

Mechanical Systems with Symmetry, Variational Principles, and Integration Algorithms

Jerrold E. Marsden*Control and Dynamical Systems 107-81
California Institute of Technology
Pasadena, CA 91125
marsden@cds.caltech.edu
and Jeffrey M. Wendlandt †Mechanical Engineering
University of California at Berkeley
Berkeley, CA 94720
wents@eecs.berkeley.edu

Reprinted from *Current and Future Directions in Applied Mathematics*
M. Alber, B. Hu, and J. Rosenthal, Eds. Birkhäuser, 1997, 219–261.

Abstract

This paper studies variational principles for mechanical systems with symmetry and their applications to integration algorithms. We recall some general features of how to reduce variational principles in the presence of a symmetry group along with general features of integration algorithms for mechanical systems. Then we describe some integration algorithms based directly on variational principles using a discretization technique of Veselov.

The general idea for these variational integrators is to directly discretize Hamilton's principle rather than the equations of motion in a way that preserves the original systems invariants, notably the symplectic form and, via a discrete version of Noether's theorem, the momentum map. The resulting mechanical integrators are second-order accurate, implicit, symplectic-momentum algorithms. We apply these integrators to the rigid body and the double spherical pendulum to show that the techniques are competitive with existing integrators.

*Research partially supported by DOE contract DE-FG03-95ER-25251 and the California Institute of Technology.

†Research partially supported by DOE contract DE-FG03-95ER-25251.
<http://robotics.eecs.berkeley.edu/~wents/>

1 Introduction

This paper begins with a brief survey of some aspects of variational principles for mechanical systems with symmetry as well as integration algorithms for mechanical systems. Our main goal is to present a systematic construction of mechanical integrators for simulating finite dimensional mechanical systems with symmetry based on a discretization of Hamilton's principle. We strive for a method that is theoretically attractive as well as numerically competitive. Our algorithms are second order accurate symplectic-momentum integrators valid for general and constrained systems. We do not claim the methods are superior in specific problems for which custom methods are available.¹ However, for many mechanical systems, it provides a good systematic, general purpose, starting point.

Reduced Variational Principles. Symmetry plays a special role in variational principles. Not only does it lead to conservation laws of Noether, but the reduced variational principle for the Euler-Poincaré equations on a general Lie algebra induced by Hamilton's principle on the corresponding Lie group was only recently found (Marsden and Scheurle [1993b] and Bloch, Krishnaprasad, Marsden and Ratiu [1996]). In fluid mechanics, such variational principles were associated with "Lin constraints", but even here it was only with work such as Seliger and Whitham [1968] and Bretherton [1970] that the situation was clarified. More generally, one can study the role of reduction in the Euler-Lagrange equations and this leads to the *reduced Euler-Lagrange equations* (Marsden and Scheurle [1993a,b]), which have played an important role in nonholonomic mechanics.² These general notions are theoretically closely related to and helped motivate the development of the variational integrators discussed here.

Mechanical Integrators. Numerical integration methods that preserve energy, momentum, or the symplectic form, are called mechanical integrators. A result of Ge and Marsden [1988] states, roughly speaking, that if the energy and momentum map include all the integrals of motion, then one cannot create integrators that are symplectic, energy preserving, *and* momentum preserving unless they coincidentally integrate the equations exactly up to a time parametrization (see §4.1 for the exact statement). Accordingly, the class of mechanical integrators divides into *symplectic-momentum* and *energy-momentum* integrators. By exploiting the structure of mechanical

¹As, for example, in symplectic integrators for the solar system; see, for example, Wisdom and Holman [1991].

²See, for example, Bloch, Krishnaprasad, Marsden and Murray [1996] and Koon and Marsden [1996].

systems, one can hope to create mechanical integrators that are not only theoretically attractive, but are more computationally efficient and have good long term simulation properties. The situation for mechanical integrators is a complex and evolving one; we refer to Marsden, Patrick and Shadwick [1996] for a recent collection of papers in the area.

Variational Integrators. We present a method to construct symplectic-momentum integrators for Lagrangian systems defined on a linear space with holonomic constraints. The constraint manifold, Q , is assumed to be embedded in a linear space V . A discrete version of the Lagrangian is formed and a discrete variational principle is applied to the discrete Lagrangian. The resulting discrete equations define a generally implicit numerical integration algorithm on $Q \times Q$ that approximates the flow of the continuous Euler-Lagrange equations on TQ . The algorithm equations are called the *discrete Euler-Lagrange (DEL)* equations or a *variational integrator (VI)*.

The DEL equations have similarities to the continuous Euler-Lagrange equations. They preserve a symplectic form and a discrete momentum map derived using a discrete Noether theorem associated with a symmetry. The value of the discrete momentum approaches the value of the continuous momentum as the step size decreases. The method need not preserve energy, but the numerical examples suggest that the energy oscillates about a constant value in many cases. The energy variations decrease and the constant value approaches the continuous energy as the step size decreases.

We treat holonomic constraints through constraint functions on the containing linear space. The constraints are satisfied at each time step through the use of Lagrange multipliers.

Dissipation is of course very important for practical simulations of mechanical systems. Our philosophy, consistent with, *e.g.*, Armero and Simo [1996], Chorin, Hughes, Marsden and McCracken [1978], is that of understanding well the ideal model first, and then one can use a time-splitting (product formula) method to interleave it with ones favorite dissipative method.

The Examples. We apply our method using a quaternionic representation of the *rigid body* with the linear space, $V = \mathbb{R}^4$, and the constraint manifold, $Q = S^3 \subset V$ regarded as a double covering of the proper rotation group. The second example is the *double spherical pendulum*. Here, the linear space is, $V = \mathbb{R}^3 \times \mathbb{R}^3$, and the constraint manifold, $Q \subset V$, is $S^2 \times S^2$. This example is motivated by our work on pattern evocation (Marsden and Scheurle [1995] and Marsden, Scheurle and Wendlandt [1996]).

For these examples, the momentum, energy, accuracy, and efficiency is

examined as well as the comparison with an energy-momentum integrator.

Some Literature. If one naively discretizes Hamilton's principle (as is sometimes done) one cannot expect to get an algorithm with good conservation properties. Our approach to the discrete variational principle is based on Veselov [1988], Veselov [1991] and Moser and Veselov [1991]. It is shown in Veselov [1988] that the DEL equations preserve a symplectic form. The same discrete mechanics procedure is derived in an abstract form in Baez and Gilliam [1995] using an algebraic approach, and they also establish a discrete Noether's theorem for infinitesimal symmetry.

Many versions of discrete mechanics have been proposed, sometimes with the motivation of constructing integrators. Maeda [1981] presents a version of discrete mechanics based on the concept of a difference space. The author shows how to derive the discrete equations from a discrete version of Hamilton's variational principle, the same discretization later used in Veselov [1988]. Maeda [1981] also presents a version of Noether's theorem. A different approach to discrete mechanics for point mass systems, but not derived from a variational principle is given in Labudde and Greenspan [1974, 1976a,b]; the corresponding algorithms preserve energy and momentum. A discussion of discretizing variational principles is given in MacKay [1992] and also in Lewis and Kostelec [1996]. It is our opinion that the approach in Veselov [1988] we adopt in this paper is the theoretically most appealing method and, in addition, is numerically competitive.

Some authors discretize the principle of least action instead of Hamilton's principle. Algorithms that conserve the Hamiltonian are derived in Itoh and Abe [1988] based on difference quotients. Differentiation is not used and the action is extremized using variational difference quotients. This development presents multistep methods with variable time steps. The least action principle is discretized in a different way in Shibberu [1994]. The resulting equations explicitly enforce energy, and the equations preserve quadratic invariants.

Various energy-momentum integrators have been developed by Simo and his co-workers; for example, Simo and Tarnow [1992] and related references cited in the bibliography. Energy-momentum integrators were derived based on discrete directional derivatives and discrete versions of Hamiltonian mechanics in Gonzalez [1996a]. Additional references on energy-momentum methods are given in Gonzalez [1996a,b]. Symplectic, momentum and energy conserving schemes for the rigid body are presented in Lewis and Simo [1995].

The literature on symplectic schemes for Hamiltonian systems is vast.

The overviews of symplectic integrators in Channell and Scovel [1990], Sanz-Serna [1991] and McLachlan and Scovel [1996] provide background and references. References related to the work in this paper are Reich [1993], Reich [1994], McLachlan and Scovel [1995], and Jay [1996]. Reich [1993] gives an integration method for Hamiltonian systems that enforces position and velocity constraints in a way making the method symplectic. It is shown in McLachlan and Scovel [1995] and in Reich [1994] that the algorithm also conserves momentum corresponding to a linear symmetry group when the constraint manifold is embedded in a linear space. For another treatment of algorithms formed by embedding the constraint manifold in a linear space, see Barth and Leimkuhler [1996a]. Leimkuhler and Patrick [1996] develop an intrinsic treatment of symplectic-momentum integrators on Riemannian manifolds using generating functions.

The algorithm presented in the present paper embeds the constraint manifold in a linear space but only enforces position constraints. We feel that the enforcement of velocity constraints in the context of our method is best done in the context of nonholonomic mechanics as developed by Bloch, Krishnaprasad, Marsden and Murray [1996].

The Verlet [1967] algorithm is common in molecular dynamics simulations; see, for example, Leimkuhler and Skeel [1994]. An extension to handle holonomic constraints is the SHAKE algorithm (Ryckaert, Ciccotti and Berendsen [1977]). SHAKE was extended to handle velocity constraints with RATTLE in Anderson [1983]. For a presentation of the symplectic nature of the *Verlet*, SHAKE, and RATTLE algorithms, see Leimkuhler and Skeel [1994]. The construction developed in the present paper, when applied to a Lagrangian with a constant mass matrix and a potential energy term, produces a method similar to the SHAKE algorithm written in terms of position coordinates, but the potential force terms differ. If one applies our construction using the discrete Lagrangian definition in Equation (51), then one reproduces the SHAKE algorithm. One recovers the *Verlet* algorithm if the Lagrangian system has no constraints. This result, due to Gillilan and Wilson [1992], is based on a discrete variational principle similar to that of Veselov [1988]. Gillilan and Wilson [1992] emphasize calculating a path given end point conditions, whereas our approach emphasizes the dynamics. Our procedure can also handle more general Lagrangians, such as the Lagrangian for the rigid body in terms of quaternions.

Accuracy and Energy as a Monitor. Our construction method produces 2-step methods that have a second order local truncation error. The position error in the numerical examples show second order convergence.

One should be able to use the methods in Yoshida [1990] to increase the order of accuracy. In the simulations, we use energy as a monitor to catch any obvious problems, as in Channell and Scovel [1990] and Simo and Gonzalez [1993]. It is still unknown if this is a generally reliable indicator of accuracy, but based on the Ge-Marsden result mentioned before, it may well be. Another indication is the analysis with energy oscillation and nearby Hamiltonian systems in Sanz-Serna [1991, p.277–278], Sanz-Serna and Calvo [1994, p. 139–140] and Sanz-Serna [1996]. We must caution, however, that energy conservation alone does not imply good performance as is shown in Ortiz [1986]. In our examples, we observe energy oscillations around a constant value, which, because the symplectic form and other integrals are also exactly conserved, we take as a good indication.

When comparing energy-momentum and symplectic-momentum methods, it should be kept in mind that energy-momentum methods should be monitored using how well they conserve the symplectic form. This is of course not so straightforward as monitoring using the energy, since the symplectic condition involves computing the derivative of the flow map (*e.g.*, using a cloud of initial conditions). While the present paper does not directly address these questions, it is important to keep them in mind.

2 Variational Principles and Symmetry

Hamilton’s principle states that one obtains the Euler-Lagrange equations from extremizing the integral of the Lagrangian subject to fixed endpoint conditions: $\delta \int_a^b L(q, \dot{q}) dt = 0$. In this principle one takes variations of a given trajectory $q(t)$ in a configuration manifold Q subject to fixed temporal endpoint conditions. See the standard texts³ for a discussion. We take it for granted that the reader is familiar with Hamilton’s principle and understands the Legendre transform and how it is used to pass to the Hamiltonian side and its symplectic formulation as well as the notions of momentum map and symmetry reduction.

Noether’s Theorem. It is of course well known how to obtain the conservation laws of Noether directly from Hamilton’s principle but it will be useful for later purposes to review this. Consider a Lie group G acting on a configuration manifold Q and lift this action to the tangent bundle TQ using the tangent operation. Given a G -invariant Lagrangian $L : TQ \rightarrow \mathbb{R}$,

³Such as Abraham and Marsden [1978], Arnold [1989] and Marsden and Ratiu [1994].

the corresponding momentum map is the mapping $\mathbf{J} : TQ \rightarrow \mathfrak{g}^*$ defined by

$$\langle \mathbf{J}(v_q), \xi \rangle = \langle \mathbb{F}L(v_q), \xi_Q(q) \rangle \quad (1)$$

where $\mathbb{F}L : TQ \rightarrow T^*Q$ is the fiber derivative, and ξ_Q denotes the infinitesimal generator associated with a Lie algebra element $\xi \in \mathfrak{g}$. In coordinates, this reads

$$J_a = \frac{\partial L}{\partial \dot{q}^i} K_a^i, \quad (2)$$

where we define the action coefficients K_a^i relative to a basis e_a of \mathfrak{g} , $a = 1, \dots, k$, by writing $\xi_Q(q) = K_a^i \xi^a \partial / \partial q^i$ with $\xi = \xi^a e_a$, and a sum on the index a is understood.

Theorem 2.1 (Classical Noether Theorem) *For a solution of the Euler-Lagrange equations, the quantity \mathbf{J} is a constant in time.*

We remark in passing that this result holds even if the Lagrangian is degenerate, that is, the fiber derivative defined by $p_i = \partial L / \partial \dot{q}^i$ is not invertible.

Noether's theorem is proven directly from Hamilton's principle by choosing a function $\phi(t, s)$ of two variables such that the conditions $\phi(a, s) = \phi(b, s) = \phi(t, 0) = 0$ hold, where a and b are the temporal endpoints of the given solution to the Euler-Lagrange equations. We consider the variation $q(t, s) = \exp(\phi(t, s)\xi) \cdot q(t)$ in Hamilton's principle. Subtracting the result from the corresponding statement of infinitesimal invariance gives the result.⁴

The Rigid Body and Reduced Variational Principles. A more subtle role is to understand how to reduce variational principles and how one can form symmetric discretizations of the original system based on the variational principle. To understand these issues it should be helpful to first outline some features for the special case of the rigid body. This is an example we will be returning to for a numerical example later on (but from a quaternionic, or Cayley-Klein point of view).

We regard an element $R \in \text{SO}(3)$ giving the configuration of the body as a map of a *reference configuration* $\mathcal{B} \subset \mathbb{R}^3$ to the current configuration $R(\mathcal{B})$ taking a reference or label point $X \in \mathcal{B}$ to a current point $x = R(X) \in R(\mathcal{B})$. For a rigid body in motion, the matrix R is time dependent and

⁴See Bloch, Krishnaprasad, Marsden and Murray [1996] for the details of this classical proof in modern language.

the velocity of a point of the body is $\dot{x} = \dot{R}X = \dot{R}R^{-1}x$. Since R is an orthogonal matrix, $R^{-1}\dot{R}$ and $\dot{R}R^{-1}$ are skew matrices, and so we can write

$$\dot{x} = \dot{R}R^{-1}x = \omega \times x, \quad (3)$$

which defines the *spatial angular velocity* vector ω . The corresponding *body angular velocity* is defined by

$$\Omega = R^{-1}\omega, \quad i.e., \quad R^{-1}\dot{R}v = \Omega \times v \quad (4)$$

so that Ω is the angular velocity relative to a body fixed frame. The kinetic energy is

$$K = \frac{1}{2} \int_{\mathcal{B}} \rho(X) \|\dot{R}X\|^2 d^3X, \quad (5)$$

where ρ is a given mass density. Since

$$\|\dot{R}X\| = \|\omega \times x\| = \|R^{-1}(\omega \times x)\| = \|\Omega \times X\|,$$

K is a quadratic function of Ω . Writing $K = \Omega^T \mathbb{I} \Omega / 2$ defines the *moment of inertia tensor* \mathbb{I} , which, if the body does not degenerate to a line, is a positive definite 3×3 matrix thought of as a quadratic form. This quadratic form, can be diagonalized, and this defines the *principal axes and moments of inertia*. In this basis, we write $\mathbb{I} = \text{diag}(I_1, I_2, I_3)$.

The well known relation between the motion in R space and in Ω space is as follows:

Theorem 2.2 *The curve $R(t) \in \text{SO}(3)$ satisfies Hamilton's principle, i.e., the Euler-Lagrange equations for*

$$L(R, \dot{R}) = \frac{1}{2} \int_{\mathcal{B}} \rho(X) \|\dot{R}X\|^2 d^3X \quad (6)$$

if and only if $\Omega(t)$ defined by $R^{-1}\dot{R}v = \Omega \times v$ for all $v \in \mathbb{R}^3$ satisfies Euler's equations: $\mathbb{I}\dot{\Omega} = \mathbb{I}\Omega \times \Omega$.

To understand how to use variational principles to prove this (of course there are many other ways as well), recall that by Hamilton's principle, $R(t)$ satisfies the Euler-Lagrange equations if and only if $\delta \int L dt = 0$, where variations are taken within the group $\text{SO}(3)$ with fixed endpoints. Let the reduced Lagrangian be defined by $l(\Omega) = (\mathbb{I}\Omega) \cdot \Omega / 2$ so that $l(\Omega) = L(R, \dot{R})$ if R and Ω are related by (4). To see how we should transform the variational

principle of L , we differentiate the relation $R^{-1}\dot{R}v = \Omega \times v$ with respect to R to get

$$-R^{-1}\delta R R^{-1}\dot{R}v + R^{-1}\delta\dot{R}v = \delta\Omega \times v. \quad (7)$$

Let the skew matrix $\hat{\Sigma}$ be defined by $\hat{\Sigma} = R^{-1}\delta R$ and define the vector Σ by $\hat{\Sigma}v = \Sigma \times v$. Note that

$$\dot{\hat{\Sigma}} = -R^{-1}\dot{R}R^{-1}\delta R + R^{-1}\delta\dot{R} \quad \text{or} \quad R^{-1}\delta\dot{R} = \dot{\hat{\Sigma}} + R^{-1}\dot{R}\hat{\Sigma}.$$

Substitution gives

$$-\hat{\Sigma}\hat{\Omega}v + \dot{\hat{\Sigma}}v + \hat{\Omega}\hat{\Sigma}v = \widehat{\delta\Omega}v \quad \text{or} \quad \widehat{\delta\Omega} = \dot{\hat{\Sigma}} + [\hat{\Omega}, \hat{\Sigma}].$$

The identity $[\hat{\Omega}, \hat{\Sigma}] = (\Omega \times \Sigma)^\wedge$ holds by Jacobi's identity for the cross product, and so

$$\delta\Omega = \dot{\Sigma} + \Omega \times \Sigma. \quad (8)$$

These calculations prove the following

Theorem 2.3 *Hamilton's principle $\delta \int_a^b L dt = 0$ on $\text{SO}(3)$ is equivalent to the reduced variational principle $\delta \int_a^b l dt = 0$ on \mathbb{R}^3 where the variations $\delta\Omega$ are of the form (8) with $\Sigma(a) = \Sigma(b) = 0$.*

To complete the proof of Theorem 2.2, it suffices to work out the equations equivalent to the reduced variational principle. This is easily done as in the calculus of variations and one indeed gets the Euler equations.

The *body angular momentum* is defined in the usual way, by $\Pi = \mathbb{I}\Omega$ so that in a principal axis frame,

$$\Pi = (\Pi_1, \Pi_2, \Pi_3) = (I_1\Omega_1, I_2\Omega_2, I_3\Omega_3).$$

Assuming that no external moments act on the body, the spatial angular momentum vector $\pi = R\Pi$ is conserved in time. This follows by general considerations of symmetry, but it can, of course, be checked directly from Euler's equations by computing $d\pi/dt$.

The Euler-Poincaré Equations and Variational Principles. There is a generalization of Theorem 2.2 to general Lie groups using the Euler-Lagrange equations and the variational principle as a starting point. (For a discussion with the links with the Lie-Poisson equations, see for example, Marsden and Ratiu [1994]; also see this reference and Bloch, Krishnaprasad, Marsden and Ratiu [1996] for the proof.)

Theorem 2.4 *Let G be a Lie group and $L : TG \rightarrow \mathbb{R}$ a left invariant Lagrangian. Let $l : \mathfrak{g} \rightarrow \mathbb{R}$ be its restriction to the identity. For a curve $g(t) \in G$, let $\xi(t) = g(t)^{-1} \cdot \dot{g}(t)$; i.e., $\xi(t) = T_{g(t)}L_{g(t)^{-1}}\dot{g}(t)$. Then the following are equivalent*

- i** $g(t)$ satisfies the Euler-Lagrange equations for L on G
- ii** Hamilton's principle holds, for variations with fixed endpoints
- iii** the **Euler-Poincaré equations** hold:

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_{\xi}^* \frac{\delta l}{\delta \xi} \quad (9)$$

iv the variational principle $\delta \int l(\xi(t))dt = 0$ holds on \mathfrak{g} , using variations of the form $\delta \xi = \dot{\eta} + [\xi, \eta]$ where η vanishes at the endpoints.

In coordinates on the Lie algebra, the Euler-Poincaré equations read as follows

$$\frac{d}{dt} \frac{\partial l}{\partial \xi^d} = C_{ad}^b \frac{\partial l}{\partial \xi^b} \xi^a, \quad (10)$$

where C_{ad}^b are the structure constants of the Lie algebra.

3 The Reduced Euler-Lagrange Equations

The discussion in the preceding section was generalized to arbitrary configuration spaces and symmetry groups in Marsden and Scheurle [1993b]. As we mentioned in the introduction, this theory has played an important role in nonholonomic systems and in questions of optimal control (see Bloch, Krishnaprasad, Marsden and Murray [1996] and Koon and Marsden [1996]).

We start with a configuration manifold Q and a Lagrangian $L : TQ \rightarrow \mathbb{R}$. Let G be a Lie group and let \mathfrak{g} be its Lie algebra. Assume that G acts on Q and lift this action to TQ by the tangent operation. Assuming that L is G invariant, there is induced a **reduced Lagrangian** $l : TQ/G \rightarrow \mathbb{R}$. We can regard TQ/G as a \mathfrak{g} bundle over TS , where $S = Q/G$. We assume that G acts freely and properly on Q , so we can regard $Q \rightarrow Q/G$ as a principal G -bundle.⁵

An important ingredient is the introduction of a connection A on the principal bundle $Q \rightarrow S = Q/G$. The particular case of the mechanical connection (see Marsden [1992] for a discussion) is often made. A connection allows one to split the variables into a horizontal and vertical part.

⁵Additional work is needed to relax this assumption, as the singular case is very important in examples. However, we shall not discuss this here.

The Hamel Equations. Next, we introduce some notation so that we can write the reduced Euler-Lagrange equations in coordinates. Let x^α , also called “internal variables”, be coordinates for shape space Q/G , ξ^a be coordinates for the Lie algebra \mathfrak{g} relative to a chosen basis, so that as before, $\xi = g^{-1}\dot{g}$, l be the reduced Lagrangian regarded as a function of the variables $x^\alpha, \dot{x}^\alpha, \xi^a$, and, as before, C_{db}^a be the structure constants of the Lie algebra \mathfrak{g} of G .

If one writes the Euler-Lagrange equations on TQ in a local principal bundle trivialization, with coordinates x^α on the base and ξ^a in the fiber, then one gets the *Hamel equations* which are the Euler-Lagrange equations for the variables x^α and the Euler-Poincaré equations for the variables ξ^a . The Hamel equations do not make global intrinsic sense as a pair of equations, unless $Q \rightarrow S$ admits a global flat connection. With a connection, one can intrinsically and globally split the original variational principle relative to horizontal and vertical variations. In addition, there are also good mechanical reasons, related to the study of stability (see Simo, Lewis and Marsden [1991]) for introducing connections.

The Reduced Euler-Lagrange Equations. One gets from the Hamel equations to the equations written in terms of a connection by means of the *velocity shift* given by replacing ξ by the vertical part of (\dot{x}, g, \dot{g}) relative to the connection:

$$\Omega^a = A_\alpha^a \dot{x}^\alpha + \xi^a \quad \text{i.e.,} \quad \Omega = A\dot{x} + g^{-1}\dot{g}.$$

Here, A_α^d are the local coordinates of the connection A . This change of coordinates is also motivated from the mechanical point of view since the variables Ω have the interpretation of the *locked angular velocity*; they are related to the momentum by means of the locked inertia tensor. One can also read the preceding equation in reverse as a reconstruction equation to reconstruct $g(t)$ in terms of \dot{x} and Ω .

Carrying out the above velocity shift gives the following *reduced Euler-Lagrange equations*:

$$\frac{d}{dt} \frac{\partial l}{\partial \dot{x}^\alpha} - \frac{\partial l}{\partial x^\alpha} = \frac{\partial l}{\partial \Omega^a} \left(-B_{\alpha\beta}^a \dot{x}^\beta + \mathcal{E}_{\alpha d}^a \Omega^d \right) \quad (11)$$

$$\frac{d}{dt} \frac{\partial l}{\partial \Omega^b} = \frac{\partial l}{\partial \Omega^a} \left(-\mathcal{E}_{ab}^a \dot{x}^\alpha + C_{db}^a \Omega^d \right). \quad (12)$$

In these equations, $B_{\alpha\beta}^a$ are the coordinates of the curvature B of A , (here we follow the curvature conventions of Bloch, Krishnaprasad, Marsden and

Murray [1996] who generalized these equations to the nonholonomic case) and $\mathcal{E}_{\alpha d}^a = C_{bd}^a A_\alpha^b$.

The variables Ω^a may be regarded as the rigid part of the variables on the original configuration space, while x^α are the internal variables. As in Simo, Lewis, and Marsden [1991], the division of variables into internal and rigid parts has deep implications for both stability theory and for bifurcation theory, continuing lines developed originally by Riemann, Poincaré and others.

One of the key results in Hamiltonian reduction theory says that the reduction of a cotangent bundle T^*Q by a symmetry group G is a bundle over T^*S , where $S = Q/G$ is shape space, and where the fiber is either \mathfrak{g}^* , the dual of the Lie algebra of G , or is a coadjoint orbit, depending on whether one is doing Poisson or symplectic reduction. (See Montgomery, Marsden, and Ratiu [1984] and Marsden [1992] for details and references). The above reduced Euler-Lagrange equations gives the analogue of this structure on the tangent bundle. These two sets of equations are coupled through the curvature of a connection on the bundle and the fact that the Lagrangian is, in general, a function of all the variables. Normally one chooses the connection to be the mechanical connection, although any choice is allowed.

The Splitting of Hamilton's Principle. The relation with reduced variational principles is as follows, which is consistent with the philosophy put forward by Lagrange-d'Alembert for nonholonomic systems.

Theorem 3.1 *The reduced Euler-Lagrange equations (11) are equivalent to Hamilton's principle for the original Lagrangian, where the variations are restricted to be horizontal relative to the given connection. Likewise, the equations (12) together with the reconstruction equations are equivalent to Hamilton's principle where the variations are constrained to be vertical.*

In other words, breaking apart Hamilton's principle into variations that are horizontal and those that are vertical leads to the important structure of the reduced Euler-Lagrange equations.

The above is the analogue of what one in Hamiltonian reduction would call Poisson reduction. In the symplectic context one normally constrains the momentum map \mathbf{J} to a specific value μ . There is a Lagrangian analogue of this too, which was known to Routh for the case of abelian groups (cyclic variables). A key ingredient to the construction of a reduced Lagrangian system in this case is the modification of the Lagrangian L to the Routhian R^μ , which is obtained from the Lagrangian by subtracting off the connection

paired with the constraining value μ of the momentum map. We refer to Marsden and Scheurle [1993a] for details.

4 Generalities on Integration Algorithms

We turn our attention now to some integration algorithms for Lagrangian systems with symmetry, giving in this section a little background material for what was discussed in the introduction and for what follows in the next sections.

Mechanical Integrators. By an *algorithm* on a phase space P we mean a collection of maps $F_\tau : P \rightarrow P$ (depending smoothly, say, on $\tau \in \mathbb{R}$ for small τ and $z \in P$). Sometimes we write $z^{k+1} = F_\tau(z^k)$ for the algorithm and we write Δt or h for the *step size* τ . We say that the algorithm is *consistent* or is *first order accurate* with a vector field X on P if

$$\left. \frac{d}{d\tau} F_\tau(z) \right|_{\tau=0} = X(z). \quad (13)$$

Higher order accuracy is defined similarly by matching higher order derivatives. One of the basic things one is interested in is *convergence* namely, when is

$$\lim_{n \rightarrow \infty} (F_{t/n})^n(z) = \varphi_t(z) \quad (14)$$

where φ_t is the flow of X , and what are the error estimates? There are some general theorems guaranteeing this, with an important hypothesis being *stability*; *i.e.*, $(F_{t/n})^n(z)$ must remain close to z for small t and all $n = 1, 2, \dots$. We refer to Chorin, Hughes, Marsden and McCracken [1978] and Abraham, Marsden and Ratiu [1988] for details. For example, the *Lie-Trotter formula*

$$e^{t(A+B)} = \lim_{n \rightarrow \infty} (e^{tA/n} e^{tB/n})^n \quad (15)$$

for the time-splitting of linear problems and their nonlinear generalizations are instances in which one has a great variety of convergence theorems.

For a Hamiltonian system on a symplectic manifold which has symmetry, an algorithm F_τ is said to be

a *symplectic-integrator* if each F_τ is symplectic,

an *energy-integrator* if $H \circ F_\tau = H$ (where $X = X_H$),

a *momentum-integrator* if $\mathbf{J} \circ F_\tau = \mathbf{J}$.

If F_τ has one or more of these properties, we call it a *mechanical integrator*. Notice that if an integrator has one of these three properties, then so does any iterate of it.

In this paper we shall be interested in the counterpart to these ideas on the Lagrangian side.

There are many different ways that have been employed to find mechanical integrators, as has been mentioned in the introduction. For example, one can search amongst existing algorithms and find ones with special algebraic properties that make them symplectic or energy-preserving. Second, one can attempt to design mechanical integrators from scratch.

Example 1. A first order explicit symplectic scheme in the plane is given by the map $(q_0, p_0) \mapsto (q, p)$ defined by

$$\begin{aligned} q &= q_0 + (\Delta t)p_0 \\ p &= p_0 - (\Delta t)V'(q_0 + (\Delta t)p_0). \end{aligned} \tag{16}$$

This map is a first order approximation to the flow of Hamilton's equations for the Hamiltonian $H = (p^2/2) + V(q)$. Here, one can verify by direct calculation that this scheme is a symplectic map.

Example 1 can, if one wishes, be based on the use of generating functions, as we shall see below. A modification of Example 1 using Poincaré's generating function, but also one that can be checked directly is:

Example 2. An implicit symplectic scheme in the plane for the same Hamiltonian as in Example 1 is

$$\begin{aligned} q &= q_0 + (\Delta t)(p + p_0)/2 \\ p &= p_0 - (\Delta t)V'((q + q_0)/2). \end{aligned} \tag{17}$$

Other examples are sometimes based on special observations. The next example shows that the second order accurate mid-point rule is symplectic (Feng [1987]). This algorithm is also useful in developing almost Poisson integrators (Austin and Krishnaprasad and Wang [1993]).

Example 3. In a symplectic vector space the *mid point rule* is symplectic:

$$\frac{z^{k+1} - z^k}{\Delta t} = X_H \left(\frac{z^k + z^{k+1}}{2} \right). \tag{18}$$

Notice that for small Δt the map defined implicitly by this equation is well defined by the implicit function theorem. One may show it is symplectic, using the fact that the *Cayley transform* $S = (\mathbf{1} - \lambda A)^{-1} (\mathbf{1} + \lambda A)$ of an infinitesimally symplectic linear map A is symplectic if $\mathbf{1} - \lambda A$ is invertible for some real λ .

Example 4. Here is an example of an implicit energy preserving algorithm from Chorin, Hughes, Marsden and McCracken [1978]. Consider a Hamiltonian system for $\mathbf{q} \in \mathbb{R}^n$ and $\mathbf{p} \in \mathbb{R}^n$:

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}. \quad (19)$$

Define the following implicit scheme

$$\mathbf{q}_{n+1} = \mathbf{q}_n + \Delta t \frac{H(\mathbf{q}_{n+1}, \mathbf{p}_{n+1}) - H(\mathbf{q}_{n+1}, \mathbf{p}_n)}{\lambda^T (\mathbf{p}_{n+1} - \mathbf{p}_n)} \lambda, \quad (20)$$

$$\mathbf{p}_{n+1} = \mathbf{p}_n - \Delta t \frac{H(\mathbf{q}_{n+1}, \mathbf{p}_n) - H(\mathbf{q}_n, \mathbf{p}_n)}{\mu^T (\mathbf{q}_{n+1} - \mathbf{q}_n)} \mu, \quad (21)$$

where

$$\lambda = \frac{\partial H}{\partial \mathbf{p}} (\alpha \mathbf{q}_{n+1} + (1 - \alpha) \mathbf{q}_n, \beta \mathbf{p}_{n+1} + (1 - \beta) \mathbf{p}_n), \quad (22)$$

$$\mu = \frac{\partial H}{\partial \mathbf{q}} (\gamma \mathbf{q}_{n+1} + (1 - \gamma) \mathbf{q}_n, \delta \mathbf{p}_{n+1} + (1 - \delta) \mathbf{p}_n), \quad (23)$$

and where $\alpha, \beta, \gamma, \delta$ are arbitrarily chosen constants in $[0, 1]$.

To see that conservation of energy holds, note that from (20), we have

$$(\mathbf{q}_{n+1} - \mathbf{q}_n)^T (\mathbf{p}_{n+1} - \mathbf{p}_n) = \Delta t (H(\mathbf{q}_{n+1}, \mathbf{p}_{n+1}) - H(\mathbf{q}_{n+1}, \mathbf{p}_n)), \quad (24)$$

and from (21)

$$(\mathbf{p}_{n+1} - \mathbf{p}_n)^T (\mathbf{q}_{n+1} - \mathbf{q}_n) = -\Delta t (H(\mathbf{q}_{n+1}, \mathbf{p}_n) - H(\mathbf{q}_n, \mathbf{p}_n)). \quad (25)$$

Subtracting (25) from (24), we obtain $H(\mathbf{q}_{n+1}, \mathbf{p}_{n+1}) = H(\mathbf{q}_n, \mathbf{p}_n)$. This algorithm is checked to be consistent. In general, however, it is not symplectic — this is in accord with Theorem 4.1 below, that one does not normally expect integrators for systems that are not integrable to be both energy preserving and symplectic.

Example 5. Let us apply the Lie-Trotter or time splitting idea to the simple pendulum. The equations are

$$\frac{d}{dt} \begin{pmatrix} \varphi \\ p \end{pmatrix} = \begin{pmatrix} p \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -\sin \varphi \end{pmatrix}.$$

Each vector field can be integrated explicitly to give maps $G_\tau(\varphi, p) = (\varphi + \tau p, p)$ and $H_\tau(\varphi, p) = (\varphi, p - \tau \sin \varphi)$, each of which is symplectic. Thus, the composition $F_\tau = G_\tau \circ H_\tau$, namely,

$$F_\tau(q, p) = (\varphi + \tau p - \tau^2 \sin \varphi, p - \tau \sin \varphi)$$

is a first order symplectic scheme for the simple pendulum. It is closely related to the *standard map*. The orbits of F_τ need not preserve energy and they may be chaotic, whereas the trajectories of the simple pendulum are of course not chaotic.

We refer to the cited references for more examples of this type, including symplectic Runge-Kutta schemes.

As we have indicated, a number of algorithms have been developed specifically for integrating Hamiltonian systems to conserve the energy integral, but without attempting to capture all of the details of the Hamiltonian structure. In fact, some of the standard energy-conservative algorithms have poor momentum behavior over even moderate time ranges, which makes them unsuitable for problems in satellite dynamics for example, where the exact conservation of a momentum integral is essential.

One can get angular momentum drift in energy-conservative simulations of, for example, rods that are free to vibrate and rotate. To control such drifts and attain the high levels of computational accuracy demanded by automated control mechanisms, one would be forced to reduce computational step sizes to such an extent that the numerical simulation would be prohibitively inefficient. Similarly, if one attempts to use a standard energy-conservative algorithm to simulate both the rotational and vibrational modes of a freely moving flexible rod, the algorithm may predict that the rotational motion will come to a virtual halt after only a few cycles! For a documented simulation of a problem with momentum conservation, see Simo and Wong [1989]. One can readily imagine that in the process of enforcing energy conservation one could upset conservation of angular momentum.

All the implicit members of the Newmark family, perhaps the most widely used time-stepping algorithms in nonlinear structural dynamics, *are not designed to conserve energy and also fail to conserve momentum*. Among

the explicit members, only the central difference method preserves momentum (see Simo, Tarnow and Wong [1991]).

A given algorithm is called *group equivariant* provided each F_h commutes with the action of the given symmetry group. Since momentum integrals in Hamiltonian systems are associated with invariance of the Hamiltonian of the given system under the symplectic action of symmetry groups, one might guess that to derive momentum-conservative algorithms, one should look for algorithms to be group equivariant. There are in fact some general theorems to this effect, and as we shall see in our context of discrete versions of Hamilton's principle, this is exactly what happens.

Restrictions on Mechanical Integrators. Given the importance of conserving integrals of motion and the important role played by the Hamiltonian structure in the reduction procedure for a system with symmetry, one might hope to find an algorithm that combines all of the desirable properties: conservation of energy, conservation of momenta (and other independent integrals), and conservation of the symplectic structure. However, one cannot do all three of these things at once unless one relaxes one or more of the conditions in a precise sense given by the next result.

Theorem 4.1 (Ge and Marsden, 1988) *Suppose that a given algorithm is energy preserving, symplectic and is momentum preserving. Consistent with this, assume also that the algorithm is group equivariant. If the dynamics is nonintegrable on the reduced space (in the sense spelled out in the proof) then the algorithm already gives the exact solution of the original dynamics problem up to a possible time reparametrization and a phase shift given by the group action.*

Proof. Suppose $F_{\Delta t}$ is the symplectic algorithm, and that it satisfies the given hypotheses. Since it is group equivariant, it induces an algorithm on the symplectic reduced phase space. We assume that the Hamiltonian H is the only integral of motion of the reduced dynamics (*i.e.*, all other integrals of the system have been found and taken out in the reduction process in the sense that any other conserved quantity (in a chosen smoothness class) is functionally dependent on H). We now work with the reduced algorithm without changing the notation.

Since $F_{\Delta t}$ is symplectic, it follows from general facts about Hamiltonian systems that it is the Δt -time map of some *time-dependent* Hamiltonian function K . Now assume that the symplectic map $F_{\Delta t}$ also conserves H for *all* (small) values of Δt . It follows that $\{H, K\} = 0 = \{K, H\}$. The

latter equation implies that K is functionally dependent on H (for each fixed t) since the flow of H (the “true dynamics”) has no other integrals of motion. The functional dependence of K on H in turn implies that their Hamiltonian vector fields are parallel (this is an easy calculation using the chain rule), so the flow of K (the approximate solution) and the flow of H (the exact solution) must lie along identical curves in the reduced phases space; in other words, the flows are equivalent up to time reparametrization. Since this holds on the reduced space and the original algorithm preserves the momentum map and is equivariant, we get the stated result. (The group phase here is that of G_μ , the same as one has in the theory of geometric phases). *QED*

Thus, it is unlikely in this sense that one can find an algorithm that simultaneously conserves the symplectic structure, the momentum map, and the Hamiltonian. It is tempting (but probably wrong) to guess from this that one can monitor accuracy by keeping track of all three.

Comment on Generating Functions. We remark that many symplectic-momentum integrators have been based on the use of generating functions. We refer to Channell and Scovel [1990] and Ge [1991a] for surveys. They are based on the (standard) fact that if $S : Q \times Q \rightarrow \mathbb{R}$ defines a diffeomorphism $(q_0, p_0) \mapsto (q, p)$ implicitly by

$$p = \frac{\partial S}{\partial q} \quad \text{and} \quad p_0 = -\frac{\partial S}{\partial q_0}, \quad (26)$$

then this diffeomorphism is symplectic. One of the basic facts about Hamilton-Jacobi theory is that the flow of Hamilton’s equations is the canonical transformation generated by the solution of the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}\right) = 0, \quad (27)$$

where $S(q_0, q, t)|_{q=q_0, t=0}$ generates the identity. (This may require singular behavior in t ; for example, consider $S = \frac{1}{2t}(q - q_0)^2$.) The strategy is to find an approximate solution of the Hamilton-Jacobi equation for small time Δt and to use this to obtain the desired algorithm using (26).

There are several other versions of the algorithm that one can also treat. For example, if specific coordinates are chosen on the phase space, one can use a generating function of the form $S(q^i, p_{0i}, t)$. In this case one can get the simple formula for a first order algorithm given in Example 1 above by using

$S = p_{0i}q^i - \Delta t H(q^i, p_{0i})$, which is easy to implement, and for Hamiltonians of the form kinetic plus potential, leads to the stated explicit symplectic algorithm. As explained in Ge [1991a], one can use other types of generating functions. For example, using the Poincaré generating function, one recovers the mid-point scheme. There are some general results which say that invariant generating functions produce symplectic-momentum algorithms. The simplest version of this states that if $S : Q \times Q \rightarrow \mathbb{R}$ is invariant under the diagonal action of G , *i.e.*, $S(gq, gq_0) = S(q, q_0)$, then the cotangent momentum map \mathbf{J} is invariant under the canonical transformation φ_S generated by S , *i.e.*, $\mathbf{J} \circ \varphi_S = \mathbf{J}$. See Ge [1991a] and Marsden [1992] and references therein for additional information.

By putting together symplectic integrators following the generating function approach on T^*G and using the reduced Hamilton-Jacobi equation *i.e.*, the *Lie-Poisson Hamilton-Jacobi equation*, Ge and Marsden [1988] constructed an interesting class of Lie-Poisson integrators. The method in particular applies to rigid body dynamics. Unfortunately, the construction of integrators using generating functions can involve complex algebra to set up and it can sometimes be awkward to use in the presence of constraints.

Comments on Energy-Momentum Integrators. Because of the preceding theorem, algorithms that conserve the Hamiltonian and the momentum map will not, in general, conserve the symplectic structure. The following is one strategy for constructing energy-momentum integrators:

- i.** Formulate an energy-preserving algorithm on the symplectic reduced phase space given by $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ or the Poisson reduced space P/G . If such an algorithm is interpreted in terms of the primitive phase space P , it becomes an iterative mapping from one orbit to another.
- ii.** In terms of canonical coordinates (q, p) on P , interpret the orbit-to-orbit mapping described above and if P/G was used, impose the constraint $\mathbf{J}(q_k, p_k) = \mathbf{J}(q_{k+1}, p_{k+1})$. The constraint does not uniquely determine the restricted mapping, so we may obtain a large class of iterative schemes.
- iii.** To determine a map in the above class, we must determine how points in one G_μ -orbit are mapped to points in another orbit. There is ambiguity about how phase space points drift along the G_μ -orbits, which is related to *geometric phases*. By discretizing the geometric phase, we can specify the shift along each G_μ -orbit associated with each iteration of the map.

Simo and Wong [1989], Krishnaprasad and Austin [1990] and Simo, Tarnow and Wong [1991] provide methods for making the choices required in Steps **ii** and **iii**. The projection from the level set of constant angular

momentum onto the surface of constant energy can be performed implicitly or explicitly leading to predictor/corrector type of algorithms. The cost involved in the construction of the projection reduces to a rapid line search. The algorithm in Simo and Wong [1989] is special in that the projection is not needed for $Q = \text{SO}(3)$: the discrete flow lies in the intersection of the level set of angular momentum and the surface of constant energy. This algorithm is singularity-free and integrates the dynamics exactly up to a time reparametrization, consistent with the restrictions on mechanical integrators given above. Extensions of these schemes to elasticity, rods and shells amenable to parallelization are given in Simo and Doblare [1991]. When one is constructing an energy-momentum integrator by a method like this, a main issue is how to deal with systems with constraints. As we shall see, variational symplectic-momentum integrators handle this question rather nicely.

5 The Discrete Variational Principle

A *discrete variational principle (DVP)* is presented in this section that leads to evolution equations that are discrete analogs to the Euler-Lagrange equations. We call the evolution equations *discrete Euler-Lagrange (DEL)* equations. Some, but not all of the results in this section are found in Veselov [1988], Veselov [1988], Moser and Veselov [1991] and Baez and Gilliam [1995] but are rederived here in the context of the notation and context for geometric mechanics we have recalled above for consistent notation, completeness, and clarity.

The Discrete Variational Principle. We define a *discrete Lagrangian* on a configuration manifold Q to be a map $\mathbb{L} : Q \times Q \rightarrow \mathbb{R}$. Fixing a positive integer N , the *action sum* is the map $\mathbb{S} : Q^{N+1} \rightarrow \mathbb{R}$ defined by

$$\mathbb{S} = \sum_{k=0}^{N-1} \mathbb{L}(q_{k+1}, q_k), \quad (28)$$

where $q_k \in Q$ and $k \in \mathbb{Z}$ is the discrete time. The *discrete variational principle* states that the evolution equations extremize the action sum given fixed end points, q_0 and q_N . Extremizing \mathbb{S} over q_1, \dots, q_{N-1} leads to the DEL equations:

$$D_2\mathbb{L}(q_{k+1}, q_k) + D_1\mathbb{L}(q_k, q_{k-1}) = 0 \quad \text{for all } k \in \{1, \dots, N-1\} \quad (29)$$

i.e.,

$$D_2\mathbb{L} \circ \Phi + D_1\mathbb{L} = 0, \quad (30)$$

where $\Phi : Q \times Q \rightarrow Q \times Q$ is defined implicitly by $\Phi(q_k, q_{k-1}) = (q_{k+1}, q_k)$. If $D_2\mathbb{L}$ is invertible, then Equation (30) defines the discrete map, Φ , which flows the system forward in discrete time.

As was observed by Patrick, this scheme is identical to the scheme obtained by using the generating function $S(q_1, q_2) = -L(q_1, q_2)$, so that some of the facts below may also be deduced from results in Ge and Marsden [1988]. However, the way the variational method deals with constraints, namely with Lagrange multipliers on an ambient linear space, makes it attractive.

Conservation of the Symplectic Structure. The symplectic structure on $Q \times Q$ is defined next and an equation for the symplectic form on $Q \times Q$ is given. We then shown that Φ preserves this symplectic form. We then derive a discrete Noether's theorem by showing that invariance of the discrete Lagrangian leads to a conserved quantity, a momentum map, for the flow of Φ .

We first define the *discrete fiber derivative* by

$$\mathbb{F}L : Q \times Q \rightarrow T^*Q; \quad (q_1, q_0) \mapsto (q_0, D_2\mathbb{L}(q_1, q_0)) \quad (31)$$

and define a 2-form ω on $Q \times Q$ by pulling back the canonical 2-form Ω_{CAN} on T^*Q :

$$\omega = \mathbb{F}L^*(\Omega_{\text{CAN}}) = \mathbb{F}L^*(-\mathbf{d}\Theta_{\text{CAN}}) = -\mathbf{d}(\mathbb{F}L^*(\Theta_{\text{CAN}})), \quad (32)$$

where \mathbf{d} is the exterior derivative. Choose coordinates, q^i , on Q and choose the canonical coordinates, (q^i, p_i) , on T^*Q . In these coordinates, $\Omega_{\text{CAN}} = dq^i \wedge dp_i$ and $\Theta_{\text{CAN}} = p_i dq^i$. The DEL equations are

$$\frac{\partial \mathbb{L}}{\partial q_k^i} \circ \Phi(q_{k+1}, q_k) + \frac{\partial \mathbb{L}}{\partial q_{k+1}^i}(q_{k+1}, q_k) = 0 \quad (33)$$

i.e.,

$$\frac{\partial \mathbb{L}}{\partial q_{k+1}^i}(q_{k+2}, q_{k+1}) + \frac{\partial \mathbb{L}}{\partial q_{k+1}^i}(q_{k+1}, q_k) = 0. \quad (34)$$

Continuing the calculations in Equation (32) gives

$$\begin{aligned}
\omega &= -\mathbf{d} \left(\frac{\partial \mathbb{L}}{\partial q_k^i} (q_{k+1}, q_k) \right) dq_k^i \\
&= -\frac{\partial^2 \mathbb{L}}{\partial q_k^i \partial q_{k+1}^j} (q_{k+1}, q_k) dq_{k+1}^j \wedge dq_k^i - \frac{\partial^2 \mathbb{L}}{\partial q_k^i \partial q_k^j} (q_{k+1}, q_k) dq_k^j \wedge dq_k^i \\
&= \frac{\partial^2 \mathbb{L}}{\partial q_k^i \partial q_{k+1}^j} (q_{k+1}, q_k) dq_k^i \wedge dq_{k+1}^j, \tag{35}
\end{aligned}$$

since the second term on the second line vanishes.

Theorem 5.1 *The map Φ preserves the symplectic form ω , i.e. $\Phi^*\omega = \omega$ where $\Phi^*\omega$ denotes the pullback of ω by the map Φ .*

This means that when the algorithm is transferred to the cotangent bundle via the discrete fiber derivative, one gets a symplectic integrator for the standard symplectic structure.

The proof of the theorem is very simple. Let $\Phi(y, x) = (u, v)$ and write $\omega = \mathbf{d}(p(y, x)dx) = D_{12}\mathbb{L}(y, x)dx \wedge dy$. In this notation, $y = v = q_{k+1}$, $x = q_k$, and $u = q_{k+2}$. To show that $\Phi^*\omega = \omega$, write

$$\begin{aligned}
\Phi^*\omega &= \Phi^* \left(-d \left(\frac{\partial \mathbb{L}}{\partial v^i} (u, v) dv^i \right) \right) = -\mathbf{d} \left(\Phi^* \left(\frac{\partial \mathbb{L}}{\partial v^i} (u, v) dv^i \right) \right) \\
&= -\mathbf{d} \left(\frac{\partial \mathbb{L}}{\partial v^i} \circ \Phi (y, x) d(v^i(y, x)) \right) = -\mathbf{d} \left(-\frac{\partial \mathbb{L}}{\partial y^i} (y, x) dy^i \right) \\
&= \frac{\partial^2 \mathbb{L}}{\partial x^j \partial y^i} dx^j \wedge dy^i = \omega \tag{36}
\end{aligned}$$

Here we have used Equation (34) and the fact that $\mathbf{d}(v(y, x)) = dy$.

The Discrete Noether Theorem. We now derive a discrete version of Noether's theorem by the same method we used in the continuous case. Assume that the discrete Lagrangian \mathbb{L} is invariant under the diagonal action of a Lie group G that acts on Q , and let $\xi \in \mathfrak{g}$ where \mathfrak{g} is the Lie algebra of G . Invariance of \mathbb{L} implies that

$$\mathbb{L}(\exp(s\xi)q_{k+1}, \exp(s\xi)q_k) = \mathbb{L}(q_{k+1}, q_k). \tag{37}$$

Differentiating (37) and setting $s = 0$ implies that

$$D_1\mathbb{L}(q_{k+1}, q_k) \cdot \xi_Q(q_{k+1}) + D_2\mathbb{L}(q_{k+1}, q_k) \cdot \xi_Q(q_k) = 0, \tag{38}$$

where ξ_Q is the infinitesimal generator. Consider the action sum, (28), where $0 < k < N - 1$ and vary q_{k+1} using a parameter $s \in \mathbb{R}$ by considering $q_{k+1}(s) = \exp(s\xi)q_{k+1}$. Since $q_{k+1}(0)$ extremizes \mathbb{S} , $d\mathbb{S}/ds = 0$ at $s = 0$, which implies that

$$D_1\mathbb{L}(q_{k+1}, q_k) \cdot \xi_Q(q_{k+1}) + D_2\mathbb{L}(q_{k+2}, q_{k+1}) \cdot \xi_Q(q_{k+1}) = 0. \quad (39)$$

Subtracting Equation (38) from Equation (39) reveals that

$$D_2\mathbb{L}(q_{k+2}, q_{k+1}) \cdot \xi_Q(q_{k+1}) - D_2\mathbb{L}(q_{k+1}, q_k) \cdot \xi_Q(q_k) = 0. \quad (40)$$

If we define the momentum map, $\mathbb{J} : Q \times Q \rightarrow \mathfrak{g}^*$, by

$$\langle \mathbb{J}(q_{k+1}, q_k), \xi \rangle \triangleq \langle D_2\mathbb{L}(q_{k+1}, q_k), \xi_Q(q_k) \rangle, \quad (41)$$

then (40) shows that the momentum map is preserved by $\Phi : Q \times Q \rightarrow Q \times Q$, and so we have proved the following:

Theorem 5.2 *Assume that the discrete Lagrangian \mathbb{L} is invariant under the diagonal action of a Lie group G and let the discrete momentum map be defined by (41). Then $\mathbb{J} \circ \Phi = \mathbb{J}$.*

The discrete momentum map \mathbb{J} is equivariant with respect to the action of G on $Q \times Q$ and the coadjoint action of G on \mathfrak{g}^* . This is proved as in the case of usual Lagrangians (see Marsden and Ratiu [1994]). Using these properties, one should be able to develop a theory of Lagrangian reduction in the discrete case, as with the continuous case. In particular, for Euler-Poincaré systems on a Lie algebra \mathfrak{g} , one expects to be able to directly discretize the constrained variational principle and arrive at the same integrator obtained by reducing the integrator we have constructed on TG .

6 Construction of Mechanical Integrators

We still must deal with how to discretize the Lagrangian with constraints. We first show how to construct integrators for Lagrangian systems with holonomic constraints by enforcing the constraints through the use of Lagrange multipliers. We then present a second, equivalent, construction procedure by choosing a set of generalized coordinates.

Constrained Formulation. Assume we have a mechanical system on a (finite dimensional) linear space V with a constraint manifold, $Q \subset V$, and that $L \rightarrow \mathbb{R}$ is the *unconstrained Lagrangian*, which, by restriction of L to TQ , defines the *constrained Lagrangian*, $L^c : TQ \rightarrow \mathbb{R}$. (See Marsden and Ratiu [1994] for basic facts about constrained systems.) We also assume that we have a vector valued constraint function, $g : V \rightarrow \mathbb{R}^k$, such that $g^{-1}(0) = Q \subset V$, with 0 a regular value of g . If $\dim V = n$, then $\dim Q = m = n - k$. We first define the *discrete, unconstrained Lagrangian*, $\mathbb{L} : V \times V \rightarrow \mathbb{R}$, to be

$$\mathbb{L}(y, x) = L \left(\frac{y+x}{2}, \frac{y-x}{h} \right), \quad (42)$$

where $h \in \mathbb{R}_+$ is the time step. The *discrete unconstrained action sum* is defined by

$$\mathbb{S} = \sum_{k=0}^{N-1} \mathbb{L}(v_{k+1}, v_k). \quad (43)$$

We then extremize $\mathbb{S} : V^{N+1} \rightarrow \mathbb{R}$ subject to the constraint that $v_k \in Q \subset V$ for $k \in \{1, \dots, N-1\}$, which via Lagrange multipliers leads to

$$D_2 \mathbb{L}(v_{k+1}, v_k) + D_1 \mathbb{L}(v_k, v_{k-1}) + \lambda_k^T Dg(v_k) = 0 \quad (44)$$

(with no sum over k), with constraints $g(v_k) = 0$, for all $k \in \{1, \dots, N-1\}$.

Given v_k and v_{k-1} in $Q \subset V$, *i.e.*, $g(v_k) = 0$ and $g(v_{k-1}) = 0$, we need to solve

$$D_2 \mathbb{L}(v_{k+1}, v_k) + D_1 \mathbb{L}(v_k, v_{k-1}) + \lambda_k^T Dg(v_k) = 0 \quad (45)$$

along with the constraints

$$g(v_{k+1}) = 0 \quad (46)$$

for v_{k+1} and λ_k .

In terms of the original, unconstrained Lagrangian, (45) reads as follows:

$$\begin{aligned} & \frac{1}{h} \left[\frac{\partial L}{\partial \dot{v}} \left(\frac{v_k + v_{k-1}}{2}, \frac{v_k - v_{k-1}}{h} \right) - \frac{\partial L}{\partial \dot{v}} \left(\frac{v_{k+1} + v_k}{2}, \frac{v_{k+1} - v_k}{h} \right) \right] \\ & + \frac{1}{2} \left[\frac{\partial L}{\partial v} \left(\frac{v_k + v_{k-1}}{2}, \frac{v_k - v_{k-1}}{h} \right) + \frac{\partial L}{\partial v} \left(\frac{v_{k+1} + v_k}{2}, \frac{v_{k+1} - v_k}{h} \right) \right] \\ & + D^T g(v_k) \lambda_k = 0 \end{aligned}$$

together with $g(v_{k+1}) = 0$. For example, if the continuous Lagrangian system is of the form

$$L(q, \dot{q}) = \frac{1}{2} \dot{q}^T M \dot{q} - V(q) \quad (47)$$

with constraints $g(q) = 0$, where M is a constant mass matrix, and V is the potential energy, then the DEL equations are

$$M \left(\frac{v_{k+1} - 2v_k + v_{k-1}}{h^2} \right) + \frac{1}{2} \left(\frac{\partial V}{\partial q} \left(\frac{v_{k+1} + v_k}{2} \right) + \frac{\partial V}{\partial q} \left(\frac{v_k + v_{k-1}}{2} \right) \right) - D^T g(v_k) \lambda_k = 0$$

together with the constraints $g(v_{k+1}) = 0$.

Intrinsic Formulation. For the intrinsic formulation (also called the generalized coordinate formulation, even though the construction is not coordinate dependent), we form the discrete Lagrangian and the action sum restricted to $Q \subset V$, and then perform the extremization directly on Q . The *constrained, discrete Lagrangian* is the map $\mathbb{L}^c : Q \times Q \rightarrow \mathbb{R}$, defined by $\mathbb{L}^c = \mathbb{L}|_{Q \times Q}$. Given a local coordinate chart, $\psi : U \subset \mathbb{R}^m \rightarrow Q \subset V$, (note that it goes *into* V) where U is an open set in \mathbb{R}^m , the constrained, discrete Lagrangian is given by

$$\begin{aligned} \mathbb{L}^c(q_{k+1}, q_k) &= \mathbb{L}(\psi(q_{k+1}), \psi(q_k)) \\ &= L \left(\frac{\psi(q_{k+1}) + \psi(q_k)}{2}, \frac{\psi(q_{k+1}) - \psi(q_k)}{h} \right). \end{aligned}$$

The *constrained action sum* is

$$\mathbb{S}^c = \sum_{k=0}^{N-1} \mathbb{L}^c(q_{k+1}, q_k). \quad (48)$$

Extremizing $\mathbb{S}^c : Q^{N+1} \rightarrow \mathbb{R}$ gives the discrete Euler-Lagrange (DEL) equations in terms of generalized coordinates,

$$D_2 \mathbb{L}^c(q_{k+1}, q_k) + D_1 \mathbb{L}^c(q_k, q_{k-1}) = 0. \quad (49)$$

In terms of the original, unconstrained Lagrangian, Equation (49) equals

$$\begin{aligned} D^T \psi(q_k) \left\{ \frac{1}{h} \left[\frac{\partial L}{\partial \dot{v}}(a_k, d_k) - \frac{\partial L}{\partial \dot{v}}(a_{k+1}, d_{k+1}) \right] \right. \\ \left. + \frac{1}{2} \left[\frac{\partial L}{\partial v}(a_k, d_k) + \frac{\partial L}{\partial v}(a_{k+1}, d_{k+1}) \right] \right\} = 0 \end{aligned} \quad (50)$$

where $a_k = \frac{1}{2}[\psi(q_k) + \psi(q_{k-1})]$ and $d_k = \frac{1}{h}[\psi(q_k) - \psi(q_{k-1})]$. We solve Equations (50) for q_{k+1} given q_k and q_{k-1} to advance the flow one time step.

Equivalence of the Formulations. The constrained and generalized coordinate formulations are in fact equivalent according to the next theorem.

Theorem 6.1 *Let g be the constraint function and ψ be the coordinate chart defined above. Let q_k and q_{k-1} be the two initial points in the coordinate chart and let $v_k = \psi(q_k)$ and $v_{k-1} = \psi(q_{k-1})$. Let $Dg(v_k)$ and $D\psi(q_k)$ be full rank. Then the generalized formulation, Equation (50), has a solution for q_{k+1} if and only if the constrained formulation has a solution for v_{k+1} and λ_k . Furthermore, $v_{k+1} = \psi(q_{k+1})$.*

This is proved by a straightforward chase of the definitions (see Wendlandt and Marsden [1996] for the details). The procedure is the discrete analogue of the well known fact that one can do constrained Lagrangian dynamics either via Lagrange multipliers or by a direct restriction of the Lagrangian (see Marsden and Ratiu [1994] for example.)

For the numerical examples presented later, we solve the DEL equations, Equation (45), using Newton-Raphson equation solvers. These solvers require the construction of a Jacobian formed by differentiating (45) and (46) with respect to v_{k+1} and λ_k . For many applications, the resulting Jacobian is nearly symmetric and sparse; this and scaling tricks can be exploited to increase the simulation efficiency. For tree structured multibody systems, one can show that the linear equations involving the Jacobian can be solved in linear time. Sparse matrix techniques and symplectic integration are also used for multibody systems in Barth and Leimkuhler [1996a].

Local Truncation Error and Solvability. To calculate the truncation error, we first insert an exact solution of the differential equations into (50), and expand the result in powers of the step size h . To calculate the expansion, we expand Equation (50) about $v_k^i = \psi^i(q_k)$ and $\dot{v}_k^i = (\partial\psi^i/\partial q_k^j)\dot{q}_k^j$, and expand the result in powers of h . This lengthy calculation, not reproduced here, shows that the local truncation error of the method is *second order*. The coefficient of the h^2 term is a lengthy expression involving second, third, and fourth partial derivatives of $L : TV \rightarrow \mathbb{R}$.

If one uses the following definition for the discrete Lagrangian:

$$\mathbb{L}(y, x) = L\left(y, \frac{y-x}{h}\right), \quad (51)$$

then the resulting DEL equations will only be first order accurate for a general Lagrangian. There is no cancellation of terms in the h^1 term as there is with the definition in Equation (42). However, in some cases, the

resulting DEL equations may be explicit while the DEL equations from the definition in Equation (42) are implicit. An example of this occurring is if the continuous Lagrangian is in the form in Equation (47), and there are no constraints.

One can show, using the implicit function theorem, that if $D_{22}L$ is non-singular and if the Jacobian of the constraints is full rank, then for a sufficiently small time step, the generalized coordinate DEL equations are solvable for q_{k+1} .

The Symplectic Form and Discrete Momentum Map. The integrators created through the construction procedure are symplectic-momentum integrators. The integrators are symplectic in that the map produced on T^*V or T^*Q is a symplectic map. Also, if the Lie group acts linearly on V , then the continuous flow of the Euler-Lagrange equations and the discrete map produced from the DEL equations preserve the same momentum map on T^*Q .

If one accurately integrates the continuous equations and uses the result to initialize the discrete equations, one will notice that the *value* of the momentum map will differ from the value of the momentum map for the continuous system. The difference arises from the difference in the assignment of the momentum coordinate in T^*V through the discrete fiber derivative. In the continuous case, the momentum is D_2L while in the discrete case, we use $-hD_2\mathbb{L}$. We multiply by a $-h$ from the definitions given in Equation (31) because $-hD_2\mathbb{L}$ converges to D_2L as $h \rightarrow 0$.

If the Lagrangian of a continuous system is invariant under the action of a group, and if the constraints are also invariant under the group action, *i.e.* if we have the identities

$$L(G \cdot v, G \cdot \dot{v}) = L(v, \dot{v}), \quad g(G \cdot v) = g(v),$$

where the action of G on $v \in V$ is represented as $G \cdot v$, then the flow of the Euler-Lagrange equations preserve the momentum map $\mathbb{J} : TV \rightarrow \mathfrak{g}^*$, where

$$\langle \mathbb{J}(v, \dot{v}), \xi \rangle \triangleq \left\langle \frac{\partial L}{\partial \dot{v}}(v, \dot{v}), \xi_V(v) \right\rangle.$$

If the group G also acts linearly on V , then the discrete Lagrangian is also invariant under the group action as is readily verified. As earlier, the momentum map $\mathbb{J} : V \times V \rightarrow \mathfrak{g}^*$ defined by the relation

$$\langle \mathbb{J}(v_{k+1}, v_k), \xi \rangle \triangleq \langle D_2\mathbb{L}(v_{k+1}, v_k), \xi_V(v_k) \rangle$$

is conserved by the flow of the DEL equations.

We now calculate $-hD_2\mathbb{L}$ and notice that

$$\begin{aligned} & -hD_2\mathbb{L}(v_{k+1}, v_k) \\ &= -h \frac{\partial}{\partial v_k} \left(L \left(\frac{v_{k+1} + v_k}{2}, \frac{v_{k+1} - v_k}{h} \right) \right) \\ &= \frac{\partial L}{\partial \dot{v}} \left(\frac{v_{k+1} + v_k}{2}, \frac{v_{k+1} - v_k}{h} \right) - \frac{h}{2} \frac{\partial L}{\partial v} \left(\frac{v_{k+1} + v_k}{2}, \frac{v_{k+1} - v_k}{h} \right). \end{aligned}$$

As $h \rightarrow 0$, the discrete momentum value, $-hD_2\mathbb{L}$, converges to the continuous momentum value, D_2L . Therefore, the quantities that depend on the discrete momentum value, such as the *discrete momentum map* defined to be $-h\mathbb{J}$, converge to their continuous counterparts as $h \rightarrow 0$.

7 Numerical Examples

We apply the construction procedure to produce mechanical integrators for the rigid body and the double spherical pendulum (DSP). We choose to use constrained coordinates instead of generalized coordinates to avoid coordinate singularities and coordinate patching. We use unit quaternions to create the rigid body algorithm, and use the position of the two masses for the double spherical pendulum.

The Rigid Body. The algorithm presented here updates quaternion variables based on the previous two quaternion variables. The configuration manifold is taken to be $Q = S^3 \subset V$ where $V = \mathbb{R}^4$. Quaternions were used instead of using $V = \mathbb{R}^9$ with the six orthogonal constraints of $SO(3)$ primarily to avoid a large number of Lagrange multipliers. The constraint function is enforced with Lagrange multipliers.

Rigid body integrators that preserve certain mechanical properties have been created by several researchers as we noted in the introduction. An energy-momentum integrator is presented in Simo and Wong [1991]. A symplectic integrator which preserves the momentum and energy is presented in Lewis and Simo [1995]. A symplectic-momentum integrator is presented in McLachlan and Scovel [1995]. A rigid body integrator based on a discrete variational principle and in terms of 3×3 matrices with constraints is presented in Moser and Veselov [1991].

We first attach a body frame to the rigid body and represent the frame as a matrix, $R \in SO(3)$, which maps vectors in the body frame, \mathcal{B} , to vectors in

the spatial (inertial) frame \mathcal{S} , consistent with our earlier description of rigid body motion. The rotation matrix is thought of as a mapping $R : \mathcal{B} \rightarrow \mathcal{S}$.

We now recall some background on quaternions in this context (see Murray, Li, and Sastry [1994]). As a set, the quaternions comprise \mathbb{R}^4 and the set of unit quaternions form the three sphere $S^3 \subset \mathbb{R}^4$. If we think of a quaternion $a \in \mathbb{R}^4$ as consisting of a scalar value, a_s , and a vector with three components denoted $a_v = (a_x, a_y, a_z)$, then **quaternionic multiplication** is given in terms of ordinary multiplication and the dot and cross products as follows:

If a and b are quaternions, then quaternionic multiplication is given by $c = a \star b$, where $c_s = a_s b_s - a_v \cdot b_v$ and $c_v = a_s b_v + b_s a_v + a_v \times b_v$.

With this multiplication, the unit sphere S^3 is a three dimensional Lie group. The conjugate of a denoted \bar{a} is given by $\bar{a} = (a_s, -a_v)$ and for unit quaternions, \bar{a} is the inverse of a , in that $a \star \bar{a} = (1, 0, 0, 0)$.

The following formula constructs an element $A \in \text{SO}(3)$ from its unit quaternion representation, $a \in S^3$:

$$A = (2a_s^2 - 1)I + 2a_s \hat{a}_v + 2a_v a_v^T, \quad (52)$$

where the hat operation ($\hat{v} \cdot w = v \times w$) is as was defined earlier in our review of rigid body dynamics.

This defines a group homomorphism of $S^3 \rightarrow \text{SO}(3)$ which is two to one. Correspondingly, if A, B , and $C \in \text{SO}(3)$ are represented by unit quaternions, a, b , and c , respectively, then $C = AB$ if and only if $c = \pm a \star b$. An additional fact about quaternions is that if $w = Av$ and a is a unit quaternion that represents A , then $(0, w) = a \star (0, v) \star \bar{a}$ where $(0, w)$ is regarded as a quaternion formed from the vector w .

If $R : \mathcal{B} \rightarrow \mathcal{S}$ is the rotation matrix representing the orientation of the rigid body, recall that the body angular velocity vector, ω_b , is given by $\hat{\omega}_b = R^T \dot{R}$. In terms of quaternions, this reads as follows: if r is the unit quaternion representing R , then $\bar{r} \star \dot{r} = (0, \omega_b/2)$.

Using the above relationship for the body angular velocity, we construct the continuous Lagrangian, $L : TV \rightarrow \mathbb{R}$, to be

$$L(q, \dot{q}) = \frac{1}{2} (2\bar{q} \star \dot{q})^T \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I} \end{bmatrix} (2\bar{q} \star \dot{q}), \quad (53)$$

where \mathbb{I} is the inertia matrix. The constraint manifold Q is defined by $q_s^2 + q_v \cdot q_v = 1$.

The Lagrangian in Equation (53) is invariant under left quaternionic multiplication, *i.e.* $L(r \star q, r \star \dot{q}) = L(q, \dot{q})$, where r is a unit quaternion. The invariance leads to conservation of angular momentum.

The discrete Lagrangian, $\mathbb{L} : TV \rightarrow \mathbb{R}$, is chosen to be

$$\mathbb{L}(y, x) = L\left(\frac{y+x}{2}, \frac{y-x}{h}\right). \quad (54)$$

A calculation shows that when restricted to Q , we have

$$\mathbb{L}(y, x) = \frac{1}{2h^2}(\bar{x} \star y - \bar{y} \star x)^T \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I} \end{bmatrix} (\bar{x} \star y - \bar{y} \star x). \quad (55)$$

The discrete Lagrangian on all of $V \times V$ is then taken to be (55). Since we are extremizing \mathbb{S} restricted to Q , the extension of \mathbb{L} to $V \setminus Q$ can be chosen arbitrarily.

The discrete Lagrangian (55) is also invariant under left quaternionic multiplication, *i.e.* $\mathbb{L}(r \star y, r \star x) = \mathbb{L}(y, x)$, where r is a unit quaternion, and the invariance leads to conservation of discrete momentum which converges to the continuous momentum as the step size decreases, as we have seen.

An Example Simulation The DEL equations for the rigid body and relevant Jacobian are created in Mathematica and exported to C-code for simulation. The initial conditions and rigid body parameters are

$$q_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \quad \omega_b = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix}; \quad \mathbb{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}. \quad (56)$$

We must first initialize the rigid body integrator by choosing two initial quaternion values. We do this by using an Euler step with $\dot{q} = q \star (0, \omega_b/2)$ with $h = 10^{-5}$ s. We then use the DVP integrator with $h = 10^{-5}$ s to set the second initial point for $h = 10^{-4}$ s, 10^{-3} s, 10^{-2} s, and 10^{-1} s. The system is simulated for 30 seconds. To calculate errors in energy, momentum, and position, we first choose a standard value. We use the energy and momentum given initially after the first Euler step at $h = 10^{-5}$ s as the standard energy and momentum values. We use the results of the 30s simulation with $h = 10^{-4}$ s as the standard position variables. We use the following formula to calculate errors for each simulation:

$$\text{error} = \frac{1}{Nm} \sum_{i=1}^N \|v_i - v_i^s\|_2, \quad (57)$$

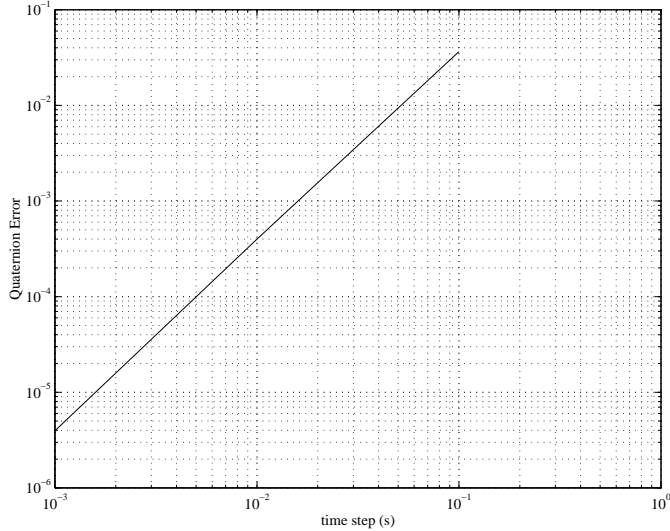


Figure 1: Quaternion Error Versus Time Step

where m is the length of the vector v_i , v_i^s is the standard value at the i th sample, and N is the number of samples. We find that the CPU time drops off nearly linearly as the time step increases; see Wendlandt and Marsden [1996] for details. The quaternion error versus time step is shown in Figure 1. The plot shows a second order relationship between error and time step, consistent with the theory.

Figure 2 compares the plot of the quaternion, q_y , versus time for the simulations at $h = 10^{-4}$ s and $h = 10^{-1}$ s. The trajectory for the large time step exhibits the same qualitative behavior as the small time step, but the deviations increase for longer simulation times.

The energy error versus time step reveals a second order relationship between energy error and time step. The energy for the $h = 10^{-4}$ s simulation deviates between 32.999999359J and 32.999999349J. The energy for the simulation at $h = 10^{-3}$ s deviates between 32.999937236J and 32.999937235J. There is no deviation in energy for the $h = 10^{-2}$ s and $h = 10^{-1}$ s simulations.

For each time step, the constant value of the discrete momentum map is conserved, and converges to the continuous momentum value as the step size decreases. There is a second order relationship between momentum error and time step.

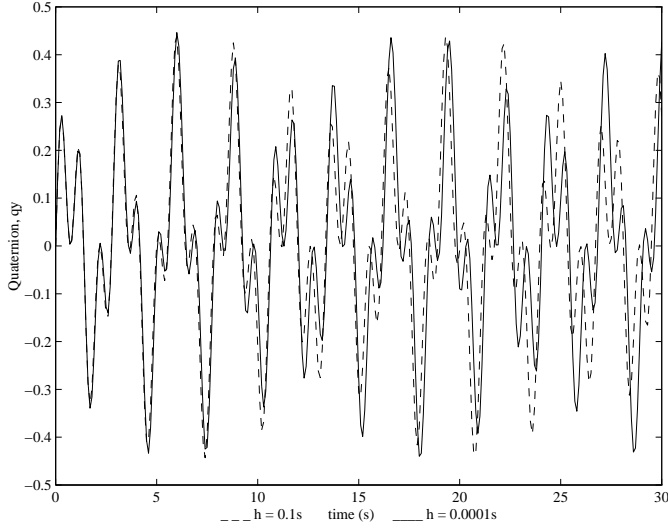


Figure 2: Quaternion Coordinate Versus Time

The Double Spherical Pendulum. The simulation in this section is motivated by our work on pattern evocation for the double spherical pendulum given in Marsden and Scheurle [1995] and Marsden, Scheurle and Wendlandt [1996].

The double spherical pendulum consists of two constrained point masses. The configuration space is $Q = S^2 \times S^2$ and the linear space is $V = \mathbb{R}^3 \times \mathbb{R}^3$. The position of the first mass is $q_1 = (x_1, y_1, z_1)$, and the position of the second mass is $q_2 = (x_2, y_2, z_2)$. The constraint function, given by the pendulum length constraints, is

$$g(v) = \begin{bmatrix} q_1 \cdot q_1 - l_1^2 \\ (q_2 - q_1) \cdot (q_2 - q_1) - l_2^2 \end{bmatrix}. \quad (58)$$

The DSP Lagrangian system is of the form in Equation (47), and the DVP algorithm for the DSP is, in this case, identical to the SHAKE algorithm:

$$\frac{1}{h} M [q^{n+1} - 2q^n + q^{n-1}] + hu - hD^T g(q^n) \lambda = 0$$

with the constraint $g(q^{n+1}) = 0$, where u is the column vector with components $(0, 0, m_1 g, 0, 0, m_2 g)$, m_1 and m_2 are the masses, and

$$M = \begin{bmatrix} m_1 I & 0 \\ 0 & m_2 I \end{bmatrix}, \quad q = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}. \quad (59)$$

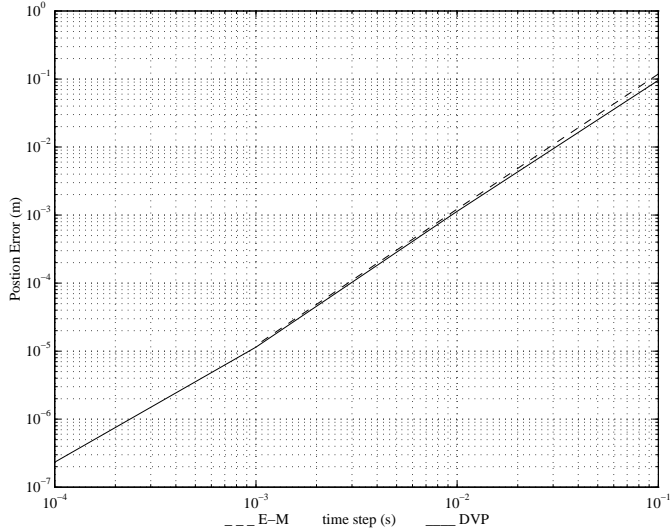


Figure 3: Position Error Versus Time Step for the DSP Simulation

Wendlandt and Marsden [1996] compare the simulation from the discrete variational principle (DVP) construction to an energy-momentum (EM) formulation based on Gonzalez [1996a], and applied to the DSP in Wendlandt [1995]. Here we summarize the results.

An Example Simulation. The following parameters are used for the DSP: $m_1 = 2.0\text{Kg}$, $m_2 = 3.5\text{Kg}$, $l_1 = 4.0\text{m}$, $l_2 = 3.0\text{m}$, and $g = 9.81\text{m/s}^2$. The initial conditions are $x_1 = 2.820\text{m}$, $y_1 = 0.025\text{m}$, $x_2 = 5.085\text{m}$, $y_2 = 0.105\text{m}$, $\dot{x}_1 = 3.381\text{m/s}$, $\dot{y}_1 = 2.506\text{m/s}$, $\dot{x}_2 = 2.497\text{m/s}$, and $\dot{y}_2 = 10.495\text{m/s}$. The position and velocity of the z -coordinate is determined from the constraints, and the z -coordinate for both masses is taken to be negative. The output of the EM simulation at a time step of 0.0001s is used as the standard and initializes the second step in the DVP simulations. The DVP simulations are slightly faster for each time step and both CPU times drop off nearly linearly with increasing time step.

The position error for the EM and DVP simulations is shown in Figure 3. Both simulations show a second order relationship between position error and time step. The error for the EM simulation is slightly greater than the error for the DVP simulation for $h \geq 10^{-3}\text{s}$.

The y position of the second mass is shown in Figure 4 for the EM and DVP simulations for $h = 0.0001\text{s}$ and $h = 0.1\text{s}$. Both the EM and DVP

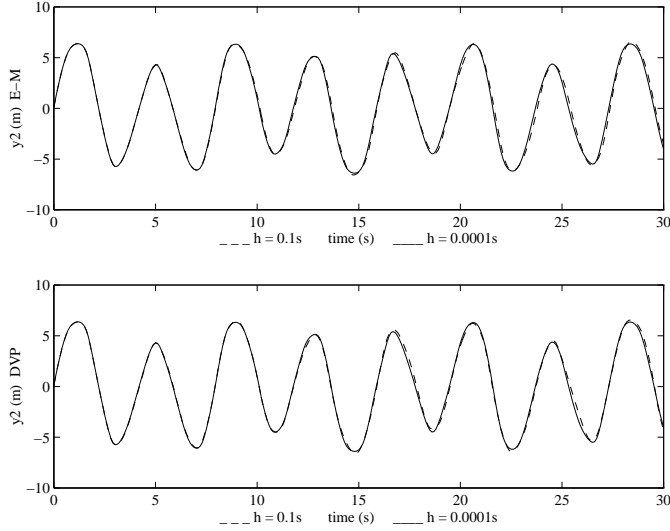


Figure 4: Position Coordinate Versus Time for the DSP Simulation

simulations at $h = 0.0001s$ overlap and cannot be distinguished when plotted on the same graph. For both the EM and DVP simulations, reasonably accurate and fast trajectories are produced at large time steps, $h = 0.1s$.

The DVP energy error appears to drop off as the square of the time step, at least for the large time steps. The energy error is zero for all time steps for the EM simulation. The energy for the DVP simulation at $h = 0.0001s$ deviates between $24.944495109J$ and $24.944499828J$ and deviates between $20.910805793J$ and $25.583335766J$ for $h = 0.1s$.

The DVP algorithm should preserve momentum but for the smallest time step, $h = 0.0001s$, the momentum varies between $199.825467170m^2/s$ to $199.825467184m^2/s$; this may be due to numerical errors. The momentum is constant for the other time steps. Again, the constant discrete momentum value approaches the value of the continuous momentum as the step size decreases.

Figure 5 shows the energy for the DVP simulations versus time for $h = 0.1s$ and $0.01s$ in the lower graph. The upper graph shows energy versus time for $h = 0.001s$ and $0.0001s$. The energy oscillates about a constant value, and the constant value approaches the true energy. The amplitude of the oscillations decrease as the step size decreases. The fluctuations in energy appear to be related to the constraint forces. The middle graph is a plot of the multipliers versus time, and the fluctuations in the multipliers is

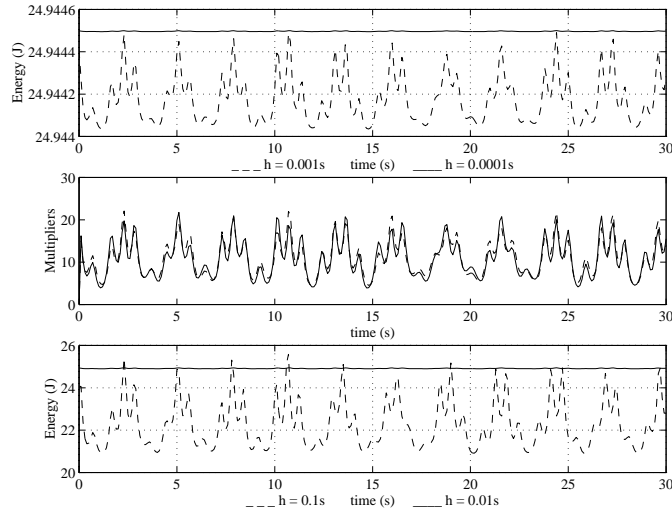


Figure 5: Energy and Multipliers Versus Time for the DSP Simulation

correlated to the fluctuations in energy (see Barth and Leimkuhler [1996a], who use variable step size to decrease the energy oscillation.)

Conclusions

This paper reviewed some of the basic theory of variational principles and integration algorithms and presented a discrete variational procedure to construct mechanical integrators for Lagrangian systems and applied it to the rigid body and the double spherical pendulum. The discrete Euler-Lagrange (DEL) equations share similarities to the continuous equations of motion and preserve a symplectic form and invariants resulting from group invariance of the Lagrangian. Some areas of future work are the following:

Energy-Momentum Integrators. One can look for an analogue of the DVP based on discretizing the principle of least action that would lead to energy-momentum integrators.

Additional Examples. It would be of interest to try out these methods on many other examples, such as the dynamics of an underwater vehicle studied by Leonard [1995a,b] and Leonard and Marsden [1996].

Discrete Lagrangian Reduction. It is natural to further investigate the dynamics on $(Q \times Q)/G$ induced by our G -invariant discrete Lagrangian analogous to the reduced Euler-Lagrange equations to produce the *discrete reduced Euler-Lagrange equations*. In particular, the relation between the

integrator for the rigid body as presented here and its induced reduced integrator (which is not just an integrator on the Lie algebra!) would be interesting to investigate.

Nonholonomic Systems. The method presented in this paper treats holonomic constraints and one would like to generalize the method to treat non-holonomic constraints, as in Bloch, Krishnaprasad, Marsden and Murray [1996]. For nonholonomic systems, the standard symplectic form is not preserved, and there are momentum equations and not conservation laws. Also, energy can be conserved in these systems. It would be of interest to develop algorithms taking into account these effects.

Multistep Methods and Time Step Control. It seems possible to modify the method to construct multistep mechanical integrators to increase the accuracy of the method. One would also like to modify the method to allow variable time steps to improve efficiency.

External Forces. It should be straightforward to add external and control forces to simulate controlled mechanical systems. The second author is currently using the techniques presented in this paper to develop a multibody simulator to simulate control systems for human models (see Wendlandt and Sastry [1996]).

Spacetime Integrators Since the method here is variational by nature and focuses on the temporal behavior, it should be helpful in the development of spacetime integrators by synthesis with existing finite element methods.

Acknowledgments

We thank Francisco Armero, Oscar Gonzalez, Abhi Jain, Ben Leimkuhler, Andrew Lewis, Robert MacKay, Richard Murray, George Patrick and Shmuel Weissman for useful discussions or comments.

References

- Abraham, R. and J.E. Marsden [1978] *Foundations of Mechanics*. Second Edition, Addison-Wesley.
- Abraham, R., J.E. Marsden, and T.S. Ratiu [1988] *Manifolds, Tensor Analysis, and Applications*. Second Edition, Applied Mathematical Sciences **75**, Springer-Verlag.
- Anderson, H. [1983]. Rattle: A velocity version of the shake algorithm for molecular dynamics calculations. *J. Comp. Phys.*, **52**, 24–34.

- Armero, F. and J.C. Simo [1996] Long-Term Dissipativity of Time-Stepping Algorithms for an Abstract Evolution Equation with Applications to the Incompressible MHD and Navier-Stokes Equations. *Comp. Meth. Appl. Mech. Eng.* **131**, 41–90.
- Arnold, V.I. [1989] *Mathematical Methods of Classical Mechanics*. Second Edition, Graduate Texts in Mathematics **60**, Springer-Verlag.
- Austin, M. and P.S. Krishnaprasad [1993] Almost Poisson Integration of Rigid Body Systems, *J. Comp. Phys.* **106**
- Baez, J. C. and J. W. Gilliam [1995]. An algebraic approach to discrete mechanics. Preprint, <http://math.ucr.edu/home/baez/ca.tex>.
- Barth, E. and B. Leimkuhler [1996a]. A semi-explicit, variable-stepsize integrator for constrained dynamics. Mathematics department preprint series, University of Kansas.
- Barth, E. and B. Leimkuhler [1996b]. Symplectic methods for conservative multibody systems. *Fields Institute Communications*, **10**, 25–43.
- Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and T.S. Ratiu [1994] Dissipation Induced Instabilities, *Ann. Inst. H. Poincaré, Analyse Nonlineaire* **11**, 37–90.
- Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and T.S. Ratiu [1996] The Euler-Poincaré equations and double bracket dissipation. *Comm. Math. Phys.* **175**, 1–42.
- Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and R. Murray [1996] Nonholonomic mechanical systems with symmetry. *Arch. Rat. Mech. An.* (to appear).
- Bretherton, F.P. [1970] A note on Hamilton’s principle for perfect fluids. *J. Fluid Mech.* **44**, 19-31.
- Channell, P. and C. Scovel [1990] Symplectic integration of Hamiltonian systems, *Nonlinearity* **3**, 231–259.
- Chorin, A.J, T.J.R. Hughes, J.E. Marsden, and M. McCracken [1978] Product Formulas and Numerical Algorithms, *Comm. Pure Appl. Math.* **31**, 205–256.
- Feng, K. [1986] Difference schemes for Hamiltonian formalism and symplectic geometry. *J. Comp. Math.* **4**, 279–289.

- Feng, K. and Z. Ge [1988] On approximations of Hamiltonian systems, *J. Comp. Math.* **6**, 88–97.
- Feng, K. and M.Z. Qin [1987] The symplectic methods for the computation of Hamiltonian equations, *Springer Lecture Notes in Math.* **1297**, 1–37.
- Ge, Z. [1990] Generating functions, Hamilton-Jacobi equation and symplectic groupoids over Poisson manifolds, *Indiana Univ. Math. J.* **39**, 859–876.
- Ge, Z. [1991a] Equivariant symplectic difference schemes and generating functions, *Physica D* **49**, 376–386.
- Ge, Z. [1991b] A constrained variational problem and the space of horizontal paths, *Pacific J. Math.* **149**, 61–94.
- Ge, Z. [1994] Generating functions, Hamilton-Jacobi equations and symplectic groupoids on poisson manifolds. *Indiana U. Math. J.* **39**, 850–875.
- Ge, Z. and J.E. Marsden [1988] Lie-Poisson integrators and Lie-Poisson Hamilton-Jacobi theory, *Phys. Lett. A* **133**, 134–139.
- Gillilan, R. and K. Wilson [1992]. Shadowing, rare events, and rubber bands: A variational Verlet algorithm for molecular dynamics. *J. Chem. Phys.*, **97**, 1757–1772.
- Gonzalez, O. [1996a] *Design and Analysis of Conserving Integrators for Nonlinear Hamiltonian Systems with Symmetry*, Thesis, Stanford University, Mechanical Engineering.
- Gonzalez, O. [1996b]. Time integration and discrete Hamiltonian systems. *Journal of Nonlinear Science*, to appear.
- Itoh, T. and K. Abe [1988]. Hamiltonian-conserving discrete canonical equations based on variational difference quotients. *Journal of Computational Physics*, **77**, 85–102.
- Jay, L. [1996]. Symplectic partitioned runge-kutta methods for constrained Hamiltonian systems. *SIAM Journal on Numerical Analysis*, **33**, 368–387.

- Koon, W.S. and J.E. Marsden [1996] Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction. *SIAM J. Control and Optim (to appear)*.
- Labudde, R. A. and D. Greenspan [1974]. Discrete mechanics-a general treatment. *Journal of Computational Physics*, **15**, 134–167.
- Labudde, R. A. and D. Greenspan [1976a]. 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.
- Labudde, R. A. and D. Greenspan [1976b]. 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.
- Leimkuhler, B. and G. Patrick [1996]. Symplectic integration on Riemannian manifolds. *J. of Nonl. Sci.* **6**, 367–384.
- Leimkuhler, B. and R. Skeel [1994]. Symplectic numerical integrators in constrained Hamiltonian systems. *Journal of Computational Physics*, **112**, 117–125.
- Leonard, N.E. [1995a] Control synthesis and adaptation for an underactuated autonomous underwater vehicle, *IEEE J. of Oceanic Eng.* **20**, 211–220.
- Leonard, N.E. [1995b] Stability of a bottom-heavy underwater vehicle, *Mechanical and Aerospace Engineering, Princeton University Technical Report 2048*, to appear in *Automatica* **33**, March 1997.
- Leonard, N.E. and J.E. Marsden [1996] Stability and Drift at Nongeneric Momenta and Underwater Vehicle Dynamics, *Mechanical and Aerospace Engineering, Princeton University Technical Report 2075*,
- Lewis, D. and J.C. Simo [1995] Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups. *J. Nonlinear Sci.* **4**, 253–299.
- Lewis, H. and Kostelec, P. [1996]. The use of Hamilton’s principle to derive time-advance algorithms for ordinary differential equations. *Comp. Phys. Commun.*, to appear.

- MacKay, R. [1992]. Some aspects of the dynamics of Hamiltonian systems. In Broomhead, D. S. and Iserles, A., eds, *The Dynamics of numerics and the numerics of dynamics*, 137–193. Clarendon Press, Oxford.
- Maeda, S. (1981). Lagrangian formulation of discrete systems and concept of difference space. *Math. Japonica*, **27**, 345–356.
- Marsden, J.E. [1992], *Lectures on Mechanics* London Mathematical Society Lecture note series, **174**, Cambridge University Press.
- Marsden, J. E., G. W. Patrick, and W. F. Shadwick [1996] *Integration Algorithms and Classical Mechanics*. Fields Inst. Commun., **10**, Am. Math. Society.
- Marsden, J.E. and T.S. Ratiu [1994] *Symmetry and Mechanics*. Texts in Applied Mathematics, **17**, Springer-Verlag.
- Marsden, J.E. and J. Scheurle [1993a] Lagrangian reduction and the double spherical pendulum, *ZAMP* **44**, 17–43.
- Marsden, J.E. and J. Scheurle [1993b] The reduced Euler-Lagrange equations, *Fields Inst. Comm.* **1**, 139–164.
- Marsden, J.E. and J. Scheurle [1995] Pattern evocation and geometric phases in mechanical systems with symmetry, *Dyn. and Stab. of Systems*. **10**, 315–338.
- Marsden, J.E., J. Scheurle and J. Wendlandt [1996] Visualization of orbits and pattern evocation for the double spherical pendulum, *ICIAM 95: Mathematical Research, Academie Verlag, Ed. by K. Kirchgässner, O. Mahrenholtz and R. Mennicken* **87**, 213–232.
- McLaughlin, R.I. and C. Scovel [1995] Equivariant constrained symplectic integration. *J. of Nonl. Sci.* **5**, 233–256.
- McLachlan, R. I. and C. Scovel [1996]. A survey of open problems in symplectic integration. *Fields Inst. Comm.*, **10**, 151–180.
- Montgomery, R., J.E. Marsden, and T.S. Ratiu [1984] Gauged Lie-Poisson structures, *Cont. Math. AMS* **28**, 101–114.
- Moser, J. and A.P. Veselov [1991] Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Comm. Math. Phys.* **139**, 217–243.

- Murray, R., Z. Li, and S. Sastry [1994]. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, Boca Raton, Fl.
- Ortiz, M. (1986). A note on energy conservation and stability of nonlinear time-stepping algorithms. *Computers and Structures*, **24**, 167–168.
- Reich, S. [1993]. Symplectic integration of constrained Hamiltonian systems by Runge-Kutta methods. Technical Report 93-13, University of British Columbia.
- Reich, S. [1994] *Symplectic integrators for systems of rigid bodies*. (preprint, Inst. Angw. Anal. Stch.)
- Reich, S. (1994). Momentum preserving symplectic integrators. *Physica D*, 76(4):375–383.
- Ryckaert, J., 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. of Comp. Phys.*, **23**, 327–341.
- Sanz-Serna, J. M. [1988] Runge-Kutta schemes for Hamiltonian systems, *BIT* **28**, 877–883.
- Sanz-Serna, J. M. [1991] Symplectic integrators for Hamiltonian problems: an overview. *Acta Num.*, **1**, 243–286.
- Sanz-Serna, J. M. [1996] Backward error analysis of symplectic integrators. *Fields Inst. Comm.*, **10**, 193–205.
- Sanz-Serna, J. M. and M. Calvo [1994] *Numerical Hamiltonian Problems*. Chapman and Hall, London.
- Seliger, R.L. and G.B. Whitham [1968] Variational principles in continuum mechanics, *Proc. Roy. Soc. Lond.* **305**, 1–25.
- Shibberu, Y. [1994]. Time-discretization of Hamiltonian systems. *Computers Math. Applic.*, **28**, 123–145.
- Simo, J.C. and F. Armero [1994] Unconditional Stability and Long-Term Behavior of Transient Algorithms for the Incompressible Navier-Stokes and Euler Equations. *Comp. Meth. Appl. Mech. Eng.* **111**, 111–154.
- Simo, J.C. and M. Doblare [1991], Nonlinear Dynamics of Three-Dimensional Rods: Momentum and Energy Conserving Algorithms, *Intern. J. Num. Methods in Eng.*

- Simo, J.C. and O. Gonzalez. [1993], Assessment of Energy-Momentum and Symplectic Schemes for Stiff Dynamical Systems. *Proc. ASME Winter Annual Meeting, New Orleans*, Dec. 1993.
- Simo, J.C., D.R. Lewis, and J.E. Marsden [1991] Stability of relative equilibria I: The reduced energy momentum method, *Arch. Rat. Mech. Anal.* **115**, 15-59.
- Simo, J.C. and J.E. Marsden [1984] On the rotated stress tensor and a material version of the Doyle Ericksen formula, *Arch. Rat. Mech. Anal.* **86**, 213-231.
- Simo, J.C., J.E. Marsden, and P.S. Krishnaprasad [1988] The Hamiltonian structure of nonlinear elasticity: The material, spatial, and convective representations of solids, rods, and plates, *Arch. Rat. Mech. Anal.* **104**, 125-183.
- Simo, J.C., T.A. Posbergh, and J.E. Marsden [1990] Stability of coupled rigid body and geometrically exact rods: block diagonalization and the energy-momentum method, *Physics Reports* **193**, 280-360.
- Simo, J.C., T.A. Posbergh, and J.E. Marsden [1991] Stability of relative equilibria II: Three dimensional elasticity, *Arch. Rat. Mech. Anal.* **115**, 61-100.
- Simo, J.C. and N. Tarnow [1992] The discrete energy momentum method. Conserving algorithms for nonlinear elastodynamics, *ZAMP* **43**, 757-792.
- Simo, J.C. and N. Tarnow [1994] A New Energy Momentum Method for the Dynamics of Nonlinear Shells, *Int. J. Num. Meth. Eng.*, **37**, 2527-2549.
- Simo, J.C. and N. Tarnow and M. Doblare [1993] Nonlinear Dynamics of Three-Dimensional Rods: Exact Energy and Momentum Conserving Algorithms. *Comp. Meth. Appl. Mech. Eng.*
- Simo, J.C., N. Tarnow, and K.K. Wong [1992] Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics *Comp. Meth. Appl. Mech. Eng.* **100**, 63-116.
- Simo, J.C. and K.K. Wong [1989] Unconditionally stable algorithms for the orthogonal group that exactly preserve energy and momentum, *Int. J. Num. Meth. Eng.* **31**, 19-52, addendum, **33**, 1321-1323, 1992.

- Verlet, L. [1967]. Computer experiments on classical fluids. *Phys. Rev.*, **159**, 98–103.
- Veselov, A.P. [1988] Integrable discrete-time systems and difference operators. *Funct. An. and Appl.* **22**, 83–94.
- Veselov, A.P. [1991] Integrable Lagrangian correspondences and the factorization of matrix polynomials. *Funct. An. and Appl.* **25**, 112–123.
- Wendlandt, J. [1995]. Pattern evocation and energy-momentum integration of the double spherical pendulum. MA thesis, University of California at Berkeley. Department of Mathematics, CPAM-656.
- Wendlandt, J.M. and J.E. Marsden [1996] Mechanical integrators derived from a discrete variational principle, *Caltech CDS Technical Report 96-013*.
- Wendlandt, J. and Sastry, S. [1996]. Recursive workspace control of multi-body systems: A planar biped example, *IEEE CDC*, Kobe, Japan, Dec., 1996.
- Wisdom, J., and M. Holman [1992] Symplectic maps for the N body problem, *Astron. J.*, **102**, 1528–1538.
- Wisdom, J., and M. Holman and J. Touma [1996] *Symplectic correctors*. *Fields Inst. Comm.*, **10**, 217–244.
- Yoshida, H. [1990]. Construction of higher order symplectic integrators. *Phys. Lett. A*, **150**, 262–268.