an arbitrary architecture.<sup>13</sup> Fig. 7 compares the *REBOUND* IAS15 integrator computational speed with our code. We spaced planets by a ratio of semimajor axis of 1.8, and with initial orbital angles separated by 1.4 radians. For the AHL21 integration, we use a step size that is 1/20 of the orbital

period of the inner planet and we integrate for 800 orbits of the inner planet. We ran both codes with a range of planets from one to ten, and we turned off the transit finding to make a fair comparison. We tried two different versions of the AHL21 integrator: with fast kicks for pairs of planets, and with fast kicks turned off. When the fast kicks are used for planet pairs, we find that the AHL21 code compares well to *REBOUND* IAS15 algorithm when no gradients are computed, either slightly faster or comparable in wall clock time for 1–10 planets (red and blue dashed lines in Fig. 7). However, when the gradient is computed, our code takes a computational time that is  $\approx 4 - 5$  times faster than *REBOUND* IAS15 for a large number of planets when fast kicks are turned off. If the fast-kicks are used for planet pairs, then *NbodyGradient* is an order of magnitude faster than the IAS15 integrator in *REBOUND*. Note that both *REBOUND* gradients and *NbodyGradient* assume the Newtonian equations of motion when computing gradients.

We also compare with the *WHFast* algorithm in *REBOUND*. This algorithm is also symplectic, but requires a central dominant mass. The *WHFast* algorithm is by far the speediest of the three: both with and without gradients it is about an order of magnitude faster than either AHL21 or IAS15. However, as it is a second-order algorithm, it may require the use of a corrector, and/or shorter step-sizes, to obtain similar precision as the AHL21 algorithm which is fourth order; this may come with extra computational cost, depending on the particular application. From a theoretical standpoint, if the Kepler solver function calls dominate the compute time, *WHFast* should only be twice as fast as AHL21 with kicks between planets in *NbodyGradient* (green-dashed curve). Since we have not achieved this, it may indicate we have not yet properly optimized our code. *JULIA*, *NbodyGradient*'s language, is believed to be able to achieve speeds comparable to C++, *WHFast*'s language. We plan to continue to optimize *NbodyGradient*.

Our primary goal in developing this *N*-body code is for modelling observational data for which the uncertainties are typically dominated by measurement errors rather than model accuracy. Hence, we

<sup>13</sup><https://github.com/hannorein/rebound>

are willing to exchange some accuracy for computational speed by using a symplectic integrator with a large time-step. Thus, AHL21 may provide a useful compromise between IAS15 and WHFast. Note that the integrator is still precise, but the symplectic Hamiltonian only approximates the real Hamiltonian, and we have not been able to derive a corrector to transform between symplectic and real coordinates; see the discussion in Deck et al. (2014) and references therein.

In sum, the AHL21 algorithm in `NbodyGradient` may gain some computation time for general  $N$ -body problems with the trade-off of interpreting the initial conditions as symplectic coordinates, not real coordinates. In addition, currently REBOUND (either IAS15 or WHFast) does not yet implement the gradients of the transit times with respect to the initial conditions, which `NbodyGradient` was designed to compute from the start.

## 7 SUMMARY AND CONCLUSIONS

The original goal of our development of this paper and code was to make possible the analytic computation of the derivatives of the times of transit with respect to the initial conditions; this is the first time this has been done in an  $N$ -body code to our knowledge. We have accomplished this with a fast and robust code written in the JULIA language. JULIA has the advantage of matching compiled C speeds if it is written in an optimum manner, which our benchmarks indicate we have achieved. Yet it allows for interactive usage and high-level coding that makes building and debugging the code more straightforward. In addition, JULIA easily allows changing of the variable types, so that higher precision computations are available simply by calling the  $N$ -body functions with coordinates and masses initialized with high-precision variables; the functions will be automatically recompiled at the first time being called with different numeric types.

We have also developed this code with generality in mind; in particular, we would like to eventually apply it to hierarchical systems such as circumbinary planets or planets hosting moons, which is why we have based the symplectic splitting on the DH17 algorithm that does not assume a dominant body (or bodies). In addition, the popular WH method assumes the perturbed approximation holds, which implies the method breaks down during close encounters. The explanation, using error analysis, is that WH has two-body error terms. However, when DH17 is used without kicks, it only has three-body terms which blow up during strong three-body encounters (Dehnen & Hernandez 2017). So our code can better handle close two-body encounters, as was shown by Dehnen & Hernandez (2017), who used it to simulate a stellar cluster.

However, a drawback of the DH17 algorithm we found was the lack of precision caused by cancellations between the backward drifts and forward Kepler steps. We have fixed this problem by carrying out analytic cancellation of these expressions with modified versions of Gauss's  $f$  and  $g$  functions. This fix creates an algorithm that is both numerically stable and precise: energy and angular momentum are conserved well for long integrations. The algorithm is accurate to fourth order in time-step, but even for large time-steps it will integrate the non-Keplerian perturbations between the bodies with sufficient accuracy for observational data (and the time-step can be decreased until the desired precision is reached; this happens rapidly thanks to the fourth-order scaling of the algorithm with time-step). This accuracy is higher order than the WH symplectic integrator or its versions with symplectic correctors, which have error terms scaling as  $h^2$ . As with any symplectic integrator, the integration coordinates are offset slightly from the real coordinates (Wisdom, Holman & Touma 1996), which causes a long-term phase shift. However, this shift is small, even for large time-steps, and should not affect the interpretation of

the state of multibody dynamical systems. In practice it results in a slight offset of the initial coordinates which is caused by introducing high-frequency terms in the physical Hamiltonian.

In order to compute the derivatives of the transit times with respect to the initial conditions, we have propagated the Jacobian of the  $N$ -body positions and velocities with respect to the initial conditions and masses throughout each time-step of the  $N$ -body integration. We have alleviated numerical cancellations in this expansion, and used series expansions for special functions when cancellation of the leading orders occurs. This has given an algorithm that yields precise derivatives and that appears to adhere to Brouwer's law for up to  $10^7$  time-steps for the problem we tested. We also find that it compares favourably in run time to the variational equations integrated by IAS15 (Rein & Tamayo 2016), with a factor of 4-10 speed-up for long time-steps in the comparison we tested.

We have found that the derivatives make possible the optimization of the masses of planets in the TRAPPIST-1 planetary system, and the results compare well with an analysis with the GENGA code (Grimm & Stadel 2014), as reported in Agol et al. (2021). In particular, we were able to use the derivatives to efficiently find the maximum likelihood and to compute the Hessian at the maximum likelihood. We have also used it to find the likelihood profile as a function of the masses of the planets and orbital parameters, as well as to run a Hamiltonian Markov Chain Monte Carlo computation in 35 dimensions to derive the posterior distribution for the system parameters, which agree well with a complementary analysis based on the code described in Grimm et al. (2018). Finally, the derivatives enabled an efficient search for an eighth planet, which required optimization over 40 free parameters; no strong evidence for an eighth planet turned up in this search (Agol et al. 2021).

These analyses depend on initial conditions which were specified in terms of orbital elements, which we plan to describe in subsequent work. We are also continuing to develop the output options, API, and documentation, and we welcome community contributions to the code repository.<sup>14</sup>

There are some limitations to our work. Due to its symplectic nature, our code does not allow for non-conservative effects to be included, such as tidal forces, drag, or general relativity. However, these could be added in future work through the machinery described by Tamayo et al. (2020). We assume that the masses of the bodies are constant. We do not compute second- or third-order derivatives, as has been implemented by Rein & Tamayo (2016). Even so, we expect this code to find application in a wide range of dynamical problems related to observation of exoplanetary systems and beyond.

## ACKNOWLEDGEMENTS

EA acknowledges support from the Guggenheim Foundation, from NSF grant AST-1615315, and from the NASA Astrobiology Institute's Virtual Planetary Laboratory Lead Team, funded through the NASA Astrobiology Institute under solicitation NNA12ZDA002C and Cooperative Agreement Number NNA13AA93A. ZL acknowledges support from the Washington NASA Space Grant Consortium Summer Undergraduate Research Program. We thank Mosé Giordano for advice on optimizing JULIA code, and we thank Hanno Rein, Dan Tamayo, and the anonymous referee for comments on the submitted version that greatly improved the paper.

<sup>14</sup><http://github.com/ericagol/NbodyGradient.jl>

## DATA AVAILABILITY

No new data were generated or analysed in support of this research. Simulated data that were used for making the figures will be made available in a github repository.

## REFERENCES

- Agol E., Deck K., 2016, *ApJ*, 818, 177  
 Agol E., Fabrycky D. C., 2017, in Deeg H. J., Belmonte J. A., eds, Handbook of Exoplanets. Springer International Publishing, Cham, p. 1  
 Agol E., Steffen J., Sari R., Clarkson W., 2005, *MNRAS*, 359, 567  
 Agol E. et al., 2021, *Planet. Sci. J.*, 2, 1  
 Bezanon J., Edelman A., Karpinski S., Shah V. B., 2017, *SIAM Rev.*, 59, 65  
 Blackford L. S. et al., 2002, *ACM Trans. Math. Softw.*, 28, 135  
 Borsato L., Marzari F., Nascimbeni V., Piotto G., Granata V., Bedin L. R., Malavolta L., 2014, *A&A*, 571, A38  
 Borsato L. et al., 2019, *MNRAS*, 484, 3233  
 Brouwer D., 1937, *AJ*, 46, 149  
 Carter J. A. et al., 2011, *Science*, 331, 562  
 Channell P., Scovel J., 1991, *Phys. D*, 50, 80  
 Deck K. M., Agol E., 2015, *ApJ*, 802, 116  
 Deck K. M., Agol E., 2016, *ApJ*, 821, 96  
 Deck K. M., Agol E., Holman M. J., Nesvorný D., 2014, *ApJ*, 787, 132  
 Dehnen W., Hernandez D. M., 2017, *MNRAS*, 465, 1201  
 Doyle L. R. et al., 2011, *Science*, 333, 1602  
 Fabrycky D. C., 2010, in Seager S., ed., Exoplanets. University of Arizona Press, Tucson, AZ, p. 217  
 Freudenthal J. et al., 2018, *A&A*, 618, A41  
 Girolami M., Calderhead B., 2011, *J. R. Stat. Soc. B*, 73, 123  
 Gonçalves Ferrari G., Boekholt T., Portegies Zwart S. F., 2014, *MNRAS*, 440, 719  
 Grimm S. L., Stadel J. G., 2014, *ApJ*, 796, 23  
 Grimm S. L. et al., 2018, *A&A*, 613, A68  
 Hadden S., Lithwick Y., 2016, *ApJ*, 828, 44  
 Hadden S., Lithwick Y., 2017, *AJ*, 154, 5  
 Hairer E., Lubich C., Wanner G., 2006, Geometric numerical integration: structure-preserving algorithms for ordinary differential equations in Springer Series in Computational Mathematics, Vol. 31. Springer Science & Business Media, Berlin  
 Hairer E., McLachlan R. I., Razakarivony A., 2008, *BIT Numer. Math.*, 48, 231  
 Hamers A. S., Portegies Zwart S. F., 2016, *MNRAS*, 459, 2827  
 Hernandez D. M., 2016, *MNRAS*, 458, 4285  
 Hernandez D. M., Bertschinger E., 2015, *MNRAS*, 452, 1934  
 Hernandez D. M., Bertschinger E., 2018, *MNRAS*, 475, 5570  
 Hernandez D. M., Dehnen W., 2017, *MNRAS*, 468, 2614  
 Holman M. J., Murray N. W., 2005, *Science*, 307, 1288  
 Holman M. J. et al., 2010, *Science*, 330, 51  
 Jontof-Hutter D., 2019, *Annu. Rev. Earth Planet. Sci.*, 47, 141  
 Jontof-Hutter D. et al., 2016, *ApJ*, 820, 39  
 Kahan W., 1965, *Commun. ACM*, 8, 40  
 Laughlin G., Chambers J. E., 2001, *ApJ*, 551, L109  
 Linial I., Gilbaum S., Sari R., 2018, *ApJ*, 860, 16  
 Lithwick Y., Xie J., Wu Y., 2012, *ApJ*, 761, 122  
 Malhotra R., Black D., Eck A., Jackson A., 1992, *Nature*, 356, 583  
 Meschiarri S., Laughlin G. P., 2010, *ApJ*, 718, 543  
 Mikkola S., Innanen K., 1999, *Celest. Mech. Dyn. Astron.*, 74, 59  
 Nesvorný D., Beaugé C., 2010, *ApJ*, 709, L44  
 Nesvorný D., Vokrouhlický D., 2014, *ApJ*, 790, 58  
 Nesvorný D., Vokrouhlický D., 2016, *ApJ*, 823, 72  
 Pál A., 2010, *MNRAS*, 409, 975  
 Peale S. J., 1993, *AJ*, 105, 1562  
 Rasio F. A., Nicholson P. D., Shapiro S. L., Teukolsky S. A., 1992, *Nature*, 355, 325  
 Rein H., Tamayo D., 2015, *MNRAS*, 452, 376  
 Rein H., Tamayo D., 2016, *MNRAS*, 459, 2275

- Tamayo D., Rein H., Shi P., Hernandez D. M., 2020, *MNRAS*, 491, 2885  
 Wisdom J., Hernandez D. M., 2015, *MNRAS*, 453, 3015  
 Wisdom J., Holman M., 1991, *AJ*, 102, 1528  
 Wisdom J., Holman M., Touma J., 1996, *Fields Inst. Commun.*, 10, 217  
 Wolfram Research Inc., Champaign, IL, 2019, *Mathematica*, Version 12.0  
 Wolszczan A., 1994, *Science*, 264, 538  
 Wolszczan A., Frail D. A., 1992, *Nature*, 355, 145  
 Yoffe G., Ofir A., Aharonson O., 2021, *ApJ*, 908, 114

## APPENDIX A: DERIVATION OF AHL21 KEPLER + DRIFT STEP

Here, we give more detail for the derivation of the combined Kepler and Drift steps that were described in Section 3.3. As a reminder, we start with two bodies  $i$  and  $j$  with relative coordinates  $\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$  and  $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ . In this section we drop the subscript  $ij$ . At the start of the time-step, the coordinates are  $\mathbf{x}_0$  and  $\mathbf{v}_0$ , while at the end of the time-step the coordinates are  $\mathbf{x}$  and  $\mathbf{v}$ . We propagate these forward with either a negative drift followed by a Kepler step, yielding a change in position and velocity of  $\Delta\mathbf{x}_{DK}$  and  $\Delta\mathbf{v}_{DK}$ , or a Kepler step followed by a negative drift, yielding  $\Delta\mathbf{x}_{KD}$  and  $\Delta\mathbf{v}_{KD}$ . In the process we need to solve Kepler's equation (13). The two pathways are shown in Fig. 1.

### A1 Drift then Kepler

As discussed in Section 3, the negative drift is taken first, yielding an intermediate position  $\hat{\mathbf{x}}_0 = \mathbf{x}_0 - h\mathbf{v}_0$ . The resulting change in position and velocity over the time-step is given by equation (15),

$$\begin{aligned}\Delta\mathbf{x}_{DK} &= \hat{\mathbf{x}} - \mathbf{x}_0 = (\hat{f} - 1)\mathbf{x}_0 + (\hat{g} - h\hat{f})\mathbf{v}_0, \checkmark \\ \Delta\mathbf{v}_{DK} &= \hat{\mathbf{v}} - \mathbf{v}_0 = \hat{f}\mathbf{x}_0 + (\hat{g} - h\hat{f} - 1)\mathbf{v}_0, \checkmark\end{aligned}\quad (\text{A1})$$

where, again,  $\hat{f}$ ,  $\hat{g}$ ,  $\hat{f}$ , and  $\hat{g}$  are all computed in terms of  $(\hat{\mathbf{x}}_0, \mathbf{v}_0, k, h)$ . This means that the scalar functions these depend on also need to be computed in terms of  $\hat{\mathbf{x}}_0$ , i.e.  $\hat{r}_0 = |\hat{\mathbf{x}}_0|$ ,  $\hat{\beta} = 2k/\hat{r}_0 - v_0^2$ ,  $\hat{\eta}_0 = \hat{\mathbf{x}}_0 \cdot \mathbf{v}_0$ , and  $\hat{G}_i = G_i(\hat{\beta}, \hat{\gamma})$ , where  $\hat{\gamma}$  can be computed with Newton's method from equation (13) evaluated using  $\hat{r}_0$ ,  $\hat{\eta}_0$ , and  $\hat{G}_i$ , and  $\hat{r}$  can be computed from (12) in the same manner.

The Gauss function terms in equation (15) are given as

$$\begin{aligned}\hat{f} - 1 &= -\frac{k}{\hat{r}_0}\hat{G}_2, \checkmark \\ \hat{g} - h\hat{f} &= k\left(\frac{h}{\hat{r}_0}\hat{G}_2 - \hat{G}_3\right), \checkmark \\ \hat{f} &= -\frac{k}{\hat{r}\hat{r}_0}\hat{G}_1, \checkmark \\ \hat{g} - h\hat{f} - 1 &= \frac{k}{\hat{r}}\left(\frac{h}{\hat{r}_0}\hat{G}_1 - \hat{G}_2\right), \checkmark\end{aligned}\quad (\text{A2})$$

where the 1's have been cancelled analytically for more accurate expressions at small time-steps.

### A2 Kepler then drift

In the case of a Kepler step followed by a negative drift, the Kepler step is computed in terms of the initial coordinates,  $\mathbf{x}_0$  and  $\mathbf{v}_0$ , yielding intermediate coordinates  $(\check{\mathbf{x}}, \check{\mathbf{v}})$ , and then the negative drift is applied resulting in  $\mathbf{x} = \check{\mathbf{x}} - h\check{\mathbf{v}}$  (Fig. 1). This yields the change in position and velocity of

$$\begin{aligned}\Delta\mathbf{x}_{KD} &= \check{\mathbf{x}} - h\check{\mathbf{v}} - \mathbf{x}_0 = (f - h\check{f} - 1)\mathbf{x}_0 + (g - h\check{g})\mathbf{v}_0, \checkmark \\ \Delta\mathbf{v}_{KD} &= \check{\mathbf{v}} - \mathbf{v}_0 = \check{f}\mathbf{x}_0 + (\check{g} - 1)\mathbf{v}_0, \checkmark\end{aligned}\quad (\text{A3})$$

where  $f, g, \dot{f}$ , and  $\dot{g}$  are all computed in terms of  $(\mathbf{x}_0, \mathbf{v}_0, k, h)$ , and the ‘KD’ indicates that the Kepler step precedes the negative drift, Kepler - drift.

These functions can also be expressed in terms of the Gauss functions as

$$\begin{aligned} f - hf - 1 &= \frac{k}{r} \left( G_2 - \frac{k}{r_0} (G_2^2 - G_1 G_3) \right), \\ g - hg &= \frac{k}{r} (r_0 (G_1 G_2 - G_0 G_3) + \eta_0 (G_2^2 - G_1 G_3)), \\ \dot{f} &= -\frac{k}{rr_0} G_1, \checkmark \\ \dot{g} - 1 &= -\frac{k}{r} G_2, \checkmark \end{aligned} \quad (\text{A4})$$

where we have used equations (13) and (12) to transform these equations. Note that in this case each of these functions depends on  $(\mathbf{x}_0, \mathbf{v}_0, k, h)$  as the Kepler step is applied *before* the drift.

Unfortunately, these equations can lead to numerical instability for small values of  $\gamma = \sqrt{|\beta|}s$ .<sup>15</sup> The offending terms involve difference of products of  $G_i$  functions:  $G_1 G_2 - G_0 G_3$  and  $G_2^2 - G_1 G_3$ . These terms have a Taylor series expansion in which the leading order terms in  $s$  cancel; this is also true for the function  $G_3$ .

So, we define two new functions,

$$\begin{aligned} H_1 &= G_2^2 - G_1 G_3, \\ H_2 &= G_1 G_2 - G_0 G_3, \end{aligned} \quad (\text{A5})$$

given in Table 1. In terms of these functions we have

$$\begin{aligned} f - hf - 1 &= \frac{k}{r} \left( G_2 - \frac{k}{r_0} H_1 \right), \checkmark \\ g - hg &= \frac{k}{r} (r_0 H_2 + \eta_0 H_1), \checkmark \end{aligned} \quad (\text{A6})$$

For large values of  $\gamma$  we evaluate these with the special function definitions, summarized in Table 1, while for small values of  $\gamma$ , we evaluate  $G_3, H_1$  and  $H_2$  in terms of the following Taylor series:

$$H_1(\gamma, \beta) = \frac{2\gamma^4}{\beta^2} \sum_{n=0}^{\infty} \frac{(\varepsilon\gamma^2)^n (n+1)}{(2n+4)!}, \checkmark \quad (\text{A7})$$

$$H_2(\gamma, \beta) = \frac{2\gamma^3}{|\beta|^{3/2}} \sum_{n=0}^{\infty} \frac{(\varepsilon\gamma^2)^n (n+1)}{(2n+3)!}, \checkmark \quad (\text{A8})$$

$$G_3(\gamma, \beta) = \frac{\gamma^3}{|\beta|^{3/2}} \sum_{n=0}^{\infty} \frac{(\varepsilon\gamma^2)^n}{(2n+3)!}, \checkmark \quad (\text{A9})$$

where  $\varepsilon = -1$  for  $\beta > 0$  (elliptic) and  $\varepsilon = 1$  for  $\beta < 0$  (hyperbolic) cases. Note that in evaluating these series expansions we compute each term recursively, and terminate the series expansion when the function matches one of the two prior partial sums (indicating that the series is converged to machine precision).

The fact that these functions have leading terms  $\propto \gamma^3$  and  $\propto \gamma^4$  is due to cancellation of lower order terms in the trigonometric representation. This cancellation can lead to round-off errors for small values of  $\gamma$ , which are commonly encountered when there are a wide range of orbital time-scales in a system. We find that higher precision is obtained by evaluating the series expressions for  $\gamma < 1/2$  out to  $\approx 6$  terms in double precision, while for  $\gamma > 1/2$ , higher precision is obtained from the full trigonometric expression.

<sup>15</sup>Note that in the elliptic Kepler’s equation,  $\gamma$  is equal to the change in eccentric anomaly over the time-step.

The other  $G_i$  functions ( $G_0, G_1$  and  $G_2$ ) we evaluate with the stable trigonometric and hyperbolic function transforms discussed in Wisdom & Hernandez (2015), also given in Table 1. With the combination of the drift and Kepler steps, it turns out that we no longer need the drift of the centre-of-mass coordinates in each Kepler step as these cancel exactly.

## APPENDIX B: DERIVATION OF DERIVATIVES OF KEPLER + DRIFT STEPS

In this appendix, we give more detail on the derivation of the derivatives of the combined Kepler and drift steps, as well as formulae for the derivatives of the scalar quantities in equations (18), (19), (20), and (21).

### B1 Differential of intermediate quantities

The differentials of  $r_0, \beta, \eta_0$  are given by

$$\delta r_0 = \frac{\mathbf{x}_0 \cdot \delta \mathbf{x}_0}{r_0}, \quad (\text{B1})$$

$$\delta \beta = \frac{\delta k}{r_0} - \frac{k \delta \mathbf{x}_0 \cdot \mathbf{x}_0}{r_0^3} - 2\mathbf{v}_0 \cdot \delta \mathbf{v}_0, \quad (\text{B2})$$

$$\delta \eta_0 = \mathbf{v}_0 \cdot \delta \mathbf{x}_0 + \mathbf{x}_0 \cdot \delta \mathbf{v}_0. \quad (\text{B3})$$

Note that these quantities are constant over a time-step, and so there is no dependence upon  $\delta h$ .

Taking the differential of the Universal Kepler equation 13, we find

$$\begin{aligned} |\beta|^{-1/2} r \delta \gamma &= \delta h + \delta k \left[ \frac{D}{r_0} - G_3 \right] \\ &\quad - \delta \mathbf{x}_0 \cdot \left[ (kD + G_1 r_0^2) \frac{\mathbf{x}_0}{r_0^3} + G_2 \mathbf{v}_0 \right] \\ &\quad - \delta \mathbf{v}_0 \cdot [D \mathbf{v}_0 + G_2 \mathbf{x}_0], \end{aligned} \quad (\text{B4})$$

$$D = \beta^{-1} [h + \eta_0 G_2 + 2k G_3]. \checkmark \quad (\text{B5})$$

Note that the  $\delta \mathbf{v}_0 \cdot \mathbf{x}_0$  and  $\delta \mathbf{x}_0 \cdot \mathbf{v}_0$  have the same derivative terms; this is due to both of these terms deriving from  $\delta \eta_0$ .

Taking the differential of the radial equation, (12), we find

$$\begin{aligned} \delta r &= \frac{\delta \gamma}{|\beta|^{1/2}} (\eta_0 G_0 + \zeta G_1) + \frac{\delta k}{\beta r_0} (r_0 - r - k G_2) \\ &\quad + \delta \mathbf{x}_0 \cdot \left[ (kr + k^2 G_2 - \zeta r_0 G_0) \frac{\mathbf{x}_0}{\beta r_0^3} + G_1 \mathbf{v}_0 \right] \\ &\quad + \delta \mathbf{v}_0 \cdot \left[ (\eta_0 G_1 + 2k G_2) \frac{\mathbf{v}_0}{\beta} + G_1 \mathbf{x}_0 \right], \end{aligned} \quad (\text{B6})$$

where we define

$$\zeta = k - r_0 \beta. \checkmark \quad (\text{B7})$$

The functions  $G_i(\beta, \gamma)$  have derivatives in terms of  $\beta$  and  $\gamma$ , which we can combine with the foregoing differentials for these quantities to obtain the derivatives with respect to the basis  $(\mathbf{x}_0, \mathbf{v}_0, k, h)$ . These intermediate derivatives of  $G_i$  (in both the elliptic and hyperbolic cases) are given by

$$\frac{\partial G_{i+1}}{\partial \gamma} = \frac{G_i}{|\beta|^{1/2}} \quad (i \geq 0), \quad (\text{B8})$$

$$\frac{\partial G_0}{\partial \gamma} = -\frac{\beta G_1}{|\beta|^{1/2}}, \quad (\text{B9})$$

and

$$\frac{\partial G_n}{\partial \beta} = -\frac{n}{2\beta} G_n. \quad (\text{B10})$$

With these intermediate derivatives in hand, we can compute the full differentials of the Gauss functions in the combined drift and Kepler terms. Since these formulae differ in the two cases, we consider each separately in turn in the following two subsections.

## B2 Differential of drift-then-Kepler step

The differential of the scalar quantities  $\delta(\hat{f} - 1)$ ,  $\delta(\hat{g} - h\hat{f})$ ,  $\delta(\hat{f})$ , and  $\delta(\hat{g} - h\hat{f} - 1)$  should also be scalars, and can be expressed in terms similar to the  $\delta\gamma$  and  $\delta r$  terms given above. Note, however, that as the Kepler step takes place *after* the negative drift, all of these functions are to be computed in terms of  $\hat{\mathbf{x}}_0$  substituted for  $\mathbf{x}_0$ , and so we need to add an extra step in the derivation to find the differentials in terms of  $\mathbf{x}_0$  in lieu of  $\hat{\mathbf{x}}_0$ .

The differential of these functions in terms of intermediate scalar quantities is given by

$$\frac{\delta(\hat{f} - 1)}{\hat{f} - 1} = \frac{\delta k}{k} - \frac{\delta \hat{r}_0}{\hat{r}_0} + \frac{\hat{G}_1}{\hat{G}_2} \frac{\delta \hat{\gamma}}{|\hat{\beta}|^{1/2}} - \frac{\delta \hat{\beta}}{\hat{\beta}}, \quad (\text{B11})$$

$$\begin{aligned} \delta(\hat{g} - h\hat{f}) &= \frac{\delta k}{k} (\hat{g} - h\hat{f}) + k \left[ \left( \frac{\delta h}{\hat{r}_0} - \frac{h\delta \hat{r}_0}{\hat{r}_0^2} \right) \hat{G}_2 \right. \\ &\quad \left. + \frac{\delta \hat{\gamma}}{|\hat{\beta}|^{1/2}} \left( \frac{h}{\hat{r}_0} \hat{G}_1 - \hat{G}_2 \right) - \frac{\delta \hat{\beta}}{\hat{\beta}} \left( \frac{h}{\hat{r}_0} - \frac{3}{2} \hat{G}_3 \right) \right], \quad (\text{B12}) \end{aligned}$$

$$\frac{\delta \hat{f}}{\hat{f}} = \frac{\delta k}{k} - \frac{\delta \hat{r}_0}{\hat{r}_0} - \frac{\delta \hat{r}}{\hat{r}} + \frac{\hat{G}_0}{\hat{G}_1} \frac{\delta \hat{\gamma}}{|\hat{\beta}|^{1/2}} - \frac{1}{2} \frac{\delta \hat{\beta}}{\hat{\beta}}, \quad (\text{B13})$$

$$\begin{aligned} \delta(\hat{g} - h\hat{f} - 1) &= \left[ \frac{\delta k}{\hat{r}_0 \hat{r}} - \frac{k\delta \hat{r}_0}{\hat{r}_0^2 \hat{r}} - \frac{k\delta \hat{r}}{\hat{r}_0 \hat{r}^2} \right] (h\hat{G}_1 - \hat{r}_0 \hat{G}_2) \\ &\quad + \frac{k}{\hat{r}_0 \hat{r}} \left[ \delta h \hat{G}_1 - \delta \hat{r}_0 \hat{G}_2 - \frac{1}{2} \frac{\delta \hat{\beta}}{\hat{\beta}} (h\hat{G}_1 - 2\hat{r}_0 \hat{G}_2) \right. \\ &\quad \left. + (h\hat{G}_0 - r_0 \hat{G}_1) \frac{\delta \hat{\gamma}}{|\hat{\beta}|^{1/2}} \right]. \quad (\text{B14}) \end{aligned}$$

In this equation we have used the ‘‘ symbol to indicate that each of these quantities is a function of  $\hat{\mathbf{x}}_0$  rather than  $\mathbf{x}_0$  (with the exception of  $h$  and  $k$ ). Into these differentials we can substitute our expressions for  $\delta \hat{\beta}$ ,  $\delta \hat{r}$ ,  $\delta \hat{r}_0$ ,  $\delta \hat{\eta}_0$ , and  $\delta \hat{\gamma}$  given above, keeping in mind that these need to be computed in terms of  $\delta \hat{\mathbf{x}}_0$  and  $\hat{\mathbf{x}}_0$  substituted for  $\delta \mathbf{x}_0$  and  $\mathbf{x}_0$ . Then, we need to replace the differential  $\delta \hat{\mathbf{x}}_0$  by  $\delta \mathbf{x}_0 - h\delta \mathbf{v}_0 - \mathbf{v}_0 \delta h$ . The  $\mathbf{v}_0 \delta h$  term leads to dot products of  $\mathbf{v}_0 \cdot \hat{\mathbf{x}}_0 = \hat{\eta}_0$  and  $\mathbf{v}_0 \cdot \mathbf{v}_0 = \frac{2k}{\hat{r}_0} - \hat{\beta}$ . The algebraic computation of these operations was aided by Mathematica (Wolfram Research 2019), yielding the following results:

$$\begin{aligned} \delta \ln(\hat{f} - 1) &= \delta k \mathcal{A}_k + \delta h \mathcal{A}_h + \delta \mathbf{x}_0 \cdot \mathbf{x}_0 \mathcal{A}_{xx} \\ &\quad + \delta(\mathbf{x}_0 \cdot \mathbf{v}_0) \mathcal{A}_{xv} + \delta \mathbf{v}_0 \cdot \mathbf{v}_0 \mathcal{A}_{vv}, \checkmark \\ \mathcal{A}_k &= \frac{1}{k} - \frac{2}{\beta r_0} + \frac{c_1 G_1}{r r_0 G_2}, \checkmark \\ \mathcal{A}_h &= \frac{G_1}{r} \left( \frac{2k}{r_0} - \beta + \frac{1}{G_2} \right) - c_{24} \eta_0, \checkmark \\ \mathcal{A}_{xx} &= c_{24}, \checkmark \\ \mathcal{A}_{xv} &= - \left[ c_{24} h + \frac{G_1}{r} \right], \checkmark \\ \mathcal{A}_{vv} &= \frac{1}{r} \left[ \frac{k H_6}{\beta G_2} - r_0 G_2 + h(2G_1 + h r c_{24}) \right], \checkmark \quad (\text{B15}) \end{aligned}$$

$$\begin{aligned} \delta(\hat{g} - h\hat{f}) &= \delta k \mathcal{B}_k + \delta h \mathcal{B}_h + \delta \mathbf{x}_0 \cdot \mathbf{x}_0 \mathcal{B}_{xx} \\ &\quad + \delta(\mathbf{x}_0 \cdot \mathbf{v}_0) \mathcal{B}_{xv} + \delta \mathbf{v}_0 \cdot \mathbf{v}_0 \mathcal{B}_{vv}, \checkmark \\ \mathcal{B}_k &= -\frac{c_9 k}{\beta r_0^2} + \frac{c_1 c_{13} k}{r r_0^2} + \frac{G_2 h - G_3 r_0}{r_0}, \checkmark \\ \mathcal{B}_h &= \frac{k G_2}{r_0} + \frac{k c_{13}}{r r_0} + \frac{k G_2 c_{13}}{r r_0} \left( \frac{2k}{r_0} - \beta \right) - \eta_0 c_{10}, \checkmark \\ \mathcal{B}_{xx} &= c_{10}, \checkmark \\ \mathcal{B}_{xv} &= - \left[ c_{10} h + \frac{c_{13} G_2 k}{r r_0} \right], \checkmark \\ \mathcal{B}_{vv} &= \frac{2hk G_2 c_{13}}{r r_0} + h^2 c_{10} \\ &\quad + \frac{k}{\beta r r_0} [r_0^2 H_8 - \beta h r_0 G_2^2 + (hk + \eta_0 r_0) H_6], \checkmark \quad (\text{B16}) \end{aligned}$$

$$\begin{aligned} \delta \ln \hat{f} &= \delta k \mathcal{C}_k + \delta h \mathcal{C}_h + \delta \mathbf{x}_0 \cdot \mathbf{x}_0 \mathcal{C}_{xx} \\ &\quad + \delta(\mathbf{x}_0 \cdot \mathbf{v}_0) \mathcal{C}_{xv} + \delta \mathbf{v}_0 \cdot \mathbf{v}_0 \mathcal{C}_{vv}, \checkmark \\ \mathcal{C}_k &= \frac{1}{k} - \frac{1}{\beta r_0} - \frac{c_{17}}{\beta r r_0} - \frac{c_1 (G_1 c_2 - G_0 r)}{r^2 r_0 G_1}, \checkmark \\ \mathcal{C}_h &= c_{22} \left( \frac{2k}{r_0} - \beta \right) - \frac{c_2}{r^2} + c_{21} \eta_0 + \frac{G_0}{G_{1r}}, \checkmark \\ \mathcal{C}_{xx} &= c_{21}, \checkmark \\ \mathcal{C}_{xv} &= c_{22} - c_{21} h, \checkmark \\ \mathcal{C}_{vv} &= c_{34} - 2h c_{22} + h^2 c_{21}, \checkmark \quad (\text{B17}) \end{aligned}$$

$$\begin{aligned} \delta(\hat{g} - h\hat{f} - 1) &= \delta k \mathcal{D}_k + \delta h \mathcal{D}_h + \delta \mathbf{x}_0 \cdot \mathbf{x}_0 \mathcal{D}_{xx} \\ &\quad + \delta(\mathbf{x}_0 \cdot \mathbf{v}_0) \mathcal{D}_{xv} + \delta \mathbf{v}_0 \cdot \mathbf{v}_0 \mathcal{D}_{vv}, \checkmark \\ \mathcal{D}_k &= \frac{1}{r r_0} \left[ -\frac{k(c_{13} - G_2 r_0)}{\beta r_0} + c_{13} \right. \\ &\quad \left. - \frac{k c_{13} c_{17}}{\beta r r_0} + \frac{k c_1 c_{12}}{r r_0} - \frac{k c_1 c_2 c_{13}}{r^2 r_0} \right], \checkmark \\ \mathcal{D}_h &= \frac{k G_1}{r r_0} + \frac{k c_{12}}{r^2 r_0} - \frac{k c_{13} c_2}{r^3 r_0} \\ &\quad - c_{26} \left( \frac{2k}{r_0} - \beta \right) - c_{25} \eta_0, \checkmark \\ \mathcal{D}_{xx} &= c_{25}, \checkmark \\ \mathcal{D}_{xv} &= c_{26} - c_{25} h, \checkmark \\ \mathcal{D}_{vv} &= c_{33} + c_{25} h^2 - 2c_{26} h, \checkmark \quad (\text{B18}) \end{aligned}$$

with auxiliary quantities defined as

$$\begin{aligned} c_1 &= D - r_0 G_3, \checkmark \\ c_2 &= \eta_0 G_0 + G_1 \zeta, \checkmark \\ c_3 &= Dk + G_1 r_0^2, \checkmark \\ c_4 &= \eta_0 G_1 + 2G_0 r_0, \checkmark \\ c_5 &= \frac{r_0 - k G_2}{r G_1}, \checkmark \\ c_6 &= \frac{r_0 G_0 - k G_2}{\beta}, \checkmark \\ c_7 &= G_2 \left( \frac{1}{G_1} + \frac{c_2}{r} \right), \checkmark \\ c_8 &= r_0^{-3} (k c_6 + r r_0 + c_3 c_5), \checkmark \\ c_9 &= 2h G_2 - 3r_0 G_3, \checkmark \\ c_{10} &= \frac{k}{r_0^4} \left[ -G_2 r_0 h + \frac{k c_9}{\beta} - \frac{c_3 c_{13}}{r} \right], \checkmark \\ c_{12} &= G_0 h - G_1 r_0, \checkmark \\ c_{13} &= G_1 h - G_2 r_0, \checkmark \end{aligned}$$

$$\begin{aligned}
 c_{17} &= r_0 - r - kG_2, \checkmark \\
 c_{18} &= \eta_0 G_1 + 2kG_2, \checkmark \\
 c_{20} &= k(kG_2 + r) - G_0 r_0 \zeta, \checkmark \\
 c_{21} &= \frac{(kG_2 - r_0)(\beta c_3 - kG_1 r)}{\beta r^2 r_0^3 G_1} + \frac{\eta_0 G_1}{r r_0^2} - \frac{2}{r_0^2}, \checkmark \\
 c_{22} &= \frac{1}{r} \left[ -G_1 - \frac{G_0 G_2}{G_1} + \frac{G_2 c_2}{r} \right], \checkmark \\
 c_{24} &= \frac{1}{r_0^3} \left[ r_0 \left( \frac{2k}{\beta r_0} - 1 \right) - \frac{G_1 c_3}{r G_2} \right], \checkmark \\
 c_{25} &= \frac{k}{r r_0^2} \left[ -G_2 + \frac{k(c_{13} - G_2 r_0)}{\beta r_0^2} - \frac{c_{13}}{r_0} - \frac{c_{12} c_3}{r r_0^2} \right. \\
 &\quad \left. + \frac{c_{13} c_2 c_3}{r^2 r_0^2} - \frac{c_{13}(k(kG_2 + r) - G_0 r_0 \zeta)}{\beta r r_0^2} \right], \checkmark \\
 c_{26} &= \frac{k}{r^2 r_0} \left[ -G_2 c_{12} - G_1 c_{13} + \frac{G_2 c_{13} c_2}{r} \right], \checkmark \\
 c_{33} &= \frac{Dk^2}{r^3 r_0} (hG_2 - r_0 G_3) \\
 &\quad + \frac{k}{\beta r^2 r_0} \left[ -\eta_0 k G_1 G_2^2 - G_1 G_2 G_3 k^2 \right. \\
 &\quad \left. - r_0 \eta_0 \beta G_1 G_2^2 - r_0 k G_1 H_2 - \beta G_2^2 G_0 r_0^2 \right], \checkmark \\
 c_{34} &= \frac{1}{\beta r^2} \left[ -\beta(\eta_0 G_2 + r_0 G_1)^2 - \eta_0 k H_8 - H_6 k^2 \right. \\
 &\quad \left. + (G_2^2 - 3G_1 G_3) \beta k r_0 \right] + \frac{\eta_0 G_2^2}{r G_1} + \frac{k H_8}{\beta r G_1}, \checkmark \tag{B19}
 \end{aligned}$$

where we have dropped the  $\hat{\phantom{x}}$  superscript on the right-hand side of equations (B15)–(B19) for legibility, but note that all *scalar* quantities in these equations which are a function of  $\mathbf{x}_0$  *must* be evaluated as a function of  $\hat{\mathbf{x}}_0$  in lieu of  $\mathbf{x}_0$ . In these equations, we have taken care to analytically cancel terms to leading order in  $\gamma$  by defining the following functions

$$\begin{aligned}
 H_3 &= G_1 G_2 - 3G_3, \checkmark \\
 H_4 &= -\beta H_1, \checkmark \\
 H_5 &= G_1 G_2 - (2 + G_0) G_3, \checkmark \\
 H_6 &= 2G_2^2 - 3G_1 G_3, \checkmark \\
 H_7 &= G_1 G_2 - 2G_0 G_1 G_2 + 3G_0^2 G_3 = \beta G_1 G_2^2 - G_0 H_8, \checkmark \\
 H_8 &= G_1 G_2 - 3G_0 G_3 = -2H_3 + 3H_5. \checkmark \tag{B20}
 \end{aligned}$$

These functions are tabulated in Table 1. As with  $G_3$ ,  $H_1$ , and  $H_2$ , for small values of  $\gamma$  we evaluate these with a series expansion:

$$H_3(\gamma, \beta) = -\frac{4\gamma^5}{\beta|\beta|^{1/2}} \sum_{n=0}^{\infty} \frac{(\varepsilon\gamma^2)^n (4^{n+1} - 1)}{(2n+5)!}, \checkmark \tag{B21}$$

$$H_5(\gamma, \beta) = -\frac{2\gamma^5}{\beta|\beta|^{1/2}} \sum_{n=0}^{\infty} \frac{(\varepsilon\gamma^2)^n (n+1)}{(2n+5)!}, \checkmark \tag{B22}$$

$$H_6(\gamma, \beta) = \frac{2\gamma^6}{\beta^2} \sum_{n=0}^{\infty} \frac{(\varepsilon\gamma^2)^n (4^{n+2} - 3n - 7)}{(2n+6)!}, \checkmark \tag{B23}$$

where  $\varepsilon = -1$  for  $\beta > 0$  (elliptic) and  $\varepsilon = 1$  for  $\beta < 0$  (hyperbolic) cases. The coefficients of these series are computed recursively for efficiency. Note that we found expressions for  $H_4$ ,  $H_7$ , and  $H_8$  in terms of the other  $G$  and  $H$  functions.

### B3 Differential of Kepler then drift step

The Kepler step followed by a negative drift is slightly simpler as the Gauss functions can be expressed in terms of  $\mathbf{x}_0$  rather than  $\hat{\mathbf{x}}_0$ . The differential of the scalar functions are given by

$$\begin{aligned}
 \delta(f - hf - 1) &= \delta k \mathcal{E}_k + \delta h \mathcal{E}_h + \delta \mathbf{x}_0 \cdot \mathbf{x}_0 \mathcal{E}_{xx} \\
 &\quad + \delta(\mathbf{x}_0 \cdot \mathbf{v}_0) \mathcal{E}_{xv} + \delta \mathbf{v}_0 \cdot \mathbf{v}_0 \mathcal{E}_{vv}, \checkmark
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{E}_k &= \frac{k}{r r_0} \left[ -\frac{c_{14} c_{17}}{\beta r r_0} - \frac{2G_2}{\beta} + \frac{4H_1 k}{\beta r_0} \right. \\
 &\quad \left. - \frac{c_1 c_{14} c_2}{r^2 r_0} + \frac{c_1(G_1 r_0 - H_2 k)}{r r_0} + \frac{c_{14}}{k} - H_1 \right], \checkmark \\
 \mathcal{E}_h &= \frac{k}{r^2} \left[ -\frac{c_{14} c_2}{r r_0} + G_1 - \frac{H_2 k}{r_0} \right], \checkmark \\
 \mathcal{E}_{xx} &= \frac{k}{\beta r^3 r_0^4} \left[ \beta c_3 (c_{14} c_2 + c_{23} r) + H_1 k r^2 r_0 \left( \beta - \frac{2k}{r_0} \right) \right. \\
 &\quad \left. + c_{14} r (G_0 r_0 \zeta + k(r - G_2 k)) \right], \checkmark \\
 \mathcal{E}_{xv} &= \frac{k}{r^2 r_0} \left[ \frac{c_{14} c_2 G_2}{r} + k(G_1 H_1 + G_2 H_2) - 2G_1 G_2 r_0 \right], \checkmark \\
 \mathcal{E}_{vv} &= \frac{k}{\beta r^2 r_0} \left[ 2\eta_0 k (G_2 G_3 - G_1 H_1) \right. \\
 &\quad + (3G_3 H_2 - 4H_1 G_2) k^2 + \beta G_2 r_0 (3H_1 k - G_2 r_0) \\
 &\quad + \frac{c_{14}}{r} \left( -\beta(G_2 \eta + G_1 r_0)^2 + \eta_0 k (2G_0 G_3 - H_2) \right. \\
 &\quad \left. + k \beta r_0 (H_1 - 2G_1 G_3) - H_6 k^2 \right)], \checkmark \tag{B24}
 \end{aligned}$$

$$\begin{aligned}
 \delta(g - hg) &= \delta k \mathcal{F}_k + \delta h \mathcal{F}_h + \delta \mathbf{x}_0 \cdot \mathbf{x}_0 \mathcal{F}_{xx} \\
 &\quad + \delta(\mathbf{x}_0 \cdot \mathbf{v}_0) \mathcal{F}_{xv} + \delta \mathbf{v}_0 \cdot \mathbf{v}_0 \mathcal{F}_{vv}, \checkmark \\
 \mathcal{F}_k &= \frac{k}{r} \left[ \frac{c_{15}}{k} - \frac{c_{15} c_{17}}{\beta r r_0} - \frac{c_{19}}{\beta r_0} - \frac{c_1 c_{15} c_2}{r^2 r_0} + \frac{c_1 c_{16}}{r r_0} \right], \checkmark
 \end{aligned}$$

$$\mathcal{F}_h = \frac{k}{r^3} [c_{16} r - c_{15} c_2], \checkmark$$

$$\begin{aligned}
 \mathcal{F}_{xx} &= \frac{k}{r r_0} \left[ -\frac{c_{15}(k(G_2 k + r) - G_0 r_0 \zeta)}{\beta r r_0^2} + \frac{c_{19} k}{\beta r_0^2} \right. \\
 &\quad \left. + \frac{c_{15} c_2 c_3}{r^2 r_0^2} - \frac{c_{16} c_3}{r r_0^2} + H_2 \right], \checkmark
 \end{aligned}$$

$$\mathcal{F}_{xv} = \frac{k}{r^2} \left[ \frac{c_{15} c_2 G_2}{r} - c_{15} G_1 - c_{16} G_2 + H_1 r \right], \checkmark$$

$$\begin{aligned}
 \mathcal{F}_{vv} &= \frac{k}{\beta r^2} \left[ 2\eta_0^2 (G_1 H_1 - G_2 G_3) + \eta_0 k (4G_2 H_1 - 3H_2 G_3) \right. \\
 &\quad + r_0 \eta_0 (4G_0 H_1 - 2G_1 G_3) \\
 &\quad + 3r_0 k ((G_1 + \beta G_3) H_1 - G_3 G_2) \\
 &\quad + (G_0 H_8 - \beta G_1 (G_2^2 + G_1 G_3)) r_0^2 \\
 &\quad - \frac{c_{15}}{r} \left( \beta G_2^2 \eta_0^2 + \eta_0 k H_8 + H_6 k^2 + \beta G_1^2 r_0^2 \right. \\
 &\quad \left. + (2\eta_0 G_1 G_2 - k(G_2^2 - 3G_1 G_3)) \beta r_0 \right)], \checkmark \tag{B25}
 \end{aligned}$$

$$\begin{aligned}
 \delta \ln \dot{f} &= \delta k \mathcal{G}_k + \delta h \mathcal{G}_h + \delta \mathbf{x}_0 \cdot \mathbf{x}_0 \mathcal{G}_{xx} \\
 &\quad + \delta (\mathbf{x}_0 \cdot \mathbf{v}_0) \mathcal{G}_{xv} + \delta \mathbf{v}_0 \cdot \mathbf{v}_0 \mathcal{G}_{vv}, \checkmark \\
 \mathcal{G}_k &= \frac{1}{k} + \frac{c_1(r_0 - kG_2)}{r_0 r^2 G_1} - \frac{r + c_{17}}{\beta r r_0}, \checkmark \\
 \mathcal{G}_h &= \frac{r_0 - kG_2}{r^2 G_1}, \checkmark \\
 \mathcal{G}_{xx} &= \frac{1}{r_0^2} \left[ -\frac{k(G_2 k - r_0)}{\beta r r_0} + \frac{c_2 c_3}{r^2 r_0} - \frac{c_3 G_0}{G_1 r r_0} + \frac{\eta_0 G_1}{r} - 2 \right], \checkmark \\
 \mathcal{G}_{xv} &= -\frac{1}{r} \left( \frac{G_0 G_2}{G_1} + \frac{r_0 G_1 + \eta_0 G_2}{r} \right), \checkmark \\
 \mathcal{G}_{vv} &= \frac{1}{\beta r^2} \left[ \frac{(\eta_0 G_2^2 \beta + k H_8)(r_0 G_0 + k G_2)}{G_1} - H_6 k^2 \right. \\
 &\quad \left. + \beta r_0 (k(H_1 - 2G_1 G_3) - 2\eta_0 G_1 G_2) - \beta G_1^2 r_0^2 \right], \checkmark \quad (\text{B26})
 \end{aligned}$$

$$\begin{aligned}
 \delta(\dot{g} - 1) &= \delta k \mathcal{H}_k + \delta h \mathcal{H}_h + \delta \mathbf{x}_0 \cdot \mathbf{x}_0 \mathcal{H}_{xx} \\
 &\quad + \delta (\mathbf{x}_0 \cdot \mathbf{v}_0) \mathcal{H}_{xv} + \delta \mathbf{v}_0 \cdot \mathbf{v}_0 \mathcal{H}_{vv}, \checkmark \\
 \mathcal{H}_k &= \frac{1}{r r_0} \left[ \frac{G_2 k (-G_2 k + r + r_0)}{\beta r} + \frac{c_1 c_2 G_2 k}{r^2} \right. \\
 &\quad \left. - \frac{c_1 G_1 k}{r} - G_2 r_0 \right], \checkmark \\
 \mathcal{H}_h &= \frac{k}{r^3} [c_2 G_2 - G_1 r], \checkmark \\
 \mathcal{H}_{xx} &= \frac{k}{r^2 r_0^3} \left[ \frac{G_2 (k(G_2 k - r) - G_0 r_0 \zeta)}{\beta} \right. \\
 &\quad \left. + \frac{c_3 (\eta_0 G_2 + G_1 r_0)}{r} \right], \checkmark \\
 \mathcal{H}_{xv} &= \frac{G_2 k}{r^3} [r G_1 + r_0 G_1 + \eta_0 G_2], \checkmark \\
 \mathcal{H}_{vv} &= \frac{k}{\beta r^3} \left[ \eta_0^2 \beta G_2^3 - \eta_0 k G_2 H_3 + 3 r_0 \eta_0 \beta G_1 G_2^2 \right. \\
 &\quad \left. + r_0 k (3 \beta G_1 G_2 G_3 - G_0 H_6) \right. \\
 &\quad \left. + \beta G_2 (G_0 G_2 + G_1^2) r_0^2 \right], \checkmark \quad (\text{B27})
 \end{aligned}$$

with additional auxiliary definitions,

$$\begin{aligned}
 c_{14} &= r_0 G_2 - k H_1, \checkmark \\
 c_{15} &= \eta_0 H_1 + H_2 r_0, \checkmark \\
 c_{16} &= \eta_0 H_2 + G_1 \gamma r_0 |\beta|^{-1/2}, \checkmark \\
 c_{17} &= r_0 - r - k G_2, \checkmark \\
 c_{19} &= 4 \eta_0 H_1 + 3 H_2 r_0, \checkmark \\
 c_{23} &= k H_2 - r_0 G_1. \checkmark \quad (\text{B28})
 \end{aligned}$$

Note that in this case, as the Kepler step takes place first, all of the scalar quantities in this equation are defined in terms of  $\mathbf{x}_0$ .

### APPENDIX C: KEPLER + DRIFT MASS DERIVATIVE EXPRESSIONS

Here are the functions  $J_1 - J_4$  used in computing the mass derivatives in equations (24) and (25):

$$\begin{aligned}
 J_1 &= \hat{r}_0 \hat{H}_4 + k \hat{H}_6, \checkmark \\
 J_2 &= \hat{H}_6 \hat{G}_3 k^2 + \hat{\eta}_0 \hat{r}_0 (\hat{H}_6 + \hat{G}_2 \hat{H}_4) + \hat{r}_0^2 \hat{G}_0 \hat{H}_5 + k \hat{\eta}_0 \hat{G}_2 \hat{H}_6 \\
 &\quad + (\hat{G}_1 \hat{H}_6 + \hat{G}_3 \hat{H}_4) k \hat{r}_0, \checkmark \\
 J_3 &= -(\hat{G}_2 k - \hat{r}_0) (\beta \hat{r}_0 (\hat{G}_3 - \hat{G}_1 \hat{G}_2) - \beta \hat{\eta}_0 \hat{G}_2^2 + k \hat{H}_3), \checkmark \\
 J_4 &= k \left( -\beta \hat{\eta}_0^2 \hat{G}_2^4 + \hat{\eta}_0 \hat{G}_2 (\hat{G}_1 \hat{G}_2^2 + \hat{G}_1^2 \hat{G}_3 - 5 \hat{G}_2 \hat{G}_3) k \right. \\
 &\quad \left. + \hat{G}_2 \hat{G}_3 \hat{H}_3 k^2 + 2 \hat{\eta}_0 \hat{r}_0 \beta \hat{G}_2^2 (\hat{G}_3 - \hat{G}_1 \hat{G}_2) \right. \\
 &\quad \left. + (4 \hat{G}_3 - \hat{G}_0 \hat{G}_3 - \hat{G}_1 \hat{G}_2) (\hat{G}_3 - \hat{G}_1 \hat{G}_2) \hat{r}_0 k \right. \\
 &\quad \left. - \beta (\hat{G}_3 - \hat{G}_1 \hat{G}_2)^2 \hat{r}_0^2 \right), \checkmark \quad (\text{C1})
 \end{aligned}$$

and the functions  $J_5 - J_8$  used in computing the mass derivatives in equations (26) and (27):

$$\begin{aligned}
 J_5 &= r \left( 2 \eta_0 k (G_1 H_1 - G_3 G_2) + (4 G_2 H_1 - 3 G_3 H_2) k^2 \right. \\
 &\quad \left. - \eta_0 r_0 \beta G_1 H_1 + (G_3 H_2 - 4 G_2 H_1) \beta k r_0 + G_2 H_1 \beta^2 r_0^2 \right) \\
 &\quad - c_{14} \left( -\eta_0^2 \beta G_2^2 - k \eta_0 H_8 - \eta_0 r_0 \beta (G_1 G_2 + G_0 G_3) \right. \\
 &\quad \left. + 2 (H_1 - G_1 G_3) \beta k r_0 - k^2 H_6 - (G_2 - \beta G_1 G_3) \beta r_0^2 \right), \checkmark \\
 J_6 &= r_0 r \left( 2 \eta_0^2 (G_3 G_2 - G_1 H_1) + \eta_0 k (3 G_3 H_2 - 4 G_2 H_1) \right. \\
 &\quad \left. + r_0 \eta_0 (\beta G_3 (G_1 G_2 + G_0 G_3) - 2 G_0 H_6) \right. \\
 &\quad \left. + (-H_6 (G_1 + \beta G_3) + G_2 (2 G_3 - H_2)) r_0 k \right. \\
 &\quad \left. + (H_7 - \beta^2 G_1 G_3^2) r_0^2 \right) \\
 &\quad - r_0 c_{15} \left( -\beta \eta_0^2 G_2^2 + \eta_0 k (-H_2 + 2 G_0 G_3) - H_6 k^2 \right. \\
 &\quad \left. - r_0 \eta_0 \beta (H_2 + 2 G_0 G_3) + 2 \beta (2 H_1 - G_2^2) r_0 k \right. \\
 &\quad \left. + \beta (\beta G_1 G_3 - G_2) r_0^2 \right), \checkmark \\
 J_7 &= (r_0 - k G_2) (-\eta_0 \beta G_2^2 + H_3 k + (G_3 - G_1 G_2) \beta r_0), \checkmark \\
 J_8 &= \beta G_1 (G_3 - G_1 G_2) r_0^2 - \beta \eta_0^2 G_2^3 + \eta_0 k G_2 H_3 \\
 &\quad + \eta_0 r_0 \beta G_2 (G_3 - 2 G_1 G_2) + (H_6 - \beta G_2^3) r_0 k. \checkmark \quad (\text{C2})
 \end{aligned}$$

### APPENDIX D: TABLE OF NOTATION

Table D1 lists the mathematical symbols used throughout this paper.

**Table D1.** Symbols used in this paper.

Symbol	Definition	Reference
$\mathbf{a}_i$	Instantaneous acceleration of body $i$ .	(38)
$\mathbf{a}_{ij}$	Difference in acceleration between bodies $i$ and $j$ .	Section 4.8
$A$	Set of pairs of bodies interacting via fast kicks.	Section 4.7, Alg. 2
$A^C$	Set of pairs of bodies interacting drift + Keplerian (complement of $A$ ).	Section 4.7, Alg. 2
$\mathcal{A}_k, \mathcal{A}_h, \mathcal{A}_{xx}, \dots, \mathcal{H}_{xv}, \mathcal{H}_{vv}$	Terms used in derivatives.	(B15)-(B27)
au	Astronomical unit.	Section 4.2
$b_{\text{sky},ijk}$	The impact parameter at the $k$ th time of transit of body $i$ in front of body $j$ .	Section 4.11
$c_1 - c_{34}$	Quantities defined for computing derivatives of drift + Kepler.	(B19),(B28)
$C$	Constant in correction term.	Section 4.8
$d$	Distance to observer.	Section 4.3
$e$	Orbital eccentricity.	Section 5.1
$D$	Constant used in computing derivatives.	(B4)
$\Delta E, E_0$	Energy error and initial total energy.	Section 5.2
$f, g, \hat{f}, \hat{g}$	Gauss's Kepler propagation functions.	Section 3.2
$\hat{f}, \hat{g}, \hat{f}, \hat{g}$	Gauss's functions for drift-then-kepler.	Section 3.3.1
$g_{\text{sky},ij}$	Sky plane velocity dotted w/ position between bodies $i$ and $j$ .	(48)
$G$	Newton's constant.	(4)
$G_i$	Functions used in universal Kepler solver.	Table 1
$\hat{G}_i$	Functions used in universal Kepler solver after initial drift.	Section A1
$h$	Symplectic integrator time-step (days).	Section 2
$H$	Hamiltonian.	(1)
$H_A, H_B$	Integrable Hamiltonians in symplectic splitting.	(2)
$H_1 - H_8$	Combinations of $G_i$ functions used in combined Kepler-drift step and derivatives.	Table 1, (A5), (B20)
$i, j$	Label for bodies $i$ and $j$ .	(1)
$J_1 - J_8$	Intermediate variables used in mass derivatives.	(C1),(C2)
$\mathbf{J}_{\text{kep}}$	Jacobian of Kepler/drift step between two bodies ( $7 \times 7$ matrix).	Section 4.4
$\mathbf{J}_{\text{substep}}, \Delta \mathbf{J}_{\text{substep}}$	Jacobian of any substep is $\mathbf{J}_{\text{substep}} = \mathbf{I} + \Delta \mathbf{J}_{\text{substep}}$ .	(44)
$\mathbf{J}_{\text{current}}, \mathbf{J}_{\text{prior}}$	Jacobians of the current and prior steps.	(44)
$\Delta \mathbf{J}_D$	Jacobian of drift step is $\mathbf{I} + \Delta \mathbf{J}_D$ .	Section 4.6
$\mathbf{J}_{\text{AHL21}}$	Jacobian of single symplectic step.	Section 4.9
$\mathbf{J}_n$	Jacobian after $n$ th step.	Section 4.12
$\Delta \mathbf{J}_{4\text{th}}$	Jacobian of fourth-order correction $\mathbf{I} + \Delta \mathbf{J}_{4\text{th}}$ .	Section 4.8
$\Delta \mathbf{J}_{\text{DK},ij}, \Delta \mathbf{J}_{\text{KD},ij}$	Jacobian of Kepler + drift substep for bodies $i$ and $j$ .	Section 4.5
$k$	Central force constant ( $=G(m_i + m_j)$ ). Also used as an index for the transit number.	(4)
$K_{ij}$	Keplerian for pair of bodies $i$ and $j$ .	(1)
$m_i$	Mass of $i$ th body.	Section 4.1
$m_{ij}$	Sum of masses of bodies $i$ and $j$ .	Section 3.2
$\mathbf{m}$	Vector of masses.	Section 4.1
$M_\odot$	Solar mass.	Section 4.2
$n$	Number of time-steps elapsed.	Section 4.11
$N$	Number of bodies in system.	Section 4.1
$N_S$	Number of time-steps.	Section 5.2, Section 5.5
$N_{\text{tr}}$	Number of transit times.	Section 5.1
$\mathbf{q}_i(t)$	Coordinates of the $i$ th body at time $t$	(17)
$\mathbf{q}(t)$	Position, velocity, mass vector of all bodies (system state at time $t$ ).	Section 4.3
$\mathbf{q}_n$	$\mathbf{q}(t)$ at $n$ th time-step.	Section 4.11
$\mathbf{q}_{\text{current}}, \mathbf{q}_{\text{prior}}, \Delta \mathbf{q}$	System state at current and prior substep, and the difference: $\Delta \mathbf{q} = \mathbf{q}_{\text{current}} - \mathbf{q}_{\text{prior}}$	Section 4.10
$r_0, r, r_{ij}$	Initial/final separation between bodies $i$ and $j$ for Kepler solver.	Section 3.2
$\hat{r}_0, \hat{r}$	Initial/final separation between bodies $i$ and $j$ for drift + Kepler solver.	Section A1
$s$	Independent variable in universal Kepler step.	Section 3.2
$t$	Current simulation time.	Section 2, Alg. 1,2
$\Delta t_{\text{init}}, \Delta t$	Initial and final fraction of a time-step to transit time.	Section 4.11
$t_0$	Initial time of integration (days).	Section 2, Alg. 1,2
$t_n$	Time after $n$ th step: $t_n = t_0 + nh$ .	Section 4.11
$t_{ijk}$	The $k$ th time of transit of body $i$ in front of body $j$ .	Section 4.11
$t_{\text{max}}$	Duration of simulation (days).	Section 2, Alg. 1,2
$\mathbf{T}_{ij}$	Correction tensor.	(37)
$T$	Total kinetic energy.	(1)
$T_{ij}$	Kinetic energy of bodies $i$ and $j$ .	(1)
$\Delta \mathbf{v}_i$	Pairwise velocity kick or correction of $i$ th body (instead of Kepler step).	(32), (36), (42)
$v_0, v$	Initial/final relative speed in Kepler problem.	Section 3.2
$v_{\text{sky},ijk}$	The sky velocity at the $k$ th time of transit of body $i$ in front of body $j$ .	Section 4.11
$\mathbf{v}_i = (\dot{x}_i, \dot{y}_i, \dot{z}_i)$	Cartesian velocities of $i$ th body	(3), Section 4.3

Table D1 – continued

Symbol	Definition	Reference
$V$	Total potential energy.	(1)
$V_{ij}$	Potential energy of bodies $i$ and $j$ .	(1)
$w_{ij}$	Intermediate quantity for kick derivative.	(33)
$\mathbf{x}_i = (x_i, y_i, z_i)$	Cartesian coordinates of $i$ th body	(3), Section 4.3
$\mathbf{x}_{ij}, \mathbf{v}_{ij}$	Relative position and velocity of bodies $i$ and $j$ ( $\mathbf{x}, \mathbf{v}$ in Section 3.2).	(4)
$\mathbf{x}_0, \mathbf{v}_0$	Relative position and velocity of bodies $i$ and $j$ at start of universal Kepler step	Section 3.2
$\hat{\mathbf{x}}_0$	Intermediate position in DK step.	(14)
$\hat{\mathbf{x}}, \hat{\mathbf{v}}$	Final position and velocity in DK step.	Section 3.3.1
$\Delta \mathbf{x}_{\text{DK}}, \Delta \mathbf{v}_{\text{DK}}$	Change in position and velocity for DK step between bodies $i$ and $j$ .	(15),(A2)
$\check{\mathbf{x}}, \check{\mathbf{v}}$	Intermediate position and velocity in KD step.	Section 3.3.2
$\Delta \mathbf{x}_{\text{KD}}, \Delta \mathbf{v}_{\text{KD}}$	Change in position and velocity for KD step between bodies $i$ and $j$ .	(16),(A3)
$\Delta \mathbf{x}_{i,\text{KD}}, \Delta \mathbf{v}_{i,\text{KD}}$	Change in position and velocity for KD step for body $i$ .	(22)
$\Delta \mathbf{x}_{i,\text{DK}}, \Delta \mathbf{v}_{i,\text{DK}}$	Change in position and velocity for DK step for body $i$ .	(22)
$x_{\text{obs}}$	Observer position.	Section 4.3
$\alpha$	Factor from DH17 algorithm(set to zero).	Section 2, Section 4.8
$\beta$	Dimensionless energy for Kepler step.	(7)
$\hat{\beta}$	Dimensionless energy for drift + Kepler step.	(A1)
$\gamma$	Variable used in defining $G_i$ functions. $= \beta ^{1/2}s$	Table 1
$\hat{\gamma}$	Value of $\gamma$ computed after an initial drift.	Section A1
$\eta_0$	Quantity used in Kepler solver.	Section 3.2
$\hat{\eta}_0$	Dot product of velocity and position after an initial drift.	Section A1
$\epsilon$	Small parameter in Hamiltonian splitting for general symplectic integrator.	(2)
$\epsilon_{\text{err}}$	Numerical error.	Section 5.2
$\epsilon_{\text{diff}}$	Fractional change in parameters for finite difference.	(63)
$\varepsilon$	Sign of $\beta$ .	(A7–A9),(B21–B23)
$\zeta$	Intermediate variable.	(B7)
$\varpi$	Longitude of periastron.	Section 5.1

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