

Variational integrators for perturbed non-canonical Hamiltonian systems

J. W. Burby^a, C. L. Ellison^a, H. Qin^{a,b}

^a*Princeton Plasma Physics Laboratory, Princeton, New Jersey 08543, USA*

^b*Dept. of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

Abstract

Finite-dimensional non-canonical Hamiltonian systems arise naturally from Hamilton's principle in phase space. We present a method for deriving variational integrators that can be applied to perturbed non-canonical Hamiltonian systems on manifolds based on discretizing this phase-space variational principle. Relative to the perturbation parameter ϵ , this type of integrator can take $O(1)$ time steps with arbitrary accuracy in ϵ by leveraging the unperturbed dynamics. Moreover, these integrators are coordinate independent in the sense that their time-advance rules transform correctly when passing from one phase space coordinate system to another.

Keywords: variational integrators, geometric mechanics, perturbation theory

1. Introduction

The most famous variational principle in classical mechanics is *Hamilton's principle* of stationary action [1]. According to Hamilton's principle, a system's path in configuration space, $q(t) \in Q$, will be a critical point of

$$\mathcal{A}(q) = \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t) dt \quad (1)$$

regarded as a functional of paths in configuration space with fixed endpoints. Here, L is the Lagrangian function associated with the mechanical system in question. The closely-related *Hamilton's principle in phase space* [2] generalizes Hamilton's principle to arbitrary exact symplectic manifolds. Specifically, if the symplectic manifold M with symplectic form [3] $-\mathbf{d}\vartheta$ is the phase

space of a mechanical system, then the phase space variational principle asserts that a system's path in phase space, $z(t) \in M$, will be a critical point of

$$S(z) = \int_{t_1}^{t_2} \vartheta_{z(t)}(\dot{z}(t)) - H(z(t), t) dt \quad (2)$$

regarded as a functional of paths in phase space with fixed endpoints. Here, H is the system's Hamiltonian. This pair of variational principles serves as the variational workhorse of modern treatments of mechanics. Hamilton's principle is often well-suited to the formulation of a problem, as evidenced by its applications in continuum mechanics [4], whereas the phase space principle is naturally adapted to perturbation theory, especially in guiding center and gyrokinetic theory [5, 6].

In this article, we will investigate variational discretizations [7] of *perturbed* non-canonical Hamiltonian systems on manifolds that obey Hamilton's principle in phase space. These perturbed systems will be specified by a phase space manifold M ; a symplectic form on M of the form $-\mathbf{d}\vartheta$; and a time-dependent Hamiltonian function $\mathcal{H}_t = H_t + \epsilon h_t$, where H_t represents an unperturbed system and ϵ is a small parameter. Trajectories of this type of system are then given as extremals of the action S given in Eq. (2). A wide variety of mechanical systems fit this mold, including all perturbed canonical Hamiltonian systems. Notably, there are perturbed Hamiltonian systems for which Hamilton's principle in phase space is the only known variational formulation [6, 8]. The variational discretizations we will be concerned with are those that exploit the small value of ϵ to enhance the accuracy of the discrete Euler-Lagrange equations; see Refs. [9, 10, 11, 12] for generic variational discretizations of Hamilton's principle in phase space and Refs. [13, 14] for specialized methods that can be applied to Hamilton's principle in phase space while working in canonical coordinates.

Discretizations of Hamilton's principle for perturbed systems have been developed already in Ref. [15]. The discrete Euler-Lagrange equations associated with these discretizations are capable of recovering previously-discovered symplectic integrators [16, 17, 18] that exploit the small value of ϵ to effectively enhance their order of accuracy. In particular, these integrators are capable of achieving up to local $O(\epsilon^2)$ accuracy (how to achieve greater accuracy in ϵ with this type of integrator is not discussed in Ref. [15]). On the other hand, it seems a method for finding discretizations of Hamilton's principle in phase space that exploits the small value of ϵ has never been

discussed. One might hope that the technique used in Ref. [15] could be easily extended to treat the phase space variational principle, but this is not the case. The derivation of the discretizations in Ref. [15] depends on the presence of an exact discrete Lagrangian for its success. Because the usual notion of exact discrete Lagrangian [7] only applies to Hamilton's principle¹, this method cannot be transcribed to produce a similar method for discretizing Hamilton's principle in phase space. Thus, there is currently not a method for developing variational integrators for perturbed guiding center motion, or any other perturbed system whose only known variational formulation is in the form of Hamilton's principle in phase space.

The purpose of this article is to formulate the first discretizations of Hamilton's principle in phase space that are adapted to perturbed problems. These discretizations are contained in our Eqs. (37), (42), (45), and (47). Amongst them are discretizations whose time-advance rules are accurate to any desired order in ϵ while allowing for $O(1)$ time steps. Each discretization can be applied to any perturbed non-canonical Hamiltonian system whose continuous-time trajectories are extremals of the action S in Eq. (2). Our method for deriving these discretizations is based on constructing an exact discrete action for Hamilton's principle in phase space, a task that is fundamentally different from constructing an exact discrete action for Hamilton's principle. The method is completely coordinate-independent, and therefore leads to time-advance rules that transform correctly when passing from one coordinate chart on the phase space M to another. This coordinate-independence is achieved by introducing an arbitrary affine connection on M . We demonstrate that by initializing these two-step integration algorithms using the smooth modified system studied in backward error analysis [19, 20, 21], the resulting discrete-time trajectories satisfy one-step algorithms that preserve symplectic forms on M , which is a sharper result on symplecticity than that provided by the theory developed in Ref. [7] or Ref. [11].

The presentation will be organized as follows. We specify the scope of our work and precisely define the notion of a discretization of Hamilton's principle in phase space in Section 2. We derive an *exact* discretization of Hamilton's principle in phase space suitable to perturbed Hamiltonian systems in Section

¹The exact discrete Lagrangian associated with Hamilton's principle is a function of pairs of points in configuration space. The phase-space action cannot be regarded as a function of pairs of points in phase space in the same way because generally there is not a solution to Hamilton's equations that connects a pair of points in phase space.

3. Using this exact discretization, we develop approximate discretizations of Hamilton's principle in phase space that can be applied to practical problems in Section 4. In section 5, we discuss the symplecticity of the integration algorithms provided by our discretizations. Section 6 contains two examples. We conclude with a discussion in Section 7.

2. Problem statement

Let M be a symplectic manifold with symplectic form $-\mathbf{d}\vartheta$. Let $\mathcal{H}_t = H_t + \epsilon h_t$ be a time-dependent real-valued function on M , where ϵ is a small parameter. The time-dependent vector field $X_{\mathcal{H}_t}$ defined by the formula

$$\mathbf{i}_{X_{\mathcal{H}_t}} \mathbf{d}\vartheta = -\mathbf{d}\mathcal{H}_t \quad (3)$$

is known as the time-dependent Hamiltonian vector field with Hamiltonian \mathcal{H}_t [3]. Fix $t_1, t_2 \in \mathbb{R}$ with $t_1 < t_2$. If $\gamma : [t_1, t_2] \rightarrow M$ is an integral curve of $X_{\mathcal{H}_t}$, i.e.

$$\gamma'(t) = X_{\mathcal{H}_t}(\gamma(t)), \quad (4)$$

then γ is a critical point of the functional $S_{(\gamma(t_1), \gamma(t_2))} : \mathcal{P}_{(\gamma(t_1), \gamma(t_2))} \rightarrow \mathbb{R}$, where

$$\mathcal{P}_{(z_1, z_2)} = \{c : [t_1, t_2] \rightarrow M \mid c(t_1) = z_1, c(t_2) = z_2\} \quad (5)$$

and

$$S_{(\gamma(t_1), \gamma(t_2))}(c) = \int_{t_1}^{t_2} \vartheta_{c(t)}(c'(t)) - \mathcal{H}_t(c(t)) dt. \quad (6)$$

Conversely, if $c \in \mathcal{P}_{(z_1, z_2)}$ is a critical point of $S_{(z_1, z_2)}$, then c must be an integral curve of the Hamiltonian vector field $X_{\mathcal{H}_t}$. The latter pair of facts is known as Hamilton's principle in phase space [2]. Note that for many choices of (z_1, z_2) , $S_{(z_1, z_2)}$ will not have *any* critical points; a necessary condition for the existence of a critical point is that z_1 and z_2 can be connected by an integral curve of $X_{\mathcal{H}_t}$.

Let $\tau \in \mathbb{R}$ be a positive $O(1)$ constant. Assume $t_1 = N_1\tau$ and $t_2 = N_2\tau$, where $N_1, N_2 \in \mathbb{Z}$. Between the continuous-time path space, $\mathcal{P}_{(z_1, z_2)}$, and the discrete-time path space with increment τ ,

$$\mathbb{P}_{(z_1, z_2)} = \{\mathbf{c} : [t_1, t_2] \cap (\tau\mathbb{Z}) \rightarrow M \mid \mathbf{c}(t_1) = z_1, \mathbf{c}(t_2) = z_2\}, \quad (7)$$

there is a natural projection map $\pi_{(z_1, z_2)} : \mathcal{P}_{(z_1, z_2)} \rightarrow \mathbb{P}_{(z_1, z_2)}$, where for each integer $N_1 \leq k \leq N_2$

$$\pi_{(z_1, z_2)}(c)(k\tau) = c(k\tau). \quad (8)$$

Our goal is to identify a functional $\mathbb{S}_{(z_1, z_2)}^\infty : \mathbb{P}_{(z_1, z_2)} \rightarrow \mathbb{R}$, which we will refer to as the exact discrete action, with three properties.

(D1) If $\gamma : [t_1, t_2] \rightarrow M$ is an integral curve of the Hamiltonian vector field $X_{\mathcal{H}_t}$ with $\gamma(t_1) = z_1$ and $\gamma(t_2) = z_2$, then γ must be a critical point of the functional $\mathbb{S}_{(z_1, z_2)}^\infty \circ \pi_{(z_1, z_2)}$.

(D2) $\mathbb{S}_{(z_1, z_2)}^\infty$ must be of the form

$$\mathbb{S}_{(z_1, z_2)}^\infty(\mathbf{c}) = \sum_{k=N_1}^{N_2-1} L_\infty(\mathbf{c}_k, \mathbf{c}_{k+1}, \tau k), \quad (9)$$

where $\mathbf{c}_k = \mathbf{c}(\tau k)$ and $L_\infty : M \times M \times \mathbb{R} \rightarrow \mathbb{R}$.

(D3) If the exact discrete action is replaced with its l 'th order Maclaurin polynomial in ϵ ,

$$\mathbb{S}_{(z_1, z_2)}^l(\mathbf{c}) = \sum_{k=N_1}^{N_2-1} L_l(\mathbf{c}_k, \mathbf{c}_{k+1}, \tau k), \quad (10)$$

then the discrete Euler-Lagrange equations [7] associated with $\mathbb{S}_{(z_1, z_2)}^l$ should function as a numerical integration algorithm with local $O(\epsilon^{l+1})$ accuracy.

Remarks on such a functional are in order. Note that the time step of the integration algorithm associated with $\mathbb{S}_{(z_1, z_2)}^l$ is τ , which is not assumed to be small. The idea at work here is that integral curves of the unperturbed vector field X_{H_t} can be used to approximate integral curves of $X_{\mathcal{H}_t}$ with $O(\epsilon)$ accuracy on $O(1)$ time intervals. Thus, when H_t describes an integrable Hamiltonian system, the practical limit on the size of τ for a fixed value of ϵ should be expected to be $\tau \ll \tau_b$, where τ_b is the perturbation's characteristic bounce time. Also note that because the discrete Euler-Lagrange equations associated with $\mathbb{S}_{(z_1, z_2)}^l$ are given by

$$\mathbf{d}^{(2)} L_l(\mathbf{c}_{k-1}, \mathbf{c}_k, \tau(k-1)) + \mathbf{d}^{(1)} L_l(\mathbf{c}_k, \mathbf{c}_{k+1}, \tau k) = 0, \quad (11)$$

they provide a *two-step* [22, 19, 20] integration algorithm for a first-order dynamical system on M . Thus, the standard theory behind variational integrators [7] implies that this algorithm preserves a symplectic structure on $M \times M$. However, the flow of the Hamiltonian vector field $X_{\mathcal{H}_t}$ preserves a symplectic form on M . In Section 5 we reconcile these qualitatively different notions of structure preservation using a minor modification of the “smooth modified system” concept developed in [19]. Finally, note that by setting $\epsilon = 0$, we move into the setting of generic Hamiltonian systems on M , i.e. those without an *a priori* perturbative structure. Thus, when $\epsilon = 0$ and τ is chosen to be small, $\mathbb{S}_{(z_1, z_2)}^\infty$ can be expanded in powers of τ to yield arbitrarily accurate variational integrators for generic Hamiltonian systems.

3. Derivation of an exact discrete action for Hamilton’s principle in phase space

In order to derive a functional $\mathbb{S}_{(z_1, z_2)}^\infty$ that satisfies properties (D1)–(D3), we will manipulate the functional $S_{(z_1, z_2)}$ into the form $\mathbb{S}_{(z_1, z_2)}^\infty \circ \pi_{(z_1, z_2)}$ while making use of the following heuristic approximation principle.

The path space approximation principle: Modifications to the functional $S_{(z_1, z_2)}$ can be made as long as they do not change the first variation of $S_{(z_1, z_2)}$ at integral curves of $X_{\mathcal{H}_t}$.

The intuitive justification of this principle is that we are mainly interested in critical points of the functional $S_{(z_1, z_2)}$ and a critical point of $S_{(z_1, z_2)}$ will also be a critical point of $S'_{(z_1, z_2)}$ provided these two functionals agree (modulo a constant) in a neighborhood of the critical point.

As a convenient first step, we will pass into the “interaction picture”. Passing into the interaction picture amounts to transforming the path space in such a way that integral curves of the unperturbed vector field X_{H_t} become trivially constant curves. Let $F_{t,s} : M \rightarrow M$ be the time-dependent flow map of the unperturbed Hamiltonian vector field X_{H_t} , i.e. the two-parameter family of mappings characterized by the relations

$$F_{t,t}(z) = z \tag{12}$$

$$\frac{d}{dt} F_{t,s}(z) = X_{H_t}(F_{t,s}(z)). \tag{13}$$

Then the interaction picture transformation from the old path space to the new path space $\mathcal{I} : \mathcal{P}_{(z_1, z_2)} \rightarrow \mathcal{P}_{(\bar{z}_1, \bar{z}_2)}$, with $(\bar{z}_1, \bar{z}_2) = (z_1, F_{t_1, t_2}(z_2))$, is given by

$$\mathcal{I}(c)(t) = F_{t_1, t}(c(t)). \quad (14)$$

After performing this change-of-path, the functional $S_{(z_1, z_2)}$ is transformed into $\bar{S}_{(\bar{z}_1, \bar{z}_2)} = \mathcal{I}_* S_{(z_1, z_2)}$, which is the pushforward of $S_{(z_1, z_2)}$ along the mapping \mathcal{I} . Given $\bar{c} \in \mathcal{P}_{(\bar{z}_1, \bar{z}_2)}$, $\bar{S}_{(\bar{z}_1, \bar{z}_2)}(\bar{c})$ is readily found to be

$$\bar{S}_{(\bar{z}_1, \bar{z}_2)}(\bar{c}) = \int_{t_1}^{t_2} \vartheta_{\bar{c}(t)}(\bar{c}'(t)) - \epsilon(F_{t, t_1}^* h_t)(\bar{c}(t)) dt + \text{const}. \quad (15)$$

The constant term does not affect the location of critical points, and so we omit it from this point forward. Also note that, by Hamilton's principle in phase space, \bar{c} is a critical point of $\bar{S}_{(\bar{z}_1, \bar{z}_2)}$ if and only if \bar{c} is an integral curve of the time-dependent Hamiltonian vector field $X_{\epsilon K_t}$, where

$$K_t = F_{t, t_1}^* h_t. \quad (16)$$

Next, we decompose the time integral in $\bar{S}_{(\bar{z}_1, \bar{z}_2)}$ as

$$\bar{S}_{(\bar{z}_1, \bar{z}_2)}(\bar{c}) = \sum_{k=N_1}^{N_2-1} \bar{S}^k(\bar{c}), \quad (17)$$

where

$$\bar{S}^k(\bar{c}) = \int_{\tau_k}^{\tau(k+1)} \vartheta_{\bar{c}(t)}(\bar{c}'(t)) - \epsilon K_t(\bar{c}(t)) dt, \quad (18)$$

and examine $\bar{S}^k(\bar{c})$ for each k . The goal of this analysis is to devise an approximation for $\bar{S}^k(\bar{c})$ that depends on \bar{c} only through \bar{c}_k and \bar{c}_{k+1} . Let $G_{t,s} : M \rightarrow M$ be the time-dependent flow map of the Hamiltonian vector field $X_{\epsilon K_t}$. As is readily verified, the identity

$$\begin{aligned} \int_{\tau_k}^{\tau(k+1)} \vartheta_{\bar{c}(t)}(\bar{c}'(t)) - \epsilon K_t(\bar{c}(t)) dt &= \int_{\bar{c}^k} \vartheta + \left(\epsilon \int_{\tau(k+1/2)}^{\tau(k+1)} G_{s, \tau(k+1)}^* l_s ds \right) (\bar{c}_{k+1}) \\ &\quad + \left(\epsilon \int_{\tau_k}^{\tau(k+1/2)} G_{s, \tau_k}^* l_s ds \right) (\bar{c}_k) \end{aligned} \quad (19)$$

holds, where

$$\bar{c}^k(t) = G_{\tau(k+1/2),t}(\bar{c}(t)), \quad \tau k \leq t \leq \tau(k+1) \quad (20)$$

and

$$l_s = \vartheta(X_{K_s}) - K_s. \quad (21)$$

Thus, $\bar{S}^k(\bar{c})$ could be determined using only the values of \bar{c} at τk and $\tau(k+1)$ were it not for the term $\int_{\bar{c}^k} \vartheta$.

Observe that when \bar{c} is a critical point of $\bar{S}_{(\bar{z}_1, \bar{z}_2)}$, the curve \bar{c}^k is constant and the integral $\int_{\bar{c}^k} \vartheta$ vanishes. When \bar{c} is infinitesimally close to a critical point, $\int_{\bar{c}^k} \vartheta$ is given by the first variation of the functional $\mathcal{F}_1 : \mathcal{P}_{(\bar{z}_1, \bar{z}_2)} \rightarrow \mathbb{R}$, where

$$\mathcal{F}_1(\bar{c}) = \int_{\bar{c}^k} \vartheta. \quad (22)$$

To calculate the variation of \mathcal{F}_1 , we first compute the variation of $\mathcal{F}_o : \mathcal{P}_{(\bar{z}_1, \bar{z}_2)} \rightarrow \mathbb{R}$ at a trivial curve $\bar{c}(t) = \text{const}$, where

$$\mathcal{F}_o(\bar{c}) = \int_{\bar{c}^k} \vartheta, \quad (23)$$

and $\bar{c}^k = \bar{c}[\tau k, \tau(k+1)]$. The result is readily found to be

$$\mathbf{d}\mathcal{F}_o \bar{c}(\delta \bar{c}) = \vartheta_{\bar{c}_{k+1}}(\delta \bar{c}_{k+1}) - \vartheta_{\bar{c}_k}(\delta \bar{c}_k). \quad (24)$$

The chain rule then implies that the first variation of \mathcal{F}_1 at a critical point of $\bar{S}_{(\bar{z}_1, \bar{z}_2)}$ is given by

$$\begin{aligned} \mathbf{d}\mathcal{F}_1 \bar{c}(\delta \bar{c}) &= \vartheta_{G_{\tau(k+1/2), \tau(k+1)}(\bar{c}_{k+1})}(TG_{\tau(k+1/2), \tau(k+1)}(\delta \bar{c}_{k+1})) \\ &\quad - \vartheta_{G_{\tau(k+1/2), \tau k}(\bar{c}_k)}(TG_{\tau(k+1/2), \tau k}(\delta \bar{c}_k)). \end{aligned} \quad (25)$$

Here, T denotes the tangent functor as defined in Ref. [3]. Notably, the first variation of \mathcal{F}_1 at a critical point of $\bar{S}_{(\bar{z}_1, \bar{z}_2)}$ is completely determined by $\delta \bar{c}$ and \bar{c} evaluated at τk and $\tau(k+1)$.

Guided by these observations and the path space approximation principle, we will now replace the term $\int_{\bar{c}^k} \vartheta$ in $\bar{S}^k(\bar{c})$ with an approximation that can

be computed using only the values of \bar{c} at τk and $\tau(k+1)$. We will choose this approximation so that it agrees with $\int_{\bar{c}^k} \vartheta$ when \bar{c} is a critical point and when \bar{c} is infinitesimally close to a critical point. Let ∇ be an arbitrary affine connection on M . It is well-known [23] that M admits an open cover $\{U_i\}_{i \in I}$ with two properties: (i) if $z_1, z_2 \in U_i$, then there is a unique geodesic segment contained in U_i with endpoints z_1 and z_2 , (ii) If $z_1, z_2 \in U_i$ and $z_1, z_2 \in U_j$, then the geodesic segment joining z_1, z_2 in U_i is equal to the geodesic segment joining z_1, z_2 in U_j . Thus, on the open neighborhood of the diagonal in $M \times M$, $\mathcal{O} = \bigcup_{i \in I} U_i \times U_i$, we can define a real-valued function

$$f(z_1, z_2) = \int_{I(z_2, z_1)} \vartheta, \quad (26)$$

where $I(z_2, z_1)$ is the unique directed geodesic segment from z_1 to z_2 contained in some U_i . If \mathcal{O} cannot be taken to be all of $M \times M$, assume that f has been smoothly extended to all of $M \times M$. In terms of this possibly-extended f our approximation for $\int_{\bar{c}^k} \vartheta$ is

$$\begin{aligned} \int_{\bar{c}^k} \vartheta &\approx f(F_{\tau k, t_1}(\bar{c}_k), F_{\tau k, t_1}(\bar{c}_{k+1})) \\ &\quad - \left(\int_{t_1}^{\tau k} F_{s, t_1}^* \mathcal{L}_s ds \right) (\bar{c}_{k+1}) \\ &\quad + \left(\int_{t_1}^{\tau k} F_{s, t_1}^* \mathcal{L}_s ds \right) (\bar{c}_k), \end{aligned} \quad (27)$$

where

$$\mathcal{L}_s = \vartheta(X_{H_s}) - H_s. \quad (28)$$

It is readily verified that this approximation is exact when \bar{c} is either a critical point of $\bar{S}_{(\bar{z}_1, \bar{z}_2)}$ or infinitesimally close to such a critical point. When \bar{c} is a critical point, both sides of Eq. (27) obviously vanish. Likewise, regarding each side of Eq. (27) as a functional of \bar{c} , the two sides' first variations at a critical point agree. The latter assertion is easy to check using Eq. (25) and

the fact that when \bar{c} is sufficiently close to a critical point, the right hand-side of Eq. (27) is given by

$$\int_{I(F_{\tau k, t_1}(\bar{c}_{k+1}), F_{\tau k, t_1}(\bar{c}_k))} F_{t_1, \tau k}^* \vartheta. \quad (29)$$

With this approximation in place, we can now easily obtain an expression in the interaction picture for a functional $\bar{\mathbb{S}}_{(\bar{z}_1, \bar{z}_2)}^\infty$ that satisfies (D1) and (D2). To see this, note that we now have $\bar{S}^k(\bar{c}) \approx \bar{L}_\infty(\bar{c}_k, \bar{c}_{k+1}, \tau k)$, where

$$\begin{aligned} \bar{L}_\infty(\bar{c}_k, \bar{c}_{k+1}, t) = & f(F_{t, t_1}(\bar{c}_k), F_{t, t_1}(\bar{c}_{k+1})) \\ & - \left(\int_{t_1}^t F_{s, t_1}^* \mathcal{L}_s ds \right) (\bar{c}_{k+1}) \\ & + \left(\int_{t_1}^t F_{s, t_1}^* \mathcal{L}_s ds \right) (\bar{c}_k) \\ & + \left(\epsilon \int_{t+\tau/2}^{t+\tau} G_{s, t+\tau}^* l_s ds \right) (\bar{c}_{k+1}) \\ & + \left(\epsilon \int_t^{t+\tau/2} G_{s, t}^* l_s ds \right) (\bar{c}_k). \end{aligned} \quad (30)$$

Thus, Eq. (17) implies

$$\bar{S}_{(\bar{z}_1, \bar{z}_2)}(\bar{c}) \approx \sum_{k=N_1}^{N_2-1} \bar{L}_\infty(\bar{c}_k, \bar{c}_{k+1}, \tau k) = \bar{\mathbb{S}}_{(\bar{z}_1, \bar{z}_2)}^\infty \circ \pi_{(\bar{z}_1, \bar{z}_2)}(\bar{c}), \quad (31)$$

where $\bar{\mathbb{S}}_{(\bar{z}_1, \bar{z}_2)}^\infty : \mathbb{P}_{\bar{z}_1, \bar{z}_2} \rightarrow \mathbb{R}$ is given by

$$\bar{\mathbb{S}}_{(\bar{z}_1, \bar{z}_2)}^\infty(\bar{\mathbf{c}}) = \sum_{k=N_1}^{N_2-1} \bar{L}_\infty(\bar{\mathbf{c}}_k, \bar{\mathbf{c}}_{k+1}, \tau k). \quad (32)$$

This says that $\bar{\mathbb{S}}_{(\bar{z}_1, \bar{z}_2)}^\infty$ satisfies (D2). Also note that, by construction, we have the following equalities when \bar{c} is a critical point of $\bar{S}_{(\bar{z}_1, \bar{z}_2)}$.

$$\bar{S}_{(\bar{z}_1, \bar{z}_2)}(\bar{c}) = \bar{\mathbb{S}}_{(\bar{z}_1, \bar{z}_2)}^\infty \circ \pi_{(\bar{z}_1, \bar{z}_2)}(\bar{c}) \quad (33)$$

$$\mathbf{d}\bar{S}_{(\bar{z}_1, \bar{z}_2)} \bar{c} = \mathbf{d}(\bar{\mathbb{S}}_{(\bar{z}_1, \bar{z}_2)}^\infty \circ \pi_{(\bar{z}_1, \bar{z}_2)})_{\bar{c}}. \quad (34)$$

Therefore, any critical point of $\bar{S}_{(\bar{z}_1, \bar{z}_2)}$ is also a critical point of $\bar{S}_{(\bar{z}_1, \bar{z}_2)}^\infty \circ \pi_{(\bar{z}_1, \bar{z}_2)}$, which says that $\bar{S}_{(\bar{z}_1, \bar{z}_2)}^\infty$ satisfies (D1).

In fact $\bar{S}_{(\bar{z}_1, \bar{z}_2)}^\infty$ satisfies (D3) in addition to (D1) and (D2). The proof of this statement is not substantially different than the proof of Theorem 2.3.1 in Ref. [7]. We have therefore succeeded in identifying an exact discrete action $\bar{S}_{(\bar{z}_1, \bar{z}_2)}^\infty$. However, we currently have this action expressed in the interaction picture. To remedy this, we will conclude this section by passing out of the interaction picture.

Passing out of the interaction picture consists of transforming the new path space $\mathcal{P}_{(\bar{z}_1, \bar{z}_2)}$ back into the old path space $\mathcal{P}_{(z_1, z_2)}$ by applying the mapping \mathcal{I}^{-1} . Upon performing this change-of-path, the functional $\bar{S}_{(\bar{z}_1, \bar{z}_2)}^\infty \circ \pi_{(\bar{z}_1, \bar{z}_2)}$ transforms into $\mathcal{I}^*(\bar{S}_{(\bar{z}_1, \bar{z}_2)}^\infty \circ \pi_{(\bar{z}_1, \bar{z}_2)})$. After some tedious, yet straightforward algebraic manipulations, we have found this pullback is given by

$$\mathcal{I}^*(\bar{S}_{(\bar{z}_1, \bar{z}_2)}^\infty \circ \pi_{(\bar{z}_1, \bar{z}_2)}) = S_{(z_1, z_2)}^\infty \circ \pi_{(z_1, z_2)}, \quad (35)$$

where

$$S_{(z_1, z_2)}^\infty(\mathbf{c}) = \sum_{k=N_1}^{N_2-1} L_\infty(\mathbf{c}_k, \mathbf{c}_{k+1}, \tau k), \quad (36)$$

and

$$\begin{aligned} L_\infty(\mathbf{c}_k, \mathbf{c}_{k+1}, t) &= \bar{L}_\infty(F_{t_1, t}(\mathbf{c}_k), F_{t_1, t}(\mathbf{c}_{k+1}), t) \\ &= f(\Phi_{t+\tau/2, t}^t(\mathbf{c}_k), \Phi_{t+\tau/2, t+\tau}^t(F_{t, t+\tau}(\mathbf{c}_{k+1}))) \\ &\quad + \mathcal{L}^t(\mathbf{c}_k) - \mathcal{L}^{t+\tau}(\mathbf{c}_{k+1}) + \left(\int_t^{t+\tau} F_{s, t+\tau}^* \mathcal{L}_s ds \right) (\mathbf{c}_{k+1}) \\ &\quad + \left(\epsilon \int_{t+\tau/2}^{t+\tau} \Phi_{s, t+\tau}^{t*} l_s^t ds \right) (F_{t, t+\tau}(\mathbf{c}_{k+1})) \\ &\quad + \left(\epsilon \int_t^{t+\tau/2} \Phi_{s, t}^{t*} l_s^t ds \right) (\mathbf{c}_k). \end{aligned} \quad (37)$$

The notation introduced in this expression is defined as follows. The mapping $\Phi_{t,s}^u$ is the time-dependent flow map of the Hamiltonian vector field with

time-dependent Hamiltonian $\epsilon F_{t,u}^* h_t$. In particular,

$$\Phi_{t,s}^u = F_{u,t_1} \circ G_{t,s} \circ F_{t_1,u}. \quad (38)$$

The function l_s^u is given by

$$l_s^u = \vartheta(X_{F_{s,u}^*} h_s) - F_{s,u}^* h_s. \quad (39)$$

Finally,

$$\mathcal{L}^k = \int_{t_1}^{\tau k} F_{s,\tau k}^* \mathcal{L}_s ds. \quad (40)$$

The functional $\mathbb{S}_{(z_1, z_2)}^\infty$ is therefore a valid exact discrete action, i.e. it satisfies properties (D1)–(D3). Note that the terms $\mathcal{L}^k(\mathbf{c}_k) - \mathcal{L}^{k+1}(\mathbf{c}_{k+1})$ in L_∞ are gauge contributions in the sense that, when summed over k , they only contribute a constant to $\mathbb{S}_{(z_1, z_2)}^\infty$. Therefore they can be omitted from L_∞ without affecting the location of critical points.

4. Truncations of $\mathbb{S}_{(z_1, z_2)}^\infty$

While the discrete Euler-Lagrange equations associated with $\mathbb{S}_{(z_1, z_2)}^\infty$ are rigorously satisfied by the integral curves of $X_{\mathcal{H}_t}$ by property (D1), they generally do not serve as a particularly useful numerical integration algorithm. This is because calculating L_∞ is generally very difficult; the flow maps $F_{t,s}$ and $\Phi_{t,s}^u$ must be known in advance. Thus, it is important to have manageable approximations for $\mathbb{S}_{(z_1, z_2)}^\infty$ on hand when developing variational integrators. To this end, we will now present general expressions for the first few terms in $\mathbb{S}_{(z_1, z_2)}^\infty$'s power series in ϵ . The utility of such expressions follows from property (D3): if $\mathbb{S}_{(z_1, z_2)}^\infty$ is replaced by its l 'th order Maclaurin polynomial in ϵ , the resulting discrete Euler-Lagrange equations will function as a two-step numerical integration algorithm with local $O(\epsilon^{l+1})$ accuracy. We will adopt the convention that $\mathbb{S}_{(z_1, z_2)}^l$ denotes $\mathbb{S}_{(z_1, z_2)}^\infty$'s l 'th order Maclaurin polynomial in ϵ .

$l = 0$:

When $l = 0$, the truncated discrete action is given by

$$\mathbb{S}_{(z_1, z_2)}^0(\mathbf{c}) = \sum_{k=N_1}^{N_2-1} L_0(\mathbf{c}_k, \mathbf{c}_{k+1}, \tau k), \quad (41)$$

where

$$L_0(\mathbf{c}_k, \mathbf{c}_{k+1}, t) = f(\mathbf{c}_k, F_{t,t+\tau}(\mathbf{c}_{k+1})) + \mathcal{L}^t(\mathbf{c}_k) - \mathcal{L}^{t+\tau}(\mathbf{c}_{k+1}) + \left(\int_t^{t+\tau} F_{s,t+\tau}^* \mathcal{L}_s ds \right) (\mathbf{c}_{k+1}). \quad (42)$$

Recall that the function f is defined in Eq. (26). Also recall that $\mathcal{L}^{\tau k}(\mathbf{c}_k) - \mathcal{L}^{\tau(k+1)}(\mathbf{c}_{k+1})$ is a gauge term and can therefore be omitted. The discrete Euler-Lagrange equations associated with L_0 provide a two-step integration algorithm for $X_{\mathcal{H}_t}$ with local $O(\epsilon)$ accuracy. In particular, $\mathbb{S}_{(z_1, z_2)}^0$ serves as an exact discrete action for the unperturbed system X_{H_t} . Therefore, L_0 can be used to derive variational integrators for *generic* non-canonical Hamiltonian systems without a perturbative structure as follows. Choose τ to be a small parameter. Then L_0 can meaningfully be expanded in a Maclaurin series in τ . From the general theory developed in Ref. [7], it follows that if L_0 is replaced with its n 'th order Maclaurin series in τ , the associated discrete Euler-Lagrange equations will serve as an $(n + 1)$ 'th order integrator.

$l = 1$:

In order to derive an expression for $\mathbb{S}_{(z_1, z_2)}^1$, it is necessary to make use of the following identity. Let $K_s^u = F_{s,u}^* h_s$. If $g : M \rightarrow M$ is an arbitrary smooth function, then

$$\begin{aligned} \Phi_{t,s}^{u*} g &= g + \epsilon \int_s^t \Phi_{a,s}^{u*} \left(L_{X_{K_a^u}} g \right) da \\ &= g + \epsilon \int_s^t L_{X_{K_a^u}} g da + \epsilon^2 \int_s^t \int_s^a L_{X_{K_b^u}} L_{X_{K_a^u}} g db da + O(\epsilon^3), \end{aligned} \quad (43)$$

where L_X denotes the Lie derivative along the vector field X . Provided $|t - s| = O(1)$, this identity can be used to obtain an asymptotic expansion of the quantity $\Phi_{t,s}^{u*} g$ in powers of ϵ .

Applying this identity to Eq. (37), we obtain

$$\mathbb{S}_{(z_1, z_2)}^1(\mathbf{c}) = \sum_{k=N_1}^{N_2-1} L_1(\mathbf{c}_k, \mathbf{c}_{k+1}, \tau k), \quad (44)$$

where

$$\begin{aligned}
L_1(\mathbf{c}_k, \mathbf{c}_{k+1}, t) = & L_0(\mathbf{c}_k, \mathbf{c}_{k+1}, t) \\
& + \epsilon \int_{t+\tau/2}^{t+\tau} l_s^t(F_{t,t+\tau}(\mathbf{c}_{k+1})) ds + \epsilon \int_t^{t+\tau/2} l_s^t(\mathbf{c}_k) ds \\
& - \left(\epsilon \int_{t+\tau/2}^{t+\tau} L_{X_{K_s^t}}^{(2)} f ds \right) (\mathbf{c}_k, F_{t,t+\tau}(\mathbf{c}_{k+1})) \\
& + \left(\epsilon \int_t^{t+\tau/2} L_{X_{K_s^t}}^{(1)} f ds \right) (\mathbf{c}_k, F_{t,t+\tau}(\mathbf{c}_{k+1})). \tag{45}
\end{aligned}$$

The discrete Euler-Lagrange equations associated with L_1 furnish a variational integrator for $X_{\mathcal{H}_t}$ with local $O(\epsilon^2)$ accuracy. Keep in mind that the time step $\tau = O(1)$. Practically speaking, for a given value of ϵ , τ should be significantly less than the perturbation's characteristic bounce time. If τ is chosen to be a small parameter, then these expressions can be expanded in powers of τ . If this expansion were to be performed and L_1 were replaced with its n 'th order Maclaurin polynomial in τ , the resulting discrete Euler-Lagrange equations would yield an integration algorithm with local $O(\tau^{n+1}\epsilon^2)$ accuracy.

$l = 2$:

Upon further application of Eq. (43), the $l = 2$ result is given by

$$\mathbb{S}_{(z_1, z_2)}^2(\mathbf{c}) = \sum_{k=N_1}^{N_2-1} L_2(\mathbf{c}_k, \mathbf{c}_{k+1}, \tau k), \tag{46}$$

where

$$\begin{aligned}
L_2(\mathbf{c}_k, \mathbf{c}_{k+1}, t) &= L_1(\mathbf{c}_k, \mathbf{c}_{k+1}, t) \\
&- \epsilon^2 \int_{t+\tau/2}^{t+\tau} \int_s^{t+\tau} L_{X_{K_a^t}} l_s^t(F_{t,t+\tau}(\mathbf{c}_{k+1})) da ds \\
&+ \epsilon^2 \int_t^{t+\tau/2} \int_t^s L_{X_{K_a^t}} l_s^t(\mathbf{c}_k) da ds \\
&+ \left(\epsilon^2 \int_{t+\tau/2}^{t+\tau} \int_s^{t+\tau} L_{X_{K_a^t}}^{(2)} L_{X_{K_s^t}}^{(2)} f da ds \right) (\mathbf{c}_k, F_{t,t+\tau}(\mathbf{c}_{k+1})) \\
&+ \left(\epsilon^2 \int_t^{t+\tau/2} \int_t^s L_{X_{K_a^t}}^{(1)} L_{X_{K_s^t}}^{(1)} f da ds \right) (\mathbf{c}_k, F_{t,t+\tau}(\mathbf{c}_{k+1})) \\
&- \left(\epsilon^2 \int_t^{t+\tau/2} \int_{t+\tau/2}^{t+\tau} L_{X_{K_s^t}}^{(1)} L_{X_{K_a^t}}^{(2)} f da ds \right) (\mathbf{c}_k, F_{t,t+\tau}(\mathbf{c}_{k+1})). \quad (47)
\end{aligned}$$

The discrete Euler-Lagrange equations associated with L_2 furnish a variational integrator for $X_{\mathcal{H}_t}$ with local $O(\epsilon^3)$ accuracy.

5. Symplecticity

Fix a non-negative integer l . In this section we will discuss the sense in which the numerical integration algorithm associated with $\mathbb{S}_{(z_1, z_2)}^l$ is symplectic. This topic is more subtle than it may first appear because the discrete Euler-Lagrange equations associated with L_l give a two-step integrator for a first-order dynamical system on M .

First we will illustrate the sense in which the integration algorithm associated with $\mathbb{S}_{(z_1, z_2)}^l$ is symplectic on $M \times M$ by applying the methods of Ref. [7] in a straightforward manner. Recall that the discrete Euler-Lagrange equations associated with L_l are given by

$$\mathbf{d}^{(2)} L_l(\mathbf{c}_{k-1}, \mathbf{c}_k, \tau(k-1)) + \mathbf{d}^{(1)} L_l(\mathbf{c}_k, \mathbf{c}_{k+1}, \tau k) = 0. \quad (48)$$

Thus, we can define a mapping $F_k^l : M \times M \rightarrow M \times M$, where

$$F_k^l(z_1, z_2) = (z_2, f_k^l(z_1, z_2)), \quad (49)$$

and $f_k^l : M \times M \rightarrow M$ is defined implicitly by the equation

$$\mathbf{d}^{(2)}L_l(z_1, z_2, \tau(k-1)) + \mathbf{d}^{(1)}L_l(z_2, f_k^l(z_1, z_2), \tau k) = 0. \quad (50)$$

F_k^l allows us to parameterize the space of solutions of the discrete Euler-Lagrange equations by $M \times M$. In particular, it can be used to define a mapping $\mathfrak{C} : M \times M \rightarrow \mathbb{P}$, where \mathbb{P} is the set of mappings $[\tau N_1, \tau N_2] \cap (\tau \mathbb{Z}) \rightarrow M$ and

$$\mathfrak{C}(z_1, z_2)(\tau N_1) = z_1 \quad (51)$$

$$\mathfrak{C}(z_1, z_2)(\tau k) = \pi_1 \circ F_k^l \circ F_{k-1}^l \circ \dots \circ F_{N_1+1}^l(z_1, z_2), \quad N_1 < k \leq N_2 \quad (52)$$

Using this parameterization, we can define the restricted action on $M \times M$, $\hat{\mathbb{S}}_{M \times M}^l : M \times M \rightarrow \mathbb{R}$, where

$$\hat{\mathbb{S}}_{M \times M}^l(z_1, z_2) = \sum_{k=N_1}^{N_2-1} L_l(\mathfrak{C}(z_1, z_2)(\tau k), \mathfrak{C}(z_1, z_2)(\tau(k+1)), \tau k). \quad (53)$$

The exterior derivative of $\hat{\mathbb{S}}_{M \times M}^l$ is readily found to be

$$\mathbf{d}\hat{\mathbb{S}}_{M \times M}^l = \theta_1^{N_1} - F_{N_2-1, N_1}^{l*} \theta_2^{N_2-1}, \quad (54)$$

where $F_{N_2-1, N_1}^l = F_{N_2-1}^l \circ \dots \circ F_{N_1+1}^l$ and the one-forms θ_1^k, θ_2^k are given by

$$\theta_1^k(z_1, z_2)(u_{(z_1, z_2)}) = \mathbf{d}L_l(z_1, z_2, \tau k)(T\pi_1(u_{(z_1, z_2)})) \quad (55)$$

$$\theta_2^k(z_1, z_2)(u_{(z_1, z_2)}) = -\mathbf{d}L_l(z_1, z_2, \tau k)(T\pi_2(u_{(z_1, z_2)})). \quad (56)$$

Likewise, the second exterior derivative of $\hat{\mathbb{S}}_{M \times M}^l$ is found to be

$$\mathbf{d}\mathbf{d}\hat{\mathbb{S}}_{M \times M}^l = 0 = \mathbf{d}\theta_1^{N_1} - F_{N_2-1, N_1}^{l*} \mathbf{d}\theta_2^{N_2-1}. \quad (57)$$

It is clear from their definitions that the one-forms θ_1^k and θ_2^k differ by an exact differential,

$$\theta_1^k(z_1, z_2) - \theta_2^k(z_1, z_2) = \mathbf{d}L_l(z_1, z_2, \tau k). \quad (58)$$

Therefore we have the conservation law

$$F_{k,N_1}^{l*} \omega^k \text{ is independent of } k, \quad (59)$$

where $\omega^k = \mathbf{d}\theta_1^k = \mathbf{d}\theta_2^k$. We have thus shown that the time-dependent flow map F_{k,N_1}^l on $M \times M$ associated with $\mathbb{S}_{(z_1, z_2)}^l$ preserves a time-dependent symplectic form on $M \times M$.

Next we will demonstrate a more refined result on the symplecticity of the integration algorithm associated with $\mathbb{S}_{(z_1, z_2)}^l$. Assume that there exists a time-dependent vector field X_t on M whose associated flow map $\mathcal{F}_{t,s}$ satisfies the discrete Euler-Lagrange equations *exactly*. In other words, $\mathcal{F}_{t,s}$ satisfies the equation

$$\begin{aligned} & \mathbf{d}^{(2)} L_l(\mathcal{F}_{\tau(k-1), \tau N_1}(z), \mathcal{F}_{\tau k, \tau N_1}(z), \tau(k-1)) \\ & + \mathbf{d}^{(1)} L_l(\mathcal{F}_{\tau k, \tau N_1}(z), \mathcal{F}_{\tau(k+1), \tau N_1}(z), \tau k) = 0. \end{aligned} \quad (60)$$

for each $N_1 \leq k \leq N_2$ and $z \in M$. Using $\mathcal{F}_{t,s}$, we can define a mapping $\mathcal{C} : M \rightarrow \mathbb{P}$, where

$$\mathcal{C}(z)(\tau k) = \mathcal{F}_{\tau k, \tau N_1}(z). \quad (61)$$

The mapping \mathcal{C} then naturally leads to the introduction of the restricted action on M , $\hat{\mathbb{S}}_M^l : M \rightarrow \mathbb{R}$, where

$$\hat{\mathbb{S}}_M^l(z) = \sum_{k=N_1}^{N_2-1} L_l(\mathcal{F}_{\tau k, \tau N_1}(z), \mathcal{F}_{\tau(k+1), \tau N_1}(z), \tau k). \quad (62)$$

The identity $\mathbf{d}\hat{\mathbb{S}}_M^l = 0$ then gives the conservation law

$$\mathcal{F}_{\tau k, \tau N_1}^* \Omega^k \text{ is independent of } k, \quad (63)$$

where $\Omega^k = \mathbf{d}\tilde{\theta}_1^k = \mathbf{d}\tilde{\theta}_2^k$ and

$$\tilde{\theta}_1^k(z) = \mathbf{d}^{(1)} L_l(z, \mathcal{F}_{\tau(k+1), \tau k}(z), \tau k) \quad (64)$$

$$\tilde{\theta}_2^k(z) = -\mathbf{d}^{(2)} L_l(\mathcal{F}_{\tau(k-1), \tau k}(z), z, \tau(k-1)). \quad (65)$$

Note that $\tilde{\theta}_1^k = \tilde{\theta}_2^k$. Thus, we see that the smooth modified system, which is specified by X_t , preserves a time-dependent symplectic form on M . The sense in which this conservation law applies to the two-step integration algorithm

specified by $\mathbb{S}_{(z_1, z_2)}^l$ is as follows. The two step algorithm specified by the discrete Euler-Lagrange equations requires a pair of initial conditions in order to produce a discrete-time trajectory. If the second initial condition is supplied by flowing along X_t for τ seconds starting from the first initial condition, then the discrete trajectory produced by solving the discrete Euler-Lagrange equations will automatically lie along an integral curve of X_t . Thus, provided the second initial condition for our two-step method is chosen carefully, the two-step method is equivalent to the one-step method given by the flow map associated with X_t , which we have just shown preserves a time-dependent symplectic form on M . This result is consistent with the discussions found in Refs. [21, 19, 20] that explain the advantages of choosing the second initial condition for a two-step integrator using the smooth modified system.

The question of whether or not the vector field X_t exists seems to be incompletely resolved. Using the methods of Hairer, who calls X_t the smooth modified system [19, 20], an asymptotic series for X_t in powers of ϵ can be developed in a straightforward manner if one assumes the ansatz

$$X_t = X_{H_t} + \epsilon Y_t^1 + \epsilon^2 Y_t^2 + \dots \quad (66)$$

While, in general, this series diverges, such divergence does not necessarily imply that an X_t with the desired properties fails to exist. It is possible, for instance, that the series (66) can be resummed in the sense of Borel [24] to give X_t . Note that this existence question is not resolved by the result proved in Ref. [25]. Indeed, two step variational integrators tend to not be absolutely stable and therefore cannot be treated with the methods of Ref. [25].

Regardless of the answer to the existence question, the asymptotic series in Eq. (66) can often be computed and then truncated at some order. While this truncated vector field will not have a flow map that exactly satisfies the discrete Euler-Lagrange equations, by truncating at a sufficiently high order, it can be made to satisfy the discrete Euler-Lagrange equations with any desired level of accuracy. This fact has already been exploited to improve the stability properties of multi-step variational integration methods in [21]. It would be interesting to also exploit the same fact to develop energy and symplecticity bounds for the discretizations of Hamilton's principle in phase space developed here. We leave this to future consideration. Note that existing proofs of bounded energy errors for variational integrators only apply to discretizations of Hamilton's principle, and not to discretizations of Hamilton's principle in phase space; when the second initial condition supplied to

a two-step discretization of Hamilton's principle in phase space is not chosen to lie along an integral curve of X_t , there are known examples of poor energy behavior [26, 27].

6. Examples

In this section we will illustrate the use of our discretizations of Hamilton's principle in phase space to develop variational integrators for a pair of perturbed Hamiltonian systems. In order to illustrate that L_∞ gives discrete Euler-Lagrange equations that are exactly satisfied by the appropriate continuous-time trajectories, we will first treat the harmonic oscillator regarded as a perturbed rigid rotor. In this example, L_∞ is simple to calculate and it is a simple matter to verify that true solutions to the harmonic oscillator differential equation satisfy L_∞ 's discrete Euler-Lagrange equations. We will then use the first-order (in ϵ) approximation for L_∞ , L_1 , to derive a variational integrator for a non-canonical Hamiltonian system that describes the nearly-integrable flow of magnetic field lines in a nominally axisymmetric geometry. This second example is less trivial than the Harmonic oscillator in the sense that the field line dynamics are non-linear and L_∞ is impossible to calculate; the computation of L_1 is already an onerous task to perform without the aid of symbolic manipulation software.

Example 1:

Consider the canonical Hamiltonian system on \mathbb{R}^2 specified by the symplectic form $-\mathbf{d}\vartheta$, with $\vartheta = y \mathbf{d}x$, and the Hamiltonian function $\mathcal{H} = y^2/2 + \epsilon x^2/2$. Clearly, this system describes the dynamics of a harmonic oscillator with frequency $\sqrt{\epsilon}$. Equivalently, we can regard \mathcal{H} as describing a perturbed rigid rotor, where $H = y^2/2$ describes the unperturbed dynamics of the rotor, and $\epsilon h = \epsilon x^2/2$ describes the perturbation. This second interpretation allows us to employ our discretizations of Hamilton's principle in phase space that are adapted to perturbed problems to develop a variational integrator for this problem.

In order to identify this integrator, we will calculate L_∞ explicitly using Eq. (37). The ingredients that enter into such a calculation are (i) introducing an affine connection on \mathbb{R}^2 , (ii) finding an expression for the function $f(z_1, z_2)$, (iii) finding an expression for the unperturbed flow map $F_{t,s}$, (iv) finding an expression for for $\Phi_{t,s}^u$, and (v) evaluating the necessary time integrals that

appear in the expression for L_∞ . We will now go through each of these steps in turn.

(i) We will use the obvious connection on \mathbb{R}^2 associated with the standard inner product $\langle u, v \rangle = u_1 v_1 + u_2 v_2$. Relative to this connection and between any pair of points in $z_1, z_2 \in \mathbb{R}^2$, there is a unique geodesic segment equal to the convex hull of $\{z_1, z_2\}$. The unique parameterization of this geodesic segment with parameter $\lambda \in [0, 1]$ and orientation $z_1 \rightarrow z_2$ is given by

$$I(z_2, z_1)(\lambda) = (1 - \lambda) z_1 + \lambda z_2. \quad (67)$$

(ii) The connection chosen in the previous step renders the computation of $f(z_1, z_2)$ analytically tractable. Indeed, we have

$$\begin{aligned} f(z_1, z_2) &= \int_{I(z_2, z_1)} y \, \mathbf{d}x \\ &= \int_0^1 (y_1 + \lambda(y_2 - y_1))(x_2 - x_1) \, d\lambda \\ &= \frac{1}{2}(x_2 - x_1)(y_1 + y_2). \end{aligned} \quad (68)$$

(iii-iv) The unperturbed flow map $F_{t,s}$ is given by

$$F_{t,s}(x, y) = (x + (t - s)y, y). \quad (69)$$

Determining this flow map is a simple matter because the Hamiltonian underlying the unperturbed dynamics is merely $H = y^2/2$. The flow map of $X_{\mathcal{H}}$, $\mathfrak{F}_{t,s}$, is also simple to identify because the dynamics of a simple harmonic oscillator with frequency $\sqrt{\epsilon}$ are very well understood. Indeed, we have

$$\begin{aligned} \mathfrak{F}_{t,s}(x, y) &= \\ &= (x \cos(\sqrt{\epsilon}(t - s)) + \frac{y}{\sqrt{\epsilon}} \sin(\sqrt{\epsilon}(t - s)), y \cos(\sqrt{\epsilon}(t - s)) - \sqrt{\epsilon}x \sin(\sqrt{\epsilon}(t - s))). \end{aligned} \quad (70)$$

On the other hand, the flow map $\Phi_{t,s}^u$ is associated with the time-dependent Hamiltonian

$$\epsilon F_{t,u}^* h_t(x, y) = \epsilon(x + (t - u)y)^2/2, \quad (71)$$

which is not the Hamiltonian for any commonly encountered dynamical system. Nevertheless, it is easy to check that $\Phi_{t,s}^u$ can be expressed in terms of $F_{t,s}$ and $\mathfrak{F}_{t,s}$ according to

$$\Phi_{t,s}^u = F_{u,t} \circ \mathfrak{F}_{t,s} \circ F_{s,u}. \quad (72)$$

Thus, between Eqs. (69) and (70) we have identified an explicit expression for $\Phi_{t,s}^u$ (that we will not display).

(v) Finally, we can calculate L_∞ by directly evaluating Eq. (37). Modulo gauge terms, L_∞ is given by

$$\begin{aligned} L_\infty(z_1, z_2) = & -\frac{1}{2}(y_2x_1 - y_1x_2) \cos(\sqrt{\epsilon}\tau) \\ & -\frac{1}{2} \left(\frac{1}{\sqrt{\epsilon}}y_2y_1 + \sqrt{\epsilon}x_2x_1 \right) \sin(\sqrt{\epsilon}\tau). \end{aligned} \quad (73)$$

It is simple to verify the the discrete Euler-Lagrange equations that follow from this expression for L_∞ are exactly satisfied by the solution to the harmonic oscillator differential equation. This is true regardless of how large or small the time step τ is chosen.

In this case, L_∞ can technically be expanded in powers of ϵ regardless of the value of τ . This follows from the fact that the radius of convergence of the Maclaurin series of either $\sin(x)$ or $\cos(x)$ is infinite. However, when $\sqrt{\epsilon}\tau \ll 1$ these series converge much more rapidly than when $\sqrt{\epsilon}\tau \geq 1$. Thus, we expect that truncating L_∞ 's Maclaurin series in ϵ after only a few terms will lead to a reasonably-accurate integrator for this perturbed rigid rotor only when $\sqrt{\epsilon}\tau \ll 1$, which is precisely the condition that τ be much less than the characteristic bounce time $1/\sqrt{\epsilon}$.

Example 2:

Next, we will summarize the application of the discretizations developed in this work to a non-trivial non-canonical perturbed Hamiltonian system. This system's phase space is \mathbb{R}^2 equipped with the non-canonical symplectic form $-\mathbf{d}\vartheta$, where

$$\vartheta = (x^2 + y^2) (y \mathbf{d}x - x \mathbf{d}y). \quad (74)$$

The Hamiltonian, which is time-dependent and periodic, is given by $\mathcal{H}_t =$

$H + \epsilon h_t$, where

$$H = \frac{2}{9}(x^2 + y^2)^3 \quad (75)$$

$$h_t = x \sin(t) + x^2 \sin(t). \quad (76)$$

This Hamiltonian system can be regarded as a model for magnetic field line flow in a nominally axisymmetric geometry with small resonant perturbations [28]. In this interpretation, the time variable is identified with the toroidal angle, while x and y are identified with a set of Cartesian coordinates in the poloidal plane centered on the unperturbed magnetic axis. We will demonstrate that the variational integrators for this system given by L_0 and L_1 have local $O(\epsilon)$ and $O(\epsilon^2)$ accuracy, respectively. Then we will make some qualitative remarks on the ability of the L_1 integrator to resolve second- and higher-order islands. The time step for the L_0 and L_1 integrators will be set equal to 2π . Thus, these integrators function as the zero'th and first order $t = 0$ Poincaré maps.

In order to find an expression for L_1 , a connection must be chosen; $f(z_1, z_2)$ must be calculated; and the unperturbed flow map must be found. With knowledge of these quantities, Eq.(45) can be evaluated directly to find the desired expression. As in the previous example, we adopt the natural flat connection on \mathbb{R}^2 . Relative to this connection, the function $f(z_1, z_2)$ is given by

$$f(z_1, z_2) = \frac{1}{3}(x_2 y_1 - x_1 y_2)(x_1^2 + x_1 x_2 + x_2^2 + y_1^2 + y_1 y_2 + y_2^2). \quad (77)$$

Finally, the flow map of the dynamical system defined by the Hamiltonian function H is given by

$$F_{t,s}(x, y) = \left(x \cos \left(\frac{1}{3}(x^2 + y^2)(t - s) \right) + y \sin \left(\frac{1}{3}(x^2 + y^2)(t - s) \right), \right. \\ \left. -x \sin \left(\frac{1}{3}(x^2 + y^2)(t - s) \right) + y \cos \left(\frac{1}{3}(x^2 + y^2)(t - s) \right) \right). \quad (78)$$

We will not display the result of using these quantities to calculate L_1 because the resulting expression has many terms, but we must emphasize that this tedious calculation is readily performed using a symbolic manipulation

tool such as *Mathematica*. The laborious task of differentiating such a complicated discrete Lagrangian can be handled by either computing the needed derivatives symbolically before running a simulation or by employing an automatic differentiation tool such as ADOLC [29, 30] at runtime. By employing automatic differentiation tools, implementation and testing of lengthy expressions appearing in the discrete Euler-Lagrange equations and nonlinear solve can be avoided. On the other hand, precomputing derivatives may lead to shorter run times.

Figure 1 shows the results of using the discrete Euler-Lagrange equations associated with L_1 to generate the $t = 0$ Poincaré section for the dynamical system specified by \mathcal{H}_t . Notably, the L_1 -integrator reproduced the first-order islands very well. This can be checked upon noting that the unperturbed frequency as a function of the distance R from the origin in \mathbb{R}^2 is given by $\omega(R) = R^2/3$; first-order perturbation theory predicts an island chain at $R_1 = \sqrt{3}$ and another at $R_2 = \sqrt{3}/2$. On the other hand, as the highlighted portion of the figure indicates, higher-order island chains were not captured correctly by L_1 . This shortcoming is to be expected in light of the fact that the flow map associated with \mathcal{H}_t only satisfies the discrete Euler-Lagrange equations associated with L_1 up to terms second-order in ϵ . The “incoherent” fine-scale structure present in the L_1 integration came as a result of the onset of parasitic modes that generally plague multistep integration methods [19]. The same parasitic modes completely destabilized the L_1 integration run shown in Figure 1 after several tens of thousands of iterations. Smaller values of epsilon can perform larger numbers of iterations before being overtaken by parasitic modes.

Figure 2 illustrates the $O(\epsilon)$ and $O(\epsilon^2)$ errors of the integrators provided by L_0 and L_1 , respectively. The L_1 integrator decreases in error quadratically with ϵ until reaching the error tolerance of the nonlinear Newton-Rhpson solver of 10^{-12} . Decreasing ϵ beyond 10^{-6} therefore shows no further improvement in the error of L_1 .

7. Discussion

In the preceding sections, we presented and applied the first discretizations of Hamilton’s principle in phase space that are adapted to perturbed noncanonical Hamiltonian systems. Notably, these discretizations function as variational integrators with $O(1)$ time steps and local $O(\epsilon^N)$ accuracy, where N is any desired nonnegative integer. Moreover, for each discretization, our

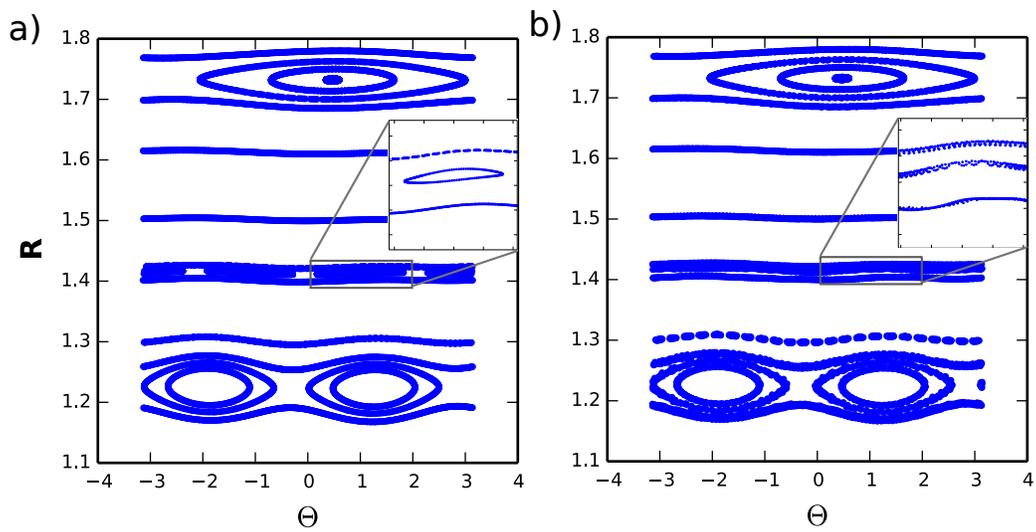


Figure 1: (a) The $t = 0$ Poincaré section for \mathcal{H}_t calculated using a Runge-Kutta integrator with very fine temporal resolution. (b) The same Poincaré section computed using the discrete Euler-Lagrange equations associated with L_1 . In each case, $\epsilon = .0075$ and the variables R and Θ denote the standard polar coordinates on \mathbb{R}^2 . The large first-order islands are reproduced well by the L_1 -integrator, while, as the highlighted portions of the figures indicate, higher-order islands are not properly reproduced [31].

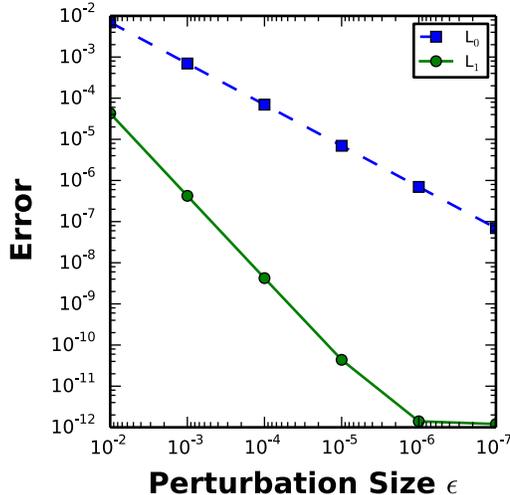


Figure 2: “Error” denotes the distance between the predictions of a well-converged 2π -second Runge-Kutta integration and a single iteration of the L_0 or L_1 integrator. Each integration method was initialized with the same initial condition.

expression for the associated discrete Lagrangian is manifestly coordinate independent. This coordinate independence has been achieved by introducing an arbitrary affine connection on the phase space. Thus, our discretizations may prove to be useful for constructing variational integrators on manifolds.

We have also shown that if these two-step integrators are initialized using the smooth modified system studied in backward error analysis, then they function as one-step algorithms that preserve symplectic forms on the phase space M . We expect that this result will play an important role in the backward error analysis of the variational discretizations developed here. For instance, using the fact that the smooth modified system preserves a symplectic form on M , it should be possible to show that the method for choosing the second initial condition for two-step methods championed in [19, 21] leads to better energy behavior than a more conventional Runge-Kutta-based initialization.

While the integrators identified here formally apply to any perturbed non-canonical Hamiltonian system, they are *much easier* to apply to nearly-integrable systems because the unperturbed flow map can often be determined analytically in these cases. When the unperturbed flow map is an-

alytically unattainable, it would have to be determined numerically. This would entail devising some numerical scheme for evaluating (at least in an approximate sense) the various integrations along unperturbed orbits that appear in the discrete Lagrangians presented in Section 4. Whether or not such a scheme exists that does not involve prohibitively large computational overhead is currently unknown to us. We leave investigating this issue to future work.

A theoretical application of the integrators developed here that we will pursue in the future is *coarse-graining* Hamilton's principle in phase space. Specifically, we would like to derive the stochastic action mentioned in Ref. [32] by directly manipulating Hamilton's principle in phase space. Our hope is that this result will follow by appropriately rescaling time and then looking at the behavior of $\mathbb{S}_{(z_1, z_2)}^2$ (Eq. (46)) as $\epsilon \rightarrow 0$.

8. Acknowledgements

The authors would like to express their gratitude to A. I. Zhmoginov for his help in editing this manuscript. This work was supported by the U.S. Department of Energy under contract DE-AC02-09CH11466.

References

- [1] L. D. Landau, E. M. Lifshitz, Mechanics, Elsevier Butterworth-Heinemann, 1976.
- [2] V. I. Arnold, Mathematical Methods of Classical Mechanics, Springer, 1989.
- [3] R. Abraham, J. Marsden, Foundations of Mechanics, AMS Chelsea publishing, American Mathematical Soc., 2008.
- [4] D. D. Holm, J. E. Marsden, T. S. Ratiu, The Euler-Poincaré equations and semidirect products with applications to continuum theories, Adv. Math 137 (1998) 1.
- [5] A. J. Brizard, T. S. Hahm, Foundations of nonlinear gyrokinetic theory, Rev. Mod. Phys. 79 (2007) 421–468.
- [6] J. Cary, A. Brizard, Hamiltonian theory of guiding-center motion, Rev. Mod. Phys. 81 (2009) 693.

- [7] J. E. Marsden, M. West, Discrete mechanics and variational integrators, *Acta Numer.* 10 (2001) 357.
- [8] H. Aref, Point vortex dynamics: A classical mathematics playground, *J. Math. Phys.* 48 (2007) 5401.
- [9] H. Qin, X. Guan, Variational symplectic integrator for long-time simulations of the guiding-center motion of charged particles in general magnetic fields, *Phys. Rev. Lett.* 100 (2008) 035006.
- [10] H. Qin, X. Guan, W. M. Tang, Variational symplectic algorithm for guiding center dynamics and its application in tokamak geometry, *Phys. Plasmas* 16 (2009) 042510.
- [11] C. W. Rowley, J. E. Marsden, Variational integrators for degenerate Lagrangians, with application to point vortices, 41st IEEE Conference on Decision and Control 40 (2002) 1521.
- [12] J. Li, H. Qin, Z. Pu, L. Xie, S. Fu, Variational symplectic algorithm for guiding center dynamics in the inner magnetosphere, *Phys. Plasmas* 18 (2011) 052902.
- [13] S. Lall, M. West, Discrete variational Hamiltonian mechanics, *J. Phys. A: Math. Gen.* 39 (2006) 5509.
- [14] M. Leok, J. Zhang, Discrete Hamiltonian variational integrators, *IMA J. Numer. Anal.* 31 (2011) 1497.
- [15] W. M. Farr, Variational integrators for almost integrable systems, *Celest. Mech. Dyn. Astron.* 103 (2009) 105.
- [16] R. I. McLachlan, Composition methods in the presence of small parameters, *BIT Numer. Math.* 35 (1995) 258.
- [17] J. E. Chambers, M. A. Murison, Pseudo-high-order symplectic integrators, *Astron. J.* 119 (2000) 425.
- [18] J. Laskar, P. Robutel, High order symplectic integrators for perturbed Hamiltonian systems, *Celest. Mech. Dyn. Astron.* 80 (2001) 39.
- [19] E. Hairer, Backward error analysis for multistep methods, *Numer. Math.* 84 (1999) 199–232.

- [20] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration, Springer, 2006.
- [21] C. L. Ellison, J. W. Burby, J. M. Finn, H. Qin, W. M. Tang, Initializing and stabilizing multistep algorithms for modeling dynamical systems, 2013. J. Comput. Phys. (to be submitted).
- [22] G. Dahlquist, Convergence and stability in the numerical integration of ordinary differential equations, Math. Scand. 4 (1956) 33–53.
- [23] M. M. Postnikov, Geometry VI: Riemannian Geometry, Encyclopaedia of Mathematical Sciences, Springer, 2001. URL: <http://books.google.com/books?id=1kyUImXf8U0C>.
- [24] J. P. Boyd, The devil’s invention: Asymptotic, superasymptotic, and hyperasymptotic series, Acta Appl. Math. 56 (1999) 98.
- [25] U. Kirchgraber, Multi-step methods are essentially one-step methods, Numer. Math 48 (1986) 85–90.
- [26] J. Squire, H. Qin, W. M. Tang, Gauge properties of the guiding center variational symplectic integrator, Phys. Plasmas 19 (2012) 052501.
- [27] J. Vankerschaver, M. Leok, A novel formulation of point vortex dynamics on the sphere: Geometrical and numerical aspects, J. Nonlinear Sci. July (2013) 01–37.
- [28] J. R. Cary, R. G. Littlejohn, Noncanonical Hamiltonian mechanics and its application to magnetic field line flow, Ann. Phys. 151 (1983) 1.
- [29] A. Griewank, A. Walther, Evaluating Derivatives: Principles and Techniques, SIAM: Frontiers in Mathematics, 2009.
- [30] A. Walther, A. Griewank, Getting started with ADOL-C, from Combinatorial Scientific Computing, Chapman-Hall CRC Computational Science, 2012.
- [31] J. D. Hunter, Matplotlib: A 2d graphics environment, Computing In Science & Engineering 9 (2007) 90–95.
- [32] J. W. Burby, A. I. Zhmoginov, H. Qin, Hamiltonian mechanics of stochastic acceleration, Phys. Rev. Lett. 111 (2013) 195001.