

Multisymplectic Geometry, Variational Integrators, and Nonlinear PDEs

Jerrold E. Marsden¹, George W. Patrick², Steve Shkoller³

¹ Control and Dynamical Systems, California Institute of Technology, 107-81, Pasadena, CA 91125, USA.
E-mail: marsden@cds.caltech.edu

² Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, SK S7N5E6, Canada.
E-mail: patrick@math.usask.ca

³ Center for Nonlinear Studies, MS-B258, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.
E-mail: shkoller@cds.caltech.edu

Received: 12 January 1998 / Accepted: 12 May 1998

Abstract: This paper presents a geometric-variational approach to continuous and discrete mechanics and field theories. Using multisymplectic geometry, we show that the existence of the fundamental geometric structures as well as their preservation along solutions can be obtained directly from the variational principle. In particular, we prove that a unique multisymplectic structure is obtained by taking the derivative of an action functional, and use this structure to prove covariant generalizations of conservation of symplecticity and Noether's theorem. Natural discretization schemes for PDEs, which have these important preservation properties, then follow by choosing a discrete action functional. In the case of mechanics, we recover the variational symplectic integrators of Veselov type, while for PDEs we obtain covariant spacetime integrators which conserve the corresponding discrete multisymplectic form as well as the discrete momentum mappings corresponding to symmetries. We show that the usual notion of symplecticity along an infinite-dimensional space of fields can be naturally obtained by making a spacetime split. All of the aspects of our method are demonstrated with a nonlinear sine-Gordon equation, including computational results and a comparison with other discretization schemes.

Contents

1	Introduction	352
2	Lagrangian Mechanics	353
3	Veselov Discretizations of Mechanics	358
4	Variational Principles for Classical Field Theory	361
5	Veselov-type Discretizations of Multisymplectic Field Theory	375
5.1	General theory	375
5.2	Numerical checks	384
6	Concluding Remarks	390
	References	394

1. Introduction

The purpose of this paper is to develop the geometric foundations for multisymplectic-momentum integrators for variational partial differential equations (PDEs). These integrators are the PDE generalizations of symplectic integrators that are popular for Hamiltonian ODEs (see, for example, the articles in Marsden, Patrick and Shadwick [1996], and especially the review article of McLachlan and Scovel [1996]) in that they are covariant spacetime integrators which preserve the geometric structures of the system.

Because of the covariance of our method which we shall describe below, the resulting integrators are spacetime localizable in the context of hyperbolic PDEs, and generalize the notion of symplecticity and symmetry preservation in the context of elliptic problems. Herein, we shall primarily focus on spacetime integrators; however, we shall remark on the connection of our method with the finite element method for elliptic problems, as well as the Gregory and Lin [1991] method in optimal control.

Historically, in the setting of ODEs, there have been many approaches devised for constructing symplectic integrators, beginning with the original derivations based on generating functions (see de Vogelaere [1956]) and proceeding to symplectic Runge-Kutta algorithms, the shake algorithm, and many others. In fact, in many areas of molecular dynamics, symplectic integrators such as the Verlet algorithm and variants thereof are quite popular, as are symplectic integrators for the integration of the solar system. In these domains, integrators that are either symplectic or which are adaptations of symplectic integrators, are amongst the most widely used.

A fundamentally new approach to symplectic integration is that of Veselov [1988], [1991] who developed a discrete mechanics based on a discretization of Hamilton's principle. This method leads in a natural way to symplectic-momentum integrators which include the shake and Verlet integrators as special cases (see Wendlandt and Marsden [1997]). In addition, Veselov integrators often have amazing properties with regard to preservation of integrable structures, as has been shown by Moser and Veselov [1991]. This aspect has yet to be exploited numerically, but it seems to be quite important.

The approach we take in this paper is to develop a Veselov-type discretization for PDE's in variational form. The relevant geometry for this situation is multisymplectic geometry (see Gotay, Isenberg, and Marsden [1997] and Marsden and Shkoller [1998]) and we develop it in a variational framework. As we have mentioned, this naturally leads to multisymplectic-momentum integrators. It is well-known that such integrators cannot in general preserve the Hamiltonian *exactly* (Ge and Marsden [1988]). However, these integrators have, under appropriate circumstances, very good energy performance in the sense of the conservation of a nearby Hamiltonian up to exponentially small errors, assuming small time steps, due to a result of Neishtadt [1984]. See also Dragt and Finn [1979], and Simo and Gonzales [1993]. This is related to backward error analysis; see Sanz-Serna and Calvo [1994], Calvo and Hairer [1995], and the recent work of Hyman, Newman and coworkers and references therein. It would be quite interesting to develop the links with Neishtadt's analysis more thoroughly.

An important part of our approach is to understand how the symplectic nature of the integrators is implied by the variational structure. In this way we are able to identify the symplectic and momentum conserving properties after discretizing the variational principle itself. Inspired by a paper of Wald [1993], we obtain a formal method for locating the symplectic or multisymplectic structures directly from the action function and its derivatives. We present the method in the context of ordinary Lagrangian mechanics, and apply it to discrete Lagrangian mechanics, and both continuous and discrete

multisymplectic field theory. While in these contexts our variational method merely uncovers the well-known differential-geometric structures, our method forms an excellent pedagogical approach to those theories.

Outline of paper.

Section 2. In this section we sketch the three main aspects of our variational approach in the familiar context of particle mechanics. We show that the usual symplectic 2-form on the tangent bundle of the configuration manifold arises naturally as the boundary term in the first variational principle. We then show that application of $d^2 = 0$ to the variational principle restricted to the space of solutions of the Euler–Lagrange equations produces the familiar concept of conservation of the symplectic form; this statement is obtained variationally in a non-dynamic context; that is, we do not require an evolutionary flow. We then show that if the action function is left invariant by a symmetry group, then Noether’s theorem follows directly and simply from the variational principle as well.

Section 3. Here we use our variational approach to construct discretization schemes for mechanics which preserve the discrete symplectic form and the associated discrete momentum mappings.

Section 4. This section defines the three aspects of our variational approach in the multisymplectic field-theoretic setting. Unlike the traditional approach of defining the canonical multisymplectic form on the dual of the first jet bundle and then pulling back to the Lagrangian side using the covariant Legendre transform, we obtain the geometric structure by staying entirely on the Lagrangian side. We prove the covariant analogue of the fact that the flow of conservative systems consists of symplectic maps; we call this result the *multisymplectic form formula*. After variationally proving a covariant version of Noether’s theorem, we show that one can use the multisymplectic form formula to recover the usual notion of symplecticity of the flow in an infinite-dimensional space of fields by making a spacetime split. We demonstrate this machinery using a nonlinear wave equation as an example.

Section 5. In this section we develop discrete field theories from which the covariant integrators follow. We define discrete analogues of the first jet bundle of the configuration bundle whose sections are the fields of interest, and proceed to define the discrete action sum. We then apply our variational algorithm to this discrete action function to produce the discrete Euler–Lagrange equations and the discrete multisymplectic forms. As a consequence of our methodology, we show that the solutions of the discrete Euler–Lagrange equations satisfy the discrete version of the multisymplectic form formula as well as the discrete version of our generalized Noether’s theorem. Using our nonlinear wave equation example, we develop various multisymplectic-momentum integrators for the sine-Gordon equations, and compare our resulting numerical scheme with the energy-conserving methods of Li and Vu-Quoc [1995] and Guo, Pascual, Rodriguez, and Vazquez [1986]. Results are presented for long-time simulations of kink-antikink solutions for over 5000 soliton collisions.

Section 6. This section contains some important remarks concerning the variational integrator methodology. For example, we discuss integrators for reduced systems, the role of grid uniformity, and the interesting connections with the finite-element methods for elliptic problems. We also make some comments on future work.

2. Lagrangian Mechanics

Hamilton’s principle. We begin by recalling a problem going back to Euler, Lagrange and Hamilton in the period 1740–1830. Consider an n -dimensional configuration man-

ifold Q with its tangent bundle TQ . We denote coordinates on Q by q^i and those on TQ by (q^i, \dot{q}^i) . Consider a Lagrangian $L : TQ \rightarrow \mathbb{R}$. Construct the corresponding action functional S on C^2 curves $q(t)$ in Q by integration of L along the tangent to the curve. In coordinate notation, this reads

$$S(q(t)) \equiv \int_a^b L \left(q^i(t), \frac{dq^i}{dt}(t) \right) dt. \quad (2.1)$$

The action functional depends on a and b , but this is not explicit in the notation. **Hamilton's principle** seeks the curves $q(t)$ for which the functional S is stationary under variations of $q(t)$ with fixed endpoints; namely, we seek curves $q(t)$ which satisfy

$$dS(q(t)) \cdot \delta q(t) \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} S(q_\epsilon(t)) = 0 \quad (2.2)$$

for all $\delta q(t)$ with $\delta q(a) = \delta q(b) = 0$, where q_ϵ is a smooth family of curves with $q_0 = q$ and $(d/d\epsilon)|_{\epsilon=0} q_\epsilon = \delta q$. Using integration by parts, the calculation for this is simply

$$\begin{aligned} dS(q(t)) \cdot \delta q(t) &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \int_a^b L \left(q_\epsilon^i(t), \frac{dq_\epsilon^i}{dt}(t) \right) dt \\ &= \int_a^b \delta q^i \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right) dt + \left. \frac{\partial L}{\partial \dot{q}^i} \delta q^i \right|_a^b. \end{aligned} \quad (2.3)$$

The last term in (2.3) vanishes since $\delta q(a) = \delta q(b) = 0$, so that the requirement (2.2) for S to be stationary yields the **Euler–Lagrange equations**

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = 0. \quad (2.4)$$

Recall that L is called **regular** when the symmetric matrix $[\partial^2 L / \partial \dot{q}^i \partial \dot{q}^j]$ is everywhere nonsingular. If L is regular, the Euler–Lagrange equations are second order ordinary differential equations for the required curves.

The standard geometric setting. The action (2.1) is independent of the choice of coordinates, and thus the Euler–Lagrange equations are coordinate independent as well. Consequently, it is natural that the Euler–Lagrange equations may be intrinsically expressed using the language of differential geometry. This intrinsic development of mechanics is now standard, and can be seen, for example, in Arnold [1978], Abraham and Marsden [1978], and Marsden and Ratiu [1994].

The **canonical 1-form** θ_0 on the $2n$ -dimensional cotangent bundle of Q , T^*Q is defined by

$$\theta_0(\alpha_q)w_{\alpha_q} \equiv \alpha_q \cdot T\pi_Q w_{\alpha_q}, \quad \alpha_q \in T_q^*Q, w_{\alpha_q} \in T_{\alpha_q}T^*Q,$$

where $\pi_Q : T^*Q \rightarrow Q$ is the canonical projection. The Lagrangian L intrinsically defines a fiber preserving bundle map $\mathbb{F}L : TQ \rightarrow T^*Q$, the **Legendre transformation**, by vertical differentiation:

$$\mathbb{F}L(v_q)w_q \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} L(v_q + \epsilon w_q).$$

We define the **Lagrange 1-form** on TQ , the Lagrangian side, by pull-back $\theta_L \equiv \mathbb{F}L^*\theta_0$, and the **Lagrange 2-form** by $\omega_L = -d\theta_L$. We then seek a vector field X_E (called the **Lagrange vector field**) on TQ such that $X_E \lrcorner \omega_L = dE$, where the **energy** E is defined by $E(v_q) \equiv \mathbb{F}L(v_q)v_q - L(v_q)$.

If $\mathbb{F}L$ is a local diffeomorphism then X_E exists and is unique, and its integral curves solve the Euler–Lagrange equations (2.4). In addition, the flow F_t of X_E preserves ω_L ; that is, $F_t^*\omega_L = \omega_L$. Such maps are **symplectic**, and the form ω_L is called a **symplectic 2-form**. This is an example of a **symplectic manifold**: a pair (M, ω) where M is a manifold and ω is closed nondegenerate 2-form.

Despite the compactness and precision of this differential-geometric approach, it is difficult to motivate and, furthermore, is not entirely contained on the Lagrangian side. The canonical 1-form θ_0 seems to appear from nowhere, as does the Legendre transform $\mathbb{F}L$. Historically, after the Lagrangian picture on TQ was constructed, the canonical picture on T^*Q emerged through the work of Hamilton, but the modern approach described above treats the relation between the Hamiltonian and Lagrangian pictures of mechanics as a mathematical tautology, rather than what it is – a discovery of the highest order.

The variational approach. More and more, one is finding that there are advantages to staying on the “Lagrangian side”. Many examples can be given, but the theory of Lagrangian reduction (the Euler–Poincaré equations being an instance) is one example (see, for example, Marsden and Ratiu [1994] and Holm, Marsden and Ratiu [1998a,b]); another, of many, is the direct variational approach to questions in black hole dynamics given by Wald [1993]. In such studies, it is the variational principle that is the center of attention.

We next show that one can derive in a natural way the fundamental differential geometric structures, including momentum mappings, directly from the variational approach. This development begins by removing the boundary condition $\delta q(a) = \delta q(b) = 0$ from (2.3). Eq. (2.3) becomes

$$dS(q(t)) \cdot \delta q(t) = \int_a^b \delta q^i \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right) dt + \left. \frac{\partial L}{\partial \dot{q}^i} \delta q^i \right|_a^b, \tag{2.5}$$

where the left side now operates on more general δq (this generalization will be described in detail in Sect. 4), while the last term on the right side does not vanish. That last term of (2.5) is a linear pairing of the function $\partial L / \partial \dot{q}^i$, a function of q^i and \dot{q}^i , with the tangent vector δq^i . Thus, one may consider it to be a 1-form on TQ ; namely the 1-form $(\partial L / \partial \dot{q}^i) dq^i$. This is exactly the Lagrange 1-form, and we can turn this into a formal theorem/definition:

Theorem 2.1. *Given a C^k Lagrangian L , $k \geq 2$, there exists a unique C^{k-2} mapping $D_{\text{EL}}L : \ddot{Q} \rightarrow T^*Q$, defined on the second order submanifold*

$$\ddot{Q} \equiv \left\{ \left. \frac{d^2 q}{dt^2}(0) \right| q \text{ a } C^2 \text{ curve in } Q \right\}$$

of TTQ , and a unique C^{k-1} 1-form θ_L on TQ , such that, for all C^2 variations $q_\epsilon(t)$,

$$dS(q(t)) \cdot \delta q(t) = \int_a^b D_{\text{EL}}L \left(\frac{d^2 q}{dt^2} \right) \cdot \delta q dt + \theta_L \left(\frac{dq}{dt} \right) \cdot \delta q \Big|_a^b, \tag{2.6}$$

where

$$\delta q(t) \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} q_\epsilon(t), \quad \hat{\delta} q(t) \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \left. \frac{d}{dt} \right|_{t=0} q_\epsilon(t).$$

The 1-form so defined is called the **Lagrange 1-form**.

Indeed, uniqueness and local existence follow from the calculation (2.3), and the coordinate independence of the action, and then global existence is immediate. Here then, is the first aspect of our method:

Using the variational principle, the Lagrange 1-form θ_L is the “boundary part” of the functional derivative of the action when the boundary is varied. The analogue of the symplectic form is the (negative of) the exterior derivative of θ_L .

For the mechanics example being discussed, we imagine a development wherein θ_L is so defined and we define $\omega_L \equiv -d\theta_L$.

Lagrangian flows are symplectic. One of Lagrange’s basic discoveries was that the solutions of the Euler–Lagrange equations give rise to a symplectic map. It is a curious twist of history that he did this without the machinery of either differential forms, of the Hamiltonian formalism or of Hamilton’s principle itself. (See Marsden and Ratiu [1994] for an account of some of this history.)

Assuming that L is regular, the variational principle then gives coordinate independent second order ordinary differential equations, as we have noted. We temporarily denote the vector field on TQ so obtained by X , and its flow by F_t . Our further development relies on a change of viewpoint: we focus on the restriction of S to the subspace \mathcal{C}_L of solutions of the variational principle. The space \mathcal{C}_L may be identified with the initial conditions, elements of TQ , for the flow: to $v_q \in TQ$, we associate the integral curve $s \mapsto F_s(v_q)$, $s \in [0, t]$. The value of S on that curve is denoted by S_t , and again called the **action**. Thus, we define the map $S_t : TQ \rightarrow \mathbb{R}$ by

$$S_t(v_q) = \int_0^t L(q(s), \dot{q}(s)) ds, \quad (2.7)$$

where $(q(s), \dot{q}(s)) = F_s(v_q)$. The fundamental Eq. (2.6) becomes

$$dS_t(v_q)w_{v_q} = \theta_L(F_t(v_q)) \cdot \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} F_t(v_q^\epsilon) - \theta_L(v_q) \cdot w_{v_q},$$

where $\epsilon \mapsto v_q^\epsilon$ is an arbitrary curve in TQ such that $v_q^0 = v_q$ and $(d/d\epsilon)|_0 v_q^\epsilon = w_{v_q}$. We have thus derived the equation

$$dS_t = F_t^* \theta_L - \theta_L. \quad (2.8)$$

Taking the exterior derivative of (2.8) yields the fundamental fact that the flow of X is symplectic:

$$0 = ddS_t = d(F_t^* \theta_L - \theta_L) = -F_t^* \omega_L + \omega_L,$$

which is equivalent to

$$F_t^* \omega_L = \omega_L.$$

This leads to the following:

Using the variational principle, the fact that the evolution is symplectic is a consequence of the equation $d^2 = 0$, applied to the action restricted to the space of solutions of the variational principle.

In passing, we note that (2.8) also provides the differential-geometric equations for X . Indeed, one time derivative of (2.8) and using (2.7) gives $dL = \mathcal{L}_X \theta_{\mathcal{L}}$, so that

$$X \lrcorner \omega_L = -X \lrcorner d\theta_L = -\mathcal{L}_X \theta_L + d(X \lrcorner \theta_L) = d(X \lrcorner \theta_L - L) = dE,$$

if we define $E \equiv X \lrcorner \theta_L - L$. Thus, we quite naturally find that $X = X_E$.

Of course, this set up also leads directly to Hamilton–Jacobi theory, which was one of the ways in which symplectic integrators were developed (see McLachlan and Scovel [1996] and references therein.) However, we shall not pursue the Hamilton–Jacobi aspect of the theory here.

Momentum maps. Suppose that a Lie group G , with Lie algebra \mathfrak{g} , acts on Q , and hence on curves in Q , in such a way that the action S is invariant. Clearly, G leaves the set of solutions of the variational principle invariant, so the action of G restricts to \mathcal{C}_L , and the group action commutes with F_t . Denoting the infinitesimal generator of $\xi \in \mathfrak{g}$ on TQ by ξ_{TQ} , we have by (2.8),

$$0 = \xi_{TQ} \lrcorner dS_t = \xi_{TQ} \lrcorner (F_t^* \theta_L - \theta_L) = F_t^*(\xi_{TQ} \lrcorner \theta_L) - \xi_{TQ} \lrcorner \theta_L. \quad (2.9)$$

For $\xi \in \mathfrak{g}$, define $J_\xi : TQ \rightarrow \mathbb{R}$ by $J_\xi \equiv \xi_{TQ} \lrcorner \theta_L$. Then (2.9) says that J_ξ is an integral of the flow of X_E . We have arrived at a version of Noether’s theorem (rather close to the original derivation of Noether):

Using the variational principle, Noether’s theorem results from the infinitesimal invariance of the action restricted to space of solutions of the variational principle. The conserved momentum associated to a Lie algebra element ξ is $J_\xi = \xi_{TQ} \lrcorner \theta_L$, where θ_L is the Lagrange one-form.

Reformulation in terms of first variations. We have just seen that symplecticity of the flow and Noether’s theorem result from restricting the action to the space of solutions. One tacit assumption is that the space of solutions is a manifold in some appropriate sense. This is a potential problem, since solution spaces for field theories are known to have singularities (see, e.g., Arms, Marsden and Moncrief [1982]). More seriously there is the problem of finding a multisymplectic analogue of the statement that the Lagrangian flow map is symplectic, since for multisymplectic field theory one obtains an evolution picture only after splitting spacetime into space and time and adopting the “function space” point of view. Having the general formalism depend either on a spacetime split or an analysis of the associated Cauchy problem would be contrary to the general thrust of this article. We now give a formal argument, in the context of Lagrangian mechanics, which shows how both these problems can be simultaneously avoided.

Given a solution $q(t) \in \mathcal{C}_L$, a **first variation** at $q(t)$ is a vector field V on Q such that $t \mapsto F_t^V \circ q(t)$ is also a solution curve (i.e. a curve in \mathcal{C}_L). We think of the solution space \mathcal{C}_L as being a (possibly) singular subset of the smooth space of all putative curves \mathcal{C} in TQ , and the restriction of V to $q(t)$ as being the derivative of some curve in \mathcal{C}_L at $q(t)$. When \mathcal{C}_L is a manifold, a first variation is a vector at $q(t)$ tangent to \mathcal{C}_L . Temporarily define $\alpha \equiv dS - \theta_L$, where by abuse of notation θ_L is the one form on \mathcal{C} defined by

$$\theta_L(q(t))\delta q(t) \equiv \theta_L(b)\delta q(b) - \theta_L(a)\delta q(a).$$

Then \mathcal{C}_L is defined by $\alpha = 0$ and we have the equation

$$dS = \alpha + \theta_L,$$

so if V and W are first variations at $q(t)$, we obtain

$$0 = V \lrcorner W \lrcorner d^2 S = V \lrcorner W \lrcorner d\alpha + V \lrcorner W \lrcorner d\theta_L. \quad (2.10)$$

We have the identity

$$d\alpha(V, W)(q(t)) = V(\alpha(W)) - W(\alpha(V)) - \alpha([V, W]), \quad (2.11)$$

which we will use to evaluate (2.10) at the curve $q(t)$. Let F_ϵ^V denote the flow of V , define $q_\epsilon^V(t) \equiv F_\epsilon^V(q(t))$, and make similar definitions with W replacing V . For the first term of (2.11), we have

$$V(\alpha(W))(q(t)) = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \alpha(W)(q_\epsilon^V),$$

which vanishes, since α is zero along q_ϵ^V for every ϵ . Similarly the second term of (2.11) at $q(t)$ also vanishes, while the third term vanishes since $\alpha(q(t)) = 0$. Consequently, symplecticity of the Lagrangian flow F_t may be written

$$V \lrcorner W \lrcorner d\theta_L = 0,$$

for all first variations V and W . This formulation is valid whether or not the solution space is a manifold, and it does not explicitly refer to any temporal notion. Similarly, Noether's theorem may be written in this way. Summarizing:

Using the variational principle, the analogue of the evolution is symplectic is the equation $d^2 S = 0$ restricted to first variations of the space of solutions of the variational principle. The analogue of Noether's theorem is infinitesimal invariance of dS restricted to first variations of the space of solutions of the variational principle.

The variational route to the differential-geometric formalism has obvious pedagogical advantages. More than that, however, it systematizes searching for the corresponding formalism in other contexts. We shall in the next sections show how this works in the context of discrete mechanics, classical field theory and multisymplectic geometry.

3. Veselov Discretizations of Mechanics

The discrete Lagrangian formalism in Veselov [1988], [1991] fits nicely into our variational framework. Veselov uses $Q \times Q$ for the discrete version of the tangent bundle of a configuration space Q ; heuristically, given some a priori choice of time interval Δt , a point $(q_1, q_0) \in Q \times Q$ corresponds to the tangent vector $(q_1 - q_0)/\Delta t$. Define a **discrete Lagrangian** to be a smooth map $L : Q \times Q = \{q_1, q_0\} \rightarrow \mathbb{R}$, and the corresponding action to be

$$S \equiv \sum_{k=1}^n L(q_k, q_{k-1}). \quad (3.1)$$

The variational principle is to extremize S for variations holding the endpoints q_0 and q_n fixed. This variational principle determines a "discrete flow" $F : Q \times Q \rightarrow Q \times Q$ by $F(q_1, q_0) = (q_2, q_1)$, where q_2 is found from the **discrete Euler-Lagrange equations** (DEL equations):

$$\frac{\partial L}{\partial q_1}(q_1, q_0) + \frac{\partial L}{\partial q_0}(q_2, q_1) = 0. \quad (3.2)$$

In this section we work out the basic differential-geometric objects of this discrete mechanics directly from the variational point of view, consistent with our philosophy in the last section.

A mathematically significant aspect of this theory is how it relates to integrable systems, a point taken up by Moser and Veselov [1991]. We will not explore this aspect in any detail in this paper, although later, we will briefly discuss the reduction process and we shall test an integrator for an integrable pde, the sine-Gordon equation.

The Lagrange 1-form. We begin by calculating dS for variations that do not fix the endpoints:

$$\begin{aligned} dS(q_0, \dots, q_n) \cdot (\delta q_0, \dots, \delta q_n) &= \sum_{k=0}^{n-1} \left(\frac{\partial L}{\partial q_1}(q_{k+1}, q_k) \delta q_{k+1} + \frac{\partial L}{\partial q_0}(q_{k+1}, q_k) \delta q_k \right) \\ &= \sum_{k=1}^n \frac{\partial L}{\partial q_1}(q_k, q_{k-1}) \delta q_k + \sum_{k=0}^{n-1} \frac{\partial L}{\partial q_0}(q_{k+1}, q_k) \delta q_k \\ &= \sum_{k=1}^{n-1} \left(\frac{\partial L}{\partial q_1}(q_k, q_{k-1}) + \frac{\partial L}{\partial q_0}(q_{k+1}, q_k) \right) \delta q_k \\ &\quad + \frac{\partial L}{\partial q_0}(q_1, q_0) \delta q_0 + \frac{\partial L}{\partial q_1}(q_n, q_{n-1}) \delta q_n. \end{aligned} \quad (3.3)$$

It is the last two terms that arise from the boundary variations (i.e. these are the ones that are zero if the boundary is fixed), and so these are the terms amongst which we expect to find the discrete analogue of the Lagrange 1-form. Actually, interpretation of the boundary terms gives the *two* 1-forms on $Q \times Q$

$$\theta_L^-(q_1, q_0) \cdot (\delta q_1, \delta q_0) \equiv \frac{\partial L}{\partial q_0}(q_1, q_0) \delta q_0, \quad (3.4)$$

and

$$\theta_L^+(q_1, q_0) \cdot (\delta q_1, \delta q_0) \equiv \frac{\partial L}{\partial q_1}(q_1, q_0) \delta q_1, \quad (3.5)$$

and we regard *the pair* (θ^-, θ^+) as being the analogue of the one form in this situation.

Symplecticity of the flow. We parameterize the solutions of the variational principle by the initial conditions (q_1, q_0) , and restrict S to that solution space. Then Eq. (3.3) becomes

$$dS = \theta_L^- + F^* \theta_L^+. \quad (3.6)$$

We should be able to obtain the symplecticity of F by determining what the equation $ddS = 0$ means for the right-hand-side of (3.6). At first, this does not appear to work, since $ddS = 0$ gives

$$F^*(d\theta_L^+) = -d\theta_L^-, \quad (3.7)$$

which apparently says that F pulls a certain 2-form back to a different 2-form. The situation is aided by the observation that, from (3.4) and (3.5),

$$\theta_L^- + \theta_L^+ = dL, \quad (3.8)$$

and consequently,

$$d\theta_L^- + d\theta_L^+ = 0. \quad (3.9)$$

Thus, there are *two* generally distinct 1-forms, but (up to sign) only *one* 2-form. If we make the definition

$$\omega_L \equiv d\theta_L^- = -d\theta_L^+,$$

then (3.7) becomes $F^*\omega_L = \omega_L$. Eq. (3.4), in coordinates, gives

$$\omega_L = \frac{\partial^2 L}{\partial q_0^i \partial q_1^j} dq_0^i \wedge dq_1^j,$$

which agrees with the discrete symplectic form found in Veselov [1988], [1991].

Noether's theorem. Suppose a Lie group G with Lie algebra \mathfrak{g} acts on Q , and hence diagonally on $Q \times Q$, and that L is G -invariant. Clearly, S is also G -invariant and G sends critical points of S to themselves. Thus, the action of G restricts to the space of solutions, the map F is G -equivariant, and from (3.6),

$$0 = \xi_{Q \times Q} \lrcorner dS = \xi_{Q \times Q} \lrcorner \theta_L^- + \xi_{Q \times Q} \lrcorner (F^*\theta_L^+),$$

for $\xi \in \mathfrak{g}$, or equivalently, using the equivariance of F ,

$$\xi_{Q \times Q} \lrcorner \theta_L^- = -F^*(\xi_{Q \times Q} \lrcorner \theta^+). \quad (3.10)$$

Since L is G -invariant, (3.8) gives $\xi_{Q \times Q} \lrcorner \theta_L^- = -\xi_{Q \times Q} \lrcorner \theta_L^+$, which in turn converts (3.10) to the conservation equation

$$\xi_{Q \times Q} \lrcorner \theta_L^+ = F^*(\xi_{Q \times Q} \lrcorner \theta^+). \quad (3.11)$$

Defining the discrete momentum to be

$$J_\xi \equiv \xi_{Q \times Q} \lrcorner \theta_L^+,$$

we see that (3.11) becomes conservation of momentum. A virtually identical derivation of this discrete Noether theorem is found in Marsden and Wendlant [1997].

Reduction. As we mentioned above, this formalism lends itself to a discrete version of the theory of Lagrangian reduction (see Marsden and Scheurle [1993a,b], Holm, Marsden and Ratiu [1998a] and Cendra, Marsden and Ratiu [1998]). This theory is not the focus of this article, so we shall defer a brief discussion of it until the conclusions.

4. Variational Principles for Classical Field Theory

Multisymplectic geometry. We now review some aspects of multisymplectic geometry, following Gotay, Isenberg and Marsden [1997] and Marsden and Shkoller [1997].

We let $\pi_{XY} : Y \rightarrow X$ be a fiber bundle over an oriented manifold X . Denote the first jet bundle over Y by $J^1(Y)$ or J^1Y and identify it with the *affine* bundle over Y whose fiber over $y \in Y_x := \pi_{XY}^{-1}(x)$ consists of $\text{Aff}(T_xX, T_yY)$, those linear mappings $\gamma : T_xX \rightarrow T_yY$ satisfying

$$T\pi_{XY} \circ \gamma = \text{Identity on } T_xX.$$

We let $\dim X = n + 1$ and the fiber dimension of Y be N . Coordinates on X are denoted $x^\mu, \mu = 1, 2, \dots, n, 0$, and fiber coordinates on Y are denoted by $y^A, A = 1, \dots, N$. These induce coordinates v^A_μ on the fibers of $J^1(Y)$. If $\phi : X \rightarrow Y$ is a section of π_{XY} , its tangent map at $x \in X$, denoted $T_x\phi$, is an element of $J^1(Y)_{\phi(x)}$. Thus, the map $x \mapsto T_x\phi$ defines a section of $J^1(Y)$ regarded as a bundle over X . This section is denoted $j^1(\phi)$ or $j^1\phi$ and is called the first jet of ϕ . In coordinates, $j^1(\phi)$ is given by

$$x^\mu \mapsto (x^\mu, \phi^A(x^\mu), \partial_\nu \phi^A(x^\mu)), \tag{4.1}$$

where $\partial_\nu = \partial/\partial x^\nu$.

Higher order jet bundles of Y , $J^m(Y)$, then follow as $J^1(\dots(J^1(Y)))$. Analogous to the tangent map of the projection $\pi_{Y, J^1(Y)}, T\pi_{Y, J^1(Y)} : TJ^1(Y) \rightarrow TY$, we may define the jet map of this projection which takes $J^2(Y)$ onto $J^1(Y)$

Definition 4.1. Let $\gamma \in J^1(Y)$ so that $\pi_{X, J^1(Y)}(\gamma) = x$. Then

$$J\pi_{Y, J^1(Y)} : \text{Aff}(T_xX, T_\gamma J^1(Y)) \rightarrow \text{Aff}(T_xX, T\pi_{Y, J^1(Y)} \cdot T_\gamma J^1(Y)).$$

We define the subbundle Y'' of $J^2(Y)$ over X which consists of second-order jets so that on each fiber

$$Y''_x = \{s \in J^2(Y)_\gamma \mid J\pi_{Y, J^1(Y)}(s) = \gamma\}.$$

In coordinates, if $\gamma \in J^1(Y)$ is given by (x^μ, y^A, v^A_μ) , and $s \in J^2(Y)_\gamma$ is given by $(x^\mu, y^A, v^A_\mu, w^A_{\mu\nu}, \kappa^A_{\mu\nu})$, then s is a second-order jet if $v^A_\mu = w^A_\mu$. Thus, the second jet of $\phi \in \Gamma(\pi_{XY})$, $j^2(\phi)$, given in coordinates by the map $x^\mu \mapsto (x^\mu, \phi^A, \partial_\nu \phi^A, \partial_\mu \partial_\nu \phi^A)$, is an example of a second-order jet.

Definition 4.2. The *dual jet bundle* $J^1(Y)^*$ is the vector bundle over Y whose fiber at $y \in Y_x$ is the set of affine maps from $J^1(Y)_y$ to $\Lambda^{n+1}(X)_x$, the bundle of $(n+1)$ -forms on X . A smooth section of $J^1(Y)^*$ is therefore an affine bundle map of $J^1(Y)$ to $\Lambda^{n+1}(X)$ covering π_{XY} .

Fiber coordinates on $J^1(Y)^*$ are (p, p_A^μ) , which correspond to the affine map given in coordinates by

$$v^A_\mu \mapsto (p + p_A^\mu v^A_\mu) d^{n+1}x, \tag{4.2}$$

where $d^{n+1}x = dx^1 \wedge \dots \wedge dx^n \wedge dx^0$.

Analogous to the canonical one- and two-forms on a cotangent bundle, there exist canonical $(n+1)$ - and $(n+2)$ -forms on the dual jet bundle $J^1(Y)^*$. In coordinates, with $d^n x_\mu := \partial_\mu \lrcorner d^{n+1}x$, these forms are given by

$$\Theta = p_A^\mu dy^A \wedge d^n x_\mu + p d^{n+1}x \quad \text{and} \quad \Omega = dy^A \wedge dp_A^\mu \wedge d^n x_\mu - dp \wedge d^{n+1}x. \quad (4.3)$$

A Lagrangian density $\mathcal{L} : J^1(Y) \rightarrow \Lambda^{n+1}(X)$ is a smooth bundle map over X . In coordinates, we write

$$\mathcal{L}(\gamma) = L(x^\mu, y^A, v^A_\mu) d^{n+1}x. \quad (4.4)$$

The corresponding covariant Legendre transformation for \mathcal{L} is a fiber preserving map over Y , $\mathbb{F}\mathcal{L} : J^1(Y) \rightarrow J^1(Y)^*$, expressed intrinsically as the first order vertical Taylor approximation to \mathcal{L} :

$$\mathbb{F}\mathcal{L}(\gamma) \cdot \gamma' = \mathcal{L}(\gamma) + \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{L}(\gamma + \varepsilon(\gamma' - \gamma)), \quad (4.5)$$

where $\gamma, \gamma' \in J^1(Y)_y$. A straightforward calculation shows that the covariant Legendre transformation is given in coordinates by

$$p_A^\mu = \frac{\partial L}{\partial v^A_\mu}, \quad \text{and} \quad p = L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu. \quad (4.6)$$

We can then define the **Cartan form** as the $(n+1)$ -form $\Theta_{\mathcal{L}}$ on $J^1(Y)$ given by

$$\Theta_{\mathcal{L}} = (\mathbb{F}\mathcal{L})^* \Theta, \quad (4.7)$$

and the $(n+2)$ -form $\Omega_{\mathcal{L}}$ by

$$\Omega_{\mathcal{L}} = -d\Theta_{\mathcal{L}} = (\mathbb{F}\mathcal{L})^* \Omega, \quad (4.8)$$

with local coordinate expressions

$$\begin{aligned} \Theta_{\mathcal{L}} &= \frac{\partial L}{\partial v^A_\mu} dy^A \wedge d^n x_\mu + \left(L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu \right) d^{n+1}x, \\ \Omega_{\mathcal{L}} &= dy^A \wedge d \left(\frac{\partial L}{\partial v^A_\mu} \right) \wedge d^n x_\mu - d \left[L - \frac{\partial L}{\partial v^A_\mu} v^A_\mu \right] \wedge d^{n+1}x. \end{aligned} \quad (4.9)$$

This is the differential-geometric formulation of the multisymplectic structure. Subsequently, we shall show how we may obtain this structure directly from the variational principle, staying entirely on the Lagrangian side $J^1(Y)$.

The multisymplectic form formula. In this subsection we prove a formula that is the multisymplectic counterpart to the fact that in finite-dimensional mechanics, the flow of a mechanical system consists of symplectic maps. Again, we do this by studying the action function.

Definition 4.3. *Let U be a smooth manifold with (piecewise) smooth closed boundary. We define the set of smooth maps*

$$\mathcal{C}^\infty = \{ \phi : U \rightarrow Y \mid \pi_{XY} \circ \phi : U \rightarrow X \text{ is an embedding} \}.$$

For each $\phi \in \mathcal{C}^\infty$, we set $\phi_X := \pi_{XY} \circ \phi$ and $U_X := \pi_{XY} \circ \phi(U)$ so that $\phi_X : U \rightarrow U_X$ is a diffeomorphism.

We may then define the infinite-dimensional manifold \mathcal{C} to be the closure of \mathcal{C}^∞ in either a Hilbert space or Banach space topology. For example, the manifold \mathcal{C} may be given the topology of a Hilbert manifold of bundle mappings, $H^s(U, Y)$, (U considered a bundle with fiber a point) for any integer $s \geq (n+1)/2$, so that the Hilbert sections $\phi \circ \phi_X^{-1}$ in Y are those whose distributional derivatives up to order s are square-integrable in any chart. With our condition on s , the Sobolev embedding theorem makes such mappings well defined. Alternately, one may wish to consider the Banach manifold \mathcal{C} as the closure of \mathcal{C}^∞ in the usual C^k -norm, or more generally, in a Holder space $C^{k+\alpha}$ -norm. See Palais [1968] and Ebin and Marsden [1970] for a detailed account of manifolds of mappings. The choice of topology for \mathcal{C} will not play a crucial role in this paper.

Definition 4.4. Let \mathcal{G} be the Lie group of π_{XY} -bundle automorphisms η_Y covering diffeomorphisms η_X , with Lie algebra \mathfrak{g} . We define the **action** $\Phi : \mathcal{G} \times \mathcal{C} \rightarrow \mathcal{C}$ by

$$\Phi(\eta_Y, \phi) = \eta_Y \circ \phi.^1$$

Furthermore, if $\phi \circ \phi_X^{-1} \in \Gamma(\pi_{U_X, Y})$, then $\Phi(\eta_Y, \phi) \in \Gamma(\pi_{\eta_X(U_X), Y})$. The **tangent space** to the manifold \mathcal{C} at a point ϕ is the set $T_\phi \mathcal{C}$ defined by

$$\{V \in C^\infty(X, TY) \mid \pi_{Y, TY} \circ V = \phi, \&T\pi_{XY} \circ V = V_X, \text{ a vector field on } X\}. \tag{4.10}$$

Of course, when these objects are topologized as we have described, the definition of the tangent space becomes a theorem, but as we have mentioned, this functional analytic aspect plays a minor role in what follows.

Given vectors $V, W \in T_\phi \mathcal{C}$ we may extend them to vector fields \mathcal{V}, \mathcal{W} on \mathcal{C} by fixing vector fields $v, w \in TY$ such that $V = v \circ (\phi \circ \phi_X^{-1})$ and $W = w \circ (\phi \circ \phi_X^{-1})$, and letting $\mathcal{V}_\rho = v \circ (\rho \circ \rho_X^{-1})$ and $\mathcal{W}_\rho = w \circ (\rho \circ \rho_X^{-1})$. Thus, the flow of \mathcal{V} on \mathcal{C} is given by $\Phi(\eta_Y^\lambda, \rho)$, where η_Y^λ covering η_X^λ is the flow of v . The definition of the bracket of vector fields using their flows, then shows that

$$[\mathcal{V}, \mathcal{W}](\rho) = [v, w] \circ (\rho \circ \rho_X^{-1}).$$

Whenever it is contextually clear, we shall, for convenience, write V for $v \circ (\phi \circ \phi_X^{-1})$.

Definition 4.5. The **action function** \mathcal{S} on \mathcal{C} is defined as follows:

$$\mathcal{S}(\phi) = \int_{U_X} \mathcal{L}(j^1(\phi \circ \phi_X^{-1})) \text{ for all } \phi \in \mathcal{C}. \tag{4.11}$$

Let $\lambda \mapsto \eta_Y^\lambda$ be an arbitrary smooth path in \mathcal{G} such that $\eta_Y^0 = e$, and let $V \in T_\phi \mathcal{C}$ be given by

$$V = \left. \frac{d}{d\lambda} \right|_{\lambda=0} \Phi(\eta_Y^\lambda, \phi), \text{ and } V_X = \left. \frac{d}{d\lambda} \right|_{\lambda=0} \eta_X^\lambda \circ \phi. \tag{4.12}$$

Definition 4.6. We say that ϕ is a **stationary point, critical point, or extremum** of \mathcal{S} if

$$\left. \frac{d}{d\lambda} \right|_{\lambda=0} \mathcal{S}(\Phi(\eta_Y^\lambda, \phi)) = 0. \tag{4.13}$$

¹ We shall also use the notation $\Phi(\eta_Y, \phi)$ to denote the section $\eta_Y \circ (\phi \circ \phi_x^{-1}) \circ \eta_x^{-1}$.

Then,

$$\begin{aligned} d\mathcal{S}_\phi \cdot V &= \left. \frac{d}{d\lambda} \right|_{\lambda=0} \int_{\eta_X^\lambda \circ \phi_X(U)} \mathcal{L}(\partial^1 \Phi(\eta_Y^\lambda, \phi)) \\ &= \left. \frac{d}{d\lambda} \right|_{\lambda=0} \int_{\phi_X(U)} j^1(\phi \circ \phi_X^{-1})^* j^1(\eta_Y^\lambda)^* \Theta_{\mathcal{L}}, \end{aligned} \quad (4.14)$$

where we have used the fact that $\mathcal{L}(z) = z^* \Theta_{\mathcal{L}}$ for all holonomic sections z of $J^1(Y)$ (see Corollary 4.2 below), and that

$$j^1(\eta_Y \circ \phi \circ \phi_X^{-1} \circ \eta_X^{-1}) = j^1(\eta_Y) \circ j^1(\phi \circ \phi_X^{-1}) \circ \eta_X^{-1}.$$

Using the Cartan formula, we obtain that

$$\begin{aligned} d\mathcal{S}_\phi \cdot V &= \int_{U_X} j^1(\phi \circ \phi_X^{-1})^* \mathfrak{L}_{j^1(V)} \Theta_{\mathcal{L}} \\ &= \int_{U_X} -j^1(\phi \circ \phi_X^{-1})^* [j^1(V) \lrcorner \Omega_{\mathcal{L}}] \\ &\quad + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(V) \lrcorner \Theta_{\mathcal{L}}]. \end{aligned} \quad (4.15)$$

Hence, a necessary condition for $\phi \in \mathcal{C}$ to be an extremum of \mathcal{S} is that the first term in (4.15) vanishes. One may readily verify that the integrand of the first term in (4.15) is equal to zero whenever $j^1(V)$ is replaced by $W \in TJ^1(Y)$ which is either $\pi_{Y, J^1(Y)}$ -vertical or tangent to $j^1(\phi \circ \phi_X^{-1})$ (see Marsden and Shkoller [1998]), so that using a standard argument from the calculus of variations, $j^1(\phi \circ \phi_X^{-1})^* [W \lrcorner \Omega_{\mathcal{L}}]$ must vanish for all vectors W on $J^1(Y)$ in order for ϕ to be an extremum of the action. We shall call such elements $\phi \in \mathcal{C}$ covering ϕ_X , solutions of the Euler–Lagrange equations.

Definition 4.7. *We let*

$$\mathcal{P} = \{\phi \in \mathcal{C} \mid j^1(\phi \circ \phi_X^{-1})^* [W \lrcorner \Omega_{\mathcal{L}}] = 0 \text{ for all } W \in TJ^1(Y)\}. \quad (4.16)$$

In coordinates, $(\phi \circ \phi_X^{-1})^A$ is an element of \mathcal{P} if

$$\frac{\partial L}{\partial y^A}(j^1(\phi \circ \phi_X^{-1})) - \frac{\partial}{\partial x^\mu} \left(\frac{\partial L}{\partial v_\mu^A}(j^1(\phi \circ \phi_X^{-1})) \right) = 0 \text{ in } U_X.$$

We are now ready to prove the multisymplectic form formula, a generalization of the symplectic flow theorem, but we first make the following remark. If \mathcal{P} is a submanifold of \mathcal{C} , then for any $\phi \in \mathcal{P}$, we may identify $T_\phi \mathcal{P}$ with the set $\{V \in T_\phi \mathcal{C} \mid j^1(\phi \circ \phi_X^{-1})^* \mathfrak{L}_{j^1(V)} [W \lrcorner \Omega_{\mathcal{L}}] = 0 \text{ for all } W \in TJ^1(Y)\}$ since such vectors arise by differentiating $\left. \frac{d}{d\epsilon} \right|_{\epsilon=0} j^1(\phi^\epsilon \circ \phi_X^{-1})^* [W \lrcorner \Omega_{\mathcal{L}}] = 0$, where ϕ^ϵ is a smooth curve of solutions of the Euler–Lagrange equations in \mathcal{P} (when such solutions exist). More generally, we do not require \mathcal{P} to be a submanifold in order to define the first variation solution of the Euler–Lagrange equations.

Definition 4.8. *For any $\phi \in \mathcal{P}$, we define the set*

$$\mathcal{F} = \{V \in T_\phi \mathcal{C} \mid j^1(\phi \circ \phi_X^{-1})^* \mathfrak{L}_{j^1(V)} [W \lrcorner \Omega_{\mathcal{L}}] = 0 \text{ for all } W \in TJ^1(Y)\}. \quad (4.17)$$

Elements of \mathcal{F} solve the first variation equations of the Euler–Lagrange equations.

Theorem 4.1 (Multisymplectic form formula). *If $\phi \in \mathcal{P}$, then for all V and W in \mathcal{F} ,*

$$\int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(V) \lrcorner j^1(W) \lrcorner \Omega_{\mathcal{L}}] = 0. \quad (4.18)$$

Proof. We define the 1-forms α_1 and α_2 on \mathcal{C} by

$$\alpha_1(\phi) \cdot V := \int_{U_X} -j^1(\phi \circ \phi_X^{-1})^* [j^1(V) \lrcorner \Omega_{\mathcal{L}}]$$

and

$$\alpha_2(\phi) \cdot V := \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(V) \lrcorner \Theta_{\mathcal{L}}],$$

so that by (4.15),

$$dS_{\phi} \cdot V = \alpha_1(\phi) \cdot V + \alpha_2(\phi) \cdot V \quad \text{for all } V \in T_{\phi}\mathcal{C}. \quad (4.19)$$

Recall that for any 1-form α on \mathcal{C} and vector fields V, W on \mathcal{C} ,

$$d\alpha(V, W) = V[\alpha(W)] - W[\alpha(V)] - \alpha([V, W]). \quad (4.20)$$

We let $\phi_{\epsilon} = \eta_Y^{\epsilon} \circ \phi$ be a curve in \mathcal{C} through ϕ , where η_Y^{ϵ} is a curve in \mathcal{G} through the identity such that

$$W = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \eta_Y^{\epsilon} \text{ and } W \in \mathcal{F},$$

and consider Eq. (4.19) restricted to all $V \in \mathcal{F}$.

Thus,

$$\begin{aligned} d(\alpha_2(V))(\phi) \cdot W &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} (\alpha_2(V)(\phi_{\epsilon})) \\ &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* j^1(\eta_Y^{\epsilon}) [j^1(V) \lrcorner \Theta_{\mathcal{L}}] \\ &= \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* \mathfrak{L}_{j^1(W)} [j^1(V) \lrcorner \Theta_{\mathcal{L}}] \\ &= \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(W) \lrcorner d(j^1(V) \lrcorner \Theta_{\mathcal{L}})] \\ &\quad + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* d[j^1(W) \lrcorner j^1(V) \lrcorner \Theta_{\mathcal{L}}], \end{aligned}$$

where the last equality was obtained using Cartan's formula. Using Stoke's theorem, noting that $\partial\partial U$ is empty, and applying Cartan's formula once again, we obtain that

$$\begin{aligned} d(\alpha_2(\phi)(V)) \cdot W &= \int_{\partial U_X} -j^1(\phi \circ \phi_X^{-1})^* [j^1(W) \lrcorner j^1(V) \lrcorner \Omega_{\mathcal{L}}] \\ &\quad + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(W) \lrcorner \mathfrak{L}_{j^1(V)} \Theta_{\mathcal{L}}], \end{aligned}$$

and

$$\begin{aligned} d(\alpha_2(\phi)(W)) \cdot V &= \int_{\partial U_X} -j^1(\phi \circ \phi_X^{-1})^*[j^1(V) \lrcorner j^1(W) \lrcorner \Omega_{\mathcal{L}}] \\ &\quad + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*[j^1(V) \lrcorner \mathfrak{L}_{j^1(W)}\Theta_{\mathcal{L}}]. \end{aligned}$$

Also, since $[j^1(V), j^1(W)] = j^1([V, W])$, we have

$$\alpha_2(\phi)([V, W]) = \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*[j^1(V), j^1(W)] \lrcorner \Theta_{\mathcal{L}}.$$

Now

$$[j^1(V), j^1(W)] \lrcorner \Theta_{\mathcal{L}} = \mathfrak{L}_{j^1(V)}(j^1(W) \lrcorner \Theta_{\mathcal{L}}) - j^1(W) \lrcorner \mathfrak{L}_{j^1(V)}\Theta_{\mathcal{L}},$$

so that

$$\begin{aligned} d\alpha_2(\phi)(V, W) &= 2 \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*[j^1(V) \lrcorner j^1(W) \lrcorner \Omega_{\mathcal{L}}] \\ &\quad + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*[j^1(V) \lrcorner \mathfrak{L}_{j^1(W)}\Theta_{\mathcal{L}} - \mathfrak{L}_{j^1(V)}(j^1(W) \lrcorner \Theta_{\mathcal{L}})]. \end{aligned}$$

But

$$\mathfrak{L}_{j^1(V)}(j^1(W) \lrcorner \Theta_{\mathcal{L}}) = d(j^1(V) \lrcorner j^1(W) \lrcorner \Theta_{\mathcal{L}}) + j^1(V) \lrcorner d(j^1(W) \lrcorner \Theta_{\mathcal{L}})$$

and

$$j^1(V) \lrcorner \mathfrak{L}_{j^1(W)}\Theta_{\mathcal{L}} = j^1(V) \lrcorner d(j^1(W) \lrcorner \Theta_{\mathcal{L}}) - j^1(V) \lrcorner j^1(W) \lrcorner \Omega_{\mathcal{L}}.$$

Hence,

$$\begin{aligned} &\int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*[j^1(V) \lrcorner \mathfrak{L}_{j^1(W)}\Theta_{\mathcal{L}} - \mathfrak{L}_{j^1(V)}(j^1(W) \lrcorner \Theta_{\mathcal{L}})] \\ &= \int_{\partial U_X} -j^1(\phi \circ \phi_X^{-1})^*(j^1(V) \lrcorner j^1(W) \lrcorner \Omega_{\mathcal{L}}) \\ &\quad - \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*d(j^1(V) \lrcorner j^1(W) \lrcorner \Theta_{\mathcal{L}}). \end{aligned}$$

The last term once again vanishes by Stokes theorem together with the fact that $\partial\partial U$ is empty, and we obtain that

$$d\alpha_2(\phi)(V, W) = \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*(j^1(V) \lrcorner j^1(W) \lrcorner \Omega_{\mathcal{L}}). \quad (4.21)$$

We now use (4.20) on α_1 . A similar computation as above yields

$$d(\alpha_1(\phi) \cdot V) \cdot W = \int_{U_X} j^1(\phi \circ \phi_X^{-1})^*\mathfrak{L}_{j^1(W)}[j^1(V) \lrcorner \Omega_{\mathcal{L}}]$$

which vanishes for all $\phi \in \mathcal{P}$ and $W \in \mathcal{F}$. Similarly, $d(\alpha_1(\phi) \cdot W) \cdot v = 0$ for all $\phi \in \mathcal{P}$ and $V \in \mathcal{F}$. Finally, $\alpha_1(\phi) = 0$ for all $\phi \in \mathcal{P}$.

Hence, since

$$0 = ddS(\phi)(V, W) = d\alpha_1(\phi)(V, W) + d\alpha_2(\phi)(V, W),$$

we obtain the formula (4.18). \square

Symplecticity revisited. Let Σ be a compact oriented connected boundaryless n -manifold which we think of as our reference Cauchy surface, and consider the space of embeddings of Σ into X , $\text{Emb}(\Sigma, X)$; again, although it is unnecessary for this paper, we may topologize $\text{Emb}(\Sigma, X)$ by completing the space in the appropriate C^k or H^s -norm.

Let B be an m -dimensional manifold. For any fiber bundle $\pi_{BK} : K \rightarrow B$, we shall, in addition to $\Gamma(\pi_{BK})$, use the corresponding script letter \mathcal{K} to denote the space of sections of π_{BK} . The space of sections of a fiber bundle is an infinite-dimensional manifold; in fact, it can be precisely defined and topologized as the manifold \mathcal{C} of the previous section, where the diffeomorphisms on the base manifold are taken to be the identity map, so that the tangent space to \mathcal{K} at σ is given simply by

$$T_\sigma \mathcal{K} = \{W : B \rightarrow VK \mid \pi_{K,TK} \circ W = \sigma\},$$

where VK denotes the vertical tangent bundle of K . We let $\pi_{K,L(VK,\Lambda^m(B))} : L(VK, \Lambda^m(B)) \rightarrow K$ be the vector bundle over K whose fiber at $k \in K_x, x = \pi_{BK}(k)$, is the set of linear mappings from $V_k K$ to $\Lambda^m(B)_x$. Then the cotangent space to \mathcal{K} at σ is defined as

$$T_\sigma^* \mathcal{K} = \{\pi : B \rightarrow L(VK, \Lambda^m(B)) \mid \pi_{K,L(VK,\Lambda^{n+1}(B))} \circ \pi = \sigma\}.$$

Integration provides the natural pairing of $T_\sigma^* \mathcal{K}$ with $T_\sigma \mathcal{K}$:

$$\langle \pi, V \rangle = \int_B \pi \cdot V.$$

In practice, the manifold B will either be X or some $(n + 1)$ -dimensional subset of X , or the n -dimensional manifold Σ_τ , where for each $\tau \in \text{Emb}(\Sigma, X)$, $\Sigma_\tau := \tau(\Sigma)$. We shall use the notation Y_τ for the bundle $\pi_{\Sigma_\tau, Y}$, and \mathcal{Y}_τ for sections of this bundle. For the remainder of this section, we shall set the manifold \mathcal{C} introduced earlier to \mathcal{Y} .

The infinite-dimensional manifold \mathcal{Y}_τ is called the τ -**configuration space**, its tangent bundle is called the τ -**tangent space**, and its cotangent bundle $T^* \mathcal{Y}_\tau$ is called the τ -**phase space**. Just as we described in Sect. 2, the cotangent bundle has a canonical 1-form θ_τ and a canonical 2-form ω_τ . These differential forms are given by

$$\theta_\tau(\varphi, \pi) \cdot V = \int_{\Sigma_\tau} \pi(T\pi_{\mathcal{Y}_\tau, T^* \mathcal{Y}_\tau} \cdot V) \text{ and } \omega_\tau = -d\theta_\tau, \tag{4.22}$$

where $(\varphi, \pi) \in \mathcal{Y}_\tau, V \in T_{(\varphi, \pi)} T^* \mathcal{Y}_\tau$, and $\pi_{\mathcal{Y}_\tau, T^* \mathcal{Y}_\tau} : T^* \mathcal{Y}_\tau \rightarrow \mathcal{Y}_\tau$ is the cotangent bundle projection map.

An infinitesimal slicing of the bundle π_{XY} consists of Y_τ together with a vector field ζ which is everywhere transverse to Y_τ , and covers ζ_X which is everywhere transverse to Σ_τ . The existence of an infinitesimal slicing allows us to invariantly decompose the temporal from the spatial derivatives of the fields. Let $\phi \in \mathcal{Y}, \varphi := \phi|_{\Sigma_\tau}$, and let $i_\tau : \Sigma_\tau \rightarrow X$ be the inclusion map. Then we may define the map β_ζ taking $j^1(\mathcal{Y})_\tau$ to $j^1(\mathcal{Y}_\tau) \times \Gamma(\pi_{\Sigma_\tau, VY_\tau})$ over \mathcal{Y}_τ by

$$\beta_\zeta(j^1(\phi) \circ i_\tau) = (j^1(\varphi), \dot{\varphi}) \text{ where } \dot{\varphi} := \mathfrak{L}_\zeta \phi. \tag{4.23}$$

In our notation, $j^1(\mathcal{Y})_\tau$ is the collection of restrictions of holonomic sections of $J^1(Y)$ to Σ_τ , while $j^1(\mathcal{Y}_\tau)$ are the holonomic sections of $\pi_{\Sigma_\tau, J^1(Y)}$. It is easy to see that β_ζ is an isomorphism; it then follows that β_ζ is an isomorphism of $j^1(\mathcal{Y})_\tau$ with $T\mathcal{Y}_\tau$, since $j^1(\varphi)$ is completely determined by φ . This bundle map is called the jet decomposition map, and its inverse is called the jet reconstruction map. Using this map, we can define the instantaneous Lagrangian.

Definition 4.9. *The instantaneous Lagrangian $L_{\tau,\zeta} : T\mathcal{Y}_\tau \rightarrow \mathbb{R}$ is given by*

$$L_{\tau,\zeta}(\varphi, \dot{\varphi}) = \int_{\Sigma_\tau} i_\tau^*[\zeta_X \lrcorner \mathcal{L}(\beta_\zeta^{-1}(j^1(\varphi), \dot{\varphi}))] \tag{4.24}$$

for all $(\varphi, \dot{\varphi}) \in T\mathcal{Y}_\tau$.

The instantaneous Lagrangian $L_{\tau,\zeta}$ has an instantaneous Legendre transform

$$\mathbb{F}L_{\tau,\zeta} : T\mathcal{Y}_\tau \rightarrow T^*\mathcal{Y}_\tau; (\varphi, \dot{\varphi}) \mapsto (\varphi, \pi)$$

which is defined in the usual way by vertical fiber differentiation of $L_{\tau,\zeta}$ (see, for example, Abraham and Marsden [1978]). Using the instantaneous Legendre transformation, we can pull-back the canonical 1- and 2-forms on $T^*\mathcal{Y}_\tau$.

Definition 4.10. *Denote, respectively, the instantaneous Lagrange 1- and 2-forms on $T\mathcal{Y}_\tau$ by*

$$\theta_\tau^L = \mathbb{F}L_{\tau,\zeta}^* \theta_\tau \text{ and } \omega_\tau^L = -d\theta_\tau^L. \tag{4.25}$$

Alternatively, we may define θ_τ^L using Theorem 2.1, in which case no reference to the cotangent bundle is necessary.

We will show that our covariant multisymplectic form formula can be used to recover the fact that the flow of the Euler-Lagrange equations in the bundle $\pi_{\text{Emb}(\Sigma, X), \cup_{\tau \in \text{Emb}(\Sigma, X)} T\mathcal{Y}_\tau}$ is symplectic with respect to ω_τ^L . To do so, we must relate the multisymplectic Cartan $(n + 2)$ -form $\Omega_{\mathcal{L}}$ on $J^1(Y)$ with the symplectic 2-form ω_τ^L on $T\mathcal{Y}_\tau$.

Theorem 4.2. *Let $\Theta_\tau^{\mathcal{L}}$ be the canonical 1-form on $j^1(\mathcal{Y})_\tau$ given by*

$$\Theta_\tau^{\mathcal{L}}(j^1(\phi) \circ i_\tau) \cdot V = \int_{\Sigma_\tau} i_\tau^* j^1(\phi)^* [V \lrcorner \Theta_{\mathcal{L}}], \tag{4.26}$$

where $j^1(\phi) \circ i_\tau \in j^1(\mathcal{Y})_\tau$, $V \in T_{j^1(\phi) \circ i_\tau} j^1(\mathcal{Y})_\tau$.

(a) *If the 2-form $\Omega_\tau^{\mathcal{L}}$ on $j^1(\mathcal{Y})_\tau$ is defined by $\Omega_\tau^{\mathcal{L}} := -d\Theta_\tau^{\mathcal{L}}$, then for $V, W \in T_{j^1(\phi) \circ i_\tau} j^1(\mathcal{Y})_\tau$,*

$$\Omega_\tau^{\mathcal{L}}(j^1(\phi) \circ i_\tau)(V, W) = \int_{\Sigma_\tau} i_\tau^* j^1(\phi)^* [W \lrcorner V \lrcorner \Omega_{\mathcal{L}}]. \tag{4.27}$$

(b) *Let the diffeomorphism $s_X : \Sigma \times \mathbb{R} \rightarrow X$ be a slicing of X such that for $\lambda \in \mathbb{R}$,*

$$\Sigma_\lambda := s_X(\Sigma \times \{\lambda\}) \text{ and } \tau_\lambda := \tau_\lambda(\Sigma),$$

where $\tau_\lambda \in \text{Emb}(\Sigma, X)$ is given by $\tau_\lambda(x) = s_X(x, \lambda)$. For any $\phi \in \mathcal{P}$, let $V, W \in T_\phi \mathcal{Y} \cap \mathcal{F}$ so that for each $\tau \in \text{Emb}(\Sigma, X)$, $j^1 V_\tau, j^1 W_\tau \in T_{j^1(\phi) \circ i_\tau} j^1(\mathcal{Y})_\tau$, and let $\tau_{\lambda_1}, \tau_{\lambda_2} \in \text{Emb}(\Sigma, X)$. Then

$$\Omega_{\tau_{\lambda_1}}^{\mathcal{L}}(j^1 V_{\tau_{\lambda_1}}, j^1 W_{\tau_{\lambda_1}}) = \Omega_{\tau_{\lambda_2}}^{\mathcal{L}}(j^1 V_{\tau_{\lambda_2}}, j^1 W_{\tau_{\lambda_2}}). \tag{4.28}$$

Proof. Part (a) follows from the Cartan formula together with Stokes theorem using an argument like that in the proof of Theorem 4.1.

For part (b), we recall that the multisymplectic form formula on \mathcal{Y} states that for any subset $U_X \subset X$ with smooth closed boundary and vectors $V, W \in T_\phi \mathcal{Y} \cap \mathcal{F}$, $\phi \in \mathcal{Y}$,

$$\int_{\partial U_X} j^1(\phi)^*[j^1(V) \lrcorner j^1(W) \lrcorner \Omega_{\mathcal{L}}] = 0. \tag{4.29}$$

Let

$$U_X = \cup_{\lambda \in [\lambda_1, \lambda_2]} \Sigma_\lambda.$$

Then $\partial U_X = \Sigma_{\lambda_1} - \Sigma_{\lambda_2}$, so that (4.29) can be written as

$$\begin{aligned} 0 &= \int_{\Sigma_{\lambda_2}} j^1(\phi \circ i_{\tau_{\lambda_2}})^*[j^1 V_{\tau_{\lambda_2}} \lrcorner j^1 W_{\tau_{\lambda_2}} \lrcorner \Omega_{\mathcal{L}}] \\ &\quad - \int_{\Sigma_{\lambda_1}} j^1(\phi \circ i_{\tau_{\lambda_1}})^*[j^1 V_{\tau_{\lambda_1}} \lrcorner j^1 W_{\tau_{\lambda_1}} \lrcorner \Omega_{\mathcal{L}}] \\ &= \Omega_{\tau_{\lambda_1}}^{\mathcal{L}}(j^1 V_{\tau_{\lambda_1}}, j^1 W_{\tau_{\lambda_1}}) - \Omega_{\tau_{\lambda_2}}^{\mathcal{L}}(j^1 V_{\tau_{\lambda_2}}, j^1 W_{\tau_{\lambda_2}}), \end{aligned}$$

which proves (4.28). \square

Theorem 4.3. *The identity $\Theta_\tau^{\mathcal{L}} = \beta_\zeta^* \theta_\tau^L$ holds.*

Proof. Let $W \in T_{j^1(\phi) \circ i_\tau} j^1(\mathcal{Y})_\tau$, which we identify with $w \circ \phi \circ i_\tau$, where w is a $\pi_{X, J^1(Y)}$ -vertical vector. Choose a coordinate chart which is adapted to the slicing so that $\partial_0|_{Y_\tau} = \zeta$. With $w = (0, W^A, W_\mu^A)$, we see that

$$\Theta_\tau^{\mathcal{L}} \cdot W = \int_{\Sigma_\tau} \frac{\partial L}{\partial v^A_0}(\phi^B, \phi^B_{,\mu}) W^A d^n x_0.$$

Now, from (4.24) we get

$$\begin{aligned} \theta_\tau^L(\varphi, \dot{\varphi}) &= \frac{\partial L_{\tau, \zeta}}{\partial \dot{y}^A} dy^A \\ &= \int_{\Sigma_\tau} \frac{\partial}{\partial \dot{y}^A} i_\tau^* [\partial_0 \lrcorner L(x^\mu, \phi^A, \phi^A_{,\mu}) d^{n+1}x \otimes dy^A] \\ &= \int_{\Sigma_\tau} \frac{\partial L}{\partial v^A_0}(\phi^B, \phi^B_{,\mu}) dy^A \otimes d^n x_0, \end{aligned}$$

where we arrived at the last equality using the fact that $\dot{y}^A = v^A_0$ in this adapted chart. Since $(T\beta_\zeta \cdot W)^A = W^A$, we see that $\Theta_\tau^{\mathcal{L}} \cdot W = \theta_\tau^L \cdot (T\beta_\zeta \cdot W)$, and this completes the proof. \square

Let the instantaneous energy $E_{\tau, \zeta}$ associated with $L_{\tau, \zeta}$ be given by

$$E_{\tau, \zeta}(\varphi, \dot{\varphi}) = \mathbb{F}L_{\tau, \zeta}(\dot{\varphi}) \cdot \dot{\varphi} - L_{\tau, \zeta}(\varphi, \dot{\varphi}), \tag{4.30}$$

and define the “time”-dependent Lagrangian vector field $X_{E_{\tau, \zeta}}$ by

$$X_{E_{\tau, \zeta}} \lrcorner \omega_\tau^L = dE_{\tau, \zeta}.$$

Since $\cup_{\tau \in \text{Emb}(\Sigma, X)} T\mathcal{Y}_\tau$ over $\text{Emb}(\Sigma, X)$ is infinite-dimensional and ω_τ^L is only weakly nondegenerate, the second-order vector field $X_{E_{\tau, \zeta}}$ does not, in general, exist. In the case that it does, we obtain the following result.

Corollary 4.1. *Assume $\bar{X}_{E_{\bar{\tau}, \zeta}}$ exists and let $F_{\bar{\tau}}$ be its semiflow, defined on some subset \mathcal{D} of the bundle $\cup_{\tau \in \text{Emb}(\Sigma, X)} T\mathcal{Y}_{\tau}$ over $\text{Emb}(\Sigma, X)$. Fix $\bar{\tau}$ so that $F_{\bar{\tau}}(\varphi_1, \dot{\varphi}_1) = (\varphi_2, \dot{\varphi}_2)$, where $(\varphi_1, \dot{\varphi}_1) \in T\mathcal{Y}_{\tau_1}$ and $(\varphi_2, \dot{\varphi}_2) \in T\mathcal{Y}_{\tau_2}$. Then $F_{\bar{\tau}}^* \omega_{\tau_2}^L = \omega_{\tau_1}^L$.*

Proof. This follows immediately from Theorem 4.2(b) and Theorem 4.3 and the fact that β_{ζ} induces an isomorphism between $j^1(\mathcal{Y})_{\tau}$ and $T\mathcal{Y}_{\tau}$. \square

Example: Nonlinear wave equation. To illustrate the geometry that we have developed, let us consider the scalar nonlinear wave equation given by

$$\frac{\partial^2 \phi}{\partial x^{02}} - \Delta \phi - N'(\phi) = 0, \quad \phi \in \Gamma(\pi_{XY}), \tag{4.31}$$

where Δ is the Laplace-Beltrami operator and N is a real-valued C^∞ function of one variable. For concreteness, fix $n=1$ so that the spacetime manifold $X := \mathbb{R}^2$, the configuration bundle $Y := \pi_{\mathbb{R}^2, \mathbb{R}}$, and the first jet bundle $J^1(Y) := \pi_{\mathbb{R}^2, \mathbb{R}^3}$.

Equation (4.31) is governed by the Lagrangian density

$$\mathcal{L} = \left\{ \frac{1}{2} \left[\frac{\partial \phi^2}{\partial x^0} - \frac{\partial \phi^2}{\partial x^1} \right] + N(\phi) \right\} dx^1 \wedge dx^0. \tag{4.32}$$

Using coordinates $(x^0, x^1, \phi, \phi_{,0}, \phi_{,1})$ for $J^1(Y)$, we write the multisymplectic 3-form for this nonlinear wave equation on \mathbb{R}^2 in coordinates as

$$\begin{aligned} \Omega_{\mathcal{L}} = & -d\phi \wedge d\phi_{,0} \wedge dx^1 - d\phi \wedge d\phi_{,1} \wedge dx^0 - N'(\phi) d\phi \wedge dx^1 \wedge dx^0 \\ & + \phi_{,0} d\phi_{,0} \wedge dx^1 \wedge dx^0 - \phi_{,1} d\phi_{,1} \wedge dx^1 \wedge dx^0; \end{aligned} \tag{4.33}$$

a short computation verifies that solutions of (4.31) are elements of \mathcal{P} , or that $j^1(\phi \circ \phi_X^{-1})^* [W \lrcorner \Omega_{\mathcal{L}}] = 0$ for all $W \in TJ^1(Y)$ (see Marsden and Shkoller [1998]).

We will use this example to demonstrate that our multisymplectic form formula generalizes the notion of symplecticity given by Bridges [1997]. Since the Lagrangian (4.32) does not explicitly depend on time, it is convenient to identify sections of Y as mappings from \mathbb{R}^2 into \mathbb{R} , and similarly, sections of $J^1(Y)$ as mappings from \mathbb{R}^2 into \mathbb{R}^3 . Thus, for $\phi \in \Gamma(\pi_{XY})$, $j^1(\phi)(x^\mu) := (\phi(x^\mu), \phi_{,0}(x^\mu), \phi_{,1}(x^\mu)) \in \mathbb{R}^3$, and if we set $p^\mu := \phi_{, \mu}$, then (4.31) can be reformulated to

$$\begin{aligned} & \mathbf{J}_0 j^1 \phi_{,0} + \mathbf{J}_1 j^1 \phi_{,1} := \\ & \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ p^0 \\ p^1 \end{bmatrix}_{,0} + \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ p^0 \\ p^1 \end{bmatrix}_{,1} = \begin{bmatrix} N'(\phi) \\ -p^0 \\ p^1 \end{bmatrix}. \end{aligned} \tag{4.34}$$

To each degenerate matrix \mathbf{J}_μ , we associate the contact form ω^μ on \mathbb{R}^3 given by $\omega^\mu(u_1, u_2) = \langle \mathbf{J}_\mu u_1, u_2 \rangle$, where $u_1, u_2 \in \mathbb{R}^3$ and $\langle \cdot, \cdot \rangle$ is the standard inner product on \mathbb{R}^3 . Bridges obtains the following conservation of symplecticity:

$$\frac{\partial}{\partial x^0} [\omega^0(j^1(\phi_{,0}), j^1(\phi_{,1}))] + \frac{\partial}{\partial x^1} [\omega^1(j^1(\phi_{,0}), j^1(\phi_{,1}))] = 0. \tag{4.35}$$

This result is interesting, but has somewhat limited scope in that the vector fields in (4.35) upon which the contact forms act are not general solutions to the first variation equations; rather, they are the specific first variation solutions $\phi_{, \mu}$. Bridges obtains this

result by crucially relying on the multi-Hamiltonian structure of (4.31); in particular, the vector $(N'(\phi), -p^0, p^1)$ on the right-hand-side of (4.34) is the gradient of a smooth multi-Hamiltonian function $H(\phi, p^0, p^1)$ (although the multi-Hamiltonian formalism is not important for this article, we refer the reader to Marsden and Shkoller [1998] for the Hamiltonian version of our covariant framework, and to Bridges [1997]). Using Eq. (4.34), it is clear that

$$H_{,0} = \omega^0(j^1(\phi_{,0}), j^1(\phi_{,1})) \text{ and } H_{,1} = -\omega^1(j^1(\phi_{,0}), j^1(\phi_{,1}))$$

so that (4.35) follows from the relation $H_{,0,1} = H_{,1,0}$.

Proposition 4.1. *The multisymplectic form formula is an intrinsic generalization of the conservation law (4.35); namely, for any $V, W \in \mathcal{F}$ that are $\pi_{X, j^1(Y)}$ -vertical,*

$$\frac{\partial}{\partial x^0} [\omega^0(j^1(V), j^1(W))] + \frac{\partial}{\partial x^1} [\omega^1(j^1(V), j^1(W))] = 0. \tag{4.36}$$

Proof. Let $j^1(V)$ and $j^1(W)$ have the coordinate expressions (V, V^0, V^1) and (W, W^0, W^1) , respectively. Using (4.33), we compute

$$j^1(W) \lrcorner j^1(V) \lrcorner \Omega_{\mathcal{L}} = (VW^0 - V^0W) dx + (VW^1 - V^1W) dt,$$

so that with Theorem 4.1 and the definition of ω^μ , we have, for $U_X \subset X$,

$$\int_{\partial U_X} \omega^0(j^1(V), j^1(W)) dx - \omega^1(j^1(V), j^1(W)) dt = 0,$$

and hence by Green's theorem,

$$\int_{U_X} \left\{ \frac{\partial}{\partial x^0} [\omega^0(j^1(V), j^1(W))] + \frac{\partial}{\partial x^1} [\omega^1(j^1(V), j^1(W))] \right\} dx^1 \wedge dx^0 = 0.$$

Since U_X is arbitrary, we obtain the desired result. \square

In general, when V is π_{XY} -vertical, $j^1(V)$ has the coordinate expression $(V, V_{,\mu} + \partial V / \partial \phi \cdot \phi_{,\mu})$, but for the special case that $V = \phi_{,\mu}$, $j^1(\phi_{,\mu}) = (j^1\phi)_{,\mu}$, and Proposition 4.1 gives

$$\frac{\partial}{\partial x^0} [\phi_0\phi_{,0,1} - \phi_1\phi_{,0,0}] - \frac{\partial}{\partial x^1} [\phi_0\phi_{,1,1} - \phi_1\phi_{,0,1}] = 0,$$

which simplifies to the trivial statement that

$$\phi_{,0}N(\phi)_{,1} - \phi_{,1}N(\phi)_{,0} = 0.$$

The variational route to the Cartan form. We may alternatively define the Cartan form by beginning with Eq. (4.14). Using the infinitesimal generators defined in (4.12), we obtain that

$$\begin{aligned} d\mathcal{S}_\phi \cdot V &= \left. \frac{d}{d\lambda} \right|_{\lambda=0} \mathcal{S}(\Phi(\eta_Y^\lambda), \phi) \\ &= \left. \frac{d}{d\lambda} \right|_{\lambda=0} \int_{\eta_X^\lambda(U_X)} \mathcal{L}(j^1(\Phi(\eta_Y^\lambda), \phi)) \\ &= \int_{U_X} \left. \frac{d}{d\lambda} \right|_{\lambda=0} \mathcal{L}(j^1(\Phi(\eta_Y^\lambda), \phi)) + \int_{U_X} \mathfrak{L}_{V_X} [\mathcal{L}(j^1(\phi \circ \phi_X^{-1}))]. \end{aligned} \tag{4.37}$$

Using the natural splitting of TY , any vector $V \in T_\phi\mathcal{C}$ may be decomposed as

$$V = V^h + V^v, \text{ where } V^h = T(\phi \circ \phi_X^{-1}) \cdot V_X \text{ and } V^v = V - V^h, \quad (4.38)$$

where we recall that $V_X = T\pi_{XY} \cdot V$.

Lemma 4.1. *For any $V \in T_\phi\mathcal{C}$,*

$$d\mathcal{S}_\phi \cdot V^h = \int_{\partial U_X} V_X \lrcorner [\mathcal{L}(j^1(\phi \circ \phi_X^{-1}))], \quad (4.39)$$

and

$$d\mathcal{S}_\phi \cdot V^v = \int_{U_X} \frac{d}{d\lambda} \Big|_{\lambda=0} \mathcal{L}(j^1(\Phi(\eta_Y^\lambda, \phi))). \quad (4.40)$$

Proof. The equality (4.40) is obvious, since the second term in (4.37) clearly vanishes for all vertical vectors. For vectors V^h , the first term in (4.37) vanishes; indeed, using the chain rule, we need only compute that

$$\frac{d}{d\lambda} \Big|_{\lambda=0} \eta_Y^\lambda \circ (\phi \circ \phi_X^{-1}) \circ \eta_X^{\lambda^{-1}} = V^h - T(\phi \circ \phi_X^{-1}) \cdot V_X,$$

which is zero by (4.38). We then apply the Cartan formula to the second term in (4.37) and note that $d\mathcal{L}$ is an $(n+2)$ -form on the $(n+1)$ -dimensional manifold U_X so that we obtain (4.39). \square

Theorem 4.4. *Given a smooth Lagrangian density $\mathcal{L} : J^1(Y) \rightarrow \Lambda^{n+1}(X)$, there exist a unique smooth section $D_{EL}\mathcal{L} \in C^\infty(Y'', \Lambda^{n+1}(X) \otimes T^*Y)$ and a unique differential form $\Theta_\mathcal{L} \in \Lambda^{n+1}(J^1(Y))$ such that for any $V \in T_\phi\mathcal{C}$, and any open subset U_X such that $\bar{U}_X \cap \partial X = \emptyset$,*

$$d\mathcal{S}_\phi \cdot V = \int_{U_X} D_{EL}\mathcal{L}(j^2(\phi \circ \phi_X^{-1})) \cdot V + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(V) \lrcorner \Theta_\mathcal{L}]. \quad (4.41)$$

Furthermore,

$$D_{EL}\mathcal{L}(j^2(\phi \circ \phi_X^{-1})) \cdot V = j^1(\phi \circ \phi_X^{-1})^* [j^1(V) \lrcorner \Omega_\mathcal{L}] \text{ in } U_X. \quad (4.42)$$

In coordinates, the action of the Euler–Lagrange derivative $D_{EL}\mathcal{L}$ on Y'' is given by

$$\begin{aligned} D_{EL}\mathcal{L}(j^2(\phi \circ \phi_X^{-1})) &= \left[\frac{\partial \mathcal{L}}{\partial y^A} (j^1(\phi \circ \phi_X^{-1})) - \frac{\partial^2 \mathcal{L}}{\partial x^\mu \partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \right. \\ &\quad - \frac{\partial^2 \mathcal{L}}{\partial y^B \partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \cdot (\phi \circ \phi_X^{-1})_{,\mu}^B \\ &\quad \left. - \frac{\partial^2 \mathcal{L}}{\partial v^B_\nu \partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \cdot (\phi \circ \phi_X^{-1})_{,\mu\nu}^B \right] dy^A \wedge d^{n+1}x, \end{aligned} \quad (4.43)$$

while the form $\Theta_\mathcal{L}$ matches the definition of the Cartan form given in (4.9) and has the coordinate expression

$$\Theta_\mathcal{L} = \frac{\partial \mathcal{L}}{\partial v^A_\mu} dy^A \wedge d^n x_\mu + \left(\mathcal{L} - \frac{\partial \mathcal{L}}{\partial v^A_\mu} v^A_\mu \right) d^{n+1}x. \quad (4.44)$$

Proof. Choose $U_X := \phi_X(U)$ small enough so that it is contained in a coordinate chart, say O_1 . In these coordinates, let $V = (V^\mu, V^A)$ so that along $\phi \circ \phi_X^{-1}$, our decomposition (4.38) may be written as

$$V_X = V^\mu \frac{\partial}{\partial x^\mu} \text{ and } V^v = (V^v)^A \frac{\partial}{\partial y^A} := \left(V^A - V^\mu \frac{\partial(\phi \circ \phi_X^{-1})^A}{\partial x^\mu} \right) \frac{\partial}{\partial y^A},$$

and Eq. (4.40) gives

$$d\mathcal{S}_\phi \cdot V^v = \int_{U_X} \left[\frac{\partial L}{\partial y^A} (j^1(\phi \circ \phi_X^{-1})) \cdot (V^v)^A + \frac{\partial L}{\partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \cdot \frac{\partial(V^v)^A}{\partial x^\mu} \right] d^{n+1}x, \quad (4.45)$$

where we have used the fact that in coordinates along $j^1(\phi \circ \phi_X^{-1})$,

$$\{j^1(V)\}^A_\mu = \partial_\mu[(V^v)^A(j^1(\phi \circ \phi_X^{-1}))].$$

Integrating (4.45) by parts, we obtain

$$\begin{aligned} d\mathcal{S}_\phi \cdot V^v = \int_{U_X} & \left\{ \left[\frac{\partial L}{\partial y^A} (j^1(\phi \circ \phi_X^{-1})) - \frac{\partial}{\partial x^\mu} \frac{\partial L}{\partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \right] \cdot V^A \right\} d^{n+1}x \\ & + \int_{\partial U_X} \left\{ \frac{\partial L}{\partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \cdot V^A d^n x_\mu \right. \\ & \left. + \frac{\partial L}{\partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \frac{\partial(\phi \circ \phi_X^{-1})^A}{\partial x^\nu} \cdot V^\nu d^n x_\mu \right\}. \end{aligned} \quad (4.46)$$

Let α be the n -form integrand of the boundary integral in (4.46); then $\int_{\partial U_X} \alpha = \int_{\partial j^1(\phi \circ \phi_X^{-1})(U_X)}$ since α is invariant under this lift. Additionally, from Eq. (4.39), we obtain the horizontal contribution

$$d\mathcal{S}_\phi \cdot V^h = \int_{\partial U_X} (V^\mu \partial_\mu) \lrcorner (L d^{n+1}x) = \int_{\partial j^1(\phi \circ \phi_X^{-1})(U_X)} V^\mu L d^n x_\mu, \quad (4.47)$$

so combining Eqs. (4.46) and (4.47), a simple computation verifies that

$$\begin{aligned} d\mathcal{S}_\phi \cdot V = \int_{U_X} & \left\{ \left[\frac{\partial L}{\partial y^A} (j^1(\phi \circ \phi_X^{-1})) - \frac{\partial}{\partial x^\mu} \frac{\partial L}{\partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \right] d^{n+1}x \otimes dy^A \right\} \cdot V \\ & + \int_{\partial j^1(\phi \circ \phi_X^{-1})(U_X)} V \lrcorner \left\{ \frac{\partial L}{\partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) dy^A \wedge d^n x_\mu \right. \\ & \left. + \left[L - \frac{\partial L}{\partial v^A_\mu} (j^1(\phi \circ \phi_X^{-1})) \frac{\partial(\phi \circ \phi_X^{-1})^A}{\partial x^\mu} \right] d^{n+1}x \right\}. \end{aligned} \quad (4.48)$$

The vector V in the second term of (4.48) may be replaced by $j^1(V)$ since $\pi_{Y, j^1(Y)}$ -vertical vectors are clearly in the kernel of the form that V is acting on. This shows that (4.43) and (4.44) hold, and hence that the boundary integral in (4.48) may be written as

$$\int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(V) \lrcorner \Theta_{\mathcal{L}}].$$

Now, if we choose another coordinate chart O_2 , the coordinate expressions of $D_{EL}\mathcal{L}$ and $\Theta_{\mathcal{L}}$ must agree on the overlap $O_1 \cap O_2$ since the left-hand-side of (4.41) is intrinsically defined. Thus, we have uniquely defined $D_{EL}\mathcal{L}$ and $\Theta_{\mathcal{L}}$ for any U_X such that $\overline{U}_X \cap \partial X = \emptyset$.

Finally, (4.42) holds, since $\Omega_{\mathcal{L}} = d\Theta_{\mathcal{L}}$ is also intrinsically defined and both sides of the equation yield the same coordinate representation, the Euler–Lagrange equations in U_X . \square

Remark. To prove Theorem 4.4 for the case $U_X = X$, we must modify the proof to take into account the boundary conditions which are prescribed on ∂X .

Corollary 4.2. *The $(n + 1)$ -form $\Theta_{\mathcal{L}}$ defined by the variational principle satisfies the relationship*

$$\mathcal{L}(z) = z^* \Theta_{\mathcal{L}}$$

for all holonomic sections $z \in \Gamma(\pi_X, J^1(Y))$.

Proof. This follows immediately by substituting (4.42) into (4.41) and integrating by parts using Cartan’s formula. \square

Remark. We have thus far focused on holonomic sections of $J^1(Y)$, those that are the first jets of sections of Y , and correspondingly, we have restricted the general splitting of TY given by

$$TY = \text{image } \gamma \oplus VY \text{ for any } \gamma \in \Gamma(J^1(Y)),$$

to $TY = T\phi \oplus VY$, $\phi \in \Gamma(Y)$ as we specified in (4.38). For general sections $\gamma \in \Gamma(J^1(Y))$, the horizontal bundle is given by image γ , and the Frobenius theorem guarantees that γ is locally holonomic if the connection is flat, or equivalently if the curvature of the connection R_γ vanishes. Since this is a local statement, we may assume that $Y = U \times \mathbb{R}^N$, where $U \subset \mathbb{R}^{n+1}$ is open, and that π_{XY} is simply the projection onto the first factor. For $\phi \in \Gamma(Y)$, and $\gamma \in \Gamma(J^1(Y))$, $\gamma(x, \phi(x)) : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^N$ is a linear operator which is holonomic if $\phi'(x) = \gamma(x, \phi(x))$, where $\phi'(x)$ is the differential of ϕ , and this is the case whenever the operator $\phi''(x)$ is symmetric. Equivalently, the operator

$$S_\gamma(x, y) \cdot (v, w) := D_1\gamma(x, y) \cdot (v, w) + D_2\gamma(x, y) \cdot (\gamma(x, y) \cdot v, w)$$

is symmetric for all $v, w \in \mathbb{R}^{n+1}$. One may easily verify that the local curvature is given by

$$R_\gamma(x, y) \cdot (v, w) := S_\gamma(x, y) \cdot (v, w) - S_\gamma(x, y) \cdot (w, v)$$

and that $\gamma = j^1(\phi)$ locally for some $\phi \in \Gamma(Y)$, if and only if $R_\gamma = 0$.

The variational route to Noether’s theorem. Suppose the Lie group \mathcal{G} acts on \mathcal{C} and leaves the action \mathcal{S} invariant so that

$$\mathcal{S}(\Phi(\eta_Y, \phi)) = \mathcal{S}(\phi) \text{ for all } \eta_Y \in \mathcal{G}. \tag{4.49}$$

This implies that for each $\eta_Y \in \mathcal{G}$, $\Phi(\eta_Y, \phi) \in \mathcal{P}$ whenever $\phi \in \mathcal{P}$. We restrict the

action of \mathcal{G} to \mathcal{P} , and let ξ_C be the corresponding infinitesimal generator on \mathcal{C} restricted to points in \mathcal{P} ; then

$$\begin{aligned} 0 &= (\xi_C \lrcorner d\mathcal{S})_\phi = \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(\xi) \lrcorner \Theta_{\mathcal{L}}] \\ &= \int_{U_X} j^1(\phi \circ \phi_X^{-1})^* [j^1(\xi) \lrcorner \Omega_{\mathcal{L}}], \end{aligned}$$

since $\mathfrak{L}_{j^1(\xi)}\Theta_{\mathcal{L}} = 0$ by (4.49) and Corollary 4.2.

We denote the covariant momentum map on $J^1(Y)$ by $\mathcal{J}^{\mathcal{L}} \in L(\mathfrak{g}, \Lambda^n(J^1(Y)))$ which we define as

$$j^1(\xi) \lrcorner \Omega_{\mathcal{L}} = d\mathcal{J}^{\mathcal{L}}(\xi). \tag{4.50}$$

Using (4.50), we find that $\int_{U_X} d[j^1(\phi \circ \phi_X^{-1})^* \mathcal{J}^{\mathcal{L}}(\xi)] = 0$, and since this must hold for all infinitesimal generators ξ_C at $\phi \in \mathcal{C}$, the integrand must also vanish so that

$$d[j^1(\phi \circ \phi_X^{-1})^* \mathcal{J}^{\mathcal{L}}(\xi)] = 0, \tag{4.51}$$

which is precisely a restatement of the covariant Noether Theorem.

5. Veselov-type Discretizations of Multisymplectic Field Theory

5.1. General theory. We now generalize the Veselov discretization given in Sect. 3 to multisymplectic field theory, by discretizing the spacetime X . For simplicity we restrict to the discrete analogue of $\dim X = 2$; i.e. $n = 1$. Thus, we take $X = \mathbb{Z} \times \mathbb{Z} = \{(i, j)\}$ and the fiber bundle Y to be $X \times \mathcal{F}$ for some smooth manifold \mathcal{F} .

Notation. The development in this section is aided by a small amount of notation and terminology. Elements of Y over the base point (i, j) are written as y_{ij} and the projection π_{XY} acts on Y by $\pi_{XY}(y_{ij}) = (i, j)$. The fiber over $(i, j) \in X$ is denoted Y_{ij} . A **triangle** Δ of X is an ordered triple of the form

$$\Delta = ((i, j), (i, j + 1), (i + 1, j + 1)).$$

The first component (i, j) of Δ is the **first vertex** of the triangle, denoted Δ^1 , and similarly for the **second** and **third vertices**. The set of all triangles in X is denoted X^Δ . By abuse of notation the same symbol is used for a triangle and the (unordered) set of its vertices. A point $(i, j) \in X$ is **touch**ed by a triangle if it is a vertex of that triangle. If $U \subseteq X$, then $(i, j) \in U$ is an **interior point** of U if U contains all three triangles of X that touch (i, j) . The **interior** $\text{int } U$ of U is the collection of the interior points of U . The **closure** $\text{cl } U$ of U is the union of all triangles touching interior points of U . A **boundary point** of U is a point in U and $\text{cl } U$ which is not an interior point. The **boundary** of U is the set of boundary points of U , so that

$$\partial U \equiv (U \cap \text{cl } U) \setminus \text{int } U.$$

Generally, U properly contains the union of its interior and boundary, and we call U **regular** if it is exactly that union. A **section** of Y is a map $\phi : U \subseteq X \rightarrow Y$ such that $\pi_{XY} \circ \phi = \text{id}_U$.

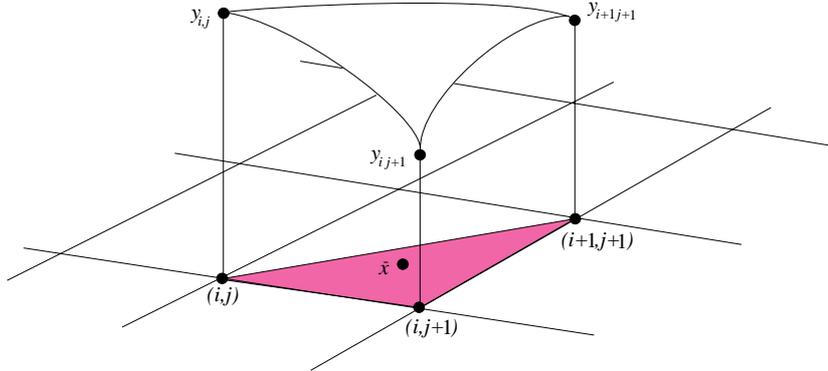


Fig. 5.1. Depiction of the heuristic interpretation of an element of $J^1 Y$ when X is discrete

Multisymplectic phase space. We define the **first jet bundle**² of Y to be

$$\begin{aligned} J^1 Y &\equiv \{(y_{ij}, y_{i, j+1}, y_{i+1, j+1}) \mid (i, j) \in X, y_{ij}, y_{i, j+1}, y_{i+1, j+1} \in \mathcal{F}\} \\ &\equiv X^\Delta \times \mathcal{F}^3. \end{aligned}$$

Heuristically (see Fig. 5.1), X corresponds to some grid of elements x_{ij} in continuous spacetime, say \tilde{X} , and $(y_{ij}, y_{i, j+1}, y_{i+1, j+1}) \in J^1 Y$ corresponds to $j^1 \phi(\bar{x})$, where \bar{x} is “inside” the triangle bounded by $x_{ij}, x_{i, j+1}, x_{i+1, j+1}$, and ϕ is some smooth section of $\tilde{X} \times \mathcal{F}$ interpolating the field values $y_{ij}, y_{i, j+1}, y_{i+1, j+1}$. The **first jet extension** of a section ϕ of Y is the map $j^1 \phi : X^\Delta \rightarrow J^1 Y$ defined by

$$j^1 \phi(\Delta) \equiv (\Delta, \phi(\Delta^1), \phi(\Delta^2), \phi(\Delta^3)).$$

Given a vector field Z on Y , we denote its restriction to the fiber Y_{ij} by Z_{ij} , and similarly for vector fields on $J^1 Y$. The **first jet extension** of a vector field Z on Y is the vector field $j^1 Z$ on $J^1 Y$ defined by

$$j^1 Z(y_{\Delta^1}, y_{\Delta^2}, y_{\Delta^3}) \equiv (Z_{\Delta^1}(y_{\Delta^1}), Z_{\Delta^2}(y_{\Delta^2}), Z_{\Delta^3}(y_{\Delta^3})),$$

for any triangle Δ .

The variational principle. Let us posit a **discrete Lagrangian** $L : J^1 Y \rightarrow \mathbb{R}$. Given a triangle Δ , define the function $L_\Delta : \mathcal{F}^3 \rightarrow \mathbb{R}$ by

$$L_\Delta(y_1, y_2, y_3) \equiv L(\Delta, y_1, y_2, y_3),$$

so that we may view the Lagrangian L as being a choice of a function L_Δ for each triangle Δ of X . The variables on the domain of L_Δ will be labeled y^1, y^2, y^3 , irrespective of the particular Δ . Let U be regular and let \mathcal{C}_U be the set of sections of Y on U , so \mathcal{C}_U is the manifold $\mathcal{F}^{|U|}$. The **action** will assign real numbers to sections in \mathcal{C}_U by the rule

$$S(\phi) \equiv \sum_{\Delta: \Delta \subseteq U} L \circ j^1 \phi(\Delta). \tag{5.1}$$

² Using three vertices is the simplest choice for approximating the two partial derivatives of the field ϕ , but may not lead to a good numerical scheme. Later, we shall also use four vertices together with averaging to define the partial derivatives of the fields.

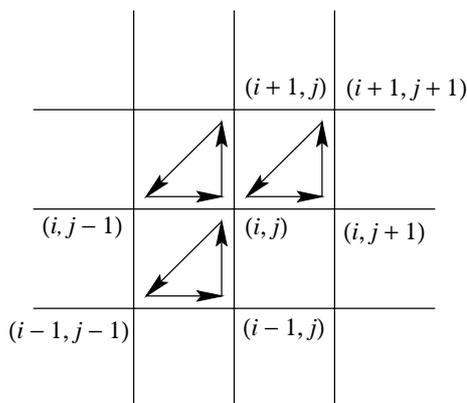


Fig. 5.2. The triangles which touch (i, j)

Given $\phi \in \mathcal{C}_U$ and a vector field V , there is the 1-parameter family of sections

$$(F_\epsilon^V \phi)(i, j) \equiv F_\epsilon^{V_{ij}}(\phi(i, j)),$$

where $F_\epsilon^{V_{ij}}$ denotes the flow of V_{ij} on \mathcal{F} . The *variational principle* is to seek those ϕ for which

$$\left. \frac{d}{d\epsilon} \right|_{\epsilon=0} S(F_\epsilon^V \phi) = 0$$

for all vector fields V .

The discrete Euler–Lagrange equations. The variational principle gives certain field equations, the *discrete Euler–Lagrange field equations* (DELf equations), as follows. Focus upon some $(i, j) \in \text{int } U$, and abuse notation by writing $\phi(i, j) \equiv y_{ij}$. The action, written with its summands containing y_{ij} explicitly, is (see Fig. 5.2)

$$S = \cdots + L(y_{ij}, y_{i+1j}, y_{i+1j+1}) + L(y_{ij-1}, y_{ij}, y_{i+1j}) + L(y_{i-1j-1}, y_{i-1j}, y_{ij}) + \cdots$$

so by differentiating in y_{ij} , the DELf equations are

$$\frac{\partial L}{\partial y^1}(y_{ij}, y_{i+1j}, y_{i+1j+1}) + \frac{\partial L}{\partial y^2}(y_{ij-1}, y_{ij}, y_{i+1j}) + \frac{\partial L}{\partial y^3}(y_{i-1j-1}, y_{i-1j}, y_{ij}) = 0,$$

for all $(i, j) \in \text{int } U$. Equivalently, these equations may be written

$$\sum_{l: \Delta^l(i, j) = \Delta^l} \frac{\partial L_\Delta}{\partial y^l}(y_{\Delta^1}, y_{\Delta^2}, y_{\Delta^3}) = 0, \tag{5.2}$$

for all $(i, j) \in \text{int } U$.

The discrete Cartan form. Now suppose we allow nonzero variations on the boundary ∂U , so we consider the effect on S of a vector field V which does not necessarily vanish on ∂U . For each $(i, j) \in \partial U$ find the triangles in U touching (i, j) . There is at least one such triangle since $(i, j) \in \text{cl } U$; there are not three such triangles since $(i, j) \notin \text{int } U$. For each such triangle Δ , (i, j) occurs as the l^{th} vertex, for one or two of $l = 1, 2, 3$, and those l^{th} expressions from the list

$$\begin{aligned} & \frac{\partial L}{\partial y^1}(y_{ij}, y_{i+1j}, y_{i+1j+1})V_{ij}(y_{ij}), \\ & \frac{\partial L}{\partial y^2}(y_{i-1j}, y_{ij}, y_{i+1j})V_{ij}(y_{ij}), \\ & \frac{\partial L}{\partial y^3}(y_{i-1j-1}, y_{i-1j}, y_{ij})V_{ij}(y_{ij}), \end{aligned}$$

yielding one or two numbers. The contribution to dS from the boundary is the sum of all such numbers. To bring this into a recognizable format, we take our cue from discrete Lagrangian mechanics, which featured *two* 1-forms. Here the above list suggests the *three* 1-forms on J^1Y , the first of which we define to be

$$\begin{aligned} \Theta_L^1(y_{ij}, y_{i+1j}, y_{i+1j+1}) \cdot (v_{y_{ij}}, v_{y_{i+1j}}, v_{y_{i+1j+1}}) \\ \equiv \frac{\partial L}{\partial y^1}(y_{ij}, y_{i+1j}, y_{i+1j+1}) \cdot (v_{y_{ij}}, 0, 0), \end{aligned}$$

Θ_L^2 and Θ_L^3 being defined analogously. With these notations, the contribution to dS from the boundary can be written $\theta_L(\phi) \cdot V$, where θ_L is the 1-form on the space of sections \mathcal{C}_U defined by

$$\theta_L(\phi) \cdot V \equiv \sum_{\Delta: \Delta \cap \partial U \neq \emptyset} \left(\sum_{l: \Delta^l \in \partial U} [(j^1\phi)^*(j^1V \lrcorner \Theta_L^l)](\Delta) \right). \tag{5.3}$$

In comparing (5.3) with (4.41), the analogy with the multisymplectic formalism of Sect. 4 is immediate.

The discrete multisymplectic form formula. Given a triangle Δ in X , we define the projection $\pi_\Delta : \mathcal{C}_U \rightarrow J^1Y$ by

$$\pi_\Delta(\phi) \equiv (\Delta, y_{\Delta^1}, y_{\Delta^2}, y_{\Delta^3}).$$

In this notation, it is easily verified that (5.3) takes the convenient form

$$\theta_L = \sum_{\Delta: \Delta \cap \partial U \neq \emptyset} \left(\sum_{l: \Delta^l \in \partial U} \pi_\Delta^* \Theta_L^l \right). \tag{5.4}$$

A *first-variation* at a solution ϕ of the DELF equations is a vertical vector field V such that the associated flow F^V maps ϕ to other solutions of the DELF equations. Set $\Omega_L^l = -d\Theta_L^l$. Since

$$\Theta_L^1 + \Theta_L^2 + \Theta_L^3 = dL, \tag{5.5}$$

one obtains

$$\Omega_L^1 + \Omega_L^2 + \Omega_L^3 = 0,$$

so that only two of the three 2-forms Ω_L^l , $l = 1, 2, 3$ are essentially distinct. Exactly as in Sect. 2, the equation $d^2S = 0$, when specialized to two first-variations V and W now gives, by taking one exterior derivative of (5.4),

$$0 = d\theta_L(\phi)(V, W) = \sum_{\Delta: \Delta \cap \partial U \neq \emptyset} \left(\sum_{l: \Delta^l \in \partial U} V \lrcorner W \lrcorner \pi_\Delta^* \Omega_L^l \right),$$

which in turn is equivalent to

$$\sum_{\Delta: \Delta \cap \partial U \neq \emptyset} \left(\sum_{l: \Delta^l \in \partial U} [(j^1 \phi)^*(j^1 V \lrcorner j^1 W \lrcorner \Omega_L^l)](\Delta) \right) = 0. \tag{5.6}$$

Again, the analogy with the multisymplectic form formula for continuous space-time (4.18) is immediate.

The discrete Noether theorem. Suppose that a Lie group G with Lie algebra \mathfrak{g} acts on F by vertical symmetries in such a way that the Lagrangian L is G -invariant. Then G acts on Y and $J^1 Y$ in the obvious ways. Since there are three Lagrange 1-forms, there are three momentum maps J^l , $l = 1, 2, 3$, each one a \mathfrak{g}^* -valued function on triangles in X , and defined by

$$J_\xi^l \equiv \xi_{J^1 Y} \lrcorner \Theta_L^l,$$

for any $\xi \in \mathfrak{g}$. Invariance of L and (5.5) imply that

$$J^1 + J^2 + J^3 = 0,$$

so, as in the case of the 1-forms, only two of the three momenta are essentially distinct. For any ξ , the infinitesimal generator ξ_Y is a first-variation, so invariance of S , namely $\xi_Y \lrcorner dS = 0$, becomes $\xi_Y \lrcorner \theta_L = 0$. By left insertion into (5.3), this becomes the discrete version of Noether's theorem:

$$\sum_{\Delta: \Delta \cap \partial U \neq \emptyset} \left(\sum_{l: \Delta^l \in \partial U} J^l(\Delta) \right) = 0. \tag{5.7}$$

Conservation in a space and time split. To understand the significance of (5.6) and (5.7) consider a discrete field theory with space a discrete version of the circle and time the real line, as depicted in Fig. 5.3, where space is split into space and time, with "constant time" being constant j and the "space index" $1 \leq i \leq N$ being cyclic. Applying (5.7) to the region $\{(i, j) \mid j = 0, 1, 2\}$ shown in the figure, Noether's theorem takes the conservation form

$$\begin{aligned} \sum_{i=1}^N J^1(y_{i0}, y_{i1}, y_{i+11}) &= - \sum_{i=1}^N (J^2(y_{i1}, y_{i2}, y_{i+12}) + J^3(y_{i1}, y_{i2}, y_{i+12})) \\ &= \sum_{i=1}^N J^1(y_{i1}, y_{i2}, y_{i+12}). \end{aligned}$$

Similarly, the discrete multisymplectic form formula also takes a conservation form. When there is spatial boundary, the discrete Noether theorem and the discrete multisymplectic form formulas automatically account for it, and thus form nontrivial generalizations of these conservation results.

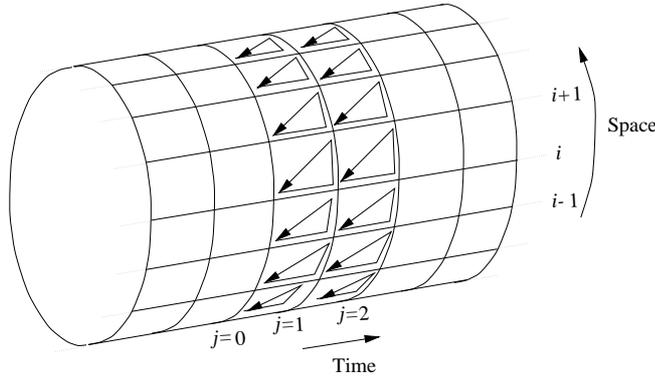


Fig. 5.3. Symplectic flow and conservation of momentum from the discrete Noether theorem when the spatial boundary is empty and the temporal boundaries agree

Furthermore, as in the continuous case, we can achieve “evolution type” symplectic systems (i.e. discrete Moser–Veselov mechanical systems) if we define Q as the space of fields at constant j , so $Q \equiv \mathcal{F}^N$, and take as the discrete Lagrangian

$$\tilde{L}([q_j^0], [q_j^1]) \equiv \sum_{i=1}^N L(q_i^0, q_i^1, q_{i+1}^1).$$

Then the Moser–Veselov DEL evolution-type equations (3.2) are equivalent to the DELF equations (5.2), the multisymplectic form formula implies symplecticity of the Moser–Veselov evolution map, and conservation of momentum gives identical results in both the “field” and “evolution” pictures.

Example: Nonlinear wave equation. To illustrate the discretization method we have developed, let us consider the Lagrangian (4.32) of Sect. 4, which describes the nonlinear sine-Gordon wave equation. This is a completely integrable system with an extremely interesting hierarchy of soliton solutions, which we shall investigate by developing for it a variational multisymplectic-momentum integrator; see the recent article by Palais [1997] for a wonderful discussion on soliton theory.

To discretize the continuous Lagrangian, we visualize each triangle Δ as having base length h and height k , and we think of the discrete jet $(y_{\Delta^1}, y_{\Delta^2}, y_{\Delta^3})$ as corresponding to the continuous jet

$$\frac{\partial \phi}{\partial x^0}(\bar{y}_{ij}) = \frac{y_{i,j+1} - y_{ij}}{h}, \quad \frac{\partial \phi}{\partial x^1}(\bar{y}_{ij}) = \frac{y_{i+1,j+1} - y_{i,j+1}}{k},$$

where \bar{y}_{ij} is the center of the triangle³. This leads to the discrete Lagrangian

$$L = \frac{1}{2} \left(\frac{y_2 - y_1}{h} \right)^2 - \frac{1}{2} \left(\frac{y_3 - y_2}{k} \right)^2 + N \left(\frac{y_1 + y_2 + y_3}{3} \right),$$

with corresponding DELF equations

³ Other discretizations based on triangles are possible; for example, one could use the value y_{ij} for insertion into the nonlinear term instead of \bar{y}_{ij} .

$$\begin{aligned}
 & \frac{y_{i+1j} - 2y_{ij} + y_{i-1j}}{k^2} - \frac{y_{ij+1} - 2y_{ij} + y_{ij-1}}{h^2} \\
 & + \frac{1}{3}N' \left(\frac{y_{ij} + y_{ij+1} + y_{i+1j+1}}{3} \right) \\
 & + \frac{1}{3}N' \left(\frac{y_{ij-1} + y_{ij} + y_{i+1j}}{3} \right) \\
 & + \frac{1}{3}N' \left(\frac{y_{i-1j-1} + y_{i-1j} + y_{ij}}{3} \right) = 0. \tag{5.8}
 \end{aligned}$$

When $N = 0$ (wave equation) this gives the explicit method

$$y_{i,j+1} = \frac{h^2}{k^2}(y_{i+1,j} - 2y_{ij} + y_{i-1,j}) + 2y_{ij} - y_{i,j-1},$$

which is stable whenever the Courant stability condition is satisfied.

Extensions: Jets from rectangles and other polygons. Our choice of discrete jet bundle is obviously not restricted to triangles, and can be extended to rectangles or more general polygons (left of Fig. 5.4). A **rectangle** is a quadruple of the form,

$$\Delta = ((i, j), (i, j + 1), (i + 1, j + 1), (i + 1, j)),$$

a point is an *interior point* of a subset U of rectangles if U contains all four rectangles touching that point, the discrete Lagrangian depends on variables y_1, \dots, y_4 , and the DELF equations become

$$\begin{aligned}
 & \frac{\partial L}{\partial y^1}(y_{ij}, y_{ij+1}, y_{i+1j+1}, y_{i+1j}) + \frac{\partial L}{\partial y^2}(y_{ij-1}, y_{ij}, y_{i+1j}, y_{i+1j-1}) \\
 & + \frac{\partial L}{\partial y^3}(y_{i-1j-1}, y_{i-1j}, y_{ij}, y_{ij-1}) + \frac{\partial L}{\partial y^4}(y_{i-1j}, y_{i-1j+1}, y_{ij+1}, y_{ij}) = 0.
 \end{aligned}$$

The extension to polygons with even higher numbers of sides is straightforward; one example is illustrated on the right of Fig. 5.4. The motivation for consideration of these

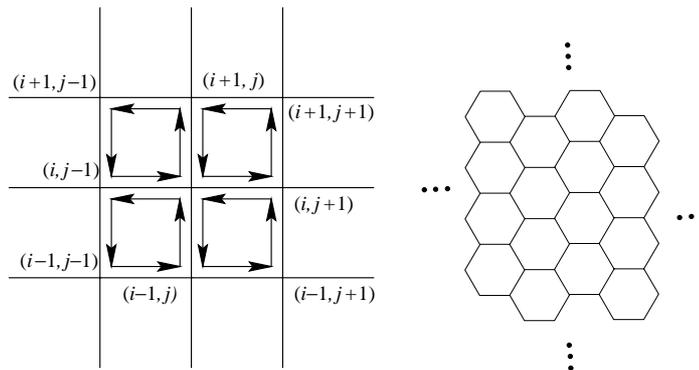


Fig. 5.4. On the left, the method based on rectangles; on the right, a possible method based on hexagons

extensions is enhancing the stability of the triangle-based method in the nonlinear wave example just above.

Example: Nonlinear wave equation, rectangles. Think of each rectangle Δ as having length h and height k , and each discrete jet $(y_{\Delta^1}, y_{\Delta^2}, y_{\Delta^3}, y_{\Delta^4})$ being associated to the continuous jet

$$\frac{\partial \phi}{\partial x^0}(p) = \frac{y_{i,j+1} - y_{i,j}}{h}, \quad \frac{\partial \phi}{\partial x^1}(p) = \frac{1}{2} \left(\frac{y_{i+1,j} - y_{i,j}}{k} + \frac{y_{i+1,j+1} - y_{i,j+1}}{k} \right),$$

where p is the center of the rectangle. This leads to the discrete Lagrangian

$$L = \frac{1}{2} \left(\frac{y_2 - y_1}{h} \right)^2 - \frac{1}{2} \left(\frac{y_4 - y_1}{2k} + \frac{y_3 - y_2}{2k} \right)^2 + N \left(\frac{y_1 + y_2 + y_3 + y_4}{4} \right). \quad (5.9)$$

If, for brevity, we set

$$\bar{y}_{ij} \equiv \frac{y_{ij} + y_{i,j+1} + y_{i+1,j+1} + y_{i+1,j}}{4},$$

then one verifies that the DELF equations become

$$\begin{aligned} & \left[\frac{1}{2} \frac{y_{i+1,j} - 2y_{i,j} + y_{i-1,j}}{k^2} + \frac{1}{4} \frac{y_{i+1,j+1} - 2y_{i,j+1} + y_{i-1,j+1}}{k^2} \right. \\ & \quad \left. + \frac{1}{4} \frac{y_{i+1,j-1} - 2y_{i,j-1} + y_{i-1,j-1}}{k^2} \right] - \left[\frac{y_{i,j+1} - 2y_{i,j} + y_{i,j-1}}{h^2} \right] \\ & \quad + \frac{1}{4} \left[N'(\bar{y}_{ij}) + N'(\bar{y}_{i,j-1}) + N'(\bar{y}_{i-1,j-1}) + N'(\bar{y}_{i-1,j}) \right] = 0, \end{aligned}$$

which, if we make the definitions

$$\begin{aligned} \partial_h^2 y_{ij} &\equiv y_{i,j+1} - 2y_{i,j} + y_{i,j-1}, & \partial_k^2 y_{ij} &\equiv y_{i+1,j} - 2y_{i,j} + y_{i-1,j}, \\ \bar{N}'(\bar{y}_{ij}) &\equiv \frac{1}{4} \left[N'(\bar{y}_{ij}) + N'(\bar{y}_{i,j-1}) + N'(\bar{y}_{i-1,j-1}) + N'(\bar{y}_{i-1,j}) \right], \end{aligned}$$

is (more compactly)

$$\frac{1}{k^2} \left[\frac{1}{4} \partial_k^2 y_{i,j+1} + \frac{1}{2} \partial_k^2 y_{i,j} + \frac{1}{4} \partial_k^2 y_{i,j-1} \right] - \frac{1}{h^2} \partial_h^2 y_{ij} + \bar{N}'(\bar{y}_{ij}) = 0. \quad (5.10)$$

These are implicit equations which must be solved for $y_{i,j+1}$, $1 \leq i \leq N$, given $y_{i,j}$, $y_{i,j-1}$, $1 \leq i \leq N$; rearranging, an iterative form equivalent to (5.10) is

$$\begin{aligned} & - \left(\frac{h^2}{2(h^2 + 2k^2)} \right) y_{i+1,j+1} + y_{i,j+1} - \left(\frac{h^2}{2(h^2 + 2k^2)} \right) y_{i-1,j+1} \\ & = \frac{h^2}{h^2 + 2k^2} \left((y_{i+1,j} - 2y_{i,j} + y_{i-1,j}) + \frac{1}{2} (y_{i+1,j-1} - 2y_{i,j-1} + y_{i-1,j-1}) \right) \\ & \quad + \frac{2k^2}{h^2 + 2k^2} (2y_{i,j} - y_{i,j-1}) \\ & \quad + \frac{h^2 k^2}{2(h^2 + 2k^2)} (N'(\bar{y}_{ij}) + N'(\bar{y}_{i,j-1}) + N'(\bar{y}_{i-1,j-1}) + N'(\bar{y}_{i-1,j})). \end{aligned}$$

In the case of the sine-Gordon equation the values of the field ought to be considered as lying in \mathbb{S}^1 , by virtue of the vertical symmetry $y \mapsto y + 2\pi$. Soliton solutions for example will have a jump of 2π and the method will fail unless field values at close-together spacetime points are differenced modulo 2π . As a result it becomes important to calculate using integral multiples of small field-dependent quantities, so that it is clear when to discard multiples of 2π , and for this the above iterative form is inconvenient. But if we define

$$\partial_h^1 y_{ij} \equiv y_{i,j+1} - y_{ij}, \quad \partial_k^1 y_{ij} \equiv y_{i+1,j} - y_{ij},$$

then there is the following iterative form, again equivalent to (5.10),

$$\begin{aligned} y_{i,j+1} &= y_{ij} + \partial_h^1 y_{ij}, \quad \text{and} \\ &- \left(\frac{h^2}{2(h^2 + 2k^2)} \right) \partial_h^1 y_{i+1,j} + \partial_h^1 y_{i,j} - \left(\frac{h^2}{2(h^2 + 2k^2)} \right) \partial_h^1 y_{i-1,j} \\ &= \frac{h^2}{h^2 + 2k^2} (3\partial_k^2 y_{ij} + \partial_k^2 y_{i,j-1}) + \frac{2k^2}{h^2 + 2k^2} \partial_h^1 y_{ij} \\ &\quad + \frac{h^2 k^2}{2(h^2 + 2k^2)} \bar{N}'(\bar{y}_{ij}). \end{aligned} \tag{5.11}$$

One can also modify (5.9) so as to treat space and time symmetrically, which leads to the discrete Lagrangian

$$\begin{aligned} L &= \frac{1}{2} \left(\frac{y_2 - y_1}{2h} + \frac{y_3 - y_4}{2h} \right)^2 - \frac{1}{2} \left(\frac{y_4 - y_1}{2k} + \frac{y_3 - y_2}{2k} \right)^2 \\ &\quad + N \left(\frac{y_1 + y_2 + y_3 + y_4}{4} \right), \end{aligned}$$

and one verifies that the DELF equations become

$$\begin{aligned} &\frac{1}{k^2} \left[\frac{1}{4} \partial_k^2 y_{i,j+1} + \frac{1}{2} \partial_k^2 y_{ij} + \frac{1}{4} \partial_k^2 y_{i,j-1} \right] \\ &- \frac{1}{h^2} \left[\frac{1}{4} \partial_h^2 y_{i+1,j} + \frac{1}{2} \partial_h^2 y_{ij} + \frac{1}{4} \partial_h^2 y_{i-1,j} \right] + \bar{N}'(\bar{y}_{ij}) = 0, \end{aligned} \tag{5.12}$$

an equivalent iterative form of which is

$$\begin{aligned} y_{i,j+1} &= y_{ij} + \partial_h^1 y_{ij}, \quad \text{and} \\ &- \left(\frac{h^2 - k^2}{2(h^2 + k^2)} \right) \partial_h^1 y_{i+1,j} + \partial_h^1 y_{i,j} - \left(\frac{h^2 - k^2}{2(h^2 + k^2)} \right) \partial_h^1 y_{i-1,j} \\ &= \frac{h^2}{2(h^2 + k^2)} (3\partial_k^2 y_{ij} + \partial_k^2 y_{i,j-1}) \\ &\quad + \frac{h^2}{2(h^2 + k^2)} (2\partial_h^1 y_{ij} + \partial_h^1 y_{i+1,j} + \partial_h^1 y_{i-1,j}) \\ &\quad + \frac{h^2 k^2}{2(h^2 + k^2)} \bar{N}'(\bar{y}_{ij}). \end{aligned} \tag{5.13}$$

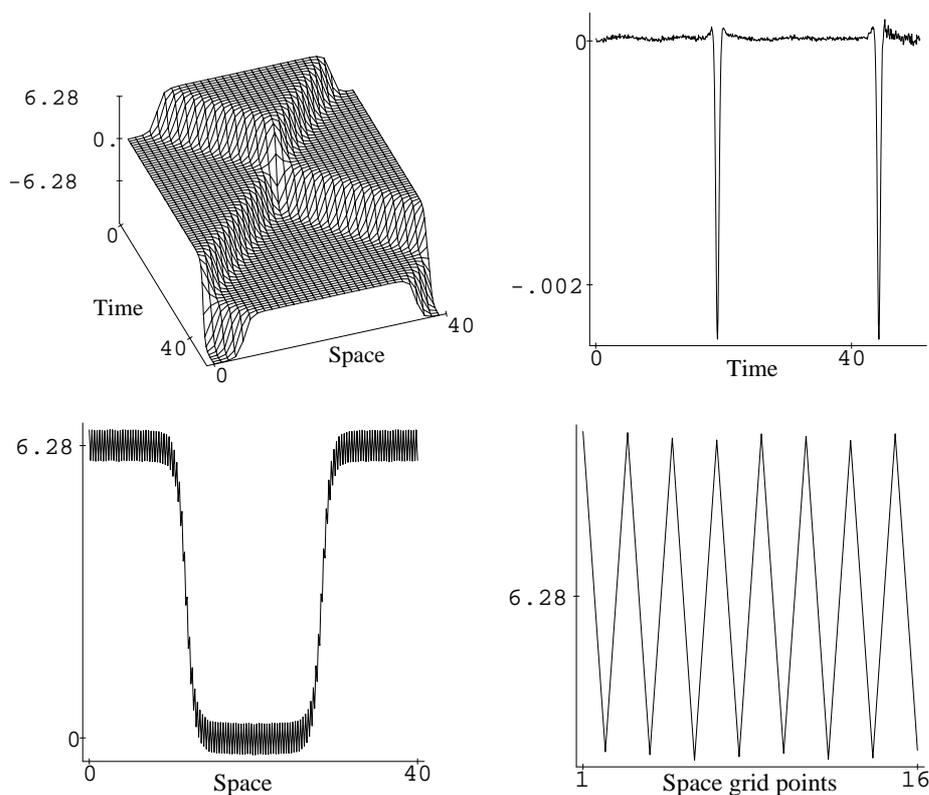


Fig. 5.5. *Top left:* the wave forms for a two soliton kink and antikink collision using (5.12). *Top right:* the energy error. *Bottom left:* the wave form at time $t \approx 11855$. *Bottom right:* the portion of the bottom left graph for spatial grid points 1 . . . 16

5.2. Numerical checks. While the focus of this article is not the numerical implementation of the integrators which we have derived, we have, nevertheless, undertaken some preliminary numerical investigations of our multisymplectic methods in the context of the sine-Gordon equation with periodic boundary conditions.

The rectangle-based multisymplectic method. The top half of Fig. 5.5 shows a simulation of the collision of “kink” and “antikink” solitons for the sine-Gordon equation, using the rectangle-based multisymplectic method (5.12). In the bottom half of that figure we show the result of running that simulation until the solitons have undergone about 460 collisions; shortly after this the simulation stops because the iteration (5.13) diverges. The anomalous spatial variations in the waveform of the bottom left of Fig. 5.5 have period 2 spatial grid divisions and are shown in finer scale on the bottom right of that figure. These variations are reminiscent of those found in Ablowitz, Herbst and Schober [1996] for the completely integrable discretization of Hirota, where the variations are attributed to independent evolution of waveforms supported on even vs. odd grid points. Observation of (5.12) indicates what is wrong: the nonlinear term N contributes to (5.12) in a way that will average out these variations, and consequently, once they have begun, (5.12) tends to continue such variations via the linear wave equation. In Ablowitz et. al., the situation is rectified when the number of spatial grid points is not even, and this is the case

for (5.12) as well. This is indicated on the left of Fig. 5.6, which shows the waveform after about 5000 soliton collisions when $N = 255$ rather than $N = 256$. Figure 5.7 summarizes the evolution of energy error⁴ for that simulation.

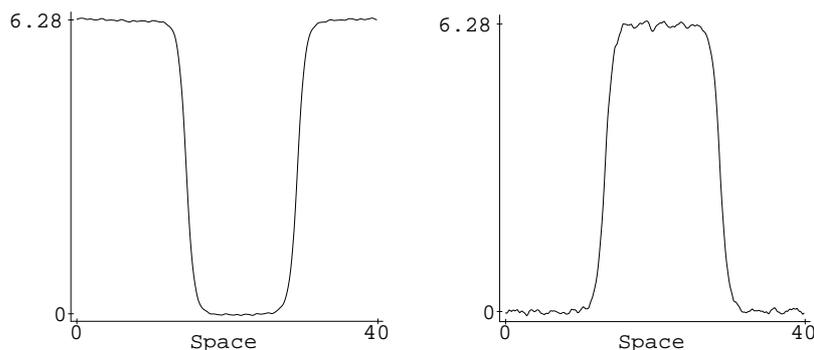


Fig. 5.6. On the *left*, the final wave form (after about 5000 soliton collisions at $t \approx 129133$) obtained using the rectangle-based multisymplectic method (5.12). On the *right*, the final waveform (at $t \approx 129145$) from the energy-conserving method (5.14) of Vu-Quoc and Li. In both simulations, temporal drift is occurring. For this reason the waveforms are inverted with respect to one another; moreover, the separate solitons are drifting at slightly different rates, as indicated by the off-center waveforms

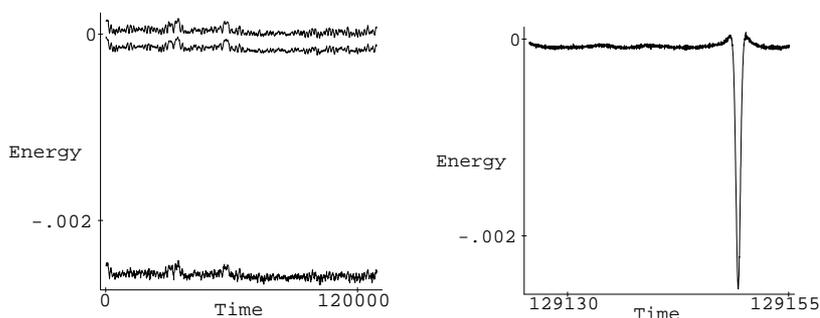


Fig. 5.7. On the *left*, the energy error corresponding to our multisymplectic method (5.13) for 5000 soliton collisions; the three graphs correspond to the minimum, average, and maximum energy error over consecutive 5000 time step regions. On the *right*, the final energy error (i.e. the energy error after 5000 soliton collisions), which can be compared with the initial energy error plot in the top left of Fig. 5.5

⁴ The discrete energy that we calculated was

$$\sum_{i=1}^N \left(\frac{1}{2} \left(\frac{y_{i,j+1} - y_{i,j}}{2h} + \frac{y_{i+1,j+1} - y_{i+1,j}}{2h} \right)^2 + \frac{1}{2} \left(\frac{y_{i+1,j} - y_{i,j}}{2k} + \frac{y_{i+1,j+1} - y_{i,j+1}}{2k} \right)^2 - N(\bar{y}_{i,j}) \right).$$

Initial data. For the two-soliton-collision simulations, we used the following initial data: $h = k/8$ (except $h = k/16$ where noted), where $k = 40/N$ and $N = 255$ spatial grid points (except Fig. 5.5 where $N = 256$). The circle that is space should be visualized as having circumference $L = 40$. Let $\kappa = 1 - \epsilon$, where $\epsilon = 10^{-6}$, $\tilde{L} = L/4 = 10$,

$$P = 2 \int_0^{1/\kappa} \frac{1}{\sqrt{1-y^2}\sqrt{1-\kappa^2y^2}} dy \approx 15.90, \quad c = \sqrt{1 - \frac{\tilde{L}^2}{\kappa^2 P^2}} \approx .7773,$$

and

$$\tilde{\phi}(x) \equiv 2 \arcsin \left(\operatorname{sn} \left(\frac{x}{\kappa\sqrt{1-c^2}}; \kappa \right) \right).$$

Then $\tilde{\phi}(x - ct)$ is a kink solution if space has a circumference of \tilde{L} . This kink and an oppositely moving antikink (but placed on the last quarter of space) made up the initial field, so that $y_{i0} = \phi(40(i - 1)/N)$, $i = 1, \dots, N$, where

$$\phi(x) \equiv \begin{cases} \phi(x) & 0 \leq x < L/4 \\ 2\pi & L/4 \leq x < 3L/4 \\ 2\pi - \phi(x - 3L/4) & 3L/4 \leq x < L \end{cases},$$

while $y_{i1} = y_{i0} + \dot{\phi}(40(i - 1)/N)h$, where

$$\dot{\phi}(x) \equiv \begin{cases} (\phi(x - hc) - \phi(x))/h & 0 \leq x < L/4 \\ 0 & L/4 \leq x < 3L/4 \\ -(\phi(x - hc) - \phi(x))/h & 3L/4 \leq x < L \end{cases}.$$

Comparison with energy-conserving methods. As an example of how our method compares with an existing method, we considered the energy-conserving method of Vu-Quoc and Li [1993], p. 354:

$$\frac{1}{k^2} \left[\frac{1}{4} \partial_k^2 y_{i,j+1} + \frac{1}{2} \partial_k^2 y_{i,j} + \frac{1}{4} \partial_k^2 y_{i,j-1} \right] - \frac{1}{h^2} \partial_h^2 y_{i,j} + \frac{1}{2} \left(\frac{N(y_{i,j+1}) - N(y_{i,j})}{y_{i,j+1} - y_{i,j}} + \frac{N(y_{i,j}) - N(y_{i,j-1})}{y_{i,j} - y_{i,j-1}} \right) = 0. \tag{5.14}$$

This has an iterative form similar to (5.13) and is quite comparable with (5.10) and (5.12) in terms of the computation required. Our method seems to preserve the soliton waveform better than (5.14), as is indicated by comparison of the left and right Fig. 5.6.

In regards to the closely related papers Vu-Quoc and Li [1993] and Li and Vu-Quoc [1995], we could not verify in our simulations that their method conserves energy, nor could we verify their *proof* that their method conserves energy. So, as a further check, we implemented the following energy-conserving method of Guo, Pascual, Rodriguez, and Vazquez [1986]:

$$\partial_k^2 y_{i,j} - \partial_h^2 y_{i,j} + \frac{N(y_{i,j+1}) - N(y_{i,j-1})}{y_{i,j+1} - y_{i,j-1}}, \tag{5.15}$$

which conserves the discrete energy

$$\sum_{j=1}^N \left(\frac{1}{2} \frac{(y_{i,j+1} - y_{i,j})(y_{i,j} - y_{i,j-1})}{h^2} + \frac{1}{2} \left(\frac{y_{i+1,j} - y_{i,j}}{k} \right)^2 - N(y_{i,j}) \right).$$

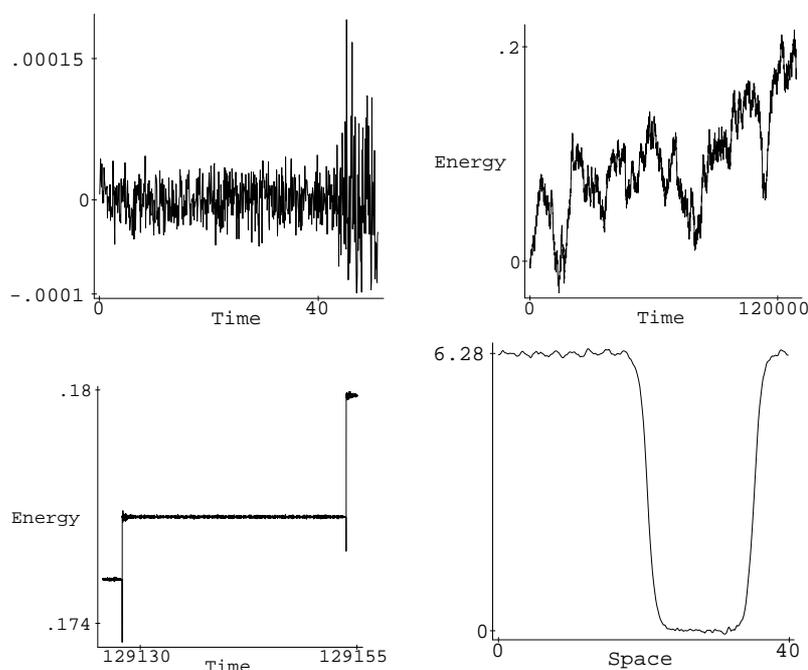


Fig. 5.8. A long-time simulation using the energy-conserving method (5.15) of Guo et al. *Above left:* the initial energy error. *Above right:* the average energy error over consecutive 5000 time step regions (the maximum and minimum closely parallel the average). *Below left:* the final energy error. *Below right:* the final waveform at $t \approx 129149$

This method diverged after just 345 soliton collisions. As can be seen from (5.15), the nonlinear potential N enters as a difference over two grid spacings, which suggests that halving the time step might result in a fairer comparison with the methods (5.12) or (5.14). With this advantage, method (5.15) was able to simulate 5000 soliton collisions, with a waveform degradation similar to the energy-conserving method (5.14), as shown at the bottom right of Fig. 5.8. The same figure also shows that, although the energy behavior of (5.15) is excellent for short time simulations, it drifts significantly over long times, and the final energy error has a peculiar appearance. Figure 5.9 shows the time evolution of the waveform through the soliton collision that occurs just before the simulation stops. Apparently, at the soliton collisions, significant high frequency oscillations are present, and these are causing the jumps in the energy error in the bottom left plot of Fig. 5.8. This error then accumulates due to the energy-conserving property of the method. In these simulations, so as to guard against the possibility that this behavior of the energy was due to inadequately solving the implicit Eq. (5.15), we imposed a minimum limit of 3 iterations in the corresponding iterative loop, whereas this loop would otherwise have converged after just 1 iteration.

Comparison with the triangle-based multisymplectic method. The discrete second derivatives in the method (5.15) are the same as in the triangle-based multisymplectic method (5.8); these derivatives are simpler than either our rectangle-based multisymplectic method (5.12) or the energy-conserving method of Vu-Quoc and Li (5.14). To explore this we implemented the triangle-based multisymplectic method (5.8). Even

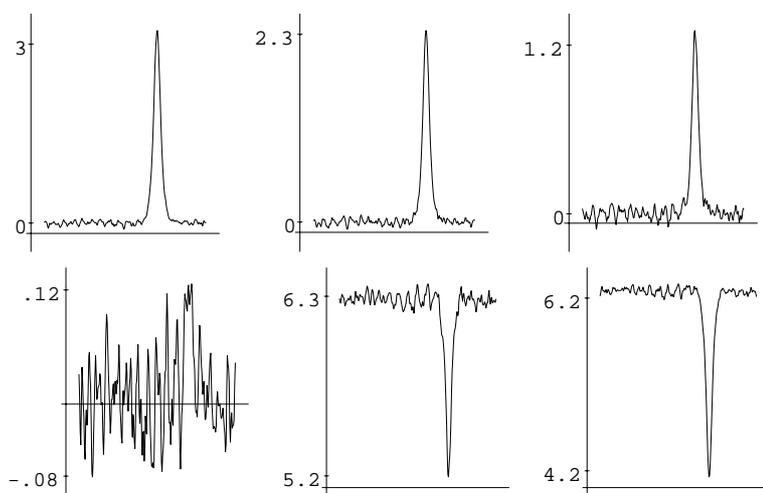


Fig. 5.9. The soliton collision at time $t \approx 129130$, after the energy-conserving method (5.15) of Guo et al. has simulated about 5000 soliton collisions. The solitons collide beginning at the top left and proceed to the top right, then to the bottom left, and finally to the bottom right. The vertical scales are not constant and visually exaggerate the high frequency oscillations, which are small on the scale 0 to 2π

with the less complicated discrete second derivatives our triangle-based multisymplectic method simulated 5000 soliton collisions with comparable energy⁵ and waveform preservation properties as the rectangle-based multisymplectic method (5.12), as shown in Fig. 5.11. Figure 5.10 shows the time evolution of the waveform through the soliton collision just before the simulation stops, and may be compared to Fig. 5.9. As can be seen, the high frequency oscillations that are present during the soliton collisions are smaller and smoother for the triangle-based multisymplectic method than for the energy-conserving method (5.15). A similar statement is true irrespective which of the two multisymplectic or two energy conserving methods we tested, and is true all along the waveform, irrespective of whether or not a soliton collision is occurring.

Summary. Our multisymplectic methods are finite difference methods that are computationally competitive with existing finite difference methods. Our methods show promise for long-time simulations of conservative partial differential equations, in that, for long-time simulations of the sine-Gordon equation, our method 1) had superior energy-conserving behavior, *even when compared with energy-conserving methods*; 2) better preserved the waveform than energy-conserving methods; and 3) exhibited superior stability, in that our methods excited smaller and more smooth high frequency oscillations than energy-conserving methods. However, further numerical investigation is certainly necessary to make any lasting conclusions about the long-time behavior of our integrator.

⁵ The discrete energy that we calculated was

$$\sum_{i=1}^N \left(\frac{1}{2} \left(\frac{y_{i,j+1} - y_{i,j}}{h} \right)^2 + \frac{1}{2} \left(\frac{y_{i+1,j} - y_{i,j}}{2k} \right)^2 - N(\bar{y}_{i,j}) \right).$$

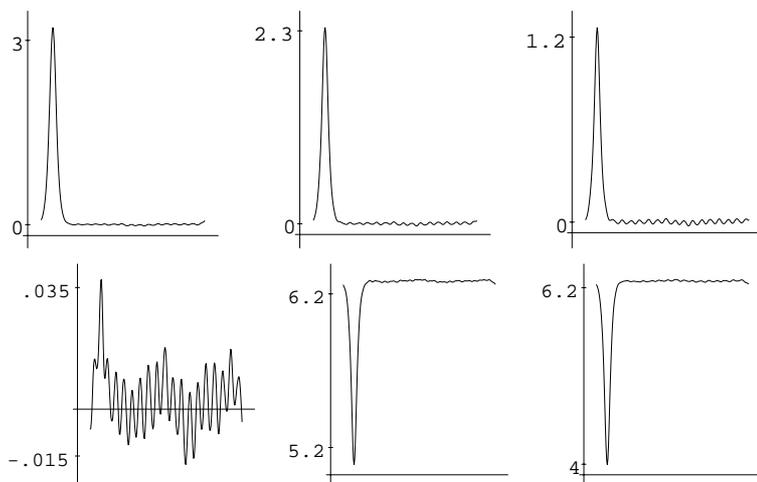


Fig. 5.10. Similar to the above plot but for our triangle-based multisymplectic method (5.8)

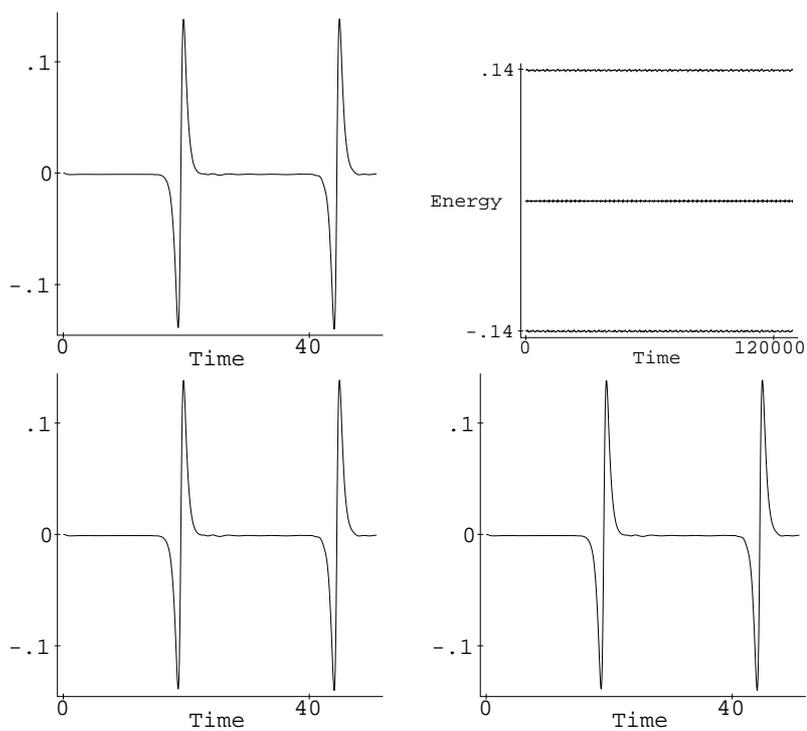


Fig. 5.11. A simulation of 5000 soliton collisions using the triangle-based multisymplectic method (5.8). Above left: The initial energy error. Above right: The minimum, average and maximum energy as in the left of Fig. 5.7. Below left: the final waveform (at $t \approx 129130$). Below right: the final energy error

The programs. The programs that were used in the preceding simulations are “C” language implementations of the various methods. A simple tridiagonal LUD method was used to solve the linear equations (e.g. the left side of (5.13)), as in Vu-Quoc and Li [1993], p. 379. An 8th order extrapolator was used to provide a seed for the implicit step. All calculations were performed in double precision while the implicit step was terminated when the fields ceased to change to single precision; the program’s output was in single precision. The extrapolation usually provided a seed accurate enough so that the methods became practically *explicit*, in that for many of the time-steps the first or second run through the iterative loops solving the implicit equations solved those equations to single precision. However, in the absence of a regular spacetime grid the expenses of the extrapolation and solving the linear equation would grow. Our programs are freely available at URL <http://www.cds.caltech.edu/shkoller/mps>.

6. Concluding Remarks

Here we make a few miscellaneous comments and remark on some work planned for the future.

Lagrangian reduction. As mentioned in the text, it is useful to have a discrete counterpart to the Lagrangian reduction of Marsden and Scheurle [1993a,b], Holm, Marsden and Ratiu [1998a] and Cendra, Marsden and Ratiu [1998]. We sketch briefly how this theory might proceed. This reduction can be done for both the case of “particle mechanics” and for field theory.

For particle mechanics, the simplest case to start with is an invariant (say left) Lagrangian on the tangent bundle of a Lie group: $L : TG \rightarrow \mathbb{R}$. The reduced Lagrangian is $l : \mathfrak{g} \rightarrow \mathbb{R}$ and the corresponding Euler–Poincaré equations have a variational principle of Lagrange d’Alembert type in that there are constraints on the allowed variations. This situation is described in Marsden and Ratiu [1994].

The discrete analogue of this would be to replace a discrete Lagrangian $\mathbb{L} : G \times G \rightarrow \mathbb{R}$ by a reduced discrete Lagrangian $\ell : G \rightarrow \mathbb{R}$ related to \mathbb{L} by

$$\ell(g_1 g_2^{-1}) = \mathbb{L}(g_1, g_2).$$

In this situation, the algorithm from $G \times G$ to $G \times G$ reduces to one from G to G and it is generated by ℓ in a way that is similar to that for \mathbb{L} . In addition, the discrete variational principle for \mathbb{L} which states that one should find critical points of

$$\mathbb{L}(g_1, g_2) + \mathbb{L}(g_2, g_3)$$

with respect to g_2 to implicitly define the map $(g_1, g_2) \mapsto (g_2, g_3)$, reduces naturally to the following principle: Find critical points of $\ell(g) + \ell(h)$ with respect to variations of g and h of the form $g\xi := L_g\xi$ and $\xi h = R_h\xi$, where L_g and R_h denote left and right translation and where $\xi \in \mathfrak{g}$. In other words, one sets to zero, the derivative of the sum $\ell(gg_\epsilon^{-1}) + \ell(g_\epsilon h)$ with respect to ϵ at $\epsilon = 0$ for a curve g_ϵ in G that passes through the identity at $\epsilon = 0$. This defines (with caveats of regularity as before) a map of G to itself, which is the reduced algorithm. This algorithm can then be used to advance points in $G \times G$ itself, by advancing each component by the reduced trajectory, reproducing the algorithm on $G \times G$. In addition, this can be used with the adjoint or coadjoint action to advance points in \mathfrak{g}^* to approximate the Lie–Poisson dynamics.

These equations for a discrete map, say $\phi_\ell : G \rightarrow G$ generated by ℓ on G are called the **discrete Euler–Poincaré equations** as they are the discrete analogue of the Euler–Poincaré equations on \mathfrak{g} . Notice that, at least in theory, computation can be done for this map first and then the dynamics on $G \times G$ is easily reconstructed by simply advancing each pair as follows: $(g_1, g_2) \mapsto (hg_1, \bar{h}g_2)$, where $\bar{h} = \phi_\ell(g_1^{-1}g_2) = \phi_\ell(h)$.

If one identifies the discrete Lagrangians with generating functions (as explained in Wendlandt and Marsden [1997]) then the reduced Lagrangian generates the reduced algorithm in the sense of Ge and Marsden [1988], and this in turn is closely related to the Lie–Poisson–Hamilton–Jacobi theory.

Next, consider the more general case of TQ with its discretization $Q \times Q$ with a group action (assumed to be free and proper) by a Lie group G . The reduction of TQ by the action of G is TQ/G , which is a bundle over $T(Q/G)$ with fiber isomorphic to \mathfrak{g} . The discrete analogue of this is $(Q \times Q)/G$ which is a bundle over $(Q/G) \times (Q/G)$ with fiber isomorphic to G itself. The projection map $\pi : (Q \times Q)/G \rightarrow (Q/G) \times (Q/G)$ is given by $[(q_1, q_2)] \mapsto ([q_1], [q_2])$ where $[\]$ denotes the relevant equivalence class. Notice that in the case in which $Q = G$ this bundle is “all fiber”. The reduced discrete Euler–Lagrange equations are similar to those in the continuous case, in which one has shape equations coupled with a version of the discrete Euler–Poincaré equations.

Of course all of the machinery in the continuous case can be contemplated here too, such as stability theory, geometric phases, etc. In addition, it would be useful to generalize this Lagrangian reduction theory to the multisymplectic case. All of these topics are planned for other papers.

Role of uniformity of the grid. Consider an autonomous, continuous Lagrangian $\mathcal{L} : TQ \rightarrow \mathbb{R}$ where, for simplicity, Q is an open submanifold of Euclidean space. Imagine some *not necessarily uniform* temporal grid (t_0, t_1, \dots) of \mathbb{R} , so that $t_0 < t_1 < t_2 < \dots$. In this situation, it is natural to consider the discrete action

$$S = \sum_{k=1}^n L_k(q_k, q_{k-1}) \equiv \sum_{k=1}^n \mathcal{L} \left(\frac{q_k + q_{k-1}}{2}, \frac{q_k - q_{k-1}}{t_k - t_{k-1}} \right) (t_k - t_{k-1}). \quad (6.1)$$

This action principle deviates from the action principle (3.1) of Sect. 3 in that the discrete Lagrangian density depends explicitly on k . Of course nonautonomous continuous Lagrangians also yield k -dependent discrete Lagrangian densities, irrespective of uniformity of the grid. Thus, nonuniform temporal grids or nonautonomous Lagrangians give rise to discrete Lagrangian densities which are more general than those we have considered in Sect. 3. For field theories, the Lagrangian in the action (5.1) depends on the spacetime variables already, through its explicit dependence on the triangle Δ . However, it is only in the context of a uniform grid that we have experimented numerically and only in that context that we have discussed the significance of the discrete multisymplectic form formula and the discrete Noether theorem.

Using (6.1) as an example, will now indicate why the issue of grid uniformity may not be serious. The DEL equations corresponding to the action (6.1) are

$$\frac{\partial L_k}{\partial q_1}(q_k, q_{k-1}) + \frac{\partial L_{k+1}}{\partial q_2}(q_{k+1}, q_k) = 0, \quad k = 1, 2, \dots, \quad (6.2)$$

and this gives evolution maps $F_{k+1,k} : Q \times Q \rightarrow Q \times Q$ defined so that

$$F_{k+1,k}(q_k, q_{k-1}) = (q_{k+1}, q_k), \quad k = 1, 2, \dots$$

when (6.2) holds. For the canonical 1-forms corresponding to (3.4) and (3.5) we have the k -dependent one forms

$$\theta_{L,k}^-(q_1, q_0) \cdot (\delta q_1, \delta q_0) \equiv \frac{\partial L_k}{\partial q_0}(q_1, q_0) \delta q_0, \quad (6.3)$$

and

$$\theta_{L,k}^+(q_1, q_0) \cdot (\delta q_1, \delta q_0) \equiv \frac{\partial L_k}{\partial q_1}(q_1, q_0) \delta q_1, \quad (6.4)$$

and Eqs. (3.7) and (3.9) become

$$F_{k+1,k}^*(d\theta_{L,k}^+) = -d\theta_{L,k+1}^-, \quad d\theta_{L,k}^- + d\theta_{L,k}^+ = 0 \quad (6.5)$$

respectively. Together, these two equations give

$$F_{k+1,k}^*(d\theta_{L,k}^+) = d\theta_{L,k+1}^+, \quad (6.6)$$

and if we set

$$F_k \equiv F_{k,k-1} \circ F_{k-1,k-2} \circ \cdots \circ F_{2,1}$$

then (6.6) chain together to imply $F_k^*(d\theta_{L,1}^+) = d\theta_{L,k}^+$. This appears less than adequate since it merely says that the pull back by the evolution of a certain 2-form is, in general, a different 2-form. The significant point to note, however, is that *this situation may be repaired at any k simply by choosing $L_k = L_1$* . It is easily verified that the analogous statement is true with respect to momentum preservation via the discrete Noether theorem.

Specifically, imagine integrating a symmetric autonomous mechanical system in a timestep adaptive way with Eqs. (6.2). As the integration proceeds, various timesteps are chosen, and if momentum is monitored it will show a dependence on those choices. *A momentum-preserving symplectic simulation may be obtained by simply choosing the last timestep to be of equal duration to the first.* This is the highly desirable situation which gives us some confidence that grid uniformity is a nonissue. There is one caveat: symplectic integration algorithms are evolutions which are high frequency perturbations of the actual system, the frequency being the inverse of the timestep, which is generally far smaller than the time scale of any process in the simulation. However, timestep adaptation schemes will make choices on a much larger time scale than the timestep itself, and then drift in the energy will appear on this larger time scale. A meaningful long-time simulation cannot be expected in the unfortunate case that the timestep adaptation makes repeated choices in a way that resonates with some process of the system being simulated.

The sphere. The sphere cannot be generally uniformly subdivided into spherical triangles; however, a good approximately uniform grid is obtained as follows: start from an inscribed icosahedron which produces a uniform subdivision into twenty spherical isosceles triangles; these are further subdivided by halving their sides and joining the resulting points by short geodesics.

Elliptic PDEs. The variational approach we have developed allows us to examine the multisymplectic structure of elliptic boundary value problems as well. For a given Lagrangian, we form the associated action function, and by computing its first variation, we obtain the unique multisymplectic form of the elliptic operator. The multisymplectic

form formula contains information on how symplecticity interacts with spatial boundaries. In the case of two spatial dimensions, $X = \mathbb{R}^2$, $Y = \mathbb{R}^3$, we see that Eq. (4.36) gives us the conservation law

$$\operatorname{div} \mathcal{X} = 0,$$

where the vector $\mathcal{X} = (\omega^0(j^1V, j^1W), \omega^1(j^1V, j^1W))$.

Furthermore, using our generalized Noether theory, we may define momentum-mappings of the elliptic operator associated with its symmetries. It turns out that for important problems of spatial complexity arising in, for example, pattern formation systems, the covariant Noether current intrinsically contains the constrained toral variational principles whose solutions are the complex patterns (see Marsden and Shkoller [1998]).

There is an interesting connection between our variational construction of multi-symplectic-momentum integrators and the finite element method (FEM) for elliptic boundary value problems. FEM is also a variationally derived numerical scheme, fundamentally differing from our approach in the following way: whereas we form a discrete action sum and compute its first variation to obtain the discrete Euler–Lagrange equations, in FEM, it is the original continuum action function which is used together with a projection of the fields and their variations onto appropriately chosen finite-dimensional spaces. One varies the projected fields and integrates such variations over the spatial domain to recover the discrete equations. In general, the two discretization schemes do not agree, but for certain classes of finite element bases with particular integral approximations, the resulting discrete equations match the discrete Euler–Lagrange equations obtained by our method, and are hence naturally multisymplectic.

To illustrate this concept, we consider the Gregory and Lin method of solving two-point boundary value problems in optimal control. In this scheme, the discrete equations are obtained using a finite element method with a basis of linear interpolants. Over each one-dimensional element, let N_1 and N_2 be the two linear interpolating functions. As usual, we define the action function by $S(q) = \int_0^T L(q(t), \dot{q}(t))dt$. Discretizing the interval $[0, T]$ into $N+1$ uniform elements, we may write the action with fields projected onto the linear basis as

$$S(q) = \sum_{k=0}^{N-1} \int_k^{k+1} L(\{N_1\phi_k + N_2\phi_{k+1}\}, \{\dot{N}_1\phi_k + \dot{N}_2\phi_{k+1}\})dt.$$

Since the Euler–Lagrange equations are obtained by linearizing the action and hence the Lagrangian, and as the functions N_i are linear, one may easily check that by evaluating the integrals in the linearized equations using a trapezoidal rule, the discrete Euler–Lagrange equations given in (3.3) are obtained. Thus, the Gregory and Lin method is actually a multisymplectic-momentum algorithm.

Applicability to fluid problems. Fluid problems are not literally covered by the theory presented here because their symmetry groups (particle relabeling symmetries) are not vertical. A generalization is needed to cover this case and we propose to work out such a generalization in a future paper, along with numerical implementation, especially for geophysical fluid problems in which conservation laws such as conservation of enstrophy and Kelvin theorems more generally are quite important.

Other types of integrators. It remains to link the approaches here with other types of integrators, such as volume preserving integrators (see, e.g., Kang and Shang [1995], Quispel [1995]) and reversible integrators (see, e.g., Stoffer [1995]). In particular since

volume manifolds may be regarded as multisymplectic manifolds, it seems reasonable that there is an interesting link.

Constraints. One of the very nice things about the Veselov construction is the way it handles constraints, both theoretically and numerically (see Wendlandt and Marsden [1997]). For field theories one would like to have a similar theory. For example, it is interesting that for fluids, the incompressibility constraint can be expressed as a pointwise constraint on the first jet of the particle placement field, namely that its Jacobian be unity. When viewed this way, it appears as a holonomic constraint and it should be amenable to the present approach. Under reduction by the particle relabeling group, such a constraint of course becomes the divergence free constraint and one would like to understand how these constraints behave under both reduction and discretization.

Acknowledgement. We would like to extend our gratitude to Darryl Holm, Tudor Ratiu and Jeff Wendlandt for their time, encouragement and invaluable input. Work of J. Marsden was supported by the California Institute of Technology and NSF grant DMS 96-33161. Work by G. Patrick was partially supported by NSERC grant OGP0105716 and that of S. Shkoller was partially supported by the Cecil and Ida M. Green Foundation and DOE. We also thank the Control and Dynamical Systems Department at Caltech for providing a valuable setting for part of this work.

References

- Arms, J.M., Marsden, J.E. and Moncrief, V.: The structure of the space solutions of Einstein's equations: II Several Killings fields and the Einstein–Yang–Mills equations. *Ann. of Phys.* **144**, 81–106 (1982)
- Ablowitz, M.J., Herbst, B.M., and Schober, C.: On the numerical solution of the Sine-Gordon equation I. Integrable discretizations and homoclinic manifolds. *J. Comp. Phys.* **126**, 299–314 (1996)
- Abraham, R. and Marsden, J.E.: *Foundations of Mechanics*. Second Edition, Reading, MA: Addison-Wesley, 1978
- Arms, J.M., Marsden, J.E. and Moncrief, V.: The structure of the space solutions of Einstein's equations: II Several Killings fields and the Einstein–Yang–Mills equations. *Ann. of Phys.* **144**, 81–106 (1982)
- Arnold, V.I.: *Mathematical Methods of Classical Mechanics*. Graduate Texts in Math. **60**, Berlin–Heidelberg–New York: Springer Verlag, 1978 (Second Edition, 1989)
- Ben-Yu, G, Pascual, P.J., Rodriguez, M.J. and Vazquez, L.: Numerical solution of the sine-Gordon equation. *Appl. Math. Comput.* **18**, 1–14 (1986)
- Bridges, T.J.: Multi-symplectic structures and wave propagation. *Math. Proc. Camb. Phil. Soc.*, **121**, 147–190 (1997)
- Calvo, M.P. and Hairer, E.: Accurate long-time integration of dynamical systems. *Appl. Numer. Math.*, **18**, 95 (1995)
- Cendra, H., Marsden, J.E. and Ratiu, T.S.: Lagrangian reduction by stages. Preprint (1998)
- de Vogelaère, R.: Methods of integration which preserve the contact transformation property of the Hamiltonian equations. Department of Mathematics, University of Notre Dame, Report No. **4** (1956)
- Dragt, A.J. and Finn, J.M. [1979]: Normal form for mirror machine Hamiltonians. *J. Math. Phys.* **20**, 2649–2660
- Ebin, D. and Marsden, J.: Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. Math.* **92**, 102–163 (1970)
- Ge, Z. and Marsden, J.E.: Lie–Poisson integrators and Lie–Poisson Hamilton–Jacobi theory. *Phys. Lett. A* **133**, 134–139 (1988)
- Gotay, M., Isenberg, J. and Marsden, J.E.: Momentum Maps and the Hamiltonian Structure of Classical Relativistic Field Theories, I. Preprint. (1997)
- Gregory, J and Lin, C.: The numerical solution of variable endpoint problems in the calculus of variations, *Lecture Notes in Pure and Appl. Math.*, **127**, 175–183 (1991)
- Guo, B. Y., Pascual, P.J., Rodriguez, M.J. and Vazquez, L.: Numerical solution of the sine-Gordon equation. *Appl. Math. Comput.* **18**, 1–14 (1986)

- Holm, D.D., Marsden, J.E. and Ratiu, T.: The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. in Math.* To appear (1998a)
- Holm, D. D., J.E. Marsden, J.E. and Ratiu, T.: The Euler–Poincaré equations in geophysical fluid dynamics. In: *Proceedings of the Isaac Newton Institute Programme on the Mathematics of Atmospheric and Ocean Dynamics*, Cambridge: Cambridge University Press (to appear) (1998b)
- Li, S. and Vu-Quoc, L.: Finite-difference calculus invariant structure of a class of algorithms for the nonlinear Klein–Gordon equation. *SIAM J. Num. Anal.* **32**, 1839–1875 (1995)
- Marsden, J.E., Patrick, G.W. and Shadwick, W.F. (Eds.): *Integration Algorithms and Classical Mechanics*. Fields Institute Communications **10**, Providence, RI: Am. Math. Society, 1996
- Marsden, J.E. and Ratiu, T.S.: *Introduction to Mechanics and Symmetry*. Texts in Applied Mathematics **17**, Berlin–Heidelberg–New York: Springer-Verlag, 1994
- Marsden, J.E. and Shkoller, S.M.: Multisymplectic geometry, covariant Hamiltonians and water waves. *Math. Proc. Camb. Phil. Soc.* **124** (1998)
- Marsden, J.E. and Wendlandt, J.M.: Mechanical systems with symmetry, variational principles and integration algorithms. *Current and Future Directions in Applied Mathematics*, Edited by M. Alber, B. Hu, and J. Rosenthal, Base–Boston: Birkhäuser, pp. 219–261 (1997)
- McLachlan, R.I. and Scovel, C.: A survey of open problems in symplectic integration. *Fields Institute Communications* **10**, 151–180 (1996)
- Moser, J. and Veselov, A.P.: Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Commun. Math. Phys.* **139**, 217–243 (1991)
- Neishtadt, A.: The separation of motions in systems with rapidly rotating phase. *P.M.M. USSR* **48**, 133–139 (1984)
- Palais, R.S.: *Foundations of global nonlinear analysis*, New York: Benjamin, 1968
- Palais, R.S.: The symmetries of solitons. *Bull. Am. Math. Soc.* **34**, 339–403 (1997)
- Quispel, G.R.W.: Volume-preserving integrators. *Phys. Lett. A.* **206**, 26 (1995)
- Sanz-Serna, J. M. and Calvo, M.: *Numerical Hamiltonian Problems*. London: Chapman and Hall, 1994
- Simo, J.C. and Gonzalez, O.: Assessment of Energy-Momentum and Symplectic Schemes for Stiff Dynamical Systems. *Proc. ASME Winter Annual Meeting*, New Orleans, Dec. 1993
- Stoffer, D.: Variable steps for reversible integration methods, *Computing* **55**, 1 (1995)
- Veselov, A.P.: Integrable discrete-time systems and difference operators. *Funkts. Anal. Prilozhen.* **22**, 1–13 (1988)
- Veselov, A.P.: Integrable Lagrangian correspondences and the factorization of matrix polynomials. *Funkts. Anal. Prilozhen.* **25**, 38–49 (1991)
- Vu-Quoc, L. and Li, S.: Invariant-conserving finite difference algorithms for the nonlinear Klein–Gordon equation. *Comput. Methods Appl. Mech. Engrg.* **107**, 341–391 (1993)
- Wald, R.M.: Variational principles, local symmetries and black hole entropy. *Proc. Lanczos Centenary volume SIAM*, 231–237 (1993)
- Wendlandt, J.M. and Marsden, J.E.: Mechanical integrators derived from a discrete variational principle. *Physica D* **106**, 223–246 (1997)

Communicated by A. Kupiainen