

Improved accuracy in degenerate variational integrators for guiding centre and magnetic field line flow

Joshua W. Burby^{1,†}, John M. Finn² and C. Leland Ellison³

¹Los Alamos National Laboratory, Los Alamos, NM 87544, USA

²Tibbar Plasma Technologies, Los Alamos, NM 87544, USA

³Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

(Received 9 March 2021; revised 27 October 2021; accepted 29 October 2021)

First-order-accurate degenerate variational integration (DVI) was introduced in Ellison *et al.* (*Phys. Plasmas*, vol. 25, 2018, 052502) for systems with a degenerate Lagrangian, i.e. one in which the velocity-space Hessian is singular. In this paper we introduce second-order-accurate DVI schemes, both with and without non-uniform time stepping. We show that it is not in general possible to construct a second-order scheme with a preserved two-form by composing a first-order scheme with its adjoint, and discuss the conditions under which such a composition is possible. We build two classes of second-order-accurate DVI schemes. We test these second-order schemes numerically on two systems having non-canonical variables, namely the magnetic field line and guiding centre systems. Variational integration for Hamiltonian systems with non-uniform time steps, in terms of an extended phase space Hamiltonian, is generalized to non-canonical variables. It is shown that preservation of proper degeneracy leads to single-step (one-step) methods without parasitic modes, i.e. to non-uniform time step DVIs. This extension applies to second-order-accurate as well as first-order schemes, and can be applied to adapt the time stepping to an error estimate.

Key word: plasma dynamics

1. Introduction

A variety of dissipation-free dynamical models in plasma physics share the property of arising from variational principles, including the guiding centre equations, magnetic field line flow and collision-free Vlasov dynamics. Variational integrators discretize the action associated with this variational principle rather than discretizing the equations of motion directly. The advantages of variational integrators are similar to those of the more specialized symplectic integrators, most of which discretize Hamiltonian systems in canonical variables. Variational integrators, especially those we study in this paper, can conveniently deal with non-canonical variables. These methods preserve exactly the

† Email address for correspondence: jburby@lanl.gov

Hamiltonian (or variational) nature of the original ordinary differential equation system, and the main advantages of these integration methods accrue when very long-time-scale behaviour is to be studied. For example, when using full-orbit simulations to assess the validity of the guiding centre approximation for runaway electrons, Liu, Wang & Qin (2016) concluded that simulations involving approximately 10^{11} time steps were required.

The task of finding a reliable variational integrator for a given variational dynamical system is generally quite challenging. The easiest case occurs when the Lagrangian underlying the variational principle is non-degenerate (see Marsden & West 2001), meaning that the velocity-space Hessian is invertible. Stable low-order and high-order variational integrators may be constructed in the non-degenerate setting using systematic procedures. The problem becomes much more challenging, however, when the velocity-space Hessian is degenerate, i.e. has a non-trivial null space. This sort of degeneracy arises, for instance, when dealing with a so-called phase-space Lagrangian (see Cary & Littlejohn 1983), either in canonical or non-canonical variables; such a Lagrangian is linear in the velocities, and therefore has Hessian equal to zero.

Ellison *et al.* (2018) showed that when the variational integrator formalism discussed in Marsden & West (2001) is applied to a phase-space Lagrangian, the resulting scheme typically performs very poorly; unphysical, potentially unstable parasitic modes arise and spoil the benefits of using a variational discretization. This flaw in the basic theory of variational integration represents a serious shortcoming as far as its applicability is concerned, because degenerate Lagrangians are commonly encountered in practice. In plasma physics phase-space Lagrangians are routinely used to model the fundamental problems of magnetic field line flow (see Cary & Littlejohn 1983) and guiding centre motion (see Littlejohn 1983) and have been shown to describe many infinite-dimensional plasma models as well (e.g. Burby 2015, 2017*a,b*).

In a more optimistic vein, Ellison *et al.* (2018) and Ellison (2016) also proposed a conceptually appealing strategy for avoiding the generic pitfalls of variational integration applied to phase-space Lagrangians. The idea was to select carefully the discretization of the Lagrangian so that it preserves the degeneracy of the continuous Lagrangian. Such properly degenerate discrete Lagrangians were shown to be free of the parasitic modes that plague generic discrete phase-space Lagrangians. What therefore emerged from this work was a refined notion of variational integration appropriate to phase-space Lagrangians (and perhaps more general degenerate Lagrangians as well) termed degenerate variational integration (DVI). In the same work, DVI was applied to magnetic field line flow under the gauge restriction that one covariant component of the magnetic field is zero. Also, DVI was applied to guiding centre motion, with the added physical restriction that the same covariant component of the magnetic field is zero. Indeed, good long-term behaviour of the orbits was observed. By exploiting a near-identity transformation of standard guiding centre theory, Burby & Ellison (2017) showed that DVI can still be applied to guiding centre dynamics if this stringent constraint on the magnetic field is lifted.

While DVI may be a promising candidate for coping with degenerate Lagrangians within a variational integration framework, DVI theory as it stands today is still in its infancy. In particular, the examples of degenerate variational integrators in Ellison *et al.* (2018) and Ellison (2016) suffer from three important drawbacks. (a) First, they start from continuous dynamics formulated in terms of either canonical variables or a restricted class of non-canonical variables. The magnetic field line and guiding centre examples of Ellison (2016) and Ellison *et al.* (2018) belong to this class. (b) Next, they only achieve first-order accuracy in time. (c) Finally, they rely on uniform time stepping. Moreover, it is presently

unclear if these drawbacks in examples are reflections of inherent limitations of the DVI concept, or only apparent limitations that might be overcome with additional insights.

The first purpose of this paper is to construct a large class of second-order-accurate degenerate variational integrators, involving the so-called processing technique described in Blanes, Casas & Murua (2004). In particular we aim to address issue (b) by formulating two related second-order-accurate DVI schemes and applying them to the field line and guiding centre problems considered in Ellison *et al.* (2018) and Burby & Ellison (2017). Our second purpose is to address issue (c) by formulating non-uniform time stepping for DVI. The method chosen is related to the well-known method of extended phase space (see Hairer, Lubich & Wanner 2006) generalized to the above class of non-canonical variables. Issue (a) will be the subject of future publications.

After providing an updated discussion of the basic elements of DVI theory in § 2, a class of second-order-accurate DVI schemes is presented in § 3. We begin by discussing these DVI schemes for systems with canonical variables, showing transparently why it is not possible in general to obtain second-order accuracy by composing a first-order scheme with its adjoint. We also show how to formulate these second-order systems in a restricted class of non-canonical systems. In § 4 we formulate these schemes for the magnetic field line integration problem and the guiding centre problem, both in this restricted class. In § 5 we report on the numerical application of these second-order DVI schemes to the field line and guiding centre examples, showing good long-time behaviour and second-order accuracy.

In § 6 we first discuss non-uniform time stepping in the context of canonical systems, showing the extended phase-space action and discretizations of it. We discuss the condition for a single-step scheme, i.e. a degenerate variational integrator, as described in § 2. This condition, as for uniform time steps, is that the discrete Hessian has the proper rank, the rank of the continuous Hessian. We proceed to show how to apply the extended phase-space method to the class of non-canonical variables described in § 3. We apply this methodology to the field line and guiding centre examples. These systems are in the special form of non-canonical variables described in § 2, and a method is described to obtain a single-step (DVI) scheme system with this special form of non-canonical variables. As before, the rank of the discrete Hessian predicts the single-step nature of the schemes. We also argue that it is straightforward to apply this extended phase-space method to second- and higher-order schemes and to adaptive time stepping.

We summarize and discuss the results of this paper in § 7.

2. Degenerate variational integration: review and recent developments

The purpose of this section is to provide details of the basic properties of DVI. While a similar discussion appears in Ellison *et al.* (2018), the ensuing discussion will reflect an improved understanding of DVI that has developed since the publication of Ellison *et al.* (2018).

Degenerate variational integration is a refinement of variational integration that applies to phase-space Lagrangians, and perhaps to more general degenerate Lagrangians. It is therefore helpful to recall briefly the basic ingredients of variational integration along the lines of Marsden & West (2001). To that end, consider a dynamical system governed by the variational principle based on the action S :

$$\delta S = \delta \int_{t_1}^{t_2} L(q(t), \dot{q}(t)) dt = 0, \quad (2.1)$$

where the dimension of q -space is m . Assume for now that the Lagrangian L is non-degenerate, which means the velocity space Hessian

$$M_{ij}(q, \dot{q}) = \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \quad (2.2)$$

is invertible for each (q, \dot{q}) . The most important consequence of non-degeneracy is that it implies the Euler–Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial q^i} \quad (2.3)$$

are a system of m second-order ordinary differential equations on q -space. Indeed, because we have

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = M_{ij} \ddot{q}^j + \frac{\partial L}{\partial \dot{q}^i \partial q^j} \dot{q}^j \quad (2.4)$$

(we use the standard summation convention for repeated indices), (2.3) may be rewritten as

$$\ddot{q}^l = (M^{-1})^{lj} \left(\frac{\partial L}{\partial q^j} - \frac{\partial^2 L}{\partial \dot{q}^j \partial q^i} \dot{q}^i \right), \quad (2.5)$$

which for a non-degenerate Hessian is a system of m second-order ordinary differential equations for q . Per the usual prescription, such equations are also equivalent to systems of $2m$ first-order ordinary differential equations advancing (q, \dot{q}) . Therefore (q, \dot{q}) -space is a suitable phase space for such a non-degenerate Lagrangian system. Another important property of non-degenerate Lagrangian systems is that it is possible, at least locally, to perform a Legendre transformation to obtain a Hamiltonian system in canonical variables, $q^i, p_i = \partial L / \partial \dot{q}^i$.

According to Marsden & West (2001), a variational integrator for a system with Lagrangian L is a time-marching algorithm that may be derived from the discrete-time approximation of (2.1), leading to a discrete variational principle based on the action S_d :

$$\delta S_d = \delta \sum_{k=N_1}^{N_2} h L_d(q_k, q_{k+1}) = 0, \quad (2.6)$$

where the discrete Lagrangian L_d is chosen as a specific approximation to the windowed time average of the Lagrangian according to

$$L_d(q_k, q_{k+1}) \approx \frac{1}{h} \int_{t_k}^{t_{k+1}} L(q(t), \dot{q}(t)) dt. \quad (2.7)$$

Here h denotes the spacing of a uniform temporal grid $t_k = kh$. The quantities q_k, q_{k+1} are evaluated at the endpoints of the interval $[t_k, t_{k+1}]$. The discrete Euler–Lagrange (DEL)

equations associated with the variational principle (2.6) are given by

$$\frac{\partial L_d}{\partial q_2^i}(q_{k-1}, q_k) + \frac{\partial L_d}{\partial q_1^i}(q_k, q_{k+1}) = 0. \quad (2.8)$$

Provided the discrete Hessian

$$\mathcal{M}_{ij}(q_0, q_1) = \frac{\partial L_d}{\partial q_0^i \partial q_1^j}(q_0, q_1) \quad (2.9)$$

is invertible for each (q_0, q_1) , the DEL equations define a mapping

$$(q_k, q_{k+1}) \mapsto (q_{k+1}, q_{k+2}^*(q_k, q_{k+1})), \quad (2.10)$$

where $q_{k+2}^*(q_k, q_{k+1})$ is the solution of (2.8) for q_{k+1} as a function of (q_{k-1}, q_k) guaranteed by the implicit function theorem.

When the continuous-time Lagrangian L is non-degenerate, i.e. the discrete Hessian is invertible, repeated application of the mapping (2.10) defined by the DEL equations (2.8) generates a sequence $k \mapsto q_k$ that approximates a solution of (2.3) sampled at the times $t_k = hk$. In particular, note that such a discrete-time trajectory requires $2m$ initial conditions to supply to (2.10), which is the same number of initial conditions required to specify a solution of the continuous-time Euler–Lagrange equations (2.3).

Suppose now that the continuous-time Lagrangian L has the general form

$$L(z, \dot{z}) = \vartheta_i(z) \dot{z}^i - H(z). \quad (2.11)$$

Lagrangians of this form in arbitrary variables are known as phase-space Lagrangians (see Cary & Littlejohn 1983), and are notable because every Hamiltonian system on an exact symplectic manifold is governed by such a Lagrangian. The class of Lagrangians in (2.11) is a generalization of phase-space Lagrangians of the form $p_i \dot{q}^i - H(q, p)$, in canonical variables. Here we have conformed with standard conventions when discussing phase-space Lagrangians in non-canonical variables by making the notational change $q \mapsto z$.

Because L is linear in the velocities, the velocity-space Hessian is zero, $M_{ij} = 0$. The Euler–Lagrange equations ((2.3) with $q \rightarrow z$) therefore cannot be equivalent to a system of $2m$ first-order ordinary differential equations on (z, \dot{z}) -space, or equivalently a system of second-order equations on m -dimensional z -space. Instead they are equivalent to the first-order system on the m -dimensional z -space given by

$$\dot{z}^j \omega_{ij} = \frac{\partial H}{\partial z^j}, \quad (2.12)$$

where the $m \times m$ antisymmetric matrix ω_{ij} is defined by

$$\omega_{ij}(z) = \frac{\partial \vartheta_i}{\partial z^j} - \frac{\partial \vartheta_j}{\partial z^i}. \quad (2.13)$$

If the functions ϑ_i have the property that ω_{ij} is invertible for each z , this yields a system of m first-order equations. It follows that the phase space for a phase-space Lagrangian has dimension m , which is half the dimension of the phase space for a non-degenerate Lagrangian. In particular, the number of initial conditions required to specify a solution of (2.12) is m instead of $2m$.

Application of variational integration to a phase-space Lagrangian is usually problematic for the following reason. Because invertible matrices are generic, most choices of discrete Lagrangian L_d will have an invertible discrete Hessian \mathcal{M} . One consequence of the invertibility of the discrete Hessian, which is suggestive that something is wrong, is that the DEL equations for L_d involve three time levels, as in (2.8), and therefore require $2m$ initial conditions to generate the discrete-time trajectory $k \mapsto z_k$. Indeed, from the perspective of the DEL equations (2.8), a discretized phase-space Lagrangian is no different from a discrete Lagrangian coming from a non-degenerate continuous-time Lagrangian. Therefore the mapping (2.10) is still well defined, which implies that the discrete system derived from a generic discrete Lagrangian requires $2m$ initial conditions, in spite of the fact that the underlying continuous-time system (2.12) requires only m initial conditions.

The preceding argument shows that typical variational integrators for phase-space Lagrangians are multi-step methods. Multi-step methods generally have parasitic modes, which may be unstable. Nevertheless, many multi-step methods have favourable numerical performance, in spite of the existence of these parasitic modes (see Hairer *et al.* 2006). In a reasonable scheme for a dissipative system, such parasitic modes damp out harmlessly in the early stages of the integration. However, the parasitic modes of multi-step variational integrators typically do not damp out. To understand why, we note that multi-step integrators arising from discrete phase-space Lagrangians must have either neutrally stable parasitic modes or one growing parasitic mode for each damped parasitic mode (see Ellison *et al.* 2018). This symmetry between damped and growing parasitic modes is a consequence of the preservation of a symplectic form on the $2m$ -dimensional space of pairs (z_1, z_2) . Thus, the parasitic modes associated with a discrete phase-space Lagrangian may be neutrally stable at best, but still susceptible to driving by nonlinear terms. Unfavourable behaviour of parasitic modes arising from examples of discrete phase-space Lagrangians is described in Ellison *et al.* (2018) and Ellison (2016).

Ellison *et al.* (2018) observed that the following special class of discrete phase Lagrangians avoid these multi-step issues in discretizing phase-space Lagrangians.

DEFINITION 2.1 (properly degenerate discrete Lagrangian). *Suppose the dimension m of phase space (coordinates z) is even. A discrete Lagrangian $L_d(z_0, z_1)$ is properly degenerate if the rank of the discrete Hessian is everywhere half-maximum. In other words the rank of $\mathcal{M}_{ij}(z_0, z_1)$ is m rather than $2m$ for all (z_0, z_1) .*

REMARK 2.2. *As discussed in Ellison (2016) and Ellison *et al.* (2018), for a system with continuous Hessian of rank zero, the condition that the discrete Hessian has rank half-maximum is exactly the condition that the DEL equations define a mapping on a space of the same dimension as the continuous system's phase space.*

REMARK 2.3. *We write the dimension as $2m$; this dimension is necessarily even when working with phase-space Lagrangians because an antisymmetric matrix like ω_{ij} always has a non-trivial null space in odd dimensions. So (2.12) cannot¹ be solved for \dot{z}^i .*

A simple example of a non-properly degenerate discrete Lagrangian in a one-degree-of-freedom ($m = 1$) system in canonical variables follows from a centred discretization of

¹An antisymmetric real matrix has pure imaginary eigenvalues occurring in complex conjugate pairs, so one eigenvalue must be zero in odd dimensions.

$S = \int (p\dot{q} - H) dt$:

$$L_d(q_0, p_0, q_1, p_1) = \frac{p_0 + p_1}{2} \frac{q_1 - q_0}{h} - H\left(\frac{q_0 + q_1}{2}, \frac{p_0 + p_1}{2}\right). \tag{2.14}$$

The DEL equations are

$$\frac{\partial L_d(q_0, p_0, q_1, p_1)}{\partial q_1} + \frac{\partial L_d(q_1, p_1, q_2, p_2)}{\partial q_1} = 0 \tag{2.15}$$

and

$$\frac{\partial L_d(q_0, p_0, q_1, p_1)}{\partial p_1} + \frac{\partial L_d(q_1, p_1, q_2, p_2)}{\partial p_1} = 0, \tag{2.16}$$

leading to

$$\frac{p_2 - p_0}{2h} = -\frac{1}{2}H_1(0, 1) - \frac{1}{2}H_1(1, 2), \tag{2.17}$$

$$\frac{q_2 - q_0}{2h} = \frac{1}{2}H_2(0, 1) + \frac{1}{2}H_2(1, 2), \tag{2.18}$$

where H_1 and H_2 are derivatives of H with respect to its first and second arguments, respectively, and the symbols $(0, 1)$, $(1, 2)$ represent $(q_0 + q_1)/2$, $(p_0 + p_1)/2$ and $(q_1 + q_2)/2$, $(p_1 + p_2)/2$, respectively. This is clearly a multi-step method, linking steps $k = 0, k = 1, k = 2$, i.e. giving second-order difference equations for q_k and p_k . (The finite difference forms for \dot{q} and \dot{p} suffer from ‘stencil spreading’.) The discrete Hessian for this system is

$$\begin{pmatrix} \partial^2 L_d / \partial q_0 \partial q_1 & \partial^2 L_d / \partial q_0 \partial p_1 \\ \partial^2 L_d / \partial p_0 \partial q_1 & \partial^2 L_d / \partial p_0 \partial p_1 \end{pmatrix} = \begin{pmatrix} -\frac{H_{11}((q_0 + q_1)/2, (p_0 + p_1)/2)}{4} & -\frac{1}{2h} \frac{H_{12}(q_0 + q_1)/2, (p_0 + p_1)/2)}{4} \\ \frac{1}{2h} \frac{H_{21}(q_0 + q_1)/2, (p_0 + p_1)/2)}{4} & -\frac{H_{22}(q_0 + q_1)/2, (p_0 + p_1)/2)}{4} \end{pmatrix}, \tag{2.19}$$

generically of full rank, showing agreement between the multi-step property of the DEL equations and the rank of the discrete Hessian for this system. And, indeed, this system requires an extra set of initial conditions and exhibits parasitic modes.

This example is to be contrasted with the first-order-accurate case arising from the discrete phase-space Lagrangian, again in one degree of freedom and in canonical variables:

$$L_d = p_k \frac{q_{k+1} - q_k}{h} - H(q_{k+1}, p_k), \tag{2.20}$$

leading to

$$q_{k+1} = q_k + hH_2(q_{k+1}, p_k), \quad p_{k+1} = p_k - hH_1(q_{k+1}, p_k), \tag{2.21a,b}$$

a form of the symplectic Euler scheme (compare with Hairer *et al.* 2006), with q updated implicitly and used in a leapfrog manner in the explicit update of p . The discrete Hessian

is

$$\begin{pmatrix} 0 & 0 \\ \frac{1}{h} - H_{21}(q_1, p_0) & 0 \end{pmatrix}, \tag{2.22}$$

of rank $m = 1$, traced to the fact that L_d in (2.20) depends on p at only one step. This result is consistent with the single-step nature of the scheme. The adjoint scheme S_h^\dagger also has a discrete Hessian with rank $m = 1$.

Ellison (2016) proves that properly degenerate discrete phase-space Lagrangians are necessarily free of parasitic modes. This result suggests, but does not directly imply, that variational integrators derived from properly degenerate discrete phase-space Lagrangians are single-step methods instead of multi-step methods. In fact, under mild technical hypotheses, properly degenerate discrete phase-space Lagrangians are indeed single-step methods.

The simplest way to understand the single-step nature of variational integrators obtained from properly degenerate discrete phase-space Lagrangians is to restrict our attention to the linearized DEL equations.

THEOREM 2.4 (linearized single-step property). *Let $L_d(z_1, z_2)$ be a properly degenerate discrete Lagrangian satisfying the mild technical hypotheses (G1)–(G2) described in Appendix A. Then the DEL equations linearized about a trajectory $k \mapsto z_k^0$ are equivalent to a single-step method.*

REMARK 2.5. *For a complete statement and proof of theorem 2.4, see theorem A.1 in Appendix A.*

Proof sketch. Let $k \mapsto z_k^\epsilon$ (the limit $\epsilon \rightarrow 0$ represents a reference trajectory) be a smooth ϵ -dependent family of solutions of the DEL equations associated with a properly degenerate discrete Lagrangian:

$$\frac{\partial L_d}{\partial z_2^\epsilon}(z_{k-1}^\epsilon, z_k^\epsilon) + \frac{\partial L_d}{\partial z_1^\epsilon}(z_k^\epsilon, z_{k+1}^\epsilon) = 0, \tag{2.23}$$

where $\partial/\partial z_1$ and $\partial/\partial z_2$ refer to derivatives with respect to the first and second arguments of L_d . Differentiating the DEL equations with respect to ϵ at $\epsilon = 0$ shows that the linearization $k \mapsto \delta z_k = \partial_\epsilon(z_k^\epsilon)_{\epsilon=0}$ of a trajectory near $k \mapsto z_k^0$ satisfies the linearized DEL equations

$$A_{ij}(k)\delta z_{k-1}^j + C_{ij}(k)\delta z_k^j + B_{ij}(k)\delta z_{k+1}^j = 0, \tag{2.24}$$

where we have introduced the convenient shorthand notation

$$A_{ij}(k) = \mathcal{M}_{ji}(z_{k-1}^0, z_k^0), \tag{2.25}$$

$$B_{ij}(k) = \mathcal{M}_{ij}(z_k^0, z_{k+1}^0), \tag{2.26}$$

$$C_{ij}(k) = \frac{\partial L_d}{\partial z_2^i \partial z_2^j}(z_{k-1}^0, z_k^0) + \frac{\partial^2 L_d}{\partial z_1^i \partial z_1^j}(z_k^0, z_{k+1}^0). \tag{2.27}$$

Note that $C_{ij}(k) = C_{ji}(k)$ is a symmetric matrix, while $A_{ij}(k+1) = B_{ji}(k)$ are transposes of one another after a time-step shift.

Let $[Z]$ denote the $m \times m$ matrix whose components are Z_{ij} . Set $X_k = \text{im}[A(k)]$ and $Y_k = \text{im}[B(k)]$, where im denotes the range/column space of the matrix. By proper degeneracy

and hypothesis (G1), $\dim X_k = \dim Y_k = m/2$ and $X_k \cap Y_k = \{0\}$. Therefore $\mathbb{R}^m = X_k \oplus Y_k$ decomposes as a direct sum for each k . Associated with this direct sum is the pair of projection matrices $[\pi_X(k)] : \mathbb{R}^m \rightarrow X_k$ and $[\pi_Y(k)] : \mathbb{R}^m \rightarrow Y_k$. Applying the projection $[\pi_X(k)]$ to the linearized DEL equations gives

$$A_{ij}(k)\delta z_{k-1}^j + (\pi_X(k))_{\bar{i}\bar{j}}C_{\bar{i}\bar{j}}(k)\delta z_k^j = 0, \quad (2.28)$$

while applying the projection $[\pi_Y(k)]$ gives

$$(\pi_Y(k))_{\bar{i}\bar{j}}C_{\bar{i}\bar{j}}(k)\delta z_k^j + B_{ij}(k)\delta z_{k+1}^j = 0. \quad (2.29)$$

In particular, by shifting (2.28) ahead by one time step we obtain the implicit linear relation between δz_k and δz_{k+1} given by

$$0 = A_{ij}(k+1)\delta z_k^j + (\pi_X(k+1))_{\bar{i}\bar{j}}C_{\bar{i}\bar{j}}(k+1)\delta z_{k+1}^j, \quad (2.30)$$

$$0 = (\pi_Y(k))_{\bar{i}\bar{j}}C_{\bar{i}\bar{j}}(k)\delta z_k^j + B_{ij}(k)\delta z_{k+1}^j. \quad (2.31)$$

To complete the proof it is enough to demonstrate that for each δz_k , there exists a unique δz_{k+1} that satisfies (2.30)–(2.31). This is done in the proof of theorem A.1 in Appendix A. \square

REMARK 2.6. *Theorem A.4 in Appendix A uses the above result and the implicit function theorem to prove that DVIs are also one-step methods at the nonlinear level.*

We therefore have the following simple explanation for the absence of parasitic modes in variational integrators derived from properly degenerate discrete phase-space Lagrangians. Because parasitic modes only arise in multi-step schemes, and DVIs are equivalent to single-step schemes by theorem A.4, parasitic modes are not generated by DVIs. Note that Ellison (2016) proves the absence of parasitic modes using less direct arguments, but does not prove that DVIs are generally single-step methods. (The single-step property was observed in examples, however.) Theorems 2.4 and A.4 therefore give a more detailed understanding of the benefits of DVI.

3. Second-order DVI

We now turn to the task of constructing degenerate variational integrators with second-order accuracy. As in Ellison (2016) and Ellison *et al.* (2018), we focus on non-canonical phase-space Lagrangians of the form

$$L(x, y, \dot{x}, \dot{y}) = f_i(x, y)\dot{x}^i - H(x, y), \quad (3.1)$$

where the dimension m is even and $i = 1, \dots, m/2$. Such a phase-space Lagrangian is a special case of the general phase-space Lagrangian in non-canonical variables in (2.11), without terms proportional to \dot{y}_i . The phase-space Lagrangian of the magnetic field line case of § 4.1, with the gauge discussed there, is of this form. The guiding centre Lagrangian of § 4.2 uses a similar gauge and is of this form for the restricted class of magnetic fields discussed in Ellison (2016) and Ellison *et al.* (2018). Burby & Ellison (2017) used toroidal regularization to transform the guiding centre phase-space Lagrangian into this form for general magnetic fields with a single nowhere-vanishing covariant component. We have not investigated integrators for phase-space Lagrangians of the form in (3.1). It is of course possible to transform to canonical variables, a special case of the Lagrangians of the form in (3.1), but the cost of the computations required for this transformation, as well as for a more general transformation producing the form in (3.1), may be prohibitive. We defer further work in this area to a future publication.

3.1. *Composing a first-order scheme with its adjoint?*

A common method for constructing a second-order-accurate integrator starting from a first-order-accurate integrator is to compose it with its adjoint (see Hairer *et al.* 2006), discussed above for the special case of the symplectic Euler scheme. To show what can go wrong with composing a first-order scheme with its adjoint in the context of variational integration, let us introduce the discrete phase-space Lagrangian for a canonical system in one degree of freedom:

$$L_d = p_k \frac{q_{k+1} - q_k}{h} - H(q_k, p_k). \tag{3.2}$$

Except for the change $H(q_{k+1}, p_k) \rightarrow H(q_k, p_k)$, this discretization is identical to that of the symplectic Euler scheme of (2.20) and (2.21a,b). (The reason for this modification of the symplectic Euler Lagrangian is that it has an $O(h)$ term in its preserved two-form, similar to the term in the magnetic field line and guiding centre examples. Ramifications of this $O(h)$ term are discussed below.) Variations with respect to p_k and q_k yield the map

$$q_{k+1} = q_k + hH_2(q_k, p_k), \quad p_{k+1} = p_k - hH_1(q_{k+1}, p_{k+1}). \tag{3.3a,b}$$

The first of these is explicit with respect to q ; the second update is implicit in p and leapfrogged in q , so that this scheme is slightly different from the updating in symplectic Euler.² The adjoint of this scheme follows from $k \leftrightarrow k + 1, h \rightarrow -h$. We find

$$L_d = p_{k+1} \frac{(q_{k+1} - q_k)}{h} - H(q_{k+1}, p_{k+1}), \tag{3.4}$$

leading to

$$p_{k+1} = p_k - hH_1(q_k, p_k), \quad q_{k+1} = q_k + hH_2(q_{k+1}, p_{k+1}). \tag{3.5a,b}$$

This scheme is explicit in p , implicit and leapfrogged in q .

By direct substitution, we find that

$$(1 + hH_{12}(q, p)) dq \wedge dp \tag{3.6}$$

is preserved by this scheme. As discussed in Ellison *et al.* (2018), the preservation of this two-form can also be shown by inspecting the first and last terms in $dS = d \sum_k hL_d$. This is the discrete analogue to the variation of $S = \int_0^T (p\dot{q} - H) dt$, and $\delta S = 0$ gives $\int_0^T [(\dot{q} - H_p) dp + (-\dot{p} - H_q) dq] dt + p dq|_0^T$. Here, the vanishing of the integral provides Hamiltonian equations, and the endpoints, with $d^2S = 0$, imply the conservation of the canonical two-form $\omega = dq \wedge dp$, showing the canonical symplectic property of the equations. Similarly, for a general non-canonical system, the endpoint terms in $S = \int (\vartheta_i(z)\dot{z}^i - H(z)) dt$ show the preservation of $\omega_{ij} dz^i \wedge dz^j$, where $\omega_{ij} = \vartheta_{i,j} - \vartheta_{j,i}$. The preservation of this form shows the non-canonical symplectic nature of the discrete map. This property is consistent with the fact that this scheme and the adjoint symplectic Euler scheme are equivalent under a non-canonical change of variables. Unlike the symplectic Euler scheme, which preserves the canonical two-form $\omega = dq \wedge dp$, this form has $\omega = \omega(h) = (1 + O(h)) dq \wedge dp$. That is, in this scheme the phase-space coordinates (q, p)

²If we define Q_{eh} as updating $q_1 = q_0 + hH_2(q_0, p_0), p_1 = p_0$ and P_{ih} by $p_1 = p_0 - hH_1(q_0, p_1), q_1 = q_0$, we find that the present scheme equals $P_{ih} \circ Q_{eh}$ and the adjoint of the symplectic Euler scheme is $Q_{eh} \circ P_{ih}$, so the present scheme and the adjoint symplectic Euler scheme are equivalent under a non-canonical change of variables, and the former is therefore symplectic under non-canonical variables.

are not canonical. By analogous arguments, or by direct inspection, the adjoint scheme preserves

$$\omega^\dagger(h) = \omega(-h) = (1 - hH_{12}(q, p)) \, dq \wedge dp, \tag{3.7}$$

and because the correction is $O(h)$, the two-forms differ by $O(h)$. From these observations, when $\omega(h) \neq \omega(-h)$ it is not clear how to determine which two-form, if any, is preserved by the composition of the scheme and its adjoint. In § 5.1 we show numerical evidence that such a composed form does not in general preserve any two-form. Also, in Appendix C we show the direct analogy to the symplectic Euler scheme in (2.21a,b) (not (3.3a,b)), and show that its preserved two-form also has $O(h)$ corrections. While we cannot rule out the existence of a first-order-accurate scheme in our class of non-canonical variables (which would allow composition with its adjoint to obtain second-order accuracy), we proceed in the next section to show two separate second-order-accurate schemes for such systems.

3.2. Centred schemes for second-order accuracy

Because composing a first-order DVI with its adjoint does not reliably produce a scheme preserving a symplectic form (although such a scheme is second-order accurate), we are naturally led to consider the problem of proceeding to higher order by identifying improved properly degenerate discrete Lagrangians. We first illustrate for a one-degree-of-freedom case ($m = 2$) in canonical variables, introducing two different staggered, centred schemes to discretize the action $S = \int L dt$ for the phase-space Lagrangian $L(q, p, \dot{q}, \dot{p}) = p\dot{q} - H(q, p)$. The first scheme has

$$L_d(q_k, q_{k+1}, p_{k+1/2}) = p_{k+1/2} \frac{(q_{k+1} - q_k)}{h} - H\left(\frac{q_k + q_{k+1}}{2}, p_{k+1/2}\right). \tag{3.8}$$

This scheme includes a staggered nature of p and a midpoint nature with respect to q . Taking variations with respect to q_k and $p_{k+1/2}$ we find

$$q_{k+1} = q_k + hH_2\left(\frac{q_k + q_{k+1}}{2}, p_{k+1/2}\right), \tag{3.9}$$

$$p_{k+1/2} = p_{k-1/2} - \frac{h}{2}H_1\left(\frac{q_k + q_{k+1}}{2}, p_{k+1/2}\right) - \frac{h}{2}H_1\left(\frac{q_{k-1} + q_k}{2}, p_{k-1/2}\right). \tag{3.10}$$

We call this the midpoint DVI (MDVI) scheme. This scheme is clearly time-centred, leading to second-order accuracy, as we discuss below. This scheme must be advanced implicitly in both variables. It can easily be shown to be properly degenerate, by showing that the rank of the 2×2 discrete Hessian is one, essentially because, as for the symplectic Euler scheme, its adjoint and the schemes of (3.2) and (3.4), the discrete Lagrangian L_d depends on p at only one time level. This degeneracy is in spite of the fact that the scheme appears to be a two-step scheme, connecting q_{k-1} , q_k and q_{k+1} (but only $p_{k-1/2}$ and $p_{k+1/2}$). As in Ellison *et al.* (2018), q_{k-1} can be expressed in terms of q_k and $p_{k-1/2}$ by the implicit equation $q_{k-1} = q_k - hH_2((q_{k-1} + q_k)/2, p_{k-1/2})$ obtained from (3.9), retarded by one step. With this substitution, we define a time-advance map from (3.9) and (3.10) at the current step of the form $(q_k, p_{k-1/2}) \mapsto (q_{k+1}, p_{k+1/2})$, obtaining a one-step method. This property of appearing to be of two steps but showing a single-step nature after some substitutions, as discussed at length in Ellison *et al.* (2018), shows the importance of the discrete Hessian test; indeed, without the assurance of the discrete Hessian test, it would be easy to miss the possibility of this substitution.

The second scheme uses

$$L_d(q_k, q_{k+1}, p_{k+1/2}) = p_{k+1/2} \frac{(q_{k+1} - q_k)}{h} - \frac{1}{2}H(q_k, p_{k+1/2}) - \frac{1}{2}H(q_{k+1}, p_{k+1/2}), \quad (3.11)$$

and variations lead to

$$q_{k+1} = q_k + \frac{h}{2}H_2(q_k, p_{k+1/2}) + \frac{h}{2}H_2(q_{k+1}, p_{k+1/2}), \quad (3.12)$$

$$p_{k+1/2} = p_{k-1/2} - \frac{h}{2}H_1(q_k, p_{k+1/2}) - \frac{h}{2}H_1(q_k, p_{k-1/2}). \quad (3.13)$$

This second scheme is also time-centred, again showing second-order accuracy. This scheme must be advanced implicitly in both variables. For this scheme, the discrete Lagrangian is obtained using a trapezoidal quadrature scheme, hence we call this scheme the trapezoidal DVI (TDVI) scheme. In this scheme, the single-step nature is evident from the DEL equations; of course, the Hessian test confirms the single-step character. Finally, backward error analysis for both the MDVI and the TDVI schemes, with $q_k = q(t = kh)$ and $p_{k+1/2} = p((k + 1/2)h)$, shows second-order accuracy. This derivation follows from a Taylor expansion about $t_k = kh$, e.g. $q_{k+1} = q(t_k) + h\dot{q}(t_k) + (h^2/2)\ddot{q}(t_k) + O(h^3)$, $p_{k+1/2} = p(t_k) + (h/2)\dot{p}(t_k) + (h^2/8)\ddot{p}(t_k) + O(h^3)$; upon substituting in either (3.9) and (3.10) or (3.12) and (3.13) and dividing by h , we find that the $O(h)$ terms all cancel, proving second-order accuracy.

In both schemes, if initial conditions q_0 and p_0 are both given at $t = 0$, the momentum variable needs to be regressed to $p_{-1/2}$ by a processing scheme, which we discuss shortly. Numerical trials using the non-reversible (cf. Appendix B) Hamiltonian $H = (p^2 + q^2)/2 + \alpha qp^3/3$ indicate second-order accuracy and the good long-time behaviour of a scheme with a preserved two-form.

For the non-canonical (but not completely general) phase-space Lagrangians of the form of (3.1), the MDVI scheme has

$$L_d(x_k, y_{k+1/2}, x_{k+1}) = f_i \left(\frac{x_k + x_{k+1}}{2}, y_{k+1/2} \right) \frac{x_{k+1}^i - x_k^i}{h} - H \left(\frac{x_k + x_{k+1}}{2}, y_{k+1/2} \right). \quad (3.14)$$

Again, the centredness suggests, and backward error analysis indeed shows, second-order accuracy. Also, the staggered-grid discrete Lagrangian (3.14) preserves a symplectic form determined, as before, by examining the endpoint terms in the variation of the action. Again, the discrete Hessian has half-rank because L_d depends on y at only one time level.

The DEL equations stemming from (3.14) are given by

$$\begin{aligned} & \frac{1}{2} (f_{i,j} (k + \frac{1}{2}) (x_{k+1}^i - x_k^i) + f_{i,j} (k - \frac{1}{2}) (x_k^i - x_{k-1}^i)) \\ & - (f_j (k + \frac{1}{2}) - f_j (k - \frac{1}{2})) - \frac{h}{2} (H_j (k + \frac{1}{2}) + H_j (k - \frac{1}{2})) = 0, \end{aligned} \quad (3.15a)$$

$$f_{i,\alpha} (k + \frac{1}{2}) (x_{k+1}^i - x_k^i) - hH_{,\alpha} (k + \frac{1}{2}) = 0, \quad (3.15b)$$

where $i, j \in \{1, \dots, m/2\}$, m is even, while $\alpha \in \{m/2 + 1, \dots, m\}$ and $(k + 1/2)$ is shorthand for the arguments $((x_{k+1} + x_k)/2, y_{k+1/2})$ and $(k - 1/2)$ is analogous. Like the canonical MDVI, the apparent dependence on x_{k-1} may be eliminated in favour of a function depending on $(x_k, y_{k-1/2})$. In practice one has access to x_{k-1} due to the state at prior iterations (except for on the first step) and the stored value may be used instead of directly solving for x_{k-1} at each new iteration.

The TDVI scheme for this class of non-canonical cases arises from

$$L_d(x_k, y_{k+1/2}, x_{k+1}) = \frac{1}{2} (f_i(x_k, y_{k+1/2}) + f_i(x_{k+1}, y_{k+1/2})) \frac{x_{k+1}^i - x_k^i}{h} - \frac{1}{2} (H(x_k, y_{k+1/2}) + H(x_{k+1}, y_{k+1/2})). \tag{3.16}$$

Again, the time-centred property leads to second-order accuracy, and its discrete Hessian shows that it is properly degenerate, again because L_d depends on y at only one time level.

Performing variations with respect to $x_k, y_{k+1/2}$ yields the TDVI scheme:

$$\begin{aligned} & \frac{1}{2} (f_{i,j}(x_k, y_{k+1/2}) (x_{k+1}^i - x_k^i) + f_{i,j}(x_k, y_{k-1/2}) (x_k^i - x_{k-1}^i)) \\ & - \frac{1}{2} (f_j(x_{k+1}, y_{k+1/2}) + f_j(x_k, y_{k+1/2}) - f_j(x_k, y_{k-1/2}) - f_j(x_{k-1}, y_{k-1/2})) \\ & - \frac{1}{2} (H_{,j}(x_k, y_{k+1/2}) + H_{,j}(x_{k+1}, y_{k+1/2})) = 0, \end{aligned} \tag{3.17a}$$

$$\begin{aligned} & \frac{1}{2} (f_{i,\alpha}(x_k, y_{k+1/2}) + f_{i,\alpha}(x_{k+1}, y_{k+1/2})) (x_{k+1}^i - x_k^i) \\ & - \frac{h}{2} (H_{,\alpha}(x_k, y_{k+1/2}) + H_{,\alpha}(x_{k+1}, y_{k+1/2})) = 0. \end{aligned} \tag{3.17b}$$

In contrast to the canonical setting, the TDVI scheme introduces dependence on x_{k-1} in general. Of course, this dependence is superficial, and can be eliminated in favour of a function of $(x_k, y_{k-1/2})$ as previously discussed. Relative to the MDVI scheme, the TDVI scheme requires more function evaluations (i.e. additional evaluations of f and H) in the update rule; this has the potential to increase the computational expense of TDVI relative to MDVI.

Of course, initial conditions for the time-marching scheme defined by (3.14) or (3.16) will be supplied at an integer time step instead of directly on the staggered grid. In order to transform integer-time-step initial conditions (x_0, y_0) to staggered-grid initial conditions $(x_0, y_{-1/2})$, it is sufficient to advance y backward in time by a half-step $h/2$ using any scheme that is accurate to first (or higher) order. This transformation is encoded in the mapping $\varphi_{-h/2} : (x_0, y_0) \mapsto (x_0, y_{-1/2})$. Similarly, at the end of a simulation, the staggered-grid data $(x_N, y_{N-1/2})$ must be collocated to the integer-grid data (x_N, y_N) . The natural way to do this is simply to apply the inverse of $\varphi_{-h/2}$, i.e. set $(x_N, y_N) = \varphi_{-h/2}^{-1}(x_N, y_{N-1/2})$, and second-order accuracy is preserved. These two conditions are special cases of enforcing the relationship

$$\varphi_{-h/2}(x_k, y_k) = (x_k, y_{k-1/2}) \tag{3.18}$$

for all k . Indeed, for displaying results during a computation, e.g. where k equals a multiple of a fundamental period M , typically with $M > 1$, these results will retain second-order accuracy if the points are collocated according to (3.18). Furthermore, this process preserves the symplectic nature of the scheme. This is because applying the processing scheme at the initial step and its inverse at the final step essentially defines the step in terms of a non-symplectic change of variables. Writing the processing step in (3.18) as Φ and the time step as T_h , this new scheme is written as $\tilde{T}_h = \Phi^{-1} \circ T_h \circ \Phi$. For a symplectic scheme in canonical variables, this defines an equivalent map in non-canonical variables; for the case under consideration here, where the symplectic scheme is already in non-canonical variables, this transformation just specifies an $O(h)$ transformation to other non-canonical variables. Regarding invariants, the original symplectic scheme may not preserve invariants exactly, but typically such a scheme will preserve invariants such as energy approximately for sufficiently small h , preventing, for example, secular or

exponential growth in the energy. The same properties hold for the modified scheme \tilde{T}_h . In Blanes *et al.* (2004), the idea of increasing the order of a low-order scheme using a map and its inverse as pre- and post-processors is explored in greater detail. It would be interesting to determine if even higher-order DVIs may be derived using more elaborate processing than the more-or-less obvious processors described here.

4. Magnetic field line and guiding centre examples

In this section we apply the discretizations of (3.14) and (3.16) to the Lagrangians for the magnetic field line problem and the guiding centre system, both described in Ellison *et al.* (2018) and Ellison (2016).

4.1. *Magnetic field line*

For the problem of tracing magnetic field lines, we take the action to be equal to the flux:

$$\Phi = \int L\left(x, \frac{dx}{dt}\right) dt = \int \mathbf{A}(x) \cdot \frac{dx}{dt} dt, \tag{4.1}$$

where \mathbf{A} is the magnetic vector potential. The invariance with respect to reparameterization of time (see Ellison *et al.* 2018) in (4.1) is consistent with its Euler–Lagrange equation, namely $(dx/dt) \times \mathbf{B} = 0$, where $\mathbf{B} = \nabla \times \mathbf{A}$ is the magnetic field: this equation determines the direction of the flow but not the speed. This time invariance is dealt with by parameterizing the field line trajectory in terms of one of the coordinates (without loss of generality, we choose x^3) instead of ‘time’ t :

$$\Phi = \int L(x^1, x^2, x^3, dx^2/dx^3, x^3) dx^3 = \int (A_2(x^1, x^2, x^3) dx^2/dx^3 + A_3(x^1, x^2, x^3)) dx^3. \tag{4.2}$$

This Lagrangian is of the form in (3.1), with $f = [0 \ A_2]$ and $H = -A_3$; the Euler–Lagrange equations for (4.2) are

$$dx^1/dx^3 = (A_{3,2} - A_{2,3})/A_{2,1} = B^1/B^3, \tag{4.3a}$$

$$dx^2/dx^3 = -A_{3,1}/A_{2,1} = B^2/B^3. \tag{4.3b}$$

The MDVI then follows from (3.15), which when expressed in terms of the magnetic vector potential becomes

$$A_{2,1}(k + \frac{1}{2})(x_{k+1}^2 - x_k^2) + hA_{3,1}(k + \frac{1}{2}) = 0, \tag{4.4a}$$

$$\begin{aligned} &\frac{1}{2}(A_{2,2}(k + \frac{1}{2})(x_{k+1}^2 - x_k^2) + A_{2,2}(k - \frac{1}{2})(x_k^2 - x_{k-1}^2)) \\ &- (A_2(k + \frac{1}{2}) - A_2(k - \frac{1}{2})) + \frac{h}{2}(A_{3,2}(k + \frac{1}{2}) + A_{3,2}(k - \frac{1}{2})) = 0, \end{aligned} \tag{4.4b}$$

where $(k + \frac{1}{2})$ denotes evaluation at $(x_{k+1/2}^1, (x_{k+1}^2 + x_k^2)/2, x_{k+1/2}^3)$. Similar to the process in Ellison (2016) and Ellison *et al.* (2018), these equations are made into an explicitly single-step process by writing the first with $k \rightarrow k - 1$, solving for x_{k-1} and substituting for x_{k-1} in the remaining equations without the $k \rightarrow k - 1$ incrementation.

Similarly, the TDVI algorithm for the magnetic field line problem follows from (3.17):

$$\begin{aligned} & \frac{1}{2} (A_{2,1}(x_{k+1/2}^1, x_{k+1}^2, x_{k+1/2}^3) + A_{2,1}(x_{k+1/2}^1, x_k^2, x_{k+1/2}^3)) (x_{k+1}^2 - x_k^2) \\ & + \frac{h}{2} (A_{3,1}(x_{k+1/2}^1, x_k^2, x_{k+1/2}^3) + A_{3,1}(x_{k+1/2}^1, x_{k+1}^2, x_{k+1/2}^3)) = 0, \end{aligned} \tag{4.5a}$$

$$\begin{aligned} & \frac{1}{2} (A_{2,2}(x_{k+1/2}^1, x_k^2, x_{k+1/2}^3)(x_{k+1}^2 - x_k^2) + A_{2,2}(x_{k-1/2}^1, x_k^2, x_{k-1/2}^3)(x_k^2 - x_{k-1}^2)) \\ & + \frac{1}{2} (A_2(x_{k-1/2}^1, x_{k-1}^2, x_{k-1/2}^3) + A_2(x_{k-1/2}^1, x_k^2, x_{k-1/2}^3)) \\ & - \frac{1}{2} (A_2(x_{k+1/2}^1, x_k^2, x_{k+1/2}^3) + A_2(x_{k+1/2}^1, x_{k+1}^2, x_{k+1/2}^3)) \\ & + \frac{h}{2} (A_{3,2}(x_{k-1/2}^1, x_k^2, x_{k-1/2}^3) + A_{3,2}(x_{k+1/2}^1, x_k^2, x_{k+1/2}^3)) = 0. \end{aligned} \tag{4.5b}$$

Numerical tests of the MDVI and TDVI schemes are presented in § 5.

4.2. Guiding centre equations

We treat the example of the guiding centre equations based on discretizing the toroidally regularized (compare with Burby & Ellison 2017) phase-space Lagrangian

$$\begin{aligned} L(z, \dot{z}) &= A_2(\mathbf{x})\dot{x}^2 + (A_3(\mathbf{x}) + u)\dot{x}^3 - H_{\text{gc}} \\ &\equiv A_2^\dagger(\mathbf{x})\dot{x}^2 + A_3^\dagger(\mathbf{x}, u)\dot{x}^3 - H_{\text{gc}}, \end{aligned} \tag{4.6}$$

where $H_{\text{gc}} = \frac{1}{2}(B^3/B)^2 u^2 + \mu B - \frac{1}{2}E_\perp^2/B^2 - u\mathbf{b} \cdot (\mathbf{E} \times \nabla x^3)/B + \varphi$ is the guiding centre Hamiltonian, in toroidally regularized non-canonical variables, u is the toroidally regularized parallel velocity, μ is the magnetic moment, B is the magnitude of the magnetic field and φ is the scalar potential for the electric field $\mathbf{E} = -\nabla\varphi$. For simplicity, we neglect time dependence in the electromagnetic fields. Toroidal regularization requires $|B^3| > 0$ everywhere. In toroidal geometries relevant to magnetic fusion energy, possible choices for coordinates include $(x^1, x^2, x^3) = (R, Z, \phi)$ or $(x^1, x^2, x^3) = (r, \theta, \phi)$, where (R, Z, ϕ) denote cylindrical coordinates and (r, θ, ϕ) denote toroidal coordinates. As in the magnetic field example, the gauge condition $A_1 = 0$ was imposed to lead to proper degeneracy. (The original form introduced in Ellison *et al.* (2018) and Ellison (2016) required that the covariant component B_1 of the magnetic field also vanish, but Burby & Ellison (2017) subsequently relaxed that condition.)

The second-order-accurate DVIs follow by establishing the correspondence with (3.1); in this case, $f = [0 \quad A_2^\dagger \quad A_3^\dagger \quad 0]$ and $H = H_{\text{gc}}$. We discretize this phase-space Lagrangian by the MDVI and TDVI schemes for second-order accuracy. For the MDVI discretization of the guiding centre system, we use the discrete Lagrangian

$$\begin{aligned} L_d(r_{k+1/2}, \theta_k, \theta_{k+1}, \phi_k, \phi_{k+1}, u_{k+1/2}) &= A_\theta(k + \frac{1}{2})(\theta_{k+1} - \theta_k) \\ &+ (A_\phi(k + \frac{1}{2}) + u_{k+1/2})(\phi_{k+1} - \phi_k) - hH(k + \frac{1}{2}). \end{aligned} \tag{4.7}$$

The MDVI scheme, following from (3.15), is

$$A_{\theta,r}(k + \frac{1}{2})(\theta_{k+1} - \theta_k) + A_{\phi,r}(k + \frac{1}{2})(\phi_{k+1} - \phi_k) - hH_{,r}(k + \frac{1}{2}) = 0, \tag{4.8a}$$

$$\begin{aligned} & \frac{1}{2}A_{\theta,\theta}(k - \frac{1}{2})(\theta_k - \theta_{k-1}) + \frac{1}{2}A_{\theta,\theta}(k + \frac{1}{2})(\theta_{k+1} - \theta_k) \\ & + \frac{1}{2}A_{\phi,\theta}(k - \frac{1}{2})(\phi_k - \phi_{k-1}) + \frac{1}{2}A_{\phi,\theta}(k + \frac{1}{2})(\phi_{k+1} - \phi_k) \\ & + A_{\theta}(k - \frac{1}{2}) + u_{k-1/2} - A_{\theta}(k + \frac{1}{2}) - u_{k+1/2} - \frac{h}{2}(H_{,\theta}(k - \frac{1}{2}) + H_{,\theta}(k + \frac{1}{2})) = 0, \end{aligned} \tag{4.8b}$$

$$\begin{aligned} & \frac{1}{2}A_{\theta,\phi}(k - \frac{1}{2})(\theta_k - \theta_{k-1}) + \frac{1}{2}A_{\theta,\phi}(k + \frac{1}{2})(\theta_{k+1} - \theta_k) \\ & + \frac{1}{2}A_{\phi,\phi}(k - \frac{1}{2})(\phi_k - \phi_{k-1}) + \frac{1}{2}A_{\phi,\phi}(k + \frac{1}{2})(\phi_{k+1} - \phi_k) \\ & + A_{\phi}(k - \frac{1}{2}) + u_{k-1/2} - A_{\phi}(k + \frac{1}{2}) - u_{k+1/2} - \frac{h}{2}(H_{,\phi}(k - \frac{1}{2}) + H_{,\phi}(k + \frac{1}{2})) = 0, \end{aligned} \tag{4.8c}$$

$$\phi_{k+1} - \phi_k - \frac{h}{2}H_{,u}(k + \frac{1}{2}) = 0, \tag{4.8d}$$

where $(k + 1/2)$ refers to $(r_{k+1/2}, (\theta_k + \theta_{k+1})/2, (\phi_k + \phi_{k+1})/2)$ and similarly for $(k - 1/2)$. The incrementation–substitution process for obtaining a scheme that is explicitly single step, introduced in Ellison (2016) and Ellison *et al.* (2018) and used in the magnetic field lines case above, is also used here.

For the TDVI scheme, we use the discrete Lagrangian in (3.16), namely

$$\begin{aligned} & L_d(r_{k+1/2}, \theta_k, \theta_{k+1}, \phi_k, \phi_{k+1}, u_{k+1/2}) \\ & = \frac{1}{2}(A_{\theta}(r_{k+1/2}, \theta_k, \phi_k) + A_{\theta}(r_{k+1/2}, \theta_{k+1}, \phi_{k+1}))(\theta_{k+1} - \theta_k) \\ & + \frac{1}{2}(A_{\phi}(r_{k+1/2}, \theta_k, \phi_k) + A_{\phi}(r_{k+1/2}, \theta_{k+1}, \phi_{k+1}) + 2u_{k+1/2})(\phi_{k+1} - \phi_k) \\ & - \frac{h}{2}(H(r_{k+1/2}, \theta_k, \phi_k, u_{k+1/2}) + H(r_{k+1/2}, \theta_{k+1}, \phi_{k+1}, u_{k+1/2})). \end{aligned} \tag{4.9}$$

The guiding centre TDVI update equations follow from straightforward variations of this discrete Lagrangian (or from (3.17)) and are omitted here for brevity.

5. Numerical tests

In this section we present numerical tests showing that the MDVI and TDVI schemes provide all the advantages of variational schemes and are second-order-accurate.

5.1. Failure of composing with the adjoint

In § 3.1 we noted an example of a scheme with a preserved two-form with $\omega(h) \neq \omega(-h)$, suggesting complications if the scheme and its adjoint are composed in an attempt to obtain second-order accuracy. In this section we give a concrete example for which the composed scheme appears not to have a preserved two-form. We consider the non-reversible Hamiltonian

$$H = \frac{p^2 + q^2}{2} + \frac{\alpha qp^3}{3}. \tag{5.1}$$

Numerical tests show that both of the schemes of (3.3a,b) and (3.5a,b) have good long-time behaviour for this Hamiltonian. The orbits of the two first-order schemes have bounded behaviour but, with first-order accuracy, have noticeable oscillations in the value of H between bounds. The composed scheme, which has smaller oscillations in H , otherwise performs poorly: the points that should stay near the constant H surfaces spiral

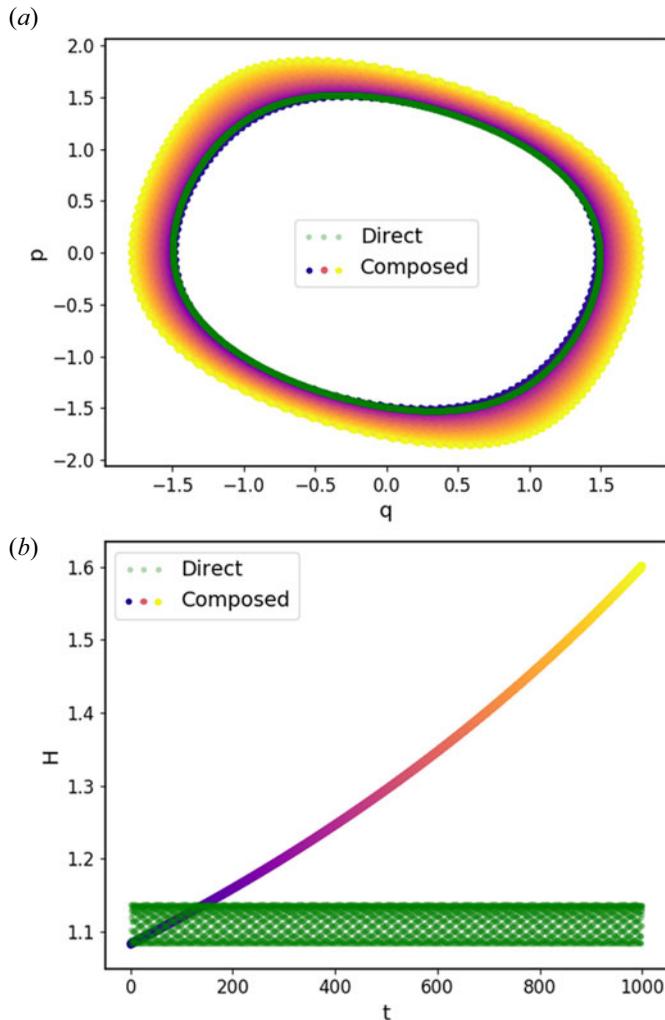


FIGURE 1. Phase portrait (a) and energy versus time (b) for two different methods of integrating the non-reversible Hamiltonian $H = (p^2 + q^2)/2 + \alpha qp^3/3$. The ‘direct’ scheme of (3.3a,b), and shown in green, is first-order-accurate and preserves a discrete symplectic structure that depends on the numerical time-step size h , with similar results for the adjoint scheme in (3.5a,b). Composing the direct method with its adjoint achieves second-order accuracy by virtue of centring in time, giving smaller oscillations in H . However, the numerical results indicate that the composed scheme leads to growth in H and therefore does not preserve a two-form.

out, as shown in figure 1. The behaviour of H in time shows exponential increase, with $\gamma = O(h^2)$. Appendix B shows an example for which composition does appear to give useful results, but this particular example is a reversible system, and this reversibility by itself appears to be responsible for the positive results.

5.2. Magnetic field line

To test the proposed algorithms in a magnetic configuration representative of those of interest to the magnetic fusion community, we use the simple analytic expression for an

axisymmetric, toroidal magnetic field presented in Qin, Guan & Tang (2009):

$$A(r, \theta, \phi) = \frac{B_0 R_0}{\cos^2 \theta} \left(r \cos \theta - R_0 \log \left(1 + \frac{r \cos \theta}{R_0} \right) \right) \nabla \theta - \frac{B_0 r^2}{2q_0} \nabla \phi, \quad (5.2)$$

where B_0 is a magnetic field amplitude, R_0 is the major radius and $q_0 = \sqrt{2}$ is the on-axis safety factor. The variables (x^1, x^2, x^3) of (4.2) are replaced by simple toroidal coordinates (r, θ, ϕ) , and ϕ takes the place of the time variable, as discussed in §4.1.

First, we demonstrate numerically that the MDVI and TDVI achieve the anticipated second-order accuracy. Next, we demonstrate that the proposed algorithms exhibit the expected qualitative behaviour of symplectic integrators. To the axisymmetric magnetic field of (5.2) we add a perturbation of the form

$$A(r, \theta, \phi) = A_0(r, \theta, \phi) - \frac{B_0 r^2}{2q_0} \sum_i \delta_i \sin(m_i \theta - n_i \phi) \nabla \phi, \quad (5.3)$$

where A_0 is given by (5.2). We choose two perturbative harmonics, $m_0 = 3, n_0 = 2$ and $m_1 = 7, n_1 = 5$, with amplitudes $\delta_0 = \delta_1 = 10^{-4}$. These perturbations lead to magnetic islands at the resonant magnetic surfaces, and small stochastic field line regions in the $(m_1, n_1) = (3, 2)$ and $(m_2, n_2) = (7, 5)$ resonant regions.

We test the order of accuracy by varying the step $\Delta\phi$ in factors of two across a range of 2^5 and comparing with a fourth-order Runge–Kutta scheme with an extremely small value of $\Delta\phi$. See figure 2. We compare the MDVI and TDVI schemes with each other and with the first-order variant described in Ellison *et al.* (2018) in which the three variables (r, θ, ϕ) are collocated, i.e. not staggered. We integrate over a large number of toroidal transits for this comparison. Second-order accuracy is confirmed for the MDVI and TDVI schemes, as is first-order accuracy for the non-staggered scheme. The MDVI and TDVI schemes exhibit relatively similar accuracy, with the MDVI scheme being more accurate for this particular example.

The Poincaré surface of section at $\phi = 0$ is shown in figure 3. Figures 3(a) and 3(b) show the second-order Runge–Kutta (RK2) scheme, with $\Delta\phi = 0.1$ and $\Delta\phi = 5 \times 10^{-4}$, respectively. Figures 3(c) and 3(d) show the MDVI scheme for the same two values of $\Delta\phi$. The RK2 results in figure 3(a) show blurriness of the Kolmogorov–Arnold–Moser (KAM) surfaces, falsely indicating a higher degree of magnetic stochasticity. The results in figure 3(b) are greatly improved. The two MDVI cases in figure 3(c,d), for very different steps $\Delta\phi$, look almost identical, showing very good preservation of KAM tori. The TDVI scheme leads to results essentially indistinguishable from those of the MDVI scheme for comparable time steps.

5.3. Guiding centre

In this section we show numerical results for the guiding centre example, with time-independent potentials and scalar potential $\varphi = 0$. That is, physically, the electric field is zero. See figure 4, confirming second-order accuracy in $h = \Delta t$ for the MDVI and TDVI schemes. Interestingly, the error in the MDVI scheme is a factor of about 3 larger, in contrast with the results shown in figure 2, where the MDVI and TDVI results are reversed. As for the magnetic field line results, this difference is due to the very long run times.

6. Non-uniform time stepping

We first review the variational form of the extended phase-space method for a Hamiltonian system in canonical variables. This method allows us to prescribe variable

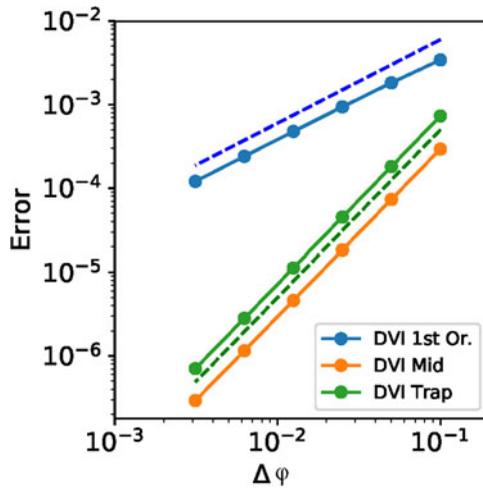


FIGURE 2. Comparison of the order of accuracy of the MDVI scheme, the TDVI scheme and the collocated first-order-accuratescheme of Ellison *et al.* (2018) for the magnetic field line problem, over a range of 2^5 in $\Delta\phi$. The dashed curves show $\text{Error} \sim h$ (blue) and $\text{Error} \sim h^2$ (green) for comparison.

time steps. We show an example of discretizing this action by the symplectic Euler scheme, and show that it is a single-step method, i.e. a degenerate variational integrator. We proceed to formulate an extended phase-space method for systems having the form of non-canonical variables prescribed in (3.1), giving examples of magnetic field line integration and the guiding centre equations. For the most straightforward first-order-accurate scheme with non-uniform time steps, we exhibit an analogue of the modified symplectic Euler schemes of (3.3a,b) and (3.5a,b). We show that this method is a DVI scheme, connecting only two time levels, according to the discrete Hessian method as well as by a direct substitution. We also discuss extending these non-uniform time-step methods to second-order accuracy, as well as using adaptive time-step control based on an error estimator.

6.1. Canonical systems

We first review the extended phase-space method in canonical variables, in one degree of freedom for transparency. The time stepping is defined by a time-step density $\rho(q, p)$, with $\rho(q, p) dt = d\zeta$. We prescribe uniform steps in the time-like variable ζ , with $\rho\Delta t \approx \Delta\zeta = h = \text{const}$. We extend the action $S = \int (p dq/dt - H) dt$ for the Hamiltonian $H = H(q, p)$ to

$$S_e = \int \left[p \frac{dq}{d\zeta} - \frac{H(q, p)}{\rho(q, p)} + \pi \left(\frac{dw}{d\zeta} - \frac{1}{\rho(q, p)} \right) \right] d\zeta. \tag{6.1}$$

Here we have written time as the dependent variable w , and the Lagrange multiplier π enforces the time-step condition as a constraint. We rewrite this as

$$S_e = \int \left[p \frac{dq}{d\zeta} + \pi \frac{dw}{d\zeta} - \frac{H(q, p) + \pi}{\rho(q, p)} \right] d\zeta, \tag{6.2}$$

leading to the extended phase-space Hamiltonian $K(q, p, \pi) = (H(q, p) + \pi)/\rho(q, p)$, with an added canonically conjugate pair, $(q, p) \rightarrow (q, p, w, \pi)$. It is clear that, since K

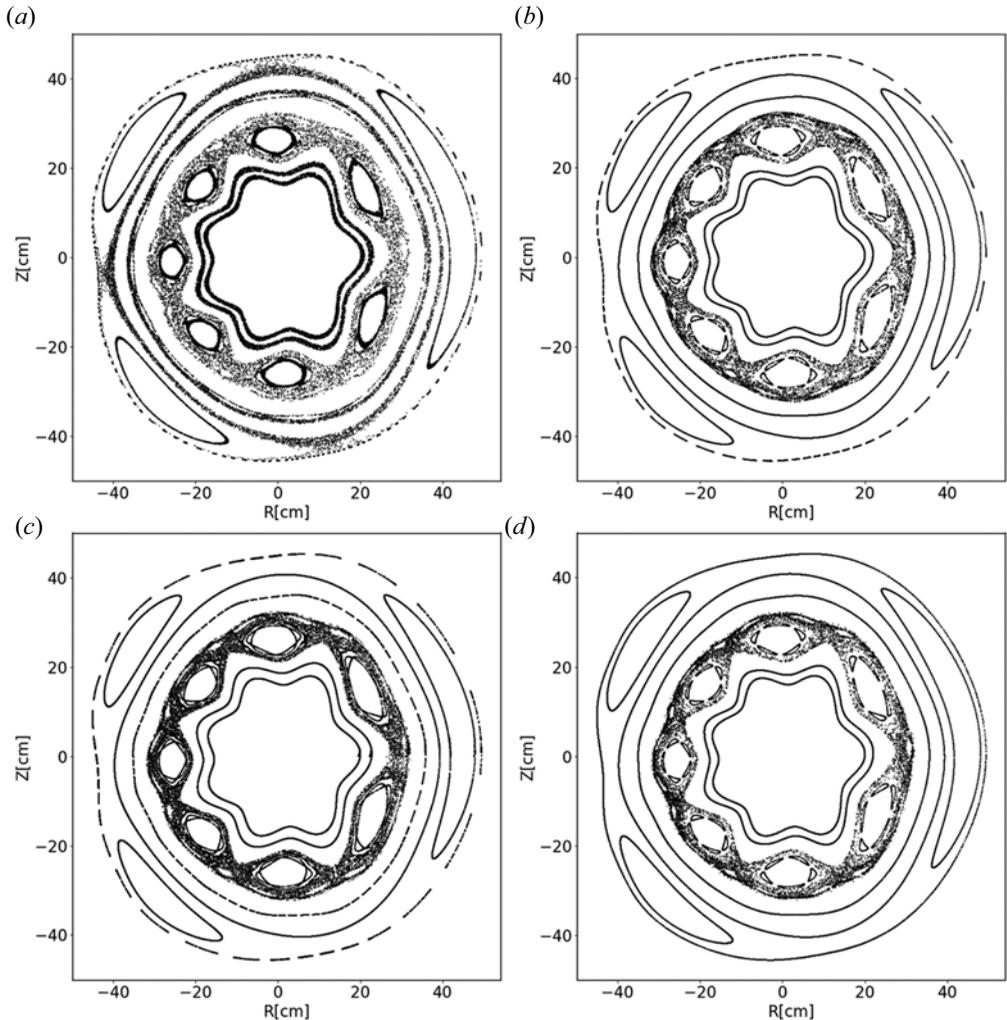


FIGURE 3. Poincaré section $\phi = 0$ for the magnetic field line problem. For comparison we show results from (a) the RK2 scheme with $\Delta\phi = 0.1$, (b) the RK2 scheme with $\Delta\phi = 0.005$, (c) the MDVI scheme with $\Delta\phi = 0.1$ and (d) the MDVI scheme with $\Delta\phi = 0.005$. In all cases we have $\phi_{\text{final}} = 3 \times 10^5$. The $(m, n) = (3, 2)$ and $(7, 5)$ island chains are evident, and in the RK2 scheme in (a) the KAM surfaces are blurred.

does not depend on ζ explicitly, K is exactly conserved. That is, if we set $\pi = -H$ initially, then $H + \pi$ remains exactly zero. Thus, the equations for (q, p, w, π) , using $H + \pi = 0$, are

$$\frac{dq}{d\zeta} = \frac{H_2(q, p, w)}{\rho(q, p)}, \quad \frac{dp}{d\zeta} = -\frac{H_1(q, p, w)}{\rho(q, p)}, \quad (6.3a,b)$$

$$\frac{dw}{d\zeta} = \frac{1}{\rho(q, p)}, \quad \frac{d\pi}{d\zeta} = 0. \quad (6.4a,b)$$

Note that the imposed time-step requirement is satisfied. The extension to a non-autonomous system, with $H(q, p, t)$, is straightforward.

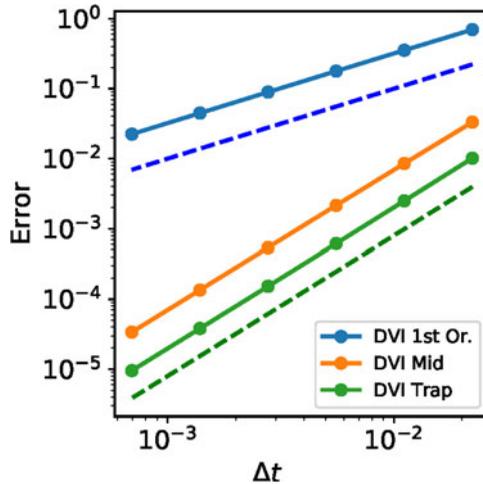


FIGURE 4. Error in the second-order MDVI and TDVI schemes and the first-order collocated scheme of Ellison *et al.* (2018), compared with results obtained with a fourth-order Runge–Kutta method with extremely small time step $h = \Delta t$. Results are shown over almost two orders of magnitude in $\Delta t = h$. First-order and second-order reference lines, in blue dashed and green dashed, respectively, are shown for comparison. Second-order accuracy for the MDVI and TDVI schemes is confirmed.

For discrete integration, it is important to keep the terms (not shown) derived from $H + \pi$ in (6.2): this quantity is not exactly zero for the discrete equations and if this quantity is set exactly equal to zero, the symplectic nature is lost, as discussed in Richardson & Finn (2011).

To illustrate, we discretize S_e as in the symplectic Euler scheme, here for two degrees of freedom, for uniform stepping in ζ , $\Delta\zeta = h$. We have

$$S_{e1} = \sum_{k=0}^{N-1} [p_k(q_{k+1} - q_k) + \pi_k(w_{k+1} - w_k)] - \sum_{k=0}^{N-1} \left[h \frac{H(q_{k+1}, p_k, w_{k+1}) + \pi_k}{\rho(q_{k+1}, p_k)} \right]. \tag{6.5}$$

The symplectic Euler scheme derived from (6.5) preserves the canonical two-degree-of-freedom two-form

$$\omega = dq \wedge dp + dw \wedge d\pi, \tag{6.6}$$

by inspection or by taking the endpoint values of dS_{e1} . This scheme is a single-step scheme, shown either by inspection or by computing the discrete Hessian.

We now consider the special class of systems in non-canonical variables with action of the form $\int (f_i \dot{x}_i - H(x, y)) dt$, as in (3.1). With time-step condition $\rho(x, y) dt = d\zeta$ and $t \rightarrow w$ we can write the analogue to (6.1) and (6.2):

$$S = \int \left[f_i(x, y) \frac{dx^i}{d\zeta} - \frac{H(x, y)}{\rho(x, y)} + \pi \left(\frac{dw}{d\zeta} - \frac{1}{\rho(x, y)} \right) \right] d\zeta, \\ = \int \left[f_i(x, y) \frac{dx^i}{d\zeta} + \pi \frac{dw}{d\zeta} - \frac{H(x, y, w) + \pi}{\rho(x, y)} \right] d\zeta. \tag{6.7}$$

Again, it is evident in the first form that π is a Lagrange multiplier enforcing the time-step condition. We can apply uniform stepping in the new time-like variable ζ , with $\Delta\zeta = h = \text{const.}$, to give the required non-uniform time stepping. In (6.7) a term involving the canonical pair (w, π) , namely $\pi dw/d\zeta$, is added to $f_i dx^i/d\zeta$. Furthermore, the new Hamiltonian is $K = (H + \pi)/\rho$, and the resulting action is of the same restricted non-canonical class of (3.1). Therefore any discretization that can be applied to the action with uniform time step can be applied to this non-canonical extended phase-space version.

6.2. *Magnetic field line integration*

For the magnetic field line integration problem, we use a gauge with $A_1 = 0$, as in § 4.1. For non-uniform time stepping, we first go back to considering x^3 to be a coordinate and put the action in the form in (6.7), leading to

$$\Phi_e = \int \left[A_2(\mathbf{x}) \frac{dx^2}{d\zeta} + \frac{A_3(\mathbf{x})}{\rho(\mathbf{x})} + \pi \left(\frac{dx^3}{d\zeta} - \frac{1}{\rho(\mathbf{x})} \right) \right] d\zeta, \tag{6.8}$$

where π is a Lagrange multiplier enforcing the time-step restriction and $\mathbf{x} = (x^1, x^2, x^3)$. This can again be put into the form

$$\Phi_e = \int \left[A_2(\mathbf{x}) \frac{dx^2}{d\zeta} + \pi \frac{dx^3}{d\zeta} - \frac{-A_3(\mathbf{x}) + \pi}{\rho(\mathbf{x})} \right] d\zeta. \tag{6.9}$$

This action is of the form in (6.7), with two-degree-of-freedom Hamiltonian equal to $(-A_3 + \pi)/\rho$.

We discretize this in a manner similar to the modified (adjoint) symplectic Euler scheme described in (3.3a,b) by forming $\Phi_{e1} = \sum_k hL_d(\mathbf{x}_k, \pi_k, \mathbf{x}_{k+1}, \pi_{k+1})$, or

$$\begin{aligned} \Phi_{e1} = & \sum_k \left[A_2(\mathbf{x}_{k+1})(x_{k+1}^2 - x_k^2) + \pi_{k+1}(x_{k+1}^3 - x_k^3) \right] \\ & + \sum_k \left[h \frac{A_3(\mathbf{x}_{k+1}) - \pi_{k+1}}{\rho(\mathbf{x}_{k+1})} \right]. \end{aligned} \tag{6.10}$$

Compared with (3.3a,b), $p_{k+1} \rightarrow A_2(\mathbf{x}_{k+1})$ and $H(q_{k+1}, p_{k+1}) \rightarrow A_3(\mathbf{x}_{k+1})$ with more direct substitutions for x_k^3 and π_k .

The DEL equations from $d\Phi_{e1} = 0$ lead to the fourth-order system

$$\begin{aligned} & A_{2,1}(\mathbf{x}_k)(x_k^2 - x_{k-1}^2) + h \frac{A_{3,1}(\mathbf{x}_k)}{\rho(\mathbf{x}_k)} \\ & + h [A_3(\mathbf{x}_k) - \pi_k] \frac{\partial}{\partial x_k^1} \frac{1}{\rho(\mathbf{x}_k)} = 0, \end{aligned} \tag{6.11}$$

$$\begin{aligned} & A_{2,2}(\mathbf{x}_k)(x_k^2 - x_{k-1}^2) + A_2(\mathbf{x}_k) - A_2(\mathbf{x}_{k+1}) \\ & + h \frac{A_{3,2}(\mathbf{x}_k)}{\rho(\mathbf{x}_k)} + h [A_3(\mathbf{x}_k) - \pi_k] \frac{\partial}{\partial x_k^2} \frac{1}{\rho(\mathbf{x}_k)} = 0, \end{aligned} \tag{6.12}$$

$$\begin{aligned}
 & A_{2,3}(\mathbf{x}_k)(x_k^3 - x_{k-1}^3) + \pi_k - \pi_{k+1} + h \frac{A_{3,3}(\mathbf{x}_k)}{\rho(\mathbf{x}_k)} \\
 & + h [A_3(\mathbf{x}_k) - \pi_k] \frac{\partial}{\partial x_k^3} \frac{1}{\rho(\mathbf{x}_k)} = 0,
 \end{aligned} \tag{6.13}$$

$$x_k^3 - x_{k-1}^3 - \frac{h}{\rho(\mathbf{x}_k)} = 0. \tag{6.14}$$

As in earlier sections, the discrete Hessian test indicates that these four equations define a single-step scheme, and therefore a mapping on a four-dimensional space. The process of incrementing and substitution leads to equations for which this single-step property is explicit. For the uniform time-stepping case of § 4.1 the assumption $\rho = \rho(\mathbf{x})$ means that the time step depends on x^3 , which takes the place of time. This suggests the possibility that complications such as parametric instabilities related to having a time-step density explicitly dependent on time might arise, as in Richardson & Finn (2011). In the case treated in this subsection, on the other hand, ζ rather than z is the independent (time-like) variable.

As occurred in Ellison *et al.* (2018), (6.12) and (6.13) appear to involve indices $k - 1$, k , $k + 1$ and therefore appear to be two-step equations, suggesting that (6.11)–(6.14) are difference equations of order higher than 4. However, the discrete Hessian can be shown to have rank 4, consistent with a first-order system in (\mathbf{x}, π) ; indeed similar substitutions to those of Ellison *et al.* (2018) lead to a fourth-order system, i.e. a single-step method. That is, writing (6.12), (6.11) and (6.13) in the compact form

$$A_{2,1}(\mathbf{x}_k)(x_k^2 - x_{k-1}^2) + hP(\mathbf{x}_k, \pi_k) = 0, \tag{6.15}$$

$$A_{2,2}(\mathbf{x}_k)(x_k^2 - x_{k-1}^2) + A_2(\mathbf{x}_k) - A_2(\mathbf{x}_{k+1}) + hQ(\mathbf{x}_k, \pi_k) = 0, \tag{6.16}$$

$$A_{2,3}(\mathbf{x}_k)(x_k^2 - x_{k-1}^2) + \pi_k - \pi_{k+1} + hR(\mathbf{x}_k, \pi_k) = 0, \tag{6.17}$$

we find $x_k^2 - x_{k-1}^2$ from (6.15) and substitute into (6.16) and (6.17). When the indices are incremented $k \rightarrow k + 1$ in (6.15) and (6.14), we find

$$\left. \begin{aligned}
 & A_{2,1}(\mathbf{x}_{k+1})(x_{k+1}^2 - x_k^2) + hP(\mathbf{x}_{k+1}, \pi_{k+1}) = 0, \\
 & - hA_{2,2}(\mathbf{x}_k) \frac{P(\mathbf{x}_k, \pi_k)}{A_{2,1}(\mathbf{x}_k)} + A_2(\mathbf{x}_k) - A_2(\mathbf{x}_{k+1}) + hQ(\mathbf{x}_k, \pi_k) = 0, \\
 & - hA_{2,3}(\mathbf{x}_k) \frac{P(\mathbf{x}_k, \pi_k)}{A_{2,1}(\mathbf{x}_k)} + \pi_k - \pi_{k+1} + hR(\mathbf{x}_k, \pi_k) = 0, \\
 & x_{k+1}^3 - x_k^3 - \frac{h}{\rho(\mathbf{x}_{k+1})} = 0.
 \end{aligned} \right\} \tag{6.18}$$

Because only indices k and $k + 1$ are involved, the single-step property predicted by the rank of the discrete Hessian is evident. Note the solvability condition $A_{2,1} = B^3 \neq 0$, necessary for z to parameterize the length along the field line. We have performed numerical tests with non-uniform time stepping, showing comparable properties with uniform time stepping of the field line equations (§ 5.2). We defer computations based on a time step adapted to an error estimator obtained by backward error analysis, as in Richardson & Finn (2011), to a future publication.

6.3. *Guiding centre equations*

Guided by the results for the magnetic field line equations, it was shown in Ellison *et al.* (2018) that if the term $A_1^\dagger \dot{x}^1$ is zero, substitutions can be made to lead to a single-step scheme. Because the requirement $A_1^\dagger(\mathbf{x}, u) = A_1(\mathbf{x}) + ub_1(\mathbf{x}) = 0$ must hold for arbitrary values of u , it requires both the gauge condition $A_1 = 0$ and the physical condition $b_1 = 0$. In Ellison *et al.* (2018), numerical tests were performed for axisymmetric fields and for coordinates such that the covariant component b_1 is zero. In Burby & Ellison (2017), it was shown that it is possible to find coordinates such that this condition is satisfied for arbitrary magnetic fields, provided one component, e.g. the toroidal component, does not change sign.

For applying non-uniform time steps to the guiding centre equations, we again introduce a new time-like variable ζ such that $\rho(\mathbf{x}, u) dt = d\zeta$ for the time-step density $\rho(\mathbf{x}, u)$. Then we make the substitution, with $w = t$ and

$$S_{gc} = \int \left[A_2^\dagger(\mathbf{x}) \frac{dx^2}{dt} + A_3^\dagger(\mathbf{x}) \frac{dx^3}{dt} - H_{gc}(\mathbf{x}, u) \right] dt \tag{6.19}$$

going to

$$S_{gc} = \int \left[A_2^\dagger(\mathbf{x}) \frac{dx^2}{d\zeta} + A_3^\dagger(\mathbf{x}) \frac{dx^3}{d\zeta} - \frac{H_{gc}(\mathbf{x}, u)}{\rho(\mathbf{x}, u)} \right] d\zeta + \int \left[\frac{dw}{d\zeta} - \frac{1}{\rho(\mathbf{x}, u)} \right] d\zeta. \tag{6.20}$$

Again this can be put in the form

$$S_{gc} = \int \left[A_2^\dagger(\mathbf{x}) \frac{dx^2}{d\zeta} + A_3^\dagger(\mathbf{x}) \frac{dx^3}{d\zeta} + \pi \frac{dw}{d\zeta} - K(\mathbf{x}, u, \pi) \right] d\zeta, \tag{6.21}$$

where $K(\mathbf{x}, u, p_0) = (H_{gc}(\mathbf{x}, u) + \pi)/\rho(\mathbf{x}, u)$. This is an action in the form of (6.7) on the extended phase space $(\mathbf{x}, u) \rightarrow (\mathbf{x}, u, w, \pi)$. We discretize this action in a manner similar to that in (6.10), namely

$$S_{gc1} = \sum_k hL_d(\mathbf{x}_k, u_k, w_k, \pi_k, \mathbf{x}_{k+1}, u_{k+1}, w_{k+1}, \pi_{k+1}), \tag{6.22}$$

with

$$hL_d = A_2^\dagger(\mathbf{x}_{k+1}, u_{k+1})(x_{k+1}^2 - x_k^2) + A_3^\dagger(\mathbf{x}_{k+1}, u_{k+1})(x_{k+1}^3 - x_k^3) + \pi_{k+1}(w_{k+1} - w_k) - hK(\mathbf{x}_{k+1}, u_{k+1}, \pi_{k+1}), \tag{6.23}$$

where

$$K(\mathbf{x}_{k+1}, u_{k+1}, \pi_{k+1}) = \frac{H_{gc}(\mathbf{x}_{k+1}, u_{k+1}) + \pi_{k+1}}{\rho(\mathbf{x}_{k+1}, u_{k+1})}. \tag{6.24}$$

The DEL equations are

$$A_{2,1}^\dagger(\mathbf{x}_k, u_k)(x_k^2 - x_{k-1}^2) + A_{3,1}^\dagger(\mathbf{x}_k, u_k)(x_k^3 - x_{k-1}^3), \tag{6.25}$$

$$-hK_1(\mathbf{x}_k, u_k, \pi_k) = 0, \tag{6.26}$$

$$A_{2,2}^\dagger(\mathbf{x}_k, u_k)(x_k^2 - x_{k-1}^2) + A_{3,2}^\dagger(\mathbf{x}_k, u_k)(x_k^3 - x_{k-1}^3) + A_2^\dagger(\mathbf{x}_k, u_k) - A_2^\dagger(\mathbf{x}_{k+1}, u_{k+1}) - hK_2(\mathbf{x}_k, u_k, \pi_k) = 0, \tag{6.27}$$

$$A_{2,3}^\dagger(\mathbf{x}_k, u_k)(x_k^2 - x_{k-1}^2) + A_{3,3}^\dagger(\mathbf{x}_k, u_k)(x_k^3 - x_{k-1}^3) + A_3^\dagger(\mathbf{x}_k, u_k) - A_3^\dagger(\mathbf{x}_{k+1}, u_{k+1}) - hK_3(\mathbf{x}_k, u_k, \pi_k) = 0, \tag{6.28}$$

$$b_{2,k}(x_k^2 - x_{k-1}^2) + b_{3,k}(x_k^3 - x_{k-1}^3) - hK_u(\mathbf{x}_k, u_k, \pi_k) = 0, \tag{6.29}$$

$$\pi_k - \pi_{k+1} = 0, \tag{6.30}$$

$$w_k - w_{k-1} = h \frac{\partial}{\partial p_{w,k}} K(\mathbf{x}_k, u_k, \pi_k) = \frac{h}{\rho(\mathbf{x}_k, u_k)}. \tag{6.31}$$

The assumed time independence of the fields leads to the simple form in (6.30). Similar to the uniform time-step case of Ellison *et al.* (2018), the discrete Hessian has rank six, consistent with a first-order system in (x, y, z, u, w, π) . We start by taking (6.26) and (6.29), written as

$$\begin{bmatrix} A_{2,1}^\dagger(\mathbf{x}_k, u_k) & A_{3,1}^\dagger(\mathbf{x}_k, u_k) \\ b_{2,k} & b_{3,k} \end{bmatrix} \begin{bmatrix} x_k^2 - x_{k-1}^2 \\ x_k^3 - x_{k-1}^3 \end{bmatrix} = \begin{bmatrix} hK_1(\mathbf{x}_k, u_k, w_k, \pi_k) \\ hK_u(\mathbf{x}_k, u_k, w_k, \pi_k) \end{bmatrix}. \tag{6.32}$$

Solving for $x_k^2 - x_{k-1}^2$ and $x_k^3 - x_{k-1}^3$, which are written in terms of quantities with index k , we substitute these into (6.27) and (6.28), incrementing $k \rightarrow k + 1$ in (6.27), (6.28) and (6.31). The resulting equations involve time steps labelled with only k and $k + 1$. That is, consistent with the discrete Hessian condition, the scheme is a single-step scheme, a DVI, and parasitic modes cannot occur. As for the magnetic field line case discussed in the last subsection, we have obtained similar numerical results to uniform time-stepping results of § 5.3, and defer adaptive integration to a future publication.

6.4. Extensions for higher accuracy

From the formulation in the last two sections, it is clear from (6.9) and (6.21) that the modification to prescribe non-uniform time stepping leads to an addition to the phase-space Lagrangian of a term $\pi dw/d\zeta$ or $\pi dz/d\zeta$ and a modification to the Hamiltonian $H \rightarrow (H + \pi)/\rho$, and these terms can be discretized in exactly the same manner as in the uniform time-step case. This means that the modifications in this section can be applied to any discretization of the phase-space Lagrangian that leads to a DVI scheme. Therefore, it should be straightforward to construct a non-uniform time-step scheme for either of the second-order-accurate DVI methods of § 3.

It is also clear that such discretizations can be applied to any time-step density ρ , so that it should be straightforward to use an optimum density ρ based on an error estimator, to minimize the integrated error over an orbit for the scheme at hand, as done in Richardson & Finn (2011) and Finn (2015). Therefore, it is possible to combine the formulations of this paper to give an adaptive second-order-accurate variational integrator. We leave further details to a future publication.

7. Summary and discussion

In Ellison *et al.* (2018) and Ellison (2016), the concept of proper degeneracy for a discrete time-stepping scheme for a degenerate variational system was introduced. In these works, the focus was on systems governed by a phase-space Lagrangian, which produces a system of first-order differential equations, the Hamiltonian equations, in canonical or non-canonical variables. This concept relates to a discretization that preserves the first-order nature of the Hamiltonian equations on phase space, i.e. is a single-step rather than a multi-step scheme. Multi-step schemes are to be avoided in variational systems because they can possess parasitic modes that can grow unphysically, as discussed in Ellison *et al.* (2018). For some examples, the single-step property can be determined by inspection simply. But it is in fact common to have a system that appears to have a multi-step nature, but can be reduced to a form where the single-step property is evident. However, finding the right substitutions is not always so straightforward. In this reference, a method of addressing this single-step versus multi-step issue in terms of the rank of the discrete Hessian was developed. In Ellison *et al.* (2018) and Ellison (2016), schemes that preserve this single-step nature were called DVI schemes.

The schemes developed in Ellison *et al.* (2018) and Ellison (2016) are all first-order-accurate. One aim of this paper is to develop second-order-accurate DVIs. A commonly used method of developing a second-order-accurate scheme from a first-order variational scheme is a special case of a composition method described for instance in Hairer *et al.* (2006). This involves composing the first-order scheme Φ_h with its adjoint $\Phi_h^\dagger = \Phi_{-h}^{-1}$, and this method works well for discretizations that preserve the two-form ω_0 of the original ordinary differential equation system; this form is independent of h . However, for other schemes the discrete equations preserve a two-form that depends on the time step, $\omega = \omega(h)$. The adjoint of such a scheme preserves $\omega(-h)$ and it is not obvious whether the composed map preserves a two-form at all if $\omega(h) \neq \omega(-h)$. In this paper we consider an example of a simple autonomous Hamiltonian system in canonical variables, i.e. preserving the two-form $\omega_0 = dq \wedge dp$, and a discretization of its phase-space Lagrangian. This scheme preserves another form $\omega(h) = \omega_0 + O(h)$, so that $\omega(h) \neq \omega(-h)$. Numerically, we find that, for some Hamiltonians, the orbits of the composed scheme spiral out with increasing t , the growth rate of the energy behaving like $\gamma = O(h^2)$, showing that composing with the adjoint does not lead to a scheme with a preserved two-form in general, and therefore does not possess the advantageous properties of variational (symplectic) integration.

In the place of the composition method, we have constructed two centred schemes, involving a processing scheme to advance some of the variables to the half time step, and centring the other variables either in a midpoint or a trapezoidal manner. We call these schemes the MDVI scheme and the TDVI scheme. We have shown these schemes to be second-order-accurate by a backward error analysis and derived the properly degenerate property by computing the rank of the discrete Hessian (as well as by inspection). We have also applied the MDVI and TDVI schemes to two systems of importance to plasma physics, namely the magnetic field line system and the guiding centre system. Both of these systems are in a restricted class of non-canonical variables. The numerical results show the anticipated positive properties, namely the benefits of degenerate variational integration, the lack of parasitic modes and second-order accuracy.

The second aim of this paper relates to using non-uniform time steps. This method has been developed for Hamiltonian systems in canonical variables in Hairer *et al.* (2006). In this paper we show how to write a variational principle in extended phase space for systems with this class of non-canonical variables. Further, using an error estimator, it is

possible to make the time step adaptive, by minimizing the total error along an orbit, as in Richardson & Finn (2011).

We have first reviewed the extended phase-space action principle for one-degree-of-freedom Hamiltonian systems in canonical variables with action $S = \int (p\dot{q} - H) dt$, allowing variable time steps. For canonical variables, this method involves a discretization of the action with a constraint related to the variation of the time stepping, producing a canonical symplectic integrator in the extended phase space $(q, p) \rightarrow (q, p, w, \pi)$, where the extra canonical pair are time and its canonical conjugate. The extension to non-canonical variables applies to the restricted class of systems discussed earlier, with variables (x, y) and an action of the form $\int (f_i(x, y)\dot{x}_i - H(x, y)) dt$. The two well-known examples of Hamiltonian systems in non-canonical variables of importance to plasma physics, namely the integration of magnetic field lines and the guiding centre equations, can be obtained via an action of this restricted non-canonical form. We have shown how to write an extended phase-space action for this class of non-canonical variables with non-uniform time stepping. We have developed discretizations that lead again to DVI schemes. The generalization of the extended phase-space method, to non-canonical variables and to the second-order-accurate DVI schemes introduced in this paper, is straightforward. This capacity for non-uniform time stepping leads immediately to the capability for adaptive time stepping, as described for symplectic integrators in Richardson & Finn (2011).

Acknowledgements

Editor H. Qin thanks the referees for their advice in evaluating this article.

Funding

This material is based upon work supported by the National Science Foundation under grant no. 1440140, while the authors J.B. and J.M.F. were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the autumn semester of 2018. Research presented in this article was supported by the Los Alamos National Laboratory LDRD programme under project number 20180756PRD4. A portion of this work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Declaration of interests

The authors report no conflict of interest.

Appendix A. Detailed proofs of the DVI single-step property

THEOREM A.1 (linearized single-step property). *Let $L_d(z_1, z_2)$ be a properly degenerate discrete Lagrangian, and introduce the $m \times m$ matrices $[A(z_1, z_2)]$, $[B(z_1, z_2)]$, $[C(z_1, z_2, z_3)]$ with components*

$$A_{ij}(z_1, z_2) = \mathcal{M}_{ji}(z_1, z_2), \quad (\text{A1})$$

$$B_{ij}(z_1, z_2) = \mathcal{M}_{ij}(z_1, z_2), \quad (\text{A2})$$

$$C_{ij}(z_1, z_2, z_3) = \frac{\partial L_d}{\partial z_2^i \partial z_2^j}(z_1, z_2) + \frac{\partial^2 L_d}{\partial z_1^i \partial z_1^j}(z_2, z_3). \quad (\text{A3})$$

Under the following transversality assumptions:

(G1) *for each $(z_1, z_2), (z'_1, z'_2) \in Z \times Z$ near the diagonal, $\text{im}[A(z_1, z_2)] \cap \text{im}[B(z'_1, z'_2)] = \{0\}$ and*

(G2) for each $(z_1, z_2), (z'_1, z'_2) \in Z \times Z$ near the diagonal in $Z \times Z$ and $(z''_1, z''_2, z''_3) \in Z \times Z \times Z$ near the diagonal in $Z \times Z \times Z$,

$$[C(z''_1, z''_2, z''_3)](\ker[B(z'_1, z'_2)]) \tag{A4}$$

is a graph over $\text{im}[A(z_1, z_2)]$,

the DEL equations linearized about a trajectory $k \mapsto z_k^0$ whose neighbouring samples satisfy $|z_{k+1}^0 - z_k^0| < \delta$ for some small $\delta > 0$ independent of k are equivalent to a single-step method.

REMARK A.2. If L_d is some properly degenerate discrete Lagrangian satisfying (G1) and (G2), then all properly degenerate discrete Lagrangians in a neighbourhood of L_d will satisfy (G1) and (G2). In practice this observation greatly simplifies the task of verifying (G1) and (G2) because the $h \rightarrow 0$ limit of a properly degenerate discrete Lagrangian is usually quite simple to analyse.

REMARK A.3. The condition $|z_{k+1}^0 - z_k^0| < \delta$ is generally satisfied provided that the time step h in a variational integrator is sufficiently small.

Proof. The proof picks up at the end of the proof sketch of theorem 2.4.

To that end, consider the linear map $\Phi : \mathbb{R}^N \rightarrow X_{k+1} \times Y_k$ given by

$$\Phi(\delta z) = ([\pi_X(k + 1)][C(k + 1)]\delta z, [B(k)]\delta z). \tag{A5}$$

By (2.30)–(2.31) it is enough to show that the kernel of Φ is trivial. To see that this is so, first note that by transversality assumption (G2) the linear space $[C(k + 1)](\ker[B(k)])$ must be of the form

$$\begin{aligned} & [C(k + 1)](\ker[B(k)]) \\ &= \{w_X + \Gamma(w_X) \mid w_X \in \text{im}[A(k + 1)]\}, \end{aligned} \tag{A6}$$

where $\Gamma : \text{im}[A(k + 1)] \rightarrow \text{im}[B(k + 1)]$ is a linear map. In particular, $\dim [C(k + 1)](\ker[B(k)]) = \dim \text{im}[A(k + 1)] = m/2$, which by the rank-nullity theorem implies that $[C(k + 1)] \mid \ker[B(k)]$ is invertible onto its image. Now suppose that $\Phi(\delta z) = 0$. This implies that δz must be in the kernel of $[B(k)]$. Therefore $[C(k + 1)]\delta z \in [C(k + 1)](\ker[B(k)])$ must have the form

$$[C(k + 1)]\delta z = w_X + \Gamma(w_X), \tag{A7}$$

for a unique $w_X \in \text{im}[A(k + 1)]$. But because $[\pi_X(k + 1)][C(k + 1)]\delta z = 0$, it must be the case that

$$\begin{aligned} 0 &= [\pi_X(k + 1)][C(k + 1)]\delta z \\ &= [\pi_X(k + 1)](w_X + \Gamma(w_X)) \\ &= w_X, \end{aligned} \tag{A8}$$

which implies that $\delta z = 0$. □

THEOREM A.4 (nonlinear single-step property). *Let $L_d(z_1, z_2)$ be a properly degenerate discrete Lagrangian that satisfies the transversality conditions (G1) and (G2) given in the*

statement of theorem 2.4. Solutions $k \mapsto z_k$ of the DEL equations near a given solution $k \mapsto z_k^0$ that satisfies $|z_{k+1}^0 - z_k^0| < \delta$ for some sufficiently small k -independent $\delta > 0$ are generated by a single-step method $\varphi : Z \rightarrow Z$. In other words,

$$z_{k+1} = \varphi(z_k), \tag{A9}$$

for each k .

Proof. First we introduce some convenient notation. Let $\alpha : Z \times Z \rightarrow \mathbb{R}^m$ and $\beta : Z \times Z \rightarrow \mathbb{R}^m$ be the functions defined by

$$\alpha_i(z_1, z_2) = \frac{\partial L_d}{\partial z_2^i}(z_1, z_2), \tag{A10}$$

$$\beta_i(z_1, z_2) = \frac{\partial L_d}{\partial z_1^i}(z_1, z_2). \tag{A11}$$

For each $\tilde{z} \in Z$, also define the related functions $\alpha_{\tilde{z}} : Z \rightarrow \mathbb{R}^m$ and $\beta_{\tilde{z}} : Z \rightarrow \mathbb{R}^m$ according to

$$\alpha_{\tilde{z}}(z) = \alpha(z, \tilde{z}), \tag{A12}$$

$$\beta_{\tilde{z}}(z) = \beta(\tilde{z}, z). \tag{A13}$$

Finally, introduce the DEL operator $E : Z \times Z \times Z \rightarrow \mathbb{R}^m$ given by

$$E(z_1, z_2, z_3) = \alpha(z_1, z_2) + \beta(z_2, z_3), \tag{A14}$$

and the associated function $E_z : Z \times Z \rightarrow \mathbb{R}^m$ given by

$$E_z(z_1, z_2) = E(z_1, z, z_2). \tag{A15}$$

In terms of these notations, the DEL equations may be written in several equivalent ways:

$$0 = E(z_{k-1}, z_k, z_{k+1}) \tag{A16}$$

$$= E_{z_k}(z_{k-1}, z_{k+1}) \tag{A17}$$

$$= \alpha_{z_k}(z_{k-1}) + \beta_{z_k}(z_{k+1}) \tag{A18}$$

$$= \alpha(z_{k-1}, z_k) + \beta(z_k, z_{k+1}). \tag{A19}$$

By the constant-rank theorem, for each z the level sets of either α_z or β_z are $m/2$ -dimensional submanifolds that foliate Z . We will call a level set of α_z an α -leaf, and a level set of β_z a β -leaf. We may choose mutually disjoint neighbourhoods U_k of each z_k^0 such that the intersection of either the α -foliation or the β -foliation with U_k is diffeomorphic to $\mathbb{R}^{m/2} \times \mathbb{R}^{m/2}$. In particular we may define smooth maps

$$\gamma_z^\alpha : \cup_k U_k \rightarrow \mathbb{R}^{m/2}, \tag{A20}$$

$$\gamma_z^\beta : \cup_k U_k \rightarrow \mathbb{R}^{m/2}, \tag{A21}$$

such that the restriction of γ_z^α (γ_z^β) to U_k is a quotient map for the α -foliation (β -foliation) intersected with U_k . Moreover, we may assume without loss of generality that $\gamma_z^\alpha(z_k^0) = 0 = \gamma_z^\beta(z_k^0)$, independent of z .

Because, for each z , α_z is constant along the α -leaves and β_z is constant along the β -leaves, the DEL operator $E(z_1, z, z_3)$ only depends on the α -leaf that contains z_1 and the β -leaf that contains z_3 . Therefore for each $z_k \in Z$ there must be a function $\varepsilon_{z_k} : \mathbb{R}^{m/2} \times \mathbb{R}^{m/2} \rightarrow \mathbb{R}^m$ defined by the relation

$$E_{z_k}(z_{k-1}, z_{k+1}) = \varepsilon_{z_k}(X, Y), \tag{A22}$$

$$X = \gamma_{z_k}^\alpha(z_{k-1}), \tag{A23}$$

$$Y = \gamma_{z_k}^\beta(z_{k+1}), \tag{A24}$$

for $z_k \in U_k$.

By hypothesis (G1) the derivative $D\varepsilon_{z_k}(X, Y) : \mathbb{R}^{m/2} \times \mathbb{R}^{m/2} \rightarrow \mathbb{R}^m$ is invertible for each (X, Y) and $z_k \in U_k$. Therefore by the inverse function theorem the function ε_{z_k} restricts to a diffeomorphism on a neighbourhood of $(X_{k-1}(z_k), Y_{k+1}(z_k)) = (\gamma_{z_k}^\alpha(z_{k-1}^0), \gamma_{z_k}^\beta(z_{k+1}^0))$. At the price of possibly shrinking the U_k , we may assume that this neighbourhood is all of $\mathbb{R}^{m/2} \times \mathbb{R}^{m/2}$.

Let $(\hat{X}_{z_k}, \hat{Y}_{z_k}) = (\varepsilon_{z_k})^{-1}$ be the inverse of the diffeomorphism $\varepsilon_{z_k} : \mathbb{R}^{m/2} \times \mathbb{R}^{m/2} \rightarrow \mathbb{R}^m$. The DEL equations $E(z_{k-1}, z_k, z_{k+1}) = 0$ imply

$$0 = \varepsilon_{z_k}(\gamma_{z_k}^\alpha(z_{k-1}), \gamma_{z_k}^\beta(z_{k+1})), \tag{A25}$$

which is equivalent to

$$\gamma_{z_k}^\alpha(z_{k-1}) = \hat{X}_{z_k}(0), \tag{A26}$$

$$\gamma_{z_k}^\beta(z_{k+1}) = \hat{Y}_{z_k}(0). \tag{A27}$$

In particular, shifting (A26) gives the m equations for the m unknowns z_{k+1} :

$$F^\alpha(z_k, z_{k+1}) = \gamma_{z_{k+1}}^\alpha(z_k) - \hat{X}_{z_{k+1}}(0) = 0, \tag{A28}$$

$$F^\beta(z_k, z_{k+1}) = \gamma_{z_k}^\beta(z_{k+1}) - \hat{Y}_{z_k}(0) = 0. \tag{A29}$$

The proof will therefore be complete if we can show that the mapping $z_{k+1} \mapsto (F^\alpha(z_k, z_{k+1}), F^\beta(z_k, z_{k+1}))$ is a diffeomorphism for fixed z_k in a neighbourhood of z_{k+1}^0 .

To that end, note that by the implicit function theorem it is enough to show that the linear map

$$\begin{aligned} \Phi : \delta z_{k+1} \\ \mapsto (D_{z_{k+1}} F^\alpha(z_k^0, z_{k+1}^0)[\delta z_{k+1}], D_{z_{k+1}} F^\beta(z_k^0, z_{k+1}^0)[\delta z_k]) \end{aligned} \tag{A30}$$

has a trivial kernel. Demonstrating that this is so amounts to reproducing the proof of theorem 2.4. The summary is the following.

Suppose that δz_{k+1} is in the kernel. Because $\gamma_{z_{k+1}}^\alpha(z_k^0) = 0$ for each z_{k+1} , δz_{k+1} must satisfy

$$0 = \delta \hat{X} = \left. \frac{d}{d\epsilon} \right|_0 \hat{X}_{z_{k+1}^0 + \epsilon \delta z_{k+1}}(0), \tag{A31}$$

$$0 = \delta \gamma^\beta = \left. \frac{d}{d\epsilon} \right|_0 \gamma_{z_k^0}^\beta(z_{k+1}^0 + \epsilon \delta z_{k+1}). \tag{A32}$$

The second equation (A32) will be satisfied if and only if δz_{k+1} tangent to the β -leaf passing through z_k^0 . This means

$$B_{ij}\delta z_{k+1}^j = 0, \tag{A33}$$

or δz_{k+1} is in the kernel of the matrix $[B]$ defined in (2.26). Moving now to (A31), note that because $\varepsilon_{z_{k+1}}(\hat{X}_{z_{k+1}}(0), \hat{Y}_{z_{k+1}}(0)) = 0$ by definition, differentiating in z_{k+1} at z_{k+1}^0 gives

$$\begin{aligned} 0 &= D_X \varepsilon_{z_{k+1}}(0, 0)[\delta \hat{X}] + D_Y \varepsilon_{z_{k+1}}(0, 0)[\delta \hat{Y}] \\ &+ \frac{d}{d\varepsilon} \Big|_0 \varepsilon_{z_{k+1} + \varepsilon \delta z_{k+1}}(0, 0), \end{aligned} \tag{A34}$$

where we have used $(\hat{X}_{z_{k+1}}(0), \hat{Y}_{z_{k+1}}(0)) = (0, 0)$ by our normalization convention for $\gamma^\alpha, \gamma^\beta$, and we have introduced

$$\delta \hat{Y} = \frac{d}{d\varepsilon} \Big|_0 \hat{Y}_{z_{k+1} + \varepsilon \delta z_{k+1}}(0). \tag{A35}$$

Each of the derivatives in (A34) may be expressed in terms of derivatives of E by implicitly differentiating (A22), which leads to

$$[A]\delta X + [C]\delta z_{k+1} + [B]\delta Y = 0, \tag{A36}$$

where δX is any vector that satisfies

$$D\gamma_{z_{k+1}}^\alpha(z_k^0)[\delta X] = \delta \hat{X}, \tag{A37}$$

δY is any vector that satisfies

$$D\gamma_{z_{k+1}}^\beta(z_{k+2}^0)[\delta Y] = \delta \hat{Y}, \tag{A38}$$

and the matrices $[A], [B], [C]$ are defined in (2.25)–(2.27). Now using $\delta \hat{X} = 0$, (A37) implies that δX must be in the kernel of $[A]$. Therefore if we apply the projection matrix $[\pi_X]$ guaranteed by transversality assumption (G1) to (A36), we obtain

$$[\pi_X][C]\delta z_{k+1} = 0. \tag{A39}$$

But because δz_{k+1} is in the kernel of $[B]$, transversality assumption (G2) implies that $\delta z_{k+1} = 0$. □

Appendix B. Reversibility: a warning

We first consider the one-degree-of-freedom Hamiltonian

$$H = \frac{p^2 + q^2}{2} + \frac{\alpha qp^2}{2}. \tag{B1}$$

Applying either the scheme in (3.3a,b) or its adjoint in (3.5a,b), we find, of course, first-order accuracy but also good long-time properties, the latter because of the preservation of the two-forms in (3.6) and (3.7). If we compose the two schemes, we also find good long-time properties, and with second-order accuracy. However, these favourable properties are traced not to the preservation of a two-form but to the

reversibility of the Hamiltonian in (B1): the symmetry $R : (q, p) \mapsto (q, -p)$ leaves H invariant, and the fixed points of this symmetry are $p = 0$. This reversibility is inherited by the exact time- h map Φ_h , i.e. Φ_h satisfies

$$R \circ \Phi_h = \Phi_h^{-1} \circ R. \tag{B2}$$

If a discrete scheme T_h also satisfies this map reversibility, it should have the favourable properties due to reversibility (see Finn 2015). In fact, neither T_h nor T_h^\dagger (where adjoint is defined as $T_h^\dagger = T_{-h}^{-1}$; see Hairer *et al.* (2006)) satisfy this map reversibility property. However, T_h does satisfy the related property weak reversibility (defined previously in Finn (2015)),

$$R \circ T_h = T_{-h} \circ R, \tag{B3}$$

and similarly for T_h^\dagger . From this property it follows that the composed scheme $\Psi_h = T_h \circ T_h^\dagger$ (or $T_h^\dagger \circ T_h$) is also weakly reversible, and because it is self-adjoint, it is also reversible. This map reversibility appears to be responsible for the observed good long-time behaviour.

As discussed in Richardson & Finn (2011) and Finn (2015), it can be misleading to evaluate a scheme by testing it on a reversible Hamiltonian system, because good results might be obtained solely due to the reversibility property and not from any property inherited from the variational nature.

The Hamiltonian $H = (p^2 + q^2)/2 + \alpha qp^3/3$ considered in § 5.1 has another symmetry $(q, p) \mapsto (-q, -p)$, but this symmetry preserves the point $(q, p) = (0, 0)$ rather than a line ($p = 0$), and such a symmetry does not endow any special properties, so we do not consider this Hamiltonian to be reversible (see Richardson & Finn 2011; Finn 2015). And indeed, the results in § 5.1 show that the orbits spiral out, showing the lack of a preserved two-form.

Appendix C. Analogue of the symplectic Euler scheme for the magnetic field line problem

Here, we consider the most direct analogue to the symplectic Euler scheme for canonical variables, applied to the class of non-canonical systems of § 3. We specialize to the magnetic field line problem for concreteness, and have

$$L_d(x_k^1, x_k^2, x_{k+1}^2) = A_2(x_k^1, x_{k+1}^2)(x_{k+1}^2 - x_k^2) + hA_3(x_k^1, x_{k+1}^2). \tag{C1}$$

If this system has a preserved two-form with $\omega(-h) = \omega(h)$, it can be composed with its adjoint to preserve ω and obtain second-order accuracy. Note that L_d depends on x^1 at only one time level and therefore, as noted in § 2, its discrete Hessian has rank 1. This shows that the system is indeed properly degenerate, and is a DVI.

Its preserved two-form is found simply by looking at the endpoint terms in dS for $k = 0, 1$:

$$S = A_2(x_0^1, x_1^2)(x_1^2 - x_0^2) + hA_3(x_0^1, x_1^2), \tag{C2}$$

from which we find

$$\begin{aligned} dS = & [A_{2,1}(x_0^1, x_1^2)(x_1^2 - x_0^2) + hA_{3,1}(x_0^1, x_1^2)] dx_0^1 \\ & + [-A_2(x_0^1, x_1^2)] dx_0^2 + [A_{2,1}(x_1^1, x_2^2)(x_2^2 - x_1^2) + hA_{3,1}(x_1^1, x_2^2)] dx_1^1 \\ & + [A_{2,2}(x_0^1, x_1^2)(x_1^2 - x_0^2) + hA_{3,2}(x_0^1, x_1^2) + A_2(x_0^1, x_1^2) - A_2(x_1^1, x_2^2)] dx_1^2 \\ & + [A_{2,2}(x_1^1, x_2^2)(x_2^2 - x_1^2) + hA_{3,2}(x_1^1, x_2^2) + A_2(x_1^1, x_2^2) - A_2(x_2^1, x_3^2)] dx_2^2 \\ & + [A_2(x_2^1, x_3^2)] dx_2^2, \end{aligned} \quad (C3)$$

with the last term subtracted in the dx_2^2 term. Because of satisfying the DEL equations, all the terms except for the endpoint terms

$$dS = A_2(x_2^1, x_3^2) dx_2^2 - A_2(x_0^1, x_1^2) dx_0^2 \quad (C4)$$

vanish, and $d^2S = 0$ leads to the preservation of the two-form is composed

$$\omega = dA_2(x_0^1, x_1^2) \wedge dx_0^2. \quad (C5)$$

Upon substituting

$$x_1^2 = x_0^2 + hu(x_0^1, x_1^2), \quad u(x_0^1, x_1^2) = -\frac{hA_{3,1}(x_0^1, x_1^2)}{A_{2,1}(x_0^1, x_1^2)}, \quad (C6a,b)$$

equal to $hB_2(x_0^1, x_1^2)/B_3(x_0^1, x_1^2)$ and from the dx_0^1 term in dS , we find the preserved two-form ω equals

$$(A_{2,1}(x_0^1, x_1^2) + hA_{2,2}(x_0^1, x_1^2)u(x_0^1, x_1^2) + O(h^2)) dx_0^1 \wedge dx_0^2, \quad (C7)$$

the flux invariant $B_3 dx_0^1 \wedge dx_0^2$ of the continuous system plus an $O(h)$ correction, proportional to $A_{2,2}$. As for the modified symplectic Euler scheme of § 3.1, the adjoint of this scheme preserves the same form but with $\omega(h) \rightarrow \omega(-h) \neq \omega(h)$, and therefore the composition of this scheme with its adjoint cannot be assured of having a preserved two-form. The essential difference between this scheme and the symplectic Euler scheme for a canonical system is the dependence of A_2 on x^2 in (C1).

REFERENCES

- BLANES, S., CASAS, F. & MURUA, A. 2004 On the numerical integration of ordinary differential equations by processed methods. *SIAM J. Numer. Anal.* **42**, 531–552.
- BURBY, J. & ELLISON, C.L. 2017 Toroidal regularization of the guiding center Lagrangian. *Phys. Plasmas* **24**, 110703.
- BURBY, J.W. 2015 Chasing Hamiltonian structure in gyrokinetic theory. PhD thesis, Princeton University. Available at <https://www.osti.gov/biblio/1341437-chasing-hamiltonian-structure-gyrokinetic-theory>.
- BURBY, J.W. 2017a Magnetohydrodynamic motion of a two-fluid plasma. *Phys. Plasmas* **24**, 082104.
- BURBY, J.W. 2017b Finite-dimensional collisionless kinetic theory. *Phys. Plasmas* **24**, 032101.
- CARY, J.R. & LITTLEJOHN, R.G. 1983 Noncanonical Hamiltonian mechanics and its application to magnetic field line flow. *Ann. Phys.* **151**, 1–34.
- ELLISON, C.L. 2016 Development of multistep and degenerate variational integrators for applications in plasma physics. Doctoral thesis, School Princeton University.
- ELLISON, C.L., FINN, J.M., BURBY, J.W., KRAUS, M., QIN, H. & TANG, W.M. 2018 Degenerate variational integrators for magnetic field line flow and guiding center trajectories. *Phys. Plasmas* **25**, 052502.

- FINN, J.M. 2015 Issues in measure-preserving three dimensional flow integrators: self-adjointness, reversibility, and non-uniform time stepping. *Phys. Plasmas* **22**, 032508.
- HAIRER, E., LUBICH, C. & WANNER, G. 2006 *Geometric Numerical Integration*. Springer.
- LITTLEJOHN, R.G. 1981 Hamiltonian formulation of guiding center motion. *Phys. Fluids* **24**, 1730–1749.
- LITTLEJOHN, R.G. 1983 Variational principles of guiding centre motion. *J. Plasma Phys.* **29**, 111–125.
- LIU, J., WANG, Y. & QIN, H. 2016 Collisionless pitch-angle scattering of runaway electrons. *Nucl. Fusion* **56**, 064002.
- MARSDEN, J.E. & WEST, M. 2001 Discrete mechanics and variational integrators. *Acta Numer.* **10**, 357.
- QIN, H., GUAN, X. & TANG, W.M. 2009 Variational symplectic algorithm for guiding center dynamics and its application in tokamak geometry. *Phys. Plasmas* **16**, 042510.
- RICHARDSON, A.S. & FINN, J.M. 2011 Symplectic integrators with adaptive time steps. *Plasma Phys. Control. Fusion* **54**, 014004.