Acta Numerica (2001), pp. 357–514 © Cambridge University Press, 2001

Discrete mechanics and variational integrators

J. E. Marsden and M. West
Control and Dynamical Systems 107-81,
Caltech, Pasadena, CA 91125-8100, USA
E-mail: marsden@cds.caltech.edu
mwest@cds.caltech.edu

This paper gives a review of integration algorithms for finite dimensional mechanical systems that are based on discrete variational principles. The variational technique gives a unified treatment of many symplectic schemes, including those of higher order, as well as a natural treatment of the discrete Noether theorem. The approach also allows us to include forces, dissipation and constraints in a natural way. Amongst the many specific schemes treated as examples, the Verlet, SHAKE, RATTLE, Newmark, and the symplectic partitioned Runge–Kutta schemes are presented.

CONTENTS

PART 1: Discrete variational mechanics

1.1 Introduction 359
1.2 Background: Lagrangian mechanics 365
1.3 Discrete variational mechanics: Lagrangian viewpoint 370
1.4 Background: Hamiltonian mechanics 376
1.5 Discrete variational mechanics: Hamiltonian viewpoint 383
1.6 Correspondence between discrete and continuous mechanics 386
1.7 Background: Hamilton–Jacobi theory 390
1.8 Discrete variational mechanics: Hamilton–Jacobi viewpoint 392

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

PART 2: Variational integrators

2.1 Introduction 394
2.2 Background: Error analysis 397
2.3 Variational error analysis 399
2.4 The adjoint of a method and symmetric methods 402
2.5 Composition methods 405
2.6 Examples of variational integrators 408

PART 3: Forcing and constraints

3.1 Background: Forced systems 421
3.2 Discrete variational mechanics with forces 423
3.3 Background: Constrained systems 430
3.4 Discrete variational mechanics with constraints 438
3.5 Constrained variational integrators 444
3.6 Background: Forced and constrained systems 452
3.7 Discrete variational mechanics with forces and constraints 456

PART 4: Time-dependent mechanics

4.1 Introduction 463
4.2 Background: Extended Lagrangian mechanics 464
4.3 Discrete variational mechanics: Lagrangian viewpoint 472
4.4 Background: Extended Hamiltonian mechanics 480
4.5 Discrete variational mechanics: Hamiltonian viewpoint 486
4.6 Correspondence between discrete and continuous mechanics 490
4.7 Background: Extended Hamilton–Jacobi theory 494
4.8 Discrete variational mechanics: Hamilton–Jacobi viewpoint 496
4.9 Time-dependent variational integrators 497

PART 5: Further topics

5.1 Discrete symmetry reduction 503
5.2 Multisymplectic integrators for PDEs 503
5.3 Open problems 504

References 507

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

PART ONE
Discrete variational mechanics

1.1. Introduction

This paper gives a survey of the variational approach to discrete mechanics and to mechanical integrators. This point of view is not confined to conservative systems, but also applies to forced and dissipative systems, so is useful for control problems (for instance) as well as traditional conservative problems in mechanics. As we shall show, the variational approach gives a comprehensive and unified view of much of the literature on both discrete mechanics as well as integration methods for mechanical systems and we view these as closely allied subjects.

Some of the important topics that come out naturally from this method are symplectic–energy-momentum methods, error analysis, constraints, forcing, and Newmark algorithms. Besides giving an account of methods such as these, we connect these techniques to other recent and exciting developments, including the PDE setting of multisymplectic spacetime integrators (also called AVI, or asynchronous variational integrators), which are being used for problems such as nonlinear wave equations and nonlinear shell dynamics. In fact, one of our points is that by basing the integrators on fundamental mechanical concepts and methods from the outset, one eases the way to other areas, such as continuum mechanics and systems with forcing and constraints.

In the last few years this subject has grown to be very large and an active area of research, with many points of view and many topics. We shall focus here on our own point of view, namely the variational view. Naturally we must omit a number of important topics, but include several of our own. We do make contact with some, but not all, of other topics in the final part of this article and in the brief history below.

As in standard mechanics, some things are easier from a Hamiltonian perspective and others are easier from a Lagrangian perspective. Regarding symplectic integrators from both viewpoints gives greater insight into their properties and derivations. We have tried to give a balanced perspective in this article.

We will assume that the configuration manifold is finite-dimensional. This means that at the outset, we will deal with the context of ordinary differential equations. However, as we have indicated, our approach is closely tied with the variational spacetime multisymplectic approach, which is the approach that is suitable for the infinite-dimensional, PDE context, so an investment in the methodology of this article eases the transition to the corresponding PDE context.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

One of the simple, but important ideas in discrete mechanics is easiest to say from the Lagrangian point of view. Namely, consider a mechanical system with configuration manifold Q. The velocity phase space is then TQ and the Lagrangian is a map $L : TQ \to \mathbb{R}$. In discrete mechanics, the starting point is to replace TQ with $Q \times Q$ and we regard, intuitively, two nearby points as being the discrete analogue of a velocity vector.

There is an important note about constraints that we would like to say at the outset. Recall from basic geometric mechanics (as in Marsden and Ratiu (1999) for instance) that specifying a constraint manifold Q means that one may already have specified constraints: for example, Q may already be a submanifold of a linear space that is specified by constraints. However, when constructing integrators in Section 2.1 we will take Q to be linear, although this is only for simplicity. One way of handling a nonlinear Q is to embed it within a linear space and use the theory of constrained systems: this point of view is developed in Section 3. This approach has computational advantages, but we will also discuss implementations of variational integrators on arbitrary configuration manifolds Q.

1.1.1. History and literature

Of course, the variational view of mechanics goes back to Euler, Lagrange and Hamilton. The form of the variational principle most important for continuous mechanics we use in this article is due, of course, to Hamilton (1834). We refer to Marsden and Ratiu (1999) for additional history, references and background on geometric mechanics.

There have been many attempts at the development of a discrete mechanics and corresponding integrators that we will not attempt to survey in any systematic fashion. The theory of discrete variational mechanics in the form we shall use it (that uses $Q \times Q$ for the discrete analogue of the velocity phase space) has its roots in the optimal control literature of the 1960s: see, for example, Jordan and Polak (1964), Hwang and Fan (1967) and Cadzow (1970). In the context of mechanics early work was done, often independently, by Cadzow (1973), Logan (1973), Maeda (1980, 1981*a*, 1981*b*), and Lee (1983, 1987), by which point the discrete action sum, the discrete Euler–Lagrange equations and the discrete Noether's theorem were clearly understood. This theory was then pursued further in the context of integrable systems in Veselov (1988, 1991) and Moser and Veselov (1991), and in the context of quantum mechanics in Jaroszkiewicz and Norton (1997*a*, 1997*b*) and Norton and Jaroszkiewicz (1998).

The variational view of discrete mechanics and its numerical implementation is further developed in Wendlandt and Marsden (1997*a*) and (1997*b*) and then extended in Kane, Marsden and Ortiz (1999*a*), Marsden, Pekarsky and Shkoller (1999*a*, 1999*b*), Bobenko and Suris (1999*a*, 1999*b*) and Kane,

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Marsden, Ortiz and West (2000). The beginnings of an extension of these ideas to a nonsmooth framework is given in Kane, Repetto, Ortiz and Marsden (1999*b*), and is carried further in Fetecau, Marsden, Ortiz and West (2001).

Other discretizations of Hamilton's principle are given in Mutze (1998), Cano and Lewis (1998) and Shibberu (1994). Other versions of discrete mechanics (not necessarily discrete Hamilton's principles) are given in (for instance) Itoh and Abe (1988), Labudde and Greenspan (1974, 1976*a*, 1976*b*), and MacKay (1992).

Of course, there have been many works on symplectic integration, largely done from other points of view than that developed here. We will not attempt to survey this in any systematic fashion, as the literature is simply too large with too many points of view and too many intricate subtleties. We give a few highlights and give further references in the body of the paper. For instance, we shall connect the variational view with the generating function point of view that was begun in De Vogelaére (1956). Generating function methods were developed and used in, for example, Ruth (1983), Forest and Ruth (1990) and in Channell and Scovel (1990). See also Berg, Warnock, Ruth and Forest (1994), and Warnock and Ruth (1991, 1992). For an overview of symplectic integration, see Sanz-Serna (1992*b*) and Sanz-Serna and Calvo (1994). Qualitative properties of symplectic integration of Hamiltonian systems are given in Gonzalez, Higham and Stuart (1999) and Cano and Sanz-Serna (1997). Long-time energy behaviour for oscillatory systems is studied in Hairer and Lubich (2000). Long-time behaviour of symplectic methods for systems with dissipation is given in Hairer and Lubich (1999). A numerical study of preservation of adiabatic invariants is given in Reich (1999*b*) and Shimada and Yoshida (1996). Backward error analysis is studied in Benettin and Giorgilli (1994), Hairer (1994), Hairer and Lubich (1997) and Reich (1999*a*). Other ideas connected to the above literature include those of Baez and Gilliam (1994), Gilliam (1996), Gillilan and Wilson (1992). For other references see the large literature on symplectic methods in molecular dynamics, such as Schlick, Skeel *et al.* (1999), and for various applications, see Hardy, Okunbor and Skeel (1999), Leimkuhler and Skeel (1994), Barth and Leimkuhler (1996) and references therein.

A single-step variational idea that is relevant for our approach is given in Ortiz and Stainier (1998), and developed further in Radovitzky and Ortiz (1999), and Kane *et al.* (1999*b*, 2000).

Direct discretizations on the Hamiltonian side, where one discretizes the Hamiltonian and the symplectic structure, are developed in Gonzalez (1996*b*) and (1996*a*) and further in Gonzalez (1999) and Gonzalez *et al.* (1999). This is developed and generalized much further in McLachlan, Quispel and Robidoux (1998) and (1999).

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Finally, we mention that techniques of geometric integration in the sense of preserving manifold or Lie group structures, as given in Budd and Iserles (1999), Iserles, Munthe-Kaas and Zanna (2000) and references therein, presumably could, and probably should, be combined with the techniques described herein for a more efficient treatment of certain classes of constraints in mechanical systems. Such an enterprise is for the future.

1.1.2. A simplified introduction

In this section we give a brief overview of how discrete variational mechanics can be used to derive variational integrators. We begin by reviewing the derivation of the Euler–Lagrange equations, and then show how to mimic this process on a discrete level.

For concreteness, consider the Lagrangian system $L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$, where M is a symmetric positive-definite mass matrix and V is a potential function. We work in $\mathbb{R}^n$ or in generalized coordinates and will use vector notation for simplicity, so $q = (q^1, q^2, \ldots, q^n)$. In the standard approach of Lagrangian mechanics, we form the action function by integrating L along a curve $q(t)$ and then compute variations of the action while holding the endpoints of the curve $q(t)$ fixed. This gives

$$\begin{aligned}\delta \int_0^T L\big(q(t),\dot{q}(t)\big)\,\mathrm{d}t &= \int_0^T \left[\frac{\partial L}{\partial q}\cdot\delta q + \frac{\partial L}{\partial \dot{q}}\cdot\delta\dot{q}\right]\mathrm{d}t \\ &= \int_0^T \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\right]\cdot\delta q\,\mathrm{d}t,\end{aligned}$$

where we have used integration by parts and the condition $\delta q(T) = \delta q(0) = 0$. Requiring that the variations of the action be zero for all δq implies that the integrand must be zero for each time t, giving the well-known Euler–Lagrange equations

$$\frac{\partial L}{\partial q}(q,\dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}(q,\dot{q})\right) = 0.$$

For the particular form of the Lagrangian chosen above, this is just

$$M\ddot{q} = -\nabla V(q),$$

which is Newton's equation: mass times acceleration equals force. It is well known that the system described by the Euler–Lagrange equations has many special properties. In particular, the flow on state space is symplectic, meaning that it conserves a particular two-form, and if there are symmetry actions on phase space then there are corresponding conserved quantities of the flow, known as momentum maps.

We will now see how discrete variational mechanics performs an analogue of the above derivation. Rather than taking a position q and velocity $\dot{q}$,

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

consider now two positions q_0 and q_1 and a time-step $h \in \mathbb{R}$. These positions should be thought of as being two points on a curve at time h apart, so that $q_0 \approx q(0)$ and $q_1 \approx q(h)$.

We now consider a discrete Lagrangian $L_d(q_0, q_1, h)$, which we think of as approximating the action integral along the curve segment between q_0 and q_1. For concreteness, consider the very simple approximation to the integral $\int_0^T L \, \mathrm{d}t$ given by using the rectangle rule[1] (the length of the interval times the value of the integrand with the velocity vector replaced by $(q_1 - q_0)/h$):

$$L_d(q_0, q_1, h) = h\left[\left(\frac{q_1 - q_0}{h}\right)^T M\left(\frac{q_1 - q_0}{h}\right) - V(q_0)\right].$$

Next consider a discrete curve of points $\{q_k\}_{k=0}^N$ and calculate the discrete action along this sequence by summing the discrete Lagrangian on each adjacent pair. Following the continuous derivation above, we compute variations of this action sum with the boundary points q_0 and q_N held fixed. This gives

$$\begin{aligned}
&\delta \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}, h) \\
&\quad = \sum_{k=0}^{N-1} \left[D_1 L_d(q_k, q_{k+1}, h) \cdot \delta q_k + D_2 L_d(q_k, q_{k+1}, h) \cdot \delta q_{k+1}\right] \\
&\quad = \sum_{k=1}^{N-1} \left[D_2 L_d(q_{k-1}, q_k, h) + D_1 L_d(q_k, q_{k+1}, h)\right] \cdot \delta q_k,
\end{aligned}$$

where we have used a discrete integration by parts (rearranging the summation) and the fact that $\delta q_0 = \delta q_N = 0$. If we now require that the variations of the action be zero for any choice of δq_k, then we obtain the *discrete Euler–Lagrange equations*

$$D_2 L_d(q_{k-1}, q_k, h) + D_1 L_d(q_k, q_{k+1}, h) = 0,$$

which must hold for each k. For the particular L_d chosen above, we compute

$$\begin{aligned}
D_2 L_d(q_{k-1}, q_k, h) &= M\left(\frac{q_k - q_{k-1}}{h}\right) \\
D_1 L_d(q_k, q_{k+1}, h) &= -\left[M\left(\frac{q_{k+1} - q_k}{h}\right) + h\nabla V(q_k)\right],
\end{aligned}$$

and so the discrete Euler–Lagrange equations are

$$M\left(\frac{q_{k+1} - 2q_k + q_{k-1}}{h^2}\right) = -\nabla V(q_k).$$

[1] As we shall see later, more sophisticated quadrature rules lead to higher-order accurate integrators.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

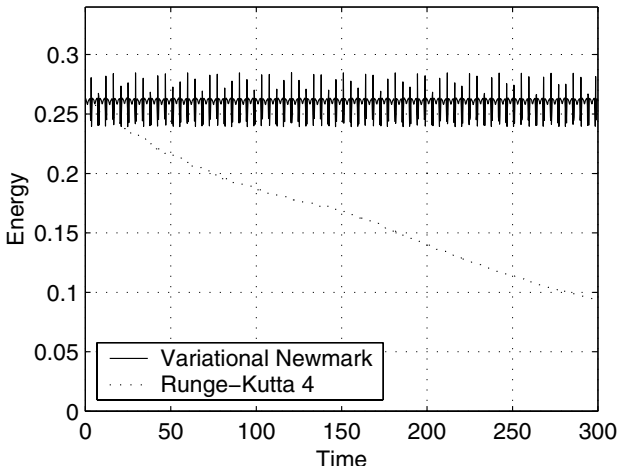


Fig. 1. Energy computed with variational Newmark (solid line) and Runge–Kutta (dashed line). Note that the variational method does not artificially dissipate energy

This is clearly a discretization of Newton's equations, using a simple finite difference rule for the derivative.

If we take initial conditions (q_0, q_1) then the discrete Euler–Lagrange equations define a recursive rule for calculating the sequence $\{q_k\}_{k=0}^{N}$. Regarded in this way, they define a map $F_{L_d} : (q_k, q_{k+1}) \mapsto (q_{k+1}, q_{k+2})$ which we can think of as a one-step integrator for the system defined by the continuous Euler–Lagrange equations.

Indeed, as we will see later, many standard one-step methods can be derived by such a procedure. An example of this is the well-known Newmark method, which for the parameter settings $\gamma = \frac{1}{2}$ and $\beta = 0$ is derived by choosing the discrete Lagrangian

$$L_d(q_0, q_1, h) = h\left[\left(\frac{q_1 - q_0}{h}\right)^T M\left(\frac{q_1 - q_0}{h}\right) - \left(\frac{V(q_0) + V(q_1)}{2}\right)\right].$$

If we use this variational Newmark method to simulate a model system and plot the energy versus time, then we obtain a graph like that in Figure 1. For comparison, this graph also shows the energy curve for a simulation with a standard stable method such as RK4 (the common fourth-order Runge–Kutta method).

The system being simulated here is purely conservative and so there should be no loss of energy over time. The striking aspect of this graph is that while the energy associated with a standard method decays due to numerical damping, for the Newmark method the energy error remains bounded. This

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

may be understood by recognizing that the integrator is symplectic, that is, it preserves the same two-form on state space as the true system.

In fact, *all variational integrators have this property*, as it is a consequence of the variational method of derivation, just as it is for continuous Lagrangian systems. In addition, they will also have the property of conserving momentum maps arising from symmetry actions, again due to the variational derivation. To understand this behaviour more deeply, however, we must first return to the beginning and consider in more detail the geometric structures underlying both continuous and discrete variational mechanics.

Of course, such sweeping statements as above have to be interpreted and used with great care, as in the precise statements in the text that follows. For example, if the integration step size is too large, then sometimes energy can behave very badly, even for a symplectic integrator (see, for example, Gonzalez and Simo (1996)). It is likewise well known that energy conservation does not guarantee accuracy (Ortiz 1986).

1.2. Background: Lagrangian mechanics

1.2.1. Basic definitions

Consider a *configuration manifold* Q, with associated *state space* given by the tangent bundle TQ, and a *Lagrangian* $L : TQ \to \mathbb{R}$.

Given an interval $[0, T]$, define the *path space* to be

$$\mathcal{C}(Q) = \mathcal{C}([0,T], Q) = \{q \colon [0,T] \to Q \mid q \text{ is a } C^2 \text{ curve}\}$$

and the *action map* $\mathfrak{G} : \mathcal{C}(Q) \to \mathbb{R}$ to be

$$\mathfrak{G}(q) \equiv \int_0^T L(q(t), \dot{q}(t))\, \mathrm{d}t. \tag{1.2.1}$$

It can be proved that $\mathcal{C}(Q)$ is a smooth manifold (Abraham, Marsden and Ratiu 1988), and $\mathfrak{G}$ is as smooth as L.

The tangent space $T_q\mathcal{C}(Q)$ to $\mathcal{C}(Q)$ at the point q is the set of C^2 maps $v_q : [0, T] \to TQ$ such that $\pi_Q \circ v_q = q$, where $\pi_Q : TQ \to Q$ is the canonical projection.

Define the *second-order submanifold* of $T(TQ)$ to be

$$\ddot{Q} \equiv \{w \in T(TQ) \mid T\pi_Q(w) = \pi_{TQ}(w)\} \subset T(TQ)$$

where $\pi_{TQ} : T(TQ) \to TQ$ and $\pi_Q : TQ \to Q$ are the canonical projections. $\ddot{Q}$ is simply the set of second derivatives $\mathrm{d}^2q/\,\mathrm{d}t^2(0)$ of curves $q : \mathbb{R} \to Q$, which are elements of the form $((q, \dot{q}), (\dot{q}, \ddot{q})) \in T(TQ)$.

Theorem 1.2.1. Given a C^k Lagrangian L, $k \geq 2$, there exists a unique C^{k-2} mapping $D_{\mathrm{EL}}L : \ddot{Q} \to T^*Q$ and a unique C^{k-1} one-form Θ_L on TQ,

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

such that, for all variations $\delta q \in T_q\mathcal{C}(Q)$ of $q(t)$, we have

$$\mathbf{d}\mathfrak{G}(q)\cdot\delta q = \int_0^T D_{\mathrm{EL}}L(\ddot{q})\cdot\delta q\,\mathrm{d}t + \Theta_L(\dot{q})\cdot\hat{\delta q}\Big|_0^T, \tag{1.2.2}$$

where

$$\hat{\delta q}(t) = \left(\left(q(t), \frac{\partial q}{\partial t}(t)\right), \left(\delta q(t), \frac{\partial \delta q}{\partial t}(t)\right)\right).$$

The mapping $D_{\mathrm{EL}}L$ is called the *Euler–Lagrange map* and has the coordinate expression

$$(D_{\mathrm{EL}}L)_i = \frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i}.$$

The one-form Θ_L is called the *Lagrangian one-form* and in coordinates is given by

$$\Theta_L = \frac{\partial L}{\partial \dot{q}^i}\mathbf{d}q^i. \tag{1.2.3}$$

Proof. Computing the variation of the action map gives

$$\begin{aligned}\mathbf{d}\mathfrak{G}(q)\cdot\delta q &= \int_0^T \left[\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\frac{\mathrm{d}}{\mathrm{d}t}\delta q^i\right]\mathrm{d}t \\ &= \int_0^T \left[\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i}\right]\cdot\delta q^i\,\mathrm{d}t + \left[\frac{\partial L}{\partial \dot{q}^i}\delta q^i\right]_0^T\end{aligned}$$

using integration by parts, and the terms of this expression can be identified as the Euler–Lagrange map and the Lagrangian one-form. □

1.2.2. Lagrangian vector fields and flows

The *Lagrangian vector field* $X_L : TQ \to T(TQ)$ is a second-order vector field on TQ satisfying

$$D_{\mathrm{EL}}L \circ X_L = 0 \tag{1.2.4}$$

and the *Lagrangian flow* $F_L : TQ \times \mathbb{R} \to TQ$ is the flow of X_L (we shall ignore issues related to global versus local flows, which are easily dealt with by restricting the domains of the flows). We shall write $F_L^t : TQ \to TQ$ for the map F_L at the frozen time t.

For an arbitrary Lagrangian, equation (1.2.4) may not uniquely define the vector field X_L and hence the flow map F_L may not exist. For now we will assume that L is such that these objects exist and are unique, and in Section 1.4.3 we will see under what conditions this is true.

A curve $q \in \mathcal{C}(Q)$ is said to be a *solution of the Euler–Lagrange equations* if the first term on the right-hand side of (1.2.2) vanishes for all variations $\delta q \in T_q\mathcal{C}(Q)$. This is equivalent to $(q, \dot{q})$ being an integral curve of X_L, and

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

means that q must satisfy the *Euler–Lagrange equations*

$$\frac{\partial L}{\partial q^i}(q,\dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}(q,\dot{q})\right) = 0 \tag{1.2.5}$$

for all $t \in (0,T)$.

1.2.3. Lagrangian flows are symplectic

Define the *solution space* $\mathcal{C}_L(Q) \subset \mathcal{C}(Q)$ to be the set of solutions of the Euler–Lagrange equations. As an element $q \in \mathcal{C}_L(Q)$ is an integral curve of X_L, it is uniquely determined by the initial condition $(q(0),\dot{q}(0)) \in TQ$ and we can thus identify $\mathcal{C}_L(Q)$ with the space of initial conditions TQ.

Defining the *restricted action map* $\hat{\mathfrak{G}} : TQ \to \mathbb{R}$ to be

$$\hat{\mathfrak{G}}(v_q) = \mathfrak{G}(q), \quad q \in \mathcal{C}_L(Q) \text{ and } (q(0),\dot{q}(0)) = v_q,$$

we see that (1.2.2) reduces to

$$\begin{aligned}\mathbf{d}\hat{\mathfrak{G}}(v_q)\cdot w_{v_q} &= \Theta_L(\dot{q}(T))((F_L^T)_*(w_{v_q})) - \Theta_L(v_q)(w_{v_q}) \\ &= ((F_L^T)^*(\Theta_L))(v_q)(w_{v_q}) - \Theta_L(v_q)(w_{v_q})\end{aligned} \tag{1.2.6}$$

for all $w_{v_q} \in T_{v_q}(TQ)$. Taking a further derivative of this expression, and using the fact that $\mathbf{d}^2\hat{\mathfrak{G}} = 0$, we obtain

$$(F_L^T)^*(\Omega_L) = \Omega_L,$$

where $\Omega_L = \mathbf{d}\Theta_L$ is the *Lagrangian symplectic form*, given in coordinates by

$$\Omega_L(q,\dot{q}) = \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j}\mathbf{d}q^i \wedge \mathbf{d}q^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\mathbf{d}\dot{q}^i \wedge \mathbf{d}q^j.$$

1.2.4. Lagrangian flows preserve momentum maps

Suppose that a Lie group G, with Lie algebra $\mathfrak{g}$, acts on Q by the (left or right) action $\Phi : G \times Q \to Q$. Consider the tangent lift of this action to $\Phi^{TQ} : G \times TQ \to TQ$ given by $\Phi_g^{TQ}(v_q) = T(\Phi_g)\cdot v_q$, which is

$$\Phi^{TQ}\big(g,(q,\dot{q})\big) = \left(\Phi^i(g,q), \frac{\partial \Phi^i}{\partial q^j}(g,q)\,\dot{q}^j\right).$$

For $\xi \in \mathfrak{g}$ define the *infinitesimal generators* $\xi_Q : Q \to TQ$ and $\xi_{TQ} : TQ \to T(TQ)$ by

$$\begin{aligned}\xi_Q(q) &= \frac{\mathrm{d}}{\mathrm{d}g}\Big(\Phi_g(q)\Big)\cdot\xi, \\ \xi_{TQ}(v_q) &= \frac{\mathrm{d}}{\mathrm{d}g}\Big(\Phi_g^{TQ}(v_q)\Big)\cdot\xi.\end{aligned}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

In coordinates these are given by

$$\xi_Q(q) = \left(q^i, \frac{\partial \Phi^i}{\partial g^m}(e,q)\xi^m \right),$$
$$\xi_{TQ}(q,\dot{q}) = \left(q^i, \dot{q}^i, \frac{\partial \Phi^i}{\partial g^m}(e,q)\xi^m, \frac{\partial^2 \Phi^i}{\partial g^m \partial q^j}(e,q)\dot{q}^j \xi^m \right).$$

We now define the *Lagrangian momentum map* $J_L : TQ \to \mathfrak{g}^*$ to be

$$J_L(v_q) \cdot \xi = \Theta_L \cdot \xi_{TQ}(v_q). \tag{1.2.7}$$

It can be checked that an equivalent expression for J_L is

$$J_L(v_q) \cdot \xi = \left\langle \frac{\partial L}{\partial \dot{q}}, \xi_Q(q) \right\rangle,$$

where $\partial L/\partial \dot{q}$ represents the Legendre transformation, discussed shortly. This equation is convenient for computing momentum maps in examples: see Marsden and Ratiu (1999).

The traditional linear and angular momenta are momentum maps, with the linear momentum $J_L : T\mathbb{R}^n \to \mathbb{R}^n$ arising from the additive action of $\mathbb{R}^n$ on itself, and the angular momentum $J_L : T\mathbb{R}^n \to \mathfrak{so}(n)^*$ coming from the action of $SO(n)$ on $\mathbb{R}^n$.

An important property of momentum maps is *equivariance*, which is the condition that the following diagram commutes.

$$\begin{array}{ccc} TQ & \xrightarrow{J_L} & \mathfrak{g}^* \\ {\scriptstyle \Phi_g^{TQ}}\downarrow & & \downarrow{\scriptstyle \mathrm{Ad}^*_{g^{-1}}} \\ TQ & \xrightarrow[J_L]{} & \mathfrak{g}^* \end{array} \tag{1.2.8}$$

In general, Lagrangian momentum maps are not equivariant, but we give here a simple sufficient condition for this property to be satisfied. Recall that a map $f : TQ \to TQ$ is said to be *symplectic* if $f^*\Omega_L = \Omega_L$. If, furthermore, f is such that $f^*\Theta_L = \Theta_L$, then f is said to be a *special symplectic map*. Clearly a special symplectic map is also symplectic, but the converse does not hold.

Theorem 1.2.2. Consider a Lagrangian system $L : TQ \to \mathbb{R}$ with a left action $\Phi : G \times Q \to Q$. If the lifted action $\Phi^{TQ} : G \times TQ \to TQ$ acts by special canonical transformations, then the Lagrangian momentum map $J_L : TQ \to \mathfrak{g}^*$ is equivariant.

Proof. Observing that $(\Phi_g^{TQ})^{-1} = \Phi_{g^{-1}}^{TQ}$, we see that equivariance is equivalent to

$$J_L(v_q) \cdot \xi = J_L \circ T\Phi_{g^{-1}}(v_q) \cdot \mathrm{Ad}_{g^{-1}}\xi.$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

We now compute the right-hand side of this expression to give

$$\begin{aligned} J_L \circ \Phi^{TQ}_{g^{-1}}(v_q) \cdot \mathrm{Ad}_{g^{-1}}\xi &= \big\langle \Theta_L\big(\Phi^{TQ}_{g^{-1}}(v_q)\big), (\mathrm{Ad}_{g^{-1}}\xi)_{TQ}\big(\Phi^{TQ}_{g^{-1}}(v_q)\big)\big\rangle \\ &= \big\langle \Theta_L\big(\Phi^{TQ}_{g^{-1}}(v_q)\big), T(\Phi^{TQ}_{g^{-1}}) \cdot \xi_{TQ}(v_q)\big\rangle \\ &= \big\langle \big((\Phi^{TQ}_{g^{-1}})^*\Theta_L\big)(v_q), \xi_{TQ}(v_q)\big\rangle \\ &= \big\langle \Theta_L(v_q), \xi_{TQ}(v_q)\big\rangle, \end{aligned}$$

which is just $J_L(v_q) \cdot \xi$, as desired. Here we used the identity $(\mathrm{Ad}_g\xi)_M = \Phi^*_{g^{-1}}\xi_M$ (Marsden and Ratiu 1999) to pass from the first to the second line. □

A Lagrangian $L : TQ \to \mathbb{R}$ is said to be *invariant* under the lift of the action $\Phi : G \times Q \to Q$ if $L \circ \Phi^{TQ}_g = L$ for all $g \in G$, and in this case the group action is said to be a *symmetry* of the Lagrangian. Differentiating this expression implies that the Lagrangian is *infinitesimally invariant*, which is the statement $\mathbf{d}L \cdot \xi_{TQ} = 0$ for all $\xi \in \mathfrak{g}$.

Observe that if L is invariant then this implies that Φ^{TQ} acts by special symplectic transformations, and so the Lagrangian momentum map is equivariant. To see this, we write $L \circ \Phi^{TQ}_g = L$ in coordinates to obtain $L(\Phi_g(q), \partial_q\Phi_g(q) \cdot \dot{q}) = L(q, \dot{q})$, and now differentiating this with respect to $\dot{q}$ in the direction δq gives

$$\frac{\partial L}{\partial \dot{q}}\big(\Phi_g(q), \partial_q\Phi_g(q) \cdot \dot{q}\big) \cdot \partial_q\Phi_g(q) \cdot \delta q = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \cdot \delta q.$$

But the left- and right-hand sides are simply $(\Phi^{TQ}_g)^*\Theta_L$ and Θ_L, respectively, evaluated on $((q, \dot{q}), (\delta q, \delta\dot{q}))$, and thus we have $(\Phi^{TQ}_g)^*\Theta_L = \Theta_L$.

We will now show that, when the group action is a symmetry of the Lagrangian, then the momentum maps are preserved by the Lagrangian flow. This result was originally due to Noether (1918), using a technique similar to the one given below.

Theorem 1.2.3. (Noether's theorem) Consider a Lagrangian system $L : TQ \to \mathbb{R}$ which is invariant under the lift of the (left or right) action $\Phi : G \times Q \to Q$. Then the corresponding Lagrangian momentum map $J_L : TQ \to \mathfrak{g}^*$ is a conserved quantity of the flow, so that $J_L \circ F^t_L = J_L$ for all times t.

Proof. The action of G on Q induces an action of G on the space of paths $\mathcal{C}(Q)$ by pointwise action, so that $\Phi_g : \mathcal{C}(Q) \to \mathcal{C}(Q)$ is given by $\Phi_g(q)(t) = \Phi_g(q(t))$. As the action is just the integral of the Lagrangian, invariance of L implies invariance of $\mathfrak{G}$ and the differential of this gives

$$\mathbf{d}\mathfrak{G}(q) \cdot \xi_{\mathcal{C}(Q)}(q) = \int_0^T \mathbf{d}L \cdot \xi_{TQ}\,\mathrm{d}t = 0.$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Invariance of $\mathfrak{G}$ also implies that Φ_g maps solution curves to solution curves and thus $\xi_{\mathcal{C}(Q)}(q) \in T_q\mathcal{C}_L$, which is the corresponding infinitesimal version. We can thus restrict $\mathbf{d}\mathfrak{G} \cdot \xi_{\mathcal{C}(Q)}$ to the space of solutions $\mathcal{C}_L$ to obtain

$$0 = \hat{\mathfrak{G}}(v_q) \cdot \xi_{TQ}(v_q) = \Theta_L(\dot{q}(T)) \cdot \xi_{TQ}(\dot{q}(T)) - \Theta_L(v_q) \cdot \xi_{TQ}(v_q).$$

Substituting in the definition of the Lagrangian momentum map J_L, however, shows that this is just $0 = J_L(F_L^T(v_q)) \cdot \xi - J_L(v_q) \cdot \xi$, which gives the desired result. □

We have thus seen that conservation of momentum maps is a direct consequence of the invariance of the variational principle under a symmetry action. The fact that the symmetry maps solution curves to solution curves will extend directly to discrete mechanics.

In fact, only infinitesimal invariance is needed for the momentum map to be conserved by the Lagrangian flow, as a careful reading of the above proof will show. This is because it is only necessary that the Lagrangian be invariant in a neighbourhood of a given trajectory, and so the global statement of invariance is stronger than necessary.

1.3. Discrete variational mechanics: Lagrangian viewpoint

Take again a configuration manifold Q, but now define the *discrete state space* to be $Q \times Q$. This contains the same amount of information as (is locally isomorphic to) TQ. A *discrete Lagrangian* is a function $L_d : Q \times Q \to \mathbb{R}$.

To relate discrete and continuous mechanics it is necessary to introduce a *time-step* $h \in \mathbb{R}$, and to take L_d to depend on this time-step. For the moment, we will take $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$, and will neglect the h dependence except where it is important. We shall come back to this point later when we discuss the context of time-dependent mechanics and adaptive algorithms. However, the idea behind this was explained in the introduction.

Construct the increasing sequence of times $\{t_k = kh \mid k = 0, \ldots, N\} \subset \mathbb{R}$ from the time-step h, and define the *discrete path space* to be

$$\mathcal{C}_d(Q) = \mathcal{C}_d(\{t_k\}_{k=0}^N, Q) = \{q_d \colon \{t_k\}_{k=0}^N \to Q\}.$$

We will identify a discrete trajectory $q_d \in \mathcal{C}_d(Q)$ with its image $q_d = \{q_k\}_{k=0}^N$, where $q_k = q_d(t_k)$. The *discrete action map* $\mathfrak{G}_d : \mathcal{C}_d(Q) \to \mathbb{R}$ is defined by

$$\mathfrak{G}_d(q_d) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}).$$

As the discrete path space $\mathcal{C}_d$ is isomorphic to $Q \times \cdots \times Q$ ($N+1$ copies), it can be given a smooth product manifold structure. The discrete action $\mathfrak{G}_d$ clearly inherits the smoothness of the discrete Lagrangian L_d.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

The tangent space $T_{q_d}\mathcal{C}_d(Q)$ to $\mathcal{C}_d(Q)$ at q_d is the set of maps $v_{q_d} : \{t_k\}_{k=0}^N \to TQ$ such that $\pi_Q \circ v_{q_d} = q_d$, which we will denote by $v_{q_d} = \{(q_k, v_k)\}_{k=0}^N$.

The discrete object corresponding to $T(TQ)$ is the set $(Q \times Q) \times (Q \times Q)$. We define the *projection operator* π and the *translation operator* σ to be

$$\pi : ((q_0, q_1), (q_0', q_1')) \mapsto (q_0, q_1),$$
$$\sigma : ((q_0, q_1), (q_0', q_1')) \mapsto (q_0', q_1').$$

The *discrete second-order submanifold* of $(Q \times Q) \times (Q \times Q)$ is defined to be

$$\ddot{Q}_d \equiv \{w_d \in (Q \times Q) \times (Q \times Q) \mid \pi_1 \circ \sigma(w_d) = \pi_2 \circ \pi(w_d)\},$$

which has the same information content as (is locally isomorphic to) $\ddot{Q}$. Concretely, the discrete second-order submanifold is the set of pairs of the form $((q_0, q_1), (q_1, q_2))$.

Theorem 1.3.1. Given a C^k discrete Lagrangian L_d, $k \geq 1$, there exists a unique C^{k-1} mapping $D_{\mathrm{DEL}}L_d : \ddot{Q}_d \to T^*Q$ and unique C^{k-1} one-forms $\Theta_{L_d}^+$ and $\Theta_{L_d}^-$ on $Q \times Q$, such that, for all variations $\delta q_d \in T_{q_d}\mathcal{C}(Q)$ of q_d, we have

$$\mathbf{d}\mathfrak{G}_d(q_d) \cdot \delta q_d = \sum_{k=1}^{N-1} D_{\mathrm{DEL}}L_d((q_{k-1}, q_k), (q_k, q_{k+1})) \cdot \delta q_k + \Theta_{L_d}^+(q_{N-1}, q_N) \cdot (\delta q_{N-1}, \delta q_N) - \Theta_{L_d}^-(q_0, q_1) \cdot (\delta q_0, \delta q_1). \quad (1.3.1)$$

The mapping $D_{\mathrm{DEL}}L_d$ is called the *discrete Euler–Lagrange map* and has coordinate expression

$$D_{\mathrm{DEL}}L_d((q_{k-1}, q_k), (q_k, q_{k+1})) = D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1}).$$

The one-forms $\Theta_{L_d}^+$ and $\Theta_{L_d}^-$ are called the *discrete Lagrangian one-forms* and in coordinates are

$$\Theta_{L_d}^+(q_0, q_1) = D_2L_d(q_0, q_1)\mathbf{d}q_1 = \frac{\partial L_d}{\partial q_1^i} dq_1^i, \quad (1.3.2a)$$

$$\Theta_{L_d}^-(q_0, q_1) = -D_1L_d(q_0, q_1)\mathbf{d}q_0 = -\frac{\partial L_d}{\partial q_0^i} dq_0^i. \quad (1.3.2b)$$

Proof. Computing the derivative of the discrete action map gives

$$\begin{aligned}\mathbf{d}\mathfrak{G}_d(q_d) \cdot \delta q_d &= \sum_{k=0}^{N-1} [D_1L_d(q_k, q_{k+1}) \cdot \delta q_k + D_2L_d(q_k, q_{k+1}) \cdot \delta q_{k+1}] \\ &= \sum_{k=1}^{N-1} [D_1L_d(q_k, q_{k+1}) + D_2L_d(q_{k-1}, q_k)] \cdot \delta q_k \\ &\quad + D_1L_d(q_0, q_1) \cdot \delta q_0 + D_2L_d(q_{N-1}, q_N) \cdot \delta q_N\end{aligned}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

using a discrete integration by parts (rearrangement of the summation). Identifying the terms with the discrete Euler–Lagrange map and the discrete Lagrangian one-forms now gives the desired result. □

Unlike the continuous case, in the discrete case there are two one-forms that arise from the boundary terms. Observe, however, that $\mathbf{d}L_d = \Theta^+_{L_d} - \Theta^-_{L_d}$ and so using $\mathbf{d}^2 = 0$ shows that

$$\mathbf{d}\Theta^+_{L_d} = \mathbf{d}\Theta^-_{L_d}.$$

This will be reflected below in the fact that there is only a single discrete two-form, which is the same as the continuous situation and is important for symplecticity.

1.3.1. Discrete Lagrangian evolution operator and mappings

A *discrete evolution operator* X plays the same role as a continuous vector field, and is defined to be any map $X : Q \times Q \to (Q \times Q) \times (Q \times Q)$ satisfying $\pi \circ X = id$. The discrete object corresponding to the flow is the *discrete map* $F : Q \times Q \to Q \times Q$ defined by $F = \sigma \circ X$. In coordinates, if the discrete evolution operator maps $X : (q_0, q_1) \mapsto (q_0, q_1, q_0', q_1')$, then the discrete map will be $F : (q_0, q_1) \mapsto (q_0', q_1')$.

We will be mainly interested in discrete evolution operators which are *second-order*, which is the requirement that $X(Q \times Q) \subset \ddot{Q}_d$. This implies that they have the form $X : (q_0, q_1) \mapsto (q_0, q_1, q_1, q_2)$, and so the corresponding discrete map will be $F : (q_0, q_1) \mapsto (q_1, q_2)$. We now consider the particular case of a discrete Lagrangian system.

The *discrete Lagrangian evolution operator* X_{L_d} is a second-order discrete evolution operator satisfying

$$D_{\mathrm{DEL}} L_d \circ X_{L_d} = 0$$

and the *discrete Lagrangian map* $F_{L_d} : Q \times Q \to Q \times Q$ is defined by $F_{L_d} = \sigma \circ X_{L_d}$.

As in the continuous case, the discrete Lagrangian evolution operator and discrete Lagrangian map are not well-defined for arbitrary choices of discrete Lagrangian. We will henceforth assume that L_d is chosen so as to make these structures well-defined, and in Section 1.5 we will give a condition on L_d which ensures that this is true.

A discrete path $q_d \in \mathcal{C}_d(Q)$ is said to be a *solution of the discrete Euler–Lagrange equations* if the first term on the right-hand side of (1.3.1) vanishes for all variations $\delta q_d \in T_{q_d}\mathcal{C}_d(Q)$. This means that the points $\{q_k\}$ satisfy $F_{L_d}(q_{k-1}, q_k) = (q_k, q_{k+1})$ or, equivalently, that they satisfy the *discrete Euler–Lagrange equations*

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \quad \text{for all } k = 1, \ldots, N-1. \quad (1.3.3)$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

1.3.2. Discrete Lagrangian maps are symplectic

Define the *discrete solution space* $\mathcal{C}_{L_d}(Q) \subset \mathcal{C}_d(Q)$ to be the set of solutions of the discrete Euler–Lagrange equations. Since an element $q_d \in \mathcal{C}_{L_d}(Q)$ is formed by iteration of the map F_{L_d}, it is uniquely determined by the initial condition $(q_0, q_1) \in Q \times Q$. We can thus identify $\mathcal{C}_{L_d}(Q)$ with the space of initial conditions $Q \times Q$.

Defining the *restricted discrete action map* $\hat{\mathfrak{G}}_d : Q \times Q \to \mathbb{R}$ to be

$$\hat{\mathfrak{G}}_d(q_0, q_1) = \mathfrak{G}_d(q_d); \quad q_d \in \mathcal{C}_{L_d}(Q) \text{ and } (q_d(t_0), q_d(t_1)) = (q_0, q_1),$$

we see that (1.3.1) reduces to

$$\begin{aligned} \mathbf{d}\hat{\mathfrak{G}}_d(v_d) \cdot w_{v_d} &= \Theta^+_{L_d}(F^{N-1}_{L_d}(v_d))((F^{N-1}_{L_d})_*(w_{v_d})) - \Theta^-_{L_d}(v_d)(w_{v_d}) \\ &= ((F^{N-1}_{L_d})^*(\Theta^+_{L_d}))(v_d)(w_{v_d}) - \Theta^-_{L_d}(v_d)(w_{v_d}) \end{aligned} \tag{1.3.4}$$

for all $w_{v_d} \in T_{v_d}(Q \times Q)$ and $v_d = (q_0, q_1) \in Q \times Q$. Taking a further derivative of this expression, and using the fact that $\mathbf{d}^2\hat{\mathfrak{G}}_d = 0$, we obtain

$$(F^{N-1}_{L_d})^*(\Omega_{L_d}) = \Omega_{L_d}$$

where $\Omega_{L_d} = \mathbf{d}\Theta^+_{L_d} = \mathbf{d}\Theta^-_{L_d}$ is the *discrete Lagrangian symplectic form*, with coordinate expression

$$\Omega_{L_d}(q_0, q_1) = \frac{\partial^2 L_d}{\partial q_0^i \partial q_1^j} \mathbf{d}q_0^i \wedge \mathbf{d}q_1^j.$$

This argument also holds if we take any subinterval of $0, \ldots, N$ and so the statement is true for any number of steps of F_{L_d}. For a single step we have $(F_{L_d})^*\Omega_{L_d} = \Omega_{L_d}$.

Given a map $f : Q \times Q \to Q \times Q$, we will say that f is *discretely symplectic* if $f^*\Omega_{L_d} = \Omega_{L_d}$. The above calculations thus prove that the discrete Lagrangian map F_{L_d} is discretely symplectic, just as we saw in the last section that the Lagrangian flow map is symplectic on TQ.

1.3.3. Discrete Noether's theorem

Consider the (left or right) action $\Phi : G \times Q \to Q$ of a Lie group G on Q, with infinitesimal generator as defined in Section 1.2. This action can be lifted to $Q \times Q$ by the product $\Phi_g^{Q\times Q}(q_0, q_1) = (\Phi_g(q_0), \Phi_g(q_1))$, which has an *infinitesimal generator* $\xi_{Q\times Q} : Q \times Q \to T(Q \times Q)$ given by

$$\xi_{Q\times Q}(q_0, q_1) = (\xi_Q(q_0), \xi_Q(q_1)).$$

The two *discrete Lagrangian momentum maps* $J^+_{L_d}, J^-_{L_d} : Q \times Q \to \mathfrak{g}^*$ are

$$\begin{aligned} J^+_{L_d}(q_0, q_1) \cdot \xi &= \Theta^+_{L_d} \cdot \xi_{Q\times Q}(q_0, q_1), \\ J^-_{L_d}(q_0, q_1) \cdot \xi &= \Theta^-_{L_d} \cdot \xi_{Q\times Q}(q_0, q_1). \end{aligned}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Using the expressions for $\Theta^{\pm}_{L_d}$ allows the discrete momentum maps to be alternatively written as

$$J^{+}_{L_d}(q_0, q_1) \cdot \xi = \langle D_2 L_d(q_0, q_1), \xi_Q(q_1) \rangle ,$$
$$J^{-}_{L_d}(q_0, q_1) \cdot \xi = \langle -D_1 L_d(q_0, q_1), \xi_Q(q_0) \rangle ,$$

which are computationally useful formulations.

As in the continuous case, it is interesting to consider when the discrete momentum maps are equivariant. This is the conditions

$$J^{+}_{L_d} \circ \Phi^{Q\times Q}_g = \mathrm{Ad}^*_{g^{-1}} \circ J^{+}_{L_d},$$
$$J^{-}_{L_d} \circ \Phi^{Q\times Q}_g = \mathrm{Ad}^*_{g^{-1}} \circ J^{-}_{L_d}.$$

In general these equations will not be satisfied; however, there is a simple sufficient condition, similar to the condition in the continuous case.

Recall that we have defined a map $f : Q \times Q \to Q \times Q$ to be discretely symplectic if $f^*\Omega_{L_d} = \Omega_{L_d}$. We now define f to be a *special discrete symplectic map* if $f^*\Theta^{+}_{L_d} = \Theta^{+}_{L_d}$ and $f^*\Theta^{-}_{L_d} = \Theta^{-}_{L_d}$. This clearly means that f is also discretely symplectic, but the reverse is not true.

Theorem 1.3.2. Take a discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$ with a (left or right) group action $\Phi : G \times Q \to Q$. If the product lifted action $\Phi^{Q\times Q} : G \times Q \times Q \to Q \times Q$ acts by special discrete symplectic maps, then the discrete Lagrangian momentum maps are equivariant.

Proof. The proof used in Theorem 1.2.2 for the continuous case can also be used here, with $J^{+}_{L_d}$ and $J^{-}_{L_d}$ being considered separately. □

If the lifted action only preserves one of $\Theta^{+}_{L_d}$ or $\Theta^{-}_{L_d}$ then only the corresponding momentum map will necessarily be equivariant.[2]

If a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$ is such that $L_d \circ \Phi^{Q\times Q}_g = L_d$ for all $g \in G$, then L_d is said to be *invariant* under the lifted action, and Φ is said to be a *symmetry* of the discrete Lagrangian. Note that invariance implies *infinitesimal invariance*, which is $\mathbf{d}L_d \cdot \xi_{Q\times Q} = 0$ for all $\xi \in \mathfrak{g}$. Also note that

$$\mathbf{d}L_d = \Theta^{+}_{L_d} - \Theta^{-}_{L_d},$$

and so when L_d is infinitesimally invariant under the lifted action the two discrete momentum maps are equal. In such cases we will use the notation $J_{L_d} : Q \times Q \to \mathfrak{g}^*$ for the unique single *discrete Lagrangian momentum map*.

[2] As in the continuous case, equivariance plays an important role in reduction theory and, in the Hamiltonian context, equivariance guarantees that the momentum map is Poisson, which is often useful.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Note that invariance of L_d under the lifted action implies that $\Phi_g^{Q\times Q}$ is a special discrete symplectic map. This can be seen by differentiating $L_d \circ \Phi_g^{Q\times Q} = L_d$ with respect to q_1 to obtain

$$D_2 L_d\big(\Phi_g^{Q\times Q}(q_0,q_1)\big)\cdot \partial_q \Phi_g(q_1)\cdot \delta q_1 = D_2 L_d(q_0,q_1)\cdot \delta q_1,$$

and observing that the left- and right-hand sides are just $(\Phi_g^{Q\times Q})^*\Theta^+_{L_d}$ and $\Theta^+_{L_d}$, respectively, applied to $(q_0,q_1,\delta q_0,\delta q_1)$. Hence $(\Phi_g^{Q\times Q})^*\Theta^+_{L_d} = \Theta^+_{L_d}$, and a similar calculation gives the result for $\Theta^-_{L_d}$.

We now give the discrete analogue of Noether's theorem, Theorem 1.2.3, which states that momentum maps of symmetries are constants of the motion.

Theorem 1.3.3. (Discrete Noether's theorem) Consider a given discrete Lagrangian system $L_d : Q\times Q \to \mathbb{R}$ which is invariant under the lift of the (left or right) action $\Phi : G\times Q \to Q$. Then the corresponding discrete Lagrangian momentum map $J_{L_d} : Q\times Q \to \mathfrak{g}^*$ is a conserved quantity of the discrete Lagrangian map $F_{L_d} : Q\times Q \to Q\times Q$, so that $J_{L_d}\circ F_{L_d} = J_{L_d}$.

Proof. We will use the same idea as in the proof of the continuous Noether's theorem, based on the fact that the variational principle is invariant under the symmetry action.

Begin by inducing an action of G on the discrete path space $\mathcal{C}_d(Q)$ by using the pointwise action. Then

$$\mathbf{d}\mathfrak{G}_d(q_d)\cdot \xi_{\mathcal{C}_d(Q)}(q_d) = \sum_{k=0}^{N-1} \mathbf{d}L_d\cdot \xi_{Q\times Q},$$

and so the space of solutions $\mathcal{C}_{L_d}(Q)$ of the discrete Euler–Lagrange equations is invariant under the lifted action of G, and the discrete Lagrangian map $F_{L_d} : Q\times Q \to Q\times Q$ commutes with the lifted action $\Phi_g : Q\times Q \to Q\times Q$.

Identifying $\mathcal{C}_{L_d}(Q)$ with the space of initial conditions $Q\times Q$ and using equation (1.3.4) gives

$$\begin{aligned}\mathbf{d}\mathfrak{G}_d(q_d)\cdot \xi_{\mathcal{C}(Q)}(q_d) &= \mathbf{d}\hat{\mathfrak{G}}_d(q_0,q_1)\cdot \xi_{Q\times Q}(q_0,q_1)\\ &= \big((F^N_{L_d})^*\big(\Theta^+_{L_d}\big) - \Theta^-_{L_d}\big)(q_0,q_1)\cdot \xi_{Q\times Q}(q_0,q_1).\end{aligned}$$

For symmetries the left-hand side is zero, and so we have

$$(\Theta^+_{L_d}\cdot \xi_{Q\times Q})\circ F^N_{L_d} = \Theta^-_{L_d}\cdot \xi_{Q\times Q},$$

which is simply the statement of preservation of the discrete momentum map, given that for symmetry actions there is only a single momentum map and that the above argument holds for all subintervals, including a single time-step. □

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

As in the continuous case, only infinitesimal invariance of the discrete Lagrangian is actually required for the discrete momentum maps to be conserved. This is due to the fact that only local invariance is used in the proof above, and global invariance is not necessary.

Note that if G is not a symmetry of L_d then the two discrete momentum maps will not be equal, and it is precisely the difference $J^+_{L_d} - J^-_{L_d}$ which describes the evolution of either momentum map during the time-step. To see this, define

$$J^\Delta_{L_d}(q_k, q_{k+1}) = J^+_{L_d}(q_k, q_{k+1}) - J^-_{L_d}(q_k, q_{k+1})$$

and observe that the discrete Euler–Lagrange equations imply

$$J^+_{L_d}(q_{k-1}, q_k) = J^-_{L_d}(q_k, q_{k+1}).$$

Combining the two above expressions shows that the two discrete momentum maps evolve according to

$$\begin{aligned} J^+_{L_d}(q_k, q_{k+1}) &= J^+_{L_d}(q_{k-1}, q_k) + J^\Delta_{L_d}(q_k, q_{k+1}), \\ J^-_{L_d}(q_k, q_{k+1}) &= J^-_{L_d}(q_{k-1}, q_k) + J^\Delta_{L_d}(q_{k-1}, q_k). \end{aligned}$$

Clearly, if L_d is invariant then $J^\Delta_{L_d} = 0$, and so the momentum maps are equal and they are conserved. If not, then these equations describe how the momentum maps evolve.

1.4. Background: Hamiltonian mechanics

1.4.1. Hamiltonian mechanics

We will only concern ourselves here with the case of a phase space that is the cotangent bundle of a configuration manifold. Although some of the elegance and power of the Hamiltonian formalism is lost in this restriction, it is simpler for our purposes, and of course is the most important case for applications.

Consider then a configuration manifold Q, and define the *phase space* to be the cotangent bundle T^*Q. The *Hamiltonian* is a function $H : T^*Q \to \mathbb{R}$. We will take local coordinates on T^*Q to be (q, p).

Define the *canonical one-form* Θ on T^*Q by

$$\Theta(p_q) \cdot u_{p_q} = \left\langle p_q, T\pi_{T^*Q} \cdot u_{p_q} \right\rangle, \tag{1.4.1}$$

where $\pi_{T^*Q} : T^*Q \to Q$ is the standard projection and $\langle \cdot, \cdot \rangle$ denotes the natural pairing between vectors and covectors. In coordinates, $\Theta(q, p) = p_i \mathbf{d}q^i$. The *canonical two-form* Ω on T^*Q is defined to be

$$\Omega = -\mathbf{d}\Theta,$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

which has coordinate expression $\Omega(q,p) = \mathbf{d}q^i \wedge \mathbf{d}p^i$. The pair (T^*Q, Ω) is an example of a *symplectic manifold* and a mapping $F : T^*Q \to T^*Q$ is said to be *canonical* or *symplectic* if $F^*\Omega = \Omega$. If $F^*\Theta = \Theta$ then F is said to be a *special symplectic map*, which clearly implies that it is also symplectic. Note that a particular case of special symplectic maps is given by cotangent lifts of maps $Q \to Q$, which automatically preserve the canonical one-form on T^*Q (see Marsden and Ratiu (1999) for details).

Given a Hamiltonian H, define the corresponding *Hamiltonian vector field* X_H to be the unique vector field on T^*Q satisfying

$$\mathbf{i}_{X_H}\Omega = \mathbf{d}H. \tag{1.4.2}$$

Writing $X_H = (X_q, X_p)$ in coordinates, we see that the above expression is

$$-X_{p_i}\mathbf{d}q^i + X_{q^i}\mathbf{d}p_i = \frac{\partial H}{\partial q^i}\mathbf{d}q^i + \frac{\partial H}{\partial p_i}\mathbf{d}p_i,$$

which gives the familiar *Hamilton's equations* for the components of X_H, namely

$$X_{q^i}(q,p) = \frac{\partial H}{\partial p_i}(q,p), \tag{1.4.3a}$$

$$X_{p_i}(q,p) = -\frac{\partial H}{\partial q^i}(q,p). \tag{1.4.3b}$$

The *Hamiltonian flow* $F_H : T^*Q \times \mathbb{R} \to T^*Q$ is the flow of the Hamiltonian vector field X_H. Note that, unlike the Lagrangian situation, the Hamiltonian vector field X_H and flow map F_H are always well-defined for any Hamiltonian.

For any fixed $t \in \mathbb{R}$, the flow map $F_H^t : T^*Q \to T^*Q$ is symplectic, as can be seen by differentiating to obtain

$$\begin{aligned}\left.\frac{\partial}{\partial t}\right|_{t=0} (F_H^t)^*\Omega = \mathcal{L}_{X_H}\Omega &= \mathbf{d}\mathbf{i}_{X_H}\Omega + \mathbf{i}_{X_H}\mathbf{d}\Omega \\ &= \mathbf{d}^2 H - \mathbf{i}_{X_H}\mathbf{d}^2\Theta = 0,\end{aligned}$$

where we have used Cartan's magic formula $\mathcal{L}_X\alpha = \mathbf{d}\mathbf{i}_X\alpha + \mathbf{i}_X\mathbf{d}\alpha$ for the Lie derivative and the fact that $\mathbf{d}^2 = 0$.

1.4.2. Hamiltonian form of Noether's theorem

Consider a (left or right) action $\Phi : G \times Q \to Q$ of G on Q, as in Section 1.2. The cotangent lift of this action is $\Phi^{T^*Q} : G \times T^*Q \to T^*Q$ given by $\Phi_g^{T^*Q}(p_q) = \Phi_{g^{-1}}^*(p_q)$, which in coordinates is

$$\Phi^{T^*Q}\big(g,(q,p)\big) = \left((\Phi_g^{-1})^i(q), p_j\frac{\partial \Phi_g^j}{\partial q^i}(q)\right).$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

This has the corresponding *infinitesimal generator* $\xi_{T^*Q} : T^*Q \to T(T^*Q)$ defined by

$$\xi_{T^*Q}(p_q) = \frac{\mathrm{d}}{\mathrm{d}g}\big(\Phi_g^{T^*Q}(p_q)\big) \cdot \xi$$

which has coordinate form

$$\xi_{T^*Q}(q,p) = \left(q^i, p_i, -\left[\left(\frac{\partial \Phi}{\partial q}\right)^{-1}\right]_j^i \frac{\partial \Phi^j}{\partial g^m}\xi^m, \right.$$
$$\left. p_j \frac{\partial^2 \Phi^j}{\partial q^i \partial g^m}\xi^m - p_j \frac{\partial^2 \Phi_j}{\partial q^i \partial q^j}\left[\left(\frac{\partial \Phi}{\partial q}\right)^{-1}\right]_k^j \frac{\partial \Phi^k}{\partial g^m}\xi^m\right),$$

where the derivatives of Φ are all evaluated at (e, q).

The *Hamiltonian momentum map* $J_H : T^*Q \to \mathfrak{g}^*$ is defined by

$$J_H(p_q) \cdot \xi = \Theta(p_q) \cdot \xi_{T^*Q}(p_q).$$

For each $\xi \in \mathfrak{g}$ we define $J_H^\xi : T^*Q \to \mathbb{R}$ by $J_H^\xi(p_q) = J_H(p_q) \cdot \xi$, which has the expression $J_H^\xi = \mathbf{i}_{\xi_{T^*Q}}\Theta$. Note that the Hamiltonian map is also given by the expression

$$J_H(p_q) \cdot \xi = \langle p_q, \xi_Q(q) \rangle,$$

which is useful for computing it in applications.

Writing the requirement for equivariance of a Hamiltonian momentum map gives the equation

$$J_H \circ \Phi_g^{T^*Q} = \mathrm{Ad}^*_{g^{-1}} \circ J_H.$$

Unlike the Lagrangian setting, however, cotangent lifted actions are *always* special symplectic maps, and so we have $(\Phi_g^{T^*Q})^*\Theta = \Theta$ irrespective of the Hamiltonian. This gives the following result.

Theorem 1.4.1. Consider a Hamiltonian system $H : T^*Q \to \mathbb{R}$ with a (left or right) group action $\Phi : G \times Q \to Q$. Then the Hamiltonian momentum map $J_H : T^*Q \to \mathfrak{g}^*$ is always equivariant with respect to the cotangent lifted action $\Phi^{T^*Q} : G \times T^*Q \to T^*Q$.

Proof. Once again, we can use exactly the same proof as for Theorem 1.2.2 in the continuous case. The only difference is that H need not be restricted to ensure that the lifted action is a special symplectic map. □

A Hamiltonian $H : T^*Q \to \mathbb{R}$ is said to be *invariant* under the cotangent lift of the action $\Phi : G \times Q \to Q$ if $H \circ \Phi_g^{T^*Q} = H$ for all $g \in G$, in which case the action is said to be a *symmetry* for the Hamiltonian. The derivative of this expression implies that such a Hamiltonian is also *infinitesimally invariant*, which is the requirement $\mathbf{d}H \cdot \xi_{T^*Q} = 0$ for all $\xi \in \mathfrak{g}$, although the converse is not generally true.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Theorem 1.4.2. (Hamiltonian Noether's theorem) Let $H : T^*Q \to \mathbb{R}$ be a Hamiltonian which is invariant under the lift of the (left or right) action $\Phi : G \times Q \to Q$. Then the corresponding Hamiltonian momentum map $J_H : T^*Q \to \mathfrak{g}^*$ is a conserved quantity of the flow; that is, $J_H \circ F_H^t = J_H$ for all times t.

Proof. Recall that $(\Phi_g^{T^*Q})^*\Theta = \Theta$ for all $g \in G$ as the action is a cotangent lift, and hence $\mathcal{L}_{\xi_{T^*Q}}\Theta = 0$. Now computing the derivative of J_H^ξ in the direction given by the Hamiltonian vector field X_H gives

$$\begin{aligned}\mathbf{d}J_H^\xi \cdot X_H &= \mathbf{d}(\mathbf{i}_{\xi_{T^*Q}}\Theta) \cdot X_H \\ &= \mathcal{L}_{\xi_{T^*Q}}\Theta \cdot X_H - \mathbf{i}_{\xi_{T^*Q}}\mathbf{d}\Theta \cdot X_H \\ &= -\mathbf{i}_{X_H}\Omega \cdot \xi_{T^*Q} \\ &= -\mathbf{d}H \cdot \xi_{T^*Q} = 0\end{aligned}$$

using Cartan's magic formula $\mathcal{L}_X\alpha = \mathbf{d}\mathbf{i}_X\alpha + \mathbf{i}_X\mathbf{d}\alpha$ and (1.4.2). As F_H^t is the flow map for X_H this gives the desired result. □

Noether's theorem still holds even if the Hamiltonian is only infinitesimally invariant, as it is only this local statement which is used in the proof.

1.4.3. Legendre transforms

To relate Lagrangian mechanics to Hamiltonian mechanics we define the *Legendre transform* or *fibre derivative* $\mathbb{F}L : TQ \to T^*Q$ by

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

which has coordinate form

$$\mathbb{F}L : (q, \dot{q}) \mapsto (q, p) = \left(q, \frac{\partial L}{\partial \dot{q}}(q, \dot{q})\right).$$

If the fibre derivative of L is locally an isomorphism then we say that L is *regular*, and if it is a global isomorphism then L is said to be *hyperregular*. We will generally assume that we are working with hyperregular Lagrangians.

The *fibre derivative* of a Hamiltonian is the map $\mathbb{F}H : T^*Q \to TQ$ defined by

$$\alpha_q \cdot \mathbb{F}H(\beta_q) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} H(\beta_q + \epsilon\alpha_q),$$

which in coordinates is

$$\mathbb{F}H : (q, p) \mapsto (q, \dot{q}) = \left(q, \frac{\partial H}{\partial p}(q, p)\right).$$

Similarly to the situations for Lagrangians, we say that H is *regular* if $\mathbb{F}H$ is a local isomorphism, and that H is *hyperregular* if $\mathbb{F}H$ is a global isomorphism.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

The canonical one- and two-forms and the Hamiltonian momentum maps are related to the Lagrangian one- and two-forms and the Lagrangian momentum maps by pullback under the fibre derivative, so that

$$\Theta_L = (\mathbb{F}L)^*\Theta, \quad \Omega_L = (\mathbb{F}L)^*\Omega, \quad \text{and} \quad J_L = (\mathbb{F}L)^* J_H.$$

If we additionally relate the Hamiltonian to the Lagrangian by

$$H(q,p) = \mathbb{F}L(q,\dot{q}) \cdot \dot{q} - L(q,\dot{q}), \tag{1.4.4}$$

where (q,p) and $(q,\dot{q})$ are related by the Legendre transform, then the Hamiltonian and Lagrangian vector fields and their associated flow maps will also be related by pullback to give

$$X_L = (\mathbb{F}L)^* X_H; \quad F_L^t = (\mathbb{F}L)^{-1} \circ F_H^t \circ \mathbb{F}L.$$

In coordinates this means that Hamilton's equations (1.4.3) are equivalent to the Euler–Lagrange equations (1.3.3). To see this, we compute the derivatives of (1.4.4) to give

$$\begin{aligned}
\frac{\partial H}{\partial q}(q,p) &= p \cdot \frac{\partial \dot{q}}{\partial q} - \frac{\partial L}{\partial q}(q,\dot{q}) - \frac{\partial L}{\partial \dot{q}}(q,\dot{q})\frac{\partial \dot{q}}{\partial q} \\
&= \frac{\partial L}{\partial q}(q,\dot{q}) && \text{(1.4.5a)} \\
&= -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}(q,\dot{q})\right) \\
&= -\dot{p}, && \text{(1.4.5b)} \\
\frac{\partial H}{\partial \dot{q}}(q,p) &= \dot{q} + p \cdot \frac{\partial \dot{q}}{\partial p} - \frac{\partial L}{\partial \dot{q}}(q,\dot{q})\frac{\partial \dot{q}}{\partial p} \\
&= \dot{q}, && \text{(1.4.5c)}
\end{aligned}$$

where $p = \mathbb{F}L(q,\dot{q})$ defines $\dot{q}$ as a function of (q,p).

A similar calculation to the above also shows that if L is hyperregular and H is defined by (1.4.4) then H will also be hyperregular and the fibre derivatives will satisfy $\mathbb{F}H = (\mathbb{F}L)^{-1}$. The converse statement also holds (see Marsden and Ratiu (1999) for more details).

The above relationship between the Hamiltonian and Lagrangian flows can be summarized by the following commutative diagram, where we recall that the symplectic forms and momentum maps are also preserved under each map.

$$\begin{array}{ccc}
TQ & \xrightarrow{F_L^t} & TQ \\
{\scriptstyle \mathbb{F}L}\downarrow & & \downarrow{\scriptstyle \mathbb{F}L} \\
T^*Q & \xrightarrow[F_H^t]{} & T^*Q
\end{array} \tag{1.4.6}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

One consequence of this relationship between the Lagrangian and Hamiltonian flow maps is a condition for when the Lagrangian vector field and flow map are well-defined.

Theorem 1.4.3. Given a Lagrangian $L : TQ \to \mathbb{R}$, the Lagrangian vector field X_L, and hence the Lagrangian flow map F_L, are well-defined if and only if the Lagrangian is regular.

Proof. This can be seen by relating the Hamiltonian and Lagrangian settings with $\mathbb{F}L$, or by computing the Euler–Lagrange equations in coordinates to give

$$\begin{aligned} 0 &= D_1 L(q,\dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t} D_2 L(q,\dot{q}) \\ &= D_1 L(q,\dot{q}) - D_1 D_2 L(q,\dot{q}) \cdot \dot{q} - D_2 D_2 L(q,\dot{q}) \cdot \ddot{q}. \end{aligned}$$

Thus, $\ddot{q}$ is well-defined as a function of $(q,\dot{q})$ if and only if $D_2 D_2 L$ is invertible, which by the implicit function theorem is equivalent to $\mathbb{F}L$ being locally invertible. □

1.4.4. Generating functions

As with Hamiltonian mechanics, a useful general context for discussing canonical transformations and generating functions is that of symplectic manifolds. Here we limit ourselves, as above, to the case of T^*Q with the canonical symplectic form Ω.

Let $F : T^*Q \to T^*Q$ be a transformation from T^*Q to itself and let $\Gamma(F) \subset T^*Q \times T^*Q$ be the graph of F. Consider the one-form on $T^*Q \times T^*Q$ defined by

$$\hat{\Theta} = \pi_2^* \Theta - \pi_1^* \Theta.$$

where $\pi_i : T^*Q \times T^*Q$ are the projections onto the two components. The corresponding two-form is then

$$\hat{\Omega} = -\mathbf{d}\hat{\Theta} = \pi_2^* \Omega - \pi_1^* \Omega.$$

Denoting the inclusion map by $i_F : \Gamma(F) \to T^*Q \times T^*Q$, we see that we have the identities

$$\pi_1 \circ i_F = \pi_1|_{\Gamma(F)}, \quad \text{and} \quad \pi_2 \circ i_F = F \circ \pi_1 \text{ on } \Gamma(F).$$

Using these relations, we have

$$\begin{aligned} i_F^* \hat{\Omega} &= i_F^* (\pi_2^* \Omega - \pi_1^* \Omega) \\ &= (\pi_2 \circ i_F)^* \Omega - (\pi_1 \circ i_F)^* \Omega \\ &= (\pi_1|_{\Gamma(F)})^* (F^* \Omega - \Omega). \end{aligned}$$

Using this last equality, it is clear that *F is a canonical transformation if and only if* $i_F^* \hat{\Omega} = 0$ or, equivalently, if and only if $\mathbf{d}(i_F^* \hat{\Theta}) = 0$. By the

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Poincaré lemma, this last statement is equivalent to there existing, at least locally, a function $S : \Gamma(F) \to \mathbb{R}$ such that $i_F^* \hat{\Theta} = \mathbf{d}S$. Such a function S is known as the *generating function* of the symplectic transformation F. Note that S is not unique.

The generating function S is specified on the graph $\Gamma(F)$, and so can be expressed in any local coordinate system on $\Gamma(F)$. The standard choices, for coordinates (q_0, p_0, q_1, p_1) on $T^*Q \times T^*Q$, are any two of the four quantities q_0, p_0, q_1 and p_1; note that $\Gamma(F)$ has the same dimension as T^*Q.

1.4.5. Coordinate expression

We will be particularly interested in the choice (q_0, q_1) as local coordinates on $\Gamma(F)$, and so we give the coordinate expressions for the above general generating function derivation for this particular case. This choice results in generating functions of the so-called *first kind* (Goldstein 1980).

Consider a function $S : Q \times Q \to \mathbb{R}$. Its differential is

$$\mathbf{d}S = \frac{\partial S}{\partial q_0}\mathbf{d}q_0 + \frac{\partial S}{\partial q_1}\mathbf{d}q_1.$$

Let $F : T^*Q \to T^*Q$ be the canonical transformation generated by S. In coordinates, the quantity $i_F^* \hat{\Theta}$ is

$$i_F^* \hat{\Theta} = -p_0 \mathbf{d}q_0 + p_1 \mathbf{d}q_1,$$

and so the condition $i_F^* \hat{\Theta} = \mathbf{d}S$ reduces to the equations

$$p_0 = -\frac{\partial S}{\partial q_0}(q_0, q_1), \tag{1.4.7a}$$

$$p_1 = \frac{\partial S}{\partial q_1}(q_0, q_1), \tag{1.4.7b}$$

which are an implicit definition of the transformation $F : (q_0, p_0) \mapsto (q_1, p_1)$. From the above general theory, we know that such a transformation is automatically symplectic, and that all symplectic transformations have such a representation, at least locally.

Note that there is not a one-to-one correspondence between symplectic transformations and real-valued functions on $Q \times Q$, because for some functions the above equations either have no solutions or multiple solutions, and so there is no well-defined map $(q_0, p_0) \mapsto (q_1, p_1)$. For example, taking $S(q_0, q_1) = 0$ forces p_0 to be zero, and so there is no corresponding map φ. In addition, one has to be careful about the special case of generating the identity transformation, as was noted in Channell and Scovel (1990) and Ge and Marsden (1988). As we will see later, this situation is identical to the existence of solutions to the discrete Euler–Lagrange equations, and, as in that case, we will assume for now that we choose generating functions and time-steps so that the equations (1.4.7) do indeed have solutions.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

1.5. Discrete variational mechanics: Hamiltonian viewpoint

1.5.1. Discrete Legendre transforms

Just as the standard Legendre transform maps the Lagrangian state space TQ to the Hamiltonian phase space T^*Q, we can define *discrete Legendre transforms* or *discrete fibre derivatives* $\mathbb{F}^+L_d, \mathbb{F}^-L_d : Q \times Q \to T^*Q$, which map the discrete state space $Q \times Q$ to T^*Q. These are given by

$$\begin{aligned}\mathbb{F}^+L_d(q_0, q_1) \cdot \delta q_1 &= D_2L_d(q_0, q_1) \cdot \delta q_1,\\ \mathbb{F}^-L_d(q_0, q_1) \cdot \delta q_0 &= -D_1L_d(q_0, q_1) \cdot \delta q_0,\end{aligned}$$

which can be written

$$\begin{aligned}\mathbb{F}^+L_d : (q_0, q_1) &\mapsto (q_1, p_1) = (q_1, D_2L_d(q_0, q_1)),\\ \mathbb{F}^-L_d : (q_0, q_1) &\mapsto (q_0, p_0) = (q_0, -D_1L_d(q_0, q_1)).\end{aligned}$$

If both discrete fibre derivatives are locally isomorphisms (for nearby q_0 and q_1), then we say that L_d is *regular*. We will generally assume that we are working with regular discrete Lagrangians. In some special cases, such as if Q is a vector space, it may be that both discrete fibre derivatives are global isomorphisms. In that case we say that L_d is *hyperregular*.

Using the discrete fibre derivatives it can be seen that the canonical one- and two-forms and Hamiltonian momentum maps are related to the discrete Lagrangian forms and discrete momentum maps by pullback, so that

$$\Theta^{\pm}_{L_d} = (\mathbb{F}^{\pm}L_d)^*\Theta, \quad \Omega_{L_d} = (\mathbb{F}^{\pm}L_d)^*\Omega, \quad \text{and} \quad J^{\pm}_{L_d} = (\mathbb{F}^{\pm}L_d)^*J_H.$$

When the discrete momentum maps arise from a symmetry action, the pullback of the Hamiltonian momentum map by either discrete Legendre transform gives the unique discrete momentum map $J_{L_d} = (\mathbb{F}^{\pm}L_d)^*J_H$.

In the continuous case there is a particular relationship between a Lagrangian and a Hamiltonian so that the corresponding vector fields and flow maps are related by pullback under the Legendre transform. Indeed, we rarely consider pairs of Lagrangian and Hamiltonian systems which are not related in this way. In the discrete case a similar relationship exists, as will be shown in Section 1.6.

Unlike the continuous case, however, we will generally be interested in discrete Lagrangian systems that do not exactly correspond to a given Hamiltonian system. In this case, the symplectic structures and momentum maps are related by pullback under the discrete Legendre transforms, but the flow maps are not. As we will see later, this is a reflection of the fact that discrete Lagrangian systems can be regarded as symplectic-momentum integrators. The relationship between the energies of a discrete Lagrangian system and a Hamiltonian system is investigated in Part 4.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

1.5.2. Momentum matching

The discrete fibre derivatives also permit a new interpretation of the discrete Euler–Lagrange equations. To see this, we introduce the notation

$$p^+_{k,k+1} = p^+(q_k, q_{k+1}) = \mathbb{F}^+ L_d(q_k, q_{k+1}),$$
$$p^-_{k,k+1} = p^-(q_k, q_{k+1}) = \mathbb{F}^- L_d(q_k, q_{k+1}),$$

for the momentum at the two endpoints of each interval $[k, k+1]$. Now observe that the discrete Euler–Lagrange equations are

$$D_2 L_d(q_{k-1}, q_k) = -D_1 L_d(q_k, q_{k+1}),$$

which can be written as

$$\mathbb{F}^+ L_d(q_{k-1}, q_k) = \mathbb{F}^- L_d(q_k, q_{k+1}), \tag{1.5.1}$$

or simply

$$p^+_{k-1,k} = p^-_{k,k+1}.$$

That is, the discrete Euler–Lagrange equations are simply enforcing the condition that the momentum at time k should be the same when evaluated from the lower interval $[k-1, k]$ or the upper interval $[k, k+1]$. This means that along a solution curve there is a unique momentum at each time k, which we denote by

$$p_k = p^+_{k-1,k} = p^-_{k,k+1}.$$

A discrete trajectory $\{q_k\}_{k=0}^N$ in Q can thus also be regarded as either a trajectory $\{(q_k, q_{k+1})\}_{k=0}^{N-1}$ in $Q \times Q$ or, equivalently, as a trajectory $\{(q_k, p_k)\}_{k=0}^N$ in T^*Q.

It will be useful to note that (1.5.1) can be written as

$$\mathbb{F}^+ L_d = \mathbb{F}^- L_d \circ F_{L_d}. \tag{1.5.2}$$

A consequence of viewing the discrete Euler–Lagrange equations as a matching of momenta is that it gives a condition for when the discrete Lagrangian evolution operator and discrete Lagrangian map are well-defined.

Theorem 1.5.1. Given a discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$, the discrete Lagrangian evolution operator X_{L_d} and the discrete Lagrange map F_{L_d} are well-defined if and only if $\mathbb{F}^- L_d$ is locally an isomorphism. The discrete Lagrange map is well-defined and invertible if and only if the discrete Lagrangian is regular.

Proof. Given $(q_0, q_1) \in Q \times Q$, the point $q_2 \in Q$ required to satisfy

$$X_{L_d}(q_0, q_1) = (q_0, q_1, q_1, q_2)$$

is defined by equation (1.5.1), and so q_2 is uniquely defined as a function of

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

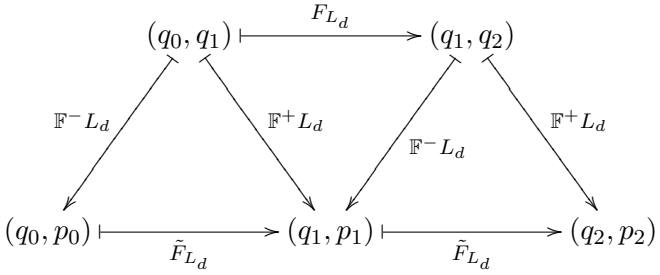
q_0 and q_1 if and only if $\mathbb{F}^- L_d$ is locally an isomorphism. From the definition of F_{L_d} it is well-defined if and only if X_{L_d} is.

The above argument only implies that F_{L_d} is well-defined as a map, however, meaning that it can be applied to map forward in time. For it to be invertible, equation (1.5.1) shows that it is necessary and sufficient for $\mathbb{F}^+ L_d$ also to be a local isomorphism, which is equivalent to regularity of L_d. □

1.5.3. Discrete Hamiltonian maps

Using the discrete fibre derivatives also enables us to push the discrete Lagrangian map $F_{L_d} : Q \times Q \to Q \times Q$ forward to T^*Q. We define the *discrete Hamiltonian map* $\tilde{F}_{L_d} : T^*Q \to T^*Q$ by $\tilde{F}_{L_d} = \mathbb{F}^{\pm} L_d \circ F_{L_d} \circ (\mathbb{F}^{\pm} L_d)^{-1}$. The fact that the discrete Hamiltonian map can be equivalently defined with either discrete Legendre transform is a consequence of the following theorem.

Theorem 1.5.2. The following diagram commutes.



(1.5.3)

Proof. The central triangle is simply (1.5.2). Assume that we define the discrete Hamiltonian map by $\tilde{F}_{L_d} = \mathbb{F}^+ L_d \circ F_{L_d} \circ (\mathbb{F}^+ L_d)^{-1}$, which gives the right-hand parallelogram. Replicating the right-hand triangle on the left-hand side completes the diagram. If we choose to use the other discrete Legendre transform then the reverse argument applies. □

Corollary 1.5.3. The following three definitions of the discrete Hamiltonian map,

$$\begin{aligned}
\tilde{F}_{L_d} &= \mathbb{F}^+ L_d \circ F_{L_d} \circ (\mathbb{F}^+ L_d)^{-1}, \\
\tilde{F}_{L_d} &= \mathbb{F}^- L_d \circ F_{L_d} \circ (\mathbb{F}^- L_d)^{-1}, \\
\tilde{F}_{L_d} &= \mathbb{F}^+ L_d \circ (\mathbb{F}^- L_d)^{-1},
\end{aligned}$$

are equivalent and have coordinate expression $\tilde{F}_{L_d} : (q_0, p_0) \mapsto (q_1, p_1)$, where

$$p_0 = -D_1 L_d(q_0, q_1), \tag{1.5.4a}$$
$$p_1 = D_2 L_d(q_0, q_1). \tag{1.5.4b}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Proof. The equivalence of the three definitions can be read directly from the diagram in Theorem 1.5.2.

The coordinate expression for $\tilde{F}_{L_d} : (q_0, p_0) \mapsto (q_1, p_1)$ can be readily seen from the definition $\tilde{F}_{L_d} = \mathbb{F}^+ L_d \circ (\mathbb{F}^- L_d)^{-1}$. Taking initial condition $(q_0, p_0) \in T^*Q$ and setting $(q_0, q_1) = (\mathbb{F}^- L_d)^{-1}(q_0, p_0)$ implies that $p_0 = -D_1 L_d(q_0, q_1)$, which is (1.5.4a). Now, letting $(q_1, p_1) = \mathbb{F}^+ L_d(q_0, q_1)$ gives $p_1 = D_2 L_d(q_0, q_1)$, which is (1.5.4b). □

As the discrete Lagrangian map preserves the discrete symplectic form and discrete momentum maps on $Q \times Q$, the discrete Hamiltonian map will preserve the pushforwards of these structures. As we saw above, however, these are simply the canonical symplectic form and canonical momentum maps on T^*Q, and so the discrete Hamiltonian map is symplectic and momentum-preserving.

We can summarize the relationship between the discrete and continuous systems in the following diagram, where the dashed arrows represent the discretization.

$$\begin{array}{ccc} TQ, F_L & \dashrightarrow & Q \times Q, F_{L_d} \\ {\scriptstyle \mathbb{F}L}\downarrow & & \downarrow{\scriptstyle \mathbb{F}L_d} \\ T^*Q, F_H & \dashrightarrow & T^*Q, \tilde{F}_{L_d} \end{array} \tag{1.5.5}$$

1.5.4. Discrete Lagrangians are generating functions

As we have seen above, a discrete Lagrangian is a real-valued function on $Q \times Q$ which defines a map $\tilde{F}_{L_d} : T^*Q \to T^*Q$. In fact, a discrete Lagrangian is simply a generating function of the first kind for the map $\tilde{F}_{L_d}$, in the sense defined in Section 1.4. This is seen by comparing the coordinate expression (1.5.4) for the discrete Hamiltonian map with the expression (1.4.7) for the map generated by a generating function of the first kind.

1.6. Correspondence between discrete and continuous mechanics

We will now define a particular choice of discrete Lagrangian which gives an *exact* correspondence between discrete and continuous systems. To do this, we must firstly recall the following fact.

Theorem 1.6.1. Consider a regular Lagrangian L for a configuration manifold Q, two points $q_0, q_1 \in Q$ and a time $h \in \mathbb{R}$. If $\|q_1 - q_0\|$ and $|h|$ are sufficiently small then there exists a unique solution $q : \mathbb{R} \to Q$ of the Euler–Lagrange equations for L satisfying $q(0) = q_0$ and $q(h) = q_1$.

Proof. See Marsden and Ratiu (1999). □

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

For some regular Lagrangian L we now define the *exact discrete Lagrangian* to be

$$L_d^E(q_0, q_1, h) = \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t))\,\mathrm{d}t \tag{1.6.1}$$

for sufficiently small h and close q_0 and q_1. Here $q_{0,1}(t)$ is the unique solution of the Euler–Lagrange equations for L which satisfies the boundary conditions $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$, and whose existence is guaranteed by Theorem 1.6.1.

We will now see that with this exact discrete Lagrangian there is an exact correspondence between the discrete and continuous systems. To do this, we will first establish that there is a special relationship between the Legendre transforms of a regular Lagrangian and its corresponding exact discrete Lagrangian. This result will also prove that exact discrete Lagrangians are automatically regular.

Lemma 1.6.2. A regular Lagrangian L and the corresponding exact discrete Lagrangian L_d^E have Legendre transforms related by

$$\begin{aligned}\mathbb{F}^+ L_d^E(q_0, q_1, h) &= \mathbb{F}L(q_{0,1}(h), \dot{q}_{0,1}(h)),\\ \mathbb{F}^- L_d^E(q_0, q_1, h) &= \mathbb{F}L(q_{0,1}(0), \dot{q}_{0,1}(0)),\end{aligned}$$

for sufficiently small h and close $q_0, q_1 \in Q$.

Proof. We begin with $\mathbb{F}^- L_d^E$ and compute

$$\begin{aligned}\mathbb{F}^- L_d^E(q_0, q_1, h) &= -\int_0^h \left[\frac{\partial L}{\partial q}\cdot\frac{\partial q_{0,1}}{\partial q_0} + \frac{\partial L}{\partial \dot{q}}\cdot\frac{\partial \dot{q}_{0,1}}{\partial q_0}\right]\mathrm{d}t\\ &= -\int_0^h \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}\right]\cdot\frac{\partial q_{0,1}}{\partial q_0}\,\mathrm{d}t - \left[\frac{\partial L}{\partial \dot{q}}\cdot\frac{\partial q_{0,1}}{\partial q_0}\right]_0^h,\end{aligned}$$

using integration by parts. The fact that $q_{0,1}(t)$ is a solution of the Euler–Lagrange equations shows that the first term is zero. To compute the second term we recall that $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$, so that

$$\frac{\partial q_{0,1}}{\partial q_0}(0) = \mathrm{Id} \qquad \text{and} \qquad \frac{\partial q_{0,1}}{\partial q_0}(h) = 0.$$

Substituting these into the above expression for $\mathbb{F}^- L_d^E$ now gives

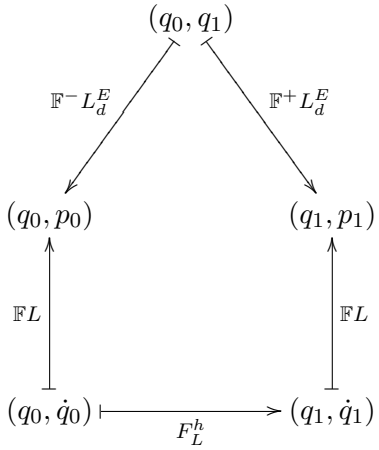
$$\mathbb{F}^- L_d^E(q_0, q_1, h) = \frac{\partial L}{\partial \dot{q}}(q_{0,1}(0), \dot{q}_{0,1}(0)),$$

which is simply the definition of $\mathbb{F}L(q_{0,1}(0), \dot{q}_{0,1}(0))$.

The result for $\mathbb{F}^+ L_d^E$ can be established by a similar computation. □

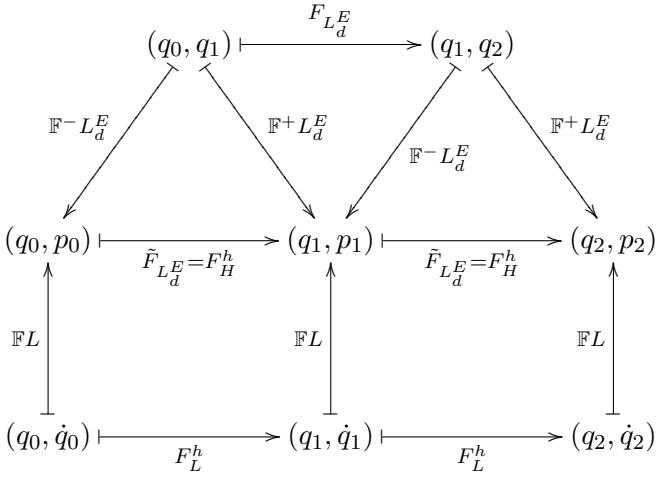
https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Since $(q_{0,1}(h), \dot{q}_{0,1}(h)) = F_L^h(q_{0,1}(0), \dot{q}_{0,1}(0))$, Lemma 1.6.2 is equivalent to the following commutative diagram.



(1.6.2)

Combining this diagram with (1.4.6) and (1.5.3) gives the following commutative diagram for the exact discrete Lagrangian.



(1.6.3)

This proves the following theorem.

Theorem 1.6.3. Consider a regular Lagrangian L, its corresponding exact discrete Lagrangian L_d^E, and the pushforward of both the continuous and discrete systems to T^*Q, yielding a Hamiltonian system with Hamiltonian H and a discrete Hamiltonian map $\tilde{F}_{L_d^E}$, respectively. Then, for a sufficiently small time-step $h \in \mathbb{R}$, the Hamiltonian flow map equals the pushforward discrete Lagrangian map:

$$F_H^h = \tilde{F}_{L_d^E}.$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

This theorem is a statement about the time evolution of the system, and can also be interpreted as saying that the diagram (1.5.5) commutes with the dashed arrows understood as samples at times $\{t_k\}_{k=0}^N$, rather than merely as discretizations.

We can also interpret the equivalence of the discrete and continuous systems as a statement about trajectories. On the Lagrangian side, this gives the following theorem.

Theorem 1.6.4. Take a series of times $\{t_k = kh,\ k = 0, \dots, N\}$ for a sufficiently small time-step h, and a regular Lagrangian L and its corresponding exact discrete Lagrangian L_d^E. Then solutions $q : [0, t_N] \to Q$ of the Euler–Lagrange equations for L and solutions $\{q_k\}_{k=0}^N$ of the discrete Euler–Lagrange equations for L_d^E are related by

$$q_k = q(t_k) \text{ for } k = 0, \dots, N, \tag{1.6.4a}$$

$$q(t) = q_{k,k+1}(t) \text{ for } t \in [t_k, t_{k+1}]. \tag{1.6.4b}$$

Here the curves $q_{k,k+1} : [t_k, t_{k+1}] \to Q$ are the unique solutions of the Euler–Lagrange equations for L satisfying $q_{k,k+1}(kh) = q_k$ and $q_{k,k+1}((k+1)h) = q_{k+1}$.

Proof. The main non-obvious issue is smoothness. Let $q(t)$ be a solution of the Euler–Lagrange equations for L and define $\{q_k\}_{k=0}^N$ by (1.6.4a). Now the discrete Euler–Lagrange equations at time k are simply a matching of discrete Legendre transforms, as in (1.5.1), but by construction and Lemma 1.6.2 both sides of this expression are equal to $\mathbb{F}L(q(t_k), \dot{q}(t_k))$. We thus see that $\{q_k\}_{k=0}^N$ is a solution of the discrete Euler–Lagrange equations.

Conversely, let $\{q_k\}_{k=0}^N$ be a solution of the discrete Euler–Lagrange equations for L_d^E and define $q : [0, t_N] \to Q$ by (1.6.4b). Clearly $q(t)$ is C^2 and a solution of the Euler–Lagrange equations on each open interval (t_k, t_{k+1}), and so we must only establish that it is also C^2 at each t_k, from which it will follow that it is C^2 and a solution on the entire interval $[0, t_N]$.

At time t_k the discrete Euler–Lagrange equations in the form (1.5.1) together with Lemma 1.6.2 reduce to

$$\mathbb{F}L(q_{k-1,k}(t_k), \dot{q}_{k-1,k}(t_k)) = \mathbb{F}L(q_{k,k+1}(t_k), \dot{q}_{k,k+1}(t_k)),$$

and, as $\mathbb{F}L$ is a local isomorphism (due to the regularity of L), we see that $q(t)$ is C^1 on $[0, t_N]$. The regularity of L also implies that

$$\ddot{q}(t) = (D_2 D_2 L)^{-1}(D_1 L - D_1 D_2 L \cdot \dot{q}(t))$$

on each open interval (t_k, t_{k+1}), and as the right-hand side only depends on $q(t)$ and $\dot{q}(t)$ this expression is continuous at each t_k, giving that $q(t)$ is indeed C^2 on $[0, t_N]$. □

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

To summarize, given Lagrangian and Hamiltonian systems with the Legendre transform mapping between them, the symplectic forms and momentum maps are always related by pullback under $\mathbb{F}L$. If, in addition, L and H satisfy the special relationship (1.4.4), then the flow maps and energy functions will also be related by pullback.

Exactly the same statements hold for the relationship between a discrete Lagrangian system and a Hamiltonian system. However, when discussing continuous systems we almost always assume that L and H are related by (1.4.4), whereas for discrete systems we generally do not assume that L_d and L or H are related by (1.6.1). This is because we are interested in using the discrete mechanics to derive integrators, and the exact discrete Lagrangian is generally not computable.

1.7. Background: Hamilton–Jacobi theory

1.7.1. Generating function for the flow

As discussed in Section 1.4, it is a standard result that the flow map F_H^t of a Hamiltonian system is a canonical map for each fixed time t. From the generating function theory, it must therefore have a generating function $S(q_0, q_1, t)$. We will now derive a partial differential equation which S must satisfy.

Consider first the time-preserving extension of F_H to the map

$$\hat{F}_H : T^*Q \times \mathbb{R} \to T^*Q \times \mathbb{R}, \quad (p_q, t) \mapsto (F_H^t(p_q), t).$$

Let $\pi_{T^*Q} : T^*Q \times \mathbb{R} \to T^*Q$ be the projection, and define the *extended canonical one-form* and the *extended canonical two-form* to be

$$\begin{aligned}\Theta_H &= i_{T^*Q}^*\Theta - i_{T^*Q}^*H \wedge \mathbf{d}t,\\ \Omega_H &= -\mathbf{d}\Theta_H = i_{T^*Q}^*\Omega - i_{T^*Q}^*\mathbf{d}H \wedge \mathbf{d}t.\end{aligned}$$

We now calculate

$$\begin{aligned}T\hat{F}_H \cdot \Big(\delta p_q, \delta t) &= \Big(TF_H^t \cdot \delta p_q + \frac{\partial}{\partial t}F_H^t(p_q) \cdot \delta t, \delta t\Big)\\ &= (TF_H^t \cdot \delta p_q + X_H \circ F_H^t \cdot \delta t, \delta t),\end{aligned}$$

using that F_H^t is the flow map of the vector field X_H, and so

$$\begin{aligned}\hat{F}_H^*\Omega_H &= (i_{T^*Q} \circ \hat{F}_H)^*\Omega - ((i_{T^*Q} \circ \hat{F}_H)^*\mathbf{d}H) \wedge (\hat{F}_H^*\mathbf{d}t)\\ &= i_{T^*Q}^*(F_H^t)^*\Omega + (i_{T^*Q}^*(F_H^t)^*\mathbf{d}H) \wedge \mathbf{d}t - ((i_{T^*Q} \circ \hat{F}_H)^*\mathbf{d}H) \wedge \mathbf{d}t\\ &= (i_{T^*Q})^*(F_H^t)^*\Omega = i_{T^*Q}^*\Omega\end{aligned}$$

as F_H^t preserves Ω for fixed t. This identity essentially states that the extended flow map pulls back the extended symplectic form to the standard symplectic form.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Consider now the space $T^*Q \times \mathbb{R} \times T^*Q$ and the projection $\pi_1 : T^*Q \times \mathbb{R} \times T^*Q \to T^*Q \times \mathbb{R}$ onto the first two components and $\pi_2 : T^*Q \times \mathbb{R} \times T^*Q \to T^*Q \times \mathbb{R}$ onto the last two components. Define the one-form

$$\hat{\Theta} = \pi_2^* \Theta_H - \pi_1^* i_{T^*Q}^* \Theta,$$

and let the corresponding two-form be

$$\hat{\Omega} = -\mathbf{d}\hat{\Theta} = \pi_2^* \Omega_H - \pi_1^* i_{T^*Q}^* \Omega.$$

The flow map of the Hamiltonian system acts as $F_H : T^*Q \times \mathbb{R} \to T^*Q$ and so the graph of F_H is a subset $\Gamma(F_H) \subset T^*Q \times \mathbb{R} \times T^*Q$. Denote the inclusion map by $i_{F_H} : \Gamma(F_H) \to T^*Q \times \mathbb{R} \times T^*Q$. We now observe that

$$\begin{aligned} \pi_1 \circ i_{F_H} &= \pi_1|_{\Gamma(F_H)}, \\ \pi_2 \circ i_{F_H} &= \hat{F}_H \circ \pi_1 \text{ on } \Gamma(F_H), \end{aligned}$$

and using these relations calculate

$$\begin{aligned} i_{F_H}^* \hat{\Omega} &= i_{F_H}^* \pi_2^* \Omega_H - i_{F_H}^* \pi_1^* i_{T^*Q}^* \Omega \\ &= (\pi_2 \circ i_{F_H})^* \Omega_H - (\pi_1 \circ i_{F_H})^* i_{T^*Q}^* \Omega \\ &= (\pi_1|_{\Gamma(F_H)})^* (\hat{F}_H^* \Omega_H - i_{T^*Q}^* \Omega) \\ &= 0. \end{aligned}$$

We have thus established that $\mathbf{d}(i_{F_H}^* \hat{\Theta}) = 0$ and so, by the Poincaré lemma, there must locally exist a function $S : \Gamma(F_H) \to \mathbb{R}$ so that $i_{F_H}^* \hat{\Theta} = \mathbf{d}S$. It is clear that restricting the above derivations to a section with fixed t simply reproduces the earlier derivation of generating functions for symplectic maps, and so the restriction $S^t : \Gamma(F_H^t) \to \mathbb{R}$ is a generating function for the map $F_H^t : T^*Q \to T^*Q$. The additional information contained in the statement $i_{F_H}^* \hat{\Theta} = \mathbf{d}S$ dictates how S depends on t.

1.7.2. Hamilton–Jacobi equation

As for the case of general generating functions discussed in Section 1.4 we will now choose a particular set of coordinates on ΓF_H and investigate the implications of $i_{F_H}^* \hat{\Theta} = \mathbf{d}S$.

Consistent with our earlier choice, we will take coordinates (q_0, q_1, t) for $\Gamma(F_H)$ and thus regard the generating function as a map $S : Q \times Q \times \mathbb{R} \to \mathbb{R}$. The differential is thus

$$\mathbf{d}S = \frac{\partial S}{\partial q_0} \mathbf{d}q_0 + \frac{\partial S}{\partial q_1} \mathbf{d}q_1 + \frac{\partial S}{\partial t} \mathbf{d}t,$$

and we also get

$$\hat{\Theta} = -p_0 \mathbf{d}q_0 + p_1 \mathbf{d}q_1 - H(q_1, p_1) \mathbf{d}t,$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

so the condition $i_{F_H}^* \hat{\Theta} = \mathbf{d}S$ is

$$p_0 = -\frac{\partial S}{\partial q_0}(q_0, q_1, t),$$
$$p_1 = \frac{\partial S}{\partial q_1}(q_0, q_1, t),$$
$$H\left(q_1, \frac{\partial S}{\partial q_1}(q_0, q_1, t)\right) = \frac{\partial S}{\partial t}(q_0, q_1, t).$$

The first two equations are simply the standard relations which implicitly specify the map F_H^t from the generating function S^t. The third equation specifies the time-dependence of S and is known as the *Hamilton–Jacobi PDE*, and can be regarded as a partial-differential equation to be solved for S.

To fully specify the Hamilton–Jacobi PDE it is necessary also to provide boundary conditions. As it is first-order in t, it is clear that specifying S as a function of q_0 and q_1 at some time t will define the solution in a neighbourhood of that time. This is equivalent to specifying the map generated by S at some time, up to an arbitrary function of t. Taking this to be the flow map for some fixed time, we see that the unique solution of the Hamilton–Jacobi PDE must be the flow map for nearby t.

1.7.3. Jacobi's solution

While it is possible in principle to solve the Hamilton–Jacobi PDE directly for S, it is generally nonlinear and a closed form solution is not normally possible. By 1840, however, Jacobi had realized that the solution is simply the action of the trajectory joining q_0 and q_1 in time t: see Jacobi (1866). This is known as *Jacobi's solution*,

$$S(q_0, q_1, t) = \int_0^t L(q_{0,1}(\tau), \dot{q}_{0,1}(\tau))\, \mathrm{d}\tau, \tag{1.7.1}$$

where $q_{0,1}(t)$ is a solution of the Euler–Lagrange equations for L satisfying the boundary conditions $q(0) = q_0$ and $q(t) = q_1$, and where L and H are related by the Legendre transform (assumed to be regular). This can be proved in the same way as Lemma 1.6.2.

1.8. Discrete variational mechanics: Hamilton–Jacobi viewpoint

As was discussed in Section 1.5, a discrete Lagrangian can be regarded as the generating function for the discrete Hamiltonian map $\tilde{F}_{L_d} : T^*Q \to T^*Q$. We then showed in Section 1.6 that there is a particular choice of discrete Lagrangian, the so-called *exact* discrete Lagrangian, which exactly generates

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

the flow map F_H of the corresponding Hamiltonian system. From the development of Hamilton–Jacobi theory in Section 1.7, it is clear that this exact discrete Lagrangian must be a solution of the Hamilton–Jacobi equation. In fact, as can be seen by comparing the definitions given in equations (1.6.1) and (1.7.1), the exact discrete Lagrangian is precisely Jacobi's solution of the Hamilton–Jacobi equation.

To summarize, discrete Lagrangian mechanics can be regarded as a variational Lagrangian derivation of the standard generating function and Hamilton–Jacobi theory. Discrete Lagrangians generate symplectic transformations, and given a Lagrangian or Hamiltonian system, one can construct the exact discrete Lagrangian which solves the Hamilton–Jacobi equation, and this will then generate the exact flow of the continuous system.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

PART TWO
Variational integrators

2.1. Introduction

We now turn our attention to considering a discrete Lagrangian system as an approximation to a given continuous system. That is, the discrete system is an integrator for the continuous system.

As we have seen, discrete Lagrangian maps preserve the symplectic structure and so, regarded as integrators, they are necessarily symplectic. Furthermore, generating function theory shows that any symplectic integrator for a mechanical system can be regarded as a discrete Lagrangian system, a fact we state here as a theorem.

Theorem 2.1.1. If the integrator $F : T^*Q \times \mathbb{R} \to T^*Q$ is symplectic then there exists[3] a discrete Lagrangian L_d whose discrete Hamiltonian map $\tilde{F}_{L_d}$ is F.

Proof. As shown above in Section 1.4, any symplectic transformation locally has a corresponding generating function, which is then a discrete Lagrangian for the method, as discussed in Section 1.5.4. □

In addition, if the discrete Lagrangian inherits the same symmetry groups as the continuous system, then the discrete system will also preserve the corresponding momentum maps. As an integrator, it will thus be a so-called *symplectic–momentum integrator*.

Just as with continuous mechanics, we have seen that discrete variational mechanics has both a variational (Lagrangian) and a generating function (Hamiltonian) interpretation. These two viewpoints are complementary and both give insight into the behaviour and derivation of useful integrators.

However, the above theorem is not literally used in the construction of variational integrators, but is rather used as the first steps in obtaining inspiration. We will obtain much deeper insight from the variational principle itself and this is, in large part, what sets variational methods apart from standard symplectic methods.

Symplectic integrators have traditionally been approached from a Hamiltonian viewpoint and there is much existing literature treating this topic (see, for example, Hairer, Nørsett and Wanner (1993), Hairer and Wanner (1996), MacKay (1992) and Sanz-Serna (1992*a*)). In this paper, we concentrate on the analysis of symplectic methods from the variational viewpoint, and we reinterpret many standard concepts from ODE integration theory in this light.

[3] The discrete Lagrangian may exist only locally, as is the case with generating functions, as was discussed in Section 1.4.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

It is also important to distinguish the two ways in which we can derive variational or generating function integrators. First, we can attempt to approximately solve the Hamilton–Jacobi PDE for a given system, such as by taking power series expansions of the generating function. This was used in some of the earliest derivations of symplectic integrators (such as De Vogelaére (1956) and Channell and Scovel (1990)). Second, the method we advocate involves trying to approximate the known Jacobi's solution to the Hamilton–Jacobi PDE: that is, we construct discrete Lagrangians that approximate the exact discrete Lagrangian. This approach is powerful not only because of the coherent and unifying underlying theory that reveals the beautiful geometry underlying discrete mechanics, but also because it leads to practical integrators.

In this section we will assume that Q, and thus also TQ and T^*Q, is a finite-dimensional vector space with an inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\| \cdot \|$. In the case that it is not a vector space, we can embed Q within a vector space and use the theory of constrained discrete systems developed below in Section 3.4 and discussed further in Section 3.5.2.

A word of caution: we must be careful about imagining that we can simply pick a coordinate chart and apply the vector space methods described below in such a chart. Doing so indiscriminately can lead to coordinate-dependent integrators that can be unattractive theoretically as well as impractical: for instance, using Euler angles for rigid body integrators has the difficulty that we may spend most of our computational time switching coordinate systems. See, for instance, Wisdom, Peale and Mignard (1984), Leimkuhler and Patrick (1996) and related papers. For some special classes of configuration manifolds, however, such as when Q is a Lie group, there may be particular global coordinate systems that can be used for this purpose.

We will also frequently consider integrators for Lagrangian systems of the form $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q)$. When dealing with such systems, we will always assume that M is a positive-definite symmetric mass matrix, so that $\mathbb{F}L(q, \dot{q}) = M\dot{q}$ and thus that L is regular.

2.1.1. Implementation of variational integrators

Although the distinction between the discrete Lagrangian map $F_{L_d} : Q \times Q \times \mathbb{R} \to Q \times Q$ and its pushforward $\tilde{F}_{L_d} : T^*Q \times \mathbb{R} \to T^*Q$ is important geometrically, for implementation purposes the two maps are essentially the same. This is because of the observation made in Section 1.5.2 that the discrete Euler–Lagrange equations that define F_{L_d} can be interpreted as matching of momenta between adjacent intervals.

In other words, given a trajectory $q_0, q_1, q_2, \ldots, q_{k-1}, q_k$ the map $F_{L_d} : Q \times Q \times \mathbb{R} \to Q \times Q$ calculates q_{k+1} according to

$$D_2 L_d(q_{k-1}, q_k, h) = -D_1 L_d(q_k, q_{k+1}, h).$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

If we now take $p_k = D_2 L_d(q_{k-1}, q_k, h)$ for each k, then this equation is simply

$$p_k = -D_1 L_d(q_k, q_{k+1}, h), \tag{2.1.1}$$

which, together with the next update

$$p_{k+1} = D_2 L_d(q_k, q_{k+1}, h), \tag{2.1.2}$$

defines the pushforward map $\tilde{F}_{L_d} : T^*Q \times \mathbb{R} \to T^*Q$. Another way to think of this is that the p_k are merely storing the values $D_2 L_d(q_{k-1}, q_k, h)$ from the last step.

For this reason it is typically easier to implement a variational integrator as the single step map $\tilde{F}_{L_d}$, as this also provides a simple method of initialization from initial values $(q_0, p_0) \in T^*Q$. Many discrete Lagrangians have pushforward maps that are simple to implement. For example, $\tilde{F}_{L_d}$ may be explicit, or it may be a Runge–Kutta method or other integrator type with standard implementation techniques.

In the general case when no special form is apparent, however, the equations (2.1.1) and (2.1.2) must be solved directly. The update $(q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ thus involves first solving the implicit equation (2.1.1) for q_{k+1} and then evaluating the explicit update (2.1.2) to give p_{k+1}.

To solve the implicit equation (2.1.1) we must typically use an iterative technique such as Newton's method. This involves computing a first guess $q_{k+1,0}$ for q_{k+1}, such as $q_{k+1,0} = 2q_k - q_{k-1}$, and then computing the sequence of approximations $q_{k+1,n}$, $n = 1, 2, \ldots$ until they converge to the solution value q_{k+1}. For Newton's method, the iteration rule is given by

$$q^i_{k+1,n+1} = q^i_{k+1,n} - A^{ij}\left[p^j_k + \frac{\partial L_d}{\partial q^j_0}(q_0, q_1, h)\right],$$

where A^{ij} is the inverse of the matrix

$$A_{ij} = \frac{\partial^2 L_d}{\partial q^i_0 \partial q^j_1}(q_0, q_1, h).$$

In the case that the Lagrangian has a simple form, such as $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q)$, then we can use an initial guess based on p_k, such as $q_{k+1,0} = q_k + hM^{-1}p_k$.

While the Newton's method outlined above typically experiences very fast convergence, it is also expensive to have to recompute A^{ij} at each iteration of the method. For this reason, it is typical to use an approximation to this matrix which can be held constant for all iterations of Newton's method. See Hairer *et al.* (1993) for details of this approach for Runge–Kutta methods.

2.1.2. Equivalence of integrators

Given two discrete Lagrangians L^1_d and L^2_d, we would like to know whether the integrators they generate are the same. Here it will be important to

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

distinguish between the discrete Lagrangian maps $Q \times Q \to Q \times Q$ and the discrete Hamiltonian maps $T^*Q \to T^*Q$. We assume that we are dealing with regular discrete Lagrangians, so that the corresponding maps are well-defined.

We say that L_d^1 is (strongly) *equivalent* to L_d^2 if their discrete *Hamiltonian* maps are equal, so that $\tilde{F}_{L_d^1} = \tilde{F}_{L_d^2}$. Using the expression $\tilde{F}_{L_d^1} = \mathbb{F}^+ L_d^1 \circ (\mathbb{F}^- L_d^1)^{-1}$, we see that if L_d^1 and L_d^2 are equivalent then their discrete Legendre transforms will be equal. This implies that the difference $L_d^\Delta = L_d^1 - L_d^2$ must be a function of h only. That is, $L_d^\Delta(q_0, q_1, h) = f(h)$ for some function f. This is clearly also a sufficient condition, as well as being necessary.

We define L_d^1 to be *weakly equivalent* to L_d^2 if their discrete *Lagrangian* maps $F_{L_d^1}$ and $F_{L_d^2}$ are equal. A sufficient (and presumably necessary) condition for this to be true is that their difference $L_d^\Delta = L_d^1 - L_d^2$ is a *null discrete Lagrangian*; that is, the discrete Euler–Lagrange equations for L_d^Δ are satisfied by any triplet (q_0, q_1, q_2). This terminology follows that of the continuous case, as in, for example, Oliver and Sivaloganathan (1988).

If L_d^Δ is a null discrete Lagrangian, then $D_2 L_d^\Delta(q_0, q_1, h)$ cannot depend on q_0 and $D_1 L_d^\Delta(q_1, q_2, h)$ cannot depend on q_2. Furthermore, these two derivatives must be the negative of each other for all q_1. We thus have that L_d^Δ is a null discrete Lagrangian if and only if it is of the form $L_d^\Delta(q_0, q_1, h) = f(q_1, h) - f(q_0, h)$ for some function f.

Using the above calculations, it is clear that strong equivalence implies weak equivalence of discrete Lagrangians. For variational integrators, weak equivalence is thus in some sense the more fundamental notion. Intuitively, if two integrators give solutions which have the same positions q_k for all time, but different momenta p_k at each step, then we would like to regard the methods as being essentially the same. This is exactly weak equivalence.

2.2. Background: Error analysis

In this section we consider a numerical method $F : T^*Q \times \mathbb{R} \to T^*Q$ which approximates the flow $F_H : T^*Q \times \mathbb{R} \to T^*Q$ of a given Hamiltonian vector field X_H. Error analysis is concerned with difference between an exact trajectory and a discrete trajectory.[4]

[4] The reader should be cautioned that in many circumstances, such as the integration of chaotic or complex systems, it may make little sense to imagine accurately computing an exact, but highly unstable, individual trajectory. Instead, we often want to accurately compute robust quantities, such as statistical measures that are insensitive to modelling errors and dynamical sensitivities. This vision, as important as it is, awaits a theory.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

2.2.1. Local error and method order

An integrator F of X_H is said to be of *order* r if there exist an open set $U \subset T^*Q$ and constants $C_l > 0$ and $h_l > 0$ so that

$$\|F(q,p,h) - F_H(q,p,h)\| \leq C_l h^{r+1} \tag{2.2.1}$$

for all $(q,p) \in U$ and $h \leq h_l$. The expression on the left-hand side of this inequality is known as the *local error*, and if a method has order of at least 1 then it is said to be *consistent*.

2.2.2. Global error and convergence

Having defined the error after one step, we now consider the error after many steps. The integrator F of X_H is said to be *convergent of order* r if there exist an open set $U \subset T^*Q$ and constants $C_g > 0$, $h_g > 0$ and $T_g > 0$ so that

$$\|(F)^N(q,p,h) - F_H(q,p,T)\| \leq C_g h^r,$$

where $h = T/N$, for all $(q,p) \in U$, $h \leq h_g$ and $T \leq T_g$. The expression on the left-hand side is the *global error* at time T.

For one-step methods such as we consider here, convergence follows from a local error bound on the method and a Lipschitz bound on X_H.

Theorem 2.2.1. Suppose that the integrator F for X_H is of order r on the open set $U \subset T^*Q$ with local error constant C_l, and assume that $L > 0$ is such that

$$\left\| \frac{\partial X_H}{\partial (q,p)} \right\| \leq L$$

on U. Then the method is consistent on U with global error constant C_g given by

$$C_g = \frac{C_l}{L}\left(e^{LT_g} - 1\right)$$

Proof. See, for example, Hairer *et al.* (1993). □

2.2.3. Order calculation

Given an integrator F for X_H, the order can be calculated by expanding both the true flow F_H and the integrator F in a Taylor series in h and then comparing terms. If the terms agree up to order r then the method will be of order r.

Here we explicitly write the first few terms of the Taylor series for the true flow for a Hamiltonian of the form $H(q,p) = \frac{1}{2}p^T M^{-1} p + V(q)$. The corresponding Hamiltonian vector field X_H is

$$\dot{q} = M^{-1}p, \tag{2.2.2a}$$

$$\dot{p} = -\nabla V(q), \tag{2.2.2b}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

and so the flow $(q(h), p(h)) = F_H(q_0, p_0, h)$ has the expansion

$$q(h) = q_0 + hM^{-1}p_0 - \frac{1}{2}h^2M^{-1}\nabla V(q_0) + \mathcal{O}(h^3), \tag{2.2.3a}$$

$$p(h) = p_0 - h\nabla V(q_0) - \frac{1}{2}h^2\nabla^2 V(q_0)M^{-1}p_0 + \mathcal{O}(h^3). \tag{2.2.3b}$$

We will see below an example of using this to calculate the order of a simple class of methods.

2.3. Variational error analysis

Rather than considering how closely the trajectory of F matches the exact trajectory given by F_H, we can alternatively consider how closely a discrete Lagrangian matches the ideal discrete Lagrangian given by the action. As we have seen in Section 1.6, if the discrete Lagrangian is equal to the action, then the corresponding discrete Hamiltonian map $\tilde{F}_{L_d}$ will exactly equal the flow F_H. We now investigate what happens when this is only an approximation.

2.3.1. Local variational order

Recall that the exact discrete Lagrangian (1.6.1) is defined by

$$L_d^E(q_0, q_1, h) = \int_0^h L(q, \dot{q})\,\mathrm{d}t,$$

where $q(t)$ is the solution of the Euler–Lagrange equations satisfying $q(0) = q_0$ and $q(h) = q_1$.

We say that a given discrete Lagrangian L_d is of *order* r if there exist an open subset $U_v \subset TQ$ with compact closure and constants $C_v > 0$ and $h_v > 0$ so that

$$\|L_d(q(0), q(h), h) - L_d^E(q(0), q(h), h)\| \leq C_v h^{r+1} \tag{2.3.1}$$

for all solutions $q(t)$ of the Euler–Lagrange equations with initial condition $(q, \dot{q}) \in U_v$ and for all $h \leq h_v$.

2.3.2. Discrete Legendre transform order

The discrete Legendre transforms $\mathbb{F}^+L_d$ and $\mathbb{F}^-L_d$ of a discrete Lagrangian L_d are said to be of *order* r if there exists an open subset $U_f \subset T^*Q$ with compact closure and constants $C_f > 0$ and $h_f > 0$ so that

$$\|\mathbb{F}^+L_d(q(0), q(h), h) - \mathbb{F}^+L_d^E(q(0), q(h), h)\| \leq C_f h^{r+1}, \tag{2.3.2a}$$

$$\|\mathbb{F}^-L_d(q(0), q(h), h) - \mathbb{F}^-L_d^E(q(0), q(h), h)\| \leq C_f h^{r+1}, \tag{2.3.2b}$$

for all solutions $q(t)$ of the Euler–Lagrange equations with initial condition $(q, \dot{q}) \in U_f$ and for all $h \leq h_f$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

The relationship between the orders of a discrete Lagrangian, its discrete Legendre transforms and its discrete Hamiltonian map is given in the following fundamental theorem.

Theorem 2.3.1. Given a regular Lagrangian L and corresponding Hamiltonian H, the following are equivalent for a discrete Lagrangian L_d:

(1) the discrete Hamiltonian map for L_d is of order r,
(2) the discrete Legendre transforms of L_d are of order r,
(3) L_d is equivalent to a discrete Lagrangian of order r.

Proof. Begin by assuming that L_d is equivalent to a discrete Lagrangian of order r, and we will show that the discrete Legendre transforms are of order r. From Section 2.1.2 we know that equivalent discrete Lagrangians have the same discrete Legendre transforms, and we may thus assume without loss that L_d is itself of order r. Now note that (2.3.1) is equivalent to there existing a function $e_v : T^*Q \times \mathbb{R} \to T^*Q$ so that

$$L_d(q(0), q(h), h) = L_d^E(q(0), q(h), h) + h^{r+1} e_v(q(0), q(h), h)$$

with $\|e_v(q(0), q(h), h)\| \leq C_v$ on U_v. Also, from Theorem 1.6.1 it is clear that we can parametrize the set U_v by either the initial condition $(q, \dot{q})$ or by the endpoints $(q(0), q(h))$.

Taking derivatives of the above expression with respect to $q(h)$ gives

$$\mathbb{F}^+ L_d(q(0), q(h), h) = \mathbb{F}^+ L_d^E(q(0), q(h), h) + h^{r+1} D_2 e_v(q(0), q(h), h),$$

and as e_v is smooth and bounded on the compact set $\mathrm{cl}(U_v)$, so too is $D_2 e_v$, giving (2.3.2a). Taking derivatives with respect to $q(0)$ now shows that the discrete Legendre transforms of L_d are of order r.

Now assume that $\mathbb{F}^+ L_d$ and $\mathbb{F}^- L_d$ are of order r, and set

$$e_v(q(0), q(h), h) = \frac{1}{h^{r+1}} \big[L_d(q(0), q(h), h) - L_d^E(q(0), q(h), h) \big].$$

Taking derivatives with respect to $q(0)$ and $q(h)$ and using (2.3.2) shows that $\|D_1 e_v\| \leq C_f$ and $\|D_2 e_v\| \leq C_f$ on $\mathrm{cl}(U_f)$, which is compact. This then implies that $e_v(\cdot, \cdot, h)$ is itself locally bounded in its first two arguments, and so there exists a function $d(h)$ and a constant C_v such that $\|e_v(q(0), q(h), h) - d(h)\| \leq C_v$. This then proves that $L_d(q(0), q(h), h) - d(h)$ has variational order r, and so L_d is equivalent to a discrete Lagrangian of order r.

We will now show the equivalence of the discrete Hamiltonian map being of order r and the discrete Legendre transforms being of order r. To do this we will make use of the following fact, which is a consequence of the implicit function theorem.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Assume that we have smooth functions related by

$$f_1(x,h) = g_1(x,h) + h^{r+1}e_1(x,h),$$
$$f_2(y,h) = g_2(y,h) + h^{r+1}e_2(y,h),$$

with e_1 and e_2 bounded on some compact sets. Then we have

$$f_2(f_1(x,h),h) = g_2(g_1(x,h),h) + h^{r+1}e_{12}(x,h), \tag{2.3.3a}$$
$$f_1^{-1}(y,h) = g_1^{-1}(y,h) + h^{r+1}\bar{e}_1(y,h), \tag{2.3.3b}$$

for some functions $e_{12}(x,h)$ and $\bar{e}_1(y,h)$ bounded on compact sets.

Now assume that $\mathbb{F}^+L_d$ and $\mathbb{F}^-L_d$ are of order r and use Corollary 1.5.3 to write

$$\tilde{F}_{L_d} = \mathbb{F}^+L_d \circ (\mathbb{F}^-L_d)^{-1},$$
$$\tilde{F}_{L_d^E} = \mathbb{F}^+L_d^E \circ (\mathbb{F}^-L_d^E)^{-1}.$$

Equation (2.3.3) gives the existence of a bounded function e_l such that

$$\tilde{F}_{L_d}(q(0),q(h),h) = \tilde{F}_{L_d^E}(q(0),q(h),h) + h^{r+1}e_l(q(0),q(h),h),$$

and thus we see that the discrete Hamiltonian map is of order r.

Finally, assume that $\tilde{F}_{L_d^E}$ is of order r, and observe that

$$(\mathbb{F}^-L_d)^{-1}(q_0,p_0) = (q_0, \pi_Q \circ \tilde{F}_{L_d}(q_0,p_0)),$$

so (2.3.3) implies (2.3.2b). But now we recall from (1.5.2) that

$$\mathbb{F}^+L_d = \tilde{F}_{L_d} \circ \mathbb{F}^-L_d,$$

and together with (2.3.3a) this gives (2.3.2a), showing that the discrete Legendre transforms are of order r. □

2.3.3. Variational order calculation

Given a discrete Lagrangian, its order can be calculated by expanding the expression for $L_d(q(0),q(h),h)$ in a Taylor series in h and comparing this to the same expansion for the exact Lagrangian. If the series agree up to r terms then the discrete Lagrangian is of order r.

We explicitly evaluate the first few terms of the expansion of the exact discrete Lagrangian to give

$$L_d^E(q(0),q(h),h) = hL(q,\dot{q}) + \frac{1}{2}h^2\left(\frac{\partial L}{\partial q}(q,\dot{q})\cdot\dot{q} + \frac{\partial L}{\partial \dot{q}}(q,\dot{q})\cdot\ddot{q}\right) + \mathcal{O}(h^3), \tag{2.3.4}$$

where $q = q(0)$, $\dot{q} = \dot{q}(0)$ and so forth. Higher derivatives of $q(t)$ are determined by the Euler–Lagrange equations.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Example 2.3.2. An illustrative class of discrete Lagrangian is given by

$$L_d^\alpha(q_0, q_1, h) = hL\left((1-\alpha)q_0 + \alpha q_1, \frac{q_1 - q_0}{h}\right)$$

for some parameter $\alpha \in [0,1]$. Calculating the expansion in h gives

$$L_d^\alpha(q(0), q(h), h) = hL(q,\dot{q}) + \frac{1}{2}h^2\left(2\alpha\frac{\partial L}{\partial q}(q,\dot{q})\cdot\dot{q} + \frac{\partial L}{\partial \dot{q}}(q,\dot{q})\cdot\ddot{q}\right) + \mathcal{O}(h^3).$$

Comparing this to the expansion (2.3.4) for the exact discrete Lagrangian shows that the method is second-order if and only if $\alpha = 1/2$; otherwise it is only consistent.

Calculating the discrete Hamiltonian map for $L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$ gives the integrator $\tilde{F}^\alpha_{L_d} : (q_0, p_0) \mapsto (q_1, p_1)$ defined implicitly by the relations

$$\frac{q_1 - q_0}{h} = M^{-1}\left(\alpha p_0 + (1-\alpha)p_1\right), \tag{2.3.5a}$$

$$\frac{p_1 - p_0}{h} = -\nabla V\left((1-\alpha)q_0 + \alpha q_1\right). \tag{2.3.5b}$$

Note that this method is explicit for $\alpha = 0$ or $\alpha = 1$ and that it is simply the midpoint rule for $\alpha = 1/2$. Expanding (2.3.5) in h gives

$$q_1 = q_0 + hM^{-1}p_0 - (1-\alpha)h^2\nabla V(q_0) + \mathcal{O}(h^3),$$
$$p_1 = p_0 - h\nabla V(q_0) - \alpha h^2\nabla^2 V(q_0)M^{-1}p_0 + \mathcal{O}(h^3),$$

and comparing this to the expansion (2.2.3) of the true flow shows that the method is second-order if and only if $\alpha = 1/2$, and otherwise it is only consistent.

The local error and the discrete Lagrangian error thus agree, as expected. ◊

Example 2.3.3. As the expansions of discrete Lagrangians are linear in L_d, if we take the symmetrized discrete Lagrangian

$$L_d^{\mathrm{sym},\alpha} = \frac{1}{2}L_d^\alpha + \frac{1}{2}L_d^{1-\alpha},$$

then the expansion will agree with that of the exact discrete Lagrangian up to terms of order h^2, so it gives a method that is second-order for any α. ◊

2.4. The adjoint of a method and symmetric methods

For a one-step method $F : T^*Q \times \mathbb{R} \to T^*Q$ the *adjoint method* is $F^* : T^*Q \times \mathbb{R} \to T^*Q$ defined by

$$(F^*)^h \circ F^{-h} = \mathrm{Id}; \tag{2.4.1}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

that is, $(F^*)^h = (F^{-h})^{-1}$. The method F is said to be *self-adjoint* if $F^* = F$. Note that we always have $F^{**} = F$.

Given a discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$, we define the *adjoint discrete Lagrangian* to be $L_d^* : Q \times Q \times \mathbb{R} \to \mathbb{R}$ defined by

$$L_d^*(q_0, q_1, h) = -L_d(q_1, q_0, -h). \tag{2.4.2}$$

The discrete Lagrangian L_d is said to be *self-adjoint* if $L_d^* = L_d$. Note that $L_d^{**} = L_d$ for any L_d.

Theorem 2.4.1. If the discrete Lagrangian L_d has discrete Hamiltonian map $\tilde{F}_{L_d}$ then the adjoint L_d^* of the discrete Lagrangian has discrete Hamiltonian map equal to the adjoint map, so that $\tilde{F}_{L_d^*} = \tilde{F}_{L_d}^*$. If the discrete Lagrangian is self-adjoint then the method is self-adjoint. Conversely, if the method is self-adjoint then the discrete Lagrangian is equivalent to a self-adjoint discrete Lagrangian.

Proof. Consider discrete Lagrangians L_d and L_d^* and the corresponding discrete Hamiltonian maps $\tilde{F}_{L_d}$ and $\tilde{F}_{L_d^*}$. For $\tilde{F}_{L_d}$ and $\tilde{F}_{L_d^*}$ to be adjoint, the definition (2.4.1) requires that $\tilde{F}_{L_d}(q_0, p_0, -h) = (q_1, p_1)$ and $\tilde{F}_{L_d^*}(q_1, p_1, h) = (q_0, p_0)$ for all (q_0, p_0). In terms of the generating functions this is

$$\begin{aligned} p_0 &= -D_1 L_d(q_0, q_1, -h), \\ p_1 &= D_2 L_d(q_0, q_1, -h), \\ p_1 &= -D_1 L_d^*(q_1, q_0, h), \\ p_0 &= D_2 L_d^*(q_1, q_0, h). \end{aligned} \tag{2.4.3}$$

Equating the expressions for p_0 and p_1 shows that this, in turn, is equivalent to

$$\begin{aligned} -D_1 L_d(q_0, q_1, -h) &= D_2 L_d^*(q_1, q_0, h), \\ D_2 L_d(q_0, q_1, -h) &= D_1 L_d^*(q_1, q_0, h). \end{aligned} \tag{2.4.4}$$

Now, if L_d and L_d^* are mutually adjoint, then the definition (2.4.2) implies (2.4.4) and so (2.4.3), thus establishing that $\tilde{F}_{L_d}$ and $\tilde{F}_{L_d^*}$ must also be mutually adjoint, which is written $\tilde{F}_{L_d^*} = \tilde{F}_{L_d}^*$. Note that this implies that $\tilde{F}_{L_d^*}^* = \tilde{F}_{L_d}$.

If L_d is self-adjoint and so $L_d = L_d^*$, then this immediately gives that $\tilde{F}_{L_d} = \tilde{F}_{L_d}^*$ and so $\tilde{F}_{L_d}$ is also self-adjoint.

Conversely, if $\tilde{F}_{L_d}$ and $\tilde{F}_{L_d^*}$ are adjoint, then (2.4.1) implies (2.4.3) which implies (2.4.4). As this simply states that the derivatives of L_d and L_d^* with respect to q_0 and q_1 satisfy the requirement (2.4.2) for adjointness it follows that L_d and L_d^* are mutually adjoint up to the addition of a function of h. Symmetry of $\tilde{F}_{L_d}$ thus implies symmetry of L_d up to a function of h, and so L_d is equivalent to a self-adjoint discrete Lagrangian. □

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

2.4.1. Exact discrete Lagrangian is self-adjoint

It is easy to verify that the exact discrete Lagrangian (1.6.1) is self-adjoint. This can be done either directly from the definition (2.4.2), or by realizing that the exact flow map F_H generated by L_d^E satisfies (2.4.1), and then using Theorem 2.4.1.

2.4.2. Order of adjoint methods

To relate the expansions of L_d and its adjoint in terms of h, it is necessary to work with the modified form

$$L_d^*\big(q(-h/2), q(h/2), h\big) = -L_d\big(q(h/2), q(-h/2), -h\big),$$

which can be used in the same way as $L_d^*(q(0), q(h), h) = -L_d(q(h), q(0), -h)$. From this it is clear that the expansion of L_d^* is the negative of the expansion of L_d with h replaced by $-h$. In other words, if L_d has expansion

$$L_d(h) = hL_d^{(1)} + \frac{1}{2}h^2L_d^{(2)} + \frac{1}{6}h^3L_d^{(3)} + \cdots$$

then L_d^* will have expansion

$$\begin{aligned} L_d^*(h) &= -(-h)L_d^{(1)} - \frac{1}{2}(-h)^2L_d^{(2)} - \frac{1}{6}(-h)^3L_d^{(3)} - \cdots \\ &= hL_d^{(1)} - \frac{1}{2}h^2L_d^{(2)} + \frac{1}{6}h^3L_d^{(3)} - \cdots \end{aligned}$$

and so the series agree on odd terms and are opposite on even terms.

This shows that the order of the adjoint discrete Lagrangian L_d^* is the same as the order of L_d. Furthermore, if L_d is self-adjoint, then all the even terms in its expansion must be zero, showing that self-adjoint discrete Lagrangians are necessarily of even order (the first nonzero term, which is $r+1$, must be odd).

These same conclusions can be also be reached by working with the discrete Hamiltonian map, and showing that its adjoint has the same order as it, and that it is of even order whenever it is self-adjoint. Theorems 2.4.1 and 2.3.1 then give the corresponding statements for the discrete Lagrangians.

Example 2.4.2. Perhaps the simplest example of adjoint discrete Lagrangians is the pair

$$\begin{aligned} L_d(q_0, q_1, h) &= hL\left(q_0, \frac{q_1 - q_0}{h}\right), \\ L_d^*(q_0, q_1, h) &= hL\left(q_1, \frac{q_1 - q_0}{h}\right), \end{aligned}$$

which clearly satisfy (2.4.2). For a Lagrangian of the form $L = \frac{1}{2}\dot{q}^T M\dot{q} -$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

$V(q)$ these two discrete Lagrangians produce the methods $\tilde{F}_{L_d}$ and $\tilde{F}_{L_d^*}$ given by

$$\tilde{F}_{L_d} \begin{cases} q_1 = q_0 + hM^{-1}p_1, \\ p_1 = p_0 - h\nabla V(q_0), \end{cases}$$

$$\tilde{F}_{L_d^*} \begin{cases} q_1 = q_0 + hM^{-1}p_0, \\ p_1 = p_0 - h\nabla V(q_1). \end{cases}$$

In the terminology of Hairer (1998) these are the two types of symplectic Euler. We can now explicitly compute:

$$\begin{aligned} (\tilde{F}_{L_d^*})^h \circ (\tilde{F}_{L_d})^{(-h)}(q_0, p_0) &= \tilde{F}_{L_d^*}(q_0 + hM^{-1}p_1, p_0 - h\nabla V(q_0), h) \\ &= (q_0, p_0), \end{aligned}$$

which shows that $\tilde{F}_{L_d}$ and $\tilde{F}_{L_d^*}$ are indeed mutually adjoint. ◊

Example 2.4.3. The discrete Lagrangians in the previous example are just L_d^α for $\alpha = 0$ and $\alpha = 1$, respectively. Extending this gives $(L_d^\alpha)^* = L_d^{1-\alpha}$, which shows that the midpoint rule (given by $\alpha = 1/2$) is self-adjoint. From this it is also clear that the symmetrized versions $L_d^{\text{sym},\alpha}$ are self-adjoint for all α. ◊

2.5. Composition methods

We now consider how to combine several discrete Lagrangians together to obtain a new discrete Lagrangian with higher order, or some other desirable property. The resulting discrete Hamiltonian map will be the composition of the maps of the component discrete Lagrangians. References on composition methods include Yoshida (1990), Qin and Zhu (1992), McLachlan (1995) and Murua and Sanz-Serna (1999).

The strength of the composition methodology can be illustrated by a few simple examples. Given a one-step method $F : T^*Q \times \mathbb{R} \to T^*Q$ with corresponding adjoint F^*, then the method $\hat{F}^h = F^{h/2} \circ (F^*)^{h/2}$ will be self-adjoint and have order at least equal to that of F. Furthermore, for a self-adjoint method F with order r, which we recall must be even, the method $\hat{F}^h = F^{\gamma h} \circ F^{(1-2\gamma)h} \circ F^{\gamma h}$ with the constant $\gamma = (2 - 2^{1/(r+1)})^{-1}$ will have order $r + 2$. This thus provides a simple way to derive methods of arbitrarily high order starting from a given low-order method. See the above references for details and more complicated examples.

Consider now discrete Lagrangians L_d^i and time-step fractions γ^i for $i = 1, \ldots, s$ satisfying $\sum_{i=1}^s \gamma^i = 1$. Note that the γ^i may each be positive or negative. We now give three equivalent interpretations of composition discrete Lagrangians.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

2.5.1. Multiple steps

Begin by taking a discrete trajectory $\{q_k\}_{k=0}^N$, dividing each step (q_k, q_{k+1}) into s substeps $(q_k = q_k^0, q_k^1, q_k^2, \ldots, q_k^s = q_{k+1})$. Rather than using the same discrete Lagrangian for each step, as we have previously always assumed, we will now use the different L_d^i on each substep in turn.

This is equivalent to taking the discrete action sum to be

$$\mathfrak{G}_d\big(\{(q_k = q_k^0, \ldots, q_k^s = q_{k+1})\}_{k=1}^N\big) = \sum_{k=0}^{N}\sum_{i=1}^{s} L_d^i(q_k^{i-1}, q_k^i, \gamma^i h). \tag{2.5.1}$$

The discrete Euler–Lagrange equations, resulting from requiring that this be stationary, pair neighbouring discrete Lagrangians together to give

$$D_2 L_d^i(q_k^{i-1}, q_k^i, \gamma^i h) + D_1 L_d^{i+1}(q_k^i, q_k^{i+1}, \gamma^{i+1} h) = 0, \qquad i = 1, \ldots, s-1, \tag{2.5.2a}$$

$$D_2 L_d^s(q_k^{s-1}, q_k^s, \gamma^s h) + D_1 L_d^1(q_{k+1}^0, q_{k+1}^1, \gamma^1 h) = 0, \tag{2.5.2b}$$

where the steps are joined with $q_k^s = q_{k+1}^0$.

Considering the L_d^i as generating functions for the discrete Hamiltonian maps $\tilde{F}_{L_d^i}$ shows that this is simply taking a step with $\tilde{F}_{L_d^1}$ of length $\gamma^1 h$, followed by a step with $\tilde{F}_{L_d^2}$ of length $\gamma^2 h$, and so on. The map over the entire time-step is thus the composition of the maps

$$\tilde{F}_{L_d^s}^{\gamma^s h} \circ \cdots \circ \tilde{F}_{L_d^2}^{\gamma^2 h} \circ \tilde{F}_{L_d^1}^{\gamma^1 h}.$$

2.5.2. Single step, multiple substeps

We now combine the discrete Lagrangians on each step into one multipoint discrete Lagrangian defined by

$$L_d(q_k^0, q_k^1, \ldots, q_k^s, h) = \sum_{i=1}^{s} L_d^i(q_k^{i-1}, q_k^i, \gamma^i h), \tag{2.5.3}$$

and we define the discrete action sum over the entire trajectory to be

$$\mathfrak{G}_d\big(\{(q_k = q_k^0, \ldots, q_k^s = q_{k+1})\}_{k=1}^N\big) = \sum_{k=0}^{N} L_d(q_k^0, q_k^1, \ldots, q_k^s, h), \tag{2.5.4}$$

which is clearly equal to (2.5.1).

Requiring that $\mathfrak{G}_d$ be stationary gives the extended set of discrete Euler–Lagrange equations

$$D_i L_d(q_k^0, q_k^1, \ldots, q_k^s, h) = 0 \qquad i = 2, \ldots, s \tag{2.5.5a}$$

$$D_{s+1} L_d(q_k^0, q_k^1, \ldots, q_k^s, h) + D_1 L_d(q_{k+1}^0, q_{k+1}^1, \ldots, q_{k+1}^s, h) = 0, \tag{2.5.5b}$$

which are equivalent to (2.5.2a) and (2.5.2b), respectively.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

2.5.3. Single step

Finally, we form a standard discrete Lagrangian which is equivalent to the above methods. Set the *composition discrete Lagrangian* to be

$$L_d(q_k, q_{k+1}, h) = \operatorname*{ext}_{(q_k^1,\ldots,q_k^{s-1})} L_d(q_k = q_k^0, q_k^1, q_k^2, \ldots, q_k^{s-1}, q_k^s = q_{k+1}, h) \tag{2.5.6}$$

which is the multipoint discrete Lagrangian defined above, evaluated on the trajectory within the step which solves (2.5.5a).

Note that the derivatives of this discrete Lagrangian satisfy

$$\begin{aligned} D_1 L_d(q_k, q_{k+1}, h) &= D_1 L_d(q_k, q_k^1, q_k^2, \ldots, q_k^{s-1}, q_{k+1}, h) \\ &\quad + \sum_{i=1}^{s-1} D_i L_d(q_k, q_k^1, q_k^2, \ldots, q_k^{s-1}, q_{k+1}, h) \cdot \frac{\partial q_k^i}{\partial q_k} \\ &= D_1 L_d(q_k, q_k^1, q_k^2, \ldots, q_k^{s-1}, q_{k+1}, h) \\ &= D_1 L_d^1(q_k, q_k^1, \gamma^1 h) \end{aligned}$$

using (2.5.5a), and similarly

$$\begin{aligned} D_2 L_d(q_k, q_{k+1}, h) &= D_{s+1} L_d(q_k, q_k^1, q_k^2, \ldots, q_k^{s-1}, q_{k+1}, h) \\ &= D_2 L_d^s(q_k^{s-1}, q_{k+1}, \gamma^s h). \end{aligned}$$

This gives the following theorem.

Theorem 2.5.1. Take discrete Lagrangians L_d^i and time-step fractions γ^i for $i = 1, \ldots, s$ satisfying $\sum_{i=1}^s \gamma^i = 1$. Define the composition discrete Lagrangian L_d by (2.5.6). Then the discrete Hamiltonian map $\tilde{F}_{L_d}$ is

$$\tilde{F}_{L_d}^h = \tilde{F}_{L_d^s}^{\gamma^s h} \circ \cdots \circ \tilde{F}_{L_d^2}^{\gamma^2 h} \circ \tilde{F}_{L_d^1}^{\gamma^1 h}$$

formed by the composition of the discrete Hamiltonian maps for each L_d^i.

Proof. The equations that define $\tilde{F}_{L_d}$ are

$$\begin{aligned} p_k &= -D_1 L_d(q_k, q_{k+1}, h) = -D_1 L_d^1(q_k, q_k^1, \gamma^1 h), \\ p_{k+1} &= D_2 L_d(q_k, q_{k+1}, h) = D_2 L_d^s(q_k^{s-1}, q_{k+1}, \gamma^s h), \end{aligned}$$

together with (2.5.5a), which is equivalent to (2.5.2a), which we write as

$$p_k^i = D_2 L_d^i(q_k^{i-1}, q_k^i, \gamma^i h) = -D_1 L_d^{i+1}(q_k^i, q_k^{i+1}, \gamma^{i+1} h)$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

for $i = 1, \ldots, s-1$. Setting $p_k^0 = p_k$ and $p_k^s = p_{k+1}$, we can group these to give

$$p_k^{i-1} = -D_1 L_d^i(q_k^{i-1}, q_k^i, \gamma^i h),$$
$$p_k^i = D_2 L_d^i(q_k^{i-1}, q_k^i, \gamma^i h),$$

for $i = 1, \ldots, s$, which are the definition of $\tilde{F}_{L_d^s}^{\gamma^s h} \circ \cdots \circ \tilde{F}_{L_d^2}^{\gamma^2 h} \circ \tilde{F}_{L_d^1}^{\gamma^1 h}$, thus giving the required equivalence. □

2.6. Examples of variational integrators

In this section we will consider a number of standard symplectic methods and show how to write them as variational integrators. Recall that we are assuming that Q is a linear space with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\| \cdot \|$. We will always assume that the Lagrangian $L : TQ \to \mathbb{R}$ is regular, so that it has a corresponding Hamiltonian $H : T^*Q \to \mathbb{R}$. In addition, we will sometimes consider the Lagrangian composed of a kinetic and a potential energy, so that it is of the form $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q)$ where M is a positive-definite symmetric matrix.

2.6.1. Midpoint rule

Given a Hamiltonian system $H : T^*Q \to \mathbb{R}$, the *midpoint rule* is an integrator $F^h : (q_0, p_0) \mapsto (q_1, p_1)$. Setting $z_0 = (q_0, p_0)$ and $z_1 = (q_1, p_1)$ the map is defined implicitly by the relation

$$\frac{z_1 - z_0}{h} = X_H\left(\frac{z_1 + z_0}{2}\right),$$

where X_H is the Hamiltonian vector field. Writing the two components separately gives

$$\frac{q_1 - q_0}{h} = \frac{\partial H}{\partial p}\left(\frac{q_1 + q_0}{2}, \frac{p_1 + p_0}{2}\right), \tag{2.6.1a}$$
$$\frac{p_1 - p_0}{h} = -\frac{\partial H}{\partial p}\left(\frac{q_1 + q_0}{2}, \frac{p_1 + p_0}{2}\right). \tag{2.6.1b}$$

The symplectic nature of the midpoint rule is often explained by using the Cayley transform (this remark is due, as far as we know, to Krishnaprasad and J. C. Simo; see, for example, Austin, Krishnaprasad and Wang (1993), Simo, Tarnow and Wong (1992) and Simo and Tarnow (1992), and related papers). See Marsden (1999) for an exposition of this method.

To write the midpoint rule as a variational integrator, assume that H is regular and that L is the corresponding regular Lagrangian defined by (1.4.4). Define the discrete Lagrangian

$$L_d^{\frac{1}{2}}(q_0, q_1, h) = hL\left(\frac{q_1 + q_0}{2}, \frac{q_1 - q_0}{h}\right).$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Evaluating the expressions (1.5.4) for the discrete Hamiltonian map gives

$$p_0 = -\frac{h}{2}\frac{\partial L}{\partial q}\left(\frac{q_1+q_0}{2}, \frac{q_1-q_0}{h}\right) + \frac{\partial L}{\partial \dot{q}}\left(\frac{q_1+q_0}{2}, \frac{q_1-q_0}{h}\right),$$

$$p_1 = \frac{h}{2}\frac{\partial L}{\partial q}\left(\frac{q_1+q_0}{2}, \frac{q_1-q_0}{h}\right) + \frac{\partial L}{\partial \dot{q}}\left(\frac{q_1+q_0}{2}, \frac{q_1-q_0}{h}\right),$$

and subtracting and adding these two equations produces

$$\frac{p_1-p_0}{h} = \frac{\partial L}{\partial q}\left(\frac{q_1+q_0}{2}, \frac{q_1-q_0}{h}\right), \tag{2.6.2}$$

$$\frac{p_1+p_0}{2} = \frac{\partial L}{\partial \dot{q}}\left(\frac{q_1+q_0}{2}, \frac{q_1-q_0}{h}\right).$$

The second of these equations is simply the statement that

$$\left(\frac{q_1+q_0}{2}, \frac{p_1+p_0}{2}\right) = \mathbb{F}L\left(\frac{q_1+q_0}{2}, \frac{q_1-q_0}{h}\right),$$

and so using (1.4.5a) shows that (2.6.2) is equivalent to (2.6.1b), while (1.4.5c) gives (2.6.1a).

For regular Lagrangian systems, the midpoint discrete Lagrangian $L_d^{1/2}$ thus has discrete Hamiltonian map which is the midpoint rule on T^*Q for the corresponding Hamiltonian system.

2.6.2. Störmer–Verlet

The Verlet method (Verlet 1967) (also known as Störmer's rule) was originally formulated for molecular dynamics problems and remains popular in that field. The derivation of Verlet as a variational integrator is in Wendlandt and Marsden (1997*a*) and is implicitly in Gillilan and Wilson (1992) as well.

Verlet is usually written for systems of the form $L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$, and was originally formulated as a map $Q\times Q \to Q\times Q$ given by $(q_k, q_{k+1}) \mapsto (q_{k+1}, q_{k+2})$ with

$$q_{k+1} = 2q_k - q_{k-1} + h^2 a_k,$$

where we use the notation $a_k = M^{-1}(-\nabla V(q_k))$. As can be readily seen, this is just the discrete Lagrangian map $F_{L_d} : Q \times Q \to Q \times Q$ for either of

$$L_d^0(q_0,q_1,h) = hL\left(q_0, \frac{q_1-q_0}{h}\right),$$

$$L_d^1(q_0,q_1,h) = hL\left(q_1, \frac{q_1-q_0}{h}\right),$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

or indeed any affine combination of these two. In particular, consider the symmetric version

$$L_d(q_0, q_1, h) = \frac{1}{2}hL\left(q_0, \frac{q_1 - q_0}{h}\right) + \frac{1}{2}hL\left(q_1, \frac{q_1 - q_0}{h}\right),$$

which gives Verlet as the corresponding F_{L_d}. Pushing this forward to T^*Q with $\mathbb{F}^{\pm}L_d$ now gives $\tilde{F}_{L_d} : T^*Q \to T^*Q$ defined by (1.5.4). Evaluating these yields

$$p_k = M\left(\frac{q_{k+1} - q_k}{h}\right) + \frac{1}{2}h\nabla V(q_k),$$
$$p_{k+1} = M\left(\frac{q_{k+1} - q_k}{h}\right) - \frac{1}{2}h\nabla V(q_{k+1}).$$

Now we subtract the first equation from the second and solve the first equation for q_{k+1} to obtain

$$q_{k+1} = q_k + hM^{-1}p_k + \frac{1}{2}h^2M^{-1}(-\nabla V(q_k)),$$
$$p_{k+1} = p_k + h\left(\frac{-\nabla V(q_k) - \nabla V(q_{k+1})}{2}\right),$$

which is the so-called *velocity Verlet* method (Swope, Andersen, Berens and Wilson 1982, Allen and Tildesley 1987) written on T^*Q. Using the Legendre transform $\mathbb{F}L(q, \dot{q}) = (q, M\dot{q})$ this can also be mapped to TQ.

We thus see that velocity Verlet will preserve the canonical two-form Ω on T^*Q, and as L_d is invariant under linear symmetries of the potential, Verlet will also preserve quadratic momentum maps such as linear and angular momentum.

2.6.3. Newmark methods

The Newmark family of integrators, originally given in Newmark (1959), are widely used in structural dynamics codes. They are usually written (see, for example, Hughes (1987)) for the system $L = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$ as maps $TQ \to TQ$ given by $(q_k, \dot{q}_k) \mapsto (q_{k+1}, \dot{q}_{k+1})$ satisfying the implicit relations

$$q_{k+1} = q_k + h\dot{q}_k + \frac{h^2}{2}\left[(1 - 2\beta)a(q_k) + 2\beta a(q_{k+1})\right], \tag{2.6.3a}$$
$$\dot{q}_{k+1} = \dot{q}_k + h\left[(1 - \gamma)a(q_k) + \gamma a(q_{k+1})\right], \tag{2.6.3b}$$
$$a(q) = M^{-1}(-\nabla V(q)), \tag{2.6.3c}$$

where the parameters $\gamma \in [0, 1]$ and $\beta \in [0, \frac{1}{2}]$ specify the method. It is simple to check that the method is second-order if $\gamma = 1/2$ and first-order otherwise, and that it is generally explicit only for $\beta = 0$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

The $\beta = 0$, $\gamma = 1/2$ case is well known to be symplectic (see, for example, Simo *et al.* (1992)) with respect to the canonical symplectic form Ω_L on TQ. This can be easily seen from the fact that this method is simply the pullback by $\mathbb{F}L$ of the discrete Hamiltonian map for $L_d^{\mathrm{sym},\alpha}$ with $\alpha = 0$ or $\alpha = 1$. Note that this method is the same as velocity Verlet.

It is also well known (for example, Simo *et al.* (1992)) that the Newmark algorithm with $\beta \neq 0$ does not preserve the *canonical* symplectic form. Nonetheless, based on a remark by Suris, it can be shown (Kane *et al.* 2000) that the Newmark method with $\gamma = 1/2$ and any β can be generated from a discrete Lagrangian. To see this, we introduce the map $\varphi^\beta : Q \times Q \to TQ$ defined by

$$\varphi^\beta(q_k, q_{k+1}) = \left(q_k, \left[\frac{q_{k+1} - q_k}{h}\right] - \frac{h}{2}\left[(1-2\beta)a(q_k) + 2\beta a(q_{k+1})\right]\right).$$

Pulling the Newmark map back by φ^β to a map $Q \times Q \to Q \times Q$ gives the map $(q_k, q_{k+1}) \mapsto (q_{k+1}, q_{k+2})$ where

$$\frac{q_{k+2} - 2q_{k+1} + q_k}{h^2} = \beta a(q_{k+2}) + (1-2\beta)a(q_{k+1}) + \beta a(q_k). \tag{2.6.4}$$

A straightforward calculation now shows that this is in fact the discrete Lagrange map $F_{L_d^\beta}$ for the discrete Lagrangian

$$L_d^\beta(q_0, q_1, h) = h\frac{1}{2}\left(\frac{\eta^\beta(q_1) - \eta^\beta(q_0)}{h}\right)^T M \left(\frac{\eta^\beta(q_1) - \eta^\beta(q_0)}{h}\right) - h\tilde{V}(\eta^\beta(q_0)),$$

where we have introduced the map $\eta^\beta : Q \to Q$ defined by

$$\eta^\beta(q) = q - \beta h^2 M^{-1} \nabla V(q)$$

and the modified potential function $\tilde{V} : Q \to \mathbb{R}$ satisfying $\nabla\tilde{V} \circ \eta^\beta = \nabla V$, which will exist for small h.

This result shows that the Newmark method for $\gamma = 1/2$ is the pullback of the discrete Hamiltonian map $\tilde{F}_{L_d^\beta}$ by the map $\mathbb{F}^+ L_d^\beta \circ (\varphi^\beta)^{-1}$. As the discrete Hamiltonian map preserves the canonical symplectic form on T^*Q, this means that Newmark preserves the two-form $[\mathbb{F}^+ L_d^\beta \circ (\varphi^\beta)^{-1}]^* \Omega$ on TQ. Note that this is not the canonical two-form Ω_L on TQ, but this is enough to explain the otherwise inexplicably good long-time behaviour of $\gamma = 1/2$ Newmark for nonlinear problems.

An alternative and independent method of analysing the symplectic members of Newmark has been given by Skeel, Zhang and Schlick (1997), including an interesting nonlinear analysis in Skeel and Srinivas (2000). This is based on the observation that if we define the map $\bar{\eta}^\beta : TQ \to TQ$ by

$$\bar{\eta}^\beta(q, v) = \left(\eta^\beta(q), v\right)$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

then the pushforward of the Newmark method by $\bar{\eta}^\beta$ is given by $(x_k, v_k) \mapsto (x_{k+1}, v_{k+1})$, where

$$x_{k+1} = x_k + hv_k + \frac{1}{2}h^2 a_k, \tag{2.6.5a}$$

$$v_{k+1} = v_k + \frac{1}{2}h(a_k + a_{k+1}), \tag{2.6.5b}$$

$$a_k = a(x_k + \beta h^2 a_k). \tag{2.6.5c}$$

This map can be shown to be symplectic with respect to the canonical two-form Ω_L on TQ, and so Newmark will preserve the two-form $(\bar{\eta}^\beta)^*\Omega_L$ on TQ.

To summarize, we have the following commutative diagram, where the map $\tilde{F}_{L_d^\beta}$ preserves the canonical two-form Ω on T^*Q, the map (2.6.5) preserves the Lagrange two-form Ω_L on TQ, and we have set $\gamma = 1/2$ in the Newmark equation (2.6.3).

$$\begin{array}{ccccccc}
T^*Q & \xleftarrow{\mathbb{F}^+L_d^\beta} & Q \times Q & \xrightarrow{\varphi^\beta} & TQ & \xrightarrow{\bar{\eta}^\beta} & TQ \\
\Big\downarrow{\scriptstyle \tilde{F}_{L_d^\beta}} & & \Big\downarrow{\scriptstyle F_{L_d^\beta}}\ (2.6.4) & & \Big\downarrow\ (2.6.3) & & \Big\downarrow\ (2.6.5) \\
T^*Q & \xleftarrow[\mathbb{F}^+L_d^\beta]{} & Q \times Q & \xrightarrow[\varphi^\beta]{} & TQ & \xrightarrow[\bar{\eta}^\beta]{} & TQ
\end{array}$$

2.6.4. Explicit symplectic partitioned Runge–Kutta methods

Symplectic integrators which are explicit partitioned Runge–Kutta methods were first used by Ruth (1983) and Forest and Ruth (1990), who constructed them as a composition of steps, each one generated by a generating function of the third kind. Using the same idea shows that these methods are also variational, at least for Hamiltonians with kinetic energy of the form $T(p) = 1/2p^T M^{-1} p$ for some constant mass matrix M.

It can be shown (Hairer *et al.* 1993) that explicit symplectic partitioned Runge–Kutta methods can always be written as the composition of a number of steps of the method $F^{a,b} : T^*Q \times \mathbb{R} \to T^*Q$ given by

$$q_1 = q_0 + ahM^{-1}p_0,$$
$$p_1 = p_0 - bh\nabla V(q_1),$$

and of its adjoint method $(F^{a,b})^*$, with each step having different parameters (a, b). Furthermore, it is simple to check that these can be chosen so that all steps have nonzero a.

We now see, however, that the method $F^{a,b}$ is the discrete Hamiltonian

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

map for the discrete Lagrangian $L_d^{a,b}$ given by

$$L_d^{a,b}(q_0, q_1, h) = h\left[b\frac{1}{2}\left(\frac{q_1 - q_0}{h}\right)^T M\left(\frac{q_1 - q_0}{h}\right) - \frac{1}{a}V(q_1)\right],$$

and from Theorem 2.4.1 it is clear that $(F^{a,b})^*$ is the discrete Hamiltonian map of the adjoint discrete Lagrangian $(L_d^{a,b})^*$.

We can thus form a composition discrete Lagrangian as in Theorem 2.5.1 whose discrete Hamiltonian map is the composition of the $F^{a,b}$ and $(F^{a,b})^*$, and is therefore the explicit symplectic partitioned Runge–Kutta method.

2.6.5. Symplectic partitioned Runge–Kutta methods

Partitioned Runge–Kutta methods are a class of integrators about which much is known and which generalize standard Runge–Kutta methods. The symplectic members of Runge–Kutta were first identified by Lasagni (1988), Sanz-Serna (1988) and Suris (1989). Symplectic partitioned Runge–Kutta methods appeared in Sanz-Serna (1992*a*) and Suris (1990). Good general references are Hairer *et al.* (1993) and Hairer and Wanner (1996). See also Geng (1995, 2000), Sofroniou and Oevel (1997) and Oevel and Sofroniou (1997) for order conditions and derivations. An explicit construction has been given by Suris (1990) for the discrete Lagrangian which generates any symplectic partitioned Runge–Kutta method. We summarize this derivation below.

Recall that a partitioned Runge–Kutta method for the regular Lagrangian system L is a map $T^*Q \times \mathbb{R} \to T^*Q$ specified by the coefficients b_i, a_{ij}, $\tilde{b}_i$, $\tilde{a}_{ij}$ for $i, j = 1, \ldots, s$, and defined by $(q_0, p_0) \mapsto (q_1, p_1)$ for

$$q_1 = q_0 + h\sum_{j=1}^{s} b_j \dot{Q}_j, \qquad p_1 = p_0 + h\sum_{j=1}^{s} \tilde{b}_j \dot{P}_j, \tag{2.6.6a}$$

$$Q_i = q_0 + h\sum_{j=1}^{s} a_{ij} \dot{Q}_j, \qquad P_i = p_0 + h\sum_{j=1}^{s} \tilde{a}_{ij} \dot{P}_j, \qquad i = 1, \ldots, s, \tag{2.6.6b}$$

$$P_i = \frac{\partial L}{\partial \dot{q}}(Q_i, \dot{Q}_i), \qquad \dot{P}_i = \frac{\partial L}{\partial q}(Q_i, \dot{Q}_i), \qquad i = 1, \ldots, s, \tag{2.6.6c}$$

where the points (Q_i, P_i) are known as the internal stages. In the special case that $a_{ij} = \tilde{a}_{ij}$ and $b_i = \tilde{b}_i$ then a partitioned Runge–Kutta method is said to be simply a Runge–Kutta method.

It is well known that the method is symplectic (that is, it preserves the canonical symplectic form Ω on T^*Q) if the coefficients satisfy

$$b_i\tilde{a}_{ij} + \tilde{b}_j a_{ji} = b_i\tilde{b}_j, \qquad i, j = 1, \ldots, s, \tag{2.6.7a}$$

$$b_i = \tilde{b}_i, \qquad i = 1, \ldots, s. \tag{2.6.7b}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

We now assume that we have coefficients satisfying (2.6.7) and write a discrete Lagrangian that generates the corresponding symplectic partitioned Runge–Kutta method. Given points $(q_0, q_1) \in Q \times Q$, we can regard (2.6.6) as implicitly defining p_0, p_1, Q_i, P_i, $\dot{Q}_i$ and $\dot{P}_i$ for $i = 1, \ldots, s$. Taking these to be so defined as functions of (q_0, q_1), we construct a discrete Lagrangian

$$L_d(q_0, q_1, h) = h \sum_{i=1}^{s} b_i L(Q_i, \dot{Q}_i). \tag{2.6.8}$$

It can now be shown (Suris 1990) that the corresponding discrete Hamiltonian map is exactly the map $(q_0, p_0) \mapsto (q_1, p_1)$, which is the symplectic partitioned Runge–Kutta method. Nonsymplectic partitioned Runge–Kutta methods will clearly not have a corresponding discrete Lagrangian formulation.

Theorem 2.6.1. The discrete Hamiltonian map generated by the discrete Lagrangian (2.6.8) is a symplectic partitioned Runge–Kutta method.

Proof. To check that the discrete Hamiltonian map defined by L_d is indeed the partitioned Runge–Kutta method specified by (2.6.6), we need only check that equations (1.5.4) are satisfied. We compute

$$\begin{aligned}\frac{\partial L_d}{\partial q_0}(q_0, q_1) &= (\Delta t) \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q} \cdot \frac{\partial Q_i}{\partial q_0} + \frac{\partial L}{\partial \dot{q}} \cdot \frac{\partial \dot{Q}_i}{\partial q_0} \right] \\ &= (\Delta t) \sum_{i=1}^{s} b_i \left[\dot{P}_i \cdot \frac{\partial Q_i}{\partial q_0} + P_i \cdot \frac{\partial \dot{Q}_i}{\partial q_0} \right],\end{aligned}$$

using the definitions for P_i and $\dot{P}_i$ in (2.6.6). Differentiating the definition for Q_i in (2.6.6b) and substituting in this and the definition of P_i in (2.6.6b) now gives

$$\begin{aligned}\frac{\partial L_d}{\partial q_0}(q_0, q_1) = (\Delta t) \sum_{i=1}^{s} b_i \Bigg[& \dot{P}_i \cdot \left(I + (\Delta t) \sum_{j=1}^{s} a_{ij} \frac{\partial \dot{Q}_j}{\partial q_0} \right) \\ & + \left(p_0 + (\Delta t) \sum_{j=1}^{s} \tilde{a}_{ij} \dot{P}_j \right) \cdot \frac{\partial \dot{Q}_i}{\partial q_0} \Bigg] \\ = (\Delta t) \sum_{i=1}^{s} b_i \Bigg[& \dot{P}_i + p_0 \cdot \frac{\partial \dot{Q}_i}{\partial q_0} \Bigg] \\ & + (\Delta t)^2 \sum_{i=1}^{s} \sum_{j=1}^{s} (b_i \tilde{a}_{ij} + b_j a_{ji}) \dot{P}_j \cdot \frac{\partial \dot{Q}_i}{\partial q_0},\end{aligned}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

and we can use the symplecticity identities (2.6.7) to obtain

$$\begin{aligned}\frac{\partial L_d}{\partial q_0}(q_0,q_1) = p_0 \cdot \left[(\Delta t)\sum_{i=1}^{s} b_i \frac{\partial \dot{Q}_i}{\partial q_0}\right] + (\Delta t)\sum_{i=1}^{s} b_i \dot{P}_i \\ + (\Delta t)\sum_{j=1}^{s} b_j \dot{P}_j \cdot \left[(\Delta t)\sum_{i=1}^{s} b_i \frac{\partial \dot{Q}_i}{\partial q_0}\right] \\ = -p_0,\end{aligned}$$

where we have differentiated the expression for q_1 in (2.6.6a) to obtain the identity

$$(\Delta t)\sum_{i=1}^{s} b_i \frac{\partial \dot{Q}_i}{\partial q_0} = -I.$$

This thus establishes that the first equation of (1.5.4) is satisfied.

Differentiating L_d with respect to q_1 and following a similar argument to that above gives the second part of (1.5.4), and shows that the discrete Hamiltonian map $\tilde{F}L_d$ generated by the discrete Lagrangian (2.6.8) is indeed the symplectic partitioned Runge–Kutta method. □

This construction thus provides a proof of the well-known fact that the restrictions (2.6.7) on the coefficients mean that the partitioned Runge–Kutta method is symplectic, as discrete Hamiltonian maps always preserve the canonical symplectic form. In addition, the linear nature of the definition of the discrete Lagrangian (2.6.8) means that it will inherit linear symmetries of the Lagrangian L, which thus proves the standard result that partitioned Runge–Kutta methods preserve quadratic momentum maps.

Another way to regard the above derivation is to say that we have written down a generating function of the first kind for the symplectic partitioned Runge–Kutta map. A generalization of this is given in Jalnapurkar and Marsden (200x), where it is shown how to construct generating functions of arbitrary type for any given symplectic partitioned Runge–Kutta method.

The above construction has also been generalized to multisymplectic partial differential equations in West (2001), thereby obtaining multisymplectic product partitioned Runge–Kutta methods (Reich 2000).

2.6.6. Galerkin methods

To obtain accurate variational integrators, the results in Section 2.3 show that the discrete Lagrangian should approximate the action over short trajectory segments. One way to do this practically is to use polynomial approximations to the trajectories and numerical quadrature to approximate the integral. This can be shown to be equivalent both to a class of continuous

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Galerkin methods and to a subset of symplectic partitioned Runge–Kutta methods.

This approach is related to the Continuous Galerkin and Discontinuous Galerkin methods, as in Estep and French (1994), Hulme (1972*a*, 1972*b*) and Thomée (1997). These methods differ in the precise choice of function space (continuous or discontinuous) and whether the position and velocity components are projected separately or the velocity projection is given by the lift of a position projection.

We know that a discrete Lagrangian should be an approximation

$$L_d(q_0, q_1, h) \approx \operatorname*{ext}_{q \in \mathcal{C}([0,h],Q)} \mathfrak{G}(q),$$

where $\mathcal{C}([0,h],Q)$ is the space of trajectories $q : [0,h] \to Q$ with $q(0) = q_0$ and $q(h) = q_1$, and $\mathfrak{G} : \mathcal{C}(0,h) \to \mathbb{R}$ is the action (1.2.1).

To approximate this quantity, we choose the particular finite-dimensional approximation $\mathcal{C}^s([0,h],Q) \subset \mathcal{C}([0,h],Q)$ of the trajectory space given by

$$\mathcal{C}^s([0,h],Q) = \{q \in \mathcal{C}([0,h],Q) \mid q \text{ is a polynomial of degree } s\},$$

and we approximate the action integral with numerical quadrature to give an approximate action $\mathfrak{G}^s : \mathcal{C}([0,h],Q) \to \mathbb{R}$ by

$$\mathfrak{G}^s(q) = h \sum_{i=1}^{s} b_i L\big(q(c_i h), \dot{q}(c_i h)\big), \tag{2.6.9}$$

where $c_i \in [0,1]$, $i = 1, \ldots, s$ are a set of quadrature points and b_i are the associated maximal order weights. We now set the Galerkin discrete Lagrangian to be

$$L_d(q_0, q_1, h) = \operatorname*{ext}_{q \in \mathcal{C}^s([0,h],Q)} \mathfrak{G}^s(q), \tag{2.6.10}$$

which can be practically evaluated. This procedure, of course, is simply performing Galerkin projection of the weak form of the ODE onto the space of piecewise polynomial trajectories. Furthermore, as we will show below, the resulting integrator is a symplectic partitioned Runge–Kutta method.

To make the above equations explicit, choose control times $0 = d_0 < d_1 < d_2 < \cdots < d_{s-1} < d_s = 1$ and control points $q_0^0 = q_0, q_0^1, q_0^2, \ldots, q_0^{s-1}, q_0^s = q_1$. These uniquely define the degree s polynomial $q_d(t; q_0^\nu, h)$ which passes through each q_0^ν at time $d_\nu h$, that is, $q_d(d_\nu h) = q_0^\nu$ for $\nu = 0, \ldots, s$. Letting $\tilde{l}_{\nu,s}(t)$ denote the Lagrange polynomials associated with the d_ν, we can express $q_d(t; q_0^\nu, h)$ as

$$q_d(\tau h; q_0^\nu, h) = \sum_{\nu=0}^{s} q_0^\nu \tilde{l}_{\nu,s}(\tau). \tag{2.6.11}$$

For $q_d(t; q_0^\nu, h)$ to be a critical point of the discrete action (2.6.9) we must

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

have stationarity with respect to variations in q_0^ν for $\nu = 1, \ldots, s-1$. Differentiating (2.6.9) and (2.6.11) implies that we have

$$0 = h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q}(c_i h) \tilde{l}_{\nu,s}(c_i) + \frac{1}{h} \frac{\partial L}{\partial \dot{q}}(c_i h) \dot{\tilde{l}}_{\nu,s}(c_i) \right] \tag{2.6.12}$$

for each $\nu = 1, \ldots, s-1$, where we denote $\frac{\partial L}{\partial q}(c_i h) = \frac{\partial L}{\partial q}(q_d(c_i h), \dot{q}_d(c_i h))$ and similarly for the other expressions.

The integration scheme $(q_0, p_0) \mapsto (q_1, p_1)$ generated by the Galerkin discrete Lagrangian (2.6.10) is now given implicitly by the relations

$$-p_0 = \frac{\partial L_d}{\partial q_0}(q_0, q_1, h), \qquad p_1 = \frac{\partial L_d}{\partial q_1}(q_0, q_1, h).$$

Evaluating these expressions and restating (2.6.12) gives the set of equations

$$E(0): -p_0 = h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q}(c_i h) \tilde{l}_{0,s}(c_i) + \frac{1}{h} \frac{\partial L}{\partial \dot{q}}(c_i h) \dot{\tilde{l}}_{0,s}(c_i) \right],$$

$$E(\nu): \quad 0 = h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q}(c_i h) \tilde{l}_{\nu,s}(c_i) + \frac{1}{h} \frac{\partial L}{\partial \dot{q}}(c_i h) \dot{\tilde{l}}_{\nu,s}(c_i) \right], \ \nu = 1, \ldots, s-1,$$

$$E(s): \quad p_1 = h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q}(c_i h) \tilde{l}_{s,s}(c_i) + \frac{1}{h} \frac{\partial L}{\partial \dot{q}}(c_i h) \dot{\tilde{l}}_{s,s}(c_i) \right],$$

which define the discrete Hamiltonian map $(q_0, p_0) \mapsto (q_1, p_1)$.

The above Galerkin discrete Lagrangian can also be interpreted as a function of several points, in a similar way to the composition discrete Lagrangians discussed in Section 2.5. Essentially, we choose a set of interior points which act as a parametrization of the space of degree s polynomials mapping $[0, h]$ to Q.

More precisely, we form the multipoint discrete Lagrangian

$$L_d(q_0^0, q_0^1, \ldots, q_0^s, h) = \mathfrak{G}^s\big(q_d(t; q_0^\nu, h)\big),$$

where we recall that $q_d(t; q_0^\nu, h)$ is the unique polynomial of degree s passing through q_0^ν at time $d_\nu h$ and $\mathfrak{G}^s$ is defined by (2.6.9). This multipoint discrete Lagrangian is the analogue of the discrete Lagrangian (2.5.3). The appropriate discrete action is then

$$\mathfrak{G}_d\big(\{(q_k = q_k^0, \ldots, q_k^s = q_{k+1})\}_{k=1}^N\big) = \sum_{k=0}^{N} L_d(q_k^0, q_k^1, \ldots, q_k^s, h),$$

and the corresponding discrete Euler–Lagrange equations are given by (2.5.5). Clearly, the discrete Lagrangian defined by extremizing the above multipoint L_d with respect to the interior points q_0^ν for $\nu = 1, \ldots, s$ is just the original Galerkin discrete Lagrangian (2.6.10), and the extended discrete

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Euler–Lagrange equations are thus equivalent to $E(\nu)$ above for $\nu = 0, \ldots, s$. This follows in the same way as the proof of Theorem 2.5.1.

We will now see that these Galerkin variational integrators can be realized as particular examples of Runge–Kutta or partitioned Runge–Kutta schemes.

Theorem 2.6.2. Take a set of quadrature points c_i with corresponding maximal order weights b_i and let L_d be the corresponding Galerkin discrete Lagrangian (2.6.10). Then the integrator generated by this discrete Lagrangian is the partitioned Runge–Kutta scheme defined by the coefficients

$$\begin{aligned} b_i &= \tilde{b}_i = \int_0^1 l_{i,s}(\rho)\mathbf{d}\rho, \\ a_{ij} &= \int_0^{c_i} l_{j,s}(\rho)\mathbf{d}\rho, \\ \tilde{a}_{ij} &= \tilde{b}_j \left(1 - \frac{a_{ji}}{b_i}\right), \end{aligned} \tag{2.6.13}$$

where the $l_{i,s}(\rho)$ are the Lagrange polynomials associated with the c_i.

Proof. Given (q_0, p_0), (q_1, p_1) and q_0^ν satisfying $E(\nu)$, $\nu = 0, \ldots, s$, we will show that they also satisfy the partitioned Runge–Kutta equations (2.6.6) written for a Lagrangian system with coefficients given by (2.6.13). We restate the defining equations here for reference:

$$q_1 = q_0 + h\sum_{j=1}^{s} b_j \dot{Q}_j, \qquad p_1 = p_0 + h\sum_{j=1}^{s} \tilde{b}_j \dot{P}_j, \tag{2.6.14a}$$

$$Q_i = q_0 + h\sum_{j=1}^{s} a_{ij} \dot{Q}_j, \quad P_i = p_0 + h\sum_{j=1}^{s} \tilde{a}_{ij} \dot{P}_j, \quad i = 1, \ldots, s, \tag{2.6.14b}$$

$$P_i = \frac{\partial L}{\partial \dot{q}}(Q_i, \dot{Q}_i), \qquad \dot{P}_i = \frac{\partial L}{\partial q}(Q_i, \dot{Q}_i), \qquad i = 1, \ldots, s. \tag{2.6.14c}$$

We will show that these equations are satisfied by the discrete Hamiltonian map.

Set $\dot{Q}_i = \dot{q}_d(c_i h; q_0^\nu, h)$ so that $\dot{q}_d(\tau h; q_0^\nu, h) = \sum_{j=1}^s \dot{Q}_j l_{j,s}(\tau)$. Integrating this expression and using the fact that $q_d(0; q_0^\nu, h) = q_0$ gives

$$q_d(\tau h; q_0^\nu, h) = q_0 + h\sum_{j=1}^{s} \dot{Q}_j \int_0^\tau l_j(\rho)\mathbf{d}\rho.$$

Setting $Q_i = q_d(c_i h; q_0^\nu, h)$ and using $q_1 = q_d(h; q_0^\nu, h)$ now gives the first parts of (2.6.14a) and (2.6.14b) for Q_i and q_1. Now define P_i and $\dot{P}_i$ according to (2.6.14c).

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Until this point we have not made use of the relations $E(\nu)$. We will now begin to do so by forming the sum of the $E(\nu)$, $\nu = 0, \ldots, s$. This gives

$$p_1 - p_0 = h \sum_{i=1}^{s} b_i \left[\frac{\partial L}{\partial q}(c_i h) \sum_{\nu=0}^{s} \tilde{l}_{\nu,s}(c_i) + \frac{1}{h} \frac{\partial L}{\partial \dot{q}}(c_i h) \sum_{\nu=0}^{s} \dot{\tilde{l}}_{\nu,s}(c_i) \right].$$

However, the Lagrange polynomials $\tilde{l}_{\nu,s}(\tau)$ sum to the identity function, and therefore the sum of their derivatives must be zero. We thus recover the second part of (2.6.14a) for p_1.

Note that the $\tilde{l}_{\nu,s+1}$ are a set of $s+1$ independent polynomials of degree s and thus are a basis for $\mathcal{P}_s$, the space of polynomials of degree s. For each $j = 1, \ldots, s$ the polynomial $l_{j,s}$ is of degree $s-1$ and so has an integral of degree s. This implies that there exist coefficients m_ν^j such that

$$\sum_{\nu=0}^{s} m_\nu^j \tilde{l}_{\nu,s+1}(\tau) = \int_0^\tau l_{j,s}(\rho) \mathbf{d}\rho - b_j.$$

Differentiating this expression with respect to τ and evaluating it at $\tau = 0$ and $\tau = 1$ gives the following three identities:

$$\sum_{\nu=0}^{s} m_\nu^j \dot{\tilde{l}}_{\nu,s+1}(\tau) = l_{j,s}(\tau),$$

$$m_s^j = \sum_{\nu=0}^{s} m_\nu^j \tilde{l}_{\nu,s+1}(1) = \int_0^1 l_{j,s}(\rho) \mathbf{d}\rho - b_j = 0,$$

$$m_0^j = \sum_{\nu=0}^{s} m_\nu^j \tilde{l}_{\nu,s+1}(0) = -b_j.$$

If we now form the sum $\sum_{\nu=0}^{s} m_\nu^j E(\nu)$ and make use of the above identities, we obtain

$$\begin{aligned} b_j p_0 &= h \sum_{i=1}^{s} b_i \left[\dot{P}_i \left(\int_0^{c_i} l_{j,s}(\rho) \mathbf{d}\rho - b_j \right) + \frac{1}{h} P_i l_{j,s}(c_i) \right] \\ &= h \sum_{i=1}^{s} \dot{P}_i \left[b_i (a_{ij} - b_j) \right] + b_j P_j, \end{aligned}$$

which can be rearranged to give the second part of (2.6.14b) for P_i. □

If the $\hat{a}_{ij}$ in Theorem 2.6.2 are equal to the a_{ij}, then the method is clearly the special case of a Runge–Kutta method, rather than the general partitioned Runge–Kutta case. Note that the definition of the $\hat{a}_{ij}$ in (2.6.13) is simply a rearrangement of the requirement (2.6.7a), and so the partitioned Runge–Kutta methods equivalent to the Galerkin variational integrators are naturally symplectic, as is clear from the symplectic nature of variational

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

integrators in general. In addition, the additive structure of the Galerkin discrete Lagrangian means that L_d will inherit linear symmetries of L, so Noether's theorem recovers the well-known fact that the partitioned Runge–Kutta methods will preserve quadratic momentum maps.

A particularly elegant symplectic Runge–Kutta method is the *collocation Gauss–Legendre rule*. In the present derivation this results simply from taking the quadrature points c_i to be those given by the Gauss–Legendre quadrature, which is the highest-order quadrature for a given number of points. The c_i produced in this manner are all strictly between 0 and 1.

If the system being integrated is stiff then better numerical performance results from having $c_s = 1$, making the integrator stiffly accurate (Hairer and Wanner 1996). If we also wish to preserve the symmetry of the discrete Lagrangian, then it is natural to seek the c_i giving the highest order quadrature rule while enforcing $c_0 = 0$ and $c_s = 1$. This is the so-called Lobatto quadrature, and the Galerkin variational integrator generated in this way is the *standard Lobatto IIIA–IIIB partitioned Runge–Kutta method*.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

PART THREE
Forcing and constraints

3.1. Background: Forced systems

Lagrangian and Hamiltonian systems with external forcing arise in many different contexts. Particular examples include control forces from actuators, dissipation and friction, and loading on mechanical systems. As we will see below, when integrating such systems it is important to take account of the geometric structure to avoid spurious numerical artifacts. One way to do this is by extending the discrete variational framework to include forcing, which we will now do.

3.1.1. Forced Lagrangian systems

A *Lagrangian force* is a fibre-preserving map $f_L : TQ \to T^*Q$ over the identity, which we write in coordinates as

$$f_L : (q, \dot{q}) \mapsto (q, f_L(q, \dot{q})).$$

Given such a force, it is standard to modify Hamilton's principle, seeking stationary points of the action, to the *Lagrange–d'Alembert principle*, which seeks curves $q \in \mathcal{C}(Q)$ satisfying

$$\delta \int_0^T L(q(t), \dot{q}(t))\,\mathrm{d}t + \int_0^T f_L(q(t), \dot{q}(t)) \cdot \delta q(t)\,\mathrm{d}t = 0, \tag{3.1.1}$$

where the δ represents variations vanishing at the endpoints. Using integration by parts shows that this is equivalent to the *forced Euler–Lagrange equations*, which have coordinate expression

$$\frac{\partial L}{\partial q}(q, \dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q})\right) + f_L(q, \dot{q}) = 0. \tag{3.1.2}$$

Note that these are the same as the standard Euler–Lagrange equations (1.2.5) with the forcing term added.

3.1.2. Forced Hamiltonian systems

A *Hamiltonian force* is a fibre-preserving map $f_H : T^*Q \to T^*Q$ over the identity. Given such a force, we define the corresponding horizontal one-form f'_H on T^*Q by

$$f'_H(p_q) \cdot u_{p_q} = \left\langle f_H(p_q), T\pi_Q \cdot u_{p_q} \right\rangle,$$

where $\pi_Q : T^*Q \to Q$ is the projection. This expression is reminiscent of the definition (1.4.1) of the canonical one-form Θ on T^*Q, and in coordinates it is $f'_H(q, p) \cdot (\delta q, \delta p) = f_H(q, p) \cdot \delta q$, so the one-form is clearly horizontal.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

The *forced Hamiltonian vector field* X_H is now defined to satisfy

$$\mathbf{i}_{X_H}\Omega = \mathbf{d}H - f'_H$$

and in coordinates this gives the well-known *forced Hamilton's equations*

$$X_q(q,p) = \frac{\partial H}{\partial q}(q,p), \tag{3.1.3a}$$

$$X_p(q,p) = -\frac{\partial H}{\partial p}(q,p) + f_H(q,p), \tag{3.1.3b}$$

which are the same as the standard Hamilton's equations (1.4.3) with the forcing term added to the momentum equation.

3.1.3. Legendre transform with forces

Given a Lagrangian L, we can take the standard Legendre transform $\mathbb{F}L : T^*Q \to TQ$ and relate Hamiltonian and Lagrangian forces by

$$f_L = f_H \circ \mathbb{F}L.$$

If we also have a Hamiltonian H related to L by the Legendre transform according to (1.4.4) then it can be shown that the forced Euler–Lagrange equations and the forced Hamilton's equations are equivalent. That is, if X_L and X_H are the forced Lagrangian and Hamiltonian vector fields, respectively, then $(\mathbb{F}L)^*(X_H) = X_L$.

3.1.4. Noether's theorem with forcing

We now consider the effect of forcing on the evolution of momentum maps that arise from symmetries of the Lagrangian $L : TQ \to \mathbb{R}$. Let $\Phi : G \times Q \to Q$ be a symmetry of L and let the Lagangian momentum map $J_L : TQ \to \mathfrak{g}^*$ be as defined in Section 1.2.4.

Evaluating the left-hand side of (3.1.1) for a variation of the form $\delta q(t) = \xi_Q(q(t))$ gives

$$\int_0^T \mathbf{d}L \cdot \xi_{TQ}\,\mathrm{d}t + \int_0^T f_L \cdot \xi_Q\,\mathrm{d}t = \int_0^T f_L \cdot \xi_Q\,\mathrm{d}t,$$

as L is assumed to be invariant. Using integration by parts as in the derivation of the forced Euler–Lagrange equations, we see that the above expression is equal to

$$\int_0^T \left[\frac{\partial L}{\partial q}(q,\dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}(q,\dot{q})\right) + f_L(q,\dot{q})\right] + \Theta_L \cdot \xi_{TQ}\big|_0^T$$
$$= (J_L \circ F_L^T)(q(0),\dot{q}(0)) \cdot \xi - J_L(q(0),\dot{q}(0)) \cdot \xi,$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

and so equating these two statements of (3.1.1) gives

$$\left[(J_L \circ F_L^T)(q(0), \dot{q}(0)) - J_L(q(0), \dot{q}(0))\right] \cdot \xi = \int_0^T f_L(q(t), \dot{q}(t)) \cdot \xi_Q(q(t))\, \mathrm{d}t. \tag{3.1.4}$$

This equation describes the evolution of the momentum map from time 0 to time T, and shows that forcing will generally alter the momentum map. In the special case that the forcing is orthogonal to the group action, the above derivation shows that Noether's theorem will still hold.

Theorem 3.1.1. (Forced Noether's theorem) Consider a Lagrangian system $L : TQ \to \mathbb{R}$ with forcing $f_L : TQ \to T^*Q$ and a symmetry action $\Phi : G \times Q \to Q$ such that $\langle f_L(q, \dot{q}), \xi_Q(q) \rangle = 0$ for all $(q, \dot{q}) \in TQ$ and all $\xi \in \mathfrak{g}$. Then the Lagrangian momentum map $J_L : TQ \to \mathfrak{g}^*$ will be preserved by the flow, so that $J_L \circ F_L^t = J_L$ for all t.

A similar result can also be derived for Hamiltonian systems, either by taking the Legendre transform of a regular forced Lagrangian system, or by working directly on the Hamiltonian side as in Section 1.4. For more details on the relationship between momentum maps and forcing see Bloch, Krishnaprasad, Marsden and Ratiu (1996*b*).

Note that, for nonzero forcing, the Lagrangian and Hamiltonian flows do not preserve the symplectic two-form. This can be seen by calculating $\mathbf{d}\hat{\mathfrak{G}}$ as was done in Section 1.2.3, and realizing that it contains a term with the integral of the force which does not vanish except when $f_L = 0$.

3.2. Discrete variational mechanics with forces

3.2.1. Discrete Lagrange–d'Alembert principle

As with other discrete structures, we take two *discrete Lagrangian forces* $f_d^+, f_d^- : Q \times Q \to T^*Q$, which are fibre-preserving in the sense that $\pi_Q \circ f_d^{\pm} = \pi_Q^{\pm}$, and which thus have coordinate expressions

$$f_d^+(q_0, q_1) = (q_1, f_d^+(q_0, q_1)),$$
$$f_d^-(q_0, q_1) = (q_0, f_d^-(q_0, q_1)).$$

We combine the two discrete forces to give a single one-form $f_d : Q \times Q \to T^*(Q \times Q)$ defined by

$$f_d(q_0, q_1) \cdot (\delta q_0, \delta q_1) = f_d^+(q_0, q_1) \cdot \delta q_1 + f_d^-(q_0, q_1) \cdot \delta q_0. \tag{3.2.1}$$

As with discrete Lagrangians, the discrete forces will also depend on the time-step h, which is important when relating the discrete and continuous mechanics. Given such forces, we modify the discrete Hamilton's principle, following Kane *et al.* (2000), to the *discrete Lagrange–d'Alembert principle*,

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

which seeks discrete curves $\{q_k\}_{k=0}^N$ that satisfy

$$\delta \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) + \sum_{k=0}^{N-1} \left[f_d^-(q_k, q_{k+1}) \cdot \delta q_k + f_d^+(q_k, q_{k+1}) \cdot \delta q_{k+1} \right] = 0 \tag{3.2.2}$$

for all variations $\{\delta q_k\}_{k=0}^N$ vanishing at the endpoints. This is equivalent to the *forced discrete Euler–Lagrange equations*

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) + f_d^+(q_{k-1}, q_k) + f_d^-(q_k, q_{k+1}) = 0, \tag{3.2.3}$$

which are the same as the standard discrete Euler–Lagrange equations (1.3.3) with the discrete forces added. These implicitly define the *forced discrete Lagrangian map* $f_d : Q \times Q \to Q \times Q$.

3.2.2. Discrete Legendre transforms with forces

Although in the continuous case we used the standard Legendre transform for systems with forcing, in the discrete case it is necessary to take the *forced discrete Legendre transforms* to be

$$\mathbb{F}^{f+} L_d : (q_0, q_1) \mapsto (q_1, p_1) = (q_1, D_2 L_d(q_0, q_1) + f_d^+(q_0, q_1)), \tag{3.2.4a}$$

$$\mathbb{F}^{f-} L_d : (q_0, q_1) \mapsto (q_0, p_0) = (q_0, -D_1 L_d(q_0, q_1) - f_d^-(q_0, q_1)). \tag{3.2.4b}$$

Using these definitions and the forced discrete Euler–Lagrange equations (3.2.3), we can see that the corresponding *forced discrete Hamiltonian map* $\tilde{F}_{L_d} = \mathbb{F}^{f\pm} L_d \circ F_{L_d} \circ (\mathbb{F}^{f\pm} L_d)^{-1}$ is given by the map $\tilde{F}_{L_d} : (q_0, p_0) \mapsto (q_1, p_1)$, where

$$p_0 = -D_1 L_d(q_0, q_1) - f_d^-(q_0, q_1), \tag{3.2.5a}$$

$$p_1 = D_2 L_d(q_0, q_1) + f_d^+(q_0, q_1), \tag{3.2.5b}$$

which is the same as the standard discrete Hamiltonian map (1.5.4) with the discrete forces added.

3.2.3. Discrete Noether's theorem with forcing

Consider a group action $\Phi : G \times Q \to Q$ and assume that the discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$ is invariant under the lifted product action, as in Section 1.3.3. We can now calculate (3.2.2) in the direction of a variation $\delta q_k = \xi_Q(q_k)$ to give

$$\sum_{k=0}^{N-1} \mathbf{d} L_d(q_k, q_{k+1}) \cdot \xi_{Q\times Q}(q_k, q_{k+1}) + \sum_{k=0}^{N-1} f_d(q_k, q_{k+1}) \cdot \xi_{Q\times Q}(q_k, q_{k+1})$$
$$= \sum_{k=0}^{N-1} f_d(q_k, q_{k+1}) \cdot \xi_{Q\times Q}(q_k, q_{k+1}),$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

or we can use a discrete integration by parts to obtain the alternative expression

$$\begin{aligned}\sum_{k=1}^{N-1} &\big[D_2L_d(q_{k-1},q_k)+D_1L_d(q_k,q_{k+1})+f_d^+(q_{k-1},q_k)+f_d^-(q_k,q_{k+1})\big]\cdot\xi_Q(q_k)\\ &+\big[D_2L_d(q_{N-1},q_N)+f_d^+(q_{N-1},q_N)\big]\cdot\xi_Q(q_N)\\ &+\big[D_1L_d(q_0,q_1)+f_d^-(q_0,q_1)\big]\cdot\xi_Q(q_0)\\ &\qquad=\mathbb{F}^{f+}L_d(q_{N-1},q_N)\cdot\xi_Q(q_N)-\mathbb{F}^{f-}L_d(q_0,q_1)\cdot\xi_Q(q_0).\end{aligned}$$

We now consider how the discrete momentum map should be defined in the presence of forcing, as there is a choice between the expressions (1.2.7) involving $\Theta^{\pm}_{L_d}$ and the expressions

$$J^{f+}_{L_d}(q_0,q_1)\cdot\xi=\big\langle\mathbb{F}^{f+}L_d(q_0,q_1),\xi_Q(q_1)\big\rangle, \tag{3.2.6a}$$

$$J^{f-}_{L_d}(q_0,q_1)\cdot\xi=\big\langle\mathbb{F}^{f-}L_d(q_0,q_1),\xi_Q(q_1)\big\rangle, \tag{3.2.6b}$$

which are based on the discrete Legendre transforms. In the unforced discrete case and in the continuous case both with and without forcing, these expressions are equal to the definition based on Θ_L and so the question does not arise. For a discrete system, however, consideration of the forced exact discrete Lagrangian defined below shows that (3.2.6) are the correct definitions. Given this, we can equate the above two forms of (3.2.2) to obtain

$$\big[J^{f+}_{L_d}\circ F^{N-1}_{L_d}-J^{f-}_{L_d}\big](q_0,q_1)\cdot\xi=\sum_{k=0}^{N-1}f_d(q_k,q_{k+1})\cdot\xi_{Q\times Q}(q_k,q_{k+1}),$$

which describes the evolution of the discrete momentum map. If the discrete forces are orthogonal to the group action, so that $\langle f_d,\xi_{Q\times Q}\rangle=0$ for all $\xi\in\mathfrak{g}$, then we have

$$0=\big\langle\mathbf{d}L_d+f_d,\xi_{Q\times Q}\big\rangle=J^{f+}_{L_d}-J^{f-}_{L_d},$$

and thus the two discrete Lagrangian momentum maps are equal. Denoting this unique map by $J^f_{L_d}:Q\times Q\to\mathfrak{g}^*$, we see that the momentum map evolution equation gives a forced Noether's theorem for discrete mechanics.

Theorem 3.2.1. (Discrete forced Noether's theorem) Consider a discrete Lagrangian system $L_d:Q\times Q\to\mathbb{R}$ with discrete forces $f_d^+,f_d^-:Q\times Q\to T^*Q$ and a symmetry action $\Phi:G\times Q\to Q$ such that $\langle f_d,\xi_{Q\times Q}\rangle=0$ for all $\xi\in\mathfrak{g}$. Then the discrete Lagrangian momentum map $J^f_{L_d}:Q\times Q\to\mathfrak{g}^*$ will be preserved by the discrete Lagrangian evolution map, so that $J^f_{L_d}\circ F_{L_d}=J^f_{L_d}$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

With the above definition of the discrete Lagrangian momentum map in the presence of forcing, we see that it will be the pullback of the Hamiltonian momentum map under the forced discrete Legendre transforms, and so the discrete forced Noether's theorem can also be stated for the forced discrete Hamiltonian map $\tilde{F}_{L_d}$ with the canonical momentum map $J_H : T^*Q \to \mathfrak{g}^*$.

As in the continuous case, a similar calculation to that given above shows that the discrete symplectic form will not be preserved in the presence of forcing.

3.2.4. Exact discrete forcing

In the unforced case, we have seen that the discrete Lagrangian should approximate the continuous action over the time-step. When forces are added, this must be modified so that the discrete Lagrange–d'Alembert principle (3.2.2) approximates the continuous expression (3.1.1).

Given a Lagrangian $L : TQ \to \mathbb{R}$ and a Lagrangian force $f_L : TQ \to T^*Q$, we define the *exact forced discrete Lagrangian* $L_d^E : Q \times Q \times \mathbb{R} \to \mathbb{R}$ and the *exact discrete forces* $f_d^{E+}, f_d^{E-} : Q \times Q \times \mathbb{R} \to T^*Q$ to be

$$L_d^E(q_0, q_1, h) = \int_0^h L(q(t), \dot{q}(t))\,\mathrm{d}t, \tag{3.2.7a}$$

$$f_d^{E+}(q_0, q_1, h) = \int_0^h f_L(q(t), \dot{q}(t)) \cdot \frac{\partial q(t)}{\partial q_1}\,\mathrm{d}t, \tag{3.2.7b}$$

$$f_d^{E-}(q_0, q_1, h) = \int_0^h f_L(q(t), \dot{q}(t)) \cdot \frac{\partial q(t)}{\partial q_0}\,\mathrm{d}t, \tag{3.2.7c}$$

where $q : [0, h] \to Q$ is the solution of the forced Euler–Lagrange equations (3.1.2) for L and f_L satisfying the boundary conditions $q(0) = q_0$ and $q(h) = q_1$.

Note that this exact discrete Lagrangian is not the same as that for the unforced system with Lagrangian L, as the curves $q(t)$ are different. In other words, the exact discrete Lagrangian depends on both the continuous Lagrangian and the continuous forces, as do the discrete forces.

Given these definitions of the exact discrete quantities and the forced discrete Legendre transforms, it is easy to check that the forced version of Lemma 1.6.2 holds, and thus so too do forced versions of Theorems 1.6.4 and 1.6.3, showing the equivalence of the exact discrete system to the continuous systems. This is of particular interest because it shows that the variational error analysis developed in Section 2.3 can also be extended to the case of forced systems in the obvious way, and that there will be a forced version of Theorem 2.3.1.

Note that, if $\Phi\colon G \times Q \to Q$ is a symmetry of L such that $\langle f_L(q, \dot{q}), \xi_Q(q)\rangle = 0$, so the forced Noether's theorem holds, then the exact discrete forces will

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

satisfy $\langle f_d, \xi_{Q\times Q}\rangle = 0$ and so the forced discrete Noether's theorem will also hold, as we would expect. This shows that (3.2.6) are the correct choice for the definition of the discrete Lagrangian momentum maps in the presence of forcing.

3.2.5. Integration of forced systems

To simulate a given forced Lagrangian or Hamiltonian system, we can choose a discrete Lagrangian and discrete forces to approximate the exact quantities given above, and then consider the resulting discrete system as an integrator for the continuous problem. We now give some simple examples of how to effect this.

Example 3.2.2. The natural discrete forces for the discrete Lagrangian L_d^α given in Example 2.3.2 are

$$f_d^{\alpha+}(q_0,q_1,h) = \alpha h f_L\left((1-\alpha)q_0+\alpha q_1, \frac{q_1-q_0}{h}\right),$$
$$f_d^{\alpha-}(q_0,q_1,h) = (1-\alpha) h f_L\left((1-\alpha)q_0+\alpha q_1, \frac{q_1-q_0}{h}\right).$$

For $L = \frac{1}{2}\dot{q}^T M\dot{q} - V(q)$, the discrete Hamiltonian map is then

$$\begin{aligned}\frac{q_1-q_0}{h} &= M^{-1}\left(\alpha p_0 + (1-\alpha)p_1\right),\\ \frac{p_1-p_0}{h} &= -\nabla V\left((1-\alpha)q_0+\alpha q_1\right)\\ &\qquad + f_H\left((1-\alpha)q_0+\alpha q_1, \alpha p_0 + (1-\alpha)p_1\right),\end{aligned}$$

which is the same as the unforced map (2.3.5) with the Hamiltonian force $f_H = (\mathbb{F}L)^{-1}\circ f_L$ added to the momentum equation. For $\alpha = 1/2$ this is once again simply the midpoint rule. ◊

A particularly interesting class of Lagrangian forces $f_L : TQ \to T^*Q$ consists of those forces that satisfy

$$\left\langle f_L(q,\dot{q}), (q,\dot{q})\right\rangle < 0,$$

for all $(q,\dot{q}) \in TQ$. Such forces are said to be (*strongly*) *dissipative*. This terminology can be justified by computing the time evolution of the energy $E_L : TQ \to \mathbb{R}$ along a solution of the forced Euler–Lagrange equations to give

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}E_L(q(t),\dot{q}(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\cdot\dot{q} + \frac{\partial L}{\partial \dot{q}}\cdot\ddot{q} - \frac{\mathrm{d}}{\mathrm{d}t}L\\ &= \left(\frac{\partial L}{\partial q} + f_L\right)\cdot\dot{q} + \frac{\partial L}{\partial \dot{q}}\cdot\ddot{q} - \frac{\partial L}{\partial q}\cdot\dot{q} - \frac{\partial L}{\partial \dot{q}}\cdot\ddot{q}\\ &= f_L(q(t),\dot{q}(t))\cdot\dot{q}(t).\end{aligned}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

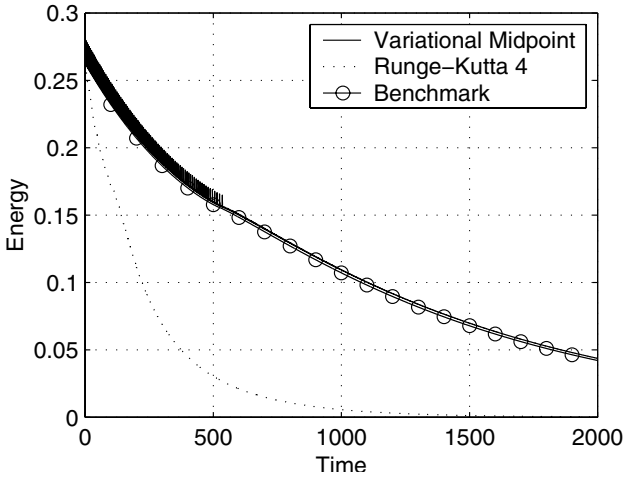


Fig. 2. Energy of a dissipative system computed with variational midpoint and Runge–Kutta. Note the accurate energy behaviour of the variational method

We thus see that dissipative forces are those for which the energy of the system always decreases. If we only have $f_L \cdot \dot{q} \leq 0$ then the force is said to be *weakly dissipative*.

Because the discrete Euler–Lagrange equations do not, in general, conserve energy, it is unlikely that, without some time-step adaptation, there is a discrete analogue of this result.

Example 3.2.3. As an example of a dissipative system, consider the movement of a unit mass particle in the plane with radial potential $V(q) = \|q\|^2(\|q\|^2 - 1)^2$ and forcing $f_L(q, \dot{q}) = -10^{-3}\dot{q}$. For this force we have $f_L \cdot \dot{q} = -10^{-3}\|\dot{q}\|^2 \leq 0$.

In Figure 2 we plot the energy behaviour of the L_d^α method with $\alpha = 1/2$ for this system. For comparison, we also plot an extremely accurate benchmark trajectory, showing the true energy of the system, and the trajectory of the standard fourth-order Runge–Kutta method.

Observe that the variational method dissipates energy due to the discrete forces added to the Euler–Lagrange equations, but this energy dissipation is of the correct amount to accurately track the true energy. In contrast, non-conservative methods such as the Runge–Kutta integrator used here artificially dissipate energy.

These effects are of particular importance when the amount of forcing or dissipation in the system is small compared to the magnitude of the conservative dynamics and the time period of integration. For an investigation

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

of the long time behaviour of symplectic methods applied to systems with dissipation, see Hairer and Lubich (1999). ◊

Example 3.2.4. (Composition methods) Consider a sequence of discrete Lagrangians L_d^i, discrete forces f_d^{i+}, f_d^{i-} and time-step fractions γ^i for $i = 1, \ldots, s$ satisfying $\sum_{i=1}^s \gamma^i = 1$. Then we can form a composition discrete Lagrangian L_d and composition discrete forces f_d^+, f_d^- in a similar way to the procedures in Section 2.5.

Given points q_0 and q_s, define q_i for $i = 1, \ldots, s-1$ to satisfy the forced discrete Euler–Lagrange equations (3.2.3) along the sequence $q_0, q_1, \ldots, q_s$. Regarding the q_i as functions of q_0 and q_1, we now define the *composition discrete Lagrangian* and *composition discrete forces* by

$$\begin{aligned}
L_d(q_0, q_1, h) &= \sum_{i=1}^{s} L_d(q_{i-1}, q_i, \gamma^i h),\\
f_d^+(q_0, q_1, h) &= f_d^{s+}(q_{s-1}, q_s, \gamma^s h)\\
&\qquad + \sum_{i=1}^{s-1} \big(f_d^{i+}(q_{i-1}, q_i, \gamma^i h) + f_d^{i-}(q_i, q_{i+1}, \gamma^{i+1} h) \big) \cdot \frac{\partial q_i}{\partial q_1},\\
f_d^-(q_0, q_1, h) &= f_d^{1-}(q_0, q_1, \gamma^1 h)\\
&\qquad + \sum_{i=1}^{s-1} \big(f_d^{i+}(q_{i-1}, q_i, \gamma^i h) + f_d^{i-}(q_i, q_{i+1}, \gamma^{i+1} h) \big) \cdot \frac{\partial q_i}{\partial q_0}.
\end{aligned}$$

With these definitions it can be shown, using a similar derivation to that in Section 2.5, that the forced discrete Hamiltonian map for L_d and f_d^+, f_d^- is the composition of the individual forced discrete Hamiltonian maps, so that

$$\tilde{F}_{L_d}^h = \tilde{F}_{L_d^s}^{\gamma^s h} \circ \tilde{F}_{L_d^{s-1}}^{\gamma^{s-1} h} \circ \cdots \circ \tilde{F}_{L_d^1}^{\gamma^1 h}.$$

In forming composition methods it is often useful to use a sequence consisting of copies of a method together with its adjoint. It is thus worth noting that the adjoint of a discrete Lagrangian and discrete forces is given by

$$\begin{aligned}
L_d^*(q_0, q_1, h) &= -L_d(q_1, q_0, -h),\\
f_d^{*+}(q_0, q_1, h) &= -f_d^-(q_1, q_0, -h),\\
f_d^{*-}(q_0, q_1, h) &= -f_d^+(q_1, q_0, -h).
\end{aligned}$$

The discrete Hamiltonian map of the adjoint discrete Lagrangian and adjoint discrete forces will be the adjoint map of the original discrete Hamiltonian map. Observe that the exact discrete Lagrangian and exact discrete forces (3.2.7) are self-adjoint. ◊

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Example 3.2.5. (Symplectic partitioned Runge–Kutta methods) Recall that the discrete Lagrangian (2.6.8) given by

$$L_d(q_0, q_1, h) = h \sum_{i=1}^{s} b_i L(Q_i, \dot{Q}_i)$$

generates symplectic partitioned Runge–Kutta methods. Reasonable choices of corresponding discrete forces are

$$f_d^+(q_0, q_1, h) = h \sum_{i=1}^{s} b_i f_L(Q_i, \dot{Q}_i) \cdot \frac{\partial Q_i}{\partial q_0}, \tag{3.2.8a}$$

$$f_d^-(q_0, q_1, h) = h \sum_{i=1}^{s} b_i f_L(Q_i, \dot{Q}_i) \cdot \frac{\partial Q_i}{\partial q_1}, \tag{3.2.8b}$$

which approximate the exact forces (3.2.7b) and (3.2.7c) in the same way that L_d approximates the exact discrete Lagrangian (3.2.7a).

With these choices of discrete forces, it can be shown that the discrete Hamiltonian map defined by (3.2.5) is exactly a partitioned Runge–Kutta method for the forced Hamiltonian system (3.1.3). ♢

In most of the other examples of variational integrators discussed above, discrete forces can be chosen in a natural way so that the discrete Hamiltonian maps give the expected integrator for the forced Hamiltonian system. In particular, this can be done for the symplectic Newmark methods (see Kane *et al.* (2000)). We can also use alternative splitting-style methods to include forcing (see Kane *et al.* (2000) for details).

3.3. Background: Constrained systems

A particularly elegant way to study many systems is to consider them as a constrained version of some larger system. This can be appealing for both theoretical reasons and, as we shall see, also on numerical grounds. Here we will only consider so-called *holonomic constraints*, which are constraints on the configuration manifold of a system.

More precisely, if we have a Lagrangian or Hamiltonian system with configuration manifold Q, we consider a *constraint function* $\phi : Q \to \mathbb{R}^d$ and constrain the dynamics to the *constraint submanifold* $N = \phi^{-1}(0) \subset Q$. Here we will always assume that $0 \in \mathbb{R}^d$ is a regular point of ϕ, so that N is truly a submanifold of Q (Abraham *et al.* 1988).

Observe that, if $i : N \to Q$ is the embedding map, then $Ti : TN \to TQ$ provides a canonical way to embed TN in TQ and we will thus regard TN as a submanifold of TQ. There is, however, no canonical way to embed the cotangent bundle T^*N in T^*Q, a fact which has important consequences for the development of constrained Hamiltonian dynamics. We will see below

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

that, in the special case when we have a regular Lagrangian or Hamiltonian, we can use this additional structure to provide a canonical embedding.

As in other areas of mechanics, we may consider constrained systems from both the Hamiltonian and the Lagrangian viewpoint. We will concentrate on the variational approach, however, as it is this formulation which readily extends to the discrete setting. The primary tool for constrained optimization problems is the Lagrange multiplier theorem, which we recall here (see Abraham *et al.* (1988) for the proof).

Theorem 3.3.1. Consider a smooth manifold $\mathcal{C}$ and a function $\Phi : \mathcal{C} \to V$ mapping to some inner product space V, such that $0 \in V$ is a regular point of Φ. Set $\mathcal{D} = \Phi^{-1}(0) \subset \mathcal{C}$. Given a function $\mathfrak{G} : \mathcal{C} \to \mathbb{R}$, define $\bar{\mathfrak{G}} : \mathcal{C} \times V \to \mathbb{R}$ by $\bar{\mathfrak{G}}(q, \lambda) = \mathfrak{G}(q) - \langle \lambda, \Phi(q) \rangle$. Then the following are equivalent:

(1) $q \in \mathcal{D}$ is an extremum of $\mathfrak{G}|_{\mathcal{D}}$;
(2) $(q, \lambda) \in \mathcal{C} \times V$ is an extremum of $\bar{\mathfrak{G}}$.

3.3.1. Constrained Lagrangian systems

Given a Lagrangian system specified by a configuration manifold Q and a Lagrangian $L : TQ \to \mathbb{R}$, consider the holonomic constraint $\phi : Q \to \mathbb{R}^d$ and the corresponding constraint submanifold $N = \phi^{-1}(0) \subset Q$. Now TN is a submanifold of TQ, and so we may restrict L to $L^N = L|_{TN}$. We are interested in the relationship of the dynamics of L^N on TN to the dynamics of L on TQ.

To consider this, we will make use of the following convenient notation. Assume that we are working on a given time interval $[0, T] \subset \mathbb{R}$, and that we have fixed endpoints $q_0, q_T \in N \subset Q$. Now set $\mathcal{C}(Q) = \mathcal{C}([0, T], Q; q_0, q_T)$ to be the space of smooth functions $q : [0, T] \to Q$ satisfying $q(0) = q_0$ and $q(T) = q_T$, and $\mathcal{C}(N)$ to be the corresponding space of curves in N. Similarly, we set $\mathcal{C}(\mathbb{R}^d) = \mathcal{C}([0, T], \mathbb{R}^d)$ to be curves $\lambda : [0, T] \to \mathbb{R}^d$ with no boundary conditions. In general $\mathcal{C}(P)$ is the space of curves from $[0, T]$ to the manifold P with the appropriate boundary conditions.

Theorem 3.3.2. Given a Lagrangian system $L : TQ \to \mathbb{R}$ with holonomic constraint $\phi : Q \to \mathbb{R}^d$, set $N = \phi^{-1}(0) \subset Q$ and $L^N = L|_{TN}$. Then the following are equivalent:

(1) $q \in \mathcal{C}(N)$ extremizes $\mathfrak{G}^N$ and hence solves the Euler–Lagrange equations for L^N;
(2) $q \in \mathcal{C}(Q)$ and $\lambda \in \mathcal{C}(\mathbb{R}^d)$ satisfy the *constrained Euler–Lagrange equations*

$$\frac{\partial L}{\partial q^i}(q(t), \dot{q}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}(q(t), \dot{q}(t))\right) = \left\langle \lambda(t), \frac{\partial \phi}{\partial q^i}(q(t)) \right\rangle, \quad (3.3.1a)$$

$$\phi(q(t)) = 0; \quad (3.3.1b)$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

(3) $(q,\lambda) \in \mathcal{C}(Q \times \mathbb{R}^d)$ extremizes $\bar{\mathfrak{G}}(q,\lambda) = \mathfrak{G}(q) - \langle \lambda, \Phi(q) \rangle$ and hence solves the Euler–Lagrange equations for the *augmented Lagrangian* $\bar{L} : T(Q \times \mathbb{R}^d) \to \mathbb{R}$ defined by

$$\bar{L}(q,\lambda,\dot{q},\dot{\lambda}) = L(q,\dot{q}) - \langle \lambda, \phi(q) \rangle .$$

Proof. We make use of the Lagrange multiplier theorem, Theorem 3.3.1. To do so, we prepare the following definitions. The full space is $\mathcal{C} = \mathcal{C}(Q)$ and the function to be extremized is the action $\mathfrak{G} : \mathcal{C}(Q) \to \mathbb{R}$. Take $V = \mathcal{C}(\mathbb{R}^d)$ with the L_2 inner product and define the constraint function $\Phi : \mathcal{C} \to V$ by $\Phi(q)(t) = \phi(q(t))$. Clearly $\Phi(q) = 0$, and hence $\phi(q(t)) = 0$ for all $t \in [0,T]$, if and only if $q \in \mathcal{C}(N)$. We thus obtain that the constraint submanifold is $\mathcal{D} = \Phi^{-1}(0) = \mathcal{C}(N)$.

Condition (1) simply means that $q \in \mathcal{C}(N) = \mathcal{D}$ is an extremum of the action for L^N, which is readily seen to be the standard action restricted to $\mathcal{C}(N)$. Thus $q \in \mathcal{D}$ is an extremum of $\mathfrak{G}|_{\mathcal{D}}$ and so, by the Lagrange multiplier theorem, this is equivalent to $(q,\lambda) \in \mathcal{C} \times V$ being an extremum of $\bar{\mathfrak{G}}(q,\lambda) = \mathfrak{G}(q) - \langle \lambda, \Phi(q) \rangle$.

Now $\mathcal{C} \times V = \mathcal{C}(Q) \times \mathcal{C}(\mathbb{R}^d)$ and so it can be identified with $\mathcal{C}(Q \times \mathbb{R}^d)$. Furthermore, we see that $\bar{\mathfrak{G}} : \mathcal{C}(Q \times \mathbb{R}^d) \to \mathbb{R}$ is

$$\begin{aligned} \bar{\mathfrak{G}}(q,\lambda) &= \mathfrak{G}(q) - \langle \lambda, \Phi(q) \rangle \\ &= \int_0^T L(q(t),\dot{q}(t))\,\mathrm{d}t - \int_0^T \langle \lambda(t), \Phi(q)(t) \rangle\,\mathrm{d}t \\ &= \int_0^T \left[L(q(t),\dot{q}(t)) - \langle \lambda(t), \phi(q(t)) \rangle \right] \mathrm{d}t, \end{aligned}$$

which is simply the action for the augmented Lagrangian $\bar{L}(q,\lambda,\dot{q},\dot{\lambda}) = L(q,\dot{q}) - \langle \lambda, \phi(q) \rangle$. As $(q,\lambda) \in \mathcal{C}(Q \times \mathbb{R}^d)$ must extremize this action, we see that it is a solution of the Euler–Lagrange equations for $\bar{L}$, which is statement (3).

Finally, we extremize $\bar{\mathfrak{G}}$ by solving $\mathbf{d}\bar{\mathfrak{G}} = 0$ to obtain the Euler–Lagrange equations. The standard integration by parts argument gives (3.3.1a) for variations with respect to q, and variations with respect to λ imply (3.3.1b), and thus we have equivalence to statement (2). □

If $i : N \to Q$ is the embedding, then by differentiating $L^N = L \circ Ti$ with respect to $\dot{q}$ we see that

$$\frac{\partial L^N}{\partial \dot{q}}(v_q) \cdot w_q = \frac{\partial L}{\partial \dot{q}}\big(Ti(v_q)\big) \cdot Ti \cdot w_q, \tag{3.3.2}$$

which means that if L is regular then so is L^N and shows that the following diagram commutes.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

$$\begin{array}{ccc} TQ|_N & \xrightarrow{\mathbb{F}L} & T^*Q|_N \\ {\scriptstyle Ti}\uparrow & & \downarrow{\scriptstyle T^*i} \\ TN & \xrightarrow[\mathbb{F}L^N]{} & T^*N \end{array} \tag{3.3.3}$$

Using this together with the fact that $\pi_Q \circ Ti = i \circ \pi_N$ for the projections $\pi_Q : TQ \to Q$ and $\pi_N : TN \to N$, we compute the pullback of the Lagrange one-form Θ_L on TQ to be

$$\begin{aligned} \big((Ti)^*\Theta_L\big)(v_q) \cdot \delta v_q &= \big\langle \mathbb{F}L\big(Ti(v_q)\big), T\pi_Q \circ T(Ti) \cdot \delta v_q \big\rangle \\ &= \big\langle \mathbb{F}L\big(Ti(v_q)\big), Ti \circ T\pi_N \cdot \delta v_q \big\rangle \\ &= \big\langle \mathbb{F}L^N(v_q), T\pi_N \cdot \delta v_q \big\rangle, \end{aligned}$$

and thus we see that $(Ti)^*\Theta_L = \Theta_{L^N}$, and so

$$(Ti)^*\Omega_L = \Omega_{L^N}. \tag{3.3.4}$$

Using the projection $T^*i : T^*Q \to T^*N$ we can reinterpret statement (2) of Theorem 3.3.2. Observe that the span of the $\nabla\phi^i$, $i = 1, \ldots, d$ is exactly the null space of T^*i, and so (3.3.1) is equivalent to

$$(T^*i)_{q(t)} \left[\frac{\partial L}{\partial q}(q(t), \dot{q}(t)) - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}}(q(t), \dot{q}(t)) \right) \right] = 0. \tag{3.3.5}$$

The above relationships hold for any Lagrangian L, irrespective of regularity. Also note that, although there is a canonical projection $T^*i : T^*Q \to T^*N$, there is no corresponding canonical embedding of T^*N into T^*Q. We will see below that when L is regular we can use the Legendre transform to define such an embedding.

3.3.2. Constrained Hamiltonian systems: Augmented approach

One can consider the Hamiltonian formulation of constrained systems by either working on the augmented space $T^*(Q \times \mathbb{R}^d)$, or working directly on T^*N, which gives the Dirac theory of constraints. We consider the former option first.

Given a Hamiltonian $H : T^*Q \to \mathbb{R}$, we define the *augmented Hamiltonian* to be

$$\bar{H}(q, \lambda, p, \pi) = H(q, p) + \langle \lambda, \phi(q) \rangle,$$

where π is the conjugate variable to λ. We now consider the *primary constraint set* $\Pi \subset T^*(Q \times \mathbb{R}^d)$ defined by $\pi = 0$. Pulling Ω back to Π gives the degenerate two-form Ω^Π, and the augmented Hamiltonian vector field $\bar{X}_{\bar{H}}$ is defined by

$$\mathbf{i}_{\bar{X}_{\bar{H}}} \Omega^\Pi = \mathbf{d}\bar{H},$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

which in coordinates is the set of *constrained Hamilton's equations*

$$\bar{X}_{q^i}(q,\lambda,p,\pi) = \frac{\partial H}{\partial p_i}, \tag{3.3.6a}$$

$$\bar{X}_{p_i}(q,\lambda,p,\pi) = -\frac{\partial H}{\partial q^i} - \left\langle \lambda, \frac{\partial \phi}{\partial q^i}(q) \right\rangle, \tag{3.3.6b}$$

$$\phi(q) = 0, \tag{3.3.6c}$$

where there is no λ equation owing to the degeneracy of Ω^{Π}. Note that for nonregular H these equations will not, in general, uniquely define the vector field $\bar{X}_{\bar{H}}$.

Consider now a regular Lagrangian L and its corresponding Hamiltonian H. Observe that the augmented Lagrangian $\bar{L}$ is degenerate, owing to the lack of dependence on $\dot{\lambda}$, and that the primary constraint manifold Π is exactly the image of $\mathbb{F}\bar{L}$. The augmented Hamiltonian and Lagrangian satisfy the equation $\bar{H} \circ \mathbb{F}\bar{L} = E_{\bar{L}}$, but this does not uniquely specify $\bar{H}$ since $\mathbb{F}\bar{L}$ need not be invertible. Nonetheless, it is simple to check that the constrained Hamilton's equations given above are equivalent to the constrained Euler–Lagrange equations (3.3.1) when we neglect the π component.

3.3.3. Constrained Hamiltonian systems: Dirac theory

As an alternative to working on the augmented space $T^*(Q \times \mathbb{R}^d)$, we can directly compare the dynamics of the constrained system on T^*N with those on T^*Q. The general form of this is the *Dirac theory of constraints* (Marsden and Ratiu 1999), but here we use only the simple case of holonomic constraints on cotangent bundles.

The main problem with this approach is that there is no canonical way to embed T^*N within T^*Q. For now we will assume that we have an embedding $\eta : T^*N \to T^*Q$ such that $\pi_Q \circ \eta = i \circ \pi_N$ and $\eta^*\Omega = \Omega^N$, where Ω and Ω^N are the canonical two-forms on T^*Q and T^*N respectively, and we will see below how to construct η given a regular Hamiltonian or Lagrangian.

Given a Hamiltonian $H : T^*Q \to \mathbb{R}$, we define $H^N : T^*N \to \mathbb{R}$ by $H^N = H \circ \eta$. The constrained Hamiltonian vector field $X_{H^N} : T^*N \to T(T^*N)$ is then defined by

$$\mathbf{i}_{X_{H^N}} \Omega^N = \mathbf{d}H^N.$$

Taking $\pi_\Omega : T(T^*Q) \to T(T^*N)$ to be the projection operator determined by using Ω to define the orthogonal complement of $T\eta \cdot T(T^*N) \subset T(T^*Q)$, leads us to the following simple relationship between the Hamiltonian vector field X_H and the constrained vector field X_{H^N}.

Theorem 3.3.3. Consider a Hamiltonian system $H : T^*Q \to \mathbb{R}$ and the corresponding constrained system $H^N : T^*N \to \mathbb{R}$ as defined above. Then

$$X_{H^N} = \pi_\Omega \cdot X_H \circ \eta.$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Proof. We have that $\eta^*\Omega = \Omega^N$. Take an arbitrary $V^N \in T(T^*N)$ and compute

$$\begin{aligned}\mathbf{i}_{(\pi_\Omega \cdot X_H \circ \eta)}\Omega^N \cdot V^N &= \Omega(T\eta \cdot \pi_\Omega \cdot X_H, T\eta \cdot V^N) \\ &= \Omega(X_H, T\eta \cdot V^N) \\ &= \mathbf{d}H \cdot T\eta \cdot V^N \\ &= \mathbf{d}H^N \cdot V^N \\ &= \mathbf{i}_{X_{H^N}}\Omega^N \cdot V^N,\end{aligned}$$

where we used the fact that $(Id - T\eta \cdot \pi_\Omega) \cdot X_H$ is Ω-orthogonal to the set $T\eta \cdot T(T^*N)$. Finally, the fact that Ω^N is nondegenerate gives the desired equivalence. □

3.3.4. Legendre transforms

Until this point we have assumed that we are using any symplectic embedding $\eta : T^*N \to T^*Q$ covering the embedding $i : N \to Q$. We now consider a hyperregular Hamiltonian H and the corresponding hyperregular Lagrangian L. Recall that *hyperregularity* of H, for example, means that $\mathbb{F}H$ is not only a local diffeomorphism (equivalent to regularity), but is a *global* diffeomorphism. Of course, if we only have regularity, these constructions may be done locally. We will show that, using this additional structure, there is a canonical way to construct η.

To do this, begin from either a hyperregular Lagrangian L or a hyperregular Hamiltonian H, and construct the corresponding L or H, which is necessarily hyperregular as well and has $\mathbb{F}L = (\mathbb{F}H)^{-1}$. This implies that L^N and H^N are also hyperregular.

We now define $\eta : T^*N \to T^*Q$ by requiring that the following diagram commutes, where $i : N \to Q$ is the embedding as before.

$$\begin{array}{ccc} TQ|_N & \xrightarrow{\mathbb{F}L} & T^*Q|_N \\ {\scriptstyle Ti}\uparrow & & \uparrow{\scriptstyle \eta} \\ TN & \xrightarrow[\mathbb{F}L^N]{} & T^*N \end{array} \tag{3.3.7}$$

Clearly $\pi_Q \circ \eta = i \circ \pi_N$, and from (3.3.4) we see that $\eta^*\Omega = \Omega^N$, and so η gives a symplectic embedding of T^*N in T^*Q. Note that, although T_qN is a linear subset of T_qQ, the map η is in general not linear and so T_q^*N is not a linear subspace of T_q^*Q. It is true, however, that $T_{p_q}(T^*N)$ is a linear subspace of $T_{p_q}(T^*Q)$.

Regarding T^*N as a submanifold of T^*Q by means of η, we have the natural embedding $T\eta : T(T^*N) \to T(T^*Q)$ and so we can regard X_{H^N} as a vector field on $\eta(T^*N)$. Using canonical coordinates (q^i, p_i) on T^*Q we

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

can derive a simple coordinate representation of this vector field:

$$\begin{aligned}\dot{q} &= \frac{\partial H}{\partial p},\\ \dot{p} &= -\frac{\partial H}{\partial q} - \lambda^T \nabla\phi(q),\\ \phi(q) &= 0.\end{aligned}$$

These equations are clearly equivalent to (3.3.6) above if we neglect the π variable there.

Consider the projection operator $\pi_{\Omega_L} : T(TQ) \to T(TN)$ defined by the Ω_L-orthogonal complement to $T(TN)$ regarded as a subspace of $T(TQ)$ by the map TTi. As $\Omega_L = (\mathbb{F}L)^*\Omega$, elements of $T(T^*Q)$ which are Ω-orthogonal pull back under $\mathbb{F}L$ to elements of $T(TQ)$ which are Ω_L-orthogonal. It follows that $T\mathbb{F}L^n \circ \pi_{\Omega_L} = \pi_\Omega \circ T\mathbb{F}L$. In addition, observe that, as both the constrained and unconstrained systems are regular, we obtain $X_L = (\mathbb{F}L)^* X_H$ and $X_{L^N} = (\mathbb{F}L^N)^* X_{H^N}$. Combining this with the statement of Theorem 3.3.3 and regarding TN and T^*N as submanifolds of TQ and T^*Q, respectively, gives the following commutative diagram.

$$\begin{array}{ccccc} T(TQ)|_{TN} & & \xrightarrow{\quad T\mathbb{F}L \quad} & & T(T^*Q)|_{T^*N} \\ & \overset{X_L}{\nwarrow} & & \overset{X_H}{\nearrow} & \\ \Big\downarrow{\scriptstyle \pi_{\Omega_L}} & & TN \xrightarrow[\mathbb{F}L^N]{(\mathbb{F}L)|_{TN}} T^*N & & \Big\downarrow{\scriptstyle \pi_\Omega} \\ & \underset{X_{L^N}}{\swarrow} & & \underset{X_{H^N}}{\searrow} & \\ T(TN) & & \xrightarrow[T\mathbb{F}L^N]{} & & T(T^*N) \end{array} \tag{3.3.8}$$

This establishes that $X_{L^N} = \pi_{\Omega_L} \circ X_L \circ Ti$, which is the Lagrangian analogue of Theorem 3.3.3. Note that this only holds for regular Lagrangians, whereas the Hamiltonian result does not require regularity.

A special case of hyperregular systems is when we have a Riemannian metric $\langle\langle \cdot, \cdot \rangle\rangle$ on Q and the Lagrangian is of the form

$$L(v_q) = \frac{1}{2} \langle\langle v_q, v_q \rangle\rangle - V \circ \pi_Q(v_q)$$

for a potential function $V : Q \to \mathbb{R}$. Computing the Legendre transform gives

$$\mathbb{F}L(v_q) \cdot w_q = \langle\langle v_q, w_q \rangle\rangle = v_q^T M(q) w_q,$$

where we introduce the symmetric positive definite mass matrix $M(q)$ as the coordinate representation of the metric. In coordinates, the Legendre transform is thus $p = M(q)\dot{q}$, and we see that the Legendre transform is

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

linear in $\dot{q}$ and so $\eta(T^*_qN)$ is a linear subspace of T^*_qQ at each $q \in N$. Note that the constrained subspaces can be expressed as

$$TN = \{(q,\dot{q}) \in TQ \mid \phi(q) = 0 \text{ and } \nabla\phi \cdot \dot{q} = 0\}, \tag{3.3.9}$$

$$\eta(T^*N) = \{(q,p) \in T^*Q \mid \phi(q) = 0 \text{ and } \nabla\phi \cdot M^{-1}(q)\dot{q} = 0\}. \tag{3.3.10}$$

We define the projection map $\mathbb{P} : T^*Q|_N \to \eta(T^*N)$ by $\mathbb{P} = \eta \circ T^*i$, and as it must satisfy $\mathbb{P}(\nabla\phi^m) = 0$ for each $m = 1, \ldots, d$ we can calculate the coordinate expression to be

$$\mathbb{P} = I - (\nabla\phi)^T[(\nabla\phi)M^{-1}(\nabla\phi)^T]^{-1}(\nabla\phi)M^{-1}, \tag{3.3.11}$$

where I is the $n \times n$ identity matrix and $\nabla\phi$ is the $d \times n$ matrix $[\nabla\phi(q)]_{mi} = \frac{\partial\phi^m}{\partial q^i}$, and all quantities are evaluated at $q \in N$.

Another way to derive this expression is to define an induced Riemannian metric on T^*Q by $\langle\langle p_q, r_q \rangle\rangle = \langle p_q, \mathbb{F}H(r_q)\rangle$, which has coordinate expression $p^TM^{-1}(q)r$. The projection $\mathbb{P}$ is then the projection onto the orthogonal subspace to the span of $\{\nabla\phi^m\}$ in the inner product given by this metric.

In this case, note that $\mathbb{P} : TQ|_N \to TN$, and so $T\mathbb{P} : T(TQ|_N) \to T(TN)$. However, observe that

$$T(TQ|_N) = \{w \in T(TQ) \mid T\pi_Q(w) \in TN\},$$

and as X_L is a second-order vector field, it satisfies $T\pi_Q \circ X_L = id$, and so we have that $X_L(v_q) \in T(TQ|_N)$ for all $v_q \in TN$. In particular, we can now show that, on the intersection of their domains, $\pi_{\Omega_L} = T\mathbb{P}$, which gives an explicit expression for the Lagrangian projection operator. This development is closely related to the expression of forces of constraint in terms of the second fundamental form (see Marsden and Ratiu (1999), Section 8.4).

3.3.5. Conservation properties

As we have seen above, the constrained systems on TN and T^*N defined by $L^N = L \circ Ti$ and $H^N = H \circ \eta$, respectively, are standard Lagrangian or Hamiltonian systems and so have the usual conservation properties.

In particular, the constrained Lagrangian system $L^N : TN \to \mathbb{R}$ will have a flow map that preserves the symplectic two-form $\Omega_{L^N} = (Ti)^*\Omega_L$, and the constrained Hamiltonian system $H^N : T^*N \to \mathbb{R}$ preserves the canonical two-form $\Omega^N = \eta^*\Omega$ on T^*N. For (hyper)regular systems, the Lagrangian and Hamiltonian two-forms are related by the Legendre transforms on both the constrained and unconstrained levels, so that $\Omega_L = (\mathbb{F}L)^*\Omega$ and $\Omega_{L^N} = (\mathbb{F}L^N)^*\Omega^N$.

Suppose that we have a group action $\Phi : G \times Q \to Q$ that leaves N invariant, that is, there is a restricted action $\Phi^N : G \times N \to N$ satisfying $i \circ \Phi^N = \Phi \circ i$. It is now a simple matter to check that the infinitesimal

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

generators are related by

$$\xi_Q \circ i = Ti \circ \xi_N,$$
$$\xi_{TQ} \circ Ti = T(Ti) \circ \xi_{TN},$$
$$\xi_{T^*Q} \circ \eta = T\eta \circ \xi_{T^*N},$$

and so the momentum maps satisfy

$$J_{L^N} = J_L \circ Ti,$$
$$J_{H^N} = J_H \circ \eta.$$

Since Noether's theorem holds for both the constrained and unconstrained systems, the above relationship shows that essentially the same momentum map is preserved on both levels. Note that if the group action does not leave the constraint submanifold N invariant, however, then in general it is not possible to define J_{L^N} or J_{H^N} and there will be no constrained Noether's theorems.

3.4. Discrete variational mechanics with constraints

We now consider a discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$ with the holonomic constraint $\phi : Q \to \mathbb{R}^d$ and corresponding constraint submanifold $N = \phi^{-1}(0) \subset Q$. As in the continuous case, the fact that $N \times N$ is naturally a submanifold of $Q \times Q$ means that we can restrict the discrete Lagrangian to $L_d^N = L_d|_{N\times N}$ to obtain a discrete Lagrangian system on $N \times N$. More precisely, we define the embedding $i^{N\times N} : N \times N \to Q \times Q$ by $i^{N\times N}(q_0, q_1) = (i(q_0), i(q_1))$.

To relate the dynamics of L_d^N to that of L_d, it is useful to introduce the notation for discrete trajectories corresponding to that used in the continuous case. Given times $\{0, h, 2h, \ldots, Nh = T\}$ and endpoints $q_0, q_T \in N$ we set $\mathbb{C}_d(Q) = \mathbb{C}_d(\{0, h, 2h, \ldots, Nh\}, Q; q_0, q_T)$ to be the set of discrete trajectories $q_d : \{0, h, 2h, \ldots, Nh\} \to Q$ satisfying $q_d(0) = q_0$ and $q_d(Nh) = q_T$, and $\mathbb{C}_d(N)$ to be the corresponding set of discrete trajectories in N.

Similarly, we denote by $\mathcal{C}_d(\mathbb{R}^d) = \mathcal{C}_d(\{h, 2h, \ldots, (N-1)h\}, \mathbb{R}^d)$ the set of maps $\lambda_d : \{h, 2h, \ldots, (N-1)h\} \to \mathbb{R}^d$ with no boundary conditions. We will see below why we do not include the boundary points 0 and Nh. In general, $\mathcal{C}_d(P)$ is the space of maps from $\{0, h, 2h, \ldots, Nh\}$ to the manifold P, and we identify such maps with their images, and write $q_d = \{q_k\}_{k=0}^N$ for $k = 0, 1, 2, \ldots, N$, and similarly for $\lambda_d = \{\lambda_k\}_{k=0}^N$.

3.4.1. Constrained discrete variational principle

As we have do not use vector fields to define the dynamics in the discrete case, and so cannot project such objects onto the constraint manifold, we turn instead to constraining the variational principle. The following theorem gives the result of this procedure.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Theorem 3.4.1. Given a discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$ with holonomic constraint $\phi : Q \to \mathbb{R}^d$, set $N = \phi^{-1}(0) \subset Q$ and $L_d^N = L_d|_{N\times N}$. Then the following are equivalent:

(1) $q_d = \{q_k\}_{k=0}^N \in \mathcal{C}_d(N)$ extremizes $\mathfrak{G}_d^N = \mathfrak{G}_d\big|_{N\times N}$ and hence solves the discrete Euler–Lagrange equations for L_d^N;

(2) $q_d = \{q_k\}_{k=0}^N \in \mathcal{C}_d(Q)$ and $\lambda_d = \{\lambda_k\}_{k=1}^{N-1} \in \mathcal{C}_d(\mathbb{R}^d)$ satisfy the *constrained discrete Euler–Lagrange equations*

$$D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1}) = \langle \lambda_k, \nabla\phi(q_k)\rangle\,, \tag{3.4.1a}$$

$$\phi(q_k) = 0; \tag{3.4.1b}$$

(3) $(q_d, \lambda_d) = \{(q_k, \lambda_k)\}_{k=0}^N \in \mathcal{C}_d(Q \times \mathbb{R}^d)$ extremizes $\bar{\mathfrak{G}}_d(q_d, \lambda_d) = \mathfrak{G}_d(q_d) - \langle \lambda_d, \Phi_d(q_d)\rangle_{l_2}$ and hence solves the discrete Euler–Lagrange equations for either of the *augmented discrete Lagrangians* $\bar{L}_d^+, \bar{L}_d^- : (Q \times \mathbb{R}^d) \times (Q \times \mathbb{R}^d) \to \mathbb{R}$ defined by

$$\bar{L}_d^+(q_k, \lambda_k, q_{k+1}, \lambda_{k+1}) = L_d(q_k, q_{k+1}) - \langle \lambda_{k+1}, \phi(q_{k+1})\rangle\,,$$
$$\bar{L}_d^-(q_k, \lambda_k, q_{k+1}, \lambda_{k+1}) = L_d(q_k, q_{k+1}) - \langle \lambda_k, \phi(q_k)\rangle\,.$$

Proof. The proof of Theorem 3.3.2 in the continuous case can be almost directly applied in the discrete case.

We take the full space to be $\mathcal{C}_d = \mathcal{C}_d(Q)$ and the function we are extremizing is the discrete action $\mathfrak{G}_d : \mathcal{C}_d(Q) \to \mathbb{R}$. The constraint is specified by setting $V_d = \mathcal{C}_d(\mathbb{R}^d)$ with the l_2 inner product, and defining the constraint function $\Phi_d : \mathcal{C}_d \to V_d$ by $\Phi_d(q_d)(kh) = \phi(q_d(kh)) = \phi(q_k)$. Thus $q_d \in \mathcal{C}_d(N)$ if and only if $\phi(q_k) = 0$ for all k, and hence if and only if $\Phi_d(q_d) = 0$. The constraint submanifold is therefore $\mathcal{D}_d = \Phi_d^{-1}(0) = \mathcal{C}_d(N)$.

As in the continuous case, statement (1) means that $q_d \in \mathcal{C}_d(N) = \mathcal{D}_d$ is an extremum of the action for L_d^N, which is the full action restricted to $\mathcal{C}_d(N)$. From the Lagrange multiplier theorem (Theorem 3.3.1), $q_d \in \mathcal{D}_d$ being an extremum of $\mathfrak{G}_d|_{\mathcal{D}_d}$ is equivalent to $(q_d, \lambda_d) \in \mathcal{C}_d \times V_d$ being an extremum of $\bar{\mathfrak{G}}_d(q_d, \lambda_d) = \mathfrak{G}_d(q_d) - \langle \lambda_d, \Phi_d(q_d)\rangle$. Computing, this gives

$$\begin{aligned}\bar{\mathfrak{G}}_d(q_d, \lambda_d) &= \mathfrak{G}_d(q_d) - \langle \lambda_d, \Phi_d(q_d)\rangle \\ &= \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) - \sum_{k=1}^{N-1} \langle \lambda_d(kh), \Phi_d(q_d)(kh)\rangle\,.\end{aligned}$$

Extremizing this function with respect to q_d now gives (3.4.1a), and extremizing with respect to λ_d recovers (3.4.1b). We therefore have equivalence to statement (2).

As we only extremize with respect to the internal points, and hold the boundary terms fixed, we may extend $\mathcal{C}_d(\mathbb{R}^d)$ to include λ_0 and λ_N. We now identify $\mathcal{C}_d \times V_d = \mathcal{C}_d(Q) \times \mathcal{C}_d(\mathbb{R}^d)$ with the space $\mathcal{C}_d(Q \times \mathbb{R}^d) =$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

$\mathcal{C}_d(\{0, h, 2h, \ldots, Nh\}, Q \times \mathbb{R}^d)$, and group the terms in the above expression for $\bar{\mathfrak{G}}_d$ to give two alternative functions $\bar{\mathfrak{G}}_d^+, \bar{\mathfrak{G}}_d^- : \mathcal{C}_d(Q \times \mathbb{R}^d) \to \mathbb{R}$ defined by

$$\bar{\mathfrak{G}}_d^+ = \sum_{k=0}^{N-1} \left[L_d(q_k, q_{k+1}) - \langle \lambda_{k+1}, \phi(q_{k+1}) \rangle \right],$$

$$\bar{\mathfrak{G}}_d^- = \sum_{k=0}^{N-1} \left[L_d(q_k, q_{k+1}) - \langle \lambda_k, \phi(q_k) \rangle \right],$$

which have the same extrema as $\bar{\mathfrak{G}}_d$ when the boundary terms are held fixed. Identifying the terms in the summations as the augmented discrete Lagrangians $\bar{L}_d^+$ and $\bar{L}_d^-$, respectively, gives equivalence to statement (3). □

Note that in Theorem 3.4.1 one can actually take any convex combination of $\bar{L}_d^+$ and $\bar{L}_d^-$, although this will not substantially alter the result.

We may also use the projection operator $T^*i : T^*Q|_N \to T^*N$ to act on statement (2) of Theorem 3.4.1, showing that (3.4.1) is equivalent to

$$(T^*i)_{q_k} \left[D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) \right] = 0. \tag{3.4.2}$$

This is the counterpart of the continuous equation (3.3.5).

3.4.2. Augmented Hamiltonian viewpoint

Just as in the continuous case, one can either work on the augmented space $T^*(Q \times \mathbb{R}^d)$ or directly on the constrained space T^*N.

The problem with trying to form the augmented discrete Hamiltonian maps $\bar{L}_d^{\pm}$ is the same as in this continuous case, namely the fact that the augmented discrete Lagrangians $\bar{L}_d^{\pm}$ are necessarily degenerate. Nonetheless, we will *define* the discrete Hamiltonian map $\tilde{F}_{\bar{L}_d^-} : (q_0, \lambda_0, p_0, \pi_0) \mapsto (q_1, \lambda_1, p_1, \pi_1)$ by the equations

$$p_0 = -D_1 L_d(q_0, q_1) + \langle \lambda_0, \nabla \phi(q_0) \rangle, \tag{3.4.3a}$$

$$\pi_0 = \phi(q_0), \tag{3.4.3b}$$

$$p_1 = D_2 L_d(q_0, q_1), \tag{3.4.3c}$$

$$\pi_1 = 0. \tag{3.4.3d}$$

Restricting to the same primary constraint set $\Pi \subset T^*(Q \times \mathbb{R}^d)$ as in the continuous case, we see that these equations are the equivalent to (3.4.1) together with the requirement $\phi(q_k) = 0$ and hence $q_d \in \mathcal{C}_d(N)$, that is, they are equivalent to statement (2) in Theorem 3.4.1.

Note that the evolution of λ is not well-defined, as in the continuous case, so that (3.4.3) do not define a map $\Pi \to \Pi$, that is, λ_0 is not a free initial condition, as it will be determined by (q_0, p_0). Note that constructing the

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

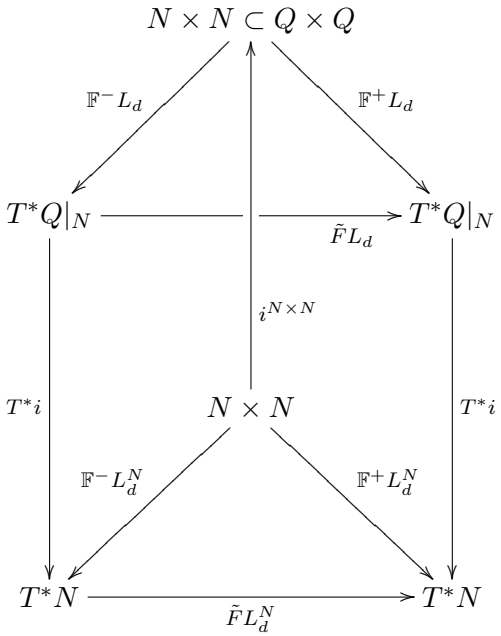
alternative map $\tilde{F}_{\bar{L}_d^+}$ does not give a well-defined forward map in general. In fact, to map forward in time it is necessary to use $\tilde{F}_{\bar{L}_d^-}$ as defined above, while $\tilde{F}_{\bar{L}_d^+}$ can be used to map backward in time.

3.4.3. Direct Hamiltonian viewpoint

Alternatively, one can neglect the augmented space and directly relate T^*N and T^*Q. To do so, we differentiate $L_d^N = L_d \circ i^{N\times N}$ with respect to q_0 and q_1 to obtain the discrete equivalents of (3.3.2), thus establishing that the following diagrams commute.

$$\begin{array}{ccc} T^*Q|_N & \xleftarrow{\mathbb{F}^- L_d} & Q\times Q|_{\{q_0\in N\}} \\ {\scriptstyle T^*i}\downarrow & & \uparrow{\scriptstyle i^{N\times N}} \\ T^*N & \xleftarrow[\mathbb{F}^- L_d^N]{} & N\times N \end{array} \qquad \begin{array}{ccc} Q\times Q|_{\{q_1\in N\}} & \xrightarrow{\mathbb{F}^+ L_d} & T^*Q|_N \\ {\scriptstyle i^{N\times N}}\uparrow & & \downarrow{\scriptstyle T^*i} \\ N\times N & \xrightarrow[\mathbb{F}^+ L_d^N]{} & T^*N \end{array} \tag{3.4.4}$$

We will henceforth assume that L_d is regular, which means that L_d^N is also regular and that the discrete Hamiltonian maps $\tilde{F}_{L_d}$ and $\tilde{F}_{L_d^N}$ are well-defined. Combining the above diagrams with the expressions $\tilde{F}_{L_d} = \mathbb{F}^+ L_d \circ (\mathbb{F}^- L_d)^{-1} : T^*Q \to T^*Q$ and $\tilde{F}_{L_d^N} = \mathbb{F}^+ L_d^N \circ (\mathbb{F}^- L_d^N)^{-1} : T^*N \to T^*N$ gives the following commutative diagram.



(3.4.5)

This proves the following theorem.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Theorem 3.4.2. Consider a regular discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$ and the constrained system $L_d^N : N \times N \to \mathbb{R}$ defined by $L_d^N = L_d \circ i^{N\times N}$. Then the discrete Hamiltonian map $\tilde{F}_{L_d^N} : T^*N \to T^*N$ has the following equivalent formulations:

(1) $\tilde{F}_{L_d^N} : (q_0, p_0) \mapsto (q_1, p_1)$ for $(q_0, p_0), (q_1, p_1) \in T^*N$ satisfying

$$p_0 = -D_1 L_d^N(q_0, q_1), \tag{3.4.6a}$$
$$p_1 = D_2 L_d^N(q_0, q_1); \tag{3.4.6b}$$

(2) $\tilde{F}_{L_d^N} : (q_0, p_0) \mapsto (q_1, p_1)$ for $(q_0, p_0), (q_1, p_1) \in T^*N$ satisfying

$$p_0 = (T^*i)_{q_0}(-D_1 L_d \circ i^{N\times N}(q_0, q_1)), \tag{3.4.7a}$$
$$p_1 = (T^*i)_{q_1}(D_2 L_d \circ i^{N\times N}(q_0, q_1)); \tag{3.4.7b}$$

(3) $\tilde{F}_{L_d^N} : \eta(T^*N) \mapsto \eta(T^*N)$ for $(q_0, p_0) \in \eta(T^*N)$ and $(q_1, p_1) \in T^*Q$ satisfying

$$p_0 = \mathbb{P}_{q_0}(-D_1 L_d \circ i^{N\times N}(q_0, q_1)), \tag{3.4.8a}$$
$$p_1 = \mathbb{P}_{q_1}(D_2 L_d \circ i^{N\times N}(q_0, q_1)), \tag{3.4.8b}$$
$$\phi(q_1) = 0. \tag{3.4.8c}$$

Here $\eta : T^*N \to T^*Q$ is any symplectic embedding covering the identity, so that $\pi_Q \circ \eta = i \circ \pi_N$, and $\mathbb{P} : T^*Q|_N \to \eta(T^*N)$ is the map defined by $\mathbb{P} = \eta \circ T^*i$.

This theorem is the discrete analogue of Theorem 3.3.3, and shows how the unconstrained Hamiltonian equations are related to the constrained equations. If we further assume that η is defined by (3.3.7) for some regular Lagrangian L with corresponding Hamiltonian H, then we can use the fact that the null space of $\mathbb{P}$ is the span of the $\nabla\phi^m$ and introduce Lagrange multipliers to write (3.4.8) as

$$p_0 = -D_1 L_d(q_0, q_1) + (\lambda^{(0)})^T \nabla\phi(q_0), \tag{3.4.9a}$$
$$p_1 = D_2 L_d(q_0, q_1) - (\lambda^{(1)})^T \nabla\phi(q_1), \tag{3.4.9b}$$
$$\phi(q_1) = 0, \tag{3.4.9c}$$
$$\nabla\phi(q_1) \cdot \frac{\partial H}{\partial p}(q_1, p_1) = 0, \tag{3.4.9d}$$

defining a map from $(q_0, p_0) \in \eta(T^*N)$ to $(q_1, p_1) \in T^*Q$ which will satisfy $(q_1, p_1) \in \eta(T^*N)$. Here the arbitrary signs on the Lagrange multipliers have been chosen to correspond to the signs for discrete forces in (3.2.5).

Now consider the special case when Q is a Riemannian manifold with metric $\langle\langle \cdot, \cdot \rangle\rangle$ having coordinate representation $M(q)$ and η is defined by (3.3.7)

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

for a Lagrangian with kinetic energy given by the metric. As we have seen in Section 3.3.4 above, $\eta(T^*N)$ and $\mathbb{P}$ are now given explicitly by (3.3.10) and (3.3.11), respectively. Using this, we can write (3.4.8) as

$$p_0 = -\big(I - (\nabla\phi)^T\big[(\nabla\phi)M^{-1}(\nabla\phi)^T\big]^{-1}(\nabla\phi)M^{-1}\big)D_1L_d(q_0,q_1), \tag{3.4.10a}$$

$$p_1 = \big(I - (\nabla\phi)^T\big[(\nabla\phi)M^{-1}(\nabla\phi)^T\big]^{-1}(\nabla\phi)M^{-1}\big)D_2L_d(q_0,q_1), \tag{3.4.10b}$$

$$\phi(q_1) = 0, \tag{3.4.10c}$$

where $\nabla\phi$ and M are evaluated at q_0 or q_1 as appropriate.

3.4.4. Conservation properties

A constrained discrete Lagrangian system on $N \times N$ and an unconstrained system on $Q \times Q$ will clearly preserve the standard discrete symplectic two-forms $\Omega_{L_d^N}$ and Ω_{L_d}, respectively. Now define the projections $\pi_Q^1 : Q \times Q \to Q$ and $\pi_N^1 : N \times N \to N$ onto the first components of $Q \times Q$ and $N \times N$. Observe that $\pi_Q^1 \circ i^{N\times N} = i \circ \pi_N^1$ and, together with the left-hand diagram in (3.4.4), a similar calculation to that preceding equation (3.3.4) will now establish that $\Theta^-_{L_d^N} = (i^{N\times N})^*\Theta^-_{L_d}$. Using the same idea for $\Theta^+_{L_d}$ and taking the exterior derivative of these expressions shows that the constrained and unconstrained discrete one- and two-forms are related by

$$\Theta^+_{L_d^N} = (i^{N\times N})^*\Theta^+_{L_d}, \qquad \Theta^-_{L_d^N} = (i^{N\times N})^*\Theta^-_{L_d}, \qquad \Omega_{L_d^N} = (i^{N\times N})^*\Omega_{L_d}.$$

Pushing all of these structures forward with the discrete Legendre transforms shows that the constrained discrete Hamiltonian map $\tilde{F}_{L_d^N}$, regarded as acting either on T^*N or $\eta(T^*N)$, preserves the canonical two-form Ω^N, while $\tilde{F}_{L_d}$ naturally preserves Ω.

If we further consider a symmetry action $\Phi : G \times Q \to Q$ which leaves N invariant, so that it covers an action $\Phi^N : G \times N \to N$, then the infinitesimal generators are related by

$$\xi_Q \circ i = Ti \circ \xi_N, \tag{3.4.11a}$$

$$\xi_{Q\times Q} \circ i^{N\times N} = T(i^{N\times N}) \circ \xi_{N\times N}. \tag{3.4.11b}$$

Using now the above relations between the constrained and unconstrained symplectic one-forms, we have that the momentum maps for the product action will be related by

$$J^+_{L_d^N} = J^+_{L_d} \circ i^{N\times N}, \tag{3.4.12a}$$

$$J^-_{L_d^N} = J^-_{L_d} \circ i^{N\times N}. \tag{3.4.12b}$$

If the group action is a symmetry of the Lagrangian then these momentum

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

maps are equal and Noether's theorem holds on both the constrained and unconstrained levels with this unique momentum map.

3.4.5. Constrained exact discrete Lagrangians

The exact discrete Lagrangian for a constrained system is not simply the standard exact discrete Lagrangian restricted to the constraint submanifold, as that would be the action along an unconstrained trajectory. Instead, the constrained exact discrete Lagrangian is the action of the constrained system, evaluated along the trajectory which lies on the constraint submanifold: that is,

$$L_d^{N,E}(q_0, q_1, h) = \int_0^h L^N(q_{0,1}(t), \dot{q}_{0,1}(t))\,\mathrm{d}t, \tag{3.4.13}$$

where $q : [0, h] \to N$ is the solution of the Euler–Lagrange equations for $L^N : TN \to \mathbb{R}$ which satisfies $q(0) = q_0$ and $q(h) = q_1$. As this discrete Lagrangian is defined on $N \times N \times \mathbb{R}$, it satisfies

$$\begin{aligned}\mathbb{F}^- L_d^{N,E}(q_0, q_1, h)) &= \mathbb{F}L^N(q_{0,1}(0), \dot{q}_{0,1}(0)),\\ \mathbb{F}^+ L_d^{N,E}(q_0, q_1, h)) &= \mathbb{F}L^N(q_{0,1}(h), \dot{q}_{0,1}(h)).\end{aligned}$$

We would like, however, to define a function on $Q \times Q \times \mathbb{R}$ whose restriction to $N \times N \times \mathbb{R}$ would give $L_d^{N,E}$. Without introducing additional structure, however, there is no canonical way to do so. Indeed, let $L_d^{Q,E} : Q \times Q \times \mathbb{R} \to \mathbb{R}$ be any smooth extension of $L_d^{N,E}$. Then from (3.3.3), (3.4.4) and the above relations we have immediately that

$$\begin{aligned}(T^*i)_{q_0}\big(\mathbb{F}^- L_d^{Q,E}(q_0, q_1, h)\big) &= (T^*i)_{q_0}\big(\mathbb{F}L(q_{0,1}(0), \dot{q}_{0,1}(0))\big),\\ (T^*i)_{q_1}\big(\mathbb{F}^+ L_d^{Q,E}(q_0, q_1, h)\big) &= (T^*i)_{q_1}\big(\mathbb{F}L(q_{0,1}(h), \dot{q}_{0,1}(h))\big),\end{aligned}$$

which is a constrained version of Lemma 1.6.2. The equivalence of the discrete and continuous systems now follows as in Section 1.6.

Note that this means that the order of accuracy of a discrete Lagrangian constrained to $N \times N$ will not, in general, be the same as the order of accuracy on $Q \times Q$: that is, if $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$ approximates the action on Q to some particular order, then the restriction $L_d^N = L_d|_{N \times N}$ will typically approximate the action of constrained solutions in N to some different order. Indeed, to derive high-order discrete Lagrangians for a constrained system, it is necessary to take account of the constraints in defining L_d^N, since a high-order L_d will typically restrict to only a first- or second-order L_d^N.

3.5. Constrained variational integrators

In this section we consider implementing the integration of a mechanical system with constraints. First we review standard geometric methods, and then we turn to variational integrators.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

3.5.1. Constrained geometric integration

There are a number of standard approaches to the numerical integration of constrained mechanical systems. These include working in local coordinates on the constraint submanifold (for example, see Bobenko and Suris (1999*b*) in the case of Lie groups, or Leimkuhler and Reich (1994)), solving a modified system on the containing space which has the constraint submanifold as a stable invariant set (for example, see Leimkuhler and Reich (1994)), and methods based in the containing space which explicitly enforce the constraints. Constrained mechanical systems are particular examples of differential algebraic systems, and many of the techniques for the numerical integration of such systems can also be applied in the mechanical setting (see Hairer and Wanner (1996) and Ascher and Petzold (1998)).

Unless the system under consideration has a particularly simple structure, working in local coordinates on the constraint submanifold suffers from a number of problems, including the fact that changing charts during the integration is not smooth, which breaks many of the nice properties of geometric integrators. In addition, local coordinate computations can be very expensive, and the equations can be very complicated, making the integrator difficult to code. For all of these reasons, it is often preferable to use integration techniques based on the containing space.

There are a number of different approaches to this, with representative samples being Gonzalez (1999), Seiler (1999, 1998*a*, 1998*b*), McLachlan and Scovel (1995) and Brasey and Hairer (1993). For a good overview of this area see Hairer (200x).

3.5.2. Variational integrators for constrained systems

Here we consider a constrained discrete Lagrangian system as an integrator for a continuous system. Given a continuous system $L : TQ \to \mathbb{R}$ and a constraint submanifold $N \subset Q$ defined by $N = \phi^{-1}(0)$ for some $\phi : Q \to \mathbb{R}^d$, we would like a discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$ so that its restriction to $N \times N \times \mathbb{R}$ approximates the exact constrained discrete Lagrangian (3.4.13). The order of this approximation is related to the order of the resulting integrator.

Given such an L_d, we can now use any of the equivalent formulations of the constrained Euler–Lagrange equations from Section 3.4.3 to obtain an integrator. As in the unconstrained case, we can regard such an integrator as defined on the product $N \times N$ or on the corresponding cotangent bundle, although the latter interpretation is typically simpler for implementation purposes.

To be explicit, we will henceforth assume that the given continuous system is regular, so that we have equivalent Lagrangian and Hamiltonian representations, and that the containing manifold Q is linear, so that it is isomorphic

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

to $\mathbb{R}^n$. We will use (3.4.9) to define the constrained discrete Hamiltonian map $\tilde{F}_{L_d^N}$ regarded as mapping $\eta(T^*N)$ to $\eta(T^*N)$, where we recall that $\eta(T^*N)$ is the embedding of T^*N in T^*Q defined by

$$\eta(T^*N) = \left\{(q,p) \in T^*Q \;\middle|\; \phi(q) = 0 \text{ and } \nabla\phi(q) \cdot \frac{\partial H}{\partial p}(q,p) = 0\right\}. \quad (3.5.1)$$

As we are now treating the discrete Lagrangian as the approximation to the exact system, it will be dependent upon a time-step h and thus have the form $L_d(q_0, q_1, h)$. Given this, we may rescale the Lagrange multipliers in (3.4.9) by h so that the constraint terms appear in the same way as discrete forces, allowing them to be interpreted as discrete forces of constraint. This gives

$$p_0 = -D_1 L_d(q_0, q_1) + h(\lambda^{(0)})^T \nabla\phi(q_0), \quad (3.5.2a)$$

$$p_1 = D_2 L_d(q_0, q_1) - h(\lambda^{(1)})^T \nabla\phi(q_1), \quad (3.5.2b)$$

$$\phi(q_1) = 0, \quad (3.5.2c)$$

$$\nabla\phi(q_1) \cdot \frac{\partial H}{\partial p}(q_1, p_1) = 0. \quad (3.5.2d)$$

To use these equations as an integrator, we must take an initial condition $(q_0, p_0) \in \eta(T^*N)$, so that q_0 and p_0 satisfy the conditions given by (3.5.1). The $2n + 2d$ system (3.5.2) must then be solved implicitly to find (q_1, p_1) and the accompanying Lagrange multipliers. Iterating this process gives the integrated trajectory.

Although this is generally the simplest way to implement a variational integrator, note that if the Lagrangian has a special form, such as being composed of kinetic and potential terms, then we could also use one of the other equivalent expressions of the discrete Hamiltonian map given previously. Alternatively, we could also choose to work directly on $N \times N$ and to use (3.4.1) as an integrator mapping each pair (q_k, q_{k+1}) to (q_{k+1}, q_{k+2}).

Using the above theory, we recall that any such methods will always be symplectic, and if the discrete Lagrangian inherits the symmetries of the continuous system, then the integrator will also conserve the corresponding momentum maps.

To implement a constrained variational integrator, it is of course necessary to choose a particular discrete Lagrangian. We give below a number of ways in which this can be done and we explicitly evaluate the defining equations (3.5.2) in several cases.

3.5.3. Low-order methods

Given a low-order discrete Lagrangian, such as L_d^α given in Example 2.3.2, one can simply restrict it to $N \times N$ to obtain an integrator for the constrained

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

system. As N will generally not be convex, the points $(1-\alpha)q_0+\alpha q_1$ will not be in N when q_0 and q_1 are. If the Lagrangian on N is the restriction of a smooth Lagrangian on Q, then this will not matter for sufficiently small stepsizes.

For a Lagrangian which is not defined off N, or which varies quickly compared to the stepsize, it is important to only evaluate L and its derivatives on N. Perhaps the simplest examples of such methods are given by L_d^0 and L_d^1, which give constrained versions of the symplectic Euler methods.

3.5.4. SHAKE and RATTLE

As we saw in Section 2.6.2, the Verlet algorithm is the discrete Lagrangian map $F_{L_d}: Q\times Q\to Q\times Q$ generated by the discrete Lagrangian

$$L_d(q_0,q_1,h)=\frac{1}{2}hL\left(q_0,\frac{q_1-q_0}{h}\right)+\frac{1}{2}hL\left(q_1,\frac{q_1-q_0}{h}\right), \tag{3.5.3}$$

where we assume that the continuous system has the form $L(q,\dot q)=\frac{1}{2}\dot q^T M\dot q - V(q)$. To form a constrained version of this method, we can simply restrict L_d to $N\times N$ and calculate the constrained discrete Euler–Lagrange equations (3.4.1). These give

$$\begin{aligned} M\left(\frac{q_{k+1}-2q_k+q_{k-1}}{h}\right)+h\nabla V(q_k)+(\lambda_k)^T\nabla\phi(q_k)&=0,\\ \phi(q_{k+1})&=0, \end{aligned}$$

which is known as the SHAKE algorithm. This was first proposed by Ryckaert, Ciccotti and Berendsen (1977) as a constrained version of Verlet.

A constrained version of the velocity Verlet integrator, RATTLE, was given by Anderson (1983). This was later shown by Leimkuhler and Skeel (1994) to be a symplectic integrator on T^*N. In fact, RATTLE *is simply the constrained discrete Hamiltonian map* $\tilde F_{L_d^N}: T^*N\to T^*N$ *associated to the discrete Lagrangian* (3.5.3). To see this, we calculate the coordinate expressions of (3.5.2) with $L(q,\dot q)=\frac{1}{2}\dot q^T M\dot q - V(q)$ to give

$$\begin{aligned} p_k &= M\left(\frac{q_{k+1}-q_k}{h}\right)+\frac{1}{2}h\nabla V(q_k)+(\lambda_k^{(0)})^T\nabla\phi(q_k),\\ p_{k+1} &= M\left(\frac{q_{k+1}-q_k}{h}\right)-\frac{1}{2}h\nabla V(q_{k+1})+(\lambda_k^{(1)})^T\nabla\phi(q_{k+1}),\\ 0 &= \phi(q_{k+1}),\\ 0 &= \nabla\phi(q_{k+1})M^{-1}p_{k+1}. \end{aligned}$$

Now we subtract the first equation from the second and solve the first equa-

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

tion for q_{k+1} to obtain

$$q_{k+1} = q_k + hM^{-1}p_k + \frac{1}{2}h^2M^{-1}(-\nabla V(q_k)) + \frac{1}{2}h^2M^{-1}(\lambda_k^{(0)})^T\nabla\phi(q_k),$$

$$p_{k+1} = p_k + h\left(\frac{-\nabla V(q_k) - \nabla V(q_{k+1})}{2}\right) + h\left(\frac{(\lambda_k^{(0)})^T\nabla\phi(q_k) + (\lambda_k^{(1)})^T\nabla\phi(q_{k+1})}{2}\right),$$

$$0 = \phi(q_{k+1}),$$

$$0 = \nabla\phi(q_{k+1})M^{-1}p_{k+1},$$

where we are assuming

$$(\nabla\phi)_{ij}(q) = \frac{\partial\phi^i}{\partial q^j}$$

and where we have scaled $\lambda_k^{(0)}$ and $\lambda_k^{(1)}$ by $-\frac{1}{2}$. This is exactly the RATTLE method.

This integrator is also the 2-stage member of the Lobatto IIIA–IIIB family (Jay 1996, 1999), which is discussed further below.

To summarize, the integrators known as Verlet, velocity Verlet, SHAKE and RATTLE are all derived from the discrete Lagrangian (3.5.3). Verlet is the discrete Lagrangian map $F_{L_d} : Q \times Q \to Q \times Q$, velocity Verlet is the discrete Hamiltonian map $\tilde{F}_{L_d} : T^*Q \to T^*Q$, SHAKE is the constrained discrete Lagrangian map $F_{L_d^N} : N \times N \to N \times N$, and RATTLE is the constrained discrete Hamiltonian map $\tilde{F}_{L_d^N} : T^*N \to T^*N$.

Thus, the variational formulation shows the natural connection between these methods, and proves in a unified way that they all conserve both the symplectic structure and quadratic momentum maps, as linear symmetries of V will be inherited by L_d.

3.5.5. Composition methods

To construct high-order integrators for a constrained system, a simple low-order constraint-preserving method can be used in a composition rule, as in Section 2.5 (Reich 1996). This approach has the advantage that the resulting method will inherit properties such as symplecticity from the base method, and will necessarily preserve the constraint.

Composing discrete Lagrangians extends directly to constrained systems. Given discrete Lagrangians L_d^i and time-step fractions γ^i for $i = 1, \ldots, s$, we can use any of the three interpretations of the composition L_d from Section 2.5. For the multiple steps method or the single step, multiple substeps method, the correct constraint to impose is that all the points q_k^i lie

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

on the constraint submanifold. This implies that the single step constrained composition discrete Lagrangian should be defined as

$$L_d(q_k, q_{k+1}, h) = \operatorname*{ext}_{q_k^i \in N} L_d(q_k, q_k^i, q_{k+1}, h),$$

which denotes the extreme value of the multipoint discrete Lagrangian over the set of interior points in the constraint submanifold N. The constrained discrete Hamiltonian map for this L_d will then be the composition of the constrained discrete Hamiltonian maps of the component L_d^i.

When composing non-self-adjoint methods, it is common to use a sequence including both the methods themselves and their adjoints. For this reason, it is worth noting that the adjoint of a constrained discrete Lagrangian is equal to the constrained version of the adjoint, that is, $(L_d^*)^N = (L_d^N)^*$. Furthermore, the associated constrained discrete Hamiltonian maps are adjoint as integrators.

3.5.6. Constrained symplectic partitioned Runge–Kutta methods

For a Hamiltonian system $H : T^*Q \to \mathbb{R}$ with holonomic constraint $\phi : Q \to \mathbb{R}^d$, a *constrained partitioned Runge–Kutta method* is a map $T^*N \to T^*N$ specified by $(q_0, p_0) \mapsto (q_1, p_1)$ where

$$q_1 = q_0 + h\sum_{j=1}^{s} b_j \dot{Q}_j, \quad p_1 = p_0 + h\sum_{j=1}^{s} \tilde{b}_j \dot{P}_j, \tag{3.5.4a}$$

$$Q_i = q_0 + h\sum_{j=1}^{s} a_{ij} \dot{Q}_j, \quad P_i = p_0 + h\sum_{j=1}^{s} \tilde{a}_{ij} \dot{P}_j, \qquad i = 1, \ldots, s, \tag{3.5.4b}$$

$$\dot{Q}_i = \frac{\partial H}{\partial p}(Q_i, P_i), \qquad \dot{P}_i = -\frac{\partial H}{\partial q}(Q_i, P_i) - \Lambda_i^T \nabla\phi(Q_i), \quad i = 1, \ldots, s, \tag{3.5.4c}$$

$$0 = \phi(Q_i), \qquad 0 = \nabla\phi(q_1) \cdot \frac{\partial H}{\partial p}(q_1, p_1), \qquad i = 1, \ldots, s. \tag{3.5.4d}$$

In addition, it is necessary to place some restrictions on the coefficients to ensure that these equations do in fact define a map on T^*N. We begin by imposing the requirement (2.6.7) of symplecticity to give

$$b_i \tilde{a}_{ij} + \tilde{b}_j a_{ji} = b_i \tilde{b}_j, \qquad i, j = 1, \ldots, s,$$
$$b_i = \tilde{b}_i, \qquad i = 1, \ldots, s.$$

We also require that the method be *stiffly accurate*: that is, $a_{si} = b_i$ for $i = 1, \ldots, s$. This means that $q_1 = Q_s$, and hence $q_1 \in N$. Further requiring that $b_i \neq 0$ for $i = 1, \ldots, s$ implies that $\tilde{a}_{is} = 0$ for each $i = 1, \ldots, s$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

To ensure that the system is not over-determined, we set $a_{1i} = 0$ for $i = 1, \ldots, s$ and so obtain $q_0 = Q_1$. Requiring that $b_i \neq 0$ for $i = 1, \ldots, s$ now implies that $\tilde{a}_{i1} = \tilde{b}_i$ for $i = 1, \ldots, s$. Given that we start from $(q_0, p_0) \in T^*N$ we thus have that $\phi(Q_1) = \phi(q_0) = 0$ is immediately satisfied.

With these restrictions, (3.5.4) is a system of $s(4n+d)+2n$ equations for the same number of unknowns, defining a map $\eta(T^*N) \to \eta(T^*N)$. It can be shown (Jay 1996) that this is a well-defined symplectic map on T^*N. Such methods are a particular example of the SPARK methods of Jay (1999), and the subset of these methods which are explicit have been analysed for constrained systems by Reich (1997).

To see how such constrained symplectic partitioned Runge–Kutta methods can be derived variationally, we proceed in a similar fashion to the unconstrained case in Section 2.6.5. Given $(q_0, q_1) \in Q \times Q$, we implicitly define $\bar{p}_0$, $\bar{p}_1$, $\bar{Q}_i$, $\bar{P}_i$, $\dot{\bar{Q}}_i$, $\dot{\bar{P}}_i$ for $i = 1, \ldots, s$, and $\bar{\Lambda}_i$ for $i = 2, \ldots, (s-1)$ by the equations

$$q_1 = q_0 + h \sum_{j=1}^{s} b_j \dot{\bar{Q}}_j, \qquad \bar{p}_1 = \bar{p}_0 + h \sum_{j=1}^{s} \tilde{b}_j \dot{\bar{P}}_j, \tag{3.5.5a}$$

$$\bar{Q}_i = q_0 + h \sum_{j=1}^{s} a_{ij} \dot{\bar{Q}}_j, \qquad \bar{P}_i = \bar{p}_0 + h \sum_{j=1}^{s} \tilde{a}_{ij} \dot{\bar{P}}_j, \quad i = 1, \ldots, s, \tag{3.5.5b}$$

$$\dot{\bar{Q}}_i = \frac{\partial H}{\partial p}(\bar{Q}_i, \bar{P}_i), \qquad i = 1, \ldots, s, \tag{3.5.5c}$$

$$\dot{\bar{P}}_i = -\frac{\partial H}{\partial q}(\bar{Q}_i, \bar{P}_i) - \bar{\Lambda}_i^T \nabla \phi(\bar{Q}_i), \quad 0 = \phi(\bar{Q}_i), \qquad i = 2, \ldots, (s-1), \tag{3.5.5d}$$

$$\dot{\bar{P}}_1 = -\frac{\partial H}{\partial q}(\bar{Q}_1, \bar{P}_1), \qquad \dot{\bar{P}}_s = -\frac{\partial H}{\partial q}(\bar{Q}_s, \bar{P}_s). \tag{3.5.5e}$$

This is a system of $4sn + (s-2)d$ equations in the same number of variables and the restrictions on the coefficients ensure that it will have a solution for sufficiently small h.

This subset of the equations (3.5.4) was chosen from the fact that $\bar{Q}_1 = q_0$ and $\bar{Q}_s = q_1$, so it is necessary to relax the constraints on these two points. Having done so, the same number of Lagrange multipliers must also then be disregarded. Given these definitions of the various quantities in terms of q_0 and q_1 we define the discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$ by

$$L_d(q_0, q_1, h) = h \sum_{i=1}^{s} b_i L(\bar{Q}_i, \dot{\bar{Q}}_i), \tag{3.5.6}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

where we assume that the coefficients satisfy all of the previous requirements. For a given continuous system (L or H) this is not the same as the corresponding expression (2.6.8) in the unconstrained case, as the equations defining $\bar{Q}_i$ and $\dot{\bar{Q}}_i$ have been modified here to take account of the constraints. We now show that the constrained discrete Hamiltonian map corresponding to (3.5.6) is indeed the constrained symplectic partitioned Runge–Kutta method.

Theorem 3.5.1. The constrained discrete Hamiltonian map for the discrete Lagrangian (3.5.6) is exactly the integrator defined by the constrained symplectic partitioned Runge–Kutta equations (3.5.4).

Proof. Differentiating $\phi(\bar{Q}_i) = 0$ for $i = 2, \ldots, s-1$ gives

$$\nabla\phi(\bar{Q}_i) \cdot \frac{\partial \bar{Q}_i}{\partial q_0} = 0, \qquad i = 2, \ldots, s-1,$$

and using this together with the definitions (3.5.5) and the same argument as in Theorem 2.6.1 shows that

$$\frac{\partial L_d}{\partial q_0} = -\bar{p}_0, \qquad \frac{\partial L_d}{\partial q_1} = \bar{p}_1.$$

We now consider a given initial condition $(q_0, p_0) \in T^*N$ and recall that the discrete Hamiltonian map will give $(q_1, p_1) \in T^*N$ which satisfy (3.5.2). To see the relation of this mapping to the symplectic partitioned Runge–Kutta map, we make the following change of variables:

$$\begin{aligned}
Q_i &= \bar{Q}_i, & P_i &= \bar{P}_i, & &i = 1, \ldots, s,\\
\dot{Q}_i &= \dot{\bar{Q}}_i, & & & &i = 1, \ldots, s,\\
\Lambda_i &= \bar{\Lambda}_i, & \dot{P}_i &= \dot{\bar{P}}_i, & &i = 2, \ldots, s-1,\\
\tilde{b}_1\Lambda_1 &= \lambda^{(0)}, & \dot{P}_1 &= \dot{\bar{P}}_1 - \Lambda_1^T\nabla\phi(Q_1), & &\\
\tilde{b}_s\Lambda_s &= \lambda^{(1)}, & \dot{P}_s &= \dot{\bar{P}}_s - \Lambda_s^T\nabla\phi(Q_s). & &
\end{aligned}$$

Recalling that the coefficients are such that $Q_1 = q_0$ and $Q_s = q_1$, we now see that (3.5.2c) and (3.5.2d), together with the restrictions (3.5.5d) on $\bar{Q}_i$, give the conditions (3.5.4d) on the non-overbar quantities.

Furthermore, (3.5.2a) and (3.5.2b) give

$$\begin{aligned}
\bar{p}_0 &= p_0 - h\tilde{b}_1\Lambda_1^T\nabla\phi(Q_1),\\
\bar{p}_1 &= p_1 + h\tilde{b}_s\Lambda_s^T\nabla\phi(Q_s).
\end{aligned}$$

Substituting these definitions into the equations (3.5.5) and using the fact that $\tilde{a}_{is} = 0$ and $\tilde{a}_{i1} = \tilde{b}_i$ for $i = 1, \ldots, s$ now shows that the non-overbar quantities satisfy (3.5.4a), (3.5.4b) and (3.5.4c). We thus have that the

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

discrete Hamiltonian map $(q_0, p_0) \mapsto (q_1, p_1)$ on $\eta(T^*N)$ is identical to the constrained symplectic Runge–Kutta map. □

3.5.7. Constrained Galerkin methods

With the insight gained from the definition of the constrained exact discrete Lagrangian (3.4.13) it is simple to extend the Galerkin discrete Lagrangians of Section 2.6.6 to include holonomic constraints.

In the particular example of polynomial trajectory approximations and numerical quadrature, the definition (2.6.10) of the Galerkin discrete Lagrangian should be modified to

$$L_d(q_0, q_1, h) = \operatorname*{ext}_{\substack{q \in \mathcal{C}^s([0,h],Q) \\ \phi(q(c_i h))=0}} \mathfrak{G}^s(q), \tag{3.5.7}$$

where $\phi : Q \to \mathbb{R}$ is the constraint function. This constrains the intermediate trajectories to intersect the constraint submanifold at each quadrature point. For such methods it is typically reasonable to require that $c_0 = 0$ and $c_s = 1$, so that the endpoints q_0 and q_1 also satisfy the constraint.

Evaluating the constrained discrete Euler–Lagrange equations for (3.5.7) shows that the associated discrete Hamiltonian map is a constrained symplectic partitioned Runge–Kutta method, in the sense of the preceding section and of Jay (1999). In particular, choosing the quadrature rule to be Lobatto quadrature results in the constrained Lobatto IIIA–IIIB method of Jay (1999).

3.6. Background: Forced and constrained systems

We now consider Lagrangian and Hamiltonian systems with *both* external forcing and holonomic constraints. The formulations and equations for such systems are straightforward combinations of the material in the preceding sections for systems with only forces or only constraints. For this reason, we will simply state the results without proof.

As before, we assume that we have a system on the unconstrained configuration manifold Q, and a holonomic constraint function $\phi : Q \to \mathbb{R}^d$ so that the constraint manifold is $N = \phi^{-1}(0) \subset Q$. The inclusion map is denoted $i : N \to Q$, and we have the natural lifts $Ti : TN \to TQ$ and $T^*i : T^*Q \to T^*N$.

3.6.1. Lagrangian systems

Given a Lagrangian force $f_L : TQ \to T^*Q$, we restrict it to $f_L^N = T^*i \circ f_L \circ Ti : TN \to T^*N$, which is then a Lagrangian force on TN. Taking the Lagrange–d'Alembert principle and restricting to the space of constrained curves gives the following theorem.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Theorem 3.6.1. Given a Lagrangian system $L : TQ \to \mathbb{R}$ with Lagrangian force $f_L : TQ \to T^*Q$ and holonomic constraint $\phi : Q \to \mathbb{R}^d$, set $N = \phi^{-1}(0) \subset Q$, $f_L^N = T^*i \circ f_L \circ Ti$, and $L^N = L|_{TN}$. Then the following are equivalent:

(1) $q \in \mathcal{C}(N)$ satisfies the Lagrange–d'Alembert principle for L^N and f_L^N and hence solves the forced Euler–Lagrange equations;

(2) $q \in \mathcal{C}(Q)$ and $\lambda \in \mathcal{C}(\mathbb{R}^d)$ satisfy the *forced constrained Euler–Lagrange equations*

$$\frac{\partial L}{\partial q^i}(q(t), \dot{q}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}(q(t), \dot{q}(t))\right) + f_L(q(t), \dot{q}(t)) = \left\langle \lambda(t), \frac{\partial \phi}{\partial q^i}(q(t)) \right\rangle, \quad (3.6.1a)$$

$$\phi(q(t)) = 0; \quad (3.6.1b)$$

(3) $(q, \lambda) \in \mathcal{C}(Q \times \mathbb{R}^d)$ satisfies the Lagrange–d'Alembert principle, and hence solves the forced Euler–Lagrange equations, for $\bar{L} : T^*(Q \times \mathbb{R}^d) \to \mathbb{R}$ and $\bar{f}_L : T(Q \times \mathbb{R}^d) \to T^*(Q \times \mathbb{R}^d)$ defined by

$$\bar{L}(q, \lambda, \dot{q}, \dot{\lambda}) = L(q, \dot{q}) - \langle \lambda, \phi(q) \rangle,$$
$$\bar{f}_L(q, \lambda, \dot{q}, \dot{\lambda}) = \pi_Q^* \circ f_L(q, \dot{q}),$$

where $\pi_Q : Q \times \mathbb{R}^d \to Q$ is the projection.

One can also project (3.6.1a) with $T^*i : T^*Q \to T^*N$ to obtain a system without λ, as in Section 3.3.

Observe that in the forced constrained Euler–Lagrange equations (3.6.1) the forcing and Lagrange multiplier terms enter in same way. For this reason, the Lagrange multiplier term is sometimes referred to as the *forces of constraint*, and we can regard it as being a force which is constructed exactly so that the solution is kept on the constraint submanifold N.

3.6.2. Hamiltonian systems

Following the development of the unforced constrained case, we can move to the Hamiltonian framework by either taking the Legendre transform of the degenerate augmented system, or by working directly on T^*N.

The former approach takes a Hamiltonian force $f_H : T^*Q \to T^*Q$ and forms the augmented Hamiltonian force $\bar{f}_H : T^*(Q \times \mathbb{R}^d) \to T^*(Q \times \mathbb{R}^d)$ by $\bar{f}_H(q, \lambda, p, \pi) = \pi_Q^* \circ f_H(q, p)$. The forced constrained Hamiltonian vector field $\bar{X}_H$ on the primary constraint set Π is defined by

$$\mathbf{i}_{\bar{X}_H} \Omega^{\Pi} = \mathbf{d}\bar{H} - \bar{f}'_H$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

where $\bar{H}$ and Ω^{Π} are as before, and $\bar{f}'_H$ is the horizontal one-form on $T^*(Q \times \mathbb{R}^d)$ corresponding to $\bar{f}_H$. In coordinates this gives the *forced constrained Hamilton equations*

$$X_{q^i}(q, \lambda, p, \pi) = \frac{\partial H}{\partial p^i},$$

$$X_{p^i}(q, \lambda, p, \pi) = -\frac{\partial H}{\partial q^i} + f_H(q, p) - \left\langle \lambda, \frac{\partial \phi}{\partial q^i}(q) \right\rangle,$$

$$\phi(q) = 0.$$

Alternatively, we can directly relate the unconstrained Hamiltonian system to the constrained system as in Section 3.3.3. To do this, we must choose a symplectic embedding $\eta : T^*N \to T^*Q$, which we will assume covers the embedding $i : N \to Q$. Given such a map, we now define the *constrained Hamiltonian force* $f_H^N : T^*N \to T^*N$ by $f_H^N = T^*i \circ f_H \circ \eta$ and we let $f_H^{N\prime}$ be the corresponding horizontal one-form on T^*N. We assume that all other structures are as in Section 3.3.3, so that the constrained Hamiltonian is $H^N = H \circ \eta$.

The forced constrained Hamiltonian vector field X_{H^N} and the forced unconstrained Hamiltonian vector field X_H are now defined by

$$\mathbf{i}_{X_{H^N}} \Omega^N = \mathbf{d}H^N - f_H^{N\prime},$$

$$\mathbf{i}_{X_H} \Omega = \mathbf{d}H - f'_H.$$

Denoting the Ω-orthogonal projection to $\eta(T^*N)$ by $\pi_\Omega : T^*Q \to T^*N$, we can show that the projection of the forced unconstrained vector field is just the forced constrained vector field.

Theorem 3.6.2. Consider a Hamiltonian system $H : T^*Q \to \mathbb{R}$ with forcing $f_H : T^*Q \to T^*Q$ and constraint submanifold $N \subset Q$ and let the constrained system $H^N : T^*N \to \mathbb{R}$ and $f_H^N : T^*N \to T^*N$ be defined as above. Then $X_{H^N} = \pi_\Omega \cdot X_H \circ \eta$.

Proof. We can use essentially the same proof as for Theorem 3.3.3 in the unforced case. The only additional requirement is to check that the one-form $f_H^{N\prime}$ is the pullback under η of f'_H, so that $f'_H(\eta(p_q)) \cdot T\eta \cdot V^N = f_H^{N\prime}(p_q) \cdot V^N$.

To see this, we recall that η covers the identity and so $\pi_Q \circ \eta = i \circ \pi_N$. Using the derivative of this expression we calculate

$$f_H^{N\prime}(p_q) \cdot V^N = \langle T^*i \circ f_H \circ \eta(p_q), T\pi_N \cdot V^N \rangle$$

$$= \langle f_H \circ \eta(p_q), Ti \circ T\pi_N \cdot V^N \rangle$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

$$= \langle f_H \circ \eta(p_q), T\pi_Q \circ T\eta \cdot V^N \rangle$$
$$= \big(\eta^*(f'_H)\big)(p_q) \cdot V^N,$$

which can then be used to modify the proof of Theorem 3.3.3, to obtain the desired result. □

3.6.3. Legendre transforms

Given a regular Lagrangian system and the corresponding regular Hamiltonian system, we have seen in Section 3.3.4 that the standard Legendre transforms provide a canonical way to construct a map $\eta : T^*N \to T^*Q$ and so to regard T^*N as a submanifold of T^*Q.

Furthermore, as we saw in Section 3.1.3, the forced Lagrangian and Hamiltonian vector fields are related by the standard Legendre transform, so this will hold for both the constrained and unconstrained systems. Note that our definitions of constrained Lagrangian and Hamiltonian forces commute with the Legendre transform, so that if $f_L = f_H \circ \mathbb{F}L$ then $f_L^N = f_H^N \circ \mathbb{F}L^N$. This can be seen by recalling that $\eta \circ \mathbb{F}L^N = \mathbb{F}L \circ Ti$ and using the definitions of the constrained forces.

We thus have that the constrained and unconstrained forced vector fields on both the Lagrangian and Hamiltonian sides are related by projection and Legendre transforms, which fully commute. In particular, we can write the projected vector field on the Hamiltonian side in coordinates to give

$$\dot{q} = \frac{\partial H}{\partial p},$$
$$\dot{p} = -\frac{\partial H}{\partial q} - \lambda^T \nabla\phi(q) + f_H(q,p),$$
$$\phi(q) = 0.$$

In the special case when the Hamiltonian depends quadratically on p then this projection is induced by the metric given on T^*Q by the kinetic energy, as in Section 3.3.4 above.

3.6.4. Conservation properties

Given a group action $\Phi : G \times Q \to Q$, we have seen in Section 3.4.4 that if Φ leaves N invariant then it can be restricted to an action Φ^N on N and the infinitesimal generators of this restricted action are related by projection to the generators of the action on Q. This then shows that the momentum maps of the constrained systems are just the appropriate restrictions of the unconstrained momentum maps.

In addition, from Section 3.1.4 we know that if the Lagrangian is invariant under the group action and the forces are orthogonal to the action, then

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Noether's theorem will still hold. In the constrained setting, observe that we have

$$\begin{aligned}\left\langle f_L^N(v_q), \xi_N(q)\right\rangle &= \left\langle T^*i \circ f_L \circ Ti(v_q), \xi_N(q)\right\rangle \\ &= \left\langle f_L \circ Ti(v_q), Ti \cdot \xi_N(q)\right\rangle \\ &= \left\langle f_L \circ Ti(v_q), \xi_Q \circ i(q)\right\rangle,\end{aligned}$$

and so if f_L is orthogonal to ξ_Q then the constrained force f_L^N will also be orthogonal to the constrained infinitesimal generator ξ_N. This gives us the following Noether's theorem.

Theorem 3.6.3. (Forced constrained Noether's theorem) Consider a Lagrangian system $L : TQ \to \mathbb{R}$ with constraint submanifold $N \subset Q$, forcing $f_L : TQ \to T^*Q$ and a symmetry action $\Phi : G \times Q \to Q$ such that $\langle f_L(q, \dot{q}), \xi_Q(q)\rangle = 0$ for all $(q, \dot{q}) \in TQ$ and $\xi \in \mathfrak{g}$. Then the constrained Lagrangian momentum map $J_{L^N} : TN \to \mathfrak{g}^*$ will be preserved by the forced constrained Lagrangian flow.

Of course, it is only necessary that the constrained force be orthogonal to the group action on the constraint submanifold and that the reduced action be a symmetry of the constrained Lagrangian. The above theorem simply gives sufficient conditions for this in terms of the unconstrained quantities.

3.7. Discrete variational mechanics with forces and constraints

We now combine the previous results for forced and constrained systems to consider discrete Lagrangian systems with *both* forcing and constraints. The definitions and results are the expected combinations of the special cases of only forcing or only constraints, and so we will not give detailed proofs.

3.7.1. Lagrangian viewpoint

Given discrete Lagrangian forces $f_d^+, f_d^- : Q \times Q \to T^*Q$, we form the restrictions $f_d^{N+}, f_d^{N-} : N \times N \to T^*N$ by $f_d^{N\pm} = T^*i \circ f_d^{\pm} \circ i^{N\times N}$, which are then discrete Lagrangian forces on N. As in the continuous Lagrangian case, we now take the discrete Lagrange–d'Alembert principle from Section 3.4 and constrain it to N, thus obtaining the following theorem.

Theorem 3.7.1. Given discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$ with discrete Lagrangian forces $f_d^+, f_d^- : Q \times Q \to T^*Q$ and holonomic constraint $\phi : Q \to \mathbb{R}^d$, set $N = \phi^{-1}(0) \subset Q$, $f_d^{N\pm} = T^*i \circ f_d^{\pm} \circ i^{N\times N}$, and $L_d^N = L_d|_{Q\times Q}$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Then the following are equivalent:

(1) $q_d = \{q_k\}_{k=0}^N \in \mathcal{C}_d(N)$ satisfies the discrete Lagrange–d'Alembert principle for L_d^N, f_d^{N+} and f_d^{N-}, and hence solves the forced discrete Euler–Lagrange equations;

(2) $q_d = \{q_k\}_{k=0}^N \in \mathcal{C}_d(Q)$ and $\lambda_d = \{\lambda_k\}_{k=1}^{N-1} \in \mathcal{C}_d(\mathbb{R}^d)$ satisfy the *forced constrained discrete Euler–Lagrange equations*

$$D_2L_d(q_{k-1},q_k) + D_1L_d(q_k,q_{k+1}) + f_d^+(q_{k-1},q_k) + f_d^-(q_k,q_{k+1}) = \langle \lambda_k, \nabla\phi(q_k)\rangle\,, \quad (3.7.1a)$$

$$\phi(q_k) = 0; \quad (3.7.1b)$$

(3) $(q_d,\lambda_d) = \{(q_k,\lambda_k)\}_{k=0}^N \in \mathcal{C}_d(Q\times\mathbb{R}^d)$ satisfies the discrete Lagrange–d'Alembert principle, and hence solves the forced discrete Euler–Lagrange equations, for either of $\bar{L}_d^+, \bar{L}_d^- : (Q\times\mathbb{R}^d)\times(Q\times\mathbb{R}^d) \to \mathbb{R}$ defined by

$$\bar{L}_d^+(q_k,\lambda_k,q_{k+1},\lambda_{k+1}) = L_d(q_k,q_{k+1}) - \langle \lambda_{k+1}, \phi(q_{k+1})\rangle\,,$$
$$\bar{L}_d^-(q_k,\lambda_k,q_{k+1},\lambda_{k+1}) = L_d(q_k,q_{k+1}) - \langle \lambda_k, \phi(q_k)\rangle\,,$$

with the discrete Lagrangian forces $\bar{f}_d^+, \bar{f}_d^- : (Q\times\mathbb{R}^d)\times(Q\times\mathbb{R}^d) \to T^*(Q\times\mathbb{R}^d)$ defined by

$$\bar{f}_d^+(q_k,\lambda_k,q_{k+1},\lambda_{k+1}) = \pi_Q^* \circ f_d^+(q_k,q_{k+1}),$$
$$\bar{f}_d^-(q_k,\lambda_k,q_{k+1},\lambda_{k+1}) = \pi_Q^* \circ f_d^-(q_k,q_{k+1}),$$

where $\pi_Q : Q\times\mathbb{R}^d \to Q$ is the projection.

Using the canonical projection operator $T^*i : T^*Q \to T^*N$, we can also write (3.7.1) without the Lagrange multipliers.

3.7.2. Discrete Hamiltonian maps

We first consider the augmented approach to constructing a discrete Hamiltonian map, despite the lack of regularity. The *forced augmented discrete Hamiltonian map* $\tilde{F}_{\bar{L}_d^-} : (q_0,\lambda_0,p_0,\pi_0) \mapsto (q_1,\lambda_1,p_1,\pi_1)$ is defined by the equations

$$p_0 = -D_1L_d(q_0,q_1) - f_d^-(q_0,q_1) + \langle \lambda_0, \nabla\phi(q_0)\rangle\,, \quad (3.7.2a)$$

$$\pi_0 = \phi(q_0), \quad (3.7.2b)$$

$$p_1 = D_2L_d(q_0,q_1) + f_d^+(q_0,q_1), \quad (3.7.2c)$$

$$\pi_1 = 0. \quad (3.7.2d)$$

Restricting to the primary constraint set $\Pi \subset T^*(Q\times\mathbb{R}^d)$ now shows that

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

these equations are equivalent to the forced constrained discrete Euler–Lagrange equations (3.7.1) together with the constraint $\phi(q_k) = 0$. As before, the evolution of λ is not well-defined.

Rather than considering the augmented systems, we can also directly relate the constrained and unconstrained systems. Here we must use the forced discrete Legendre transforms (3.2.4), which we recall are

$$\begin{aligned}\mathbb{F}^{f+}L_d : (q_0, q_1) &\mapsto (q_1, p_1) = (q_1, D_2L_d(q_0, q_1) + f_d^+(q_0, q_1)),\\ \mathbb{F}^{f-}L_d : (q_0, q_1) &\mapsto (q_0, p_0) = (q_0, -D_1L_d(q_0, q_1) - f_d^-(q_0, q_1)).\end{aligned}$$

These depend on both the discrete Lagrangian and discrete forces. From (3.4.4) we have the relations

$$\begin{aligned}D_2L_d^N &= T^*i \circ D_2L_d \circ i^{N\times N},\\ -D_1L_d^N &= T^*i \circ (-D_1L_d) \circ i^{N\times N},\end{aligned}$$

and, combining these with the definitions of the constrained discrete forces f_d^{N+} and f_d^{N-}, we have the following commutative diagrams, where the discrete Legendre transforms are those which include the forcing.

$$\begin{array}{ccc} T^*Q|_N & \xleftarrow{\mathbb{F}^{f-}L_d} & Q\times Q|_{\{q_0\in N\}} \\ {\scriptstyle T^*i}\downarrow & & \uparrow{\scriptstyle i^{N\times N}} \\ T^*N & \xleftarrow[\mathbb{F}^{f-}L_d^N]{} & N\times N \end{array} \qquad \begin{array}{ccc} Q\times Q|_{\{q_1\in N\}} & \xrightarrow{\mathbb{F}^{f+}L_d} & T^*Q|_N \\ {\scriptstyle i^{N\times N}}\uparrow & & \downarrow{\scriptstyle T^*i} \\ N\times N & \xrightarrow[\mathbb{F}^{f+}L_d^N]{} & T^*N \end{array} \tag{3.7.3}$$

This is the equivalent of (3.4.4) in the unforced case, and using this we now have the equivalent of diagram (3.4.5) for the forced discrete Legendre transforms, proving the following theorem.

Theorem 3.7.2. Consider a regular discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$ with constraint submanifold $N \subset Q$ and forcing $f_d^+, f_d^- : Q \times Q \to T^*Q$. Then the forced constrained discrete Hamiltonian map $\tilde{F}_{L_d^N} : T^*N \to T^*N$ has the following equivalent formulations:

(1) $\tilde{F}_{L_d^N} : (q_0, p_0) \mapsto (q_1, p_1)$ for $(q_0, p_0), (q_1, p_1) \in T^*N$ satisfying

$$p_0 = -D_1L_d^N(q_0, q_1) - f_d^{N-}(q_0, q_1), \tag{3.7.4a}$$
$$p_1 = D_2L_d^N(q_0, q_1) + f_d^{N+}(q_0, q_1); \tag{3.7.4b}$$

(2) $\tilde{F}_{L_d^N} : (q_0, p_0) \mapsto (q_1, p_1)$ for $(q_0, p_0), (q_1, p_1) \in T^*N$ satisfying

$$p_0 = (T^*i)_{q_0}\big((-D_1L_d - f_d^-) \circ i^{N\times N}(q_0, q_1)\big), \tag{3.7.5a}$$
$$p_1 = (T^*i)_{q_1}\big((D_2L_d + f_d^+) \circ i^{N\times N}(q_0, q_1)\big); \tag{3.7.5b}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

(3) $\tilde{F}_{L_d^N} : \eta(T^*N) \mapsto \eta(T^*N)$ for $(q_0, p_0) \in \eta(T^*N)$ and $(q_1, p_1) \in T^*Q$ satisfying

$$p_0 = \mathbb{P}_{q_0}\big((-D_1 L_d - f_d^-) \circ i^{N\times N}(q_0, q_1)\big), \tag{3.7.6a}$$

$$p_1 = \mathbb{P}_{q_1}\big((D_2 L_d + f_d^+) \circ i^{N\times N}(q_0, q_1)\big), \tag{3.7.6b}$$

$$\phi(q_1) = 0. \tag{3.7.6c}$$

Here $\eta : T^*N \to T^*Q$ is any symplectic embedding covering the identity, so that $\pi_Q \circ \eta = i \circ \pi_N$, and $\mathbb{P} : T^*Q|_N \to \eta(T^*N)$ is the map defined by $\mathbb{P} = \eta \circ T^*i$.

These equations are clearly the combination of the constrained equations from Theorem 3.4.2 with the forced equations (3.2.5).

Now assume that η is constructed from the Legendre transforms of some regular Lagrangian according to (3.3.7). Introducing Lagrange multipliers allows us to rewrite (3.7.6) as

$$p_0 = -D_1 L_d(q_0, q_1) - f_d^-(q_0, q_1) + (\lambda^{(0)})^T \nabla\phi(q_0), \tag{3.7.7a}$$

$$p_1 = D_2 L_d(q_0, q_1) + f_d^+(q_0, q_1) - (\lambda^{(1)})^T \nabla\phi(q_1), \tag{3.7.7b}$$

$$\phi(q_1) = 0, \tag{3.7.7c}$$

$$\nabla\phi(q_1) \cdot \frac{\partial H}{\partial p}(q_1, p_1) = 0, \tag{3.7.7d}$$

where (q_0, p_0) are in $\eta(T^*N)$. As before, we have chosen the signs on the Lagrange multipliers to correspond with the conventions of the discrete forces.

This form of the forced constrained discrete Hamiltonian map shows clearly that one can interpret the Lagrange multiplier terms as *discrete forces of constraint*. That is, the additional terms due to the constraints enter the equations in exactly the same way as the forcing terms. Indeed, the constraint terms can be regarded as forces which have exactly the correct action to keep the discrete trajectory on the constraint submanifold N.

If we are working with a particular form of Lagrangian, such as one involving a quadratic kinetic energy, then we can explicitly write the projection form of the discrete Hamiltonian map as was done in Section 3.4.3.

3.7.3. Exact forced constrained discrete Lagrangian

Given a Lagrangian system with forces and constraints, we can combine the ideas from Sections 3.2.4 and 3.4.5 to define the appropriate exact discrete Lagrangian and exact discrete forces.

Begin by considering the constrained system $L^N : TN \to \mathbb{R}$ with constrained force $f_L^N : TN \to T^*N$. Recall that the exact forced discrete

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Lagrangian $L_d^{N,E} : N \times N \times \mathbb{R}$ is the action (3.2.7a) along a solution of the forced Euler–Lagrange equations, and that the exact discrete forces $f_d^{N,E+}, f_d^{N,E-} : N \times N \times \mathbb{R} \to T^*N$ are the integrals of the forces (3.2.7b), (3.2.7c) along the variations of such a solution.

Having constructed these functions on $N \times N \times \mathbb{R}$, we take any smooth extension to functions $L_d^{Q,E} : Q \times Q \times \mathbb{R}$ and $f_d^{Q,E+}, f_d^{Q,E-} : Q \times Q \times \mathbb{R} \to \mathbb{R}$, as in Section 3.4.5. The same argument as used there now shows that

$$(T^*i)_{q_0}\big(\mathbb{F}^{f-}L_d^{Q,E}(q_0,q_1,h)\big) = (T^*i)_{q_0}\big(\mathbb{F}L(q_{0,1}(0),\dot{q}_{0,1}(0))\big),$$
$$(T^*i)_{q_1}\big(\mathbb{F}^{f+}L_d^{Q,E}(q_0,q_1,h)\big) = (T^*i)_{q_1}\big(\mathbb{F}L(q_{0,1}(h),\dot{q}_{0,1}(h))\big),$$

for all $q_0, q_1 \in N$ and the corresponding solutions $q : [0,h] \to N$ of the forced constrained Euler–Lagrange equations.

Using the above definitions, it is clear that to derive high-order discrete Lagrangians and discrete forces in the presence of constraints, both the discrete Lagrangian and the discrete forces will have to depend upon the continuous Lagrangian, the continuous forces and also the constraints. We will see examples of this below.

3.7.4. Noether's theorem

Consider a group action $\Phi : G \times Q \to Q$ and assume that it leaves N invariant, so that it restricts to $\Phi^N : G \times N \to N$. In the presence of forcing we saw in Section 3.2.3 that it is necessary to use the forced Legendre transforms to define the discrete momentum maps by (3.2.6). For the unconstrained system this gives

$$J_{L_d}^{f+}(q_0,q_1)\cdot\xi = \big\langle \mathbb{F}^{f+}L_d(q_0,q_1), \xi_Q(q_1)\big\rangle, \tag{3.7.8a}$$
$$J_{L_d}^{f-}(q_0,q_1)\cdot\xi = \big\langle \mathbb{F}^{f-}L_d(q_0,q_1), \xi_Q(q_1)\big\rangle, \tag{3.7.8b}$$

while the constrained forced momentum maps are

$$J_{L_d^N}^{f+}(q_0,q_1)\cdot\xi = \big\langle \mathbb{F}^{f+}L_d^N(q_0,q_1), \xi_N(q_1)\big\rangle, \tag{3.7.9a}$$
$$J_{L_d^N}^{f-}(q_0,q_1)\cdot\xi = \big\langle \mathbb{F}^{f-}L_d^N(q_0,q_1), \xi_N(q_1)\big\rangle. \tag{3.7.9b}$$

Recalling that the forced discrete Legendre transforms satisfy (3.7.3), we can use the relations (3.4.11) between the constrained and unconstrained infinitesimal generators to show that

$$J_{L_d^N}^{f+} = J_{L_d}^{f+} \circ i^{N\times N}, \tag{3.7.10a}$$
$$J_{L_d^N}^{f-} = J_{L_d}^{f-} \circ i^{N\times N}, \tag{3.7.10b}$$

which is the forced equivalent of (3.4.12). If the group action is a symmetry of the discrete Lagrangian then these momentum maps will be equal. In general Noether's theorem does not hold in the presence of forcing, except in

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

the special case when the forces are orthogonal to the group action. We will now see how this occurs in the presence of constraints.

Recall that, given discrete forces f_d^+ and f_d^-, we can construct a one-form f_d on $Q \times Q$ by (3.2.1), which gives

$$f_d^N(q_0, q_1) \cdot (\delta q_0, \delta q_1) = f_d^{N+}(q_0, q_1) \cdot \delta q_1 + f_d^{N-}(q_0, q_1) \cdot \delta q_0,$$
$$f_d(q_0, q_1) \cdot (\delta q_0, \delta q_1) = f_d^+(q_0, q_1) \cdot \delta q_1 + f_d^-(q_0, q_1) \cdot \delta q_0,$$

and so we have the relation $f_d^N = T^*(i^{N\times N}) \circ f_d \circ i^{N\times N}$. Using this, we compute

$$\begin{aligned}\left\langle f_d^N(q_0, q_1), \xi_{N\times N}(q_0, q_1)\right\rangle &= \left\langle T^*(i^{N\times N}) \circ f_d \circ i^{N\times N}(q_0, q_1), \xi_{N\times N}(q_0, q_1)\right\rangle \\ &= \left\langle f_d \circ i^{N\times N}(q_0, q_1), T(i^{N\times N}) \circ \xi_{N\times N}(q_0, q_1)\right\rangle \\ &= \left\langle f_d \circ i^{N\times N}(q_0, q_1), \xi_{Q\times Q} \circ i^{N\times N}(q_0, q_1)\right\rangle,\end{aligned}$$

where we have used the fact that $\xi_{Q\times Q} \circ i^{N\times N} = T(i^{N\times N}) \circ \xi_{N\times N}$. This shows that if f_d is orthogonal to $\xi_{Q\times Q}$, so that $\langle f_d, \xi_{Q\times Q}\rangle = 0$, then f_d^N will be orthogonal to $\xi_{N\times N}$. We thus have a Noether's theorem in this case.

Theorem 3.7.3. (Discrete forced constrained Noether's theorem) Consider a discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$ with constraint submanifold $N \subset Q$, discrete forces $f_d^+, f_d^- : Q \times Q \to T^*Q$ and a symmetry action $\Phi : G \times Q \to Q$ such that $\langle f_d, \xi_{Q\times Q}\rangle = 0$ for all $\xi \in \mathfrak{g}$. Then the constrained Lagrangian momentum map $J^f_{L_d^N} : N \times N \to \mathfrak{g}^*$ is preserved by the forced constrained discrete Hamiltonian map.

As in the continuous case with forcing and constraints, this only provides a sufficient condition as it is enough to just have orthogonality and invariance on N.

3.7.5. Variational integrators with forces and constraints

Consider a Lagrangian system $L : TQ \to \mathbb{R}$ with a constraint submanifold $N \subset Q$ specified by $N = \phi^{-1}(0)$ for some $\phi : Q \to \mathbb{R}^d$ and a Lagrangian force $f_L : TQ \to T^*Q$. We would now like to construct a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$ and discrete forces $f_d^+, f_d^- : Q \times Q \to T^*Q$ which approximate an extension of the exact discrete Lagrangian and exact forces. The discrete Hamiltonian map will then be an integrator for the continuous system.

We will assume here that the Lagrangian is regular, so that it has an equivalent Hamiltonian formulation, and also that Q is linear and isomorphic to $\mathbb{R}^n$. Regularity of the Lagrangian also provides a canonical embedding $\eta : T^*N \to T^*Q$, and we will use the Lagrange multiplier formulation (3.7.7) of the forced constrained Hamiltonian map. As in Section 3.5.2, we will

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

rescale the Lagrange multipliers by the time-step to give

$$p_0 = -D_1 L_d(q_0, q_1) - f_d^-(q_0, q_1) + h(\lambda^{(0)})^T \nabla\phi(q_0), \tag{3.7.11a}$$

$$p_1 = D_2 L_d(q_0, q_1) + f_d^+(q_0, q_1) - h(\lambda^{(1)})^T \nabla\phi(q_1), \tag{3.7.11b}$$

$$\phi(q_1) = 0, \tag{3.7.11c}$$

$$\nabla\phi(q_1) \cdot \frac{\partial H}{\partial p}(q_1, p_1) = 0, \tag{3.7.11d}$$

where the initial condition (q_0, p_0) is in $\eta(T^*N)$, and we solve over $(q_1, p_1) \in T^*Q$. The last two equations ensure that the solution (q_1, p_1) will also lie in $\eta(T^*N)$. Of course, we could also use one of the alternative formulations from Theorem 3.7.2 or we could use the forced constrained discrete Euler–Lagrange equations (3.7.1) and work directly on $N \times N$.

To construct discrete Lagrangians and discrete forces we can use any of the techniques discussed previously. Here we give a few examples.

Example 3.7.4. (Low-order methods) For a low-order discrete Lagrangian and discrete forces, such as the L_d^α and $f_d^{\alpha,\pm}$ from Example 3.2.2, we can simply restrict them to $N \times N$, as in Section 3.5.3. This yields a simple method that remains on the constraint manifold and includes the forcing. ◊

Example 3.7.5. (Composition methods) As we have seen in several examples already, composition methods provide a particularly elegant method to construct high-order methods from a given low-order integrator. In the case of systems with both forcing and constraints, the appropriate composed discrete forces and discrete Lagrangians are given by the combination of the definitions for the forced and constrained cases. ◊

Example 3.7.6. (Symplectic partitioned Runge–Kutta methods) Combining the definitions of the discrete forces (3.2.8) with the constrained formulation of the discrete Lagrangian (3.5.6), we arrive at discrete forces and a discrete Lagrangian for which the discrete Hamiltonian map is a constrained symplectic partitioned Runge–Kutta method with forcing. ◊

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

PART FOUR
Time-dependent mechanics

4.1. Introduction

In this part we address the issues of nonautonomous systems and discrete energies, which are closely related. In doing so, we rederive the symplectic–energy–momentum-conserving integrators of Kane *et al.* (1999*a*) in a purely variational way. We stress, however, that the theory developed here has many applications, aside from deriving such integrators. These applications include a deeper understanding of the behaviour of the symplectic–momentum integrators discussed in previous parts, as well as apparently unrelated areas such as nonsmooth variational mechanics (Fetecau *et al.* 2001).

The basic methodology used here is that of variational mechanics and variational discretizations. Unlike the standard discrete variational mechanics discussed in earlier parts of this work, however, we extend the framework to include time variations in addition to the usual configuration variable variations, as in Lee (1983) and (1987).

For continuous Lagrangian dynamics, these extra variations, which produce conservation of energy, do not contribute any new information to the Euler–Lagrange equations. In the discrete setting, however, we see that we obtain an extra equation which exactly ensures preservation of a quantity we can identify as the discrete energy. In this way, both the definition of the discrete energy and the fact that it is preserved arise naturally from the variational principle.

This same approach is also used in variational multisymplectic mechanics, where a configuration is regarded as a section of a fibre bundle over spacetime and we distinguish between vertical (configuration) variations and horizontal (spacetime) variations. In that theory the nature of the dual state space and fully covariant momentum maps can be properly defined. The interested reader is referred to Gotay, Isenberg and Marsden (1997) and Marsden, Patrick and Shkoller (1998), as here we use only the limited subset of the formalism which is sufficient for our purposes. To keep the exposition as simple and direct as possible, we have had to compromise with the general theory just a little: for example, in discussing momentum maps, one should really use the *affine dual* rather than the *linear dual* as we do here. This distinction becomes crucial in the multisymplectic, or PDE context.

We also investigate the links between the discrete variational mechanics and Hamilton–Jacobi theory. In particular, we see that requiring the discrete and continuous energies to be equal is exactly the Hamilton–Jacobi PDE, which proves that the symplectic–energy–momentum variational integrators will only conserve the continuous energy exactly if they also exactly integrate the continuous flow.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

This result is clearly consistent with the well-known theorem of Ge and Marsden (1988) and shows the advantage of identifying the correct discrete energy. While we cannot hope to conserve the continuous energy for an arbitrary non-integrable system, we can exactly conserve the corresponding discrete quantity.

This is similar to the situation with the Newmark integrator, discussed in Kane *et al.* (2000), where it was long thought that the integrator did not conserve the momenta, but where the problem was simply that the wrong discretization of the momenta was being used. Once the correct discretization was chosen, it was seen that the algorithm was actually momentum-conserving (and indeed symplectic).

4.2. Background: Extended Lagrangian mechanics

4.2.1. Basic definitions

Consider a *configuration manifold* Q and the *time space* $\mathbb{R}$. Define the *extended configuration manifold* to be $\bar{Q} = \mathbb{R} \times Q$ and the corresponding *extended state space* to be $\mathbb{R} \times TQ$. Take an *extended Lagrangian* $L : \mathbb{R} \times TQ \to \mathbb{R}$.

The *extended path space* is

$$\bar{\mathcal{C}} = \{c : [a_0, a_f] \to \bar{Q} \mid c \text{ is a } C^2 \text{ curve and } c_t'(a) > 0\},$$

and we denote the two components of c by $c(a) = (c_t(a), c_q(a))$. Given a path $c(a)$, we define the *initial time* $t_0 = c_t(a_0)$ and *final time* $t_f = c_t(t_f)$ and we form the *associated curve* $q : [t_0, t_f] \to Q$ by

$$q(t) = c_q(c_t^{-1}(t)).$$

It is simple to check that two paths $c^1(a)$ and $c^2(a)$ have the same associated curve if and only if they are reparametrizations of each other: that is, if there exists a smooth monotone increasing isomorphism $h : [a_0, a_f] \to [a_0, a_f]$ such that $c^1 \circ h = c^2$.

Given an associated curve, it will be useful to define

$$\hat{\dot{q}}(t) = \big(t, q(t), \dot{q}(t)\big) \in \mathbb{R} \times TQ,$$
$$\hat{\ddot{q}}(t) = \big(\big(t, q(t), \dot{q}(t)\big), \big(1, \dot{q}(t), \ddot{q}(t)\big)\big) \in T(\mathbb{R} \times TQ).$$

Now define the *extended action map* $\bar{\mathfrak{G}} : \bar{\mathcal{C}} \to \mathbb{R}$ to be

$$\bar{\mathfrak{G}}(c) = \int_{t_0}^{t_f} L(\hat{\dot{t}})\, \mathrm{d}t, \tag{4.2.1}$$

where $q(t)$ is the associated curve to $c(a)$ and t_0 and t_f are the initial and

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

final times, respectively, for c. Differentiating $q(t) = c_q(c_t^{-1}(t))$ gives

$$\dot{q}(t) = \frac{c_q'(c_t^{-1}(t))}{c_t'(c_t^{-1}(t))}, \tag{4.2.2}$$

where $\dot{q}$ denotes the derivative with respect to t and c' denotes the derivative with respect to a. With this, $\bar{\mathfrak{G}}$ can be written as

$$\bar{\mathfrak{G}}(c) = \int_{a_0}^{a_f} L\left(c_t(a), c_q(a), \frac{c_q'(a)}{c_t'(a)}\right) c_t'(a)\,\mathrm{d}a. \tag{4.2.3}$$

The tangent space $T_c\bar{\mathcal{C}}$ to the extended path space $\bar{\mathcal{C}}$ at the path c is the set of all C^2 maps $\delta c : [a_0, a_f] \to T\bar{Q}$ such that $\pi_{\bar{Q}} \circ \delta c = c$, where $\pi_{\bar{Q}} : T\bar{Q} \to \bar{Q}$ is the canonical projection map.

Define the *extended second-order submanifold* of $T(\mathbb{R} \times TQ)$ to be

$$\bar{\ddot{Q}} = \{w \in T(\mathbb{R} \times TQ) \mid T\pi_Q(w) = \pi_{TQ}(w)\} \subset T(\mathbb{R} \times TQ), \tag{4.2.4}$$

where $\pi_Q : \mathbb{R} \times TQ \to Q$ and $\pi_{TQ} : T(\mathbb{R} \times TQ) \to TQ$ are the canonical projections. $\bar{\ddot{Q}}$ is the set of points of the form $((t, q, q'), (t', q', q''))$, where the third and fifth entries are equal.

4.2.2. Variations of the action

Now that we have defined the spaces we are working on and have formed the action function, we are ready to derive the Euler–Lagrange equations of motion and prove a number of facts about solutions of these equations. The basic methodology here is variational, that is, we derive the equations and symplectic forms from taking variations of the action with respect to the path. To begin this process, we state the following fundamental theorem.

Theorem 4.2.1. Given a C^k extended Lagrangian $L : \mathbb{R} \times TQ \to \mathbb{R}$, $k \geq 2$, there exists a unique C^{k-2} mapping $\bar{D}_{\mathrm{EL}}L : \bar{\ddot{Q}} \to T^*\bar{Q}$ and a unique C^{k-1} one-form $\bar{\Theta}_L$ on the extended state space $\mathbb{R} \times TQ$, such that for all variations $\delta c \in T_c\bar{\mathcal{C}}$ of $c \in \bar{\mathcal{C}}$ we have

$$\mathbf{d}\bar{\mathfrak{G}}(c) \cdot \delta c = \int_{t_0}^{t_f} \bar{D}_{\mathrm{EL}}L(\hat{\bar{q}}) \cdot \delta c \;\mathrm{d}t + \bar{\Theta}_L(\hat{\bar{q}}) \cdot \hat{\delta} c\Big|_{t_0}^{t_f}, \tag{4.2.5}$$

where

$$\hat{\delta} c(a) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \left(c_t^\epsilon(a), c_q^\epsilon(a), \frac{(c_q^\epsilon)'(a)}{(c_t^\epsilon)'(a)}\right),$$

and $c^\epsilon(a)$ is such that $\delta c(a) = \frac{\partial}{\partial \epsilon}\big|_{\epsilon=0} c^\epsilon(a)$ and $q(t)$ is the associated curve to $c(a)$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

The one-form $\bar{\Theta}_L$ is called the *extended Lagrangian one-form* and the mapping $\bar{D}_{\mathrm{EL}}L$ is called the *extended Euler–Lagrange map*. In coordinates they have the form

$$\bar{D}_{\mathrm{EL}}L(c'') = \left[\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right)\right]\mathbf{d}c_q^i + \left[\frac{\partial L}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L\right)\right]\mathbf{d}c_t, \tag{4.2.6}$$

$$\bar{\Theta}_L(c') = \left[\frac{\partial L}{\partial \dot{q}^i}\right]\mathbf{d}c_q^i - \left[\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L\right]\mathbf{d}c_t \tag{4.2.7}$$

where the various quantities are evaluated at either t or $a = c_t^{-1}(t)$ as appropriate.

Proof. Take the derivative of equation (4.2.3) in the direction δc to obtain

$$\begin{aligned}\mathbf{d}\bar{\mathfrak{G}}(c)\cdot\delta c &= \int_{a_0}^{a_f}\left[\frac{\partial L}{\partial t}\delta c_t + \frac{\partial L}{\partial q}\delta c_q + \frac{\partial L}{\partial \dot{q}}\left(\frac{\delta c_q'}{c_t'} - \frac{c_q'\delta c_t'}{(c_t')^2}\right)\right]c_t'\,\mathrm{d}a \\ &\quad + \int_{a_0}^{a_f} L\delta c_t'\,\mathrm{d}a \\ &= \int_{a_0}^{a_f}\left[\frac{\partial L}{\partial q}c_t' - \frac{\mathrm{d}}{\mathrm{d}a}\frac{\partial L}{\partial \dot{q}}\right]\delta c_q\,\mathrm{d}a \\ &\quad + \int_{a_0}^{a_f}\left[\frac{\partial L}{\partial t}c_t' + \frac{\mathrm{d}}{\mathrm{d}a}\left(\frac{\partial L}{\partial \dot{q}}\frac{c_q'}{c_t'} - L\right)\right]\delta c_t\,\mathrm{d}a \\ &\quad + \left[\frac{\partial L}{\partial \dot{q}}\delta c_q\right]_{a_0}^{a_f} + \left[-\left(\frac{\partial L}{\partial \dot{q}}\frac{c_q'}{c_t'} - L\right)\delta c_t\right]_{a_0}^{a_f},\end{aligned}$$

where integration by parts has been used on the $\delta c_q'$ and $\delta c_t'$ terms.

Now change coordinates so that the integrals are taken with respect to t rather than a. In doing this, use the facts that $\mathrm{d}t = c_t'(a)\,\mathrm{d}a$ and $\frac{\mathrm{d}}{\mathrm{d}a} = c_t'(a)\frac{\mathrm{d}}{\mathrm{d}t}$. This gives the desired expression (4.2.5) with the Euler–Lagrange derivative (4.2.6) and the Lagrangian one-form (4.2.7). □

The fact that the extended Euler–Lagrange operator and extended Lagrangian one-form are functions only of the associated curve $q(t)$ and not of the full path $c(a)$ is a reflection of the fact that the extended action $\bar{\mathfrak{G}}$ is itself only a function of $q(t)$. This will be very important in what follows.

4.2.3. Euler–Lagrange equations

Hamilton's principle of critical action now seeks those paths $c \in \bar{\mathcal{C}}$ which are critical points of the action. More precisely, define the *space of solutions* $\bar{\mathcal{C}}_L \subset \bar{\mathcal{C}}$ to be all those paths c that satisfy $\mathbf{d}\bar{\mathfrak{G}}(c)\cdot\delta c = 0$ for all variations $\delta c \in T_c\bar{\mathcal{C}}$ which are zero at the boundary points a_0 and a_f.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Using equation (4.2.5) it is clear that c is a solution if and only if the extended Euler–Lagrange map (4.2.6) is zero on the associated curve to c for all $t \in (t_0, t_f)$. For any time t this statement is

$$\bar{D}_{\mathrm{EL}} L(\hat{\bar{q}}) = 0 \tag{4.2.8}$$

and is called the *extended Euler–Lagrange equations*.

The fact that the extended Euler–Lagrange map depends only on the associated curve, and not on the path c itself, means that any other paths which have the same associated curve will be solutions if and only if c is a solution. As remarked earlier, this is equivalent to the paths being reparametrizations of each other. We can thus group the solution space $\bar{\mathcal{C}}_L$ into equivalence classes of paths, each of which correspond to a single associated curve that satisfies the extended Euler–Lagrange equations. We will use this identification later when deriving an extended Lagrangian vector field and flow map, and in proving the conservation properties of these structures.

Of course, for an arbitrary Lagrangian the extended Euler–Lagrange equations will not necessarily uniquely define the associated curve, either. For the remainder of this section, however, we will assume that L is chosen such that $q(t)$ is uniquely determined, and in Section 4.4.4 we will consider conditions under which this is true.

Considering the expression of the extended Euler–Lagrange equations in more detail, we can break it up into the two components to give

$$\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = 0, \tag{4.2.9a}$$

$$\frac{\partial L}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L\right) = 0. \tag{4.2.9b}$$

In fact, only the first component of the Euler–Lagrange equations (4.2.9) is necessary, as it implies the second. To see this, consider an associated curve q satisfying (4.2.9a) for all $t \in (t_0, t_f)$. We now compute the second component (4.2.9b) of the Euler–Lagrange equations to be

$$\begin{aligned}
\frac{\partial L}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L\right) &= \frac{\partial L}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right)\dot{q}^i + \frac{\partial L}{\partial \dot{q}^i}\ddot{q}^i - \frac{\mathrm{d}L}{\mathrm{d}t} \\
&= \left[\frac{\partial L}{\partial t} + \frac{\partial L}{\partial q^i}\dot{q}^i + \frac{\partial L}{\partial \dot{q}^i}\ddot{q}^i\right] - \frac{\mathrm{d}L}{\mathrm{d}t} \\
&= 0,
\end{aligned}$$

where we used (4.2.9a) to pass from the first to the second line. The space of solutions $\bar{\mathcal{C}}_L$ is thus the space of paths whose associated curves satisfy (4.2.9a) for all t.

This redundancy in the equations is a reflection of the fact that they only depend on the associated curve, and so the equations cannot determine

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

the map $c_t(a)$. There must therefore be some functional dependency in the system, as is explicitly shown above.

The second part of the extended Euler–Lagrange equation (4.2.9) is actually the statement of energy evolution for an extended Lagrangian system. The *energy* $E_L : \mathbb{R} \times TQ \to \mathbb{R}$ is defined to be

$$E_L(t, q, \dot{q}) = \frac{\partial L}{\partial \dot{q}}(t, q, \dot{q}) \cdot \dot{q} - L(t, q, \dot{q}) \tag{4.2.10}$$

and is constant along solutions if L does not explicitly depend on time. This is clear from the fact that the second part (4.2.9b) of the extended Euler–Lagrange equations can be written as

$$\frac{\partial L}{\partial t} + \frac{\mathrm{d}E_L}{\mathrm{d}t} = 0.$$

Another reason that only the first part (4.2.9a) of the extended Euler–Lagrange equations is needed is that it already has the energy evolution built into it, making the second part redundant. If it did not, then the second part would ensure energy conservation (for L time-independent) or the correct energy change (for L time-varying). This will be especially important when we consider discrete Lagrangian systems.

Note that the definition of the energy function (4.2.10) also allows us to write the extended Lagrangian one-form in the compact notation

$$\bar{\Theta}_L = \frac{\partial L}{\partial \dot{q}} \mathbf{d}q - E_L \mathbf{d}t, \tag{4.2.11}$$

where we use (q, t) to refer to the two components of c. This expression for $\bar{\Theta}_L$ will be useful when we consider the corresponding discrete object.

4.2.4. Lagrangian vector fields and flow maps

As we have already seen, the solution paths are only uniquely defined up to reparametrizations in a. Equivalently, this means that only the associated curve is uniquely defined by the extended Euler–Lagrange equations. For this reason, there is no unique vector field on $T\bar{Q}$ for which solution paths are integral curves, but there is a unique vector field on $\mathbb{R} \times TQ$ for which the associated curves are integral curves, assuming that the time evolution is fixed as the identity.

More precisely, define *extended Lagrangian vector field* $\bar{X}_L : \mathbb{R} \times TQ \to T(\mathbb{R} \times TQ)$ to be the unique second-order vector field satisfying

$$\bar{D}_{\mathrm{EL}} L \circ \bar{X}_L = 0,$$
$$T\pi_{\mathbb{R}} \circ \bar{X}_L(t, q, \dot{q}) = (t, 1),$$

where $\pi_{\mathbb{R}} : \mathbb{R} \times TQ \to \mathbb{R}$ is the projection.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

The *extended Lagrangian flow* $\bar{F}_L : \mathbb{R} \times (\mathbb{R} \times TQ) \to \mathbb{R} \times TQ$ is the flow map of $\bar{X}_L$, and we write $\bar{F}_L^h : \mathbb{R} \times TQ \to \mathbb{R} \times TQ$ for some fixed time h.

4.2.5. Conservation of the extended symplectic form

Using the definition of $\bar{F}_L$, it is clear that setting $q(t) = \pi_Q \circ \bar{F}_L^{t-t_0}(t_0, q_0, \dot{q}_0)$ for some initial condition $(t_0, q_0, \dot{q}_0) \in \mathbb{R} \times TQ$ means that $q(t)$ will satisfy the extended Euler–Lagrange equations and will thus be an element of $\bar{\mathcal{C}}_L$. Also, given any $q(t) \in \bar{\mathcal{C}}_L$, it will satisfy $q(t) = \pi_Q \circ \bar{F}_L^{t-t_0}(t_0, q(t_0), \dot{q}(t_0))$.

Choosing an elapsed time $h \in \mathbb{R}$, we may thus define the *restricted extended action map* $\bar{\hat{\mathfrak{G}}} : \mathbb{R} \times TQ \to \mathbb{R}$ to be

$$\bar{\hat{\mathfrak{G}}}^h(t_0, q_0, \dot{q}_0) = \bar{\mathfrak{G}}(c),$$

where $c \in \mathcal{C}_L(Q)$ is any solution satisfying

$$\left(c_t(a_0), c_q(a_0), \frac{c_q'(a_0)}{c_t'(a_0)}\right) = (t_0, q_0, \dot{q}_0),$$
$$c_t(a_f) = t_0 + h.$$

We now wish to calculate the derivative of this expression. Begin by considering a variation derived from a family of initial conditions by

$$\big((t_0, q_0, \dot{q}_0), (\delta t_0, \delta q_0, \delta \dot{q}_0)\big) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} (t_0^\epsilon, q_0^\epsilon, \dot{q}_0^\epsilon) \in T(\mathbb{R} \times TQ),$$

and let $c^\epsilon \in \bar{\mathcal{C}}_L$ be a corresponding family of solutions satisfying

$$\left(c_t^\epsilon(a_0), c_q^\epsilon(a_0), \frac{(c_q^\epsilon)'(a_0)}{(c_t^\epsilon)'(a_0)}\right) = (t_0^\epsilon, q_0^\epsilon, \dot{q}_0^\epsilon),$$
$$c_t^\epsilon(a_f) = t_0^\epsilon + h.$$

Observe that we thus have

$$\bar{F}_L^{t-t_0^\epsilon}(t_0^\epsilon, q_0^\epsilon, \dot{q}_0^\epsilon) = \left(c_t^\epsilon(a), c_q^\epsilon(a), \frac{(c_q^\epsilon)'(a)}{(c_t^\epsilon)'(a)}\right), \qquad a = (c_t^\epsilon)^{-1}(t),$$

and so

$$\hat{\delta} c(a_0) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \bar{F}_L^0(t_0^\epsilon, q_0^\epsilon, \dot{q}_0^\epsilon) = (\delta t_0, \delta q_0, \delta \dot{q}_0),$$
$$\hat{\delta} c(a_f) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \bar{F}_L^h(t_0^\epsilon, q_0^\epsilon, \dot{q}_0^\epsilon) = T\bar{F}_L^h(t_0, q_0, \dot{q}_0) \cdot (\delta t_0, \delta q_0, \delta \dot{q}_0).$$

To calculate $\mathbf{d}\bar{\hat{\mathfrak{G}}}$ we now use equation (4.2.5) for the derivative of the extended action together with the above expressions for the variations and the

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

fact that the extended Euler–Lagrange map is zero on solutions to give

$$\mathbf{d}\bar{\hat{\mathfrak{G}}}^h(t_0,q_0,\dot{q}_0)\cdot(\delta t_0,\delta q_0,\delta\dot{q}_0)=\bar{\Theta}_L\big(\bar{F}_L^h(t_0,q_0,\dot{q}_0)\big)\cdot T\bar{F}_L^h\cdot(\delta t_0,\delta q_0,\delta\dot{q}_0)$$
$$-\bar{\Theta}_L(t_0,q_0,\dot{q}_0)\cdot(\delta t_0,\delta q_0,\delta\dot{q}_0),$$

which implies that

$$\mathbf{d}\bar{\hat{\mathfrak{G}}}^h=(\bar{F}_L^h)^*\bar{\Theta}_L-\bar{\Theta}_L. \tag{4.2.12}$$

Taking a further derivative of (4.2.12), and using the fact that $\mathbf{d}^2=0$, now gives conservation of the extended symplectic structure

$$(F_L^h)^*\Omega_L=\Omega_L,$$

where $\Omega_L=-\mathbf{d}\Theta_L$ is the *extended Lagrangian symplectic form* on $\mathbb{R}\times TQ$.

4.2.6. Conservation of momentum maps

Consider a Lie group G which acts on $\bar{Q}$ by the (left or right) action $\Phi^{\bar{Q}}: G\times\bar{Q}\to\bar{Q}$ in such a way that it covers an action $\Phi^t: G\times\mathbb{R}\to\mathbb{R}$, which is assumed to be monotone ($\partial_t\Phi^t(t)\neq 0$). In coordinates this is

$$\Phi_g^{\bar{Q}}(t,q)=\big(\Phi^t(t),\Phi^q(t,q)\big).$$

The tangent lift of this action to $\Phi^{T\bar{Q}}: G\times T\bar{Q}\to T\bar{Q}$ is defined by $\Phi_g^{T\bar{Q}}=T\Phi_g$, and the lift to the extended state space $\Phi^{\mathbb{R}\times TQ}: G\times(\mathbb{R}\times TQ)\to\mathbb{R}\times TQ$ is defined by

$$\Phi_g^{\mathbb{R}\times TQ}(t,q,\dot{q})=\left(\Phi_g^t(t),\Phi_g^q(t,q),\frac{\partial_t\Phi_g^q(t,q)+\partial_q\Phi_g^q(t,q)\cdot\dot{q}}{\partial_t\Phi_g^t(t)}\right).$$

This is defined so that, if a path $c(a)$ is transformed by the pointwise action of $\Phi_g^{T\bar{Q}}$, then the associated curve and its first derivative will transform pointwise with $\Phi_g^{\mathbb{R}\times TQ}$.

The *infinitesimal generators* corresponding to these actions are $\xi_{\bar{Q}}:\bar{Q}\to T(\bar{Q})$, $\xi_{T\bar{Q}}:T\bar{Q}\to T(T\bar{Q})$ and $\xi_{\mathbb{R}\times TQ}:\mathbb{R}\times TQ\to T(\mathbb{R}\times TQ)$, given by

$$\xi_{\bar{Q}}(t,q)=\frac{\mathrm{d}}{\mathrm{d}g}\big(\Phi_g^{\bar{Q}}(t,q)\big)\cdot\xi,$$
$$\xi_{T\bar{Q}}(t,q,\delta t,\delta q)=\frac{\mathrm{d}}{\mathrm{d}g}\big(\Phi_g^{T\bar{Q}}(t,q,\delta t,\delta q)\big)\cdot\xi,$$
$$\xi_{\mathbb{R}\times TQ}(t,q,\dot{q})=\frac{\mathrm{d}}{\mathrm{d}g}\big(\Phi_g^{\mathbb{R}\times TQ}(t,q,\dot{q})\big)\cdot\xi.$$

The *extended Lagrangian momentum map* $\bar{J}_L:\mathbb{R}\times TQ\to\mathfrak{g}^*$ is now defined by

$$\bar{J}_L(t,q,\dot{q})\cdot\xi=\bar{\Theta}_L\cdot\xi_{\mathbb{R}\times TQ}(t,q,\dot{q}).$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Denoting the two components of $\xi_{\bar{Q}}$ by $\xi_{\bar{Q}}(t,q) = (\xi^t_{\bar{Q}}(t,q), \xi^q_{\bar{Q}}(t,q))$, it can be checked that an equivalent expression for $\bar{J}_L$ is

$$\bar{J}_L(t,q,\dot{q}) \cdot \xi = \left\langle \frac{\partial L}{\partial \dot{q}}, \xi^q_{\bar{Q}}(t,q) \right\rangle - E_L(t,q,\dot{q}) \xi^t_{\bar{Q}}(t,q),$$

which is useful in applications as it does not involve the lifted action.

If $\Phi^{\bar{Q}}$ acts by extended special symplectic maps, so that $(\Phi^{\bar{Q}}_g)^* \bar{\Theta_L} = \Theta_L$, then it can be proved that the lifted action on the extended state space is equivariant, so that $\bar{J}_L \circ \Phi^{\mathbb{R}\times TQ}_g = \mathrm{Ad}^*_{g^{-1}} \circ \bar{J}_L$. The proof of this is essentially identical to that of Theorem 1.2.2.

In the autonomous setting, invariance of the Lagrangian implies invariance of the action. As the action is an integral with respect to time, however, and extended group actions can reparametrize time, this is no longer sufficient. Instead, we say that the group action is a *symmetry* if the one-form $L\mathbf{d}t$ is preserved, so that

$$(\Phi^{\mathbb{R}\times TQ}_g)^*(L\mathbf{d}t) = L\mathbf{d}t.$$

In fact, this same condition ensures that the group action is by extended special symplectic maps, and hence the momentum maps are equivariant. To see this, we differentiate the above expression with respect to $\dot{q}$ in the direction δq to obtain

$$\frac{\partial L}{\partial \dot{q}} \circ \Phi^{\mathbb{R}\times TQ}_g \cdot \partial_q \Phi^q_g \cdot \delta q = \frac{\partial L}{\partial \dot{q}} \cdot \delta q,$$

and substituting this into the expression for $(\Phi^{\mathbb{R}\times TQ}_g)^*(\bar{\Theta}_L)(t,q,\dot{q}) \cdot (\delta t, \delta q, \delta \dot{q})$ we can rearrange to obtain $\bar{\Theta}_L \cdot (\delta t, \delta q, \delta \dot{q})$.

Note that in the language of multisymplectic mechanics we have required that the Lagrangian density be invariant, and the extended Lagrangian momentum map defined above is the fully covariant Lagrangian momentum map. See Gotay *et al.* (1997) for more details.

Theorem 4.2.2. (Extended Noether's theorem) Consider an extended Lagrangian system $L : \mathbb{R} \times TQ \to \mathbb{R}$ with a (left or right) symmetry action $\Phi : G \times \bar{Q} \to \bar{Q}$. Then the corresponding (left or right) extended Lagrangian momentum map $\bar{J}_L : \mathbb{R} \times TQ \to \mathfrak{g}^*$ is a conserved quantity of the flow, so that $\bar{J}_L \circ \bar{F}^h_L = \bar{J}_L$ for all elapsed times h.

Proof. First define the lifted action on paths $\Phi^{\bar{\mathcal{C}}} : G \times \bar{\mathcal{C}} \to \bar{\mathcal{C}}$ by $\Phi^{\bar{\mathcal{C}}}_g(c)(a) = \Phi^{T\bar{Q}}_g(c(a))$. If two paths c_1 and c_2 have the same associated curve q, then the transformed paths $\bar{c}_1 = \Phi^{\bar{\mathcal{C}}}_g(c_1)$ and $\bar{c}_2 = \Phi^{\bar{\mathcal{C}}}_g(c_2)$ will also map to a single associated curve $\bar{q}$. We can thus define the lifted action of G on the set of associated curves by

$$\Phi^{\bar{\mathcal{C}}}_g(q) = \Phi^{\mathbb{R}\times TQ}_g \circ q \circ (\Phi^t_g)^{-1},$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

so that if q is the associated curve to c, then $\Phi_g^{\bar{\mathcal{C}}}(q)$ is the associated curve to $\Phi_g^{\bar{\mathcal{C}}}(c)$. The infinitesimal generators corresponding to the action on the space of paths is $\xi_{\bar{\mathcal{C}}} : \bar{\mathcal{C}} \to T\bar{\mathcal{C}}$ defined by

$$\xi_{\bar{\mathcal{C}}}(c) = \frac{\mathrm{d}}{\mathrm{d}g}\left(\Phi_g^{\bar{\mathcal{C}}}(c)\right) \cdot \xi.$$

If the one-form $L\mathbf{d}t$ is invariant, then taking an integral of this along an associated curve shows that the action $\bar{\mathfrak{G}} : \bar{\mathcal{C}} \to \mathbb{R}$ is also invariant, which is $\bar{\mathfrak{G}} \circ \Phi_g^{\bar{\mathcal{C}}} = \bar{\mathfrak{G}}$.

We thus have that solution curves map to solution curves and so $\xi_{\bar{\mathcal{C}}}(c) \in T_c\bar{\mathcal{C}}$. Restricting the infinitesimal invariance of the action to the space of solutions gives

$$0 = \mathbf{d}\bar{\mathfrak{G}}(c) \cdot \xi_{\bar{\mathcal{C}}}(c) = \bar{\Theta}_L(\hat{q}) \cdot \hat{\delta}c\Big|_{t_0}^{t_f},$$

where

$$\begin{aligned}\hat{\delta}c(a) &= \frac{\mathrm{d}}{\mathrm{d}g}\left(\Phi_g^t(c_t(a)), \Phi_g^q(c_t(a), c_q(a)), \frac{\partial_a \Phi_g^q(c_t(a), c_q(a))}{\partial_a \Phi_g^t(c_t(a))}\right) \cdot \xi \\ &= \frac{\mathrm{d}}{\mathrm{d}g}\Phi_g^{\mathbb{R}\times TQ}(\hat{q}(t)) \cdot \xi \qquad \text{where } t = c_t(a) \\ &= \xi_{\mathbb{R}\times TQ}(\hat{q}(t)).\end{aligned}$$

Recall now that, for an initial condition $(t_0, q_0, \dot{q}_0)$ with corresponding solution $q(t)$, we have $\hat{q}(t_0) = (t_0, q_0, \dot{q}_0)$ and $\hat{q}(t_f) = \bar{F}_L^h(t_0, q_0, \dot{q}_0)$. Combining this with the above expressions now gives

$$\begin{aligned}0 = \bar{\Theta}_L\big(\bar{F}_L^h(t_0, q_0, \dot{q}_0)\big) \cdot \xi_{\mathbb{R}\times TQ}\big(\bar{F}_L^h(t_0, q_0, \dot{q}_0)\big) \\ - \bar{\Theta}_L(t_0, q_0, \dot{q}_0) \cdot \xi_{\mathbb{R}\times TQ}(t_0, q_0, \dot{q}_0)\end{aligned}$$

which is exactly the statement $\bar{J}_L \circ \bar{F}_L^h = \bar{J}_L$. □

An interesting example of an extended momentum map arises for the time translation action of $\mathbb{R}$ on $\bar{Q}$. The corresponding momentum map is $\bar{J}_L(t, q, \dot{q}) = -E_L(t, q, \dot{q})\mathbf{d}\xi$ and if the Lagrangian is time-independent then Noether's theorem recovers the statement of energy conservation.

4.3. Discrete variational mechanics: Lagrangian viewpoint

We now turn to constructing a discrete time-dependent Lagrangian mechanics. As in the continuous case, the basic idea is to use a variational principle to derive the equations of motion, which then automatically guarantees conservation of discrete symplectic forms and discrete momentum maps.

The difference from standard discrete variational mechanics is that we work here with time-varying systems and take variations with respect to time, and not only with respect to the configuration variables. This gives

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

us two components to the discrete Euler–Lagrange equations, the second of which ensures the preservation of a discrete energy. This is different to the continuous case, where the first component of the Euler–Lagrange equations implies the second. In the discrete setting, both are independent and necessary.

4.3.1. Basic definitions

Consider once again a configuration manifold Q, time space $\mathbb{R}$ and extended configuration manifold $\bar{Q} = \mathbb{R} \times Q$. Define the *extended discrete Lagrangian state space* to be $\bar{Q} \times \bar{Q}$ and take an *extended discrete Lagrangian* $L_d : \bar{Q} \times \bar{Q} \to \mathbb{R}$.

The *extended discrete path space* is

$$\bar{\mathcal{C}}_d = \{c : \{0, \dots, N\} \to \bar{Q} \mid c_t(k+1) > c_t(k) \text{ for all } k\},$$

and as before we denote the two components of c by $c(k) = (c_t(k), c_q(k))$. Given a discrete path c, we form the *associated discrete curve*

$$q : \{c_t(0), \dots, c_t(N)\} \to Q$$

by

$$q(c_t(k)) = c_q(k),$$

and we denote the images of the time component of c by $t_k = c_t(k)$ and the images of the configuration component by $q_k = q(c_t(k)) = c_q(k)$. An extended discrete path is thus a sequence $\{(t_k, q_k)\}_{k=0}^{N}$ of points $(t_k, q_k) \in \bar{Q}$.

Observe that, unlike the autonomous situation, the extended discrete Lagrangian state space $\bar{Q} \times \bar{Q} = \mathbb{R} \times Q \times \mathbb{R} \times Q$ is not locally isomorphic to the extended Lagrangian state space $\mathbb{R} \times TQ$, as it has one extra dimension. Regarding a discrete Lagrangian system as an integrator, this is a reflection of the fact that nonautonomous discrete Lagrangian systems require adaptive time-stepping, which implies that it is necessary to keep track of both the current time and the current time-step.

In the case when the system being modelled is indeed autonomous, one could take the discrete system to be $Q \times \mathbb{R} \times Q$, with each element (q_0, h, q_1) consisting of two points q_0 and q_1 and a time-step h. As we will see later, however, this is not essentially different from the more general case of (t_0, q_0, t_1, q_1), and so we will only consider this more general formulation.

The *extended discrete action sum* $\bar{\mathfrak{G}}_d : \mathcal{C}_d \to \mathbb{R}$ is given by

$$\bar{\mathfrak{G}}_d(c) = \sum_{k=0}^{N-1} L_d(c(k), c(k+1)), \tag{4.3.1a}$$

$$\bar{\mathfrak{G}}_d(c) = \sum_{k=0}^{N-1} L_d(t_k, q_k, t_{k+1}, q_{k+1}). \tag{4.3.1b}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Equation (4.3.1a) is the discrete equivalent of (4.2.3) and equation (4.3.1b) is the equivalent of (4.2.1).

The tangent space $T_c\bar{\mathcal{C}}_d$ to the discrete path space $\bar{\mathcal{C}}_d$ at the path c is the set of all maps $\delta c : \{0, \ldots, N\} \to T\bar{Q}$ such that $\pi_{\bar{Q}} \circ \delta c = c$.

Consider the space $(\bar{Q} \times \bar{Q}) \times (\bar{Q} \times \bar{Q})$ and define the *extended projection operator* π and the *extended translation operator* σ by

$$\pi : ((c_0, c_1), (c_0', c_1')) \mapsto (c_0, c_1),$$
$$\sigma : ((c_0, c_1), (c_0', c_1')) \mapsto (c_0', c_1').$$

We define the *extended discrete second-order submanifold* of $(\bar{Q}\times\bar{Q})\times(\bar{Q}\times\bar{Q})$ to be

$$\bar{\ddot{Q}}_d = \{w_d \in (\bar{Q} \times \bar{Q}) \times (\bar{Q} \times \bar{Q}) \mid \pi_1 \circ \sigma(w_d) = \pi_2 \circ \pi(w_d)\},$$

where $\pi_1, \pi_2 : \bar{Q} \times \bar{Q} \to \bar{Q}$ are the projections onto the first and second components, respectively. $\bar{\ddot{Q}}_d$ is thus the set of points of the form $((t_0, q_0, t_1, q_1), (t_1, q_1, t_2, q_2))$, with the second and third values in $\bar{Q}$ being equal.

In contrast to the autonomous case, the extended discrete second-order submanifold $\bar{\ddot{Q}}_d$ is not isomorphic to the extended continuous second-order submanifold $\ddot{\bar{Q}}$. This will be reflected in the choice of initial conditions for a solution curve, as we will see below.

4.3.2. Variations of the discrete action

As in the continuous case, the discrete equations of motion and conservation laws are derived from a stationarity principle: that is, we take variations of the discrete action sum with respect to the discrete path, as stated in the following theorem.

Theorem 4.3.1. Given a C^k extended discrete Lagrangian $L_d : \bar{Q} \times \bar{Q} \to \mathbb{R}$, $k \geq 1$, there exists a unique C^{k-1} mapping $\bar{D}_{\mathrm{DEL}} L_d : \bar{\ddot{Q}}_d \to T^*\bar{Q}$ and unique C^{k-1} one-forms $\bar{\Theta}^-_{L_d}$ and $\bar{\Theta}^+_{L_d}$ on the discrete Lagrangian state space $\bar{Q} \times \bar{Q}$, such that, for all variations $\delta c \in T_c\bar{\mathcal{C}}_d$ of $c \in \bar{\mathcal{C}}_d$, we have

$$\begin{aligned}\mathbf{d}\bar{\mathfrak{G}}_d(c) \cdot \delta c = \sum_{k=1}^{N-1} & \bar{D}_{\mathrm{DEL}} L_d(t_{k-1}, q_{k-1}, t_k, q_k, t_{k+1}, q_{k+1}) \cdot (\delta q_k, \delta t_k) \\ & + \bar{\Theta}^+_{L_d}(t_{N-1}, q_{N-1}, t_N, q_N) \cdot (\delta t_{N-1}, \delta q_{N-1}, \delta t_N, \delta q_N) \\ & - \bar{\Theta}^-_{L_d}(t_0, q_0, t_1, q_1) \cdot (\delta t_0, \delta q_0, \delta t_1, \delta q_1), \end{aligned} \tag{4.3.2}$$

where we write $t_k = c_t(k)$ and $q_k = c_q(k)$ and similarly for the variations. The map $\bar{D}_{\mathrm{DEL}} L_d$ is called the *extended discrete Euler–Lagrange map* and the one-forms $\bar{\Theta}^+_{L_d}$ and $\bar{\Theta}^-_{L_d}$ are the *extended discrete Lagrangian one-forms*.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

In coordinates these have the expressions

$$\begin{aligned}\bar{D}_{\mathrm{DEL}}L_d(t_{k-1},q_{k-1},t_k,q_k,t_{k+1},q_{k+1}) \qquad (4.3.3)\\ = \big[D_4L_d(t_{k-1},q_{k-1},t_k,q_k)+D_2L_d(t_k,q_k,t_{k+1},q_{k+1})\big]\mathbf{d}q\\ +\big[D_3L_d(t_{k-1},q_{k-1},t_k,q_k)+D_1L_d(t_k,q_k,t_{k+1},q_{k+1})\big]\mathbf{d}t,\\ \bar{\Theta}^+_{L_d}(t_k,q_k,t_{k+1},q_{k+1}) \qquad (4.3.4)\\ = D_4L_d(t_k,q_k,t_{k+1},q_{k+1})\mathbf{d}q_{k+1}+D_3L_d(t_k,q_k,t_{k+1},q_{k+1})\mathbf{d}t_{k+1},\\ \bar{\Theta}^-_{L_d}(t_k,q_k,t_{k+1},q_{k+1}) \qquad (4.3.5)\\ = -D_2L_d(t_k,q_k,t_{k+1},q_{k+1})\mathbf{d}q_k-D_1L_d(t_k,q_k,t_{k+1},q_{k+1})\mathbf{d}t_k.\end{aligned}$$

Proof. Taking the derivative of (4.3.1a) at c in the direction δc yields

$$\begin{aligned}\mathbf{d}\bar{\mathfrak{G}}_d(c)\cdot\delta c &= \sum_{k=0}^{N-1}\big[D_1L_d(c(k),c(k+1))\cdot\delta c(k)\\ &\qquad +D_2L_d(c(k),c(k+1))\cdot\delta c(k+1)\big]\\ &= \sum_{k=1}^{N-1}\left[D_2L_d(c(k-1),c(k))+D_1L_d(c(k),c(k+1))\right]\cdot\delta c(k)\\ &\qquad +D_2L_d(c(N-1),c(N))\cdot\delta c(N)\\ &\qquad +D_1L_d(c(0),c(1))\cdot\delta c(0),\end{aligned}$$

where we have rearranged and grouped terms to isolate internal and boundary terms, as integration by parts achieved in the continuous setting.

Now we change notations to $(t_k,q_k)=(c_t(k),c_q(k))=c(k)$, and similarly for the variations, and we obtain the desired expressions. □

4.3.3. Discrete Euler–Lagrange equations

We now apply the discrete Hamilton's principle, and seek discrete paths $c\in\bar{\mathcal{C}}_d$ which are critical points of the discrete action. That is, we define the *discrete space of solutions* $\bar{\mathcal{C}}_{L_d}\subset\bar{\mathcal{C}}_d$ to be all those paths which satisfy $\mathbf{d}\bar{\mathfrak{G}}_d(c)\cdot\delta c=0$ for all variations $\delta c\in T_c\bar{\mathcal{C}}_d$ which are zero at the boundary points 0 and N.

As in the continuous case, it is clear from (4.3.2) that c is a solution if and only if the discrete Euler–Lagrange derivative is zero at all points other than the endpoints 0 and N. This statement at k reads

$$\bar{D}_{\mathrm{DEL}}L_d(t_{k-1},q_{k-1},t_k,q_k,t_{k+1},q_{k+1})=0 \qquad (4.3.6)$$

and is known as the *extended discrete Euler–Lagrange equations*. It can be separated into the configuration and time components to give

$$D_4L_d(t_{k-1},q_{k-1},t_k,q_k)+D_2L_d(t_k,q_k,t_{k+1},q_{k+1})=0, \qquad (4.3.7a)$$

$$D_3L_d(t_{k-1},q_{k-1},t_k,q_k)+D_1L_d(t_k,q_k,t_{k+1},q_{k+1})=0. \qquad (4.3.7b)$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Unlike the continuous case, however, paths which satisfy the first component (4.3.7a) of the Euler–Lagrange equations do not automatically satisfy the second component (4.3.7b) as well: that is, both components contribute restrictions on the space of solutions, and both are necessary. This is because the extended discrete action is a function of the entire extended path, not just of some discrete associated curve.

The interpretation of the discrete Euler–Lagrange equations and the discrete Lagrangian one-forms can be aided by defining the *discrete energies* to be

$$\begin{aligned} E^+_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}) &= -D_3 L_d(t_k, q_k, t_{k+1}, q_{k+1}), \\ E^-_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}) &= D_1 L_d(t_k, q_k, t_{k+1}, q_{k+1}). \end{aligned} \tag{4.3.8}$$

With these definitions we see that the second component of the discrete Euler–Lagrange equations (4.3.7b) is simply

$$E^+_{L_d}(t_{k-1}, q_{k-1}, t_k, q_k) = E^-_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}), \tag{4.3.9}$$

and reflects the evolution of the discrete energy.

If the discrete Lagrangian is time-invariant, so that $L_d(t_k + \tau, q_k, t_{k+1} + \tau, q_{k+1}) = L_d(t_k, q_k, t_{k+1}, q_{k+1})$ for all $\tau \in \mathbb{R}$, then taking derivatives with respect to τ shows that $D_1 L_d + D_3 L_d = 0$, and thus

$$E^-_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}) = E^+_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}). \tag{4.3.10}$$

Combining this with the time component of the Euler–Lagrange equations (4.3.7b) written as (4.3.9) then yields

$$\begin{aligned} E^+_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}) &= E^+_{L_d}(t_{k-1}, q_{k-1}, t_k, q_k), \\ E^-_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}) &= E^-_{L_d}(t_{k-1}, q_{k-1}, t_k, q_k), \end{aligned}$$

which shows that time invariance of the discrete Lagrangian leads to conservation of the discrete energies. This is a special case of the extended discrete Noether's theorem, as we will see below. If the Lagrangian is not time-invariant, then the equivalent of equation (4.3.10) will indicate how the discrete energies evolve.

4.3.4. Extended discrete Lagrangian evolution operator and mappings

The *extended discrete Lagrangian evolution operator* $\bar{X}_{L_d}$ is the second-order discrete evolution operator defined by

$$\bar{D}_{\mathrm{DEL}} L_d \circ \bar{X}_{L_d} = 0.$$

If this expression uniquely determines $\bar{X}_{L_d}$ then we will say that the extended discrete Lagrangian is *discretely well-posed*. In this case, the *extended discrete Lagrangian map* $\bar{F}_{L_d} : \bar{Q} \times \bar{Q} \to \bar{Q} \times \bar{Q}$ is specified by $\bar{F}_{L_d} = \sigma \circ X_{L_d}$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

For the rest of this section we will assume that we are working with an extended discrete Lagrangian which is discretely well-posed. Note that this is equivalent to the extended discrete Euler–Lagrange equations having unique solutions for t_2 and q_2 in terms of the other variables. Later we will investigate under what conditions L_d is in fact discretely well-posed.

Note that here, and in the material which follows, we are implicitly making use of the restriction on the extended discrete path space that $t_{k+1} > t_k$ for each k. This rules out spurious solutions which do not uniformly increase in time.

An extended discrete path $c_d \in \bar{\mathcal{C}}_d$ is a solution of the extended discrete Euler–Lagrange equations if the points $\{(t_k, q_k)\}_{k=0}^N$ satisfy

$$\bar{F}_{L_d}(t_{k-1}, q_{k-1}, t_k, q_k) = (t_k, q_k, t_{k+1}, q_{k+1})$$

for each $k = 1, \ldots, N-1$. That is, the extended discrete Lagrangian map defines a trajectory forward in time.

4.3.5. Extended discrete Lagrangian maps are symplectic

The extended discrete space of solutions $\bar{\mathcal{C}}_{L_d}$ defined above can, by means of the extended discrete Lagrangian map, be identified with the space of initial conditions $\bar{Q} \times \bar{Q}$. Restricting the extended discrete action to this space gives the *restricted extended discrete action* $\bar{\hat{\mathfrak{G}}}_d : \bar{Q} \times \bar{Q} \to \mathbb{R}$ defined by

$$\bar{\hat{\mathfrak{G}}}_d(t_0, q_0, t_1, q_1) = \bar{\mathfrak{G}}_d(c_d),$$

where $c_d \in \bar{\mathcal{C}}_{L_d}$ is the solution of the extended discrete Euler–Lagrange equations satisfying $c(0) = (t_0, q_0)$ and $c(1) = (t_1, q_1)$. We now use (4.3.2) to calculate

$$\mathbf{d}\bar{\hat{\mathfrak{G}}}_d = (\bar{F}_{L_d}^{N-1})^*(\Theta_{L_d}^+) - \Theta_{L_d}^-,$$

and so taking another derivative and using the fact that $\mathbf{d}^2 = 0$ gives the conservation law

$$(\bar{F}_{L_d}^{N-1})^*(\Omega_{L_d}) = \Omega_{L_d},$$

where the *extended discrete Lagrangian two-form* Ω_{L_d} is defined by

$$\Omega_{L_d} = -\mathbf{d}\Theta_{L_d}^+ = -\mathbf{d}\Theta_{L_d}^-.$$

Taking any subinterval of $0, \ldots, N$ and using the same argument gives the same conservation law for any number of steps of $\bar{F}_{L_d}$.

4.3.6. Extended discrete Noether's theorem

Take a Lie group G with a (left or right) action $\Phi^{\bar{Q}}$ on $\bar{Q}$, and consider the lift to the extended discrete state space $\bar{Q} \times \bar{Q}$ defined by the product action

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

$\Phi_g^{\bar{Q}\times\bar{Q}}(t_0,q_0,t_1,q_1) = (\Phi_g^{\bar{Q}}(t_0,q_0),\Phi_g^{\bar{Q}}(t_0,q_0))$. The corresponding *infinitesimal generator* is then $\xi_{\bar{Q}\times\bar{Q}} : \bar{Q}\times\bar{Q} \to T(\bar{Q}\times\bar{Q})$, defined by

$$\xi_{\bar{Q}\times\bar{Q}}(t_0,q_0,t_1,q_1) = (\xi_{\bar{Q}}(t_0,q_0),\xi_{\bar{Q}}(t_1,q_1))$$

and the *extended discrete Lagrangian momentum maps* $\bar{J}^+_{L_d}, \bar{J}^-_{L_d} : \bar{Q}\times\bar{Q} \to \mathfrak{g}^*$ are

$$\bar{J}^+_{L_d}(t_0,q_0,t_1,q_1)\cdot\xi = \bar{\Theta}^+_{L_d}\cdot\xi_{\bar{Q}\times\bar{Q}}(t_0,q_0,t_1,q_1),$$
$$\bar{J}^-_{L_d}(t_0,q_0,t_1,q_1)\cdot\xi = \bar{\Theta}^-_{L_d}\cdot\xi_{\bar{Q}\times\bar{Q}}(t_0,q_0,t_1,q_1).$$

Equivalent expressions for these are

$$\begin{aligned}
&\bar{J}^+_{L_d}(t_0,q_0,t_1,q_1)\cdot\xi \\
&\quad = \left\langle D_4L_d(t_0,q_0,t_1,q_1),\xi^q_{\bar{Q}}(t_1,q_1)\right\rangle + D_3L_d(t_0,q_0,t_1,q_1)\xi^t_{\bar{Q}}(t_1), \\
&\bar{J}^-_{L_d}(t_0,q_0,t_1,q_1)\cdot\xi \\
&\quad = \left\langle -D_2L_d(t_0,q_0,t_1,q_1),\xi^q_{\bar{Q}}(t_0,q_0)\right\rangle - D_1L_d(t_0,q_0,t_1,q_1)\xi^t_{\bar{Q}}(t_0).
\end{aligned}$$

If the lifted action on $\bar{Q}\times\bar{Q}$ acts by extended special discrete symplectic maps, which requires $(\Phi_g^{\bar{Q}\times\bar{Q}})^*(\bar{\Theta}^+_{L_d}) = \bar{\Theta}^+_{L_d}$ and $(\Phi_g^{\bar{Q}\times\bar{Q}})^*(\bar{\Theta}^-_{L_d}) = \bar{\Theta}^-_{L_d}$, then the extended discrete Lagrangian momentum maps will be equivariant: that is,

$$\bar{J}^+_{L_d}\circ\Phi_g^{\bar{Q}\times\bar{Q}} = \mathrm{Ad}^*_{g^{-1}}\circ\bar{J}^+_{L_d},$$
$$\bar{J}^-_{L_d}\circ\Phi_g^{\bar{Q}\times\bar{Q}} = \mathrm{Ad}^*_{g^{-1}}\circ\bar{J}^-_{L_d}.$$

This can be proved in the same way as Theorem 1.2.2.

An extended discrete Lagrangian $L_d : \bar{Q}\times\bar{Q} \to \mathbb{R}$ is said to be *invariant* under the lifted group action if $L_d(g\cdot(t_0,q_0),g\cdot(t_1,q_1)) = L_d(t_0,q_0,t_1,q_1))$ for all g and all t_0, q_0, t_1 and q_1, and then G is said to be *symmetry* of L_d. If a discrete Lagrangian is invariant then it is necessarily *infinitesimally invariant*, which is the requirement that $\mathbf{d}L_d\cdot\xi_{\bar{Q}\times\bar{Q}} = 0$ for all $\xi\in\mathfrak{g}$.

In the case that L_d is infinitesimally invariant, the fact that $\mathbf{d}L_d = \bar{\Theta}^+_{L_d} - \bar{\Theta}^-_{L_d}$ implies that the two extended discrete momentum maps are equal, and we will denote the single *extended discrete Lagrangian momentum map* by $\bar{J}_{L_d} : \bar{Q}\times\bar{Q} \to \mathfrak{g}^*$.

As the discrete Lagrangian plays an analogous role to the continuous action, it is not surprising that for the extended continuous Noether's theorem we had to require invariance of the one-form $L\mathbf{d}t$, which gives invariance of the action, while for the discrete version action invariance is directly implied by invariance of the extended discrete Lagrangian itself.

If the extended discrete Lagrangian is invariant, then the lifted group action acts by extended special discrete symplectic maps. This can be seen in a similar way to the corresponding autonomous statement, by differentiating

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

$L_d \circ \Phi_g^{\bar{Q}\times\bar{Q}}(t_0, q_0, t_1, q_1) = L_d(t_0, q_0, t_1, q_1)$ with respect to t_1 and q_1 and using the resulting identities to show $(\Phi_g^{\bar{Q}\times\bar{Q}})^*(\bar{\Theta}^+_{L_d}) = \bar{\Theta}^+_{L_d}$, and similarly for $\bar{\Theta}^+_{L_d}$.

Theorem 4.3.2. (Extended discrete Noether's theorem) Consider an extended discrete Lagrangian system $L_d : \bar{Q} \times \bar{Q} \to \mathbb{R}$ which is invariant under the lift of the (left or right) group action $\Phi^{\bar{Q}} : G \times \bar{Q} \to \bar{Q}$. Then the corresponding extended discrete Lagrangian momentum map $\bar{J}_{L_d} : \bar{Q} \times \bar{Q} \to \mathfrak{g}^*$ is a conserved quantity of the extended discrete Lagrangian map $\bar{F}_{L_d} : \bar{Q} \times \bar{Q} \to \bar{Q} \times \bar{Q}$, in the sense that $\bar{J}_{L_d} \circ \bar{F}_{L_d} = \bar{J}_{L_d}$.

Proof. The extended discrete Noether's theorem can be proved in exactly the same way as the standard discrete Noether's theorem (Theorem 1.3.3): that is, we compute

$$\begin{aligned} 0 &= \mathbf{d}\bar{\hat{\mathfrak{G}}}_d(t_0, q_0, t_1, q_1) \cdot \xi_{\bar{Q}\times\bar{Q}}(t_0, q_0, t_1, q_1) \\ &= \big((\bar{F}^{N-1}_{L_d})^*(\bar{\Theta}^+_{L_d}) - \bar{\Theta}^-_{L_d}\big)(t_0, q_0, t_1, q_1) \cdot \xi_{\bar{Q}\times\bar{Q}}(t_0, q_0, t_1, q_1), \end{aligned}$$

which gives the desired expression, using the fact that we may take N to be arbitrary. □

As in earlier Noether's theorems, we have required more than is actually necessary for the proof. In particular, the result will still hold if L_d is only infinitesimally invariant.

Example 4.3.3. If we have a discrete Lagrangian which is autonomous, in the sense that it is invariant with respect to the additive action of $\mathbb{R}$ on the time component of $\bar{Q}$, then the extended discrete Noether's theorem recovers the statement of discrete energy conservation.

To see this, consider the group action $g \cdot (t, q) = (t + g, q)$ for $g \in \mathbb{R}$, which has infinitesimal generator $\xi_{\bar{Q}}(t, q) = ((t, q), (\xi, 0))$ for $\xi \in \mathfrak{g}^* \cong \mathbb{R}$. Using this, we compute the extended discrete Lagrangian momentum map to be $\bar{J}_{L_d} = -E^+_{L_d}\mathbf{d}t_1$, or equivalently $\bar{J}_{L_d} = -E^-_{L_d}\mathbf{d}t_0$. Noether's theorem thus gives $E^+_{L_d}(k, k+1) = E^+_{L_d}(k-1, k)$, or $E^-_{L_d}(k, k+1) = E^-_{L_d}(k-1, k)$, which are the statements of discrete energy conservation. This is directly analogous to the continuous case, where energy conservation is also recovered by Noether's theorem in the case when L is autonomous. ◊

A difference from the continuous case arises because the proof of the extended discrete Noether's theorem relied upon the fact that both components of the extended discrete Euler–Lagrange equations (4.3.7) are satisfied at intermediate (non-boundary) points, unlike the continuous case where only the first component was required to be satisfied. Weakening this requirement to only insist on (4.3.7a) being satisfied means that full extended discrete Noether's theorem will only be true in the case that the symmetry action

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

is autonomous and time-preserving, which shows that the discrete energy would not then be preserved.

4.3.7. Autonomous discrete Lagrangians

The previous example is also interesting because this is the case in which one could formulate a discrete theory using a state space of $Q \times \mathbb{R} \times Q$ with elements (q_0, h, q_1) consisting of two points q_0 and q_1 and a time-step h. In such a theory there would be only a single discrete energy E_{L_d} rather than $E^+_{L_d}$ and $E^-_{L_d}$. Given such a discrete Lagrangian $L'_d(q_0, h, q_1)$, however, we can also simply define an extended discrete Lagrangian by

$$L_d(t_0, q_0, t_1, q_1) = L'_d(q_0, t_1 - t_0, q_1)$$

and proceed to use the extended discrete Lagrangian theory derived in this section. Such an L_d will automatically have the time translation symmetry, and so the statement of infinitesimal invariance $\mathbf{d}L_d \cdot \xi_{\bar{Q}\times\bar{Q}} = 0$ can be seen to be just

$$\begin{aligned} 0 &= \mathbf{d}L_d(t_0, q_0, t_1, q_1) \cdot (\xi, 0, \xi, 0) \\ &= D_1 L_d(t_0, q_0, t_1, q_1) \cdot \xi + D_3 L_d(t_0, q_0, t_1, q_1) \cdot \xi \\ &= \big(E^-_{L_d}(t_0, q_0, t_1, q_1) - E^+_{L_d}(t_0, q_0, t_1, q_1)\big) \cdot \xi, \end{aligned}$$

and so the two discrete energies in this case are equal. This is just the single discrete energy $E_{L_d} = E^\pm_{L_d}$, which would appear in a theory based on $Q \times \mathbb{R} \times Q$.

This same calculation can also be carried through for the discrete one- and two-forms and the other discrete structures, and shows that a $Q \times \mathbb{R} \times Q$ theory would not be essentially different from that based on $\bar{Q} \times \bar{Q}$.

4.4. Background: Extended Hamiltonian mechanics

As in the autonomous setting, the Hamiltonian picture of mechanics is complementary to the Lagrangian, and additional insight can be gained for variational integrators if the connection is made with Hamiltonian mechanics and, in particular, with extended Hamilton–Jacobi theory.

4.4.1. Basic definitions

Consider a configuration manifold Q, time space $\mathbb{R}$ and extended configuration manifold $\bar{Q} = \mathbb{R} \times Q$. Define the *extended phase space* to be $\mathbb{R} \times T^*Q$, that is, the usual phase space augmented with time, and consider an *extended Hamiltonian* $H : \mathbb{R} \times T^*Q \to \mathbb{R}$.

Given an extended Hamiltonian H, we define the *extended canonical one-form* $\bar{\Theta}_H$ and the *extended canonical two-form* or *extended symplectic*

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

form $\bar{\Omega}_H$ to be one- and two-forms respectively on the extended phase space $\mathbb{R} \times T^*Q$ given by

$$\bar{\Theta}_H(t,q,p) = p_i \mathbf{d}q^i - H(t,q,p)\mathbf{d}t, \tag{4.4.1}$$

$$\bar{\Omega}_H(t,q,p) = -\mathbf{d}\bar{\Theta}_H = \mathbf{d}q^i \wedge \mathbf{d}p_i + \mathbf{d}H(t,q,p) \wedge \mathbf{d}t. \tag{4.4.2}$$

If $\phi : \mathbb{R} \times T^*Q \to \mathbb{R} \times T^*Q$ is a map from the extended phase space with Hamiltonian H_0 to the extended phase space with Hamiltonian H_1, then we say that ϕ is *symplectic* or *canonical* if ϕ pulls the extended symplectic form back to the appropriate other extended symplectic form, so that

$$\phi^*(\bar{\Omega}_{H_1}) = \bar{\Omega}_{H_0}.$$

4.4.2. Hamiltonian vector fields and flow maps

For an extended Hamiltonian $H : \mathbb{R} \times T^*Q$ we define the *extended Hamiltonian vector field* $\bar{X}_H$ on the extended phase space $\mathbb{R} \times T^*Q$ to be the unique vector field with unit time flow satisfying

$$\mathbf{i}_{\bar{X}_H} \bar{\Omega}_H = 0. \tag{4.4.3}$$

Note that this uniquely defines $\bar{X}_H$ for any H. Writing $\bar{X}_H = (1, X_q, X_p)$ in coordinates, we see that the above expression is

$$\mathbf{i}_{\bar{X}_H} \bar{\Omega}_H = \left[\frac{\partial H}{\partial q} \cdot X_q + \frac{\partial H}{\partial p} \cdot X_p\right] \mathbf{d}t + \left[-X_p - \frac{\partial H}{\partial q}\right] \mathbf{d}q + \left[X_q - \frac{\partial H}{\partial p}\right] \mathbf{d}p,$$

and requiring that this be zero gives the familiar *extended Hamilton's equations*

$$X_q(t,q,p) = \frac{\partial H}{\partial p}(t,q,p), \tag{4.4.4a}$$

$$X_p(t,q,p) = -\frac{\partial H}{\partial q}(t,q,p). \tag{4.4.4b}$$

Notice that the time component of (4.4.3) does not contribute an equation, because the other two equations automatically imply that

$$\frac{\partial H}{\partial q} \cdot X_q + \frac{\partial H}{\partial p} \cdot X_p = \frac{\partial H}{\partial q} \cdot \left(\frac{\partial H}{\partial p}\right) + \frac{\partial H}{\partial p} \cdot \left(-\frac{\partial H}{\partial q}\right) = 0,$$

which is equivalent to the fact that time variations in the Lagrangian case do not give additional constraints beyond the Euler–Lagrange equations.

The time component of (4.4.3) is a statement about the time evolution of the Hamiltonian. Computing

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t} \cdot X_t + \frac{\partial H}{\partial q} \cdot X_q + \frac{\partial H}{\partial p} \cdot X_p = \frac{\partial H}{\partial t},$$

we see that the derivative of the Hamiltonian with respect to time simply

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

reflects the explicit dependence of H on t. If this is zero, so that the extended Hamiltonian is time-independent, then H is constant along the flow of $\bar{X}_H$.

We denote the *extended flow map* of the extended Hamiltonian vector field $\bar{X}_H$ by $\bar{F}_H : \mathbb{R} \times (\mathbb{R} \times T^*Q) \to \mathbb{R} \times T^*Q$, defined as

$$\bar{F}_H^h(t_0, q_0, p_0) = \bar{F}_H(h, t_0, q_0, p_0) = (t_0 + h, q_1, p_1).$$

The flow map $\bar{F}_H^h$ for any fixed h is a symplectic map from $\mathbb{R} \times T^*Q$ to itself, as can be readily seen by taking the time derivative to obtain

$$\begin{aligned} \left.\frac{\partial}{\partial h}\right|_{h=0} (\bar{F}_H^h)^* \bar{\Omega}_H &= \mathcal{L}_{\bar{X}_H} \bar{\Omega}_H \\ &= \mathbf{d}\mathbf{i}_{\bar{X}_H} \bar{\Omega}_H + \mathbf{i}_{\bar{X}_H} \mathbf{d}\bar{\Omega}_H \\ &= 0, \end{aligned}$$

where we have used Cartan's magic formula $\mathcal{L}_X \alpha = \mathbf{d}\mathbf{i}_X \alpha + \mathbf{i}_X \mathbf{d}\alpha$, and then the facts that $\mathbf{i}_{\bar{X}_H} \bar{\Omega}_H = 0$ is the definition of the extended Hamiltonian vector field, and that $\mathbf{d}^2 = 0$ implies $\mathbf{d}\bar{\Omega}_H = -\mathbf{d}^2 \bar{\Theta}_H = 0$.

4.4.3. Extended Hamiltonian Noether's theorem

Consider a (left or right) action $\Phi : G \times \bar{Q} \to \bar{Q}$ of G on the extended configuration manifold $\bar{Q}$, as in Section 4.2.6. The lift of this action to the extended phase space $\mathbb{R} \times T^*Q$ is denoted by $\Phi^{\mathbb{R}\times T^*Q} : G \times (\mathbb{R} \times T^*Q) \to \mathbb{R} \times T^*Q$ and is defined by

$$\left\langle \Phi_g^{\mathbb{R}\times T^*Q}(t, q, p), (t, q, \dot{q}) \right\rangle = \left\langle (t, q, p), (\Phi_g^t(t), \Phi_g^q(t, q), \partial_q \Phi_g^q(t, q) \cdot \dot{q}) \right\rangle.$$

This has the corresponding *infinitesimal generator* $\xi_{\mathbb{R}\times T^*Q} : \mathbb{R} \times T^*Q \to T(\mathbb{R} \times T^*Q)$ defined by

$$\xi_{\mathbb{R}\times T^*Q}(t, q, p) = \frac{\mathrm{d}}{\mathrm{d}g}\left(\Phi_g^{\mathbb{R}\times T^*Q}(t, q, p)\right) \cdot \xi.$$

The *extended Hamiltonian momentum map* $\bar{J}_H : \mathbb{R} \times T^*Q \to \mathfrak{g}^*$ is defined by

$$\bar{J}_H(t, q, p) \cdot \xi = \bar{\Theta}_H(t, q, p) \cdot \xi_{\mathbb{R}\times T^*Q}(t, q, p).$$

For each $\xi \in \mathfrak{g}$ we define $\bar{J}_H^\xi : \mathbb{R} \times T^*Q \to \mathbb{R}$ by $\bar{J}_H^\xi(t, q, p) = \bar{J}_H(t, q, p) \cdot \xi$, which has expression $\bar{J}_H^\xi = \mathbf{i}_{\xi_{\mathbb{R}\times T^*Q}} \bar{\Theta}_H$. Note that the Hamiltonian map is also given by the expression

$$\bar{J}_H(t, q, p) \cdot \xi = \left\langle (t, q, p), \xi_{\bar{Q}}^q(t, q) \right\rangle - H(t, q, p) \xi_{\bar{Q}}^t(t, q),$$

where the two components of $\xi_{\bar{Q}}$ are denoted by $\xi_{\bar{Q}} = (\xi_{\bar{Q}}^t, \xi_{\bar{Q}}^q)$.

As in the case of the extended Lagrangian system, invariance of the Hamiltonian is not the correct requirement to ensure conservation of the momentum maps. Instead, we define a group action to be a *symmetry* if its

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

lift acts by extended special symplectic maps, which is the requirement that $(\Phi_g^{\mathbb{R}\times T^*Q})^*\bar{\Theta}_H = \bar{\Theta}_H$ for all $g \in G$.

With this requirement the same proof as in Theorem 1.2.2 shows that the extended Hamiltonian momentum map is equivariant, in the sense that $\bar{J}_H \circ \bar{T}^*\Phi_g = \mathrm{Ad}^*_{g^{-1}} \circ \bar{J}_H$.

Theorem 4.4.1. (Extended Hamiltonian Noether's theorem) Let $H : \mathbb{R}\times T^*Q \to \mathbb{R}$ be an extended Hamiltonian system that is invariant under the lift of the (left or right) action $\Phi : G \times \bar{Q} \to \bar{Q}$. Then the corresponding extended Hamiltonian momentum map $\bar{J}_H : \mathbb{R} \times T^*Q \to \mathfrak{g}^*$ is a conserved quantity of the flow, so that $\bar{J}_H \circ \bar{F}_H^h = \bar{J}_H$ for all times h.

Proof. By assumption we have that $(\bar{T}^*\Phi_g)^*\bar{\Theta}_H = \bar{\Theta}_H$ for all $g \in G$ and hence $\mathcal{L}_{\xi_{\mathbb{R}\times T^*Q}}\bar{\Theta}_H = 0$. Now, computing the derivative of $\bar{J}_H^\xi$ in the direction given by the extended Hamiltonian vector field $\bar{X}_H$ gives

$$\begin{aligned}\mathbf{d}\bar{J}_H^\xi \cdot \bar{X}_H &= \mathbf{d}(\mathbf{i}_{\xi_{\mathbb{R}\times T^*Q}}\bar{\Theta}_H)\cdot \bar{X}_H \\ &= \mathcal{L}_{\xi_{\mathbb{R}\times T^*Q}}\bar{\Theta}_H \cdot \bar{X}_H - \mathbf{i}_{\xi_{\mathbb{R}\times T^*Q}}\mathbf{d}\bar{\Theta}_H \cdot \bar{X}_H \\ &= -\mathbf{i}_{\bar{X}_H}\bar{\Omega}_H \cdot \xi_{\mathbb{R}\times T^*Q} \\ &= 0\end{aligned}$$

using Cartan's magic formula $\mathcal{L}_X\alpha = \mathbf{d}\mathbf{i}_X\alpha + \mathbf{i}_X\mathbf{d}\alpha$ and (4.4.3). As $\bar{F}_H^h$ is the flow map for $\bar{X}_H$ this gives the desired result. □

Observe that in this extended case the assumption that $\bar{\Theta}_H$ is preserved by the lifted group action plays the same role as invariance of H did in the autonomous case, when invariance of Θ was a result of the fact that the action was a lift.

4.4.4. Extended Legendre transform

To relate the extended Lagrangian picture developed previously to that of extended Hamiltonian mechanics we define the *extended Legendre transform* $\bar{\mathbb{F}}L : \mathbb{R} \times TQ \to \mathbb{R} \times T^*Q$ to be

$$\bar{\mathbb{F}}L : (t, q, \dot{q}) \mapsto (t, q, p) = \left(t, q, \frac{\partial L}{\partial \dot{q}}(t, q, \dot{q})\right).$$

If $\bar{\mathbb{F}}L$ is a local isomorphism then we say that L is *regular*, and if is a global isomorphism then we say that L is *hyperregular*.

Using the definitions, it is simple to check that the Lagrangian and Hamiltonian symplectic forms and momentum maps are related by

$$\bar{\Theta}_L = (\bar{\mathbb{F}}L)^*\bar{\Theta}_H, \quad \bar{\Omega}_L = (\bar{\mathbb{F}}L)^*\bar{\Omega}_H, \quad \text{and} \quad \bar{J}_L = (\bar{\mathbb{F}}L)^*\bar{J}_H.$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Given a regular extended Lagrangian L, we define the associated extended Hamiltonian by

$$H = E_L \circ (\bar{\mathbb{F}}L)^{-1},$$

which in coordinates is

$$H(t, q, p) = \mathbb{F}L(t, q, \dot{q}) \cdot \dot{q} - L(t, q, \dot{q}),$$

since (t, q, p) and $(t, q, \dot{q})$ are related by the extended Legendre transform. With this additional assumption on H and L we can use calculations similar to those in Section 1.4.3 to see that the Hamiltonian and Lagrangian vector fields and flow maps are related by

$$\bar{X}_L = (\bar{\mathbb{F}}L)^* \bar{X}_H, \quad \bar{F}_L^h = (\bar{\mathbb{F}}L)^{-1} \circ \bar{F}_H^h \circ \bar{\mathbb{F}}L.$$

The definition of regularity of an extended Lagrangian can also be used to give a characterization of when the Euler–Lagrange equations uniquely define solution curves.

Theorem 4.4.2. Given an extended Lagrangian $L : \mathbb{R} \times TQ \to \mathbb{R}$, the extended Lagrangian vector field $\bar{X}_L$, and hence the extended flow map $\bar{F}_L$, is well-defined if and only if the Lagrangian is regular.

Proof. The same coordinate proof as is given for Theorem 1.4.3 also works here. □

4.4.5. Extended generating functions

Consider two extended phase spaces with Hamiltonians H_0 and H_1 and extended canonical two-forms $\bar{\Omega}_{H_0} = -\mathbf{d}\bar{\Theta}_{H_0}$ and $\bar{\Omega}_{H_1} = -\mathbf{d}\bar{\Theta}_{H_1}$ respectively, and let $\varphi : \mathbb{R} \times T^*Q \to \mathbb{R} \times T^*Q$ be a time-dependent transformation. Set $\Gamma(\varphi) \subset (\mathbb{R} \times T^*Q) \times (\mathbb{R} \times T^*Q)$ to be the graph of φ. Consider the one-form on $(\mathbb{R} \times T^*Q) \times (\mathbb{R} \times T^*Q)$ defined by

$$\bar{\Theta}_{H_0,H_1} = \pi_2^* \bar{\Theta}_{H_0} - \pi_1^* \bar{\Theta}_{H_1},$$

where $\pi_i : (\mathbb{R} \times T^*Q) \times (\mathbb{R} \times T^*Q) \to \mathbb{R} \times T^*Q$ are the projections. The corresponding two-form is then

$$\bar{\Omega}_{H_0,H_1} = -\mathbf{d}\bar{\Theta}_{H_0,H_1} = -\pi_2^* \bar{\bar{\Omega}}_{H_0} + \pi_1^* \bar{\bar{\Omega}}_{H_1}.$$

Denoting the inclusion map by $i_\varphi : \Gamma(\varphi) \to (\mathbb{R} \times T^*Q) \times (\mathbb{R} \times T^*Q)$, we see that we have the identities

$$\pi_1 \circ i_\varphi = \pi_1|_{\Gamma(\varphi)}, \tag{4.4.5}$$

$$\pi_2 \circ i_\varphi = \varphi \circ \pi_1 \text{ on } \Gamma(\varphi). \tag{4.4.6}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Using these relations we have

$$\begin{aligned} i_\varphi^*\bar{\Omega}_{H_0,H_1} &= i_\varphi^*(\pi_1^*\bar{\Omega}_{H_0} - \pi_2^*\bar{\Omega}_{H_1}) \\ &= (\pi_1 \circ i_\varphi)^*\bar{\Omega}_{H_0} - (\pi_2 \circ i_\varphi^*\bar{\Omega}_{H_1} \\ &= (\pi_1|_{\Gamma(\varphi)})^*(\bar{\Omega}_{H_0} - \varphi^*\bar{\Omega}_{H_1}), \end{aligned}$$

where we have used (4.4.5) and (4.4.6).

This last relation shows that φ is an extended symplectic transformation if and only if $i_\varphi^*\bar{\Omega}_{H_0,H_1} = 0$ or, equivalently, if and only if $\mathbf{d}(i_\varphi^*\bar{\Theta}_{H_0,H_1}) = 0$. By the Poincaré lemma, this last statement is equivalent to there existing a function $\bar{S} : \Gamma(\varphi) \to \mathbb{R}$ so that, locally, $i_\varphi^*\bar{\Theta}_{H_0,H_1} = \mathbf{d}\bar{S}$. Such a function $\bar{S}$ is known as the *extended generating function* of the extended symplectic transformation φ. Note that $\bar{S}$ is not unique, as it is only defined up to a constant.

The extended generating function $\bar{S}$ is specified on the graph $\Gamma(\varphi)$, and so can be expressed in any local coordinate system on $\Gamma(\varphi)$.

4.4.6. Coordinate expression

We will be particularly interested in the choice (q_0, q_1, t_0) as local coordinates on $\Gamma(\varphi)$, and so we give the coordinate expressions for the above general extended generating function derivation for this particular case. This choice results in extended generating functions of the so-called *first kind*.

Consider a function $\bar{S} : Q \times Q \times \mathbb{R} \to \mathbb{R}$. Now the differential will be

$$\mathbf{d}\bar{S} = \frac{\partial \bar{S}}{\partial q_0}\mathbf{d}q_0 + \frac{\partial \bar{S}}{\partial q_1}\mathbf{d}q_1 + \frac{\partial \bar{S}}{\partial t_0}\mathbf{d}t_0.$$

Let $\varphi : \mathbb{R} \times T^*Q \to \mathbb{R} \times T^*Q$ be the extended symplectic transformation generated by $\bar{S}$. In coordinates, the quantity $i_\varphi^*\bar{\Theta}_{H_0,H_1}$ is

$$i_\varphi^*\bar{\Theta}_{H_0,H_1} = \left(-p_0 - H_1\frac{\partial t_1}{\partial q_0}\right)\mathbf{d}q_0 + \left(p_1 - H_1\frac{\partial t_1}{\partial q_1}\right)\mathbf{d}q_1 + \left(H_0 - H_1\frac{\partial t_1}{\partial t_0}\right)\mathbf{d}t_0$$

and so the condition $i_\varphi^*\bar{\Theta}_{H_0,H_1} = \mathbf{d}\bar{S}$ reduces to the equations

$$\frac{\partial \bar{S}}{\partial q_0} = -p_0 - H_1\frac{\partial t_1}{\partial q_0}, \tag{4.4.7a}$$

$$\frac{\partial \bar{S}}{\partial q_1} = p_1 - H_1\frac{\partial t_1}{\partial q_1}, \tag{4.4.7b}$$

$$\frac{\partial \bar{S}}{\partial t_0} = H_0 - H_1\frac{\partial t_1}{\partial t_0}, \tag{4.4.7c}$$

which are an implicit definition of the transformation $\varphi : (t_0, q_0, p_0) \mapsto (t_1, q_1, p_1)$. From the above general theory, we know that such a transforma-

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

tion is automatically symplectic in an extended sense, and that all extended symplectic transformations have such a representation.

Observe that for some choices of $\bar{S}$ the above equations do not define a single map φ, as was also the case for autonomous generating functions and the extended discrete Euler–Lagrange equations. As before, we will assume for now that $\bar{S}$ is chosen so that there is a single well-defined map φ which satisfies (4.4.7), and we will investigate this issue further in Sections 4.5.4 and 4.9.4.

4.5. Discrete variational mechanics: Hamiltonian viewpoint

4.5.1. Extended discrete Legendre transforms

The fact that the extended discrete state space $\bar{Q} \times \bar{Q}$ is larger than the extended state space $\mathbb{R} \times TQ$ is particularly important when it comes to defining extended discrete Legendre transforms. We will see this below, where the Legendre transform to the Hamiltonian phase space $\mathbb{R} \times T^*Q$ will not be a local isomorphism, and so it will be necessary to define another map to $\mathbb{R} \times (\mathbb{R} \times T^*Q)$ in order to push the extended discrete Lagrangian map forward.

We begin by defining the *extended discrete Legendre transforms* $\bar{\mathbb{F}}^{\pm} L_d : \bar{Q} \times \bar{Q} \to \mathbb{R} \times T^*Q$ to be

$$\bar{\mathbb{F}}^{+} L_d(t_0, q_0, t_1, q_1) = (t_1, q_1, D_4 L_d(t_0, q_0, t_1, q_1)),$$
$$\bar{\mathbb{F}}^{-} L_d(t_0, q_0, t_1, q_1) = (t_0, q_0, -D_2 L_d(t_0, q_0, t_1, q_1)).$$

As $\bar{Q} \times \bar{Q}$ is larger than $\mathbb{R} \times T^*Q$, these maps cannot be even local isomorphisms. If they are both onto, however, then we say that the extended discrete Lagrangian is *regular*, and if they are both globally onto then we say that L_d is *hyperregular*, which will typically require that Q be a linear space.

Note that, in general, symplectic forms and momentum maps do not pull back to their discrete counterparts. If a momentum map arises from a vertical action, however, then we essentially reduce to the autonomous case and it can be seen that the associated extended Hamiltonian momentum map will indeed pull back to the extended discrete momentum maps.

4.5.2. Momentum and energy matching

Just as we earlier defined the discrete energies, we can also define the discrete momenta to be the image of the extended discrete Legendre transforms: that is, we set

$$p^{+}_{k,k+1} = p^{+}(t_k, q_k, t_{k+1}, q_{k+1}) = D_4 L_d(t_k, q_k, t_{k+1}, q_{k+1}),$$
$$p^{-}_{k,k+1} = p^{-}(t_k, q_k, t_{k+1}, q_{k+1}) = -D_2 L_d(t_k, q_k, t_{k+1}, q_{k+1}),$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

or equivalently

$$\bar{\mathbb{F}}^+ L_d(t_k, q_k, t_{k+1}, q_{k+1}) = (t_{k+1}, q_{k+1}, p^+_{k,k+1}),$$
$$\bar{\mathbb{F}}^- L_d(t_k, q_k, t_{k+1}, q_{k+1}) = (t_k, q_k, p^-_{k,k+1}).$$

In other words, $p^+_{k,k+1}$ is the discrete momentum at the right endpoint of the interval $[t_k, t_{k+1}]$, while $p^-_{k,k+1}$ is the discrete momentum at the left endpoint. We also introduce the notation

$$E^+_{k,k+1} = E^+_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}),$$
$$E^-_{k,k+1} = E^-_{L_d}(t_k, q_k, t_{k+1}, q_{k+1}).$$

Using these definitions, we see that the extended discrete Euler–Lagrange equations can be written

$$p^+_{k-1,k} = p^-_{k,k+1},$$
$$E^+_{k-1,k} = E^-_{k,k+1},$$

and thus can be interpreted as a matching of momenta and energies at each time t_k. This is a generalization of the autonomous case, where only the discrete momenta are matched.

The definitions above also allow us to write the discrete Lagrangian one-forms in the more compact form

$$\bar{\Theta}^+_{L_d}(k, k+1) = p^+_{k,k+1}\mathbf{d}q_{k+1} - E^+_{k,k+1}\mathbf{d}t_{k+1},$$
$$\bar{\Theta}^-_{L_d}(k, k+1) = p^-_{k,k+1}\mathbf{d}q_k - E^-_{k,k+1}\mathbf{d}t_k,$$

which makes the analogy to the continuous extended Lagrangian one-form (4.2.11) even more apparent.

Note that, unlike the continuous case, regularity of an extended discrete Lagrangian is not sufficient to ensure that the extended discrete Euler–Lagrange equations have unique solutions, or indeed any solutions at all. We will investigate this issue further in Section 4.9.4.

4.5.3. Extended discrete Hamiltonian maps

The extended discrete Legendre transforms defined above clearly cannot be used to push the extended discrete Lagrangian map $\bar{F}_{L_d} : \bar{Q} \times \bar{Q} \to \bar{Q} \times \bar{Q}$ forward to the Hamiltonian phase space, as these Legendre transforms are not injective. Another way of saying this is that we need to augment the Hamiltonian phase space with time-step information to give a well-defined map.

This results in the space $(\mathbb{R} \times T^*Q) \times \mathbb{R}$ where an element (t, q, p, h) is interpreted as being a point (t, q, p) in phase space together with a time-step h. Given a sequence $\{(t_k, q_k, p_k, h_k)\}_{k=0}^N$, we regard h_k as being the time-step $t_{k+1} - t_k$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Now define the map $\tilde{\mathbb{F}}^- L_d : \bar{Q} \times \bar{Q} \to (\mathbb{R} \times T^*Q) \times \mathbb{R}$ by

$$\tilde{\mathbb{F}}^- L_d : (t_0, q_0, t_1, q_1) \mapsto (t_0, q_0, p_0, h_0) = (t_0, q_0, -D_2 L_d(t_0, q_0, t_1, q_1), t_1 - t_0)$$

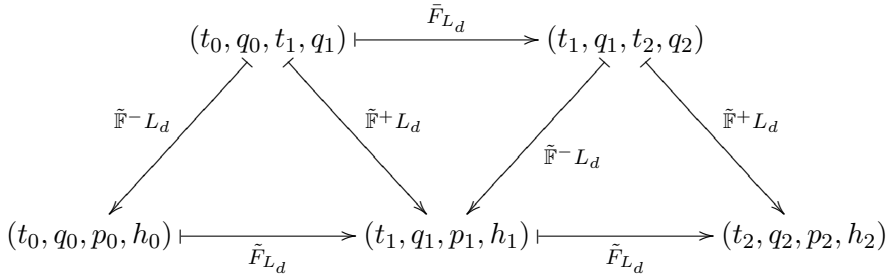
and define the *extended discrete Hamiltonian map* $\tilde{F}_{L_d} : (\mathbb{R} \times T^*Q) \times \mathbb{R} \to (\mathbb{R} \times T^*Q) \times \mathbb{R}$ to be

$$\tilde{F}_{L_d} : (t_0, q_0, p_0, h_0) \mapsto (t_1, q_1, p_1, h_1),$$
$$\tilde{F}_{L_d} = (\tilde{\mathbb{F}}^- L_d) \circ \bar{F}_{L_d} \circ (\tilde{\mathbb{F}}^- L_d)^{-1},$$

which is equivalent to the following commutative diagram.

$$\begin{array}{ccc} (t_0, q_0, t_1, q_1) & \xmapsto{\bar{F}_{L_d}} & (t_1, q_1, t_2, q_2) \\ \Big\downarrow{\scriptstyle \tilde{\mathbb{F}}^- L_d} & & \Big\downarrow{\scriptstyle \tilde{\mathbb{F}}^- L_d} \\ (t_0, q_0, p_0, h_0) & \xmapsto[\tilde{F}_{L_d}]{} & (t_1, q_1, p_1, h_1) \end{array}$$

Using this we can define the map $\tilde{\mathbb{F}}^+ L_d : \bar{Q} \times \bar{Q} \to \mathbb{R} \times T^*Q$ by the following commutative diagram.



Now that we have a discrete system on an extension of the Hamiltonian phase space, it is natural to ask what structure is preserved by the evolution map. Unlike the autonomous case, the pushforward of the extended discrete symplectic form does not give any canonical structure on $(\mathbb{R} \times T^*Q) \times \mathbb{R}$. Of course, it does define a two-form there, which will be preserved by $\tilde{F}_{L_d}$, and we will see below one way in which this is related to $\bar{\Omega}_H$ on $\mathbb{R} \times T^*Q$.

An alternative to mapping to a point $(t, q, p) \in \mathbb{R} \times T^*Q$ and an associated time-step h would be to map to (t, q, p, E), where E is a discrete energy. This is reminiscent of the structure found in the continuous formulation of multisymplectic mechanics, but such discrete systems will often fail to behave well, as is further investigated in Section 4.9.4.

4.5.4. Extended discrete Lagrangians are extended generating functions

Although the extended discrete Hamiltonian map is not a map on the Hamiltonian phase space, we will see that a particular restriction of it is in fact

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

generated by a restriction of the extended discrete Lagrangian, in the sense of Section 4.4.5.

Given $L_d : \bar{Q} \times \bar{Q} \to \mathbb{R}$, choose a fixed time-step $h \in \mathbb{R}$ and define the restriction of the extended discrete Hamiltonian map to $\varphi^h : \mathbb{R} \times T^*Q \to \mathbb{R} \times T^*Q$ by

$$\varphi^h : (t_0, q_0, p_0) \mapsto (t_1, q_1, p_1),$$

where $(t_1, q_1, p_1, h_1) = \tilde{F}_{L_d}(t_0, q_0, p_0, h)$. This provides an implicit definition of (t_0, q_0, t_1, q_1) in terms of either (t_0, q_0, p_0) or (t_1, q_1, p_1). Given this, we define the Hamiltonians H_0^h and H_1^h to be

$$\begin{aligned} H_0^h(t_0, q_0, p_0) &= E_{L_d}^-(t_0, q_0, t_1, q_1), \\ H_1^h(t_1, q_1, p_1) &= E_{L_d}^+(t_0, q_0, t_1, q_1), \end{aligned}$$

and the extended generating function $\bar{S}^h : Q \times Q \times \mathbb{R} \to \mathbb{R}$ to be

$$\bar{S}^h(q_0, q_1, t_0) = L_d(t_0, q_0, t_1, q_1).$$

With these definitions we have the following result.

Theorem 4.5.1. Take a fixed time-step $h \in \mathbb{R}$ and let the functions $\bar{S}^h$, φ^h, H_0^h and H_1^h be defined as above. Then the map φ^h from the space $\mathbb{R} \times T^*Q$ with Hamiltonian H_0^h to the space $\mathbb{R} \times T^*Q$ with Hamiltonian H_1^h is generated by the extended generating function $\bar{S}^h$.

Proof. Observe that the definition of φ^h implies that $t_1 = t_0 + h$. Computing the derivatives of $\bar{S}^h$ and using the definitions above we thus obtain

$$\begin{aligned} \frac{\partial \bar{S}}{\partial q_0} &= D_2 L_d(t_0, q_0, t_1, q_1) = -p_0, \\ \frac{\partial \bar{S}}{\partial q_1} &= D_4 L_d(t_0, q_0, t_1, q_1) = p_1, \\ \frac{\partial \bar{S}}{\partial t_0} &= D_1 L_d(t_0, q_0, t_1, q_1) + D_3 L_d(t_0, q_0, t_1, q_1) \\ &= H_0^h(t_0, q_0, p_0) - H_1^h(t_1, q_1, p_1). \end{aligned}$$

However, these are simply the equations (4.4.7) which define the map generated by $\bar{S}^h$, and so φ^h must be this map. □

As maps generated by extended generating functions must be symplectic, this construction shows that $(\varphi^h)^* \Omega_{H_1^h} = \Omega_{H_0^h}$. This provides a way in which the discrete Hamiltonian map $\tilde{F}_{L_d}$ can be viewed as preserving canonical structures on $\mathbb{R} \times T^*Q$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

4.6. Correspondence between discrete and continuous mechanics

We will now investigate the choice of extended discrete Lagrangian which gives an exact correlation between a continuous and a discrete system, for which we need the following result.

Theorem 4.6.1. Consider a regular extended Lagrangian $L : \mathbb{R} \times TQ \to \mathbb{R}$, two points $q_0, q_1 \in Q$ and two times $t_0, t_1 \in \mathbb{R}$. If $\|q_1 - q_0\|$ and $|t_1 - t_0|$ are sufficiently small then there exists a unique solution $q : [t_0, t_1] \to Q$ of the Euler–Lagrange equations for L satisfying $q(t_0) = q_0$ and $q(t_1) = q_1$.

Proof. Essentially the same proof as that given in Marsden and Ratiu (1999), Section 7.4, for the autonomous case holds in the extended setting as well. □

Given a regular extended Lagrangian $L : \mathbb{R} \times TQ \to \mathbb{R}$, define the *exact extended discrete Lagrangian* to be

$$L_d^E(t_0, q_0, t_1, q_1) = \int_{t_0}^{t_1} L(t, q_{0,1}(t), \dot{q}_{0,1}(t))\,\mathrm{d}t, \tag{4.6.1}$$

where $q_{0,1} : [t_0, t_1] \to Q$ is the unique solution of the Euler–Lagrange equations satisfying $q_{0,1}(t_0) = q_0$ and $q_{0,1}(t_1) = q_1$.

Lemma 4.6.2. A regular extended Lagrangian $L : \mathbb{R} \times TQ \to \mathbb{R}$ and the corresponding exact extended discrete Lagrangian $L_d^E : \bar{Q} \times \bar{Q} \to \mathbb{R}$ satisfy the relations

$$\begin{aligned}
\bar{\mathbb{F}}^+ L_d^E(t_0, q_0, t_1, q_1) &= \bar{\mathbb{F}} L(t_1, q_{0,1}(t_1), \dot{q}_{0,1}(t_1)),\\
E^+_{L_d^E}(t_0, q_0, t_1, q_1) &= E_L(t_1, q_{0,1}(t_1), \dot{q}_{0,1}(t_1)),\\
\bar{\mathbb{F}}^- L_d^E(t_0, q_0, t_1, q_1) &= \bar{\mathbb{F}} L(t_0, q_{0,1}(t_0), \dot{q}_{0,1}(t_0)),\\
E^-_{L_d^E}(t_0, q_0, t_1, q_1) &= E_L(t_0, q_{0,1}(t_0), \dot{q}_{0,1}(t_0)),
\end{aligned}$$

for sufficiently close $q_0, q_1 \in Q$ and $t_0, t_1 \in \mathbb{R}$.

Proof. The calculation for the Legendre transforms is essentially the same as in the time-independent case. For the energies, we calculate

$$\begin{aligned}
\frac{\partial L_d^E}{\partial t_0} &= -L(t_0, q_{0,1}(t_0), \dot{q}_{0,1}(t_0)) + \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial q} \cdot \frac{\partial q_{0,1}}{\partial t_0} + \frac{\partial L}{\partial \dot{q}} \cdot \frac{\partial \dot{q}_{0,1}}{\partial t_0}\right] \mathrm{d}t\\
&= -L(t_0, q_{0,1}(t_0), \dot{q}_{0,1}(t_0)) - \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\right] \cdot \frac{\partial q_{0,1}}{\partial t_0}\,\mathrm{d}t\\
&\qquad + \left[\frac{\partial L}{\partial \dot{q}}(t, q_{0,1}(t), \dot{q}_{0,1}(t)) \cdot \frac{\partial q_{0,1}}{\partial t_0}(t)\right]_{t_0}^{t_1},
\end{aligned}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

using integration by parts. However, $q_{0,1}(t)$ is a solution of the Euler–Lagrange equations for L and so the middle term is zero. Now note that we have

$$\left.\frac{\partial q_{0,1}(t;t_0,q_0,t_1,q_1)}{\partial t_0}\right|_{t=t_0} = \dot{q}_{0,1}(t_0;t_0,q_0,t_1,q_1),$$
$$\left.\frac{\partial q_{0,1}(t;t_0,q_0,t_1,q_1)}{\partial t_0}\right|_{t=t_1} = 0,$$

and using this gives

$$\begin{aligned}\frac{\partial L_d^E}{\partial t_0} &= -L(t_0,q_{0,1}(t_0),\dot{q}_{0,1}(t_0)) + \left[\frac{\partial L}{\partial \dot{q}}(t,q_{0,1}(t),\dot{q}_{0,1}(t))\cdot\frac{\partial q_{0,1}}{\partial t_0}(t)\right]_{t_0}^{t_1}\\ &= -L(t_0,q_{0,1}(t_0),\dot{q}_{0,1}(t_0)) + \frac{\partial L}{\partial \dot{q}}(t_0,q_{0,1}(t_0),\dot{q}_{0,1}(t_0))\cdot\dot{q}_{0,1}(t_0)\\ &= E_L(t_0,q_{0,1}(t_0),\dot{q}_{0,1}(t_0)).\end{aligned}$$

The result for $D_4L_d^E$ can be established by a similar calculation, using the fact that

$$\left.\frac{\partial q_{0,1}(t;t_0,q_0,t_1,q_1)}{\partial t_1}\right|_{t=t_0} = 0,$$
$$\left.\frac{\partial q_{0,1}(t;t_0,q_0,t_1,q_1)}{\partial t_0}\right|_{t=t_1} = -\dot{q}_{0,1}(t_1;t_0,q_0,t_1,q_1). \qquad \square$$

One interesting consequence of this lemma is that, for an exact extended discrete Lagrangian, the extended discrete Euler–Lagrange equations are always functionally dependent. Indeed, the second equation (4.3.7b) becomes a consequence of the first equation (4.3.7a), just as in the continuous case. Before we prove this, however, we give a theorem relating discrete and continuous solution curves.

Theorem 4.6.3. Given a regular extended Lagrangian $L : \mathbb{R}\times TQ \to \mathbb{R}$, let $L_d^E : \bar{Q}\times\bar{Q}\to\mathbb{R}$ be the associated exact extended discrete Lagrangian. Consider a solution $q : [t_0,t_N]\to Q$ of the extended Euler–Lagrange equations for L, and take any sequence $\{t_k\}_{k=0}^N \subset [0,T]$ with sufficiently small $|t_{k+1}-t_k|$. Setting $q_k = q(t_k)$, we now have that $\{(t_k,q_k)\}_{k=0}^N$ is a solution of the extended discrete Euler–Lagrange equations for L_d^E.

Conversely, given any solution $\{(t_k,q_k)\}_{k=0}^N$ of the extended discrete Euler–Lagrange equations for L_d^E, define a curve $q : [t_0,t_N]\to Q$ by $q(t) = q_{k,k+1}(t)$ for $t\in[t_k,t_{k+1}]$, where $q_{k,k+1} : [t_k,t_{k+1}]\to Q$ is the unique solution of the extended Euler–Lagrange equations for L satisfying $q_{k,k+1}(t_k) = q_k$ and $q_{k,k+1}(t_{k+1}) = q_{k+1}$. Then $q(t)$ is a solution of the extended Euler–Lagrange equations for L on $[t_0,t_N]$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Proof. The proof of this theorem is essentially identical to that of Theorem 1.6.4. Forming $\{(t_k, q_k)\}_{k=0}^N$ from a given solution $q(t)$, we see that the discrete Euler–Lagrange equations, which are just a matching of discrete momenta and energies at each t_k, are satisfied because the discrete quantities are equal to the continuous ones.

In the reverse direction, the first part (4.3.7a) of the extended discrete Euler–Lagrange equations, together with Lemma 4.6.2, implies that the first part (4.2.9a) of the Euler–Lagrange equations is satisfied for $q(t)$. It can now be checked, in the same way as the proof of Theorem 1.6.4, that $q(t)$ is C^2. As the second part of the Euler–Lagrange equations is dependent on the first, this means that $q(t)$ automatically satisfies the full extended Euler–Lagrange equations for L. □

Note that in the last part of the above proof the fact that the discrete curve also satisfied the second part (4.3.7b) of the extended discrete Euler–Lagrange equations was not used. This allows us to prove the following.

Corollary 4.6.4. Given a regular extended Lagrangian $L : \mathbb{R} \times TQ \to \mathbb{R}$, let $L_d^E : \bar{Q} \times \bar{Q} \to \mathbb{R}$ be the associated exact extended discrete Lagrangian. Then the second part (4.3.7b) of the extended discrete Euler–Lagrange equations for L_d^E is satisfied whenever the first part (4.3.7a) is satisfied.

Proof. Consider points $(t_0, q_0, t_1, q_1, t_2, q_2)$ which satisfy the first part of the extended discrete Euler–Lagrange equations for L_d^E. Now define $q_{0,1,2}(t) : [t_0, t_2] \to Q$ by $q_{0,1,2}(t) = q_{k,k+1}(t)$ for $t \in [t_k, t_{k+1}]$, as in Theorem 4.6.3. From the proof of that theorem it is clear that $q(t)$ is a solution of the Euler–Lagrange equations for L on $[t_0, t_2]$, and thus

$$\begin{aligned} E_L(t_1, q_{0,1}(t_1), \dot{q}_{0,1}(t_1)) &= E_L(t_1, q_{0,1,2}(t_1), \dot{q}_{0,1,2}(t_1)) \\ &= E_L(t_1, q_{1,2}(t_1), \dot{q}_{1,2}(t_1)). \end{aligned}$$

By Lemma 4.6.2, however, the left- and right-hand parts of this expression give

$$E^+_{L_d^E}(t_0, q_0, t_1, q_1) = E^-_{L_d^E}(t_1, q_1, t_2, q_2)$$

which is exactly the second part of the extended discrete Euler–Lagrange equations for L_d^E. □

Indeed, as we will see in Section 4.8, the above statement is actually both necessary and sufficient for L_d to be an exact extended discrete Lagrangian.

Note that this corollary means that the exact extended discrete Lagrangian is not discretely well-posed, and thus does not define an extended discrete Lagrangian map $\bar{F}_{L_d} : \bar{Q} \times \bar{Q} \to \bar{Q} \times \bar{Q}$. Indeed, any time-step forward will give a valid solution. This means that the statements about

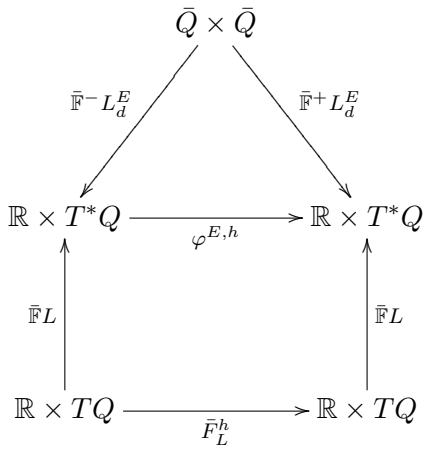
https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

symplecticity and momentum conservation do not hold literally in this case. Instead, we must consider a generalized first variation interpretation of symplecticity or momentum conservation, which would hold for any tangent vectors to the set of solution curves. See Marsden *et al.* (1998) for the outline of this idea, although it is used there for a different reason.

As we have seen above, it is exactly because any time-step gives a solution of the extended discrete Euler–Lagrange equations for L_d^E that we cannot define $\bar{F}_{L_d}$ or $\tilde{F}_{L_d}$. For a given time-step h, however, we can define the map $\varphi^{E,h} : \mathbb{R} \times T^*Q \to \mathbb{R} \times T^*Q$ by the conditions

$$\begin{aligned} \varphi^{E,h} \circ \bar{\mathbb{F}}^- L_d^E &= \bar{\mathbb{F}}^+ L_d^E, \\ t_1 &= t_0 + h, \end{aligned}$$

where $(t_1, q_1, p_1) = \varphi^{E,h}(t_0, q_0, p_0)$. This map plays the same role as the restriction of $\tilde{F}_{L_d}$ to φ^h defined in Section 4.5.4, although $\varphi^{E,h}$ is not the restriction of anything. Using now the fact that $\bar{F}_L^h(t_0, q_{0,1}(0), \dot{q}_{0,1}(0)) = (t_1, q_{0,1}(h), \dot{q}_{0,1}(h))$ and combining the definition of $\varphi^{E,h}$ with Lemma 4.6.2 shows that $\varphi^{E,h}$ satisfies the following commutative diagram.



This proves the following theorem.

Theorem 4.6.5. Consider a regular extended Lagrangian $L : \mathbb{R} \times TQ$ with corresponding extended Hamiltonian $H : \mathbb{R} \times T^*Q$ and exact extended discrete Lagrangian $L_d^E : \bar{Q} \times \bar{Q} \to \mathbb{R}$. Then $\varphi^{E,h}$ defined as above is exactly the extended Hamiltonian flow map $\bar{F}_H^h$.

The above construction will be useful when we consider the relationship between exact extended discrete Lagrangians and extended Hamilton–Jacobi theory in Section 4.8.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

4.7. Background: Extended Hamilton–Jacobi theory

4.7.1. Extended generating function for the flow

As we have already shown, the flow map $\bar{F}_H^h : (t_0, q_0, p_0) \mapsto (t_1, q_1, p_1)$ of an extended Hamiltonian system is an extended symplectic map for each fixed h. From the preceding theory, it must therefore have a representation in terms of some extended generating function $\bar{S}^h$. The extended Hamilton–Jacobi equation is a PDE that defines this generating function.

Considering that the flow map acts

$$\bar{F}_H : (\mathbb{R} \times T^*Q) \times \mathbb{R} \to (\mathbb{R} \times T^*Q)$$

we calculate the tangent map of this to be

$$\begin{aligned} T\bar{F}_H((t,q,p),h)\cdot((\delta t,\delta q,\delta p),\delta h) \\ &= T\bar{F}_H^h(t,q,p)\cdot(\delta t,\delta q,\delta p) + \frac{\partial}{\partial h}\bar{F}_H^h(t,q,p)\cdot\delta h \\ &= T\bar{F}_H^h(t,q,p)\cdot(\delta t,\delta q,\delta p) + \bar{X}_H \circ \bar{F}_H^h(t,q,p)\cdot\delta h, \end{aligned}$$

using the fact that $\bar{F}_H$ is the flow of the vector field $\bar{X}_H$. This then shows that

$$\bar{F}_H^*(\bar{\Omega}_H) = (\pi_{\mathbb{R}\times T^*Q})^*(\bar{\Omega}_H),$$

which follows from (4.4.3). Consider now the space $(\mathbb{R}\times T^*Q)\times\mathbb{R}\times(\mathbb{R}\times T^*Q)$ with projections

$$\begin{aligned} \pi_1 &: (\mathbb{R} \times T^*Q) \times \mathbb{R} \times (\mathbb{R} \times T^*Q) \to (\mathbb{R} \times T^*Q) \times \mathbb{R}, \\ \pi_2 &: (\mathbb{R} \times T^*Q) \times \mathbb{R} \times (\mathbb{R} \times T^*Q) \to \mathbb{R} \times T^*Q \end{aligned}$$

onto the first and second parts, respectively. Define the one-form

$$\hat{\Theta} = \pi_2^*\bar{\Theta}_H - \pi_1^*\pi_{\mathbb{R}\times T^*Q}^*\bar{\Theta}_H$$

and the corresponding two-form

$$\hat{\Omega} = -\mathbf{d}\hat{\Theta} = \pi_1^*\pi_{\mathbb{R}\times T^*Q}^*\bar{\Omega}_H - \pi_2^*\bar{\Omega}_H.$$

The graph of the flow map is a subset $\Gamma(\bar{F}_H) \subset (\mathbb{R}\times T^*Q)\times\mathbb{R}\times(\mathbb{R}\times T^*Q)$, and we denote the corresponding inclusion map by $i_{\bar{F}_H} : \Gamma(\bar{F}_H) \to (\mathbb{R} \times T^*Q) \times \mathbb{R} \times (\mathbb{R} \times T^*Q)$. Now note that

$$\begin{aligned} \pi_1 \circ i_{\bar{F}_H} &= \pi_1|_{\Gamma(\bar{F}_H)}, \\ \pi_2 \circ i_{\bar{F}_H} &= \bar{F}_H \circ \pi_1 \text{ on } \Gamma(\bar{F}_H), \end{aligned}$$

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

with which we calculate

$$\begin{aligned}(i_{\bar{F}_H})^*\hat{\Omega} &= (i_{\bar{F}_H})^*(\pi_1)^*(\pi_{\mathbb{R}\times T^*Q})^*(\bar{\Omega}_H) - (i_{\bar{F}_H})^*(\pi_2)^*(\bar{\Omega}_H)\\ &= (\pi_{\mathbb{R}\times T^*Q}\circ\pi_1\circ i_{\bar{F}_H})^*(\bar{\Omega}_H) - (\pi_2\circ i_{\bar{F}_H})^*(\bar{\Omega}_H)\\ &= (\pi_{\mathbb{R}\times T^*Q}\circ\pi_1\circ i_{\bar{F}_H})^*(\bar{\Omega}_H) - (\bar{F}_H\circ\pi_1\circ i_{\bar{F}_H})^*(\bar{\Omega}_H)\\ &= (\pi_{\mathbb{R}\times T^*Q}\circ\pi_1\circ i_{\bar{F}_H})^*(\bar{\Omega}_H) - (\pi_{\mathbb{R}\times T^*Q}\circ\pi_1\circ i_{\bar{F}_H})^*(\bar{\Omega}_H)\\ &= 0.\end{aligned}$$

This thus establishes that $\mathbf{d}(i^*_{\bar{F}_H}\hat{\Theta}) = 0$, and so there is locally a function $\bar{S} : \Gamma(\bar{F}_H) \to \mathbb{R}$ with $i^*_{\bar{F}_H}\hat{\Theta} = \mathbf{d}\bar{S}$.

4.7.2. Extended Hamilton–Jacobi equation

We now choose a particular set of coordinates on $\Gamma(\bar{F}_H)$ and derive a coordinate expression for $i^*_{\bar{F}_H}\hat{\Theta} = \mathbf{d}\bar{S}$. Taking coordinates on $(\mathbb{R}\times T^*Q)\times\mathbb{R}\times(\mathbb{R}\times T^*Q)$ of $((t_0,q_0,p_0),h,(t_1,q_1,p_1))$, we will take (t_0,q_0,t_1,q_1) as coordinates on $\Gamma(\bar{F}_H)$. The function $\bar{S}$ is thus a map $\bar{S} : \mathbb{R}\times Q\times\mathbb{R}\times Q\to\mathbb{R}$ and has differential

$$\mathbf{d}\bar{S} = \frac{\partial\bar{S}}{\partial t_0}\mathbf{d}t_0 + \frac{\partial\bar{S}}{\partial q_0}\mathbf{d}q_0 + \frac{\partial\bar{S}}{\partial t_1}\mathbf{d}t_1 + \frac{\partial\bar{S}}{\partial q_1}\mathbf{d}q_1.$$

In these coordinates we also have

$$\hat{\Omega} = (p_1\mathbf{d}q_1 - H(t_1,q_1,p_1)\mathbf{d}t_1) - (p_0\mathbf{d}q_0 - H(t_0,q_0,p_0)\mathbf{d}t_0).$$

Equating coefficients, we now see that the equation $i^*_{\bar{F}_H}\hat{\Theta} = \mathbf{d}\bar{S}$ is

$$p_0 = -\frac{\partial\bar{S}}{\partial q_0}(t_0,q_0,t_1,q_1), \tag{4.7.1a}$$

$$H\left(t_0,q_0,-\frac{\partial\bar{S}}{\partial q_0}(t_0,q_0,t_1,q_1)\right) = \frac{\partial\bar{S}}{\partial t_0}(t_0,q_0,t_1,q_1), \tag{4.7.1b}$$

$$p_1 = \frac{\partial\bar{S}}{\partial q_1}(t_0,q_0,t_1,q_1), \tag{4.7.1c}$$

$$H\left(t_1,q_1,\frac{\partial\bar{S}}{\partial q_1}(t_0,q_0,t_1,q_1)\right) = -\frac{\partial\bar{S}}{\partial t_1}(t_0,q_0,t_1,q_1). \tag{4.7.1d}$$

Of these four equations, (4.7.1a), (4.7.1c) and (4.7.1d) can be regarded as the generating function equations (4.4.7) for the map $\bar{F}_H$, while (4.7.1b) is the equation which must be satisfied if $\bar{F}_H$ is to be the Hamiltonian flow. Note that the first three equations do not actually specify $\bar{F}_H$ uniquely, as any time reparametrization of the flow will satisfy them. To ensure uniqueness, we must augment the above equations with the condition $t_1 = t_0 + h$.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

The last equation which $\bar{S}$ must satisfy to be an extended generating function for the flow is known as the *extended Hamilton–Jacobi equation* and is given by

$$H\left(t_0, q_0, -\frac{\partial \bar{S}}{\partial q_0}(t_0, q_0, t_1, q_1)\right) = \frac{\partial \bar{S}}{\partial t_0}(t_0, q_0, t_1, q_1). \tag{4.7.2}$$

As in the autonomous case, it is necessary to specify boundary conditions for this partial differential equation.

4.7.3. Jacobi's solution

The solution of the extended Hamilton–Jacobi equation can be written in terms of the action associated to the Lagrangian which is the Legendre transform of H. This is known as the *extended Jacobi's solution* and is

$$\bar{S}(t_0, q_0, t_1, q_1) = \int_{t_0}^{t_1} L(t, q(t), \dot{q}(t))\,\mathrm{d}t, \tag{4.7.3}$$

where $q : [t_0, t_1] \to Q$ is the unique solution of the extended Euler–Lagrange equations for L satisfying $q(t_0) = q_0$ and $q(t_1) = q_1$. The proof that this is indeed a solution is essentially identical to the proof of Lemma 4.6.2.

4.8. Discrete variational mechanics: Hamilton–Jacobi viewpoint

For a fixed time-step h, we saw in Section 4.5.4 that the restriction of $\tilde{F}_{L_d}$ to the extended phase space $\mathbb{R} \times T^*Q$ is generated by the restriction of the discrete Lagrangian to the set $t_1 - t_0 = h$. We then showed that there is a particular choice of extended discrete Lagrangian, called the exact extended discrete Lagrangian, for a particular restricted map is exactly the extended Hamiltonian flow map.

In the preceding section we have seen that the generating function for the extended Hamiltonian flow map must satisfy the extended Hamilton–Jacobi equation. It is clear that this is the case for the exact extended discrete Lagrangian, as it is simply the standard Jacobi's solution to the extended Hamilton–Jacobi equation.

Note that the story is not as simple as in the autonomous case, however. The extended discrete Lagrangian in fact defines a map $\tilde{F}_{L_d}$ on the enlarged space $(\mathbb{R}\times T^*Q)\times\mathbb{R}$, and it is only upon restriction that we have a generating function construction. This restriction essentially discards the equation for updating the time-step.

The definition of $\tilde{F}_{L_d}$ relied upon L_d being discretely well-posed, which fails for the exact discrete Lagrangian. Nonetheless, we saw that by adding the condition $t_1 = t_0 + h$ then the *restriction* of $\tilde{F}_{L_d}$ still makes sense even

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

when L_d is not discretely well-posed, and it is this map which is exactly the Hamiltonian flow. It is precisely this same extra equation which was also necessary to specify the map generated by the solution to the Hamilton–Jacobi equation.

Another way of viewing the situation is to consider a Hamiltonian system $H : \mathbb{R} \times T^*Q \to \mathbb{R}$ and to attempt to find a discrete Lagrangian for which $(\bar{\mathbb{F}}^{\pm} L_d)^* H = E^{\pm}_{L_d}$ or, equivalently, $(\bar{\mathbb{F}}^{\pm} L_d)^* \Theta_H = \Theta_{L_d}$. Such a discrete Lagrangian would have the appealing property that $\tilde{F}_{L_d}$ would exactly preserve the true energy, symplectic form and momentum maps.

Considering the statement $(\bar{\mathbb{F}}^{+} L_d)^* H = E^{+}_{L_d}$, however, we see that this is exactly the extended Hamilton–Jacobi PDE: that is, the only extended discrete Lagrangian for which the pullback of the Hamiltonian is the discrete energy is the exact extended discrete Lagrangian. The same statement also holds with energy replaced by canonical one-forms.

Note that this does not say that the standard autonomous discrete Hamiltonian map will preserve H if and only if it exactly integrates the flow, as H could be preserved even though $E^{\pm}_{L_d}$ was not equal to $(\bar{\mathbb{F}}^{\pm} L_d)^* H$. Indeed, as we saw above, this would be the case for reparametrizations of the flow map, as discussed in Ge and Marsden (1988).

4.9. Time-dependent variational integrators

Just as we can regard autonomous discrete variational systems as integrators for a given autonomous Lagrangian system, so too can we consider an extended discrete Lagrangian system as an integrator for an extended Lagrangian system $L : \mathbb{R} \times TQ \to \mathbb{R}$. For this to be the case, we need that the discrete Lagrangian $L_d : \bar{Q} \times \bar{Q} \to \mathbb{R}$ satisfies

$$L_d(t_0, q(t_0), t_1, q(t_1)) = \int_{t_0}^{t_1} L\big(t, q(t), \dot{q}(t)\big)\, \mathrm{d}t + \mathcal{O}(|t_1 - t_0|^{r+1}),$$

where $q : [t_0, t_1] \to Q$ is any solution of the extended Euler–Lagrange equations for L. Here r is known as the *order* of the discrete Lagrangian, and we require $r \geq 1$ for L_d to be *consistent*.

Having chosen an extended discrete Lagrangian of some order, we can then use the extended discrete Hamiltonian map $\tilde{F}_{L_d} : (\mathbb{R} \times T^*Q) \times \mathbb{R} \to (\mathbb{R} \times T^*Q) \times \mathbb{R}$ as an integrator for the Hamiltonian system associated to L. This maps from the point (t_k, q_k, p_k) and the time-step h_k to give a new point $(t_{k+1}, q_{k+1}, p_{k+1})$, as described in Section 4.5.3.

Although this map, based on updating time-steps, was useful for understanding the relationship of extended discrete variational mechanics to extended Hamilton–Jacobi theory, it turns out to be more convenient to implement an equivalent method which updates points in conjunction with

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

energies, rather than time-steps. To construct this, we define the maps

$$\hat{\mathbb{F}}^+ L_d : (t_0, q_0, t_1, q_1) \mapsto (t_1, q_1, p_1 = D_4 L_d, E_1 = -D_3 L_d),$$
$$\hat{\mathbb{F}}^- L_d : (t_0, q_0, t_1, q_1) \mapsto (t_0, q_0, p_0 = -D_2 L_d, E_0 = D_1 L_d),$$

and define the alternative extended discrete Hamiltonian map $\hat{F}_{L_d} : \mathbb{R} \times T^*Q \times \mathbb{R} \to \mathbb{R} \times T^*Q \times \mathbb{R}$ by

$$\hat{F}_{L_d} = \hat{\mathbb{F}}^+ L_d \circ (\hat{\mathbb{F}}^- L_d)^{-1}.$$

Regarding $\hat{F}_{L_d}$ as an integrator, it maps

$$\hat{F}_{L_d} : (t_k, q_k, p_k, E_k) \mapsto (t_{k+1}, q_{k+1}, p_{k+1}, E_{k+1})$$

and is defined by the relations

$$p_k = -D_2 L_d(t_k, q_k, t_{k+1}, q_{k+1}), \tag{4.9.1a}$$
$$E_k = D_1 L_d(t_k, q_k, t_{k+1}, q_{k+1}), \tag{4.9.1b}$$
$$p_{k+1} = D_4 L_d(t_k, q_k, t_{k+1}, q_{k+1}), \tag{4.9.1c}$$
$$E_{k+1} = -D_3 L_d(t_k, q_k, t_{k+1}, q_{k+1}), \tag{4.9.1d}$$

together with the requirement that $t_{k+1} > t_k$, which is built in to the definition of the extended discrete path space. Computationally, the implicit equations (4.9.1a) and (4.9.1b) must be solved simultaneously for t_{k+1} and q_{k+1} under the restriction $t_{k+1} > t_k$, and then the explicit equations (4.9.1c) and (4.9.1d) must be evaluated to give p_{k+1} and E_{k+1}. Note that the calculations to produce a trajectory with $\tilde{F}_{L_d}$ are the same as those for calculating the same trajectory with $\hat{F}_{L_d}$, but the presentation is clearer in this form.

As we have seen in previous sections, the method defined by (4.9.1) preserves an extended discrete symplectic form and extended discrete momentum maps and it satisfies a discrete energy evolution equation which exactly conserves the discrete energy if the discrete Lagrangian is autonomous.

4.9.1. Initial conditions

To actually use $\hat{F}_{L_d}$ as an integrator, it is necessary to choose initial conditions. Given $(t_0, q_0, p_0) \in \mathbb{R} \times T^*Q$, this reduces to the question of choosing an initial energy $E_0 \in \mathbb{R}$.

While the natural choice may, at first, appear to be the Hamiltonian H evaluated at (t_0, q_0, p_0), this is not a feasible option. As we have seen, requiring that the Hamiltonian coincides with the discrete energies is equivalent to the extended Hamilton–Jacobi PDE, and so using this as an initial condition will generally make the first step of $\hat{F}_{L_d}$ ill-defined. Even when this equation is solvable, this approach provides no control over the size of the initial time-step.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Instead, a better choice is to select an initial time-step $h_0 \in \mathbb{R}$ and then to set $t_1 = t_0 + h_0$ and to solve the equation

$$p_0 = -D_2 L_d(t_0, q_0, t_1, q_1)$$

for q_1. We can then evaluate $E_0 = E^-_{L_d}(t_0, q_0, t_1, q_1)$ to obtain the initial condition for the discrete energy.

To see the difficulties encountered with choosing $E_h = H(t_0, q_0, p_0)$, consider the following example.

Example 4.9.1. Taking the discrete Lagrangian L^α_d from Example 2.3.2 with $\alpha = 1$ and extending it to the time-dependent case gives

$$L_d(t_0, q_0, t_1, q_1) = (t_1 - t_0) L\left(t_1, q_1, \frac{q_1 - q_0}{t_1 - t_0}\right),$$

and for the case of the autonomous Lagrangian $L(t, q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q)$ we can calculate

$$\bar{\mathbb{F}}^- L_d(t_0, q_0, t_1, q_1) = \left(t_0, q_0, M\left(\frac{q_1 - q_0}{t_1 - t_0}\right)\right),$$

$$E^-_{L_d}(t_0, q_0, t_1, q_1) = \frac{1}{2}\left(\frac{q_1 - q_0}{t_1 - t_0}\right)^T M \left(\frac{q_1 - q_0}{t_1 - t_0}\right) + V(q_1),$$

and hence

$$\big((\mathbb{F}^- L_d)^* H\big)(t_0, q_0, t_1, q_1) = \frac{1}{2}\left(\frac{q_1 - q_0}{t_1 - t_0}\right)^T M \left(\frac{q_1 - q_0}{t_1 - t_0}\right) + V(q_0).$$

If we now take $E_0 = H(t_0, q_0, p_0)$ and calculate the first step of $\hat{F}_{L_d}$, then equations (4.9.1a) and (4.9.1b) are

$$(t_0, q_0, p_0) = \bar{\mathbb{F}}^- L_d(t_0, q_0, t_1, q_1),$$
$$\big((\mathbb{F}^- L_d)^* H\big)(t_0, q_0, t_1, q_1) = E^-_{L_d}(t_0, q_0, t_1, q_1),$$

which reduce to

$$q_1 = q_0 + (t_1 - t_0) M^{-1} p_0,$$
$$V(q_1) = V(q_0).$$

For a nontrivial potential function $V : Q \to \mathbb{R}$ these equations will not have a solution (t_1, q_1), except for special choices of p_0. ◊

4.9.2. Order of accuracy and local errors

As we have seen while considering the choice of initial conditions, and as we will discuss in more detail below, existence of solutions for an extended discrete Lagrangian system can be problematic for their use as integrators. For this reason we will not discuss the order of accuracy of the extended discrete

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

Hamiltonian map $\tilde{F}_{L_d}$ or the variant $\hat{F}_{L_d}$ here, but instead we will consider the relationship between the order of accuracy of the extended discrete Lagrangian and its discrete energies and extended Legendre transforms.

Theorem 4.9.2. Consider an extended Lagrangian system $L : \mathbb{R} \times TQ$ and a corresponding extended discrete Lagrangian system $L_d : \bar{Q} \times \bar{Q} \to \mathbb{R}$ of order r. Then the extended discrete Legendre transforms $\bar{\mathbb{F}}^{\pm} L_d : \bar{Q} \times \bar{Q} \to \mathbb{R} \times T^*Q$ and the discrete energies $E^{\pm}_{L_d} : \bar{Q} \times \bar{Q} \to \mathbb{R}$ are also of order r, in the sense that

$$\begin{aligned}
\bar{\mathbb{F}}^{+} L_d(t_0, q(t_0), t_1, q(t_1)) &= \bar{\mathbb{F}} L(t_1, q(t_1), \dot{q}(t_1)) + \mathcal{O}(|t_1 - t_0|^{r+1}),\\
E^{+}_{L_d}(t_0, q(t_0), t_1, q(t_1)) &= E_L(t_1, q(t_1), \dot{q}(t_1)) + \mathcal{O}(|t_1 - t_0|^{r+1}),\\
\bar{\mathbb{F}}^{-} L_d(t_0, q(t_0), t_1, q(t_1)) &= \bar{\mathbb{F}} L(t_0, q(t_0), \dot{q}(t_0)) + \mathcal{O}(|t_1 - t_0|^{r+1}),\\
E^{-}_{L_d}(t_0, q(t_0), t_1, q(t_1)) &= E_L(t_0, q(t_0), \dot{q}(t_0)) + \mathcal{O}(|t_1 - t_0|^{r+1}),
\end{aligned}$$

for any solution $q : [t_0, t_1] \to Q$ of the extended Euler–Lagrange equations for L.

Proof. This can be proved in the same way as Theorem 2.3.1, using the results of Lemma 4.6.2. □

4.9.3. Autonomous discrete Lagrangians

If we have an autonomous discrete Lagrangian $L'_d : Q \times Q \to \mathbb{R}$, then we can either treat it using the autonomous theory and obtain the integrator $\tilde{F}_{L'_d} : T^*Q \to T^*Q$, or we can form the corresponding extended discrete Lagrangian by

$$L_d(t_0, q_0, t_1, q_1) = L'_d(q_0, q_1, t_1 - t_0),$$

as was considered in Section 4.3.7. As we saw there, the time translation invariance of L_d means that the two discrete energies are equal, and using the map $\hat{F}_{L_d}$ will integrate the system while preserving the discrete energy. Indeed, this method is exactly the symplectic-energy-momentum-conserving integrator presented in Kane *et al.* (1999*a*), written in a slightly different form. That paper gives some simple specific examples that show the advantage in using these types of integrators.

Although using the autonomous integrator $\tilde{F}_{L'_d}$ for this system will not exactly preserve the discrete energy or the Hamiltonian, backward error analysis techniques can be used to show that the Hamiltonian, and hence the discrete energies, will be nearly preserved for exponentially long times (see, for example, Hairer (1994), Hairer and Lubich (1997) and Reich (1999*a*)).

For autonomous systems, it will thus generally be preferable to use a standard variational integrator, and not to enforce exact energy conservation. Even in this case, however, the theory developed above for non-autonomous systems gives much insight into the geometric structure.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

4.9.4. Existence of solutions

The question of the existence of solutions to the extended discrete Euler–Lagrange equations is, in general, rather subtle. There are two main problems which may arise.

First, it may be that for certain choices of initial condition (t_0, q_0, p_0, E_0), or correspondingly (t_0, q_0, t_1, q_1), that the extended discrete Euler–Lagrange equations simply have no solutions.

For specific choices of discrete Lagrangian it is sometimes possible to analyse this situation in more detail. For example, in Kane *et al.* (1999*a*) it is proved that, for Lagrangians of the form $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q)$ and the discrete Lagrangian L_d^α from Example 2.3.2, then the extended discrete Euler–Lagrange equations have unique solutions away from points where the kinetic energy is near zero. This criterion of being away from turning points is illustrated in that paper with numerical examples.

A particular choice of discrete Lagrangian, albeit in a different setting, is also analysed in Lew, Marsden, Ortiz and West (2001), where the existence of solutions to a time-adaptive variational PDE integrator is investigated.

The second problem which may arise is that the solution times t_k may not be unbounded, that is, they may converge to some finite value $t_k \to T \in \mathbb{R}$. This is similar to the issue of completeness of the differential equations in the continuous case, where a solution may escape to infinity in finite time.

4.9.5. Examples of extended variational integrators

To actually construct an extended discrete Lagrangian given an extended Lagrangian, most of the methods discussed in Section 2.6 can be readily extended to the nonautonomous case. Typically the time-step h should be replaced by $t_{k+1} - t_k$ and the continuous Lagrangian should be evaluated at the appropriate interpolated times. We give here a number of examples of this.

Example 4.9.3. (Low-order methods) Consider the discrete Lagrangian L_d^α from Example 2.3.2, and extend it to a nonautonomous Lagrangian by

$$\bar{L}_d^\alpha(t_0, q_0, t_1, q_1) = (t_1 - t_0) L\left((1-\alpha)t_0 + \alpha t_1, (1-\alpha)q_0 + \alpha q_1, \frac{q_1 - q_0}{t_1 - t_0}\right).$$

Similarly we can extend $L_d^{\mathrm{sym},\alpha}$ to

$$\bar{L}_d^{\mathrm{sym},\alpha} = \frac{1}{2}\bar{L}_d^\alpha + \frac{1}{2}\bar{L}_d^{1-\alpha}.$$

It is straightforward to check that the order of these discrete Lagrangians is the same as in the autonomous case, namely $\bar{L}_d^\alpha$ is first-order unless $\alpha = \frac{1}{2}$, when it is second-order, and $\bar{L}_d^{\mathrm{sym},\alpha}$ is always second-order.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

This is reflected in the expression for the discrete energy $E^-_{\bar{L}^1_d}$ given in Example 4.9.1 above, which is clearly a first-order approximation to the Hamiltonian, as implied by Theorem 4.9.2. ◊

Example 4.9.4. (Symplectic partitioned Runge–Kutta methods) The symplectic partitioned Runge–Kutta methods discussed in Section 2.6.5 have a standard extension to nonautonomous systems, although they will clearly not be symplectic in the normal sense on T^*Q in that case. Taking coefficients c_i, b_i, a_{ij}, $\tilde{c}_i$, $\tilde{b}_i$, $\tilde{a}_{ij}$ for $i, j = 1, \ldots, s$, then the corresponding partitioned Runge–Kutta method for a regular extended Lagrangian system $L : \mathbb{R} \times TQ \to \mathbb{R}$ is a map $(t_0, q_0, p_0, h) \mapsto (t_1 = t_0 + h, q_1, p_1)$ defined by

$$q_1 = q_0 + (t_1 - t_0) \sum_{j=1}^{s} b_j \dot{Q}_j, \quad p_1 = p_0 + (t_1 - t_0) \sum_{j=1}^{s} \tilde{b}_j \dot{P}_j, \tag{4.9.2a}$$

$$Q_i = q_0 + (t_1 - t_0) \sum_{j=1}^{s} a_{ij} \dot{Q}_j, \quad P_i = p_0 + (t_1 - t_0) \sum_{j=1}^{s} \tilde{a}_{ij} \dot{P}_j, \quad i = 1, \ldots, s, \tag{4.9.2b}$$

$$P_i = \frac{\partial L}{\partial \dot{q}}(t_i, Q_i, \dot{Q}_i), \qquad \dot{P}_i = \frac{\partial L}{\partial q}(\tilde{t}_i, Q_i, \dot{Q}_i), \qquad i = 1, \ldots, s, \tag{4.9.2c}$$

where $t_i = (1 - c_i)t_0 + c_i t_1$ and $\tilde{t}_i = (1 - \tilde{c}_i)t_0 + \tilde{c}_i t_1$ are the interpolated times. As in the autonomous case, we can regard (4.9.2) as defining p_0, p_1, Q_i, P_i, $\dot{Q}_i$ and $\dot{P}_i$ for $i = 1, \ldots, s$ as functions of (t_0, q_0, t_1, q_1). Assuming this, we define the extended discrete Lagrangian

$$L_d(t_0, q_0, t_1, q_1) = (t_1 - t_0) \sum_{i=1}^{s} b_i L\big(t_i, Q_i, \dot{Q}_i\big),$$

and, if the coefficients satisfy the extended conditions

$$b_i \tilde{a}_{ij} + \tilde{b}_j a_{ji} = b_i \tilde{b}_j, \qquad i, j = 1, \ldots, s, \tag{4.9.3a}$$

$$b_i = \tilde{b}_i, \qquad i = 1, \ldots, s, \tag{4.9.3b}$$

$$c_i = \tilde{c}_i, \qquad i = 1, \ldots, s, \tag{4.9.3c}$$

then it is clear from the result in Section 2.6.5 that the partitioned Runge–Kutta method is exactly the restriction φ^h of the extended discrete Hamiltonian map $\tilde{F}_{L_d}$ defined in Section 4.5.4.

If we instead use the full map $\tilde{F}_{L_d}$ or the alternative $\hat{F}_{L_d}$ then the theory developed above shows that we will have a symplectic integrator in the extended sense. ◊

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

PART FIVE
Further topics

In this part we briefly discuss some additional topics and future directions that are related to those covered in this review. We cite relevant references in the literature for further information.

5.1. Discrete symmetry reduction

The theory and applications of reduction of mechanical systems with symmetry both from the Hamiltonian and Lagrangian perspectives have been an active area of investigation for quite some time. For a recent review, see Marsden, Ratiu and Scheurle (2000). It is natural to investigate the discrete counterpart of this theory and such an investigation has begun. The discrete analogue of the Euler–Poincaré equations and reduction theory (rigid body and fluid-type equations on Lie algebras) lead to the DEP (discrete Euler–Poincaré) equations: see Marsden *et al.* (1999*a*, 1999*b*), and Bobenko and Suris (1999*a*, 1999*b*). The latter references also make intriguing links with discrete integrable systems. Other intriguing links between discrete mechanics, rigid body systems and optimal control are given in Bloch, Crouch, Marsden and Ratiu (1998).

The DEP context assumes that the configuration space of the system is a Lie group. The development of a more general reduction theory for group actions on more general configuration manifolds has begun, with the case of abelian group actions given in Jalnapurkar, Leok, Marsden and West (2000). The general nonabelian case as well as discrete analogues of singular reduction are also of considerable interest.

It would also be of interest to combine these variational methods for systems with symmetry with general methods that preserve symmetry structure to take advantage of both approaches: see Iserles, McLachlan and Zanna (1999) and Iserles *et al.* (2000), and literature cited therein.

5.2. Multisymplectic integrators for PDEs

The basic extension of the methods of this paper to the context of PDEs was given in Marsden *et al.* (1998), and Marsden and Shkoller (1999). These papers lay down the variational discretization of PDEs in a multisymplectic context. The examples in these papers were fairly simple, while more interesting examples in the context of continuum mechanics are given in Marsden, Pekarsky, Shkoller and West (2001*a*, 2001*b*).

An important point about the variational methods for multisymplectic PDEs is that they do not require that the PDE and multisymplectic structure be discretized separately (as in Bridges and Reich (200x)). Instead,

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

the discrete multisymplectic structure arises directly from the discrete variational principle, which also immediately gives a discrete multisymplectic form formula: that is, the PDE analogue of the symplectic nature of variational integrators.

This procedure is illustrated in West (2001), where the method of Suris is extended to construct discrete Lagrangians for product Runge–Kutta discretizations of multisymplectic PDEs. From the variational principle we derive the discrete multisymplectic structure, which turns out to coincide with that proposed in Reich (2000).

Lew *et al.* (2001) develop and apply the theory of asynchronous variational integrators (AVIs) in the context of shell dynamics. This is the PDE analogue of integrators that use adaptive time-steps to achieve exact energy and momentum balance as well as being multisymplectic. We expect that similar integrators can be developed for other problems in continuum mechanics, including fluids.

5.3. Open problems

5.3.1. Reversibility

It would be interesting to make closer links with methods that have been developed for reversible systems (see, for example, Cano and Sanz-Serna (1997), Barth, Leimkuhler and Reich (1999), McLachlan, Quispel and Turner (1998*b*) and references therein). We should keep in mind that there are many interesting systems that are Lagrangian but not time-reversible (such as gyroscopic systems, including particles in magnetic fields) as well as systems that are reversible but not Lagrangian. See McLachlan, Quispel and Robidoux (1998*a*) for a discussion of the general classification of these systems as well as those with a first integral, volume-preserving integrators, *etc.*

5.3.2. Variational backward error analysis

Given a symplectic integrator F for a Hamiltonian vector field, traditional backward error analysis finds a modified vector field $\bar{X}_H$ which is exactly integrated by F. It is then shown that this is in fact the Hamiltonian vector field of a modified Hamiltonian $\bar{H}$, so that $\bar{X}_H = X_{\bar{H}}$, which is therefore exactly preserved by F. As H and $\bar{H}$ are close, this shows that F almost preserves H for long times.

We could alternatively do such an analysis on the Lagrangian side. Given a discrete Lagrangian that approximates the action of a given Lagrangian system $L_d \approx \int L \, \mathrm{d}t$, we seek a modified Lagrangian whose action is exactly equal to the discrete Lagrangian, so that $L_d = \int \bar{L} \, \mathrm{d}t$.

To do this, follow the same idea as for traditional backward error analysis and expand $\bar{L} = L + hL^1 + h^2L^2 + \dots$. Substituting this into $L_d = \int \bar{L} \, \mathrm{d}t$,

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

expanding both sides in h, and equating terms allows us to solve for the L^i. To make this rigorous we would need to investigate the convergence of the sum, which will undoubtedly require optimal truncation techniques.

The primary advantage of a Lagrangian backward error analysis is that it may be able to be extended to Lagrangian PDEs, via the techniques of variational multisymplectic integrators, and to situations where the Hamiltonian vector field is not defined, such as nonsmooth contact problems.

5.3.3. Discrete multisymplectic reduction

A natural extension of the discrete reduction theory for ODEs would be its multisymplectic counterpart. Some modest progress has been made in this direction, for instance in Castrillon Lopez, Ratiu and Shkoller (2000), but much remains to be done. The examples of Maxwell's equations (where reduction theory is understood from the infinite-dimensional function space perspective, as in Marsden and Weinstein (1982)) and fluids are challenging, but progress seems likely, both with the theory and, eventually, the associated numerics.

5.3.4. Splitting methods

If $L = \sum L^i$, what can be said about constructing a discrete Lagrangian for L given discrete Lagrangians L_d^i for the L^i? The answer to such a question would be an interesting analogue of the corresponding question for Hamiltonian systems, as in, for example, McLachlan and Scovel (1996) and references therein. What makes the Hamiltonian case in a sense easier is that the equations are linear in the Hamiltonian, whereas this is not the case with the Euler–Lagrange equations (when written in the form $\dot{x} = f(x)$). On the other hand, Hamilton's variational principle is linear in the Lagrangian, so we should be able to exploit that structure.

5.3.5. Evolution of conserved quantities for forced systems

When forcing or dissipation is added to a Lagrangian or Hamiltonian system, then the symplectic form, momentum maps and energy are no longer preserved by the flow. It is often of importance in applications, however, to be able to correctly simulate the amount by which these various quantities change over time.

The variational framework for discrete systems with forcing given in Section 3.2 offers a way in which this evolution can be studied, both for the discrete system and for the true Lagrangian system. Numerical results in Kane *et al.* (2000) for simple systems indicate that, for weakly damped systems, correctly estimating energy and momentum decay requires the use of

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

conservative integrators. It may be possible to make these statements rigorous using the ideas of variational backward error analysis and variational integrators with forcing.

5.3.6. Nonsmooth mechanics

Although we have concentrated on mechanical systems which follow smooth trajectories, there are many physical situations which demand nonsmooth models, such as collision and fragmentation problems. In such cases it can be very profitable to directly use nonsmooth techniques, as described in Kane *et al.* (1999*b*), rather than try to take smooth approximations.

When dealing with truly nonsmooth mechanics, many of the conventional definitions and derivations no longer apply, relying as they do on standard calculus. The variational approach, and in particular discrete variational mechanics, can be extended to nonsmooth situations with little difficulty, however, and provides a useful tool for the analysis of nonsmooth systems and numerical integrators for them (see Kane, Ortiz and Marsden (1998) and Fetecau *et al.* (2001)).

5.3.7. Systems with nonholonomic constraints

A nonholonomic constraint is, loosely speaking, a constraint on both the position and velocity variables. This can be defined by a constraint function $\phi : TQ \to \mathbb{R}$, and we seek trajectories $q(t)$ so that $\phi(q(t), \dot{q}(t)) = 0$ for all t. The variational principle which gives the correct equations of motion in this case is the Lagrange–d'Alembert principle used in Section 3.1 to add forcing (Bloch, Krishnaprasad, Marsden and Murray 1996*a*).

We could extend the discrete variational mechanics to systems with nonholonomic constraints by using the same discrete Lagrange–d'Alembert principle as in Section 3.2. This would yield a constrained discrete Hamiltonian map which would approximate the continuous constrained system, and preserve the nonholonomic constraints.

As nonholonomic systems do not, in general, preserve the standard symplectic structure, we would not expect the discrete system to preserve the discrete symplectic structure. It would be interesting, however, to see whether the evolution was qualitatively correct, as discussed above for the energy decay of dissipative systems. See McLachlan and Perlmutter (200x) for steps in this direction.

5.3.8. Other Galerkin methods

In Section 2.6.6 we considered variational integrators derived by taking polynomial approximations to the trajectory segments, and we saw that this is equivalent to Galerkin projection onto the space of polynomials.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

While polynomials are a good choice for general smooth mechanical systems, there are many problems for which some other choice of curves may be more appropriate. For example, in systems with highly oscillatory components, variational integrators based on sinusoidal trajectory approximations may have superior accuracy or stability properties.

Acknowledgements

We thank many colleagues for their explicit and implicit help in putting this article together, including Razvan Fetecau, Arieh Iserles, Sameer Jalnapurkar, Couro Kane, Melvin Leok, Adrian Lew, Ben Leimkuhler, Michael Ortiz, George Patrick, Sergey Pekarsky, Reinout Quispel, Sebastian Reich, Steve Shkoller, and Robert Skeel. This work was supported by the California Institute of Technology and NSF/KDI grant ATM-9873133, as well as NSF grant DMS-9874082.

REFERENCES

R. Abraham, J. E. Marsden and T. Ratiu (1988), *Manifolds, Tensor Analysis, and Applications*, Vol. 75 of *Applied Mathematical Sciences*, 2nd edn, Springer, New York.

M. P. Allen and D. J. Tildesley (1987), *Computer Simulation of Liquids*, Oxford University Press, New York.

H. Anderson (1983), 'RATTLE: A velocity version of the SHAKE algorithm for molecular dynamics calculations', *J. Comput. Phys.* **52**, 24–34.

U. M. Ascher and L. R. Petzold (1998), *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, SIAM, Philadelphia, PA.

M. A. Austin, P. S. Krishnaprasad and L. S. Wang (1993), Almost Poisson integration of rigid body systems, *J. Comput. Phys.* **107**, 105–117.

J. C. Baez and J. W. Gilliam (1994), An algebraic approach to discrete mechanics, *Lett. Math. Phys.* **31**, 205–212.

E. Barth and B. Leimkuhler (1996), 'Symplectic methods for conservative multibody systems', in *Integration Algorithms and Classical Mechanics* (Toronto, ON, 1993), American Mathematical Society, Providence, RI, pp. 25–43.

E. Barth, B. Leimkuhler and S. Reich (1999), 'A time-reversible variable-stepsize integrator for constrained dynamics', *SIAM J. Sci. Comput.* **21**, 1027–1044 (electronic).

G. Benettin and A. Giorgilli (1994), 'On the Hamiltonian interpolation of near-to-the-identity symplectic mappings with application to symplectic integration algorithms', *J. Statist. Phys.* **74**, 1117–1143.

J. S. Berg, R. L. Warnock, R. D. Ruth and E. Forest (1994), 'Construction of symplectic maps for nonlinear motion of particles in accelerators', *Phys. Rev. E* **49**, 722–739.

A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden and R. M. Murray (1996*a*), 'Nonholonomic mechanical systems with symmetry', *Arch. Rational Mech. Anal.* **136**, 21–99.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden and T. Ratiu (1996*b*), 'The Euler–Poincaré equations and double bracket dissipation', *Comm. Math. Phys.* **175**, 1–42.

A. M. Bloch, P. Crouch, J. E. Marsden and T. S. Ratiu (1998), 'Discrete rigid body dynamics and optimal control', *Proc. CDC* **37**, 2249–2254. Longer manuscript in preparation.

A. I. Bobenko and Y. B. Suris (1999*a*), 'Discrete Lagrangian reduction, discrete Euler–Poincaré equations, and semidirect products', *Lett. Math. Phys.* **49**, 79–93.

A. I. Bobenko and Y. B. Suris (1999*b*), 'Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top', *Comm. Math. Phys.* **204**, 147–188.

V. Brasey and E. Hairer (1993), 'Symmetrized half-explicit methods for constrained mechanical systems', *Appl. Numer. Math.* **13**, 23–31.

T. J. Bridges and S. Reich (200x), Multi-symplectic integrators: Numerical schemes for Hamiltonian PDEs that conserve symplecticity, *Phys. Lett. A*, to appear.

C. J. Budd and A. Iserles (1999), 'Geometric integration: Numerical solution of differential equations on manifolds', *Philos. Trans. Royal Soc. London Ser. A, Math. Phys. Eng. Sci.* **357**, 945–956.

J. A. Cadzow (1970), 'Discrete calculus of variations', *Internat. J. Control.* **11**, 393–407.

J. A. Cadzow (1973), *Discrete-Time Systems: An Introduction with Interdisciplinary Applications*, Prentice-Hall.

B. Cano and J. Lewis (1998), A comparison of symplectic and Hamilton's principle algorithms for autonomous and non-autonomous systems of ordinary differential equations, Technical report, Departamento de Matemática Aplicada y Computación, Universidad de Valladolid.

B. Cano and J. M. Sanz-Serna (1997), 'Error growth in the numerical integration of periodic orbits, with application to Hamiltonian and reversible systems', *SIAM J. Numer. Anal.* **34**, 1391–1417.

M. Castrillón López, T. S. Ratiu and S. Shkoller (2000), Reduction in principal fiber bundles: Covariant Euler–Poincaré equations, *Proc. Amer. Math. Soc.* **128**, 2155–2164.

P. J. Channell and C. Scovel (1990), 'Symplectic integration of Hamiltonian systems', *Nonlinearity* **3**, 231–259.

R. De Vogelaére (1956), Methods of integration which preserve the contact transformation property of the Hamiltonian equations, University of Notre Dame preprint.

D. Estep and D. French (1994), 'Global error control for the continuous Galerkin finite element method for ordinary differential equations', *RAIRO Modelisation Mathématique et Analyse Numérique* **28**, 815–852

R. Fetecau, J. E. Marsden, M. Ortiz and M. West (2001), Nonsmooth Lagrangian mechanics. Preprint.

E. Forest and R. D. Ruth (1990), '4th-order symplectic integration', *Physica D* **43**, 105–117.

Z. Ge and J. E. Marsden (1988), 'Lie–Poisson integrators and Lie–Poisson Hamilton–Jacobi theory', *Phys. Lett. A* **133**, 134–139.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

S. Geng (1995), 'Construction of high-order symplectic PRK methods', *J. Comput. Math.* **13**, 40–50.

S. Geng (2000), 'A simple way of constructing symplectic Runge–Kutta methods', *J. Comput. Math.* **18**, 61–68.

J. W. Gilliam (1996), Lagrangian and symplectic techniques in discrete mechanics, PhD thesis, University of California, Riverside, Department of Mathematics. Available from: `http://math.ucr.edu/home/baez`

R. Gillilan and K. Wilson (1992), 'Shadowing, rare events and rubber bands: A variational Verlet algorithm for molecular dynamics', *J. Chem. Phys.* **97**, 1757–1772.

H. Goldstein (1980), *Classical Mechanics*, 1950, 2nd edn 1980, Addison-Wesley.

O. Gonzalez (1996*a*), Design and analysis of conserving integrators for nonlinear Hamiltonian systems with symmetry, PhD thesis, Stanford University, Department of Mechanical Engineering.

O. Gonzalez (1996*b*), 'Time integration and discrete Hamiltonian systems', *J. Nonlin. Sci.* **6**, 449–467.

O. Gonzalez (1999), 'Mechanical systems subject to holonomic constraints: Differential-algebraic formulations and conservative integration', *Physica D* **132**, 165–174.

O. Gonzalez and J. C. Simo (1996), 'On the stability of symplectic and energy-momentum algorithms for nonlinear Hamiltonian systems with symmetry', *Comput. Meth. Appl. Mech. Eng.* **134**, 197–222.

O. Gonzalez, D. J. Higham and A. M. Stuart (1999), 'Qualitative properties of modified equations', *IMA J. Numer. Anal.* **19**, 169–190.

M. Gotay, J. Isenberg and J. E. Marsden (1997), Momentum maps and classical relativistic fields, Part I: Covariant field theory. Unpublished; available from: `http://www.cds.caltech.edu/ marsden/`

E. Hairer (1994), 'Backward analysis of numerical integrators and symplectic methods', *Ann. Numer. Math.* **1**, 107–132.

E. Hairer (1997), 'Variable time step integration with symplectic methods', *Appl. Numer. Math.* **25**, 219–227.

E. Hairer (1998), 'Numerical geometric integration'. Notes available from: `http://www.unige.ch/math/folks/hairer/polycop.html`.

E. Hairer (200x), Geometric integration of ordinary differential equations on manifolds, *BIT*, to appear.

E. Hairer and C. Lubich (1997), 'The life-span of backward error analysis for numerical integrators', *Numer. Math.* **76**, 441–462.

E. Hairer and C. Lubich (1999), 'Invariant tori of dissipatively perturbed Hamiltonian systems under symplectic discretization', *Appl. Numer. Math.* **29**, 57–71.

E. Hairer and C. Lubich (2000), 'Long-time energy conservation of numerical methods for oscillatory differential equations', *SIAM J. Numer. Anal.* **38**, 414–441.

E. Hairer and G. Wanner (1996), *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, Vol. 14 of *Springer Series in Computational Mathematics*, 2nd edn, Springer.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

E. Hairer, S. P. Nørsett and G. Wanner (1993), *Solving Ordinary Differential Equations I: Nonstiff Problems*, Vol. 8 of *Springer Series in Computational Mathematics*, 2nd edn, Springer.

W. R. Hamilton (1834), On a general method in dynamics, *Philos. Trans. Royal Soc. London* Part II, 247–308; Part I for 1835, pp. 95–144.

D. J. Hardy, D. I. Okunbor and R. Skeel (1999), 'Symplectic variable step size integration for N-body problems', in *Proceedings of the NSF/CBMS Regional Conference on Numerical Analysis of Hamiltonian Differential Equations* (Golden, CO, 1997), *Appl. Numer. Math.* **29**, 19–30.

T. J. R. Hughes (1987), *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Prentice-Hall.

B. Hulme (1972*a*), 'One-step piecewise polynomial Galerkin methods for initial value problems', *Math. Comput.* **26**, 415–426.

B. Hulme (1972*b*), 'Discrete Galerkin and related one-step methods for ordinary differential equations', *Math. Comput.* **26**, 881–891.

C. L. Hwang and L. T. Fan (1967), 'A discrete version of Pontryagin's maximum principle'. *Oper. Res.* **15**, 139–146.

A. Iserles, R. I. McLachlan and A. Zanna (1999), 'Approximately preserving symmetries in the numerical integration of ordinary differential equations', *Europ. J. Appl. Math.* **10**, 419–445.

A. Iserles, H. Munthe-Kaas and A. Zanna (2000), 'Lie-group methods', in *Acta Numerica*, Vol. 9, Cambridge University Press, pp. 215–365.

T. Itoh and K. Abe (1988), 'Hamiltonian-conserving discrete canonical equations based on variational difference equations', *J. Comput. Phys.* **77**, 85–102.

C. G. K. Jacobi (1866), *Vorlesungen über Dynamik*, Verlag G. Reimer.

S. M. Jalnapurkar and J. E. Marsden (200x), 'Discretization of Hamiltonian Systems'. In preparation.

S. M. Jalnapurkar, M. Leok, J. E. Marsden and M. West (2000), Discrete Routh reduction. Preprint.

G. Jaroszkiewicz and K. Norton (1997*a*), 'Principles of discrete time mechanics, I: Particle systems', *J. Phys. A* **30**, 3115–3144.

G. Jaroszkiewicz and K. Norton (1997*b*), 'Principles of discrete time mechanics, II: Classical field theory', *J. Phys. A* **30**, 3145–3163.

L. O. Jay (1996), 'Symplectic partitioned Runge–Kutta methods for constrained Hamiltonian systems', *SIAM J. Numer. Anal.* **33**, 368–387.

L. O. Jay (1999), 'Structure preservation for constrained dynamics with super partitioned additive Runge–Kutta methods', *SIAM J. Sci. Comput.* **20**, 416–446.

B. W. Jordan and E. Polak (1964), 'Theory of a class of discrete optimal control systems'. *J. Electron. Control* **17**, 697–711.

C. Kane, M. Ortiz and J. E. Marsden (1998), The convergence of collision algorithms. Notes.

C. Kane, J. E. Marsden and M. Ortiz (1999*a*), 'Symplectic energy-momentum integrators', *J. Math. Phys.* **40**, 3353–3371.

C. Kane, E. A. Repetto, M. Ortiz and J. E. Marsden (1999*b*), 'Finite element analysis of nonsmooth contact', *Comput. Meth. Appl. Mech. Eng.* **180**, 1–26.

C. Kane, J. E. Marsden, M. Ortiz and M. West (2000), 'Variational integrators and

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

the Newmark algorithm for conservative and dissipative mechanical systems', *Internat. J. Numer. Math. Eng.* **49**, 1295–1325.
R. A. Labudde and D. Greenspan (1974), 'Discrete mechanics: A general treatment', *J. Comput. Phys.* **15**, 134–167.
R. A. Labudde and D. Greenspan (1976*a*), 'Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion, I: Motion of a single particle', *Numer. Math.* **25**, 323–346.
R. A. Labudde and D. Greenspan (1976*b*), 'Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion, II: Motion of a system of particles', *Numer. Math.* **26**, 1–16.
F. Lasagni (1988), 'Canonical Runge–Kutta methods', *ZAMP* **39**, 952–953.
T. D. Lee (1983), 'Can time be a discrete dynamical variable?', *Phys. Lett. B* **122**, 217–220.
T. D. Lee (1987), 'Difference equations and conservation laws', *J. Stat. Phys.* **46**, 843–860
B. Leimkuhler and G. Patrick (1996), Symplectic integration on Riemannian manifolds, *J. Nonlin. Sci.* **6**, 367–384.
B. Leimkuhler and S. Reich (1994), 'Symplectic integration of constrained Hamiltonian systems', *Math. Comput.* **63**, 589–605.
B. J. Leimkuhler and R. Skeel (1994), 'Symplectic numerical integrators in constrained Hamiltonian systems', *J. Comput. Phys.* **112**, 117–125.
A. Lew, J. E. Marsden, M. Ortiz and M. West (2001), 'Asynchronous variational integrators'. Preprint.
J. D. Logan (1973), 'First integrals in the discrete calculus of variations', *Aequationes Mathematicae* **9**, 210–220.
R. MacKay (1992), Some aspects of the dynamics of Hamiltonian systems, in *The Dynamics of Numerics and the Numerics of Dynamics* (D. Broomhead and A. Iserles, eds), Clarendon Press, Oxford, pp. 137–193.
R. I. McLachlan (1995), 'On the numerical integration of ordinary differential equations by symmetric composition methods', *SIAM J. Sci. Comput.* **16**, 151–168.
R. I. McLachlan and M. Perlmutter (200x), Geometric integration of nonholonomic mechanical systems. In preparation.
R. I. McLachlan and C. Scovel (1995), 'Equivariant constrained symplectic integration', *J. Nonlin. Sci.* **5**, 233–256.
R. I. McLachlan and C. Scovel (1996), 'A survey of open problems in symplectic integration', *Fields Inst. Commun.* **10**, 151–180.
R. I. McLachlan, G. R. W. Quispel and N. Robidoux (1998*a*), 'Unified approach to Hamiltonian systems, Poisson systems, gradient systems, and systems with Lyapunov functions or first integrals', *Phys. Rev. Lett.* **81**, 2399–2403.
R. I. McLachlan, G. R. W. Quispel and G. Turner (1998*b*), 'Numerical integrators that preserve symmetries and reversing symmetries', *SIAM J. Numer. Anal.* **35**, 586–599.
R. I. McLachlan, G. R. W. Quispel and N. Robidoux (1999), 'Geometric integration using discrete gradients', *Philos. Trans. Royal Soc. London Ser. A, Math. Phys. Eng. Sci.* **357**, 1021–1045.
S. Maeda (1980), 'Canonical structure and symmetries for discrete systems', *Math. Japonica* **25**, 405–420.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

S. Maeda (1981*a*), 'Extension of discrete Noether theorem', *Math. Japonica* **26**, 85–90

S. Maeda (1981*b*), 'Lagrangian formulation of discrete systems and concept of difference space', *Math. Japonica* **27**, 345–356.

J. E. Marsden (1999), 'Park City lectures on mechanics, dynamics and symmetry', in *Symplectic Geometry and Topology* (Y. Eliashberg and L. Traynor, eds), American Mathematical Society, Providence, RI, Vol. 7 of *IAS/Park City Math. Ser.*, pp. 335–430.

J. E. Marsden and T. Ratiu (1999), *Introduction to Mechanics and Symmetry*, Vol. 17 of *Texts in Applied Mathematics*, 2nd edn, Springer.

J. E. Marsden and S. Shkoller (1999), 'Multisymplectic geometry, covariant Hamiltonians, and water waves', *Math. Proc. Cambridge Philos. Soc.* **125**, 553–575.

J. E. Marsden and A. Weinstein (1982), The Hamiltonian structure of the Maxwell–Vlasov equations, *Physica D* **4**, 394–406.

J. E. Marsden, G. W. Patrick and S. Shkoller (1998), 'Multisymplectic geometry, variational integrators, and nonlinear PDEs', *Comm. Math. Phys.* **199**, 351–395.

J. E. Marsden, S. Pekarsky and S. Shkoller (1999*a*), 'Discrete Euler–Poincaré and Lie–Poisson equations', *Nonlinearity* **12**, 1647–1662.

J. E. Marsden, S. Pekarsky and S. Shkoller (1999*b*), Symmetry reduction of discrete Lagrangian mechanics on Lie groups, *J. Geom. Phys.* **36**, 140–151.

J. E. Marsden, T. Ratiu and J. Scheurle (2000), Reduction theory and the Lagrange–Routh equations, *J. Math. Phys.* **41**, 3379–3429.

J. E. Marsden, S. Pekarsky, S. Shkoller and M. West (2001*a*), 'Variational methods, multisymplectic geometry and continuum mechanics', *J. Geom. Phys.*, to appear.

J. E. Marsden, S. Pekarsky, S. Shkoller and M. West (2001*b*), 'Multisymplectic continuum mechanics in Euclidean spaces and multisymplectic discretizations'. Preprint.

J. Moser and A. P. Veselov (1991), 'Discrete versions of some classical integrable systems and factorization of matrix polynomials', *Comm. Math. Phys.* **139**, 217–243.

A. Murua and J. M. Sanz-Serna (1999), 'Order conditions for numerical integrators obtained by composing simpler integrators', *Philos. Trans. Royal Soc. London Ser. A, Math. Phys. Eng. Sci.* **357**, 1079–1100.

U. Mutze (1998), Predicting classical motion directly from the action principle. Preprint.

N. N. Newmark (1959), 'A method of computation for structural dynamics', *ASCE J. Eng. Mech. Div.* **85**, 67–94.

E. Noether (1918), 'Invariante Variationsprobleme', *Kgl. Ges. Wiss. Nachr. Göttingen. Math. Physik* **2**, 235–257.

K. Norton and G. Jaroszkiewicz (1998), 'Principles of discrete time mechanics, III: Quantum field theory', *J. Phys. A* **31**, 977–1000.

W. Oevel and M. Sofroniou (1997), Symplectic Runge–Kutta schemes, II: Classification of symmetric methods. Preprint.

P. J. Oliver and J. Sivaloganathan (1988), The structure of null Lagrangians, *Nonlinearity* **1**, 389–398.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

M. Ortiz (1986), 'A note on energy conservation and stability of nonlinear time-stepping algorithms', *Comput. Structures* **24**, 167–168.
M. Ortiz and L. Stainier (1999), 'The variational formulation of viscoplastic constitutive updates', *Comput. Meth. Appl. Mech. Eng.* **171**, 419–444.
M. Qin and W. J. Zhu (1992), 'Construction of higher order symplectic schemes by composition', *Computing* **27**, 309–321.
R. Radovitzky and M. Ortiz (1999), 'Error estimation and adaptive meshing in strongly nonlinear dynamic problems', *Comput. Meth. Appl. Mech. Eng.* **172**, 203–240.
S. Reich (1996), 'Symplectic integration of constrained Hamiltonian systems by composition methods', *SIAM J. Numer. Anal.* **33**, 475–491.
S. Reich (1997), 'On higher-order semi-explicit symplectic partitioned Runge–Kutta methods for constrained Hamiltonian systems', *Numer. Math.* **76**, 231–247.
S. Reich (1999a), 'Backward error analysis for numerical integrators', *SIAM J. Numer. Anal.* **36**, 1549–1570.
S. Reich (1999*b*), 'Preservation of adiabatic invariants under symplectic discretization', *Appl. Numer. Math.* **29**, 45–55.
S. Reich (2000), Multi-symplectic Runge–Kutta collocation methods for Hamiltonian wave equations, *J. Comput. Phys.* **157**, 473–499.
R. D. Ruth (1983), 'A canonical integration technique', *IEEE Trans. Nuclear Sci.* **30**, 2669–2671.
J. Ryckaert, G. Ciccotti and H. Berendsen (1977), 'Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes', *J. Comput. Phys.* **23**, 327–341.
J. M. Sanz-Serna (1988), 'Runge–Kutta schemes for Hamiltonian systems', *BIT* **28**, 877–883.
J. M. Sanz-Serna (1992*a*), The numerical integration of Hamiltonian systems, in *Computational Ordinary Differential Equations* (J. Cash and I. Gladwell, eds), Clarendon Press, Oxford, pp. 437–449.
J. M. Sanz-Serna (1992*b*), 'Symplectic integrators for Hamiltonian problems: An overview', in *Acta Numerica*, Vol. 1, Cambridge University Press, pp. 243–286.
J. M. Sanz-serna and M. P. Calvo (1994), *Numerical Hamiltonian Problems*, Chapman and Hall, London.
T. Schlick, R. Skeel *et al.* (1999), 'Algorithmic challenges in computational molecular biophysics', *J. Comput. Phys.* **151**, 9–48.
W. M. Seiler (1998*a*), 'Numerical analysis of constrained Hamiltonian systems and the formal theory of differential equations', *Math. Comput. Simulation* **45**, 561–576.
W. M. Seiler (1998*b*), 'Position versus momentum projections for constrained Hamiltonian systems', *Numer. Algorithms* **19**, 223–234.
W. M. Seiler (1999), 'Numerical integration of constrained Hamiltonian systems using Dirac brackets', *Math. Comput.* **68**, 661–681.
Y. Shibberu (1994), 'Time-discretization of Hamiltonian systems', *Comput. Math. Appl.* **28**, 123–145.
M. Shimada and H. Yoshida (1996), 'Long-term conservation of adiabatic invariants by using symplectic integrators', *Publ. Astronomical Soc. Japan* **48**, 147–155.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press

J. Simo and N. Tarnow (1992), 'The discrete energy momentum method: Conserving algorithms for nonlinear elastodynamics', *ZAMP* **43**, 757–792.

J. C. Simo, N. Tarnow and K. K. Wong (1992), 'Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics', *Comput. Meth. Appl. Mech. Eng.* **100**, 63–116.

R. D. Skeel and K. Srinivas (2000), 'Nonlinear stability analysis of area-preserving integrators', *SIAM J. Numer. Anal.* **38**, 129–148.

R. D. Skeel, G. Zhang and T. Schlick (1997), 'A family of symplectic integrators: Stability, accuracy, and molecular dynamics applications', *SIAM J. Sci. Comput.* **18**, 203–222.

M. Sofroniou and W. Oevel (1997), 'Symplectic Runge–Kutta schemes, I: Order conditions', *SIAM J. Numer. Anal.* **34**, 2063–2086.

Y. B. Suris (1989), 'The canonicity of mappings generated by Runge–Kutta type methods when integrating the system $\ddot{x} = -\partial u/\partial x$', *USSR Comput. Math. Phys.* **29**, 138–144.

Y. B. Suris (1990), 'Hamiltonian methods of Runge–Kutta type and their variational interpretation', *Math. Simulation* **2**, 78–87.

W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson (1982), 'A computer-simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules: Application to small water clusters', *J. Chem. Phys.* **76**, 637–649.

V. Thomée (1997), *Galerkin Finite Element Methods for Parabolic Problems*, Springer, New York.

L. Verlet (1967), 'Computer experiments on classical fluids', *Phys. Rev.* **159**, 98–103.

A. P. Veselov (1988), 'Integrable discrete-time systems and difference operators', *Funct. Anal. Appl.* **22**, 83–93.

A. P. Veselov (1991), 'Integrable Lagrangian correspondences and the factorization of matrix polynomials', *Funct. Anal. Appl.* **25**, 112–122.

R. L. Warnock and R. D. Ruth (1991), 'Stability of nonlinear Hamiltonian motion for a finite but very long-time', *Phys. Rev. Lett.* **66**, 990–993.

R. L. Warnock and R. D. Ruth (1992), 'Long-term bounds on nonlinear Hamiltonian motion', *Physica D* **56**, 188–215.

J. M. Wendlandt and J. E. Marsden (1997*a*), 'Mechanical integrators derived from a discrete variational principle', *Physica D* **106**, 223–246.

J. M. Wendlandt and J. E. Marsden (1997*b*), Mechanical systems with symmetry, variational principles and integration algorithms, in *Current and Future Directions in Applied Mathematics* (M. Alber, B. Hu and J. Rosenthal, eds), Birkhäuser, pp. 219–261.

M. West (2001), Variational Runge-Kutta methods for ODEs and PDEs. Preprint.

J. Wisdom, S. J. Peale and F. Mignard (1984), The chaotic rotation of Hyperion, *Icarus* **58**, 137–152.

H. Yoshida (1990), 'Construction of higher-order symplectic integrators', *Phys. Lett. A* **150**, 262–268.

https://doi.org/10.1017/S096249290100006X Published online by Cambridge University Press