Time-symmetric integration in astrophysics

Citation: Hernandez, David M, and Edmund Bertschinger. "Time-Symmetric Integration in Astrophysics." Monthly Notices of the Royal Astronomical Society 475, no. 4 (January 24, 2018): 5570–5584. © 2017 The Authors

As Published: http://dx.doi.org/10.1093/MNRAS/STY184

Publisher: Oxford University Press (OUP)

Persistent URL: http://hdl.handle.net/1721.1/121023

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

MNRAS **000**, 000–000 (0000) Preprint 22 January 2018 Compiled using MNRAS LATEX style file v3.0

Time-symmetric integration in astrophysics

David M. Hernandez ⋆ and Edmund Bertschinger ⋆
Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, Massachusetts 02139, USA

22 January 2018

ABSTRACT
Calculating the long term solution of ordinary differential equations, such as those of the N-body problem, is central to understanding a wide range of dynamics in astrophysics, from galaxy formation to planetary chaos. Because generally no analytic solution exists to these equations, researchers rely on numerical methods which are prone to various errors. In an effort to mitigate these errors, powerful symplectic integrators have been employed. But symplectic integrators can be severely limited because they are not compatible with adaptive stepping and thus they have difficulty accommodating changing time and length scales. A promising alternative is time-reversible integration, which can handle adaptive time stepping, but the errors due to time-reversible integration in astrophysics are less understood. The goal of this work is to study analytically and numerically the errors caused by time-reversible integration, with and without adaptive stepping. We derive the modified differential equations of these integrators to perform the error analysis. As an example, we consider the trapezoidal rule, a reversible non-symplectic integrator, and show it gives secular energy error increase for a pendulum problem and for a Hénon–Heiles orbit. We conclude that using reversible integration does not guarantee good energy conservation and that, when possible, use of symplectic integrators is favored. We also show that time-symmetry and time-reversibility are properties that are distinct for an integrator.

Key words: methods: numerical—celestial mechanics—globular clusters: general— planets and satellites: dynamical evolution and stability— galaxies

1 INTRODUCTION

Obtaining solutions to initial value problems of ordinary differential equations (ODEs) over long time periods is central to dynamical calculations in astrophysics. These ODEs might represent problems such as the N-body problem, N point particles interacting through pairwise forces, or the problem of particle orbits in a time-independent galactic potential. The ODEs are frequently described by a time-dependent or time-independent Hamiltonian.

Obtaining a solution to the N-body problem is essential for many purposes, from calculating the evolution of dark matter in the Universe to understanding stability and chaos of orbits in planetary systems. Different techniques, relying on different assumptions, have been developed to obtain approximate N-body solutions. The N-body problem is generally chaotic and non-integrable, so we rely on these approximations to obtain its solutions. But, in general N-body cases, it is unknown how reliable such approximations are. In fact, the approximations themselves give rise to chaos, separate from the physical chaos of the problem itself. If the numerical method itself can be responsible for chaos, then the error from the original trajectory can grow exponentially, leading to call into question the validity of the calculated solution.

Galactic potentials usually have only a few degrees of freedom, but can still be chaotic and non-integrable, and suffer from the same problems described above. In fact, much of the study of chaos began with the study of the Hénon–Heiles problem, which was motivated by the study of galactic potentials.

It would appear numerical approximation to chaotic ODEs should be suspect, but fortunately, geometric numerical integration, integration aimed at respecting the geometry of the underlying ODEs, has been developed and helps restore confidence in these numerical solutions (Channell & Scovel 1990). Depending on the equations, geometric properties include the Hamiltonian flow, time-reversibility, and quadratic and linear invariants in the phase space. In the last 30 years, astrophysics researchers have made geometric integration a standard in various fields of dynamics, including planets (Wisdom & Holman 1991; Chambers 1999; Duncan et al. 1998; Hernandez 2016), stellar clusters (Kokubo et al. 1998; Hut et al. 1995; Hernandez & Bertschinger 2015; Dehnen & Hernandez 2017), or galaxy formation (Springel 2005).

Geometric integrators that respect Hamiltonian flow are also called symplectic integrators, and they conserve general-

⋆ Email: dmhernan@mit.edu (DMH); edbert@mit.edu (EB)

izations of volumes in phase space, also known as Poincaré invariants. The theory of symplectic integration is well developed (Hairer et al. 2006). The citations above (Wisdom & Holman 1991; Chambers 1999; Duncan et al. 1998; Hernandez 2016; Kokubo et al. 1998; Hut et al. 1995; Hernandez & Bertschinger 2015; Dehnen & Hernandez 2017; Springel 2005), are all concerned with time-independent Hamiltonian problems, so a symplectic integrator is ideal. However, symplectic integrators applied to the above problems have a limitation; if the step sizes are chosen as a function of the phase space, the evolution of the trajectory is no longer Hamiltonian. This limitation is severe for the N-body problem because gravity has no length scale: two-body relaxation is affected by close and far encounters. Thus, the range of time and length scales is large, posing a severe challenge for fixed time step integration.

Thus, some researchers (Pelupessy et al. 2012; Hut et al. 1995; Kokubo et al. 1998) have abandoned the requirement of a symplectic integrator and instead focused on integrators that preserve time-reversibility, if the underlying equations have this symmetry. It is important to note there exist irreversible conservative differential equations— see Section 2.2. However, many important problems such as the N-body problem are conservative and reversible. Time-reversible integration appears to be less studied than symplectic integration, but it has been observed that time-reversible integrators generally can reduce errors for integrations in astrophysics. An explanation for such behavior is sometimes not provided. It is possible to adapt time steps while still preserving time-reversibility (Makino et al. 2006; Funato et al. 1996), so clearly we would like to abandon symplecticity if possible and if time-reversible integration is good enough. But a clear error analysis is needed with these integrators in order for one to have confidence in their use.

The goal of this paper is to provide that error analysis, and to use it to show that the behavior of a time-reversible integrator in astrophysics can be worse than a symplectic integrator. This suggests that researchers in astrophysics should use symplectic integrators when possible. To perform the error analysis, we derive the modified differential equations (MDEs) obeyed by these methods using adaptive time steps, and we use these equations to calculate how well the methods conserve energy. We study a simple pendulum problem and Hénon–Heiles orbits. We show how various reversible integrators do not conserve energy to all orders. It was already noted by Faou et al. (2004) that some fixed-step Runge–Kutta reversible methods do not conserve an energy.

Section 2.1 shows the tools necessary for deriving the MDE. Section 3 derives the MDE for the trapezoidal rule, a non-symplectic but symmetric second order Runge–Kutta method. We also derive the MDE for the trapezoidal rule with adaptive steps. We derive various properties of the trapezoidal rule. In Section 4, we apply our numerical analysis to understand the error in energy of the modified pendulum problem and Hénon–Heiles orbits and to find that time-symmetric integration can give energy drift. While the analysis in this section is limited to the trapezoidal rule, there is no reason to believe other time symmetric integrators would not suffer from energy drift. In the Appendices, we show that for Runge–Kutta methods, time-symmetry, reversibility, and symplecticity are independent concepts. We conclude in Section 5.

2 THE MODIFIED DIFFERENTIAL EQUATION

2.1 Time-symmetric integration

A system of autonomous ordinary differential equations can be written

$$\dot{\boldsymbol{y}} = \boldsymbol{f}(\boldsymbol{y}) \tag{1}$$

where $\boldsymbol{y}$ and $\boldsymbol{f}$ are both vectors of length n. We are concerned with the case where $\boldsymbol{y}$ is a vector of positions and velocities or the phase space defined by canonical coordinates and momenta: in equations, $\boldsymbol{y} = (\boldsymbol{q}, \boldsymbol{v})$ or $\boldsymbol{y} = (\boldsymbol{q}, \boldsymbol{p})$. The problem is to find $\boldsymbol{y}(t)$ given $\boldsymbol{y}(0) \equiv \boldsymbol{y}_0$. We assume that the system is autonomous, so that $\boldsymbol{f}$ depends only on $\boldsymbol{y}$ and not on t.

A numerical one-step method estimates the solution at $t = h$, $\boldsymbol{y}(h) \approx \boldsymbol{y}_1$, where

$$\boldsymbol{y}_1 = \boldsymbol{y}_0 + h\boldsymbol{G}(\boldsymbol{y}_0, h) \tag{2}$$

for some $\boldsymbol{G}$ that is related to $\boldsymbol{f}$. The method may be iterated to estimate the solution at $t = 2h$, $3h$, etc. For now, we assume that h is a constant independent of $\boldsymbol{y}_0$ and t.

A goal of numerical analysis is to find a $\boldsymbol{G}$ that is inexpensive to evaluate so that $|\boldsymbol{y}_1 - \boldsymbol{y}(h)|$ is smaller than a specified tolerance. The study of the errors $|\boldsymbol{y}_1 - \boldsymbol{y}(h)|$ is known as forward error analysis. One can turn around the problem. Given $\boldsymbol{G}$, find a modified differential equation whose exact solution is $\boldsymbol{y}(h) = \boldsymbol{y}_1$. That modified differential equation is written

$$\dot{\boldsymbol{y}} = \boldsymbol{F}(\boldsymbol{y}, h)\ . \tag{3}$$

The goal then becomes to minimize $|\boldsymbol{F}(\boldsymbol{y},h) - \boldsymbol{f}(\boldsymbol{y})|$. This is done by determining $\boldsymbol{F}(\boldsymbol{y},h)$ from $\boldsymbol{G}(\boldsymbol{y},h)$. Studying the errors $|\boldsymbol{F}(\boldsymbol{y},h) - \boldsymbol{f}(\boldsymbol{y})|$ is known as backward error analysis.

A symmetric one-step integrator is one for which a forward step h followed by a backward step $-h$ restores the initial conditions. The requirement is

$$\boldsymbol{G}(\boldsymbol{y}_1, -h) = \boldsymbol{G}(\boldsymbol{y}_0, h) = \frac{1}{h}(\boldsymbol{y}_1 - \boldsymbol{y}_0)\ . \tag{4}$$

The associated modified differential equation is even: $\boldsymbol{F}(\boldsymbol{y}, -h) = \boldsymbol{F}(\boldsymbol{y}, h)$

ρ-reversibility (Hairer et al. 2006, Section V.1) means that if we change the sign of velocities, while keeping the position coordinates constant, the solution trajectory must stay the same— only the direction of motion is inverted. Let $\boldsymbol{\rho}$ be an invertible linear transformation that changes the signs of velocities: $\boldsymbol{\rho y} = \boldsymbol{\rho}(\boldsymbol{q}, \boldsymbol{v}) = (\boldsymbol{q}, -\boldsymbol{v})$. All autonomous Newtonian physics problems are described by positions and velocities and can be written as a system of first order ODE's: $\dot{\boldsymbol{q}} = \boldsymbol{f}(\boldsymbol{q}, \boldsymbol{v})$ and $\dot{\boldsymbol{v}} = \boldsymbol{g}(\boldsymbol{q}, \boldsymbol{v})$. They are not all reversible; if they are, then,

$$\begin{aligned} &\boldsymbol{f}(\boldsymbol{q}, -\boldsymbol{v}) = -\boldsymbol{f}(\boldsymbol{q}, \boldsymbol{v}), \qquad \text{and} \\ &\boldsymbol{g}(\boldsymbol{q}, -\boldsymbol{v}) = \boldsymbol{g}(\boldsymbol{q}, \boldsymbol{v}). \end{aligned} \tag{5}$$

While many problems satisfy this requirement, not all do. For example, the system of differential equations for a charged particle moving in a magnetic field are

$$\begin{aligned} &\dot{\boldsymbol{v}} = \frac{e}{m}(\boldsymbol{v} \times \boldsymbol{B}(\boldsymbol{q}, t)), \qquad \text{and} \\ &\dot{\boldsymbol{q}} = \boldsymbol{v}, \end{aligned} \tag{6}$$

where e is the charge of the particle, m is the mass of the particle, and $\boldsymbol{B}$ is the external magnetic field. These equations do not satisfy (5); the solution trajectory is different when we switch the sign of

the velocities (the resolution of this apparent irreversibility is that the sign of $\boldsymbol{B}$ changes if we reverse the currents causing it).

If we use a one-step method to solve a ρ-reversible set of differential equations, the symmetric integrator is ρ-reversible. The ρ-reversibility condition for an integrator is connected to (5):

$$\boldsymbol{\rho}\phi_h\boldsymbol{y} = \phi_{-h}\boldsymbol{\rho}\boldsymbol{y}, \tag{7}$$

which implies $\phi_h\boldsymbol{\rho}\phi_h = \phi_h\phi_{-h}\boldsymbol{\rho}$. This only holds if the integrator is time-symmetric, or $\phi_h\phi_{-h} = I$. Thus, in what follows, until Section 3.4, 'symmetric' one-step methods will be equivalent to'time-reversible' one-step methods because we are only concerned with ρ-reversible differential equations. However, it is important to bear in mind that a symmetric method is not necessarily the same as a time-reversible method; one way to break the equivalency is by letting the step h vary as a function of phase space.

2.2 Derivation of modified differential equation

Our goal is to understand time-symmetric integrators. Unlike symplectic methods, symmetric integrators generally have no surrogate Hamiltonian (Hairer et al. 2006, Section IX.8) which informs us of the dynamics, so we instead derive the differential equations the integrator obeys. We call this the modified differential equations (MDEs), and its study has been referred to as backward error analysis (Hairer et al. 2006, Chapter IX).

Proceed as follows: first write the formally exact solution of equation (3) with initial condition $\boldsymbol{y} = \boldsymbol{y}_0$,

$$\boldsymbol{y}_1 = \exp(h\tilde{D})\boldsymbol{y}_0 = \boldsymbol{y}_0 + \sum_{n=1}^{\infty}\frac{h^n}{n!}\tilde{D}^{n-1}\boldsymbol{F}(\boldsymbol{y}_0,h)\,,\quad \tilde{D}\equiv\boldsymbol{F}(\boldsymbol{y}_0,h)\cdot\frac{\partial}{\partial\boldsymbol{y}_0}\,. \tag{8}$$

This is just the usual Taylor expansion solution. Next expand $\boldsymbol{F}(\boldsymbol{y}_0,h)$ and $\boldsymbol{G}(\boldsymbol{y}_0,h)$ in power series in h:

$$\boldsymbol{F}(\boldsymbol{y}_0,h) = \sum_{n=0}^{\infty}h^n\boldsymbol{f}_n(\boldsymbol{y}_0)\,,\quad \boldsymbol{G}(\boldsymbol{y}_0,h) = \sum_{n=0}^{\infty}h^n\boldsymbol{g}_n(\boldsymbol{y}_0)\,. \tag{9}$$

$\boldsymbol{f}_0$ is the $\boldsymbol{f}$ from (1). Use (9) to expand the derivative operator

$$\tilde{D} = \sum_{n=0}^{\infty}h^nD_n\,,\quad D_n\equiv\boldsymbol{f}_n(\boldsymbol{y}_0)\cdot\frac{\partial}{\partial\boldsymbol{y}_0}\,. \tag{10}$$

Combining equations (2) and (8)–(10) gives, for $n\leqslant 4$,

$$\begin{aligned}
\boldsymbol{g}_0 &= \boldsymbol{f}_0\\
\boldsymbol{g}_1 &= \boldsymbol{f}_1 + \frac{1}{2}D_0\boldsymbol{f}_0\\
\boldsymbol{g}_2 &= \boldsymbol{f}_2 + \frac{1}{2}(D_0\boldsymbol{f}_1 + D_1\boldsymbol{f}_0) + \frac{1}{6}D_0^2\boldsymbol{f}_0\\
\boldsymbol{g}_3 &= \boldsymbol{f}_3 + \frac{1}{2}(D_0\boldsymbol{f}_2 + D_1\boldsymbol{f}_1 + D_2\boldsymbol{f}_0) + \frac{1}{6}(D_0^2\boldsymbol{f}_1 + D_0D_1\boldsymbol{f}_0 + D_1D_0\boldsymbol{f}_0)\\
&\quad + \frac{1}{24}D_0^3\boldsymbol{f}_0\\
\boldsymbol{g}_4 &= \boldsymbol{f}_4 + \frac{1}{2}(D_0\boldsymbol{f}_3 + D_1\boldsymbol{f}_2 + D_2\boldsymbol{f}_1 + D_3\boldsymbol{f}_0)\\
&\quad + \frac{1}{6}(D_0^2\boldsymbol{f}_2 + D_0D_2\boldsymbol{f}_0 + D_2D_0\boldsymbol{f}_0 + D_0D_1\boldsymbol{f}_1 + D_1D_0\boldsymbol{f}_1 + D_1^2\boldsymbol{f}_0)\\
&\quad + \frac{1}{24}(D_0^3\boldsymbol{f}_1 + D_0^2D_1\boldsymbol{f}_0 + D_0D_1D_0\boldsymbol{f}_0 + D_1D_0^2\boldsymbol{f}_0)\\
&\quad + \frac{1}{120}D_0^4\boldsymbol{f}_0\,.
\end{aligned} \tag{11}$$

Our goal is to obtain $\boldsymbol{F}$ from $\boldsymbol{G}$. One way is to solve equations (11) recursively, starting with $\boldsymbol{f}_0 = \boldsymbol{g}_0$ substituting into $\boldsymbol{f}_1 = \boldsymbol{g}_1 - \frac{1}{2}D_0\boldsymbol{f}_0$, and so on. This is useful for determining $\boldsymbol{f}_n$ for small n.

As an example, consider the explicit Euler method

$$\boldsymbol{y}_1 = \boldsymbol{y}_0 + h\boldsymbol{f}(\boldsymbol{y}_0)\,, \tag{12}$$

for which $\boldsymbol{g}_0 = \boldsymbol{f}$, $\boldsymbol{g}_1 = \boldsymbol{g}_2 = \boldsymbol{g}_3 = 0$. This method is first order because $\boldsymbol{F}(\boldsymbol{y},h) = \boldsymbol{f}(\boldsymbol{y}) - \frac{1}{2}hD_0\boldsymbol{f}(\boldsymbol{y}) + O(h^2)$. For an nth order method,

$$\boldsymbol{g}_k = \frac{1}{(k+1)!}D_0^k\boldsymbol{f}\,,\quad 0\leqslant k\leqslant n-1\,. \tag{13}$$

Recursive solution is impractical to extend to high order. An alternative approach (which may also be difficult, but is conceptually appealing) is to sum the series for $\boldsymbol{F}$ in equation (8) by defining the differential operator

$$\tilde{G}(\boldsymbol{y}_0,h)\equiv\boldsymbol{G}(\boldsymbol{y}_0,h)\cdot\frac{\partial}{\partial\boldsymbol{y}_0}. \tag{14}$$

Then

$$h\tilde{D} = \ln(1+h\tilde{G}) = h\tilde{G} - \frac{1}{2}(h\tilde{G})^2 + \frac{1}{3}(h\tilde{G})^3 - \frac{1}{4}(h\tilde{G})^4 + \cdots\,. \tag{15}$$

The logarithm of an operator is defined by its series expansion. Applying the operators to $\boldsymbol{y}_0$ gives $\boldsymbol{F}(\boldsymbol{y}_0,h) = \tilde{D}\boldsymbol{y}_0$ and $\tilde{G}\boldsymbol{y}_0 = \boldsymbol{G}(\boldsymbol{y}_0,h)$ so that

$$\begin{aligned}
\boldsymbol{F} &= \boldsymbol{G} - \frac{1}{2}h\tilde{G}\boldsymbol{G} + \frac{1}{3}h^2\tilde{G}^2\boldsymbol{G} - \frac{1}{4}h^3\tilde{G}^3\boldsymbol{G} + \frac{1}{5}h^4\tilde{G}^4\boldsymbol{G} - \cdots\\
&= (h\tilde{G})^{-1}\ln(1+h\tilde{G})\boldsymbol{G}.
\end{aligned} \tag{16}$$

The relation $\tilde{G}^{-1}\tilde{G} = 1$ defines $\tilde{G}^{-1}$.

3 A STUDY OF THE TRAPEZOIDAL RULE: A TIME-SYMMETRIC BUT NON-SYMPLECTIC INTEGRATOR

3.1 Relating the trapezoidal and midpoint rule

We introduce several one-step integrators. Let $\phi_h^T, \phi_h^M, \phi_h^E$, and ϕ_h^I indicate the trapezoidal, midpoint, explicit Euler, and implicit Euler one-step integration methods, respectively. The midpoint rule is symplectic, while the trapezoidal rule is not, but they have a close connection. The two integrators are defined by

$$\phi_h^T\boldsymbol{y}_0 = \boldsymbol{y}_1 = \boldsymbol{y}_0 + \frac{h}{2}\left[\boldsymbol{f}(\boldsymbol{y}_0) + \boldsymbol{f}(\boldsymbol{y}_1)\right], \tag{17}$$

and

$$\phi_h^M\boldsymbol{y}_0 = \boldsymbol{y}_1 = \boldsymbol{y}_0 + h\boldsymbol{f}\left(\frac{\boldsymbol{y}_0+\boldsymbol{y}_1}{2}\right). \tag{18}$$

The explicit and implicit Euler methods are first-order, not time-symmetric, and non-symplectic. They are

$$\begin{aligned}
&\phi_h^E\boldsymbol{y}_0 = \boldsymbol{y}_1 = \boldsymbol{y}_0 + h\boldsymbol{f}(\boldsymbol{y}_0), \qquad \text{and}\\
&\phi_h^I\boldsymbol{y}_0 = \boldsymbol{y}_1 = \boldsymbol{y}_0 + h\boldsymbol{f}(\boldsymbol{y}_1).
\end{aligned} \tag{19}$$

We see that

$$\phi_h^T = \phi_{h/2}^I\phi_{h/2}^E, \qquad \text{and} \qquad \phi_h^M = \phi_{h/2}^E\phi_{h/2}^I, \tag{20}$$

so that $\phi_h^T = (\phi_{h/2}^E)^{-1}\phi_h^M\phi_{h/2}^E$. Thus, the trapezoidal and midpoint rules are said to be *conjugate* (Hairer et al. 2006, Section VI.8) to each other. To get a trapezoidal orbit, we need only apply a correction at the beginning and ending of a midpoint rule integration. This means the trapezoidal rule solution should have similar error properties to a symplectic method like the midpoint rule; we will show this more carefully in Section 3.3.

3.2 Runge–Kutta methods

The numerical algorithm (2) is a mapping of the vector space $\{\boldsymbol{y}\}$ onto itself. A broad class of integrators, that encompasses various common algorithms including the ones of Section 3.1, defines the mapping $\boldsymbol{y}_0 \to \boldsymbol{y}_1 = \boldsymbol{y}_0 + h\boldsymbol{G}(\boldsymbol{y}_0, h)$ using only $\boldsymbol{f}(\boldsymbol{y}_0)$ and derivative operators that are scalars under coordinate transformations of $\boldsymbol{y}$. They are called Runge–Kutta (RK) methods. An RK method of s stages is defined by constants a_{ij} and b_i for $1 \leqslant i, j \leqslant s$ when there is no explicit time-dependence in the governing ODE's:

$$\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{y}_0 + h\sum_{i=1}^{s} b_i \boldsymbol{k}_i, \qquad \text{and} \\ \boldsymbol{k}_i &= \boldsymbol{f}(\boldsymbol{y}_0 + h\sum_{j=1}^{s} a_{ij}\boldsymbol{k}_j), \end{aligned} \tag{21}$$

which is explicit, and thus less computationally expensive, if and only if $a_{ij} = 0$ for $j \geqslant i$ (a strictly triangular matrix). For this method, $\boldsymbol{G}(\boldsymbol{y}, h)$ depends on $\boldsymbol{f}$ and on differential operators like $D_0 \equiv \boldsymbol{f}(\boldsymbol{y}) \cdot (\partial/\partial \boldsymbol{y})$ that are scalars under general coordinate transformations $\boldsymbol{y} \to \boldsymbol{y}'$. The popular leapfrog method is not an RK method—it is known as a partitioned Runge–Kutta method— because it uses a different rule for updating the positions and momenta. In the partitioned RK case, the differential operators defining the $\boldsymbol{g}_k$ are no longer covariant under general linear transformations of the full space.

The methods (17) and (18) are RK methods. We can check that for the former, $s = 2$, $b_1 = b_2 = 1/2$, $a_{21} = a_{22} = 1/2$, and $a_{11} = a_{12} = 0$. For the latter, $s = 1$, $a_{11} = 1/2$, and $b_1 = 1/2$. Both are implicit and thus will need to be solved through iteration, whether fixed-point or Newton-Rhapson (Press et al. 2002).

It is easy to see both the implicit midpoint (as opposed to explicit midpoint, a different RK method) and trapezoidal rule are time-symmetric (and reversible, cf. Section 2.1) **if** used with fixed time step. Take a step forwards from $\boldsymbol{y}_0$ to obtain $\boldsymbol{y}_1$ and a step backwards to obtain $\boldsymbol{y}'$. The rules require $\boldsymbol{y}' = \boldsymbol{y}_0$.

Next, we investigate whether the methods are symplectic. It has been shown (Hairer et al. 2006, Section VI.4), that if and only if an RK method conserves quadratic invariants in the phase space variables of the underlying differential equations, it is symplectic. The reason is related to the fact that the functions of $\boldsymbol{y}$, $\boldsymbol{SJS}^\dagger$, defined by equations (A2), which must be invariant for symplecticity to hold, are first integrals of the variational equations. One such typical quadratic invariant in some problems is the angular momentum. Any quadratic invariant can be written $Q(\boldsymbol{y}) = \boldsymbol{y}^\dagger \boldsymbol{C}\boldsymbol{y}$, with $\boldsymbol{C}$ a symmetric matrix. Write the implicit midpoint rule as

$$\boldsymbol{y}_1 - \boldsymbol{y}_0 = h\boldsymbol{f}\left(\frac{\boldsymbol{y}_1 + \boldsymbol{y}_0}{2}\right). \tag{22}$$

Multiply from the left by $(\boldsymbol{y}_1 + \boldsymbol{y}_0)^\dagger \boldsymbol{C}$. The left hand side gives

$$\boldsymbol{y}_1^\dagger \boldsymbol{C}\boldsymbol{y}_1 - \boldsymbol{y}_1^\dagger \boldsymbol{C}\boldsymbol{y}_0 + \boldsymbol{y}_0^\dagger \boldsymbol{C}\boldsymbol{y}_1 - \boldsymbol{y}_0^\dagger \boldsymbol{C}\boldsymbol{y}_0 = \boldsymbol{y}_1^\dagger \boldsymbol{C}\boldsymbol{y}_1 - \boldsymbol{y}_0^\dagger \boldsymbol{C}\boldsymbol{y}_0, \tag{23}$$

which follows from the fact that the transpose of a scalar is the scalar. The right hand is zero because $\dot{Q}((\boldsymbol{y}_1 + \boldsymbol{y}_0)/2) = 0$. Thus, we are left with $\boldsymbol{y}_1^\dagger \boldsymbol{C}\boldsymbol{y}_1 - \boldsymbol{y}_0^\dagger \boldsymbol{C}\boldsymbol{y}_0 = 0$, which means the implicit midpoint rule conserves quadratic invariants and is thus symplectic. Any numerical experiment with a symplectic integrator that is an RK method will show conservation of all quadratic invariants; an example is the angular momentum, for differential equations that have this symmetry, such as the Kepler problem. We show a more direct proof of the symplecticity of the midpoint rule in Appendix A.

Write the trapezoidal rule as

$$\boldsymbol{y}_1 - \boldsymbol{y}_0 = \frac{h}{2}\left[\boldsymbol{f}(\boldsymbol{y}_0) + \boldsymbol{f}(\boldsymbol{y}_1)\right]. \tag{24}$$

If we multiply on the left by $(\boldsymbol{y}_1 + \boldsymbol{y}_0)^\dagger \boldsymbol{C}$, we find that $\boldsymbol{y}_1^\dagger \boldsymbol{C}\boldsymbol{y}_1 - \boldsymbol{y}_0^\dagger \boldsymbol{C}\boldsymbol{y}_0 \neq 0$ and is generally not conserved, meaning quadratic invariants are not conserved, and the trapezoidal rule is not symplectic. Numerical experiments indeed show the trapezoidal rule does not conserve quadratic invariants such as the angular momentum.

We will largely focus on the trapezoidal rule for the remainder of the paper, because we are interested in a time-symmetric, but non-symplectic integrator, and this method is one of the simplest examples of this. Some researchers have used leapfrog, which is symplectic when using fixed time step, with reversible steps. Once the steps are adapted, however, the symplectic property is lost, so there is no advantage from this standpoint to use leapfrog. On the other hand, even when used with adaptive steps, leapfrog conserves angular momentum exactly, while the trapezoidal rule does not. However, the tests we consider in what follows have no angular momentum invariant. Also, the trapezoidal rule has a related invariant for every quadratic, in the phase space, invariant in the underlying equations; see Section 3.3. Both leapfrog and trapezoidal rule conserve linear invariants, such as the total linear momentum, exactly (all RK methods do). A disadvantage of the trapezoidal rule is that it requires solving implicit equations, unlike leapfrog. But any time-symmetric Runge–Kutta method is implicit. We focus our efforts on Runge–Kutta methods because their properties have been well established, and they treat all phase space components with the same functional update rule, which will simplify our analysis of their symplecticity and energy conservation properties in Appendices C and D. An advantage of the trapezoidal rule, as shown in Section 3.1, is its connection to a symplectic method.

3.3 A conserved quantity for the trapezoidal rule

Consider the broad class of separable Hamiltonians,

$$H_0 = \sum_{i=1}^{n} \frac{p_i^2}{2m_i} + U(\boldsymbol{q}), \tag{25}$$

where $2n$ is the phase space dimension, and define

$$U_i \equiv \frac{\partial U}{\partial q_i}. \tag{26}$$

Additional derivatives of U with respect to the q_i are denoted by more U subscript indices. In Section 3, let $\boldsymbol{y}^{\mathrm{T}} = (\boldsymbol{q}^{\mathrm{T}}, \boldsymbol{p}^{\mathrm{T}})$ and $\boldsymbol{y}^{\mathrm{M}} = (\boldsymbol{q}^{\mathrm{M}}, \boldsymbol{p}^{\mathrm{M}})$ refer to $\boldsymbol{y}_1$ from the trapezoidal and implicit midpoint rule, respectively. Let other functions be functions of $\boldsymbol{y}_0$. As an example, $q_i^T = q_i + hp_i - h^2/2U_i + O(h^3)$

In Appendix B, we derive the modified differential equations for the trapezoidal and implicit midpoint rule, and the Hamiltonian for the implicit midpoint rule. Using the results from Appendix B, we find

$$\begin{aligned} q_i^{\mathrm{T}} &= q_i^{\mathrm{M}} + O(h^4), \\ p_i^{\mathrm{T}} &= p_i^{\mathrm{M}} - \frac{h^3}{8}\sum_{j,k=1}^{n} \frac{p_j p_k}{m_j m_k} U_{ijk} + O(h^4), \\ \dot{q}_i^{\mathrm{T}} &= \frac{\partial \tilde{H}}{\partial p_i} + O(h^4), \\ \dot{p}_i^{\mathrm{T}} &= -\frac{\partial \tilde{H}}{\partial q_i} - \frac{h^2}{8}\sum_{j,k=1}^{n} \frac{p_j p_k}{m_j m_k} U_{ijk} + O(h^4), \end{aligned} \tag{27}$$

Table 1. The midpoint and trapezoidal rules, and KDK and DKD leapfrogs, have a conserved energy to second order described by (31). They only differ in the values of the coefficients of a and b, whose absolute value is either $1/12$ or $1/24$, and we list them here.

Method	a	b
Midpoint	$-\frac{1}{24}$	$-\frac{1}{24}$
Trapezoidal	$+\frac{1}{12}$	$+\frac{1}{12}$
KDK Leapfrog	$-\frac{1}{24}$	$+\frac{1}{12}$
DKD Leapfrog	$+\frac{1}{12}$	$-\frac{1}{24}$

where $\tilde{H}$ is the midpoint Hamiltonian given by (B11). Note we do not follow Einstein summation convention, but we could restore the convention, for example, by substituting $\partial H/\partial p_i$ for p_i/m_i. Using this information, we can compute that along the trapezoidal trajectory,

$$\frac{d}{dt}\tilde{H} = \sum_{i=1}^{n}\left(\dot{p}_i^{\mathrm{T}}\frac{\partial\tilde{H}}{\partial p_i} + \dot{q}_i^{\mathrm{T}}\frac{\partial\tilde{H}}{\partial q_i}\right) = -\frac{h^2}{8}\sum_{i,j,k=1}^{n}\frac{p_i p_j p_k}{m_i m_j m_k}U_{ijk} + O(h^4). \quad (28)$$

This equation describes the energy drift of trapezoidal rule. The $O(h^2)$ term can be integrated with respect to time, and we find that

$$\frac{d}{dt}\tilde{E}_2 = O(h^4), \quad (29)$$

where

$$\tilde{E}_2 = H_0 + \frac{h^2}{12}\left(\sum_{i,j=1}^{n}U_{ij}\frac{p_i p_j}{m_i m_j} + \sum_{i=1}^{n}\frac{1}{m_i}U_i^2\right): \quad (30)$$

the trapezoidal rule has a conserved energy at least to second order. We will check this numerically in Section 4.1. This means a time-symmetric, non-symplectic method can also have a conserved energy at some order, but this fact may not in of itself be useful. The trapezoidal and midpoint rule, and DKD and KDK leapfrog all have a conserved energy to second order, which, for Hamiltonian (25) has form

$$\tilde{E}_2 = H_0 + h^2\left(\sum_{i,j=1}^{n}aU_{ij}\frac{p_i p_j}{m_i m_j} + \sum_{i=1}^{n}b\frac{1}{m_i}U_i^2\right), \quad (31)$$

and their coefficients a and b are shown in Table 1. a and b differ from each other for the leapfrog methods because they are partitioned RK methods, as mentioned in Section 3.2.

We can do better and show that trapezoidal rule has a conserved energy to at least fourth order. Substituting its MDE (B8) into equations (C9), reveals,

$$\tilde{E} = H + \frac{h^2}{12}(\hat{D}_{21}H) - \frac{h^4}{720}(3\hat{D}_{40} + 6\hat{D}_{41} - \hat{D}_{43})H + O(h^6)\,. \quad (32)$$

For a conventional Hamiltonian (25), this becomes

$$\begin{aligned}\tilde{E} = H &+ \frac{h^2}{12}\left(\sum_{i=1}^{n}\frac{1}{m_i}U_i^2 + \sum_{i,j=1}^{n}U_{ij}\frac{p_i p_j}{m_i m_j}\right)\\ &- \frac{h^4}{240}\left[\sum_{i,j=1}^{n}\frac{U_i U_j U_{ij}}{m_i m_j} - \sum_{i,j,k=1}^{n}\left(\frac{p_i p_k U_{jk}U_{ij}}{m_i m_j m_k} + 2\frac{p_i p_j U_k U_{ijk}}{m_i m_j m_k}\right)\right.\\ &\left. + \frac{1}{3}\sum_{i,j,k,l=1}^{n}\frac{p_i p_j p_k p_l U_{ijkl}}{m_i m_j m_k m_l}\right]\\ &+ O(h^6).\end{aligned} \quad (33)$$

In fact, we are able to show that the trapezoidal rule conserves an energy function to all orders in h, and we can write it down. Rewrite the trapezoidal rule as a sequence of three RK steps:

$$\begin{aligned}\boldsymbol{y}_{-1/2} &= \boldsymbol{y}_0 - \frac{1}{2}h\boldsymbol{f}(\boldsymbol{y}_0)\\ \boldsymbol{y}_{1/2} &= \boldsymbol{y}_{-1/2} + h\boldsymbol{f}\left(\frac{1}{2}\boldsymbol{y}_{-1/2} + \frac{1}{2}\boldsymbol{y}_{1/2}\right) = \boldsymbol{y}_{-1/2} + h\boldsymbol{f}(\boldsymbol{y}_0)\\ \boldsymbol{y}_1 &= \boldsymbol{y}_{1/2} + \frac{1}{2}h\boldsymbol{f}(\boldsymbol{y}_1) = \boldsymbol{y}_0 + \frac{1}{2}h[\boldsymbol{f}(\boldsymbol{y}_0) + \boldsymbol{f}(\boldsymbol{y}_1)]\,.\end{aligned} \quad (34)$$

The first step is a backwards explicit Euler step, the second is a symplectic midpoint method, and the third is an implicit Euler step. Because the implicit midpoint rule has a conserved Hamiltonian (assuming convergence of the series), it is natural to assume that the trapezoidal rule respects an energy function with the same functional form, but with shifted initial conditions. Indeed, let

$$E_{\mathrm{trap}}(\boldsymbol{y}) = \tilde{H}_{\mathrm{midpoint}}\left[\boldsymbol{y} - \frac{1}{2}h\boldsymbol{f}(\boldsymbol{y})\right]. \quad (35)$$

Then, we can check $E_{\mathrm{trap}}(\boldsymbol{y}_0) = E_{\mathrm{trap}}(\boldsymbol{y}_1)$, which implies that the trapezoidal rule conserves the energy function $E_{\mathrm{trap}}(\boldsymbol{y})$. If the underlying equations have a quadratic invariant Q, we also see the trapezoidal rule has a related invariant,

$$Q_{\mathrm{trap}}(\boldsymbol{y}) = Q\left[\boldsymbol{y} - \frac{1}{2}h\boldsymbol{f}(\boldsymbol{y})\right]. \quad (36)$$

To fourth order, (35) agrees with equation (32), but it is exact to all orders. We will derive in Appendix C that there exist time-symmetric methods which are not energy conserving to all orders. These results are summarized in Table E1. This means we can find energy drift with a symmetric integrator with fixed time step— this result has already been discussed by Faou et al. (2004) and others.

3.3.1 *An example: the simple harmonic oscillator*

We derive the conserved energy of the trapezoidal rule for the simple harmonic oscillator (SHO). The Hamiltonian for the SHO is

$$H(q,p) = \frac{1}{2}\left(q^2 + p^2\right). \quad (37)$$

For this Hamiltonian, the trapezoidal rule becomes explicit, since the coordinate derivatives are linear in coordinates. Also, in this case, the implicit midpoint rule gives an identical rule. The rules say

$$\begin{aligned}q_1 &= q_0 + \frac{h}{2}(p_0 + p_1) \qquad \text{and}\\ p_1 &= p_0 - \frac{h}{2}(q_0 + q_1)\,.\end{aligned} \quad (38)$$

When solved for q_1 and p_1, they say

$$\begin{aligned} q_1 &= aq_0 + bp_0 \qquad \text{and} \\ p_1 &= ap_0 - bq_0, \end{aligned} \tag{39}$$

where

$$a = \left(\frac{1-\delta}{1+\delta}\right), \qquad b = \frac{h}{(1+\delta)}, \qquad \text{and} \qquad \delta = \frac{h^2}{4}. \tag{40}$$

(39) is also the exact trajectory after time h for a Hamiltonian

$$\tilde{H} = AH, \tag{41}$$

so long as

$$\begin{aligned} \cos(Ah) &= a, \qquad \text{and} \\ \sin(Ah) &= b, \end{aligned} \tag{42}$$

implying

$$A = \frac{1}{h}\tan^{-1}\left(\frac{h}{1-\delta}\right). \tag{43}$$

$0 < A < 1$ for $0 < h < 2$, so the numerical value of the modified Hamiltonian is smaller than H. When $h \geqslant 2$, an A satisfying (42) does not exist, so the governing equations are no longer Hamiltonian. Thus, the trapezoidal and implicit midpoint rules' MDEs are governed by (41). This implies they exactly conserve the energy of the SHO, as one can verify numerically.

For a general Hamiltonian (e.g. the Hénon–Heiles problem), these simple exact results no longer hold. However, for a time-independent Hamiltonian, symplectic methods always have a conserved energy, and so do conjugate methods like the trapezoidal rule.

3.4 Modified differential equation with adaptive time steps

In previous sections and the Appendix, we discuss integrators with fixed step-sizes, but for fixed step-sizes, there already exist excellent symplectic integrators in astrophysics, starting with leapfrog. Time-symmetric integrators are popular in astrophysics due to their ability to accommodate adaptive stepping. An exactly time-symmetric integrator was proposed by Hut et al. (1995), and approximately time-symmetric integrators have been developed by Pelupessy et al. (2012) and Kokubo et al. (1998). We focus on the proposal by Hut et al. (1995), because it is exactly time-symmetric, under certain conditions we describe. In conjunction with leapfrog, they propose to write the time step as an implicit equation,

$$h = \frac{\epsilon}{2}\left[\sigma(\boldsymbol{y}_0) + \sigma(\boldsymbol{y}_1)\right]. \tag{44}$$

$\sigma(\boldsymbol{y})$ is a function that we can specify using a priori knowledge about the solution trajectory (e.g., the relevant timescales) or even without this knowledge (Stoffer 1995). An implicit step criterion can be used with an implicit one-step method, like the trapezoidal rule, not necessarily resulting in more iterations when solving the update equations. (44) can be written as an explicit infinite series in ϵ,

$$h = \epsilon s(\boldsymbol{y}_0, \epsilon) = \epsilon s_0(\boldsymbol{y}_0) + \epsilon^2 s_1(\boldsymbol{y}_0) + \ldots \tag{45}$$

The direction of time is now determined by the sign of ϵ. Of course, the s_i depend on the method. Letting $\sigma \equiv \sigma(\boldsymbol{y}_0)$, for trapezoidal rule,

$$\begin{aligned} s_0 &= \sigma, \\ s_1 &= \frac{1}{2}\sum_{i=1}^{2n} \sigma f_i \partial_i \sigma, \\ s_2 &= \frac{1}{4}\left[\sigma \sum_{i=1}^{2n} (f_i \partial_i \sigma)^2 + \sigma^2 \sum_{i,j=1}^{2n} \left((f_j \partial_j f_i)\partial_i \sigma + f_j f_i \partial_j \partial_i \sigma\right)\right], \\ &\vdots \end{aligned} \tag{46}$$

where $2n$ is the phase space dimension. f_i is the ith component of $\boldsymbol{f}$ in eq. (1) and $\partial_i \equiv \partial/\partial y_i$. Symmetry requires

$$s(\boldsymbol{y}_1, -\epsilon) = s(\boldsymbol{y}_0, \epsilon). \tag{47}$$

For criterion (44), this requirement is automatically satisfied. ρ-reversibility would require $s(\boldsymbol{\rho y}_1, \epsilon) = s(\boldsymbol{y}_0, \epsilon)$. For steps (44) this means (Hairer et al. 2006, Section VIII.3)

$$\sigma(\boldsymbol{\rho y}) = \sigma(\boldsymbol{y}). \tag{48}$$

This condition is easy to satisfy, but it is not always satisfied. Hut et al. (1995) were interested in the N-body problem and proposed a σ that is the minimum of the close encounter and free fall times. This satisfies (48) if we are taking the absolute values of relative velocities. We will explore what happens when eq. (48) is not obeyed in Section 4. The equations (44), (47), and (48) apply whether the underlying method is an RK method, like trapezoidal rule, or a partitioned Runge–Kutta method, like leapfrog. But the underlying method must be time-symmetric for either (47) or (48) to hold.

The stepping rule (44) is implicit, which is more cumbersome to analyze than an explicit rule. However, we use an implicit criterion for the following reasons:

- The trapezoidal rule is already implicit, so choice (44) does not necessarily add more computational work.
- (44) has already been used by Hut et al. (1995) and others.
- Dehnen (2017) studies explicit stepping criteria with step sizes that can only take certain values. The discreteness of the step sizes breaks the time-symmetry and reversibility symmetries. We want to construct exactly symmetric and reversible methods for our tests to be able to conclude that errors are not due to breaks in these symmetries.
- The explicit stepping criteria in (Dehnen 2017, Sections 4 and 5), even in the continuous, non-discrete case, risk becoming unsynchronized with σ, leading to stepping of questionable efficiency and accuracy. This stepping can be regarded as a multistep method.

We can construct the MDE with adaptive time steps, following the procedure of Section 2.2 and using the form (45), so that the

series are now written in terms of ϵ. Now, instead of eq. (27), we have

$$\begin{aligned}
q_i^{\rm T} &= q_i^{\rm M} - \frac{1}{12}(\epsilon s_0)^3 \sum_{j=1}^{n} \frac{p_j}{m_i m_j} U_{ij} + O(\epsilon^4),\\
p_i^{\rm T} &= p_i^{\rm M} - \frac{1}{12}(\epsilon s_0)^3 \left(\sum_{j,k=1}^{n} \frac{p_j p_k}{m_j m_k} U_{ijk} - \sum_{j=1}^{n} U_{ij} U_j \right) + O(\epsilon^4),\\
\dot{q}_i^{\rm T} &= \frac{\partial H_0}{\partial p_i} - \frac{1}{12}(\epsilon s_0)^2 \sum_{j=1}^{n} \frac{p_j}{m_i m_j} U_{ij} + O(\epsilon^3),\\
\dot{p}_i^{\rm T} &= -\frac{\partial H_0}{\partial q_i} - \frac{1}{12}(\epsilon s_0)^2 \left(\sum_{j,k=1}^{n} \frac{p_j p_k}{m_j m_k} U_{ijk} - \sum_{j=1}^{n} U_{ij} U_j \right) + O(\epsilon^3).
\end{aligned} \tag{49}$$

As in eq. (28), we can calculate the energy drift along the trapezoidal orbit as

$$\frac{d}{dt} H_0 = -\frac{1}{12}\epsilon^2 [\sigma(\boldsymbol{q},\boldsymbol{p})]^2 \sum_{i,j,k=1}^{n} U_{ijk} \frac{p_i p_j p_k}{m_i m_j m_k} + O(\epsilon^3). \tag{50}$$

If we let $\sigma = 1$ and $\epsilon = h$, this expression is just (28). This shows that the energy drift is a function of the problem and the choice of σ. In general, this ϵ^2 term cannot be integrated in terms of elementary functions. There exist reversible σ (cf. eq. 48) which lead to secular drift and irreversible σ which lead to no energy drift, as we will show in Section 4. (50) holds for any $h(\epsilon, \boldsymbol{y}_0)$, not just (44), so long as to lowest order in ϵ, $h = \epsilon\sigma(\boldsymbol{y}_0)$. For example, (50) applies to the geometric mean time step,

$$h = \epsilon \left[\sigma(\boldsymbol{y}_0)\sigma(\boldsymbol{y}_1)\right]^{1/2}. \tag{51}$$

4 NUMERICAL DEMONSTRATION

In this section we apply the error analysis of Section 3 to see that energy conservation is violated in a number of situations, even when a method is symmetric and reversible. For the tests, we consider the pendulum solution and Hénon–Heiles orbits.

4.1 The modified pendulum

In the following, we consider the trapezoidal rule (17) along with the adaptive step criteria (44). Consider the simple pendulum, with a modified potential:

$$H = \frac{p^2}{2} - \cos q + \frac{1}{5}\sin(2q). \tag{52}$$

This modified pendulum was considered by Faou et al. (2004). The reason for choosing this potential with $1/5\sin(2q)$ will become apparent below. This Hamiltonian is ρ-reversible: $\frac{\partial H(q,-p)}{\partial p} = -\frac{\partial H(q,p)}{\partial p}$, and $\frac{\partial H(q,-p)}{\partial q} = +\frac{\partial H(q,p)}{\partial q}$.

The potential is not symmetric in q over the periodic range of q, as seen in Fig. 1. The minimum is $U \approx -1.069$ and occurs at $q \approx 5.959$. First, we choose $\sigma(\boldsymbol{y}) = 1$, so that the step is constant. Because this σ satisfies (48), the integrator is reversible. We choose $h = \epsilon = 2\pi/100 \approx 0.63$: there are roughly 100 steps per period. We let $t_{\rm final} = 100$. As initial conditions, we choose $p = 2.5$ and $q = 0$ so that $H = 2.125$. So for the exact solution (and in all our numerical solutions), the sign of the momentum does not change (the orbit is circulating). In Fig. 2, we show that the change in various phase space quantities in time mimics the behavior of a symplectic integrator. $\dot{H}_{02}$ is given by the second order ϵ^2 term of eq. (50):

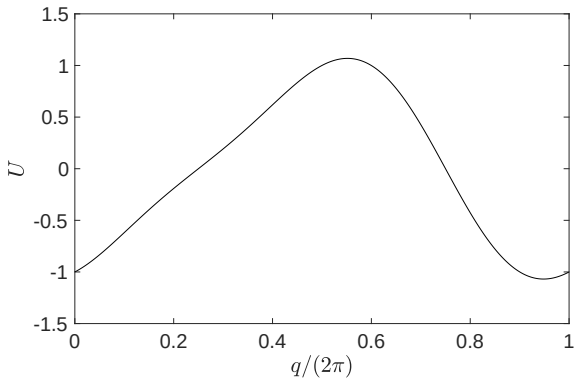

Figure 1. Potential as a function of the periodic range of q for the modified pendulum Hamiltonian (52). There is no symmetry in the potential and it has a minimum of $U \approx -1.069$.

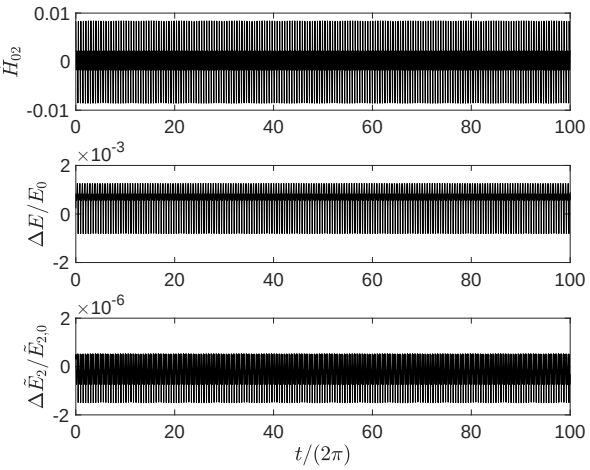

Figure 2. The evolution of some phase space quantities as a function of time when we integrate the modified pendulum with the symmetric non-symplectic trapezoidal rule. We use a constant $h \approx 0.63$ and initial conditions $p = 2.5$ and $q = 0$. The top panel gives the change in energy as given by (50). The middle panel gives the energy error, and the bottom panel gives the error of the conserved second order energy. No energy drift is observed in the middle panel, in agreement with the top panel. The second order energy is conserved better than the energy, as expected.

$$\dot{H}_{02} = -\frac{1}{12}\epsilon^2 [\sigma(\boldsymbol{q},\boldsymbol{p})]^2 \sum_{i,j,k=1}^{n} U_{ijk} \frac{p_i p_j p_k}{m_i m_j m_k}, \tag{53}$$

which oscillates symmetrically around 0. We also show the energy and $\tilde{E}_2$ error, from (31). E_0 and $\tilde{E}_{2,0}$ are the initial energy and $\tilde{E}_2$ values, respectively. $\tilde{E}_2$ is conserved better than E, supporting the finding that a conserved energy exists. We checked that, for a fixed integration time, $\Delta E/E \propto h^2$ and $\Delta\tilde{E}_2/\tilde{E}_2 \propto h^4$.

We next choose an adaptive step strategy. We choose $\sigma(\boldsymbol{y}) = U + 1.5$, so that $\sigma(\boldsymbol{y}) > 0$. This choice is both reversible and time-symmetric, according to the discussion above (48). We checked if we integrate forwards, change the sign of p, and integrate the same number of steps backwards, we recover the initial conditions up to roundoff error. We choose $\epsilon = 2\pi/(100 \times 1.63)$ so that the average time step is approximately still the same as the previous test. The initial phase space coordinates remain the same in this test: $p = 2.5$ and $q = 0$. We initialize the integration with guess for the initial

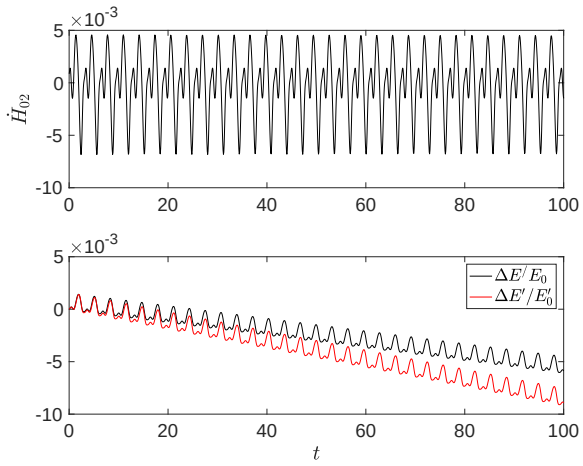

Figure 3. Same as Figure 2, but now with a reversible, time-symmetric step size strategy, $\sigma(\boldsymbol{y}) = U + 1.5$. $\epsilon \approx 0.039$. We observe a linear energy drift, which roughly agrees with the prediction from the top panel, given by the red curve in the lower panel. This integration is fully symmetric and reversible, yet shows a linear energy error drift.

.

Table 2. Description of symplectic integrators used in tests of the modified pendulum Hamiltonian.

Method	Symmetric?	Order	Classification
Leapfrog DKD and KDK	Yes	2	Partitioned Runge–Kutta
Gauss-Legendre	Yes	4	Runge–Kutta
Symplectic Euler	No	1	Runge–Kutta

step, $h_0 = \epsilon\sigma(\boldsymbol{y}_0)$, and thereafter we use the previous time step as the initial guess. We integrate for the same total time as Fig. 2, and show the results in Fig. 3. $\dot{H}_{02}$ now is not symmetric around 0. This leads to a linear drift in energy error as seen on the bottom subplot. If $\dot{H}_{02,i}$ is the value of $\dot{H}_0$ at time step number i, and h_i is the value of the time step, define

$$E'_m = E_0 + \sum_{i=1}^{m} h_i \dot{H}_{02,i}, \tag{54}$$

which we expect to be close to E_m, the energy at step m. We see in Fig. 3 this is the case at small time, but the approximation breaks down for larger time. As we decrease ϵ, the difference between the two curves becomes undetectable on the same type of plot. We checked the slope of the $\Delta E/E$ curve scales as ϵ^2, confirming the energy error is $O(t\epsilon^2)$. All other $\sigma(\boldsymbol{y})$ we tested, reversible and irreversible, gave a linear drift in energy for this problem. For the geometric mean time step (51), the errors are similar, as expected. The final energy error at $t = 100\times 2\pi$ changes by less than 1%.

We integrate (52) with one-step symplectic methods of different properties. We use stepsize $h = 2\pi/100$ and initial conditions $p = 2.5$ and $q = 0$ (as in Section 4.1). The methods are described in Table 2, where we write the method's name, order, Runge–Kutta classification, and whether the method is time-symmetric. None of the methods yield energy drift.

By contrast, it has been reported that Lobatto IIIA and IIIB, two fourth order symmetric and reversible, but non-symplectic methods show energy drift when used with fixed step size. The tests were performed in Faou et al. (2004), using $h = 0.16$ and $t_{\max} = 1600$. Faou et al. (2004) differs from our work in that it considered only fixed time-steps. We confirmed the leading order error for Lobatto IIIB is $O(th^4)$ while the leading order error for Lobatto IIIA is $O(h^4)$ and has no error drift (at leading order), except for roundoff error contributions. Lobatto IIIA and IIIB have the same symmetries as the trapezoidal rule with adaptive symmetric and reversible steps: time-symmetry and time-reversibility.

The simplified Takahashi–Imada method is symmetric and volume preserving. This means it conserves one Poincaré invariant, which does not guarantee symplecticity. It was reported in Hairer et al. (2009) that this method gives secular drift in a function close to the energy for Hamiltonian (52). They used $h = 0.2$ and $t_{\max} \approx 1900$. In our own tests, we found the method gives energy drift. Volume preservation is equivalent to symplecticity for one-degree-of-freedom problems such as Hamiltonian (52), so this would appear to be an example of a symplectic integrator giving energy drift. But this integrator is generally non-symplectic.

The only example we found of a reversible, symmetric, and non-symplectic method conserving energy for this problem is studied in Fig. 2. We have not tested symplectic integrators with adaptive steps because symplecticity is not conserved and this advantage of the method is lost.

When the orbits of these initial conditions are computed with symplectic integrators, we have observed that different steps along the orbit produce increases or decreases in energy. The net increase is zero. For the orbit of Fig. 3, the net increase is negative over an orbit, leading to energy drift.

For the unmodified pendulum,

$$H = \frac{p^2}{2} - \cos(q), \tag{55}$$

consider the same initial conditions given above. The sign of p is still invariant in all tests. Both reversible and irreversible σ, such as $\sigma(\boldsymbol{y}) = 1.5 - \cos(q)$ and $\sigma(\boldsymbol{y}) = ap + b$ with a and b constants, give no drift in energy error. However, we again get drift if we let $\sigma(\boldsymbol{y})$ be an asymmetric function of q, such as the $\sigma(\boldsymbol{y}) = 1.5 - \cos(q) + 1/5\sin(2q)$ we used above, even though this σ is reversible. These experiments show that time reversibility and energy conservation are independent concepts. To summarize, for circulating pendulum orbits, all time-symmetric methods, except the fixed time trapezoidal rule, gave undesirable error behavior. This exception may be related to the fact it is a conjugate symplectic method (see Section 3.1). All tested symplectic methods except the simplified Takahashi-Imada method yielded desirable energy conservation. The simplified Takahashi-Imada method is generally non-symplectic.

4.2 Hénon–Heiles orbits

The Hénon–Heiles problem Hamiltonian (Henon & Heiles 1964) is a two-degree-of-freedom problem—a simplified model of a galactic potential. The Hamiltonian is

$$H = \frac{1}{2}(p_x^2 + p_y^2) + U(x,y) \tag{56}$$

with $U(x,y) = \frac{1}{2}(x^2+y^2) + x^2 y - \frac{y^3}{3}$. This Hamiltonian allows both chaotic and regular trajectories and is ρ-reversible. We consider a regular orbit with initial conditions $x = 0$, $y = 0.2$, $p_y = 0.3$, $p_x = 0.125413095187199$, and $H = 0.07019755555555$ (although

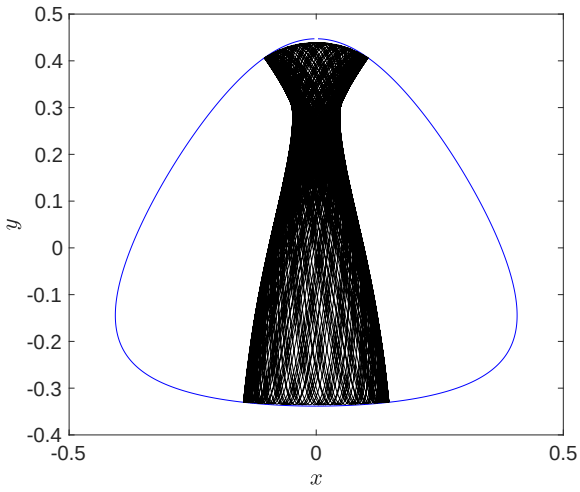

Figure 4. A regular box Hénon–Heiles orbit. The initial conditions are $x = 0$, $y = 0.2$, $p_y = 0.3$, and $H = 0.07019755555555$. We also plot the bounding equipotential. The integration is run until $t_{\max} = 100 \times 2\pi$ with a constant step-size trapezoidal method.

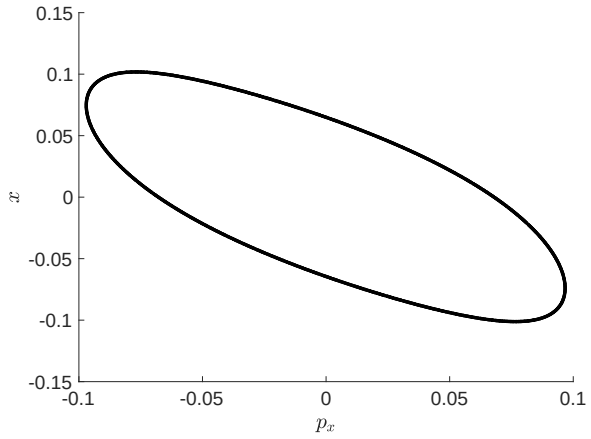

Figure 5. Surface of section plot for the orbit of Fig. 4. A point is plotted everytime $y = 0$ is crossed with $p_y > 0$. The symmetry in p_x indicates a box orbit, and the closed curve indicates a regular orbit far from resonance. The surface of section is computed up to time $t = 10^5$ with a high accuracy Runge–Kutta method.

.

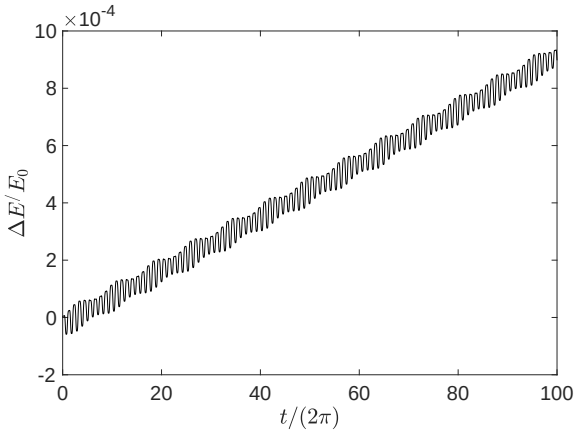

Figure 6. Energy error over time for the orbit of Fig's. 5 and 4. We run a trapezoidal rule with step selector $\sigma = 10^{-3} p_y + 0.01$ and $\epsilon = 2.5$. This integrator is time-symmetric, but not reversible, and shows a linear drift in energy. The energy drift is predicted accurately by (50).

we only need to keep three significant figures to get the same qualitative results). Using $\sigma = 1$, and $\epsilon = h = 0.1$, we show the trajectory in Fig. 4 with $t_{\max} = 100 \times 2\pi$. Also plotted is the bounding equipotential curve, $U = H$. The trajectory does not span the entire allowed area, which tells us it is a regular orbit. We can verify this in a surface of section plot. In Fig. 5, we plot a point in the x–p_x plane every time $y = 0$ is crossed with $p_y > 0$, up to time $t = 10^5$. We use a fifth and sixth order pair of Runge–Kutta methods, in what's known as Verner's embedded Runge–Kutta method (Verner 1978), for this plot. This is an adaptive step method: two methods allow an estimate of the local truncation error which is then used to determine a step size. The final energy error is $\approx 3.2 \times 10^{-14}$. This orbit is a box orbit: the sign (and magnitude) of the angular momentum oscillates. If we change the sign of the initial momenta, the trajectory is confined to the same bounding curve, which means the second isolating integral besides the energy (for these initial conditions) does not depend on the sign of the momenta. This is a consequence of the ρ-reversibility of the equations due to (56). We checked this by running the trajectory with a sign change in the initial momenta and checking that the minimum and maximum x of the trajectory is the same to 15 significant figures.

For this Hamiltonian, for eq. (50), we have

$$\frac{d}{dt} H_0 = -\frac{1}{12} \epsilon^2 \sigma^2(\mathbf{y}) \left[2 p_y (3 p_x^2 - 2 p_y^2) \right]. \tag{57}$$

Note the asymmetry in p_y and p_x due to the potential. For the orbit of Figure 4, $\overline{y} \approx 0.07$. So the centroid is non-zero for this orbit. We choose $\sigma(\mathbf{y}) = a p_y + b$, where $a = 10^{-3}$ and $b = 10^{-2}$, to ensure $\sigma(\mathbf{y}) > 0$. This σ is irreversible (below we will explore other σ, reversible and irreversible). We let $\epsilon = 2.5$ and $t_{\max} = 628$. We can estimate the typical time step by using

$$\overline{h} \approx \epsilon (0.01 + 0.001 \overline{p}_y). \tag{58}$$

We measure experimentally a time weighted average of p_y of $\overline{p_y} \approx -5 \times 10^{-4}$, which gives $\overline{h} \approx 0.025$, in agreement with experiment. In Fig. 6, we show the error in energy over time—it has a linear drift with slope about 1.4×10^{-6}. We checked, by varying ϵ, that the slope scales with ϵ^2. If we plot the error in E' on the same plot, it is nearly indistinguishable from the error in E. If $\delta = \Delta E/E - \Delta \tilde{E}/\tilde{E}$, $\overline{\delta} = 4.3 \times 10^{-7}$. For $\sigma(\mathbf{y}) = a p_y^n + b$, there will only be a drift for n odd. When n is even, the integrator is again reversible and drift is eliminated for this problem.

We show this integrator is time-symmetric but not reversible in Fig. 7. Here, after choosing an ϵ, we run forwards for some given time. Then, we switch the sign of ϵ and run forwards the same number of steps. We repeat the experiment for various ϵ, and plot ϵ vs the energy error. For the second curve, labelled after running forwards, we change the sign of $\boldsymbol{p}$ instead of ϵ. The errors of the first experiment are small, given by roundoff error, and indicated as the "Time symmetry" error in Fig. 7. We measure the error energy of this operation. In the second case, we change the sign of $\boldsymbol{p}$ and run forward the same number of steps. The first experiment gives a small error, at the level of roundoff. The dashed blue line shows a slope $t^{1/2}$, which is the expected error growth, based on Brouwer's Law (Brouwer 1937). Even though the Brouwer's Law analysis only works for fixed time steps, in a run with $t \approx 39.77$, the standard deviation in the time step lengths is 6.3×10^{-4}, so this

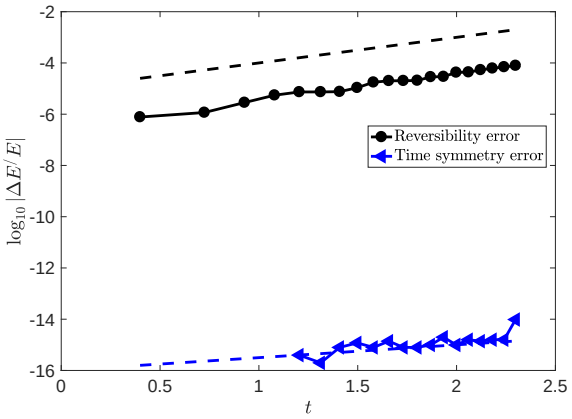

Figure 7. The time-symmetry and time-reversibility error for the orbit of Fig. 6. To compute the time symmetry error, we integrate forward for t, switch the sign of ϵ and integrate forward the same number of steps. For the reversibility error, we change the sign of $\boldsymbol{p}$ instead. The former error indicates the integrator is time-symmetric and the error grows as $t^{1/2}$ as expected. The latter error grows as t and shows reversibility is broken.

approximation is valid. The reversibility error shows that this integrator is not reversible. The dashed black line indicates a slope of t^1, as expected from the linear drift of Fig. 6 (there is a similar linear drift in Fig. 6 if we initialize with reversed $\boldsymbol{p}$). It is also possible to develop an integrator that is reversible, but not time-symmetric. For example, modify (44) to,

$$h = \frac{\epsilon^2}{2}\left[\sigma(\boldsymbol{y}_0) + \sigma(\boldsymbol{y}_1)\right], \tag{59}$$

with $\sigma(\boldsymbol{y})$ reversible. According to (47), this breaks time symmetry, but one might argue this choice is not sensible since it does not allow changing the sign of h with ϵ. But this time-symmetry break does not cause any new linear error drift. To get the analogue of (50) for step (59): replace ϵ^2 with ϵ^4 in (50).

Note we have not proved that all bound Hénon–Heiles orbits computed with a reversible integrator show no energy drift. It is possible that for some initial conditions, the solution of the MDEs is a bound orbit such that the energy has a secular increase or decrease, as was the case of Hamiltonian (52). All tested unbound orbits resulting from reversible and symmetric trapezoidal methods as well as from symplectic methods give secular energy change.

Now we repeat the experiment for different choices of $\sigma(\boldsymbol{y})$. In Fig. 8, we plot the error in energy vs. time, analogously to Fig. 6, but for these different choices of $\sigma(\boldsymbol{y})$, one irreversible and two reversible. No linear energy drift is observed in any case. We also used an explicit second order Runge–Kutta method to integrate the orbit, the explicit midpoint rule. In the notation of Section 3.2, this method has $c = 1$, $b_1 = 0$, $b_2 = 1$, $a_{21} = 1/2$, and $s = 2$. Using $h = 0.1$, we get linear drift in energy error, as expected from standard numerical analysis. These experiments demonstrate that reversibility or symplecticity of a method is not a requirement for energy conservation. As supported in the experiments and shown in Appendix C, time-symmetry and time-reversibility are properties independent to energy conservation.

We tested another regular orbit with $H = 1/12$ and initial conditions $p_x = \sqrt{1/6}$, $p_y = x = y = 0$. This time, the centroid (mean position in the x–y plane) is at 0. We plot the trajectory and equipotential curve, $U = 1/12$ in Fig. 9, using $\sigma = 1$, $\epsilon = 0.1$, and $t_{\max} = 200\pi$. The trajectory plot indicates this orbit is not far from a periodic resonance orbit, at the boundary between regular and chaotic orbits. We plot the surface of section in Fig. 10, again using the pair of Runge–Kutta methods with adaptive stepping, for $t = 10^5$, which gives a 7.4×10^{-14} energy error. This is a loop orbit: the angular momentum is less than 0 for all time. For this orbit, we did not find any reasonable $\sigma(\boldsymbol{y})$ that yields drift, whether σ is reversible or irreversible.

All chaotic orbits we tested show drift in energy, whether σ is constant or not. The drift increases as t^x where $0 < x \leqslant 1$. Chaotic orbits have been investigated elsewhere (Hairer et al. 2009; McLachlan & Perlmutter 2004; Hut et al. 1995); typically random walk behavior in the error ($\propto t^{1/2}$) is observed. We did not find any case in which a chaotic orbit gave a long term linear drift in energy error, as in the previous experiments with regular orbits.

These experiments for Hénon–Heiles show that good energy behavior is possible even with an irreversible integrator. They also show the range of appropriate step criterions :they depend on the orbit and problem and are not necessarily restrictive. To summarize, box orbit initial conditions mapped with a time-symmetric but irreversible integrator resulted in energy drift. But other irreversible integrators yielded no energy drift for this box orbit. A loop orbit did not give error drift for any tested integrator. All chaotic orbits gave some form of energy drift in all tests.

5 CONCLUSION

This work provides the error analysis needed to understand energy errors of symmetric integrators with adaptive steps used in astrophysics. We show how to study integrators using their modified differential equations (MDEs) and use this machinery to derive the MDEs for the trapezoidal rule with adaptive steps. The trapezoidal rule is a time-symmetric, but non-symplectic, integrator. Other authors have used the leapfrog method with adaptive steps; we do not study this because there is no advantage to using leapfrog as far as error properties are concerned, as discussed in Section 3.2 .We find that the trapezoidal rule, with adaptive steps, does not conserve the energy well for some problems, and we use the MDEs to explain this result. We cannot make broad statements about the energy conservation of time-symmetric methods because they have different MDEs from each other. The error of a symmetric integrator depends on the integrator, the differential equation, and the initial values. But there is no reason to think other methods will not suffer from the same shortcomings of the trapezoidal rule.

We also note that time-symmetry and reversibility are distinct concepts for an integrator(Hairer et al. 2006, Section VIII.3). Time symmetry means that if we reverse the sign of the time step, we can recover the initial conditions, while reversibility means that if we switch the sign of velocities and integrate forwards, we will recover the initial conditions. For the N-body problem with pairwise forces, which is reversible, there are integrators which are correctly both reversible and symmetric, neither reversible nor symmetric, or only symmetric or reversible; we study several of these combinations and the errors they lead to. For example, we find reversibility does not have to be a requirement for conserving the energy of a Hénon–Heiles orbit.

We conclude that while time-symmetric integration has often been observed to yield small errors over long time-scales when used for the N-body problem, it is not always the case that a time-symmetric integration will work successfully. We suggest that caution be used when deciding to use a time-symmetric method, and that preference should still be given to symplectic integrators. In

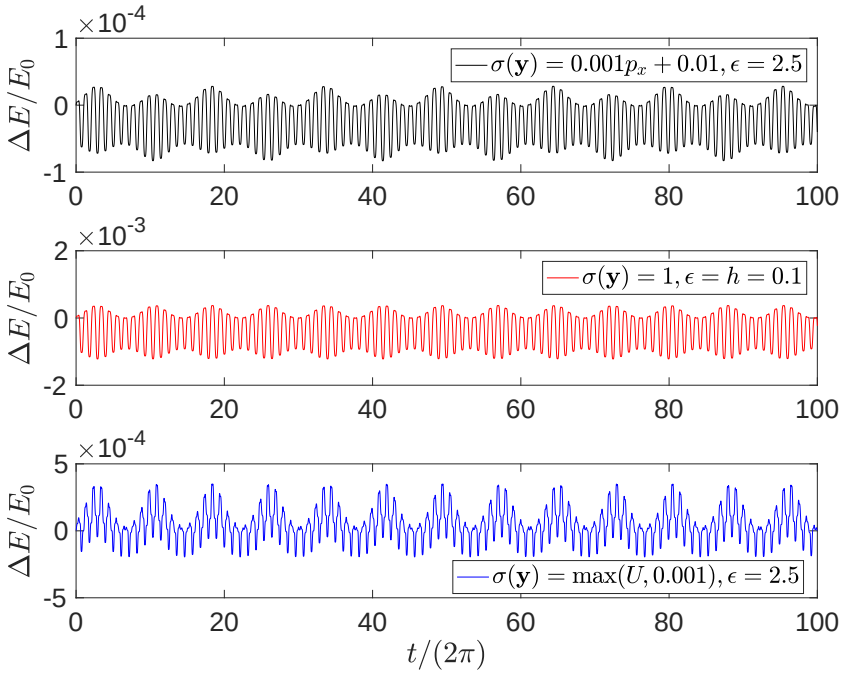

Figure 8. A repeat of the experiment of Fig. 6. The initial conditions are the same, but we vary ϵ and the step criteria σ. In no case do we observe energy drift. All panels show a time-symmetric integration, but only the second and third panel show reversible integration. This example shows that irreversibility does not imply energy drift.

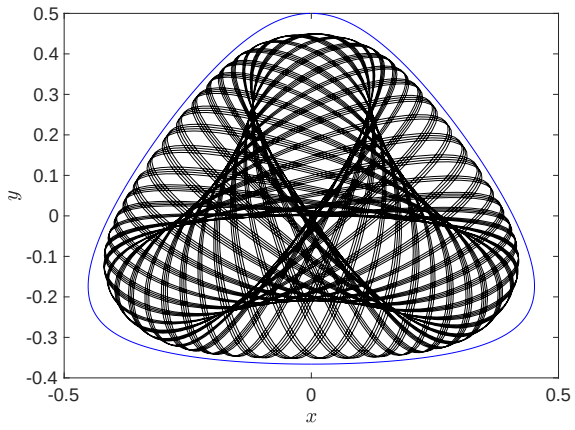

Figure 9. A regular loop orbit of the Hénon–Heiles problem and the bouding equipotential curve. The initial conditions are $p_x = \sqrt{1/6}$, $p_y = x = y = 0$ (and $H = 1/12$). The orbit is plotted until $t_{\max} = 100 \times 2\pi$.

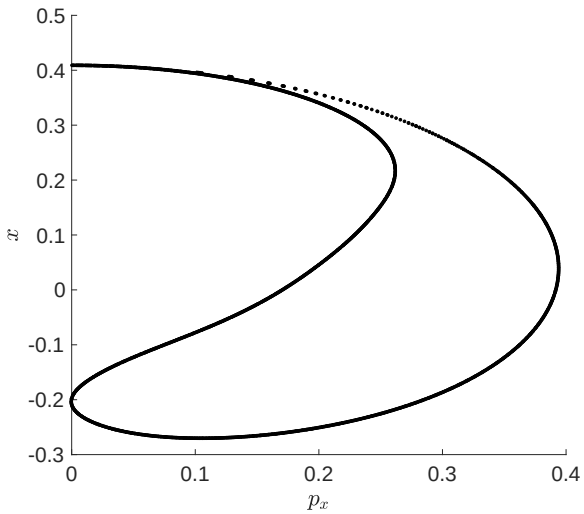

Figure 10. Surface of section plot for the orbit of Fig. 9. The plot is constructed in a similar way to Fig. 5. The discrete islands indicate the orbit is regular and near a resonance, and the asymmetry in p_x indicates a loop orbit.

general, time-symmetric methods are not guaranteed to conserve energy, unlike symplectic integrators, assuming convergence in the Hamiltonian (Hairer et al. 2006, Section IX.8).

6 ACKNOWLEDGEMENTS

We thank Walter Dehnen for stimulating discussions and detailed comments on the manuscript, the anonymous referee for comments that strengthened the paper, and Ernst Hairer for a helpful discussion.

APPENDIX A: SYMPLECTICITY OF IMPLICIT MIDPOINT RULE

The Jacobian is

$$\boldsymbol{S} \equiv \frac{\partial \boldsymbol{y}'}{\partial \boldsymbol{y}} \,. \tag{A1}$$

For the ordering $\boldsymbol{y} = (\boldsymbol{q}, \boldsymbol{p})$ (which we can choose without loss of

generality), the symplectic condition is a matrix equation with $2n^2+n$ independent constraints,

$$\boldsymbol{J} = \boldsymbol{S}\boldsymbol{J}\boldsymbol{S}^\dagger\ ,\quad \boldsymbol{J} \equiv \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I}_n \\ -\boldsymbol{I}_n & \boldsymbol{0} \end{pmatrix}. \tag{A2}$$

Here, $\boldsymbol{y}$ has $2n$ elements while $\boldsymbol{I}_n$ is the $n\times n$ identity matrix. Let $\boldsymbol{I} \equiv \boldsymbol{I}_{2n}$. Differentiating the implicit midpoint rule, (18), with respect to $\boldsymbol{y}$ gives,

$$\begin{aligned}\boldsymbol{S} &= \boldsymbol{I} + h\frac{\partial}{\partial \boldsymbol{y}} f\left(\frac{\boldsymbol{y}+\boldsymbol{y}'}{2}\right) \\ &= \boldsymbol{I} + \frac{1}{2}h\boldsymbol{J}\boldsymbol{H}(\boldsymbol{I}+\boldsymbol{S})\ ,\quad \boldsymbol{H} \equiv \boldsymbol{\partial}\boldsymbol{\partial} H_0\end{aligned} \tag{A3}$$

where H_0 is the original Hamiltonian and $\boldsymbol{H}$ is its Hessian matrix evaluated at the midpoint. We can solve for S to get,

$$\boldsymbol{S} = (\boldsymbol{I}-\boldsymbol{A})^{-1}(\boldsymbol{I}+\boldsymbol{A})\ ,\quad \boldsymbol{A} \equiv \frac{1}{2}h\boldsymbol{J}\boldsymbol{H}\ . \tag{A4}$$

Next, note that,

$$(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{J}(\boldsymbol{I}-\boldsymbol{A}^\dagger) = (\boldsymbol{I}+\boldsymbol{A})\boldsymbol{J}(\boldsymbol{I}+\boldsymbol{A}^\dagger). \tag{A5}$$

From this we find that,

$$\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}^\dagger = (\boldsymbol{I}-\boldsymbol{A})^{-1}(\boldsymbol{I}+\boldsymbol{A})\boldsymbol{J}(\boldsymbol{I}+\boldsymbol{A}^\dagger)(\boldsymbol{I}-\boldsymbol{A}^\dagger)^{-1} = \boldsymbol{J}, \tag{A6}$$

which implies the midpoint rule is symplectic.

APPENDIX B: MODIFIED DIFFERENTIAL EQUATION FOR TRAPEZOIDAL AND IMPLICIT MIDPOINT RULE

We derive the MDEs for the symmetric trapezoidal rule in order to understand its error properties and whether it conserves energy. Although Hamilton's equations split the configuration space into coordinates and momenta, the numerical integration algorithms need not do so. For RK-methods, in particular, the update rules depend on scalar operators formed from $\boldsymbol{f}$ and $\boldsymbol{\partial} \equiv \partial/\partial \boldsymbol{y}$. These operators are defined so that

$$\boldsymbol{f}_n = \sum_{m=0}^{M_{n-1}} f_{nm}\hat{D}_{nm}\boldsymbol{f}\ , \tag{B1}$$

where $\boldsymbol{f}_n(\boldsymbol{y})$ is the nth-order contribution to the modified differential equation, f_{nm} are constants, and $\hat{D}_{nm}$ are scalar differential operators. f_{nm} is not a component of $\boldsymbol{f}_n$. M_n is the number of unlabeled rooted trees with n nodes (Hairer et al. 1993, Table 2.1) Note that $\hat{D}_{nm}\boldsymbol{f}$ provide a basis for the Hilbert space of $\boldsymbol{F}$. Expansion (B1) represents the function $F(\boldsymbol{y},h)$ by a set of constants f_{nm}.

We now show how to obtain all such operators recursively in powers of h. At first order ($n=1$) there is only one operator,

$$\hat{D}_{10} \equiv \boldsymbol{f}\cdot\frac{\partial}{\partial \boldsymbol{y}} \equiv f_i\partial_i. \tag{B2}$$

The implied summation of i is from 1 to $2n$. The subscripts f_i indicate components of $\boldsymbol{f}$ in (1) in equations (B2)–(B8), they are not the indices of (9). There is no other scalar operator that can be formed from $\boldsymbol{f}$ and $\partial/\partial\boldsymbol{y}$ that has units of $\boldsymbol{f}/h$, hence $M_1 = 0$. In equations (10)–(11), $\hat{D}_{10}$ was written as D_0.

At second order ($n=2$), there are two linearly independent scalar operators with the correct units:

$$\hat{D}_{20} \equiv (\hat{D}_{10}f_i)\partial_i\ ,\ \hat{D}_{21} \equiv f_if_j\partial_i\partial_j\ . \tag{B3}$$

Note that $\hat{D}_{10}^2 = \hat{D}_{20} + \hat{D}_{21}$. We exclude $f_if_i\partial_j\partial_j$ and similar operators, even though they are scalars, because they do not arise in the series expansion of Runge–Kutta methods. At third order, there are four operators:

$$\begin{aligned}&\hat{D}_{30} \equiv (\hat{D}_{20}f_i)\partial_i\ ,\ \hat{D}_{31} \equiv (\hat{D}_{21}f_i)\partial_i\ ,\ \hat{D}_{32} \equiv f_i(\hat{D}_{10}f_j)\partial_i\partial_j\ , \\ &\hat{D}_{33} \equiv f_if_jf_k\partial_i\partial_j\partial_k\ .\end{aligned} \tag{B4}$$

At fourth order, there are 9 operators:

$$\begin{aligned}&\hat{D}_{40} \equiv (\hat{D}_{30}f_i)\partial_i\ ,\ \hat{D}_{41} \equiv (\hat{D}_{31}f_i)\partial_i\ ,\ \hat{D}_{42} \equiv (\hat{D}_{32}f_i)\partial_i\ ,\ \hat{D}_{43} \equiv (\hat{D}_{33}f_i)\partial_i\ , \\ &\hat{D}_{44} \equiv f_i(\hat{D}_{20}f_j)\partial_i\partial_j\ ,\ \hat{D}_{45} \equiv f_i(\hat{D}_{21}f_j)\partial_i\partial_j\ ,\ \hat{D}_{46} \equiv (\hat{D}_{10}f_i)(\hat{D}_{10}f_j)\partial_i\partial_j\ , \\ &\hat{D}_{47} \equiv f_if_j(\hat{D}_{10}f_k)\partial_i\partial_j\partial_k\ ,\ \hat{D}_{48} \equiv f_if_jf_kf_l\partial_i\partial_j\partial_k\partial_l\ .\end{aligned} \tag{B5}$$

The pattern becomes clear: at order n, the first M_{n-1} operators are formed from the operators of order $(n-1)$ acting on f_i combined with ∂_i while the remaining operators are formed from operators of order $n-2, n-3, \ldots, 0$ and additional derivative operators. At fifth order there are a total of $M_5 = 20$ operators; the first 9 are $\hat{D}_{5m} = (\hat{D}_{4m}f_i)\partial_i$. Note that the units of D_{nm} are h^{-n}.

Using equations (9) and (B1), the time evolution operator is now

$$\hat{D} = \boldsymbol{F}\cdot\boldsymbol{\partial} = \sum_{n=0}^{\infty} h^n \sum_{m=0}^{M_n} f_{nm}(\hat{D}_{nm}\boldsymbol{f})\cdot\boldsymbol{\partial} = \sum_{n=0}^{\infty} h^n \sum_{m=0}^{M_n} f_{nm}\hat{D}_{n+1,m}\ . \tag{B6}$$

For the trapezoidal rule, (17), Taylor expanding $\boldsymbol{f}(\boldsymbol{y}_1)$ about $\boldsymbol{y}_0$ gives,

$$\begin{aligned}\boldsymbol{g}_0 &= \boldsymbol{f} \\ \boldsymbol{g}_1 &= \frac{1}{2}D_0\boldsymbol{f} = \frac{1}{2}\hat{D}_{10}\boldsymbol{f} \\ \boldsymbol{g}_2 &= \frac{1}{4}D_0^2\boldsymbol{f} = \frac{1}{4}(\hat{D}_{20}+\hat{D}_{21})\boldsymbol{f} \\ \boldsymbol{g}_3 &= \frac{1}{12}D_0^3\boldsymbol{f} + \frac{1}{24}(D_0^2f_j)(\partial_j\boldsymbol{f}) = \frac{1}{8}\left(\hat{D}_{30}+\hat{D}_{31}+2\hat{D}_{32}+\frac{2}{3}\hat{D}_{33}\right)\boldsymbol{f} \\ \boldsymbol{g}_4 &= \frac{1}{48}\left\{D_0^4\boldsymbol{f} + (D_0^2f_j)\left[\partial_j(D_0\boldsymbol{f})\right] + D_0\left[(D_0^2f_j)(\partial_j\boldsymbol{f})\right]\right\} \\ &= \frac{1}{16}\left(\hat{D}_{40}+\hat{D}_{41}+2\hat{D}_{42}+\frac{2}{3}\hat{D}_{43}+2\hat{D}_{44}+2\hat{D}_{45}+\hat{D}_{46}\right. \\ &\quad \left.+2\hat{D}_{47}+\frac{1}{3}\hat{D}_{48}\right)\boldsymbol{f}.\end{aligned} \tag{B7}$$

Substituting into (11) and solving for $\boldsymbol{f}_n$ gives

$$\begin{aligned}\boldsymbol{f}_0 &= \boldsymbol{f} \\ \boldsymbol{f}_1 &= 0 \\ \boldsymbol{f}_2 &= \frac{1}{12}D_0^2\boldsymbol{f} = \frac{1}{12}(\hat{D}_{20}+\hat{D}_{21})\boldsymbol{f} \\ \boldsymbol{f}_3 &= 0 \\ \boldsymbol{f}_4 &= -\frac{1}{720}D_0^4\boldsymbol{f} + \frac{1}{144}\left\{(D_0^2f_j)\left[\partial_j(D_0\boldsymbol{f})\right] + D_0\left[(D_0^2f_j)(\partial_j\boldsymbol{f})\right]\right\} \\ &= \frac{1}{240}\times \\ &\left(3\hat{D}_{40}+3\hat{D}_{41}+4\hat{D}_{42}+\frac{4}{3}\hat{D}_{43}+2\hat{D}_{44}+2\hat{D}_{45}-\hat{D}_{46}-2\hat{D}_{47}-\frac{1}{3}\hat{D}_{48}\right)\boldsymbol{f}\end{aligned} \tag{B8}$$

For the implicit midpoint rule, (18),

$$\begin{aligned}
\boldsymbol{g}_0 &= \boldsymbol{f} \\
\boldsymbol{g}_1 &= \frac{1}{2}\hat{D}_{10}\boldsymbol{f} \\
\boldsymbol{g}_2 &= \frac{1}{4}\left(\hat{D}_{20}+\frac{1}{2}\hat{D}_{21}\right)\boldsymbol{f} \\
\boldsymbol{g}_3 &= \frac{1}{8}\left(\hat{D}_{30}+\frac{1}{2}\hat{D}_{31}+\hat{D}_{32}+\frac{1}{6}\hat{D}_{33}\right)\boldsymbol{f} \\
\boldsymbol{g}_4 &= \frac{1}{16}\left(\hat{D}_{40}+\hat{D}_{42}+\hat{D}_{44}+\frac{1}{2}(\hat{D}_{41}+\hat{D}_{45}+\hat{D}_{46}+\hat{D}_{47})+\frac{1}{6}\hat{D}_{43}+\frac{1}{24}\hat{D}_{48}\right)\boldsymbol{f}
\end{aligned} \tag{B9}$$

which leads to

$$\begin{aligned}
&\boldsymbol{f}_0 = \boldsymbol{f}\ ,\ \ \boldsymbol{f}_1 = 0\ ,\ \ \boldsymbol{f}_2 = \frac{1}{12}\left(\hat{D}_{20}-\frac{1}{2}\hat{D}_{21}\right)\boldsymbol{f}\ ,\ \ \boldsymbol{f}_3 = 0\ , \\
&\boldsymbol{f}_4 = \frac{1}{480}\times \\
&\left(6(\hat{D}_{40}-\hat{D}_{44})+\hat{D}_{41}-2\hat{D}_{42}-\hat{D}_{45}+3\hat{D}_{46}+\hat{D}_{47}+\frac{7}{12}(-4\hat{D}_{43}+\hat{D}_{48})\right)\boldsymbol{f}\ .
\end{aligned} \tag{B10}$$

To derive the Hamiltonian for the midpoint rule to fourth order, we use the procedure of Appendix C. It is

$$\tilde{H} = H - \frac{h^2}{24}(\hat{D}_{21}H) + \frac{h^4}{480}\left(3\hat{D}_{40}+\hat{D}_{41}-\frac{7}{12}\hat{D}_{43}\right)H + O(h^6)\ , \tag{B11}$$

where H is the original Hamiltonian.

APPENDIX C: RUNGE–KUTTA METHODS AND ENERGY CONSERVATION

We now obtain some general results concerning energy conservation for Runge–Kutta methods based on Hamiltonian systems. In this section we do not assume that energy conservation implies canonical transformation, even though the reverse is true (canonical transformation implies the existence of a local Hamiltonian, hence energy conservation for a time-independent Hamiltonian).

We consider conservative systems, for which equations (1) take the form of Hamilton's equations,

$$\frac{dq^I}{dt} = \frac{\partial H}{\partial p_I} \equiv H^I\ ,\ \ \frac{dp_I}{dt} = -\frac{\partial H}{\partial q^I} \equiv -H_I\ ,\ \ H = H(\boldsymbol{q},\boldsymbol{p})\ . \tag{C1}$$

We use superscript and lowerscript indices to distinguish derivatives with respect to coordinates and momenta. Einstein summation convention is also used. We use these notations only in this Appendix to simplify results. For a configuration space of n coordinates and n momenta, indices range from 1 to n.

We now ask under what conditions RK methods applied to a conservative Hamiltonian system have a conserved energy

$$E(\boldsymbol{y},h) = H_0(\boldsymbol{y}) + hH_1(\boldsymbol{y}) + h^2H_2(\boldsymbol{y}) + h^3H_3(\boldsymbol{y}) + h^4H_4(\boldsymbol{y}) + \cdots \tag{C2}$$

such that E is constant for solutions of the modified differential equation. (Note that we are not requiring the integrator to be symplectic; the relationship between symplectic and energy-conserving integrators will be clarified later.) In other words, the solutions must obey $\hat{D}E = F_i\partial_i E = 0$. Applying (B6) to (C2) gives

$$\sum_{k=0}^{n}\sum_{m=0}^{M_{k-1}} f_{km}\hat{D}_{k+1,m}H_{n-k} = 0 \tag{C3}$$

for all $n \geqslant 0$, with $M_0 = f_{00} = 1$.

Let's examine this order by order. For $n = 0$, equation (C3) is automatically satisfied because $\hat{D}_{10}H_0 = \hat{D}_{10}H = \{H,H\} = 0$: we are numerically integrating a Hamiltonian system. For $n = 1$, energy conservation requires that there exist a $H_1(\boldsymbol{y})$ satisfying

$$\hat{D}_{10}H_1 = -f_{10}\hat{D}_{20}H_0 \tag{C4}$$

For an RK method, H_n can be formed only from H_0 and scalar derivative operators D_{nm}. For $n = 1$, there is only one such operator, $\hat{D}_{10}$, and $\hat{D}_{10}H_0 = 0$. Therefore, *no first-order RK method has a conserved energy.* Examples are the explicit and implicit Euler methods, which usually exhibit a growth in the absolute value of the energy error that is linear in time. This behavior is explained with other numerical analysis. The only possibility that allows a conserved energy is $f_{10} = 0$, i.e. $\boldsymbol{f}_1 = 0$ and the integration method is at least second order.

As with the function $\boldsymbol{F}(\boldsymbol{y},h)$, for an RK-method we must represent $E(\boldsymbol{y},h)$ using scalar operators and the unique scalar function corresponding to $\boldsymbol{f}$, namely H_0. Thus, in equation (C3) we write

$$H_n = \sum_{m=0}^{M_{n-1}} e_{nm}\hat{D}_{nm}H_0\ . \tag{C5}$$

The following results are obtained (after much algebra) from equations (B2)–(B5):

$$\begin{aligned}
&\hat{D}_{10}H_0 = 0\ ,\ \hat{D}_{20}H_0 = -\hat{D}_{21}H_0\ ,\ \hat{D}_{30}H_0 = \hat{D}_{32}H_0 = 0\ , \\
&\hat{D}_{31}H_0 = -\hat{D}_{33}H_0\ , \\
&\hat{D}_{40}H_0 = -\hat{D}_{44}H_0 = D_{46}H_0 = (\hat{D}_{10}H^I)(\hat{D}_{20}H_I) - (\hat{D}_{10}H_I)(\hat{D}_{20}H^I)\ , \\
&\hat{D}_{41}H_0 = -\hat{D}_{42}H_0 = -\hat{D}_{45}H_0 = \hat{D}_{47}H_0 \\
&\qquad = (\hat{D}_{10}H^I)(\hat{D}_{21}H_I) - (\hat{D}_{10}H_I)(\hat{D}_{21}H^I)\ , \\
&\hat{D}_{43}H_0 = -\hat{D}_{48}H_0\ , \\
&\hat{D}_{50}H_0 = \hat{D}_{55}H_0 = 0\ ,\ \hat{D}_{51}H_0 = \hat{D}_{54}H_0 \\
&\qquad = (\hat{D}_{20}H_I)(\hat{D}_{21}H^I) - (\hat{D}_{20}H^I)(\hat{D}_{21}H_I)\ , \\
&\hat{D}_{52}H_0 = -D_{56}H_0 = (\hat{D}_{10}H^I)(\hat{D}_{32}H_I) - (\hat{D}_{10}H_I)(\hat{D}_{32}H^I)\ , \\
&\hat{D}_{53}H_0 = -\hat{D}_{57}H_0 = (\hat{D}_{10}H^I)(\hat{D}_{33}H_I) - (\hat{D}_{10}H_I)(\hat{D}_{33}H^I)\ .
\end{aligned} \tag{C6}$$

On account of these results, many of the dimensionless coefficients e_{nm} can be set to zero without loss of generality, so that

$$\begin{aligned}
E(\boldsymbol{y},h) = H_0 &+ h^2e_{21}\hat{D}_{21}H_0 + h^3e_{31}\hat{D}_{31}H_0 \\
&+ h^4\left(e_{40}\hat{D}_{40}+e_{41}\hat{D}_{41}+e_{43}\hat{D}_{43}\right)H_0 + O(h^5)\ .
\end{aligned} \tag{C7}$$

The task is now to find expressions for the e_{nm} in terms of the f_{nm}, as well as any conditions on the f_{nm} that must be satisfied in order to have energy conservation.

The following identities are also useful:

$$\begin{aligned}
&\hat{D}_{10}\hat{D}_{21}H_0 = -\hat{D}_{31}H_0\ ,\ \hat{D}_{10}\hat{D}_{31}H_0 = (-3\hat{D}_{41}+\hat{D}_{43})H_0\ , \\
&\hat{D}_{30}\hat{D}_{20}H_0 = -\hat{D}_{30}\hat{D}_{21}H_0 = \frac{1}{2}\hat{D}_{31}\hat{D}_{20}H_0 = -\frac{1}{2}\hat{D}_{31}\hat{D}_{21}H_0 = \hat{D}_{51}H_0\ , \\
&\hat{D}_{10}\hat{D}_{40}H_0 = (2\hat{D}_{51}+\hat{D}_{52})H_0\ ,\ \ \hat{D}_{10}\hat{D}_{41}H_0 = (-\hat{D}_{51}+2\hat{D}_{52}+\hat{D}_{53})H_0\ , \\
&\hat{D}_{10}\hat{D}_{43}H_0 = (\hat{D}_{58}-4\hat{D}_{53})H_0
\end{aligned} \tag{C8}$$

Combining these results gives the conditions for energy conserva-

tion up to fourth order:

$$
\begin{aligned}
&O(h^1):\ f_{10}=0\\
&O(h^2):\ e_{21}=f_{21}\\
&O(h^3):\ f_{30}=0\ ,\ f_{31}-f_{32}+3f_{33}=0\ ,\ e_{31}=-f_{33}\\
&O(h^4):\ f_{41}+f_{44}-2(f_{42}-f_{46})+5(f_{43}-f_{47}+4f_{48})-f_{21}(f_{20}+2f_{21})=0\ ,\\
&\qquad e_{40}=-f_{42}+f_{46}+2(f_{43}-f_{47}+4f_{48})\ ,\\
&\qquad e_{41}=-f_{43}+f_{47}-4f_{48}\ ,\ e_{43}=-f_{48}\ .
\end{aligned}
\tag{C9}
$$

The equations involving no e_{nm} are constraints on the numerical method in order that it have a conserved energy. At second order, there is no constraint: every second-order RK method has a conserved energy to second order, regardless whether the method is symplectic. For example, the explicit midpoint method typically shows linear growth in the absolute value of the energy. It is a second order RK method, but the slope of the linear drift scales as h^3, as we can check. At third order, there are two constraints on the four coefficients f_{3m}, so that most third-order y-methods do not have third-order energy conservation property. At fourth order, there is one constraint on the nine coefficients f_{4m} in order that a conserved energy result.

Kutta's third order method violates energy conservation at third order, while the classic Runge–Kutta fourth order method violates energy conservation at fourth order.

We will see in Appendix D that symplectic methods have additional constraints beyond those given above. Symmetric integrators are purely even in h, so that $f_{nm}=0$ for odd n. Not all symmetric integrators have a conserved energy, but all symplectic ones do. Thus, the set of symplectic integrators is a subset of the set of energy-conserving ones, and the set of symmetric integrators overlaps with both. Recall for non-adaptive one-step methods, time-symmetry and time-reversibility are equivalent.

APPENDIX D: SYMPLECTIC RUNGE–KUTTA-METHODS

Symplectic integrators are ones for which the mapping $\boldsymbol{y}_0\to\boldsymbol{y}_1$ is a canonical transformation. In this case the modified differential equation (3) is equivalent to Hamilton's equations (C1) with modified Hamiltonian $H(\boldsymbol{y},h)$. The modified Hamiltonian is expanded in power series exactly the same as $E(\boldsymbol{y},h)$ in equation (C2); we will use the same coefficients, with the expectation that requiring the integrator to be symplectic will yield different constraints than equations (C9). Enforcing symplecticity requires using the following identities,

$$
\begin{aligned}
&\partial_i(\hat{D}_{20}H)=(2\hat{D}_{20}-\hat{D}_{21})H_i\ ,\ \partial_i(\hat{D}_{31}H)=(3\hat{D}_{31}-\hat{D}_{33})H_i\ ,\\
&\partial_i(\hat{D}_{40}H)=\left[2(\hat{D}_{40}-\hat{D}_{44})+\hat{D}_{46}\right]H_i\ ,\ \partial_i(\hat{D}_{43}H)=(4\hat{D}_{43}-\hat{D}_{48})H_i\ ,\\
&\partial_i(\hat{D}_{41}H)=(\hat{D}_{41}-2\hat{D}_{42}-\hat{D}_{45}+\hat{D}_{47})H_i\ .
\end{aligned}
\tag{D1}
$$

Applying these gives the following conditions for symplecticity of RK-integrators, up to fourth order in h:

$$
\begin{aligned}
&O(h^1):\ f_{10}=0\\
&O(h^2):\ f_{20}=-2f_{21},\ e_{21}=f_{21}\\
&O(h^3):\ f_{30}=f_{32}=0,\ f_{31}=-3f_{33},\ e_{31}=\frac{1}{3}f_{31}\\
&O(h^4):\ f_{40}=-f_{44}=2f_{46},\ f_{41}=-\frac{1}{2}f_{42}=-f_{45}=f_{47},\ f_{43}=-4f_{48},\\
&\qquad e_{40}=\frac{1}{2}f_{40},\ e_{41}=f_{41},\ e_{43}=\frac{1}{4}f_{43}\ .
\end{aligned}
\tag{D2}
$$

Notice that these conditions include, but are stronger than, the energy-conserving conditions (C9). Symplectic integrators for an autonomous Hamiltonian system are always energy-conserving. However, the set of energy-conserving integrators is larger: up to fourth order, there are energy-conserving Runge–Kutta methods that are not symplectic, such as Lobatto IIIA (but note Lobatto IIIA does not conserve energy at higher orders according to Faou et al. (2004)). Lobatto IIIB, to fourth order, is neither symplectic nor conserves energy. There also exist third-order symplectic integrators, which are not symmetric.

APPENDIX E: GENERAL RUNGE–KUTTA INTEGRATORS

The general s-stage Runge–Kutta method can be written

$$
\boldsymbol{y}'=\boldsymbol{y}+h\sum_{i=1}^{s}b_i\boldsymbol{k}_i\ ,\ \ \boldsymbol{k}_i=\boldsymbol{f}(\boldsymbol{y}+h\boldsymbol{q}_i)\ ,\ \ \boldsymbol{q}_i\equiv[a_{ij}\boldsymbol{k}_j]\equiv\sum_{j=1}^{s}a_{ij}\boldsymbol{k}_j\ .
\tag{E1}
$$

Square brackets indicate a sum over the repeated indices inside the sum, e.g.

$$
[a_{ij}c_j^2]\equiv\sum_{j=1}^{s}a_{ij}c_j^2\ ,\ \ [a_{ij}a_{jk}c_k]\equiv\sum_{j=1}^{s}\sum_{k=1}^{s}a_{ij}a_{jk}c_k\ .
\tag{E2}
$$

We also define

$$
c_i\equiv\sum_{j=1}^{s}a_{ij}\ .
\tag{E3}
$$

The equation (E1) for $\boldsymbol{k}_i$ is recursive. Expanding in power series in h gives $\boldsymbol{k}_i=\hat{K}_i\boldsymbol{f}$, where the propagator is

$$
\begin{aligned}
\hat{K}_i=1&+hc_i\hat{D}_{10}+h^2[a_{ij}c_j]\hat{D}_{20}+\frac{1}{2}h^2c_i^2\hat{D}_{21}+h^3[a_{ij}a_{jk}c_k]\hat{D}_{30}\\
&+\frac{1}{2}h^3[a_{ij}c_j^2]\hat{D}_{31}+h^3c_i[a_{ij}c_j]\hat{D}_{32}+\frac{1}{6}h^3c_i^3\hat{D}_{33}+h^4[a_{ij}a_{jk}a_{kl}c_l]\hat{D}_{40}\\
&+\frac{1}{2}h^4[a_{ij}a_{jk}c_k^2]\hat{D}_{41}+h^4[a_{ij}a_{jk}c_jc_k]\hat{D}_{42}+\frac{1}{6}h^4[a_{ij}c_j^3]\hat{D}_{43}\\
&+h^4c_i[a_{ij}a_{jk}c_k]\hat{D}_{44}+\frac{1}{2}h^4c_i[a_{ij}c_j^2]\hat{D}_{45}+\frac{1}{2}h^4[a_{ij}c_j]^2\hat{D}_{46}\\
&+\frac{1}{2}h^4c_i^2[a_{ij}c_j]\hat{D}_{47}+\frac{1}{24}h^4c_i^4\hat{D}_{48}+O(h^5)\ .
\end{aligned}
\tag{E4}
$$

The integrator method is now

$$
\boldsymbol{G}=\sum_{i=1}^{s}b_i\hat{K}_i\boldsymbol{f}
\tag{E5}
$$

This can be used for various methods to check the order of an integrator, its energy conservation properties, and its symplecticity, order by order.

Symmetric integrators are a special class of integrators for which equation (4) holds. For $s=1$, the implicit midpoint method is the only symmetric Runge–Kutta integrator. For $s=2$, the general class is defined by two parameters (a_{11},a_{12}) through the Runge–Kutta matrix

$$
A_{\mathrm{symm2}}=\begin{pmatrix}a_{11} & a_{12}\\ \frac{1}{2}-a_{12} & \frac{1}{2}-a_{11}\end{pmatrix}\ ,\ \ b_{\mathrm{symm2}}=\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\end{pmatrix}\ .
\tag{E6}
$$

Elements of A_{symm2} and b_{symm2} are, respecitvely, the a_{ij} and b_i from eq. (21). These integrators are all at least second order because symmetry implies $\boldsymbol{f}_1=\boldsymbol{f}_3=0$. They are not, in general, symplectic. There is one choice of (a_{11},a_{12}) for which the integrator is fourth

Table E1. Properties of implicit Runge–Kutta integrators. For various methods, we state the number of stages (s in eq. (21)), the order, and whether to all orders the methods are symmetric, energy conserving, and symplectic.

Method	Stages	Order	Symm.	Econs	Symp
Midpoint	1	2	yes	yes	yes
Trapezoidal	2	2	yes	yes	no
Symmetric	2	$\geqslant 2$	yes	no	no
Gauss–Legendre	2	4	yes	yes	yes
Lobatto IIIA	3	4	yes	no	no
Lobatto IIIB	3	4	yes	no	no
Symmetric	3	$\geqslant 2$	yes	no	no
Gauss–Legendre	3	6	yes	yes	yes

order and symplectic (symplecticity to all orders is proved elsewhere), namely the Gauss–Legendre case

$$a_{11} = \frac{1}{4}\,,\quad a_{12} = \frac{1}{4} - \frac{\sqrt{3}}{6}\,. \tag{E7}$$

The general $s = 3$ symmetric integrator has Runge–Kutta matrix and weight vector

$$A_{\rm symm3} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & \frac{1}{2}b_2 & b_1 - a_{21} \\ b_1 - a_{13} & b_2 - a_{12} & b_1 - a_{11} \end{pmatrix}\,,\quad b_{\rm symm3} = \begin{pmatrix} b_1 & b_2 & b_1 \end{pmatrix} \tag{E8}$$

with $b_2 = 1 - 2b_1$. In all cases this integrator is at least second order; in general it is not symplectic. The integrator is at least fourth order if the parameters obey the following two relations:

$$\begin{aligned} a_{11} + a_{12} + a_{13} &= \frac{1}{2} \pm (24 b_1)^{-1/2}\,, \\ b_1(a_{11} - a_{13}) + b_2\left(a_{21} - \frac{1}{2}b_2\right) &= \mp\left(\frac{b_1}{24}\right)^{1/2}\,. \end{aligned} \tag{E9}$$

There is one choice of parameters for which the integrator is sixth order and (at least to sixth order) symplectic, namely the Gauss–Legendre case

$$\begin{aligned} &a_{11} = \frac{5}{36}\,,\quad a_{12} = \frac{2}{9} - \frac{\sqrt{15}}{15}\,,\quad a_{33} = \frac{5}{36} - \frac{\sqrt{15}}{30}\,,\quad a_{21} = \frac{5}{36} + \frac{\sqrt{15}}{24}\,, \\ &b_1 = \frac{5}{18}\,. \end{aligned} \tag{E10}$$

Note that Gauss–Legendre methods have twice the order of truncation error expected from a naive count of function evaluations (e.g., 6 versus 3). In the context of Runge–Kutta methods this arises naturally because of symmetry: all odd terms vanish in the truncation error of the modified differential equation. Table E1 summarizes various properties of implicit Runge–Kutta integrators.

It has been shown that the general conditions for symplecticity of any Runge–Kutta integrator are (Hairer et al. 2006, Chapter VI)

$$b_i a_{ij} + b_j a_{ij} = b_i b_j \text{ for all } i, j \text{ such that } 1 \leqslant i, j \leqslant s. \tag{E11}$$

These conditions are satisfied by Gauss–Legendre integrators but not by Lobatto III integrators. (Hairer et al. 2006, Chapter VI) shows that Gauss-collocation methods (including the Gauss–Legendre methods above) are symplectic.

REFERENCES

Brouwer D., 1937, AJ, 46, 149

Chambers J. E., 1999, MNRAS, 304, 793

Channell P. J., Scovel C., 1990, Nonlinearity, 3, 231

Dehnen W., 2017, MNRAS, 472, 1226

Dehnen W., Hernandez D. M., 2017, MNRAS, 465, 1201

Duncan M. J., Levison H. F., Lee M. H., 1998, AJ, 116, 2067

Faou E., Hairer E., Pham T.-L., 2004, BIT Numerical Mathematics, 44, 699

Funato Y., Hut P., McMillan S., Makino J., 1996, AJ, 112, 1697

Hairer E., Wanner G., Nørsett S. P., 1993, Solving Ordinary Differential Equations I, 2nd edn. Springer Verlag, Berlin

Hairer E., Lubich C., Wanner G., 2006, Geometrical Numerical Integration, 2nd edn. Springer Verlag, Berlin

Hairer E., McLachlan R. I., Skeel R. D., 2009, ESAIM: Mathematical Modelling and Numerical Analysis, 43, 631

Henon M., Heiles C., 1964, AJ, 69, 73

Hernandez D. M., 2016, MNRAS, 458, 4285

Hernandez D. M., Bertschinger E., 2015, MNRAS, 452, 1934

Hut P., Makino J., McMillan S., 1995, ApJ, 443, L93

Kokubo E., Yoshinaga K., Makino J., 1998, MNRAS, 297, 1067

Makino J., Hut P., Kaplan M., Saygin H., 2006, New Astronomy, 12, 124

McLachlan R. I., Perlmutter M., 2004, Journal of Physics A: Mathematical and General, 37, L593

Pelupessy F. I., Jänes J., Portegies Zwart S., 2012, New Astronomy, 17, 711

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2002, Numerical recipes in C++ : the art of scientific computing

Springel V., 2005, MNRAS, 364, 1105

Stoffer D., 1995, Computing, 55, 1

Verner J. H., 1978, SIAM, 4, 772

Wisdom J., Holman M., 1991, AJ, 102, 1528