Numer. Math. (2009) 113:243–264
DOI 10.1007/s00211-009-0245-3

Numerische Mathematik

Error analysis of variational integrators of unconstrained Lagrangian systems

George W. Patrick · Charles Cuell

Received: 9 July 2008 / Revised: 26 April 2009 / Published online: 8 July 2009

Abstract An error analysis of variational integrators is obtained, by blowing up the discrete variational principles, all of which have a singularity at zero time-step. Divisions by the time step lead to an order that is one less than observed in simulations, a deficit that is repaired with the help of a new past–future symmetry.

Mathematics Subject Classification (2000) 65L05 · 49S05 · 70H · 37J

1 Introduction

We consider a regular Lagrangian system $L\colon \boldsymbol{T}Q \to \mathbb{R}$, with Euler-Lagrange equations

$$-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} + \frac{\partial L}{\partial q^i} = 0.$$

Constraints, if present, are holonomic and are incorporated into the configuration manifold Q. Standard numerical integrators are insensitive to the properties of such systems because they discretize Euler-Lagrange equations as they would any other differential equations.

G. W. Patrick is funded by the Natural Sciences and Engineering Reseach Council, Canada.

G. W. Patrick (✉) · C. Cuell
Department of Mathematics and Statistics, Applied Mathematics and Mathematical Physics,
University of Saskatchewan, Saskatoon, Saskatchewan, S7N 5E6, Canada
e-mail: patrick@math.usask.ca

C. Cuell
e-mail: cuell@math.usask.ca

Variational integrators [7,8,12] rather discretize Hamilton's variational principle

$$\delta \int_a^b L\left(q'(t)\right)\, dt = 0, \quad q(a) \ \text{and} \ \ q(b) \ \text{constant}$$

($q'(t) \in \boldsymbol{T}Q$ includes both coordinates q^i and v^i in this notation). Such discretizations can be obtained by fixing a time step $h > 0$ and approximating the integral with a finite sum, resulting in a finite dimensional constrained optimization problem—a *discrete Hamilton's principle*. The function under the sum, called the *discrete Lagrangian*, is an approximation to the type 1 generating function

$$S_h(q^+, q^-) = \int_0^h L\left(F_t\left(\Delta_h(q^+, q^-)\right)\right)\, dt$$

where $\Delta_h(q^+, q^-)$ is the initial velocity at q^- that arrives at $q^+ \approx q^-$ after time h, and F_t is the flow of the Euler-Lagrange equations. The critical points of the discrete Hamilton's principle satisfy discrete versions of the Euler-Lagrange equations, which then define a numerical integrator for the original Lagrangian system.

Marsden and West [8] consider the local existence and uniqueness, and the local error analysis, of variational integrators. But the error analysis is problematic: the map $\Delta_h(q^+, q^-)$ is singular as $h \to 0^+$, since arbitrarily large velocities are required to go from fixed q^- to fixed q^+ in vanishingly small time. We show (Sect. 2) that terms not controlled by the definitions of the order of a discrete Lagrangian system, enter in the discrete Euler-Lagrange equations, and require additional argumentation beyond that in [8].

Here we provide a variational proof of existence and uniqueness, and complete the error analysis, for variational integrators of (regular) Lagrangian systems $L\colon \boldsymbol{T}Q \to \mathbb{R}$. To systematically treat the singularity, we use the discretizations of Cuell and Patrick [5], which

1. use finite curve segments to discretize the tangent bundle of Q; and
2. place the discrete variational principle in the velocity phase space $\boldsymbol{T}Q$.

This shifts the singularity at $h = 0$ to a degeneracy in the map that sends the curve segments to their endpoints. To prove existence and uniqueness (Theorem 3.7), we blow up the variational principle at $h = 0$, converting it to a smooth perturbation of the trivial, nonsingular optimization problem, with objective $L(v_q) + L(w_q)$ and constraint $v_q + w_q = \text{constant}$. Divisions by h lead the arguments to an order that is one less than observed in simulations. But the past and the future occur symmetrically in the blown-up variational principle, giving a new $\mathbb{Z}_2$ symmetry (exchange of v_q and w_q), from which we prove the observed order follows by a nontrivial cancellation (Theorem 4.7). A similar cancellation occurs in the error analysis in [9], but that development is restricted to type 1 generating functions of kinetic-plus-potential systems, and it is framed non-variationally, on cotangent bundles.

For discrete Lagrangian systems, we obtain semi-global existence and uniqueness: Theorem 3.7 asserts a well defined discrete evolution for arbitrarily high velocities, if h is sufficiently small i.e. the discrete evolution is defined for (h, v) in an open neighborhood of $\{0\} \times \boldsymbol{T}\mathcal{Q}$. We anchor our arguments directly to the discrete variational principles, avoiding use of auxiliary constructions, such as the discrete Euler-Lagrange equations, or canonical contexts on the cotangent bundles. These are important and useful, but it is also important and useful to obtain the discrete variational picture stand-alone, as a coherent whole. For all these purposes, and also to make clear the geometry underlying the error analysis, our development is coordinate free.

Viewing discrete tangent vectors as curve segments is generally consistent with viewing discretizations in general as attaching to a manifold finite rather than infinitesimal objects [1,2]. Our discrete tangent bundles are similar to the groupoid oriented constructions of discrete phase spaces for Lagrangian systems in [11]. Placing variational principles in phase space, as opposed to configuration spaces is also advanced as the Hamilton-Pontryagin variational principles of [3,13,14]. In this work, the extra kinematic freedom in the discrete tangent bundle phase space is necessary for existence of the limit of the discrete Lagrangian system as $h \to 0^+$. The same idea was exploited in [10] to obtain local existence of and uniqueness Lagrangian systems directly from the variational principles.

We extensively use the development in [4]. In this paper, all manifolds are assumed paracompact and Hausdorff.

2 Preview by example

The following example-oriented preview provides a good understanding of our motivations, and of the results of the general theory in Sects. 3 and 4.

2.1 Definition

We consider $\mathcal{Q} \equiv \mathbb{R} = \{q\}$, with $\boldsymbol{T}\mathcal{Q} = \mathbb{R}^2 = \{(q, v)\}$ and Lagrangian

$$L \equiv \frac{1}{2} g(q) v^2 - V(q), \tag{1}$$

where $g(q) > 0$ and $V(q)$ are smooth functions. Such systems describe a particle moving in Euclidean space under the influence of a potential, and constrained to a curve $x(q)$, where, as is easily verified, $g(q) = |dx/dq|^2$.

The discrete phase space for this system is $\mathcal{Q} \times \mathcal{Q} = \{(q^+, q^-)\}$, and a discrete Lagrangian may be defined by

$$L_h(q^+, q^-) \equiv hL\left(q^-, \frac{q^+ - q^-}{h}\right) + bh(q^+ - q^-), \tag{2}$$

where b is a constant that is necessary for the purpose of the example. The discrete action on a sequence $q_0, q_1, q_2 \in \mathcal{Q}$ is

$$S_h \equiv L_h(q_1, q_0) + L_h(q_2, q_1). \tag{3}$$

This data defines a variational integrator: given an initial discrete state (q_1, q_0), a time h advanced state (q_2, q_1) is computed from the requirement that (q_0, q_1, q_2) is a critical point of S_h, under the constraints that q_0 and q_2 are not varied. This is the standard setup of discrete mechanics [8,12].

2.2 Order

By standard definition, a numerical integrator

$$y_{i+1} = F(y_i, t_i), \quad t_{i+1} = t_i + h,$$

of differential equation $y' = f(y, t)$ is order r if the local truncation error $y_1(y_0, t_0 + h) - y(y_0, t_0 + h)$ is $O(h^{r+1})$ as $h \to 0$.

Substituting into L_h the flow of Euler-Lagrange equations i.e.

$$q^- = q, \quad q^+ = q + vt + O(t^2)$$

gives, after putting $t = h$,

$$L_h(q, v) = hL(q, v) + bh^2 v + O(h^2) = hL(q, v) + O(h^2). \tag{4}$$

The integral of L along the flow is

$$\int_0^h L\left(q + O(t), v + O(t)\right) dt = hL(q, v) + O(h^2). \tag{5}$$

According to Equation 2.3.1 of [8], L_h is by definition order 1 since (4) and (5) agree through order h^1. Item (1) of Theorem 2.3.1 of [8] asserts that the variational integrator defined by L_h should be order 1.

Indeed, the critical points of S_h are exactly the solutions of the *discrete Euler-Lagrange equations*

$$\boldsymbol{F}^- L_h(q_2, q_1) = \boldsymbol{F}^+ L_h(q_1, q_0) \tag{6}$$

where

$$\boldsymbol{F}^+ L_h \equiv \frac{\partial L_h}{\partial q^+}, \quad \boldsymbol{F}^- L_h \equiv -\frac{\partial L_h}{\partial q^-}$$

are the *discrete Legendre transforms*. One computes

$$\begin{aligned} \boldsymbol{F}^{+}L_h &= \frac{1}{h}g(q^-)(q^+ - q^-) + bh, \\ \boldsymbol{F}^{-}L_h &= \frac{1}{h}g(q^-)(q^+ - q^-) + hV'(q^-) + bh - \frac{1}{2h}g'(q^-)(q^+ - q^-)^2, \end{aligned}$$

after which Eq. (6) becomes

$$\frac{1}{h}g(q_1)(q_2 - q_1) + hV'(q_1) + bh - \frac{1}{2h}g'(q_1)(q_2 - q_1)^2 = \frac{1}{h}g(q_0)(q_1 - q_0) + bh. \tag{7}$$

Assuming $q_1 - q_0 = O(h)$, this implies

$$\frac{q_2 - 2q_1 + q_0}{h^2} = \frac{-1}{g(q_0)}\left(V'(q_0) + \frac{1}{2h^2}g'(q_0)(q_1 - q_0)^2\right) + O(h),$$

which is evidently order 1, after comparison with the Euler-Lagrange equations of L defined by (1), which are

$$\frac{d^2q}{dt^2} = \frac{-1}{g(q)}\left(V'(q) + \frac{1}{2}g'(q)v^2\right).$$

Item (1) in Theorem 2.3.1 of [8] asserts that the discrete evolution should be order 1, and it is.

2.3 Discrete Legendre transforms

Something, however, is not quite right: the term $bh(q^+ - q^-)$ in the discrete Lagrangian L_h is not controlled by the definition of an order 1 discrete Lagrangian, because it enters Eq. (4) only at order 2. In the discrete Euler Lagrange equations (7) such terms appeared twice, and at the same order as the potential, which they could have significantly altered. They appeared symmetrically on both sides of (7), and therefore did not affect the consistency of the variational integrator. Nevertheless, they did affect the separate terms appearing as discrete Legendre transforms in the discrete Euler-Lagrange equations. Item 2 in Theorem 2.3.1 of [8] asserts that the discrete Legendre transforms $\boldsymbol{F}^{+}L_h$ and $\boldsymbol{F}^{-}L_h$ should, after the substitutions $q = q^-$, $q = q^- + hv$, be order 1 consistent with the exact Legendre transform of L. They are not.

In the discrete Lagrangian, the substitutions $q = q^-$, $q = q^- + hv$ shift the term $bh(q^+ - q^-)$ to order 2, because they enforce $q^+ - q^- = O(h)$. So *that* term does not in the definition affect the order 1 consistency of the discrete Lagrangians. However, *that* term does affect the discrete Legendre transforms at order 1, because the derivatives in the Legendre transform remove $q^+ - q^-$. Therefore the definition of an order 1 discrete Lagrangian does not control the order 1 consistency of the discrete Legendre transforms. *The fault may be assigned to the* $1/h$ *singularity:* It is necessary to clear that singularity in the definition of the order of discrete Lagrangians, because the definition of order is necessarily about $h = 0$. When the singularity is cleared, the term $bh(q^+ - q^-)$ is knocked out of the definition, but it affects the discrete equations of motion. The situation is saved only because bh occurs in the discrete equations of motion as a pair of identical twins.

Suppose we view variational integrators just as the discrete Euler-Lagrange equations, discarding the variation principle altogether. Possibly for some purpose of numerical implementation, imagine being motivated to use different discrete Lagrangians in the separate terms of the discrete Euler-Lagrange equations i.e. to use rather than (6), the equations

$$\boldsymbol{F}^- L_h(q_2, q_1) = \boldsymbol{F}^+ \tilde{L}_h(q_1, q_0),$$

where $L \neq \tilde{L}$. Then the order of the corresponding numerical integrator would almost certainly be reduced by 1, despite the fact that each Lagrangian would separately satisfy all prescribed order conditions. And, in the case of order 1 Lagrangians, the result would almost certainly be an inconsistent integrator. For example, choose L_h and $\tilde{L}_h$ as Eq. (2), with different values of a.

2.4 Example analysis: blow up

To analyze the discrete variational principle as $h \to 0$, change to new variables $q, v, \tilde{q}, \tilde{v}$ defined by

$$q \equiv q_0, \quad v \equiv \frac{q_1 - q_0}{h}, \quad \tilde{q} \equiv q_1, \quad \tilde{v} \equiv \frac{q_2 - q_1}{h}. \tag{8}$$

There are four new variables but originally there were only q_0, q_1, q_2, so there will be a new constraint. Defining

$$\partial_h^-(q, v) \equiv q, \quad \partial_h^+(q, v) \equiv q + hv, \tag{9}$$

one has, from (8),

$$\partial_h^-(q, v) = q_0, \quad \partial_h^+(q, v) = q_1 = \partial_h^-(\tilde{q}, \tilde{v}), \quad \partial_h^+(\tilde{q}, \tilde{v}) = q_2. \tag{10}$$

The new constraint is the middle of these; the variations in the discrete variational principle are restricted to annihilate the derivatives of the outer two. The action becomes

$$S_h \equiv L_h(q, v) + L_h(\tilde{q}, \tilde{v}), \quad L_h(q, v) = hL(q, v) + bh^2 v, \tag{11}$$

which, together with the constraints (10), is an equivalent variational principle on $\boldsymbol{T}Q \times \boldsymbol{T}Q = \{(q, v), (\tilde{q}, \tilde{v})\}$.

We will desingularized the variational principle defined by (10) and (11), in three stages:

1. The critical points are unchanged by multiplication of S_h by any constant. So L_h in (11) can be replaced with its division by h. This results in a replacement of S_h with its division by h, although Formula (11) for S_h is unchanged. The variational principle, objective and constraints, becomes

$$\begin{cases} \hat{S}_h \equiv \hat{L}_h(q, v) + \hat{L}_h(\tilde{q}, \tilde{v}), \quad \hat{L}_h(q, v) \equiv L(q, v) + bh\, v, \\ \partial_h^-(q, v) = q_0, \quad \partial_h^+(q, v) = q_1 = \partial_h^-(\tilde{q}, \tilde{v}), \quad \partial_h^+(\tilde{q}, \tilde{v}) = q_2. \end{cases} \tag{12}$$

2. The middle constraint of (12) is $q + hv = \tilde{q}$. This constraint can be imposed by substituting $\tilde{q}$, resulting in a principle on the variables $(q, v, \tilde{v})$. The variational principle becomes

$$\begin{cases} \hat{S}_h \equiv \hat{L}_h(q, v) + \hat{L}_h(q + hv, \tilde{v}), \quad \hat{L}_h(q, v) \equiv L(q, v) + bh\, v, \\ q = q_0, \quad q + hv + h\tilde{v} = q_2. \end{cases} \tag{13}$$

3. At $h = 0$, the fixed endpoint constraints of (13) degenerate to the same function, namely $(q, v, \tilde{v}) \mapsto q$. To remove this degeneracy, these can be post-composed by any smooth bijection, such as

$$\bar{q} = \frac{q_0 + q_2}{2}, \quad z = \frac{q_2 - q_0}{h}.$$

   The variational principle becomes

$$\begin{cases} \hat{S}_h \equiv \hat{L}_h(q, v) + \hat{L}_h(q + hv, \tilde{v}), \quad \hat{L}_h(q, v) \equiv L(q, v) + bh\, v, \\ q + \dfrac{h}{2}(v + \tilde{v}) = \bar{q}, \quad v + \tilde{v} = z. \end{cases} \tag{14}$$

The desingularization is complete because the principle (14) is smooth and, as will be seen, nonsingular, through $h = 0$. The *blown-up variational principle* is obtained from (14) by substituting $h = 0$ i.e.

$$\hat{S}_0 \equiv L(\bar{q}, v) + L(\bar{q}, \tilde{v}), \quad v + \tilde{v} = z. \tag{15}$$

For small h, the variational principle (14) may be regarded as a continuous perturbation of (15).

2.5 Example analysis: discrete existence and uniqueness

Re-introducing the details of the example Lagrangian (1), the blown-up variational principle becomes

$$\hat{S}_0 = \frac{g(\bar{q})}{2}(v^2 + \tilde{v}^2) - 2V(\bar{q}), \quad v + \tilde{v} = z. \tag{16}$$

Substituting $\tilde{v} = z - v$ and differentiating gives, since $\bar{q}$ is constant in (16),

$$\frac{d\hat{S}_0}{dv} = g(\bar{q})(2v - z) = 0,$$

so, given z, there is the solution $v = \tilde{v} = z/2$. The second derivative at this critical point is

$$\frac{d^2\hat{S}_0}{dv^2} = 2g(\bar{q}),$$

which is positive, so the critical point is nondegenerate. Nondegenerate critical points vary smoothly with parameters and constraint values. In particular, for any $\bar{q}_0$ and z_0, the original principle (10) and (11) has a unique critical point $(v, \tilde{v}) = \gamma(h, \bar{q}, z)$ for all sufficiently small h, and all $\bar{q} \approx \bar{q}_0$ and $z \approx z_0$. γ is smooth and at $h = 0$ its graph is an open set of the diagonal of $\boldsymbol{T}Q \times \boldsymbol{T}Q$ i.e. the graph of the identity map. So, for sufficiently small h the graph of γ, and hence the variational principle, locally defines a near-identity time-advance map of $\boldsymbol{T}Q$ to itself i.e. a discrete evolution. For the general results, see Definitions 3.1 and 3.3, and Theorem 3.7.

2.6 Example analysis: order

Error analyses of variational integrators has purpose to give conditions on the basic data of the discrete variational principles that are sufficient to imply that the corresponding numerical integrators are order r. This follows if:

1. exact discrete variational principles are obtained that give the exact solution of the continuous Lagrangian system; and
2. two variational principles that are order $r + 1$ consistent have solutions that are order $r + 1$ consistent.

A discrete variational principle is then defined order $r + 1$ if its objective and constraints are order $r + 1$ consistent with those of the exact discrete variational principle, and then Item (2) implies that the corresponding numerical integrator is order r.

Item (1) seems not to be clarified by specializing to Example (2), but rather seems more clear in the geometric setting; see Lemma 4.2 and Theorem 4.3. Essentially, in gross aspect, solutions of an exact discrete variational principle correspond to cornered solutions of the continuous variational principle. However, by the well known Weierstrass-Erdman conditions [6], regular variational principles do not have corners i.e. critical points are unaffected if the space of solutions is enlarged to allow corners. Therefore, solutions of the continuous variational principle are solutions of the exact discrete variational principle. But then the critical points of both principles are the same because of uniqueness i.e. the solutions of the discrete principle are exact.

Item (2) *is* clarified by the example context. The desingularized principle (14) has discrete action

$$\hat{S}_h = L(q, v) + L(q + hv, \tilde{v}) + bh\, v + bh\, \tilde{v}.$$

At a solution $h = 0$ and $v = \tilde{v}$, $\hat{S}_h$ is symmetric under the exchange $v \leftrightarrow \tilde{v}$, not just at order h^0, but also at order h^1, because the original minus the exchanged is

$$\begin{aligned}\hat{S}_h(q, v, \tilde{v}) - \hat{S}_h(q, \tilde{v}, v) &= L(q, v) + L(q + hv, \tilde{v}) - L(q, \tilde{v}) - L(q + h\tilde{v}, v)\\ &= L(q, v) - L(q, \tilde{v}) + h\frac{\partial L}{\partial q}(q, \tilde{v})v - h\frac{\partial L}{\partial q}(q, v)\tilde{v} + O(h^2),\end{aligned}$$

which is $O(h^2)$ at $\tilde{v} = v$. So under the constraint $\tilde{q} = q + hv$ of (10), the term $bh(q^+ - q^-)$ of (2) affects the solutions of the discrete variational principles only symmetrically at order h^1. Considering the solution of the discrete principle as a perturbation from $h = 0$ i.e. a perturbation of the graph of the identity map, the effect on the solutions is symmetric to order h^1. But such a symmetric alteration does not change, at the same order, the function defined by the graph. In particular, the solutions of the variational principle are unaffected, at order h^1, by the term $bh(q^+ - q^-)$. For the general results, see Definitions 4.1 and 4.6, and Theorem 4.8.

3 Discrete existence and uniqueness

Our discretizations of Lagrangian systems depend on discretizations of tangent bundles of manifolds $\mathcal{M}$, by assignment of curve segments in $\mathcal{M}$ to tangent vectors of $\mathcal{M}$ [5]. We require a parameter h such that $\boldsymbol{T}\mathcal{M}$ is obtained in the limit $h \to 0^+$, so we posit a map $\psi(h, t, v_m)$, $m \in \mathcal{M}$, $v_m \in \boldsymbol{T}\mathcal{M}$, with values in $\mathcal{M}$, and obtain the curve segments $t \mapsto \psi(h, t, v_m)$. The definition builds in some flexibility. The curve segments are generated as the variable t ranges over intervals of length h, with otherwise unrestrained endpoints: for example, $[0, h]$ and $[-h/2, h/2]$ are common choices which are both accommodated.

Definition 3.1 A C^k *discretization of* $\boldsymbol{T}\mathcal{M}$, $k \geq 1$, is a tuple $(\psi, \alpha^+, \alpha^-)$, where

$$\psi : U \subseteq \mathbb{R}^2 \times \boldsymbol{T}\mathcal{M} \to \mathcal{M}, \quad \alpha^+ : [0, a) \to \mathbb{R}_{\geq 0}, \quad \alpha^- : [0, a) \to \mathbb{R}_{\leq 0},$$

are such that

1. ψ is continuous, U is open, and $\{0\} \times \{0\} \times \boldsymbol{T}\mathcal{M} \subseteq U$;
2. α^+, α^- are C^1, and $\alpha^+(h) - \alpha^-(h) = h$;
3. $\psi(h, 0, v_m) = m$, and $\dfrac{\partial \psi}{\partial t}(h, 0, v_m) = v_m$;
4. the *boundary maps* defined by

$$\partial_h^-(v_m) \equiv \psi\left(h, \alpha^-(h), v_m\right), \quad \partial_h^+(v_m) \equiv \psi\left(h, \alpha^+(h), v_m\right), \tag{17}$$

   are C^k in (h, v_m) and

$$\left.\frac{d}{dh}\right|_{h=0} \partial_h^+(v_m) = \dot{\alpha}^+ v_m, \quad \left.\frac{d}{dh}\right|_{h=0} \partial_h^-(v_m) = \dot{\alpha}^- v_m \tag{18}$$

where

$$\dot{\alpha}^+ \equiv \frac{d\alpha^+}{dh}(0), \quad \dot{\alpha}^- \equiv \frac{d\alpha^-}{dh}(0).$$

Remark 3.2 Putting $h = 0$ in $\alpha^+(h) - \alpha^-(h) = h$ gives $\alpha^+(0) = \alpha^-(0) = 0$ because $\alpha^+(0) \geq 0$ and $\alpha^-(0) \leq 0$. If ψ is a C^1 map in all its variables then Eq. (18) are superfluous because they follow by differentiating Eq. (17). Note that at $h = 0$, $\partial_h^+ = \partial_h^- = \tau_Q$, where $\tau_Q : \boldsymbol{T}Q \to Q$ is the canonical projection, and differentiating $\alpha^+(h) - \alpha^-(h) = h$ at $h = 0$ gives $\dot{\alpha}^+ - \dot{\alpha}^- = 1$. Definition 3.1 corresponds to Eq. (9) of the example in Sect. 2 if $\psi(h, t, (q, v)) \equiv q + tv$, $\alpha^-(h) = 0$, and $\alpha^+(h) = h$, so that

$$\partial_h^-(q, v) = q + \alpha^-(h)v = q, \quad \partial_h^+(q, v) = q + \alpha^+(h)v = q + hv.$$

□

Given a discretization of the tangent bundle of configuration space, one only need add an appropriate discrete Lagrangian to obtain a discretization of a Lagrangian system. See [5] for more explanation, and a general development along the lines used in this paper, of discrete Lagrangian mechanics and discretizations of Lagrangian systems, including the nonholonomic case. Here we do not require the full generality, and what is required, specialized to the holonomic case, is collected in Definition 3.3.

Definition 3.3 A C^k *discretization*, $k \geq 1$, of a Lagrangian system $L : \boldsymbol{T}Q \to \mathbb{R}$, is a tuple $(L_h, \psi, \alpha^+, \alpha^-)$ where

1. $(\psi, \alpha^+, \alpha^-)$ is a C^k discretization of $\boldsymbol{T}Q$; and
2. $L_h : \boldsymbol{T}Q \to \mathbb{R}$ is C^k in (h, v_q) and satisfies $L_h(v_q) = hL(v_q) + O(h^2)$.

$(h, v_0, \tilde{v}_0) \in \mathbb{R} \times \boldsymbol{T}Q \times \boldsymbol{T}Q$ is *critical* if $v = v_0$, $\tilde{v} = \tilde{v}_0$ is a critical point of the variational principle

$$\begin{cases} S_h \equiv L_h(v) + L_h(\tilde{v}), \\ \partial_h^-(v) \text{ and } \partial_h^+(\tilde{v}) \text{ constant, and } \partial_h^+(v) = \partial_h^-(\tilde{v}). \end{cases}$$

A *discrete evolution* is a map F defined on an open subset of $\mathbb{R} \times \boldsymbol{T}Q$ such that $(h, v, F(h, v))$ is critical for all (h, v) in the domain of F.

Remark 3.4 Precisely, $(h, v_0, \tilde{v}_0)$ is critical if

$$\boldsymbol{d}S_h(v, \tilde{v})(\delta v, \delta\tilde{v}) = 0 \quad \text{and} \quad \partial_h^+(v) = \partial_h^-(\tilde{v})$$

for all $\delta v \in \boldsymbol{T}_v Q$ and $\delta\tilde{v} \in \boldsymbol{T}_{\tilde{v}} Q$ such that

$$\boldsymbol{T}\partial_h^-(\delta v) = 0, \quad \boldsymbol{T}\partial_h^+(\delta\tilde{v}) = 0, \quad \boldsymbol{T}\partial_h^+(\delta v) = \boldsymbol{T}\partial_h^-(\delta\tilde{v}).$$

We have placed the discrete variational principle in sequences of velocity phase space, so there is a discrete analogue of the *first order constraint* $q(t)' = v(t)$ where

$q(t) = \tau_Q \circ v(t)$ i.e. the successive curve segments must join to make a continuous whole. The fixed endpoint constraints correspond to $\partial_h^-(v)$ and $\partial_h^+(\tilde{v})$ constant, which affect only the variations; the actual constraint values are not specified, providing, as is usual in both discrete and continuous Lagrangian mechanics, the freedom to accommodate initial conditions. □

Remark 3.5 Suppose that $L_h^{Q\times Q}(q^+, q^-)$ is a discrete Lagrangian as defined in [8] and let ψ be a discretization of Q as in Definition 3.3. As shown in [5], $\Psi_h(v) \equiv (\partial_h^+(v), \partial_h^-(v))$ is a diffeomorphism from an open subset of $\boldsymbol{T}Q$ to an open subset of $Q \times Q$. $L_h \equiv L_h^{Q\times Q} \circ \Psi_h$ is a discretization as in Definition 3.3, and conversely, any such discretization defines a discrete Lagrangian as in [8] by $L_h^{Q\times Q} \equiv L_h \circ \Psi_h^{-1}$. In this way, the discrete mechanics on $\boldsymbol{T}Q$ and $Q \times Q$ are equivalent because they are conjugated by Ψ. For more details, see Sect. 5. □

At $h = 0$, the discrete action is singular, and the constraints are degenerate, because

1. at $h = 0$, $L_h(v) = 0$, so $S_h(v, \tilde{v}) = L_h(v) + L_h(\tilde{v}) = 0$; and
2. at $h = 0$, $\partial_h^+(v) = \partial_h^-(\tilde{v})$ is $\tau_Q(v) = \tau_Q(\tilde{v})$, and, on that constraint, the fixed endpoint constraints $\partial_h^-(v) = q^-$ and $\partial_h^+(\tilde{v}) = \tilde{q}^+$ are replicates.

The necessary blow-ups rely on a technical result of [4] that we recall as Proposition 3.6 below. This is an invariant version of the elementary calculus fact, a version of L'Hospital's rule, that if $\hat{f}(x) \equiv f(x)/h(x)$ where $f(x)$ and $h(x)$ are C^1 functions such that $f(0) = h(0) = 0$ and $h'(0) \neq 0$, then $\hat{f}(x)$ can be continuously extended through $x = 0$ by defining $\hat{f}(0) = f'(0)/h'(0)$.

If $\pi : E \to \mathcal{M}$ is a vector bundle, and $z \in \boldsymbol{T}_{0_m}E$, then we denote the horizontal and vertical parts of z by hor $z \in \boldsymbol{T}_m\mathcal{M}$ and vert $z \in E_m$, respectively. We denote the zero section of E by $0(E)$. Also, the statement of Proposition 3.6 uses the convention that a pair $(\mathcal{M}, h_\mathcal{M})$ is called a *manifold* when $\mathcal{M}$ is a manifold and $h_\mathcal{M} : \mathcal{M} \to \mathbb{R}$ is a submersion.

Proposition 3.6 *Let $(\mathcal{M}, h_\mathcal{M})$ and $\mathcal{N}$ be a manifolds, and let $\pi : E \to \mathcal{N}$ be a vector bundle. Suppose that $f : U \subseteq \mathcal{M} \to E$ is C^k, $k \geq 1$, and that $f(m) \in 0(E)$ whenever $h_\mathcal{M}(m) = 0$. Then for all m such that $h_\mathcal{M}(m) = 0$, there is a unique $e(m) \in E_{\pi(f(m))}$ such that*

$$\text{vert}\, \boldsymbol{T}_m f(v_m) = (dh_\mathcal{M}(m) v_m)\, e(m), \quad v_m \in \boldsymbol{T}_m\mathcal{M}.$$

Moreover, the function $\hat{f} : U \to E$ defined by

$$\hat{f}(m) \equiv \begin{cases} \dfrac{f(m)}{h_\mathcal{M}(m)}, & h_\mathcal{M}(m) \neq 0, \\ e(m), & h_\mathcal{M}(m) = 0, \end{cases}$$

is C^{k-1}.

Theorem 3.7 is our main result on discrete existence and uniqueness. This goes beyond just finding regularity conditions on the discrete Lagrangian that imply existence and uniqueness, because it only relies on regularity of the Lagrangian of the continuous system—it is an analysis of the limit $h \to 0$. For the discrete analogue of regularity in our context, and the corresponding existence and uniqueness result, see [5], and in the standard context on $\mathcal{Q} \times \mathcal{Q}$, Theorem 1.5.1 of [8].

Recall that the first and second fiber derivatives of $L\colon \boldsymbol{T}\mathcal{Q} \to \mathbb{R}$ are $\boldsymbol{F}L(v_q) = \boldsymbol{D}(L|\boldsymbol{T}_q\mathcal{Q})(v_q)$ and $\boldsymbol{F}^2L(v_q) = \boldsymbol{D}^2(L|\boldsymbol{T}_q\mathcal{Q})(v_q)$, and that L is called *regular* if $\boldsymbol{F}^2L$ has values only in the nondegenerate quadratic forms on the fibers of $\boldsymbol{T}\mathcal{Q}$. If A is a set, then we denote by $\Delta(A \times A)$ the diagonal of $A \times A$.

Theorem 3.7 *Let $(L_h, \psi, \alpha^+, \alpha^-)$ be a C^k discretization of a regular Lagrangian system $L\colon \boldsymbol{T}\mathcal{Q} \to \mathbb{R}$, $k \geq 2$. Then there are neighborhoods $W \subseteq \mathbb{R} \times \boldsymbol{T}\mathcal{Q}$ of $\{0\} \times \boldsymbol{T}\mathcal{Q}$ and $U \subseteq \mathbb{R} \times \boldsymbol{T}\mathcal{Q} \times \boldsymbol{T}\mathcal{Q}$ of $\{0\} \times \Delta(\boldsymbol{T}\mathcal{Q} \times \boldsymbol{T}\mathcal{Q})$ such that, for all $(h, v) \in W$, $h > 0$, there is a unique $\tilde{v} \in \boldsymbol{T}\mathcal{Q}$ such that $(h, v, \tilde{v}) \in U$ and $(h, v, \tilde{v})$ is critical. Moreover, U and W can be chosen such that $F\colon W \to \boldsymbol{T}\mathcal{Q}$ defined by*

$$F(h, v) \equiv \begin{cases} \tilde{v}, & h > 0, \\ v, & h = 0, \end{cases}$$

is C^{k-1}.

Proof The blow-up of L_h is immediate: set

$$\hat{L}(h, v_q) = \begin{cases} \dfrac{1}{h} L_h(v_q), & h \neq 0, \\ L(v_q), & h = 0. \end{cases}$$

$\hat{L}$ is C^{k-1} by Proposition 3.6.

The constraints are blown-up by imposing $\partial_h^+(v) = \partial_h^-(\tilde{v})$, after which $\partial_h^-(v)$ and $\partial_h^+(\tilde{v})$ are $O(h)$ close and their difference can be usefully divided by h. For this, observe that both ∂_h^+ and ∂_h^- are submersions on $\boldsymbol{T}\mathcal{Q}$ when $h = 0$, since $\partial_0^+(v_q) = q$ and $\partial_0^-(v_q) = q$, so there is a neighborhood $A \supseteq \{0\} \times \boldsymbol{T}\mathcal{Q}$ on which both ∂_h^+ and ∂_h^- are submersions. Consequently

$$\mathcal{C} \equiv \left\{(h, v, \tilde{v}) : (h, v) \in A, (h, \tilde{v}) \in A, \partial_h^+(v) = \partial_h^-(\tilde{v})\right\}$$

is a submanifold of $\mathbb{R} \times \boldsymbol{T}\mathcal{Q} \times \boldsymbol{T}\mathcal{Q}$. Also, there is a tubular neighborhood

$$\zeta : W^{0(E)} \subset E \to W^{\mathcal{Q}\times\mathcal{Q}} \subset \mathcal{Q} \times \mathcal{Q} = \left\{(q^+, q^-)\right\}$$

of the normal bundle $E \equiv \left\{(v_q, -v_q) : v_q \in \boldsymbol{T}\mathcal{Q}\right\}$ to the diagonal $\Delta(\mathcal{Q} \times \mathcal{Q})$ of $\mathcal{Q} \times \mathcal{Q}$, which satisfies

$$\operatorname{vert} \boldsymbol{T}\zeta^{-1}(v_q^+, v_q^-) = \left(\frac{1}{2}(v_q^+ - v_q^-), \frac{1}{2}(v_q^- - v_q^+)\right). \tag{19}$$

The purpose of ζ^{-1} is to compute the difference between two nearby elements of $\mathcal{Q}$. For example, if $\mathcal{Q} = \mathbb{R}^n$ we can use $\zeta(v_q, -v_q) \equiv (q, q) + (v_q, -v_q)$, and then

$$\zeta^{-1}(q^+, q^-) = (v_q, -v_q), \quad v_q = \left(\frac{q^+ + q^-}{2}, \frac{q^+ - q^-}{2}\right).$$

i.e. the fiber part of E corresponds to the difference. Just below, in the definition of $\hat{\varphi}$, scalar multiplication of E by $1/h$ will be used to blow up the difference. See the proof Proposition 1.9 of [5] for more details about arranging Eq. (19).

Define $\hat{\varphi} \colon \mathcal{C} \to \mathbb{R} \times E$ by

$$\hat{\varphi}(h, v, \tilde{v}) = \begin{cases} \left(h, \dfrac{1}{h}\zeta^{-1}\left(\partial_h^+(\tilde{v}), \partial_h^-(v)\right)\right), & h \neq 0, \\ \left(h, \dfrac{1}{2}(v + \tilde{v}, -v - \tilde{v})\right), & h = 0. \end{cases}$$

and define φ and $h_{\mathcal{C}}$ on $\mathcal{C}$ by

$$\varphi(h, v, \tilde{v}) \equiv \zeta^{-1}\left(\partial_h^+(\tilde{v}), \partial_h^-(v)\right), \quad h_{\mathcal{C}}(h, v, \tilde{v}) \equiv h.$$

If $(h, v, \tilde{v}) \in \mathcal{C}$ and $h = 0$ then $\tau_{\mathcal{Q}}(v) = \partial_0^+(v) = \partial_0^-(\tilde{v}) = \tau_{\mathcal{Q}}(\tilde{v})$, so for all $(h, v, \tilde{v}) \in \mathcal{C}$, $\varphi(h, v, \tilde{v}) = 0$ if $h_{\mathcal{C}}(h, v, \tilde{v}) = 0$. By Proposition 3.6, for all $v, \tilde{v}$ there is a unique $e(v, \tilde{v})$ such that

$$\begin{aligned} \operatorname{vert} \boldsymbol{T}\varphi(0, v, \tilde{v})(\delta h, \delta v, \delta\tilde{v}) &= (e(v, \tilde{v}), -e(v, \tilde{v}))\, \boldsymbol{d}h_{\mathcal{C}}(0, v, \tilde{v})(\delta h, \delta v, \delta\tilde{v}) \\ &= (e(v, \tilde{v}), -e(v, \tilde{v}))\, \delta h, \end{aligned}$$

for all $(\delta h, \delta v, \delta\tilde{v}) \in \boldsymbol{T}_{(0,v,\tilde{v})}\mathcal{C}$. By Items 3 and 4 of Definition 3.1, $(\delta h, \delta v, \delta\tilde{v}) \in \boldsymbol{T}_{(0,v,\tilde{v})}\mathcal{C}$ if and only if

$$\boldsymbol{T}\tau_{\mathcal{Q}}(\delta v) + \delta h\, \dot{\alpha}^+ v = \boldsymbol{T}\tau_{\mathcal{Q}}(\delta\tilde{v}) + \delta h\, \dot{\alpha}^- \tilde{v} \tag{20}$$

and using Eq. (19), and the definition of φ,

$$\operatorname{vert} \boldsymbol{T}\varphi(0, v, \tilde{v})(\delta h, \delta v, \delta\tilde{v}) = \frac{1}{2}(w, -w),$$

where

$$w \equiv \boldsymbol{T}\tau_{\mathcal{Q}}(\delta\tilde{v}) + \delta h\, \dot{\alpha}^+ \tilde{v} - \boldsymbol{T}\tau_{\mathcal{Q}}(\delta v) - \delta h\, \dot{\alpha}^- v.$$

It follows that $e(v, \tilde{v})$ can be found by imposing

$$\frac{1}{2}\left(\boldsymbol{T}\tau_{\mathcal{Q}}(\delta\tilde{v}) + \delta h\, \dot{\alpha}^+ \tilde{v} - \boldsymbol{T}\tau_{\mathcal{Q}}(\delta v) - \delta h\, \dot{\alpha}^- v\right) = \delta h\, e(v, \tilde{v}) \tag{21}$$

for all $\delta h, \delta v$, and $\delta\tilde v$ which satisfy Eq. (20). Using (20) to replace $\boldsymbol{T}\tau_Q(\delta\tilde v)-\boldsymbol{T}\tau_Q(\delta v)$ in Eq. (21) gives

$$\frac{1}{2}\left(\delta h\,\dot\alpha^+ v-\delta h\,\dot\alpha^-\tilde v+\delta h\,\dot\alpha^+\tilde v-\delta h\,\dot\alpha^- v\right)=\frac{1}{2}\delta h\,(v+\tilde v)=\delta h\,e(v,\tilde v)$$

so $e(v,\tilde v)=\frac{1}{2}(\tilde v+v)$, after which the definition of φ at $h=0$, and Proposition 3.6 applied to φ, imply $\hat\varphi$ is C^{k-1}.

If $h>0$, then finding the critical points of the discrete variational principle is equivalent to finding the critical points of $\hat L|\hat\varphi^{-1}(h,z_q,-z_q)$. If $h=0$, then the latter is the problem of finding the critical points $(v_0,\tilde v_0)$ of $L(v)+L(\tilde v)$, $v,\tilde v\in \boldsymbol{T}_qQ$ subject to the constraint $\frac{1}{2}(v+\tilde v)=z_q$. These are the v and $\tilde v=2z_q-v$ such that $L(v)+L(2z_q-v)$ has a critical point at v i.e. such that

$$\boldsymbol{F}L(v)-\boldsymbol{F}L(\tilde v)=0,\quad \tilde v=2z_q-v.$$

This has the solution $v=\tilde v=z_q$, at which the Hessian of $L(v)+L(2z_q-v)$ is $2\boldsymbol{F}^2L(v)$, which is nondegenerate. Thus there is a manifold of nondegenerate critical points parametrized by $z_q\in\boldsymbol{T}Q$. Semiglobal persistence of these critical points follows by Theorem 2 of [4] i.e. there are neighborhoods $\hat U\supseteq\{0\}\times\triangle(\boldsymbol{T}Q\times\boldsymbol{T}Q)$ and $\hat V\supseteq\{0\}\times E$, and a C^{k-1} map $\hat\gamma\colon\hat V\to\hat U$, such that for all $(h,z_q,-z_q)\in\hat V$, $\hat\gamma(h,z_q,-z_q)$ is the unique critical point in $\hat U$ of $\hat L|\hat\varphi^{-1}(h,z_q,-z_q)$.

At $h=0$, $\hat\gamma(h,z_q,-z_q)=(0,z_q,z_q)$, and the image of $\hat\gamma$ forms the graph of the identity map of $\boldsymbol{T}Q$. Consequently, for small h, $\hat\gamma$ determines a map F because $\hat\gamma$ has image a graph. The technical statements in the Theorem to this effect are immediate from Proposition 5 of [4], applied to the map $\pi_{23}\circ\hat\gamma$, where $\pi_{23}(h,v,\tilde v)$ $=(v,\tilde v)$. □

Remark 3.8 The proof of Theorem 3.7 shows the blow-up at $h=0$ of the discrete variational principles gives the variational principles with action $L(v)+L(\tilde v)$, where v and $\tilde v$ are constrained (1) to be in the same fiber of $\boldsymbol{T}Q$, and (2) such that $v+\tilde v$ is constant. The blown-up variational principle is past-future symmetric i.e. symmetric under the exchange of v and $\tilde v$. □

4 Order

If the curve segments of the discretizations of the tangent bundle of Q are obtained from the base integral curves of the Euler-Lagrange vector field X_E, and the discrete Lagrangian is the classical action, then we obtain the *exact discretizations* of Marsden and West [8]. Exact discretizations are important because they exactly generate the flow of the Euler-Lagrange equation of the continuous Lagrangian system (Theorem 1.6.4 of [8]), so that the order of a discretization can be controlled by reference to its order of consistency with an exact discretization.

Definition 4.1 An *exact discretization* of a Lagrangian system $L\colon\boldsymbol{T}Q\to\mathbb{R}$ is a discretization $(L_h,\psi,\alpha^+,\alpha^-)$ where ψ and L_h satisfy

1. $\psi(h, t, v_q) \equiv \tau_Q(F_t^{X_E}(v_q))$, where $F_t^{X_E}$ is the flow of X_E; and
2. $L_h(v_q) \equiv \int_{\alpha^-(h)}^{\alpha^+(h)} L \circ \frac{\partial \psi}{\partial t}(h, t, v_q)\, dt$.

Some understanding of exact discretizations can be anchored to the variational principles. For an exact discretization, curve segment discretizations lead to a picture of piecewise smooth solutions built of segments of the exact flow. The difference between an exact discrete and continuous variational principles is that the discrete principle allows for piecewise smooth solutions—it is, after all, constructed from segments. But the well known Weierstrass-Erdman conditions [6] imply that additional critical curves are *not* obtained by allowing corners. For example, corners do *not* occur in Riemannian geodesics because they form local triangles for which there would be a shorter path along the hypotenuse than two of the sides. Lemma 4.2 and Theorem 4.3 provide the formal statements and proofs along these variational lines. These results also follow from Theorem 1.6.4 of [8], and the equivalence of our context and the standard $Q \times Q$ context, as discussed in Sect. 5.

Lemma 4.2 *Let* $(L_h, \psi, \alpha^+, \alpha^-)$ *be a* C^k *exact discretization of the* C^k *Lagrangian system* $L\colon \boldsymbol{T}Q \to \mathbb{R}$, $k \geq 2$, *suppose that the integral curve of* X_E *through* $v \in \boldsymbol{T}Q$ *is defined for times in* $[\alpha^-(h), \alpha^-(h) + 2h]$, *and set* $\tilde{v} \equiv F_{\alpha^+(h)}^{X_E}(v)$. *Then* $(h, v, \tilde{v})$ *is critical.*

Proof The variational derivative of the action

$$S \equiv \int_a^b L\left(q'(t)\right)\, dt$$

can be written [5,7] as

$$\boldsymbol{d}S\left(q(t)\right) \cdot \delta q(t) = \int_a^b \delta L\left(\frac{d^2 q}{dt^2}\right) \cdot \delta q\, dt + \boldsymbol{F}L\left(\frac{dq}{dt}\right) \delta q(t)\bigg|_a^b \tag{22}$$

where δL is locally

$$\delta L = \left(-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} + \frac{\partial L}{\partial q^i}\right) dq^i.$$

Assuming δv and $\delta v'$ satisfy the constraints

$$\boldsymbol{T}\partial_h^-(\delta v) = 0, \quad \boldsymbol{T}\partial_h^+(\delta \tilde{v}) = 0, \quad \boldsymbol{T}\partial_h^+(\delta v) = \boldsymbol{T}\partial_h^-(\delta \tilde{v}), \tag{23}$$

and remembering that $\delta L = 0$ along a solution, one applies Eq. (22) to each of the integrals in

$$S_h(v,\tilde v) = \int_{\alpha^-(h)}^{\alpha^+(h)} L \circ \left(F_t^{X_E}(v)\right) dt + \int_{\alpha^-(h)}^{\alpha^+(h)} L \circ \left(F_t^{X_E}(\tilde v)\right) dt,$$

obtaining, from Eq. (23) and $F^{X_E}_{\alpha^+(h)}(v) = F^{X_E}_{\alpha^-(h)}(\tilde v)$,

$$\boldsymbol{d}S_h(v,\tilde v)(\delta v,\delta\tilde v) = \boldsymbol{F}L\left(F^{X_E}_{\alpha^+(h)}(v)\right)\boldsymbol{T}\partial_h^+(\delta v) - \boldsymbol{F}L\left(F^{X_E}_{\alpha^-(h)}(\tilde v)\right)\boldsymbol{T}\partial_h^-(\delta\tilde v) = 0.$$

□

Theorem 4.3 *Let $(L_h,\psi,\alpha^+,\alpha^-)$ be an exact discretization of a regular Lagrangian system $L\colon \boldsymbol{T}Q \to \mathbb{R}$. Then there is a neighborhood $U \subseteq \mathbb{R}\times\boldsymbol{T}Q\times\boldsymbol{T}Q$ of $\{0\}\times\triangle(\boldsymbol{T}Q\times\boldsymbol{T}Q)$ such that, for all $(h,v,\tilde v)\in U$ with $h>0$, $\tilde v = F_h^{X_E}(v)$ if and only if $(h,v,\tilde v)$ is critical.*

Proof Let $W \subseteq \mathbb{R}\times\boldsymbol{T}Q$ and $U \subseteq \mathbb{R}\times\boldsymbol{T}Q\times\boldsymbol{T}Q$ be as in the statement of Theorem 3.7. Possibly by shrinking W, one can assume that

1. for all $(h,v)\in W$, the integral curve of X_E through v is defined for times in $[\alpha^-(h),\alpha^-(h)+2h]$; and
2. $(h,v,F_h^{X_E}(v))\in U$ for all $(h,v)\in W$.

If $(h,v,\tilde v)\in U$ and $\tilde v = F_h^{X_E}(v)$ then $(h,v,\tilde v)$ is critical by Lemma 4.2. Conversely, if $(h,v,\tilde v)$ is critical then so is $(h,v,F_h^{X_E}(v))$, so $\tilde v = F_h^{X_E}(v)$ by the uniqueness in Theorem 3.7. □

We will avoid coordinates in order computations on manifolds by using the calculus of residuals developed in [4]. Suppose $(\mathcal{M},h_{\mathcal{M}})$ and $\mathcal{N}$ are manifolds. If $f_i\colon\mathcal{M}\to\mathcal{N}$, $i=1,2$, are such that $f_1=f_2$ on $h_{\mathcal{M}}^{-1}(0)$, then define $f_2 = f_1 + O(h^r_{\mathcal{M}})$, $r\ge 1$ if, for all $m_0\in h_{\mathcal{M}}^{-1}(0)$, there is a chart ν at $n_0\equiv f_i(m_0)\in\mathcal{N}$, and there is a function $(\delta f)_\nu$ defined near m_0, and continuous at m_0, such that

$$\nu\left(f_2(m)\right) - \nu\left(f_1(m)\right) = h_{\mathcal{M}}(m)^r(\delta f)_\nu(m),$$

for all m in some neighborhood of m_0. As is easily shown, $(\delta f)_\nu(m_0)$ transforms as a tangent vector as ν is varied, and therefore defines an element $\operatorname{res}^r(f_2,f_1)(m_0)\in\boldsymbol{T}_{f(m_0)}\mathcal{N}$, called the *residual*. Proposition 4.4, which is a specialization of Proposition 3 of [4] is the key result used to compute residuals without the invocation of local charts.

Proposition 4.4 *Let $(\mathcal{M},h_{\mathcal{M}})$, $(\mathcal{N},h_{\mathcal{N}})$, and $\mathcal{P}$ be manifolds, and suppose $f_i\colon\mathcal{M}\to\mathcal{N}$ and $g_i\colon\mathcal{N}\to\mathcal{P}$, $i=1,2$ are C^1 and satisfy $h_{\mathcal{N}}\circ f_i = h_{\mathcal{M}}$, $f_2 = f_1 + O(h^r_{\mathcal{M}})$, and $g_2 = g_1 + O(h^r_{\mathcal{N}})$. Then $g_2\circ f_2 = g_1\circ f_1 + O(h^r_{\mathcal{M}})$. Moreover, if $h_{\mathcal{M}}(m)=0$ and $n\equiv f_i(m)$, then*

$$\operatorname{res}^r(g_2\circ f_2, g_1\circ f_1)(m) = \operatorname{res}^r(g_2,g_1)(n) + \boldsymbol{T}_n g_1\operatorname{res}^r(f_2,f_1)(m).$$

As has been stated, the central issue is a decrease in the order of accuracy, essentially due to a division by h. Proposition 4.4, which is yet another result of [4], tracks this in the context of Proposition 3.6.

Proposition 4.5 *Let* $(\mathcal{M}, h_{\mathcal{M}})$ *and* $\mathcal{N}$ *be a manifolds, let* $\pi : E \to \mathcal{N}$ *a vector bundle, and suppose* f_i *and* $\hat{f}_i$ *are as in Proposition* 3.6*, with* $k \geq r$*. Then* $\hat{f}_2 = \hat{f}_1 + O(h_{\mathcal{M}}^{r-1})$ *if* $f_2 = f_1 + O(h_{\mathcal{M}}^{r})$*,* $r \geq 2$*. Moreover,* $\operatorname{res}^r(f_2, f_1)$ *takes values in the vertical bundle of* E *and* $\operatorname{res}^{r-1}(\hat{f}_2, \hat{f}_1) = \operatorname{res}^r(f_2, f_1)$.

We come to the main objective, accuracy, which is the order to which an evolution map of a given discretization of a Lagrangian system agrees with the continuous flow of that Lagrangian system. This can be approached by analyzing the order that two discretizations agree, since, by Theorem 4.3, the continuous flow is obtained from the exact discretizations. Together with the equivalences of Sect. 5, this suffices to repair the proof of (3) implies (1) in Theorem 2.3.1 of [8]

Definition 4.6 Two discretizations $(L_h^i, \psi^i, \alpha^+, \alpha^-)$, $i = 1, 2$ have *order r contact* if $\psi^2(h, t, v) = \psi^1(h, t, v) + O(t^{r+1})$ and $L_h^2(v) = L_h^1(v) + O(h^{r+1})$.

Theorem 4.7 *Let* F_h^i *be evolution maps of two discretizations* $(L_h^i, \psi^i, \alpha^+, \alpha^-)$*,* $i = 1, 2$ *of a regular Lagrangian system* $L\colon \boldsymbol{T}Q \to \mathbb{R}$*. Then* $F_h^2(v) = F_h^1(v) + O(h^{r+1})$ *if* $(L_h^i, \psi^i, \alpha^+, \alpha^-)$ *have order r contact.*

Proof Assume the context and notations of the proof of Theorem 3.7: in summary,

$$\hat{L}^i(h, v_q) = \begin{cases} \dfrac{1}{h} L_h^i(v_q), & h \neq 0, \\ L(v_q), & h = 0, \end{cases}$$

are C^{k-1}, $U^i \subseteq \mathbb{R} \times \boldsymbol{T}Q \times \boldsymbol{T}Q$ is open, and

$$\mathcal{C}^i = \left\{(h, v, \tilde{v}) \in U^i : \partial_h^{i+}(v) = \partial_h^{i-}(\tilde{v})\right\}$$

are submanifolds of $\mathbb{R} \times \boldsymbol{T}Q \times \boldsymbol{T}Q$. Also, E and ζ are defined, and $\hat{\varphi}^i : \mathcal{C}^i \to \mathbb{R} \times E$ by

$$\hat{\varphi}^i(h, v, \tilde{v}) = \begin{cases} \left(h, \dfrac{1}{h}\zeta^{-1}\left(\partial_h^{i+}(\tilde{v}), \partial_h^{i-}(v)\right)\right), & h \neq 0, \\ \left(h, \dfrac{1}{2}(v + \tilde{v}, -v - \tilde{v})\right), & h = 0. \end{cases}$$

$\hat{\gamma}^i : V^i \to U^i$, where $V^i \subseteq \mathbb{R} \times E$ is open, and $\hat{\gamma}^i(h, z_q, -z_q)$ is the unique critical point in $\mathcal{C}^i$ of $\hat{L}|\hat{\varphi}^{-1}(h, z_q, -z_q)$. The maps F^i are constructed as the graphs of $\pi_{23} \circ \hat{\gamma}^i$ where $\pi_{23} = (\pi_2, \pi_3)$ and $\pi_2, \pi_3 \colon \mathbb{R} \times \boldsymbol{T}Q \times \boldsymbol{T}Q \to \boldsymbol{T}Q$ are the projections onto the second and third factors, respectively. Proposition 6 of [4] implies maps agree to the same order as their graphs, and one order higher if their residuals are symmetric.

So it suffices to show that $\pi_{23}\circ\hat\gamma^2 = \pi_{23}\circ\hat\gamma^1 + O(h^r)$ and that $\mathrm{res}^r(\pi_{23}\circ\hat\gamma^2, \pi_{23}\circ\hat\gamma^1)$ is symmetric.

To establish the contact order of $\pi_{23}\circ\hat\gamma^i$, the basic data of the corresponding critical point problems has to be compared. For that it is inconvenient that the manifolds $\mathcal{C}^i$ depend on i. Let Θ^i be maps from $\mathbb{R}\times \boldsymbol{T}\mathcal{Q}\times\boldsymbol{T}\mathcal{Q}$ to itself which have the following properties:

1. $\Theta^2 = \Theta^1 + O(h^{r+1})$;
2. Θ^i, $i = 1, 2$ are the identity on $\{0\}\times\boldsymbol{T}\mathcal{Q}\times\boldsymbol{T}\mathcal{Q}$;
3. Θ^i, $i = 1, 2$ have nonsingular derivatives on $\{0\}\times\boldsymbol{T}\mathcal{Q}\times\boldsymbol{T}\mathcal{Q}$;
4. $\tau_{\mathcal{Q}}\circ\pi_2\circ\Theta^i(h, v, \tilde v) = \partial_h^{i+}(v)$ and $\tau_{\mathcal{Q}}\circ\pi_3\circ\Theta^i(h, v, \tilde v) = \partial_h^{i-}(\tilde v)$.

For example, we can use a metric on $\mathcal{Q}$ the parallel transport $\mathbb{P}_{q_2,q_1}$ along geodesics between nearby points of $\mathcal{Q}\times\mathcal{Q}$ to define

$$\Theta^i(h, v, \tilde v) \equiv \left(h, \mathbb{P}_{\partial_h^{i+}(v),\tau_{\mathcal{Q}}(v)}(v), \mathbb{P}_{\partial_h^{i-}(\tilde v),\tau_{\mathcal{Q}}(\tilde v)}(\tilde v)\right).$$

The purpose of the maps Θ^i is to normalize the submanifolds $\mathcal{C}^i$. In particular, by transporting the vectors defining the curve segments to common connection points, each Θ^i maps $\mathcal{C}^i$ diffeomorphically to an open submanifold of $\mathbb{R}\times\boldsymbol{T}\mathcal{Q}\oplus\boldsymbol{T}\mathcal{Q}$ where

$$\boldsymbol{T}\mathcal{Q}\oplus\boldsymbol{T}\mathcal{Q} = \left\{(v, \tilde v)\in\boldsymbol{T}\mathcal{Q}\times\boldsymbol{T}\mathcal{Q} : \tau_{\mathcal{Q}}(v) = \tau_{\mathcal{Q}}(\tilde v)\right\}$$

is the Whitney direct sum.

Since $\Theta^i(0, v, \tilde v) = (0, v, \tilde v)$ and Θ^i is a local diffeomorphism on $\{0\}\times\boldsymbol{T}\mathcal{Q}\times\boldsymbol{T}\mathcal{Q}$, Theorem 1 of [4] implies that Θ^i can be assumed to be a diffeomorphism between neighborhoods $\bar U^i, \hat U^i \supseteq \{0\}\times\boldsymbol{T}\mathcal{Q}\times\boldsymbol{T}\mathcal{Q}$. Set

$$\bar S^i \equiv \hat S^i\circ(\Theta^i)^{-1}, \quad \bar\varphi^i \equiv \hat\varphi^i\circ(\Theta^i)^{-1}, \quad \bar\gamma^i \equiv \Theta^i\circ\hat\gamma^i,$$

where $\hat S(h, v, \tilde v) = \hat L(h, v) + \hat L(h, \tilde v)$. $\bar\gamma(h, z_q, -z_q)$ is the unique critical point of $\bar S^i|(\mathbb{R}\times(\boldsymbol{T}\mathcal{Q}\oplus\boldsymbol{T}\mathcal{Q}))$ in $\bar U^i$ subject to the constraint $\bar\varphi = (h, z_q, -z_q)$. By the consistency of L_h^i, the consistency of ψ^i, the blow-up constructions of Theorem 3.7, and by Proposition 4.5, $\bar S_2 = \bar S_1 + O(h^r)$ and $\bar\varphi = \bar\varphi + O(h^r)$, so $\Theta^2\circ\bar\gamma^2 = \Theta^1\circ\bar\gamma^1 + O(h^r)$. Also, since $\Theta^2 = \Theta^1 + O(h^{r+1})$, $\mathrm{res}^r(\hat\gamma^2, \hat\gamma^1) = \mathrm{res}^r(\bar\gamma^2, \bar\gamma^1)$. So it is sufficient to show that $\mathrm{res}^r(\pi_{23}\circ\bar\gamma^2, \pi_{23}\circ\bar\gamma^1)$ is symmetric i.e. that

$$\boldsymbol{T}\pi_2\,\mathrm{res}^r(\bar\gamma^2, \bar\gamma^1) = \boldsymbol{T}\pi_3\,\mathrm{res}^r(\bar\gamma^2, \bar\gamma^1).$$

From Remark 3.8, setting $h = 0$, $\Theta^i\circ\bar\gamma^i(0, z_q, -z_q)$ is the solution of the variational problem of finding the critical points of $L(v) + L(\tilde v)$ with the constraints $v, \tilde v \in \boldsymbol{T}_q\mathcal{Q}$ and $\frac{1}{2}(v+\tilde v) = z_q$. This variational problem admits the $\mathbb{Z}_2$ symmetry $(v, \tilde v)\mapsto(\tilde v, v)$ and the solution is $\Theta^i\circ\hat\gamma^i(0, z_q, -z_q) = (0, z_q, z_q)$ (and therefore all solutions occur on the fixed point set of the $\mathbb{Z}_2$ action). So it suffices to show that

$\operatorname{res}^r(\bar{S}^2, \bar{S}^1)$ and $\operatorname{res}^r(\bar{\varphi}^2, \bar{\varphi}^1)$ are symmetric i.e.

$$\begin{aligned}\operatorname{res}^r(\bar{S}^2, \bar{S}^1)(0, v, \tilde{v}) &= \operatorname{res}^r(\bar{S}^2, \bar{S}^1)(0, \tilde{v}, v),\\ \operatorname{res}^r(\bar{\varphi}^2, \bar{\varphi}^1)(0, v, \tilde{v}) &= \operatorname{res}^r(\bar{\varphi}^2, \bar{\varphi}^1)(0, \tilde{v}, v).\end{aligned} \tag{24}$$

The first of (24) is immediate: $\operatorname{res}^r(\hat{S}^2, \hat{S}^1)$ is symmetric since $\hat{S}^1$ and $\hat{S}^2$ are, and

$$\begin{aligned}\operatorname{res}^r(\hat{S}^2, \hat{S}^1) &= \operatorname{res}^r(\bar{S}^2 \circ \Theta^2, \bar{S}^1 \circ \Theta^1)\\ &= \operatorname{res}^r(\bar{S}^2, \bar{S}^1) \circ \Theta^1 + \boldsymbol{T}\bar{S}^1 \operatorname{res}^r(\Theta^2, \Theta^1)\\ &= \operatorname{res}^r(\bar{S}^2, \bar{S}^1) \circ \Theta^1.\end{aligned}$$

The proof of the second of (24) begins with the observation that

$$(\mathbf{1}, \tau_Q, \tau_Q) \circ \Theta^i(h, v, \tilde{v}) = \left(h, \partial_h^{i+}(v), \partial_h^{i-}(\tilde{v})\right),$$

so that, after defining the involution

$$\sigma(v, \tilde{v}) \equiv (\tilde{v}, v),$$

we have

$$\left(h, \partial_h^{i+}(\tilde{v}), \partial_h^{i-}(v)\right) = (\mathbf{1}, \tau_Q, \tau_Q) \circ \Theta^i \circ (\mathbf{1}, \sigma)(h, v, \tilde{v}),$$

and hence (by abuse of notation $\pi_{23}(h, q^+, q^-) = (q^+, q^-)$)

$$\pi_{23} \circ \bar{\varphi}^i = \frac{1}{h}\zeta^{-1} \circ (\tau_Q \circ \pi_2, \tau_Q \circ \pi_3) \circ \Theta^i \circ (\mathbf{1}, \sigma) \circ (\Theta^i)^{-1}.$$

Using Proposition 4.5,

$$\begin{aligned}&\operatorname{vert}\operatorname{res}^r(\pi_2 \circ \bar{\varphi}^2, \pi_{23} \circ \bar{\varphi}^1)(0, v, \tilde{v})\\ &\quad= \operatorname{vert}\operatorname{res}^{r+1}\Big(\zeta^{-1} \circ (\tau_Q \circ \pi_2, \tau_Q \circ \pi_3) \circ \Theta^2 \circ (\mathbf{1}, \sigma) \circ (\Theta^2)^{-1},\\ &\qquad \zeta^{-1} \circ (\tau_Q \circ \pi_2, \tau_Q \circ \pi_3) \circ \Theta^1 \circ (\mathbf{1}, \sigma) \circ (\Theta^1)^{-1}\Big)(0, v, \tilde{v})\\ &\quad= \operatorname{vert}\boldsymbol{T}\zeta^{-1}\,\boldsymbol{T}(\tau_Q \circ \pi_2, \tau_Q \circ \pi_3)\Big(\operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, \tilde{v}, v)\\ &\qquad -\boldsymbol{T}(\mathbf{1}, \sigma)\operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, v, \tilde{v})\Big).\end{aligned}$$

To compute the outer part of this, note that, for δq^+ and δq^- in the same fiber of $\boldsymbol{T}Q$,

$$\begin{aligned}\operatorname{vert}\boldsymbol{T}\zeta^{-1}(\delta q^+, \delta q^-) &= \frac{1}{2}(\delta q^+ - \delta q^-, -\delta q^+ + \delta q^-)\\ &= \frac{1}{2}(\mathbf{1} - \sigma)(\delta q^+, \delta q^-),\end{aligned} \tag{25}$$

and since

$$\begin{aligned}
&\sigma \circ \boldsymbol{T}(\tau_Q \circ \pi_2, \tau_Q \circ \pi_3) \left(\operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, \tilde{v}, v) - \boldsymbol{T}(\mathbf{1}, \sigma) \operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, v, \tilde{v})\right) \\
&\quad = \boldsymbol{T}(\tau_Q \circ \pi_2, \tau_Q \circ \pi_3) \boldsymbol{T}(\mathbf{1}, \sigma) \left(\operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, \tilde{v}, v)\right. \\
&\qquad \left. - \boldsymbol{T}(\mathbf{1}, \sigma) \operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, v, \tilde{v})\right) \\
&\quad = T(\tau_Q \circ \pi_2, \tau_Q \circ \pi_3) \left(\boldsymbol{T}(\mathbf{1}, \sigma) \operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, \tilde{v}, v) - \operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, v, \tilde{v})\right),
\end{aligned}$$

it follows from (25) that

$$\begin{aligned}
&\operatorname{vert} \operatorname{res}^r(\pi_2 \circ \bar{\varphi}^2, \pi_3 \circ \bar{\varphi}^1)(0, v, \tilde{v}) \\
&\quad = \frac{1}{2} \boldsymbol{T}(\tau_Q \circ \pi_2, \tau_Q \circ \pi_3) \Big(\operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, v, \tilde{v}) + \operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, \tilde{v}, v) \\
&\qquad - \boldsymbol{T}(\mathbf{1}, \sigma) \operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, v, \tilde{v}) - \boldsymbol{T}(\mathbf{1}, \sigma) \operatorname{res}^{r+1}(\Theta^2, \Theta^1)(0, \tilde{v}, v)\Big),
\end{aligned}$$

which is symmetric i.e.

$$\operatorname{vert} \operatorname{res}^r(\pi_2 \circ \bar{\varphi}^2, \pi_3 \circ \bar{\varphi}^1)(0, v, \tilde{v}) = \operatorname{vert} \operatorname{res}^r(\pi_2 \circ \bar{\varphi}^2, \pi_3 \circ \bar{\varphi}^1)(0, \tilde{v}, v).$$

This implies that $\operatorname{res}^r(\pi_2 \circ \bar{\varphi}^2, \pi_3 \circ \bar{\varphi}^1)(0, \tilde{v}, v)$ is symmetric, since by Proposition 4.5 that is vertical anyway, so equality of the vertical parts is sufficient for equality. Thus, $\operatorname{res}^r(\bar{\varphi}^2, \bar{\varphi}^1) = (0, \operatorname{res}^r(\pi_2 \circ \bar{\varphi}^2, \pi_3 \circ \bar{\varphi}^1))$ is symmetric, as required. □

Theorem 4.8 *If F_h is an evolution map of an order r discretization $(L_h, \psi, \alpha^+, \alpha^-)$ of a regular Lagrangian system $L\colon \boldsymbol{T}Q \to \mathbb{R}$, then $F_h(v) = F_h^{X_E}(v) + O(h^{r+1})$.*

Proof Combine Theorems 4.3 and 4.7. □

5 Relations to $Q \times Q$

In this Section, we provide the relations between the discrete mechanics of Definition 3.3, with discrete phase space $\boldsymbol{T}Q$, and the standard discrete mechanics, with discrete phase space $Q \times Q$. By standard discrete mechanics we mean:

1. The discrete Lagrangian is of the form $L_h^{Q\times Q}\colon Q \times Q \to \mathbb{R}$.
2. $(h, (q^+, q^-), (\tilde{q}^+, \tilde{q}^-))$ is *critical* if $((q^+, q^-), (\tilde{q}^+, \tilde{q}^-))$ is a critical point of the variational principle

$$\begin{cases} S_h^{Q\times Q}\left((q^+, q^-), (\tilde{q}^+, \tilde{q}^-)\right) \equiv L_h^{Q\times Q}(q^+, q^-) + L_h^{Q\times Q}(\tilde{q}^+, \tilde{q}^-), \\ q^- \text{ and } \tilde{q}^+ \text{ constant, and } q^+ = \tilde{q}^-. \end{cases}$$

3. The discrete evolution is defined to advance from (q^+, q^-) to $(\tilde{q}^+, \tilde{q}^-)$ such that $(h, (q^+, q^-), (\tilde{q}^+, \tilde{q}^-))$ is critical.

4. Given a (continuous) Lagrangian system $L\colon \boldsymbol{T}\mathcal{Q} \to \mathbb{R}$, a discrete Lagrangian is *order* r if

$$L(v) = L_h^{\mathcal{Q}\times\mathcal{Q}}\left(\tau_{\mathcal{Q}} \circ F_h^{X_E}(v), \tau_{\mathcal{Q}}(v)\right) + O(h^{r+1}),$$

in which, of course, any order r accurate approximation $F_h^{X_E} + O(h^{r+1})$ may be substituted for the exact flow $F_h^{X_E}$.

This is the variational principle used in the example context of Sect. 2, and also in [8].

The equivalence of the two formalisms is based on Lemma 5.1, which is obtained in the proof of Proposition 2.9 of [4].

Lemma 5.1 *Let $\mathcal{M}$ be a manifold and let $(\psi, \alpha^+, \alpha^-)$ be a discretization of the tangent bundle of $\mathcal{M}$. Then $\Psi(h, v) \equiv (h, \partial_h^+(v), \partial_h^-(v))$ is a diffeomorphism between open neighborhoods $U \setminus (\{0\} \times \boldsymbol{T}\mathcal{M})$ and $W \setminus (\{0\} \times (\mathcal{M} \times \mathcal{M}))$.*

Given $L_h^{\mathcal{Q}\times\mathcal{Q}}$, choose any order r discretization $(\psi, \alpha^+, \alpha^-)$ of $\mathcal{Q}$, define

$$L_h(v) \equiv L_h^{\mathcal{Q}\times\mathcal{Q}}\left(\partial_h^+(v), \partial_h^-(v)\right), \tag{26}$$

and consider the discrete Lagrangian system $(L_h, \psi, \alpha^+, \alpha^-)$. This is order r by Definition 4.6, and Ψ intertwines the objectives and constraints of the discrete variational principles on $\boldsymbol{T}\mathcal{Q}$ and $\mathcal{Q}\times\mathcal{Q}$. Therefore Ψ is a bijective correspondence between their critical points. Conversely, if $(L_h, \psi, \alpha^+, \alpha^-)$ is a discrete Lagrangian system, then define $L_h^{\mathcal{Q}\times\mathcal{Q}}$ such that (26) holds, and the same correspondence is obtained. Thus, the version of discrete mechanics on $\boldsymbol{T}\mathcal{Q}$ and the standard version on $\mathcal{Q} \times \mathcal{Q}$ are entirely equivalent, in an order-preserving way.

References

1. Bobenko, A.I., Schröder, P., Sullivan, J.M., Ziegler, G.M. (eds.): Discrete Differential Geometry. Birkhäuser Verlag, Boston (2008)
2. Bobenko, A.I., Suris, Y.B.: Discrete differential geometry. Consistency as integrability. arXiv:math.DG/0504358v1, (2005)
3. Bou-Rabee, N., Marsden, J.E.: Hamilton–Pontryagin integrators on Lie groups part I: Introduction and structure-preserving properties. Found. Comput. Math. **9**, 197–219 (2009)
4. Cuell, C., Patrick, G.W.: Skew critical problems. Regul. Chaotic Dyn. **12**, 589–601 (2007)
5. Cuell, C., Patrick, G.W.: Geometric discrete analogues of tangent bundles and constrained Lagrangian systems. J. Geom. Phys. **59**, 976–997 (2009)
6. Gelfand, I.M., Fomin, S.V.: Calculus of Variations. Prentice-Hall Inc., New Jersey (1963). [Revised English edition translated and edited by Richard A. Silverman]
7. Marsden, J.E., Patrick, G.W., Shkoller, S.: Multisymplectic geometry, variational integrators, and nonlinear PDEs. Comm. Math. Phys. **199**, 351–395 (1998)
8. Marsden, J.E., West, M.: Discrete mechanics and variational integrators. Acta Numer. **10**, 357–514 (2001)
9. Patrick, G.W.: Two axially symmetric coupled rigid bodies: relative equilibria, stability, bifurcations, and a momentum preserving symplectic integrator. PhD thesis, University of California at Berkeley (1991)
10. Patrick, G.W.: Lagrangian mechanics without ordinary differential equations. Rep. Math. Phys. **57**, 437–443 (2006)

11. Weinstein, A.: Lagrangian mechanics and groupoids. Fields Inst. Comm. **7**, 207–231 (1996)
12. Wendlandt, J.M., Marsden, J.E.: Discrete integrators derived from a discrete variational principle. Phys. D **106**, 233–246 (1997)
13. Yoshimura, H., Marsden, J.E.: Dirac structures in Lagrangian mechanics. I. Implicit Lagrangian systems. J. Geom. Phys. **57**, 133–156 (2006)
14. Yoshimura, H., Marsden, J.E.: Dirac structures in Lagrangian mechanics. II. Variational structures. J. Geom. Phys. **57**, 209–250 (2006)