A Brief Introduction to Variational Integrators

Adrián J. Lew and Pablo Mata A

Abstract In this chapter, a brief introduction to the formulation of variational methods for finite-dimensional Lagrangian systems is presented. To this end, the first two sections focus on describing the Lagrangian and Hamiltonian points of view of mechanics for systems evolving on manifolds. Special attention is paid to the construction of the Lagrangian function and to the role of Hamilton's variational principle in the deduction of the balance equations. The relation between the symmetries of the Lagrangian function and the existence of invariants of the dynamics along with the symplectic nature of the flow are also addressed. In the third section, the discussion turns towards the formulation of a time-discrete analogue of the theory. The cornerstone of such a construction is given by a discrete analogue of Hamilton's variational principle which provides a systematic procedure to construct discrete approximations to the exact trajectory of a mechanical system on both the configuration space and the phase space. The approximation properties and the geometric characteristics of the resulting discrete trajectories are explained. Finally, we apply the variational methodology to construct symplectic and momentum-conserving time integrators for two problems of practical interest in engineering and science.

A.J. Lew (✉)
Department of Mechanical Engineering, Stanford University, Stanford, CA 94305-4040, USA
e-mail: lewa@stanford.edu

P. Mata A
Centro de Investigación En Ecosistemas de la Patagonia, CIEP Universidad Austral de Chile, Km 4.5 Camino Aysén, Coyhaique, Chile
e-mail: pmata@ciep.cl

P. Betsch (ed.), *Structure-preserving Integrators in Nonlinear Structural Dynamics and Flexible Multibody Dynamics*, CISM International Centre for Mechanical Sciences 565, DOI 10.1007/978-3-319-31879-0_5

1 Introduction

1.1 Overview

A bumper sticker explaining how to construct variational integrators (VI) would read

> "Approximate the action instead of the equations of motion."

This simple idea turns out to be very powerful, as we shall have the opportunity to explore in these notes. In fact, it has underpinned the solutions to elastostatics problems with the finite element method for 60 years now. Why, when, and how an approximation of a fundamental object in classical mechanics, the action, gives rise to a convergent scheme to integrate the equations of motion are questions that we shall address.

To explain the implications of the above bumper sticker, in the following sections we briefly review the Lagrangian formulation of the mechanics of a conservative system, and then we mimic this process at the discrete level to construct variational integrators.

1.2 Perspective

Early approaches toward the creation of time integrators for ordinary differential equations (ODE's) consisted in constructing a suitable discretization of equations, without accounting for additional structure these equations might have. See for example, (Hughes 1987; Bathe 1996; Hairer et al. 1993; Hairer and Wanner 1996; Holmes 2007) among others.

An alternative point of view is given by the formulation of the so-called *structure-preserving* methods. Those methods are designed to preserve the geometric properties of the flow of the differential equations. This category includes but is not limited to:

- Methods that conserve the invariants of the dynamics such as energy-conserving integrators applied to conservative systems (Bayliss and Issacson 1975; Labudde and Greenspan 1976) or energy-momentum conserving methods (Simo et al. 1992; Simo and Wong 1991) which conserve energy as well as linear and angular momentum. These ideas have been also applied to problems described by partial differential equations (PDE's), such as for example to the dynamics of nonlinear solids (Gonzalez 2000). The list of contributions in this area is long, with for example (Armero and Romero 2001; Armero and Petoz 1999; Bauchau and Bottasso 1999; Betsch and Uhlar 2007; Borri et al. 2001), to name only a few of them.
- Numerical methods for dynamical systems evolving on general manifolds rather than on linear spaces. The key feature of these methods consists in that the resulting discrete trajectory belongs to the same configuration manifold as the time-continuous system (Iserles 1997; Desbrun et al. 2014). Many important problems

in physics are described in terms of dynamical systems evolving on *Lie-groups* (see Sect. A.2) such as for example, the dynamics in space of N-interacting rigid bodies, the dynamics of slender structures such as rods and filaments, the dynamics of shells and the motion of incompressible and inviscid fluids among others. Lie-group methods are numerical integrators specifically formulated for dynamical problems on Lie groups or on manifolds acted upon by Lie groups (Iserles et al. 2000; Celledoni et al. 2014).The application of these methods to the full-body problem can be consulted in (Lee et al. 2007; Celledoni and Owren 2003), to the dynamics of elastic and inelastic rods in (Mata 2015; Mata et al. 2008, 2009; Romero and Armero 2002, Simo et al. 1995) and to the dynamics of shells e.g., in (Simo and Tarnow 1994; Sansour and Wagner 2003). An error analysis of Lie-group methods can be consulted in (Faltinsen 2000).

- Symmetric methods for reversible problems. Reversible dynamical systems are characterized by the fact that inverting the direction of the velocity vector while keeping the initial position fixed, results in an inversion of the solution trajectory. Conservative mechanical systems are reversible. Numerical methods able to generate reversible numerical flows when applied to a reversible differential equations constitute an active field of research in geometric time integration (Hairer et al. 2006, Chap. V; Cano and Sanz-Serna 1988).
- Symplectic methods for Lagrangian/Hamiltonian problems. The Lagrangian/Hamiltonian systems are among the most important dynamical systems in sciences and engineering. As noted in (Feng and Qin 2010) any conservative real physical process can be formulated as a Hamiltonian system, whether they have finite or infinite degrees of freedom. An outstanding property of Hamiltonian systems is the symplectic nature of their flows on the phase space. See Sects. 2.3 and 3.5 for a precise definition of continuous and discrete symplecticity. Examples of Hamiltonian systems appearing in science and engineering include but are not limited to the structural biology (Gay-Balmaz et al. 2009), molecular dynamics (Stavros 2014; Manning and Maddocks 1999), mathematical models in ecosystem dynamics (Kirwan 2008), superconductivity (Bogolyubov 1972), plasma physics (Larsson 1996), celestial mechanics and cosmology (Arnold et al. 2006), fluid mechanics (Desbrun et al. 2014; Gawlik et al. 2011), mechanics of materials and structures (Simo et al. 1988), theoretical physics (Esposito et al. 2004; Marsden 1988), aerospace engineering (Kasdin et al. 2005), satellite dynamics and control (Kuang et al. 2003; Koon et al. 2011), kinematics and dynamics of mechanisms and robots (Macchelli et al. 2009; Chen 1990) and other areas of seismic (Luo et al. 2013), mechanical and electrical (Clemente-Gallardo and Scherpen 2003) engineering. Symplectic integrators are methods specially formulated to produce a symplectic flows on the phase space. This property is intimately related to the ability of these methods to reproduce the long-time structure of the solutions of Hamiltonian ODE's (e.g., limit cycles, attractors, invariant manifolds, etc.) as it has been reported in several occasions e.g., (Bou-Rabee and Marsden 2009). A survey on symplectic time integration of Hamiltonian ODE's and Hamiltonian PDE's can be consulted in (Leimkuhler and Reich 2005; Feng and Qin 2010) and (Hairer et al. 2006, Chap. VI). The nonlinear stability of symplectic integrators

is considered in (McLachlan et al. 2004). The role of symplectic integration in optimal control is reviewed in (Chyba et al. 2009).

Therefore, formulating numerical methods able to preserve the geometric structure of the solutions of Hamiltonian systems will have extremely broad applications. This chapter focuses on describing a particular methodology, based on the discretization of a fundamental principle in mechanics, to construct structure-preserving methods for Lagrangian systems with symmetries.

1.3 Variational Integration

Variational integrators (VI's) constitute a more recent approach toward the creation of structure-preserving methods for finite-dimensional Lagrangian systems or for appropriate discretizations of some Hamiltonian continuum systems. Their construction is based on the formulation of a discrete analogue to Hamilton's variational principle. The basic idea consists in constructing a time-discrete approximation of the action integral called the discrete action sum. Stationary points of the discrete action sum are then the discrete-in-time trajectories of the mechanical system, and can be proved to approximate the exact trajectories as the time-step goes to zero. These ideas were originally explored in the works of (Maeda 1980; Maeda 1982) and of (Veselov 1988; Moser and Veselov 1991) in the context of integrable systems in mechanics.

The procedure used to construct a time-discrete trajectory defines a variational time integrator that shows a number of remarkable properties among which are:

(i) It conserves the invariants of the dynamics associated to the symmetries of the original mechanical system if the discrete action sum is designed to preserve the same symmetries. See Sects. 2.3 and 3.4. This is also known as a discrete version of Noether's theorem, see e.g., (Marsden and West 2001; Lew et al. 2004, 2003).
(ii) A discrete version of the Legendre transform allows to construct an alternative but otherwise equivalent form of the method that defines a discrete symplectic flow over the phase space. See Sect. 3.5.
(iii) The discrete trajectory displays an outstanding energy behavior. To be more precise, the value of the energy computed over the discrete trajectory remains close to its initial value for very long times, provided the time step is small enough, see e.g., (Leimkuhler and Reich 2005; Hairer et al. 2006).
(iv) Symplectic and momentum-conserving methods of arbitrarily high order of accuracy for dynamical systems evolving on nonlinear manifolds can be systematically constructed following the standard methodology of variational integration.

The essential aspects of VI's can be reviewed in (Wendlandt and Marsden 1997; Marsden and Wendlandt 1997; Marsden and West 2001) and in (Leok and Shingel 2012a; Leok 2005) general techniques for constructing variational integrators are provided. Spectral variational integrators are described in (Hall and Leok 2014a) and prolongation–collocation methods in (Leok and Shingel 2012b).

The variational methodology has been successfully applied to a broad range of fields such as for example:

- *Mechanical problems with multisymplectic geometry*. In (Marsden et al. 1998) a geometric-variational approach to continuous and discrete field theories is described and in (Marsden et al. 2001) the authors present a variational and multisymplectic formulation of both compressible and incompressible models of continuum mechanics. Asynchronous variational time integrators for finite element discretizations of deformable solids and field theories are formulated in (Lew et al. 2004, 2003; Lew 2003; Kale and Lew 2006; De León et al. 2008). An analysis of the stability properties of asynchronous VI's can be found in (Fong et al. 2008). See (Focardi and Maria-Mariano 2008) for a converge analysis of asynchronous VI's in linear elastodynamics and (Patrick and Cuell 2009) for a complete error analysis.

- *Dynamical systems evolving on nonlinear manifolds*. Discrete analogues of Euler–Poincaré and Lie–Poisson reduction theory for systems on finite-dimensional Lie groups with symmetries are developed in (Marsden et al. 1999). A Lie Poisson structure for a discrete mechanical system evolving on a Lie group is deduced in (Marsden et al. 2000). Lie group VI's applied to the full-body problem are formulated in (Lee et al. 2007) and VI's on two-spheres in (Lee et al. 2009). The extension of spectral variational integrators to Lie groups is developed in (Hall and Leok 2014b).

- *Structural elements: beams, rods, plates, and shells*. An explicit, second-order accurate VI that can be identified with a Lie-group, symplectic, partitioned Runge–Kutta method for finite element discretizations of geometrically exact rods is presented in (Mata 2015). The formulation of VI's for spatial beams and plates is carried out in (Demoures 2012; Demoures et al. 2014). In (Nichols and Murphey 2008) a VI for simulating the dynamics of cable structures is formulated. A discrete model for shells is formulated in (Grispun et al. 2003).

- *Contact and impact*. In (Fetecau et al. 2003b; Fetecau 2003) the classical theory of (smooth) Lagrangian mechanics is extended to the nonsmooth case in order to include collisions and the foundations of the multisymplectic formulation of nonsmooth continuum mechanics are presented in Fetecau et al. (2003a). An example of asynchronous collision integrators can be consulted in Wolff and Bucher (2013) and an application to polymer chains in (Leyendecker et al. 2012). In (Ryckman and Lew 2010, 2011, 2012) a new explicit dynamic contact algorithm that takes advantage of a variational asynchronous time integrator is formulated. A variational formulation of contact is formulated in (Harmon et al. 2009) and an

optimization of this method is carried out in (Ainsley et al. 2012). See also (Cirak and West 2005).

- *Multibody dynamics and control*. A comparison between the numerical performance of VI's and energy-momentum schemes when applied to the numerical simulation of flexible multibody dynamics is presented in Betsch et al. (2010). The solvability of some geometric integrators for multibody systems is analyzed in (Kobilarov 2014). A discontinuous version of VI's is formulated in (Johnson et al. 2014) to treat collisions in multibody systems. In (Jiménez et al. 2013) numerical methods for optimal control of mechanical systems in the Lagrangian setting are formulated.
- *Stochastic differential equations*. In (Bou-Rabee and Owhadi 2009) a continuous and discrete Lagrangian theory for stochastic Hamiltonian systems on manifolds is presented. See also (Wang et al. 2009; Wang 2007). The long-time statistical properties of a Lie–Trotter splitting for inertial Langevin equations are presented in (Bou-Rabee and Owhadi 2010). The splitting is defined as a composition of a variational integrator with an Ornstein–Uhlenbeck flow (Van Bargena and Dimitroff 2009). Further material about geometric integrators for stochastic dynamical systems can be found in Tao et al. (2010). Variational integrators for constrained, stochastic mechanical systems are presented in Bou-Rabee and Owhadi (2007).
- *Constrained and forced problems*. The formulation of variational integrators applied to Lagrangian systems subjected to holonomic constraints or forces can be consulted in (Marsden and West 2001; West 2004). A variational discrete null space method is proposed in (Leyendecker et al. 2008). See also (Leyendecker et al. (2007)). A study of the Γ-convergence of VI's to the corresponding continuum action functional and the convergence properties of the discrete trajectories to stationary points of the continuum problem is presented in Schmidt et al. (2009). The use of a discrete version of Lagrange–D'Alembert principle allows to include non-conservative generalized forces. This is particularly useful for weakly dissipative systems (Kane et al. 1999, 2000).
- *Dynamics of fluids*. A geometric theory for fluid dynamics can be found in (Marsden and Ratiu 1999; Arnold and Khesin 1998; Pavlov 2009). Traditionally, numerical methods for fluid dynamics have been rarely designed to preserve the geometric structure of the solution trajectories, resulting in the introduction of spurious numerical artifacts. In contrast, in (Pavlov et al. 2011) discrete equations of motion for fluid dynamics are derived from first principles in Eulerian form. In (Gawlik et al. 2011) a variational discretization of continuum theories arising in fluid dynamics, magnetohydrodynamics and the dynamics of complex fluids is presented and in (Desbrun et al. 2014) a structure-preserving scheme for the dynamics of rotating and/or stratified fluids is formulated.
- *Thermoelasticity and nonequilibrium thermodynamics*. The dynamics of systems undergoing irreversible processes has motivated the formulation of structure-preserving integrators able to satisfy the first (energy conservation) and second

(nondecreasing entropy of an isolated system) laws of thermodynamics along with the conservation of the invariants associated to the symmetries of the system. Those methods are frequently called *thermodynamically consistent*. See e.g., (Romero 2009; Bargmann and Steinmann 2008a, b). The simplest type of such systems are likely to be the thermomechanical systems. A Lagrangian/Hamiltonian formulation of thermoelasticity is obtained by introducing the concept of thermal displacements (Maugin 2000; Maugin and Kalpakides 2002) and (Green and Naghdi 1993, 1991, 1995). In this formulation, the temperatures are obtained as the time derivatives of the thermal displacements and the entropy is the conjugate momentum to the temperature. In (Mata and Lew 2011) a class of variational integrators for finite-dimensional adiabatic thermoelastic is formulated. The same ideas are applied in (Mata and Lew 2014) to develop thermodynamically consistent methods for finite element discretizations of deformable elastic solids with second sound. Unfortunately, it is not possible to construct a Hamiltonian formulation of thermoelasticity with heat conduction of Fourier type.[1] In (Mata and Lew 2012), an entropy flux term of Fourier type is added as a dissipative perturbation to the Hamiltonian form of the balance equations of adiabatic thermoelasticity and a discrete version of D'Alembert's principle is used to formulate structure-preserving methods. A similar approach has been followed in (Kern et al. 2014).

Other applications of variational integration can be reviewed, e.g., in (Kharevych et al. 2006; Kraus 2013; Ober-Blöbauma et al. 2013; Stern and Grinspun 2009).

1.4 Objectives and Layout of the Chapter

This chapter presents an introduction to the formulation of variational methods for finite-dimensional Lagrangian dynamical systems.

In Sect. 2 we revisit the Lagrangian and Hamiltonian points of view of mechanics for systems evolving on manifolds. To this end, we first introduce the concepts of generalized coordinates, configuration space, tangent space and the consistent computation of variations. Every concept is illustrated through some simple examples. Special attention is given to the construction of Lagrangian functions and to the role of Hamilton's variational principle in the deduction of the balance equations. In Sect. 2.3, we introduce the concept of (group) symmetries of the Lagrangian functional and we explain how they are related with the existence invariants of the dynamics. The symplectic nature of the flow is also discussed. Section 2.4 introduces the

[1] Alternative variational principles have been proposed for thermoelasticity with heat conduction, see e.g., (Yang et al. 2006; Vujanovic and Djukic 1971; Gambar and Markus 1994; Hutter and Tervoort 2007; Cannarozzi and Ubertini 2001). Structure-preserving methods may be consulted in (Armero and Simo 1992; Gross and Betsch 2006, 2007; Simo and Miehe 1992) to name only a few of them. Outstanding among the most recent approaches are the methods formulated by Romero (2009, 2010) which may be considered as energy-momentum methods applied to irreversible thermoelastic systems.

Legendre transform to compute the conjugated momenta along with the Hamiltonian form of the balance equations presented as a dynamical system evolving over the phase space.

The discrete version of Lagrangian mechanics is the main topic of Sect. 3. The standard methodology to construct variational integrators is explained in Sect. 3.1 and the position-momentum of the methods is given in Sect. 3.2. Section 3.3 is devoted to the implementation of the algorithms. The approximation properties and the geometric characteristics of the resulting discrete trajectories, including a discrete version of the symplecticity, are studied in Sect. 3.5.

Finally, in Sect. 4, we exemplify the usage of the variational methodology to construct structure-preserving methods for two problems of practical interest. First, we develop a VI based on the trapezoidal rule for a free-flying body that is able to undergo arbitrarily large rotations and displacements in space. The second example corresponds to the formulation of an explicit, second-order accurate variational integrator for the finite element discretization of geometrically exact rods.

This chapter is further complemented with the appendices A, B, and C.

2 Lagrangian and Hamiltonian Mechanics for Finite-Dimensional Systems

In the following, we will introduce (or review) the basic concepts in Lagrangian and Hamiltonian Mechanics. We shall make the abstractions concrete by applying them to two working examples, a particle in a hoop, and two particles joined by a rigid rod moving in a plane.

A particle in a hoop. Consider a particle of mass m that can slide without friction on a rigid circular hoop of radius R. The hoop is rigidly attached to an inertial frame, see Fig. 1.

Two particles joined by a rigid rod in a plane. Consider two particles of mass m joined by a rigid *massless* rod of length $2L$ which can freely move in $\mathbb{R}^2$.

As we shall see, these simple examples contain a lot of the concepts we discuss next.

2.1 Basic Concepts

For the first part of these notes we are going to consider mechanical systems for which all possible positions or *configurations* of the system can be identified with a finite-dimensional *c*onfiguration space or, more generally, *configuration manifold* Q (see also Sect. A). The configuration manifold Q is a datum of the mechanical system, or at most, a modeling assumption. Prototypical systems of this type are multibody

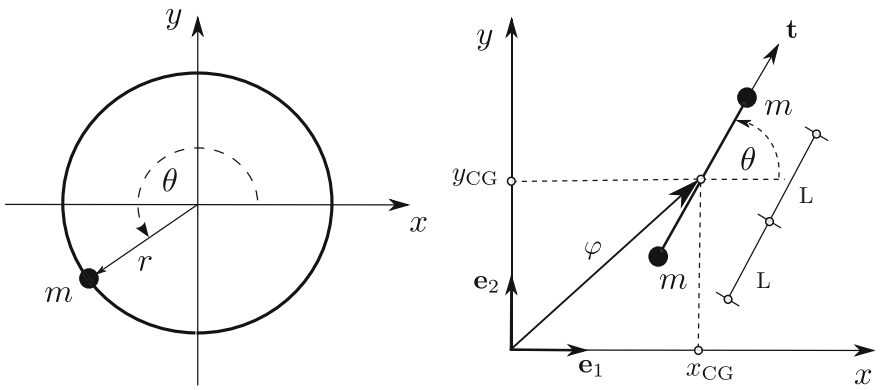


Fig. 1 Sketch of a particle in a hoop (*left*), and two particles joined by a rigid rod (*right*)

dynamical systems, involving a finite number of particles and rigid bodies moving together in $\mathbb{R}^n$, $n \in \mathbb{N}$.

To identify points in Q, we adopt a set of generalized coordinates

$$\mathbf{q} = (q_1, \ldots, q_d),$$

(lengths, angles, etc.), where $d \in \mathbb{N}$ is the dimension of Q. A *trajectory* of the mechanical system over the time interval $[0, T]$ is a map

$$\mathbf{q}(\cdot)\colon [0, T] \to Q.$$

In generalized coordinates, a trajectory is indicated with maps $q_i(t)$, $i = 1, \ldots, d$.

For mechanical systems consisting of m particles moving in some subset of $\mathbb{R}^k$, we can regard the configuration manifold as a subset of $\mathbb{R}^{k\times m}$. This means that we can define $k \times m$ maps $x_{\alpha,r}(q_1, \ldots, q_d)$, $r = 1, \ldots, k$, $\alpha = 1, \ldots, m$ that return the rth Cartesian coordinate of particle α for the given configuration of the system.

Particle in a hoop. $Q = S^1$, or the unit circle in $\mathbb{R}^2$. This is a one-dimensional manifold. The single generalized coordinate could be chosen as $q_1 = \theta$, the angle shown in Fig. 1. Notice that we could have chosen as a generalized coordinate any bijective and monotone function of θ, such as $q_1 = -\theta^3$. The choice of generalized coordinates is not unique. An example trajectory of the particle is $\theta(t) = \cos t$. For this system, we can define $x_{1,1}(\theta) = r\cos\theta$, $x_{1,2}(\theta) = r\sin\theta$, to recover the Cartesian coordinates of this particle in $\mathbb{R}^2$. Henceforth, we set $q_1 = \theta$ for this example.

Two particles joined by a rigid rod. $Q = \mathbb{R}^2 \times S^1$, which is a three-dimensional manifold, $d = 3$. A possible set of generalized coordinates is

$$(q_1, q_2, q_3) = (x_{CG}, y_{CG}, \theta),$$

where (x_{CG}, y_{CG}) are the Cartesian coordinates of the center of mass of the system and θ is the angle shown in Fig. 1. As before, other choices of generalized coordinates are possible, such as $(q_1, q_2, q_3) = (x_1, y_1, \theta)$, where (x_1, y_1) are the Cartesian coordinates of the position of one the masses. Henceforth, we choose the former set of generalized coordinates for this example. An example trajectory of the two particles is

$$(x_{CG}(t), y_{CG}(t), \theta(t)) = (t, 2t, t^2).$$

The Cartesian coordinates of this system in $\mathbb{R}^4$ are

$$\begin{aligned} x_{1,1}(x_{CG}, y_{CG}, \theta) &= x_{CG} + L\cos\theta \\ x_{1,2}(x_{CG}, y_{CG}, \theta) &= y_{CG} + L\sin\theta \\ x_{2,1}(x_{CG}, y_{CG}, \theta) &= x_{CG} - L\cos\theta \\ x_{2,2}(x_{CG}, y_{CG}, \theta) &= y_{CG} - L\sin\theta. \end{aligned}$$

□

Given a trajectory $\mathbf{q}(\cdot)$, the *generalized velocity* of the system at time t is $\dot{\mathbf{q}}(t)$. Given a point $\mathbf{q} \in Q$, the set of all possible generalized velocities of the system at $\mathbf{q}$ is called the *tangent space* $T_{\mathbf{q}}Q$, which is a vector space. The union of all points in Q with the tangent spaces attached is the *tangent bundle* TQ, and it is also a manifold. An element of the tangent bundle is, roughly speaking, a point $\mathbf{q} \in Q$ with a generalized velocity vector $\dot{\mathbf{q}}$ attached to it. Coordinates on TQ are denoted by

$$(\mathbf{q}, \dot{\mathbf{q}}) = (q_1, \ldots, q_d, \dot{q}_1, \ldots, \dot{q}_d).$$

For a system of m particles in $\mathbb{R}^k$, we can recover the traditional Cartesian components of the velocities as

$$v_{\alpha,r}(q_1, \ldots, q_d, \dot{q}_1, \ldots, \dot{q}_d) = \sum_{i=1}^{d} \dot{q}_i \frac{\partial x_{\alpha,r}}{\partial q_i}(q_1, \ldots, q_d).$$

Particle in a hoop. $TQ = S^1 \times \mathbb{R}$. A point in TQ has coordinates $(\theta, \dot{\theta})$. For example, for the trajectory $\theta(t) = \cos t$, and the coordinates of the point of TQ in which the system is at $t = 1$ is $(\cos 1, -\sin 1)$. The tangent space at $\theta = \cos 1$, $T_{\cos 1}Q$, is the line tangent to the circle at such point, with origin at $\theta = \cos 1$. This is the space to which possible velocities of the particle at that point belong, so velocities are always tangent to the hoop. Figure 2 shows the position vector of the particle in space is given by

$$r^i = r\left(\cos\theta^i e_1 + \sin\theta^i e_2\right),$$

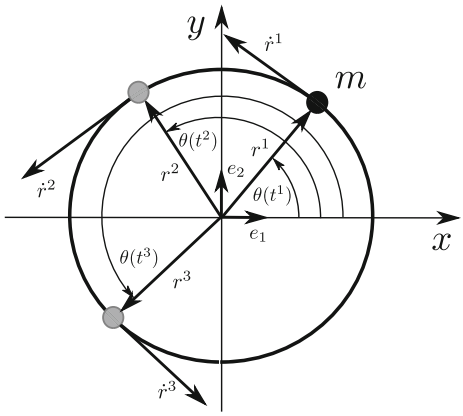


Fig. 2 Position vector, r^i, and velocity vector, $\dot{r}^i$, of the particle at times t^i $(i = 1, 2, 3)$

and the corresponding velocity vector by

$$\dot{r}^i = r\ \dot{\theta}^i \left(\cos\theta^i e_2 - \sin\theta^i e_1\right),$$

where θ^i and $\dot{\theta}^i$ denote the values of $\theta(t)$ and $\dot{\theta}(t)$ at the time t^i $(i = 1, 2, 3)$, respectively. Note that we may think the motion of the particle in a hoop as a two-dimensional motion restricted by the condition $r^i \cdot \dot{r}^i = 0$.

Two particles joined by a rigid rod. $TQ = (S^1 \times \mathbb{R}^2) \times (\mathbb{R} \times \mathbb{R}^2)$, which has coordinates $(x_{CG}, y_{CG}, \theta, \dot{x}_{CG}, \dot{y}_{CG}, \dot{\theta})$. For the trajectory $(t, 2t, t^2)$, the generalized velocities at time $t = 1$ are $(1, 2, 2)$, and the coordinates in TQ are $(1, 2, 1, 1, 2, 2)$. A graphical depiction of the tangent space is difficult here, because we should be thinking about the tangent space to the surface defined by the configuration manifold when embedded in $\mathbb{R}^4$. □

Given $\mathbf{q}(\cdot)\colon [0, T] \to Q$, we consider a one-parameter family of trajectories $\mathbf{q}^\epsilon(\cdot) : [0, T] \to Q$ such that $\mathbf{q}_0(\cdot) = \mathbf{q}(\cdot)$ for all $\epsilon \in (-\varepsilon, \varepsilon)$, for some $\varepsilon > 0$. A *variation* $\delta\mathbf{q}(\cdot)$ of $\mathbf{q}(\cdot)$ is defined as

$$\delta\mathbf{q}(t) = \frac{d}{d\epsilon}\mathbf{q}^\epsilon(t)\Big|_{\epsilon=0}. \tag{1}$$

Coordinates of a variation are $(\delta q_1, \ldots, \delta q_d)$. Different one-parameter families of trajectories generally give rise to different variations, but of course, multiple one-parameter families of trajectories give rise to the same variation. Clearly, $(\mathbf{q}(t), \delta\mathbf{q}(t)) \in T_{\mathbf{q}(t)}Q$ for each t. An intuitive graphical interpretation of a variation is shown in Fig. 3.

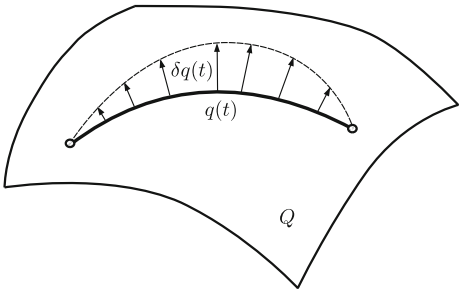


Fig. 3 Sketch of a trajectory $\mathbf{q}(\cdot)$ over Q, and a variation $\delta\mathbf{q}(\cdot)$. The variation at each time t, $\delta\mathbf{q}(t)$, is tangent to Q at $\mathbf{q}(t)$

For a system of m particles in $\mathbb{R}^k$, we can compute the Cartesian components of the variation as

$$\delta x_{\alpha,r} = \frac{d}{d\epsilon} x_{\alpha,r}(q_1^\epsilon, \ldots, q_d^\epsilon)\Big|_{\epsilon=0} = \sum_{i=1}^{d} \frac{\partial x_{\alpha,r}(q_1, \ldots, q_d)}{\partial q_i} \delta q_i.$$

Particle in a hoop. Consider the one-parameter family of trajectories $\theta^\epsilon(t) = \cos t + \epsilon\Delta\theta(t)$, for some $\Delta\theta : [0, T] \to \mathbb{R}$. Then, $\delta\theta(t) = \Delta\theta(t)$ at all times. The Cartesian components of the variation are

$$\delta x_{1,1}(t) = -r \sin\theta(t)\Delta\theta(t) \qquad \delta x_{1,2}(t) = r\cos\theta(t)\Delta\theta(t),$$

which clearly shows that $\delta q(t)$ is tangent to the circle at $q(t)$, for all times.

Another example of a variation follows by selecting $\theta^\epsilon(t) = \cos(t + \beta\epsilon)$, for $\beta \in \mathbb{R}$. In this case $\delta\theta(t) = -\beta \sin t$ at all times.

□

A *functional* is defined as a map from a set S to $\mathbb{R}$. Scalar-valued functions are functionals. More interesting functionals, however, are found when the set S contains functions, for example,

$$S[y(\cdot)] = \int_a^b y(t)\, dt \tag{2}$$

is a functional that takes values over the set S of integrable functions over $[a, b]$.

We are going to be interested in functionals that take values over sets of trajectories. The *variation of a functional* S at a trajectory $\mathbf{q}(\cdot)$ for a variation $\delta\mathbf{q}(\cdot)$ is defined as

$$\left\langle \delta S\left[\mathbf{q}(\cdot)\right], \delta\mathbf{q}\right\rangle = \frac{d}{d\epsilon} S\left[\mathbf{q}^\epsilon(\cdot)\right]\Big|_{\epsilon=0}, \tag{3}$$

where $\mathbf{q}^\epsilon(\cdot)$ is any of the one-parameter families that defines $\delta\mathbf{q}$. This is also called the Gâteaux derivative of S at $\mathbf{q}(\cdot)$ in direction $\delta\mathbf{q}(\cdot)$.

2.2 Lagrangian Mechanics

The starting point for Lagrangian mechanics is the definition of the *Lagrangian* $L : TQ \to \mathbb{R}$, $L(\mathbf{q}, \dot{\mathbf{q}})$, or in coordinates, $L(q_1, \ldots, q_d, \dot{q}_1, \ldots, \dot{q}_d)$. Notice that the Lagrangian returns a real number for each point of the tangent bundle TQ.

Different physical theories give rise to different Lagrangians. For mechanical systems, the Lagrangian has the general form $L = K - U$, where $K : TQ \to \mathbb{R}$ is the kinetic energy of the system, and $U : Q \to \mathbb{R}$ is the potential energy of the system.

Particle in a hoop. The Lagrangian for this system, in the presence of gravity in the negative-y direction is

$$L(\theta, \dot{\theta}) = \frac{m}{2} r^2 \dot{\theta}^2 - mgr \sin\theta. \tag{4}$$

Two particles joined by a rigid rod. In this case, in the absence of gravity,

$$L(x_{CG}, y_{CG}, \theta, \dot{x}_{CG}, \dot{y}_{CG}, \dot{\theta}) = m(\dot{x}_{CG}^2 + \dot{y}_{CG}^2 + L^2 \dot{\theta}^2). \tag{5}$$

General multibody systems. A general class of Lagrangians obtained in multibody dynamics has the form

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{q}} \cdot \mathbf{M}(\mathbf{q}) \dot{\mathbf{q}} - U(\mathbf{q}), \tag{6}$$

where for each $\mathbf{q} \in Q$, $\mathbf{M}(\mathbf{q})$ is the symmetric and positive-definite $d \times d$ *mass matrix* of the system at $\mathbf{q}$. The particle in the hoop and the two particles connected by a rigid rod are particular cases of this Lagrangian.

Thermoelastic systems. Finite-dimensional and adiabatic thermoelastic systems may be constructed following (Maugin and Kalpakides 2002; Romero 2009; Mata and Lew 2011, 2012, 2014). We consider N masses connected by M thermoelastic springs, such as those shown in Fig. 4. The spatial position of the masses at time t is described by $\mathbf{q}(t) = (q_1(t), .., q_d(t)) \in Q_S$, where Q_S is the a d-dimensional manifold. Additionally, each thermoelastic spring is assigned a time-dependent *thermal displacement* $\Phi^i(t) \in \mathbb{R}$, $i = 1, ..., M$, such that the empirical temperatures are computed as

$$\boldsymbol{\theta}(t) = \frac{d}{dt} \boldsymbol{\Phi}(t) = \left(\dot{\Phi}_1(t), ..., \dot{\Phi}_M(t) \right).$$

The configuration manifold for this system is $Q = Q_S \times \mathbb{R}^M$, and it is specified by points of the form $(\mathbf{q}, \boldsymbol{\Phi})$. Trajectories of the system are time-dependent functions $(\mathbf{q}(t), \boldsymbol{\Phi}(t))$, and the generalized velocities are $(\dot{\mathbf{q}}(t), \boldsymbol{\theta}(t))$; therefore, in this system the temperatures of the springs are generalized velocities. The thermoelastic behavior of each spring is described by a Helmholtz free-energy

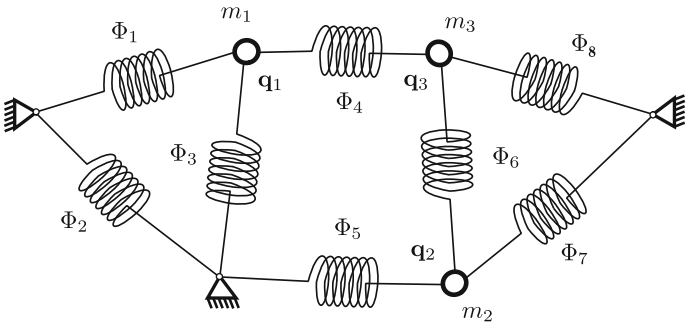


Fig. 4 A typical thermoelastic system: an assembly of three masses connected by eight thermoelastic springs and subjected to boundary conditions

function $\mathsf{A}_i(\mathbf{q}, \theta^i)$, $i \in \{1, ..., M\}$ so that the Helmholtz free energy of the system follows as

$$\mathsf{A}(\mathbf{q}, \boldsymbol{\theta}) = \sum_{i=1}^{M} \mathsf{A}_i(\mathbf{q}, \theta^i).$$

A Lagrangian for this system is constructed as

$$L(\mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\Phi}, \boldsymbol{\theta}) = \frac{1}{2}\dot{\mathbf{q}} \cdot \mathbf{M}(\mathbf{q})\dot{\mathbf{q}} - \mathsf{A}(\mathbf{q}, \boldsymbol{\theta}). \tag{7}$$

□

The second step in Lagrangian mechanics is the definition of the *action functional* over the time interval $[0, T]$

$$S\left[\mathbf{q}(\cdot)\right] = \int_0^T L\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)\, dt. \tag{8}$$

In Lagrangian mechanics the physical trajectories, namely, those that satisfy Newton's laws whenever the acceleration is well-defined, are obtained from a variational principle, *Hamilton's principle*. This principle states that: *The physical trajectory* $\mathbf{q}(\cdot)$ *is such that*

$$\left\langle S\left[\mathbf{q}(\cdot)\right], \delta\mathbf{q}\right\rangle = 0, \tag{9}$$

for all variations $\delta\mathbf{q}$ *that satisfy* $\delta\mathbf{q}(0) = \delta\mathbf{q}(T) = 0$. The set of all variations that satisfy these last conditions receive the name of *admissible variations*. Because the variation of S is the Gâteaux derivative of S, we also say that the action functional is stationary at the physical trajectory with respect to all admissible variations.

This variational principle completely characterizes trajectories of the system. It even gives meaning to physical trajectories when accelerations are not defined, and hence when Newton's second law cannot be applied. This is the case of impacts, for example. In principle, there would be no need to go any further to characterize trajectories. However, an alternative characterization of the stationary points of S is given by a system of ordinary differential equations that we shall find next. These are the equations of the trajectory of the system.

To find these equations, we proceed first by computing the variation of S for smooth enough $L(\mathbf{q}, \dot{\mathbf{q}})$, $\mathbf{q}(\cdot)$, $\delta\mathbf{q}(\cdot)$:

$$\begin{aligned}
\langle \delta S[\mathbf{q}(\cdot)], \delta\mathbf{q}\rangle &= \frac{d}{d\epsilon} S[\mathbf{q}^\epsilon(\cdot)]\Big|_{\epsilon=0} \\
&= \frac{d}{d\epsilon} \int_0^T L(\mathbf{q}^\epsilon(t), \dot{\mathbf{q}}^\epsilon(t))\, dt\Big|_{\epsilon=0} \\
&= \sum_{i=1}^{d} \int_0^T \frac{\partial L}{\partial q_i}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, \delta q_i(t) + \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, \delta \dot{q}_i(t)\, dt \\
&= \sum_{i=1}^{d} \int_0^T \left[\frac{\partial L}{\partial q_i}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \right) \right] \delta q_i(t)\, dt \\
&\qquad + \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, \delta q_i(t)\Big|_0^T . \qquad (10)
\end{aligned}$$

In Hamilton's principle $\delta\mathbf{q}(0) = \delta\mathbf{q}(T) = 0$, so the last term of the last expression is identically zero. Then, (9) implies that[2]

$$0 = \frac{\partial L}{\partial q_i}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \right) \qquad (11)$$

for all $t \in (0, T)$, and $i = 1, \ldots, d$. These are the Euler–Lagrange (E–L) equations of the system, and can be regarded as Newton's laws in terms of the chosen generalized coordinates. This is precisely part of the beauty of Lagrangian mechanics: The equations of motion are written as in (11) for *all* choices of generalized coordinates.

Particle in a hoop. The equations of motion are

$$0 = -mgr\cos\theta - \frac{d}{dt}\left(mr^2\dot{\theta}\right).$$

Two particles joined by a rigid rod. The equations of motion are:

$$m\frac{d}{dt}\dot{x}_{CG} = 0, \quad m\frac{d}{dt}\dot{y}_{CG} = 0 \quad \text{and} \quad mL^2\frac{d}{dt}\dot{\theta} = 0.$$

[2]The rigorous justification of this step requires the careful definition of the set of trajectories and corresponding variations, and then the use of some version of the fundamental lemma of the calculus of variations.

General multibody system. The equations of motion in this case take the form

$$0 = \frac{1}{2}\dot{\mathbf{q}} \cdot \frac{\partial \mathbf{M}}{\partial \mathbf{q}}(\mathbf{q})\dot{\mathbf{q}} - \frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}) - \frac{d}{dt}\left(\mathbf{M}(\mathbf{q})\dot{\mathbf{q}}\right).$$

Thermoelastic systems. In this case, the equations of motion are given by

$$\frac{d}{dt}\left(\mathbf{M}(\mathbf{q})\dot{\mathbf{q}}\right) - \frac{1}{2}\dot{\mathbf{q}} \cdot \frac{\partial \mathbf{M}}{\partial \mathbf{q}}(\mathbf{q})\dot{\mathbf{q}} = -\frac{\partial \mathsf{A}}{\partial \mathbf{q}}(\mathbf{q}, \boldsymbol{\theta}),$$
$$\frac{d}{dt}\left(-\frac{\partial \mathsf{A}}{\partial \boldsymbol{\theta}}(\mathbf{q}, \boldsymbol{\theta})\right) = 0.$$

To interpret these equations, it is useful to recall the thermodynamic relations (see, e.g., Coleman and Noll 1963),

$$\mathbf{f}(\mathbf{q}, \boldsymbol{\theta}) = -\frac{\partial \mathsf{A}}{\partial \mathbf{q}}(\mathbf{q}, \boldsymbol{\theta}) \quad \text{and} \quad \boldsymbol{\eta}(\mathbf{q}, \boldsymbol{\theta}) = -\frac{\partial \mathsf{A}}{\partial \boldsymbol{\theta}}(\mathbf{q}, \boldsymbol{\theta}), \tag{12}$$

where $\mathbf{f}$ is the vector of thermoelastic forces on the particles, and $\boldsymbol{\eta}$ is the vector of entropies, one entropy component per spring. Thus, the first equation states Newton's second law for the system in the case of a configuration-dependent mass matrix. The second equation states that the entropy of each one of the springs remains constant in time, as it should when no heat is transferred between springs. □

2.3 Conservation Properties: Lagrangian Point of View

A fundamental realization by E. Noether almost a century ago (Noether 1918) was that a (variation of a) motion that leaves the value of the Lagrangian invariant defines associated conserved quantities. Such motions are called symmetries of the Lagrangian, and we show some examples below. Typical examples of these conserved quantities are linear and angular momenta, and by adopting time as an independent coordinate as well, it is possible to regard energy as one such quantity as well (see, e.g., Kane et al. 1999). In summary, Noether's theorem states that to each symmetry of the Lagrangian corresponds a conserved quantity.

The simplest invariance or symmetry of the Lagrangian we may find is when for some i,

$$\frac{\partial L}{\partial q_i}(\mathbf{q}, \dot{\mathbf{q}}) = 0,$$

for all $(\mathbf{q}, \dot{\mathbf{q}}) \in TQ$. In this case a trajectory $\widehat{\mathbf{q}}$ that satisfies the E–L equations, satisfies that

$$\frac{\partial L}{\partial \dot{q}_i}(\widehat{\mathbf{q}}, \dot{\widehat{\mathbf{q}}})$$

is constant in time, and hence it is a conserved quantity for the motion $\widehat{\mathbf{q}}$. In such case, coordinate q_i is called a *cyclic coordinate*.

Of course, the definition of the symmetry is independent of the choice of coordinates. To keep the discussion at an intuitive level, it is useful to think about a symmetry of the Lagrangian in the following way: Given a trajectory $\mathbf{q}(t)$ and some $\varepsilon > 0$, we say that a one-parameter family of curves $\mathbf{q}^\epsilon(t)$ with $\mathbf{q}^0(t) = \mathbf{q}(t)$ is a symmetry of the Lagrangian if at all times t

$$L\left(\mathbf{q}^\epsilon(t), \dot{\mathbf{q}}^\epsilon(t)\right) = L\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right) \tag{13}$$

for any $\epsilon \in (-\varepsilon, \varepsilon)$. The variation of this symmetry,

$$\boldsymbol{\xi}(t) = (\delta q_1(t), ..., \delta q_d(t))$$

computed according to (1), is called an *infinitesimal symmetry direction*, see e.g., (Marsden and West 2001; Lew et al. 2004).

If follows from integrating (13) in time that

$$S[\mathbf{q}^\epsilon(\cdot)] = S[\mathbf{q}(\cdot)] \tag{14}$$

for all $\epsilon \in (-\varepsilon, \varepsilon)$. Clearly, the variation of S in the infinitesimal symmetry direction is equal to zero, a result of computing the derivative with respect to ϵ on both sides of (14). Then, for a trajectory $\mathbf{q}(t)$ that satisfies the Euler–Lagrange equations (11) the only nonzero terms in (10) are the boundary terms, i.e.,

$$\begin{aligned} 0 &= \frac{d}{d\epsilon}\left(\int_0^T L(\mathbf{q}^\epsilon(t), \dot{\mathbf{q}}^\epsilon(t) dt\right)\Bigg|_{\epsilon=0} \\ &= \sum_{i=1}^{d}\left(\frac{\partial L}{\partial \dot{q}_i}\big(\mathbf{q}(T), \dot{\mathbf{q}}(T)\big)\, \xi_i(T) - \frac{\partial L}{\partial \dot{q}_i}\big(\mathbf{q}(0), \dot{\mathbf{q}}(0)\big)\, \xi_i(0)\right). \end{aligned} \tag{15}$$

The above equations are a formal statement of Noether's theorem which show that the initial and final values of the *momentum*

$$\mathbf{p}(t) = \left(\frac{\partial L}{\partial \dot{q}_1}\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right), ..., \frac{\partial L}{\partial \dot{q}_d}\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)\right),$$

are equal in the $\boldsymbol{\xi}(t)$ direction,

$$\mathbf{p}(T) \cdot \boldsymbol{\xi}(T) = \mathbf{p}(0) \cdot \boldsymbol{\xi}(0). \tag{16}$$

The following examples illustrate the application of Noether's theorem.

Particle in a hoop. Consider the particle in the hoop in the *a*bsence of gravity and a one-parameter family of curves of the form $\theta^\epsilon(t) = \theta(t) + \epsilon\theta_0$ where θ_0 is a constant but otherwise arbitrary increment superposed onto $\theta(t)$. In words, we are considering trajectories identical to $\theta(t)$ that are simply rotated by a constant angle θ_0. A simple inspection reveals that the Lagrangian is invariant

$$L(\dot{\theta}^\epsilon) = L(\dot{\theta}),$$

and thus, according to Noether's theorem a conserved quantity exists. The infinitesimal symmetry direction in this case is

$$\delta\theta(t) = \theta_0,$$

and the momentum is

$$p_\theta(t) = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}, \tag{17}$$

which is the angular momentum of the particle with respect to the center of the hoop.
From (16), we conclude that

$$p_\theta(T) = mr^2\dot{\theta}(T)\theta_0 = mr^2\dot{\theta}(0)\theta_0 = p_\theta(0),$$

for any θ_0, and hence that

$$mr^2\dot{\theta}(T) = mr^2\dot{\theta}(0),$$

which shows that in the absence of external potentials breaking the symmetry of the Lagrangian, the angular momentum is a constant of the motion.

Two particles joined by a rigid rod. For this example we consider the one-parameter family of curves

$$\begin{aligned}\mathbf{z}^\epsilon(t) &= \left(x^\epsilon_{CG}(t), y^\epsilon_{CG}(t), \theta^\epsilon(t)\right)\\ &= \left(x_{CG}(t) + \epsilon\bar{x}, y_{CG}(t) + \epsilon\bar{y}, \theta(t) + \epsilon\bar{\theta}\right)\\ &= \mathbf{z}(t) + \epsilon\,\boldsymbol{\chi},\end{aligned}$$

where $\boldsymbol{\chi} = (\bar{x}, \bar{y}, \bar{\theta})$ is an arbitrary vector in $\mathbb{R}^3$. The first two components of $\boldsymbol{\chi}$, $\mathbf{c} = (\bar{x}, \bar{y})$, represent an imposed rigid body translation in space, and $\bar{\theta}$ represents an imposed rigid body rotation, see Fig. 5. This family of curves is a symmetry of the Lagrangian, since it is simple to check that

$$L(\mathbf{z}^\epsilon, \dot{\mathbf{z}}^\epsilon) = L(\mathbf{z}, \dot{\mathbf{z}}).$$

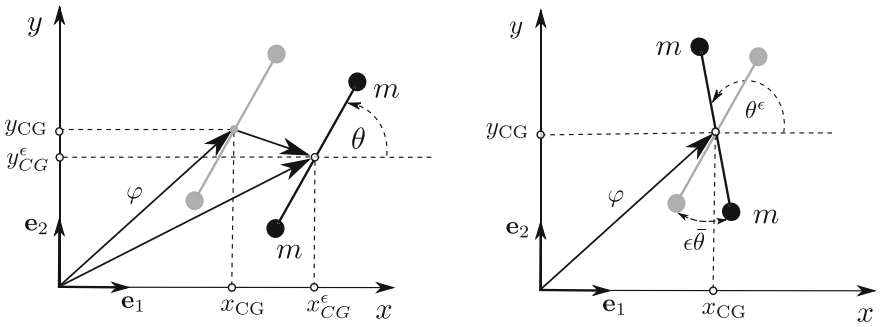


Fig. 5 Two particles joined by a rigid rod: rigid body translation (*left*), and rigid body rotation (*right*)

with infinitesimal symmetry direction

$$\delta \mathbf{z} = \frac{d}{d\epsilon} \mathbf{z}^{\epsilon}(t) \Big|_{\epsilon=0} = \boldsymbol{\chi}.$$

The momentum in this case is computed as

$$\mathbf{p}(t) = 2m \left(\dot{x}_{CG}(t),\ \dot{y}_{CG}(t),\ L^2 \dot{\theta}(t) \right).$$

It then follows from Noether's theorem that

$$\mathbf{p}(0) \cdot \boldsymbol{\chi} = \mathbf{p}(T) \cdot \boldsymbol{\chi}, \tag{18}$$

for any such $\boldsymbol{\chi}$. In particular, this implies that each component of $\mathbf{p}$ is conserved (choose $\boldsymbol{\chi} = (1, 0, 0)$, $\boldsymbol{\chi} = (0, 1, 0)$, and $\boldsymbol{\chi} = (0, 0, 1)$). Thus, the linear momentum of the system is constant in time

$$2m\dot{x}_{CG}(0) = 2m\dot{x}_{CG}(T), \qquad 2m\dot{y}_{CG}(0) = 2m\dot{y}_{CG}(T),$$

as it is the angular momentum of the system,

$$2mL^2\dot{\theta}(0) = 2mL^2\dot{\theta}(T).$$

Thermoelastic system. We consider the thermoelastic system described in Sect. 2.2 with constant mass matrix $\mathbf{M} = \mathrm{diag}(m, ..., m)$, $m \in \mathbb{R}$ and for which $Q_S \equiv \mathbb{R}^d$. For concreteness, we will let $\mathbf{q}_i$ denote the Cartesian coordinates of the ith particle in $\mathbb{R}^3$, and set $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_N) = (q_1, \ldots, q_d) \in \mathbb{R}^d$. Moreover, we assume that a Helmholtz energy function of the form

$$\mathsf{A}(\mathbf{q}, \boldsymbol{\theta}) = \sum_{i=1}^{M} \mathsf{A}_i \left(l_i(\mathbf{q}), \boldsymbol{\theta}\right),$$

where $l_i(\mathbf{q})$ denotes the distance between the two masses connected by the i^{th} thermoelastic spring.
We first consider a one-parameter family of curves $(\mathbf{q}^\epsilon, \boldsymbol{\Phi}^\epsilon)$ of the form

$$\begin{aligned} \mathbf{q}_i^\epsilon(t) &= \mathbf{q}_i(t) + \epsilon \mathbf{v} \qquad i = 1, \ldots, N \\ \boldsymbol{\Phi}^\epsilon(t) &= \boldsymbol{\Phi}(t) + \epsilon \mathbf{k}, \end{aligned}$$

where $\mathbf{v} \in \mathbb{R}^3$ and $\mathbf{k} \in \mathbb{R}^M$ are constant but otherwise arbitrary vectors and $\epsilon \in \mathbb{R}$. This family of curves results from applying arbitrary rigid body displacements onto both the mechanical and the thermal positions of the system. As we see below, this is a symmetry of the Lagrangian, with infinitesimal symmetry direction

$$\boldsymbol{\xi} = (\mathbf{v}_N, \mathbf{k}),$$

where $\mathbf{v}_N = (\underbrace{\mathbf{v}, \ldots, \mathbf{v}}_{\text{N times}}) \in \mathbb{R}^d$. To check that this is a symmetry of the Lagrangian (7), note that

$$\dot{\mathbf{q}}^\epsilon(t) = \dot{\mathbf{q}}(t) \quad \text{and} \quad \dot{\boldsymbol{\Phi}}^\epsilon(t) = \boldsymbol{\theta}^\epsilon(t) = \boldsymbol{\theta}(t),$$

and therefore the kinetic energy is invariant upon changing ϵ, i.e.,

$$\frac{1}{2}\dot{\mathbf{q}}^\epsilon \cdot \mathbf{M}\dot{\mathbf{q}}^\epsilon = \frac{1}{2}\dot{\mathbf{q}} \cdot \mathbf{M}\dot{\mathbf{q}}.$$

Additionally, considering that the distance between masses is conserved by rigid body translations in space, i.e., $l_i(\mathbf{q}^\epsilon) = l_i(\mathbf{q})$, and that the constant translations of the thermal displacements do not change the temperature, as stated above, we have that

$$\mathsf{A}\left(\mathbf{q}^\epsilon, \boldsymbol{\theta}^\epsilon\right) = \mathsf{A}\left(\mathbf{q}, \boldsymbol{\theta}\right),$$

from where it follows that the Lagrangian is invariant as well, and hence that $(\mathbf{q}^\epsilon, \boldsymbol{\Phi}^\epsilon)$ is one of its symmetries.
The momentum vector has components

$$\mathbf{p}(t) = \left(m\dot{\mathbf{q}}(t),\ \boldsymbol{\eta}\big(\mathbf{q}(t), \boldsymbol{\theta}(t)\big)\right). \tag{19}$$

It follows from Noether's theorem that

$$\mathbf{p}(T) \cdot \boldsymbol{\xi} = \mathbf{p}(0) \cdot \boldsymbol{\xi}, \tag{20}$$

for any $\boldsymbol{\xi} = (\mathbf{v}_N, \mathbf{k}) \in \mathbb{R}^d \times \mathbb{R}^M$. Equivalently,

$$\left(\sum_{i=1}^{N} m\dot{\mathbf{q}}_i(T)\right) \cdot \mathbf{v} = \left(\sum_{i=1}^{N} m\dot{\mathbf{q}}_i(0)\right) \cdot \mathbf{v}, \tag{21}$$

Thus, the linear momentum of the system is conserved (we can, for example, choose $\mathbf{v} = \boldsymbol{e}_i$ for $i = 1, 2, 3$ to conclude this, where $\{\boldsymbol{e}_i\}_i$ is a basis in $\mathbb{R}^3$), namely,

$$\sum_{i=1}^{N} m\dot{\mathbf{q}}_i(T) = \sum_{i=1}^{N} m\dot{\mathbf{q}}_i(0). \tag{22}$$

Similarly, we can conclude that the entropy of each spring $\eta_i = -\partial A_i/\partial\theta^i$ is conserved (this follows by choosing $\boldsymbol{\xi}^j = (\mathbf{0}, \mathbf{k}^j)$ for $j = 1, \ldots, M$, where $\mathbf{k}_i^j = \delta_i^j$), namely,

$$\eta_i(0) = \eta_i(T),$$

for $i = 1, \ldots, M$. This is precisely what is expected from a system without heat conduction, and it follows as a consequence of the symmetry of the Lagrangian upon rigid translations of the thermal displacements.

Next, we consider a second family of one-parameter curves, which involve rigid rotations of the mechanical displacements and leave the thermal ones unaltered. The family of curves is

$$\begin{aligned} \mathbf{q}_i^{\epsilon}(t) &= \exp(\epsilon\widetilde{\boldsymbol{\omega}})\mathbf{q}_i(t), \qquad i = 1, ..., N, \\ \boldsymbol{\Phi}^{\epsilon}(t) &= \boldsymbol{\Phi}(t), \end{aligned}$$

where

$$\widetilde{\boldsymbol{\omega}} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} = \text{skew}[\widehat{\boldsymbol{\omega}}],$$

is a skew-symmetric but otherwise constant tensor, $\widehat{\boldsymbol{\omega}} = (\omega_1, \omega_2, \omega_3)$ is the axial vector of $\widetilde{\boldsymbol{\omega}}$, which satisfies $\widetilde{\boldsymbol{\omega}}\mathbf{v} = \widehat{\boldsymbol{\omega}} \times \mathbf{v}$ for any $\mathbf{v} \in \mathbb{R}^3$, and $\exp[\cdot]$ is the tensor exponential operator; see Sect. B for a brief introduction to finite rotations.

Because this family of curves rigidly rotates the trajectory, it is simple to verify that the magnitude of the velocity of each particle does not change with ϵ, and neither does the distance between any two particles. Therefore, this family of curves is also a symmetry of the Lagrangian, with infinitesimal symmetry direction given by

$$\boldsymbol{\xi}_Q(t) = \left(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N, \mathbf{0}_M\right),$$

where $\mathbf{0}_M$ is an M-dimensional vector of zeroes and

$$\boldsymbol{\xi}_i(t) = \left.\frac{d\mathbf{q}_i^{\epsilon}}{d\epsilon}(t)\right|_{\epsilon=0} = \widetilde{\boldsymbol{\omega}}\,\mathbf{q}_i(t) = \widehat{\boldsymbol{\omega}} \times \mathbf{q}_i(t), \quad i = 1, ..., N.$$

With this new infinitesimal symmetry direction in Noether's theorem (20) using the momentum (19) we obtain that

$$\sum_{i=1}^{N} \left(\mathbf{q}^i(T) \times m\dot{\mathbf{q}}^i(T)\right) \cdot \widehat{\boldsymbol{\omega}} = \sum_{i=1}^{N} \left(\mathbf{q}^i(0) \times m\dot{\mathbf{q}}^i(0)\right) \cdot \widehat{\boldsymbol{\omega}}.$$

Again, since this holds for any $\widehat{\boldsymbol{\omega}} \in \mathbb{R}^3$, we can conclude that the angular momentum of the system

$$\mathbf{A}(t) := \sum_{i=1}^{N} \mathbf{q}^i(t) \times m\dot{\mathbf{q}}^i(t)$$

is conserved, namely,

$$\mathbf{A}(0) = \mathbf{A}(T).$$

□

2.3.1 Conservation of Energy

When the Lagrangian is a convex function of the generalized velocities,[3] such as when it is a quadratic function of $\dot{\mathbf{q}}$, the energy of the system is defined as

$$E(\mathbf{q}, \dot{\mathbf{q}}) = \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}). \tag{23}$$

This is the case for a general multibody system, whose Lagrangian is (6). In this case, the energy takes the form

$$E(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\dot{\mathbf{q}} \cdot \mathbf{M}(\mathbf{q})\dot{\mathbf{q}} + U(\mathbf{q}). \tag{24}$$

The energy of the system is conserved along its solution trajectory. This can easily be seen by computing its time derivative and using the Euler–Lagrange equations. However, this result can also be obtained if we note that for an autonomous Lagrangian the following relation hods

[3] And as assumed here, the Lagrangian does not depend explicitly on time. Such Lagrangian is said to be autonomous.

$$\int_0^T L\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)\, dt = \int_{0+\epsilon}^{T+\epsilon} L\left(\mathbf{q}(s-\epsilon), \dot{\mathbf{q}}(s-\epsilon)\right)\, ds = I(\epsilon)$$

for any $\epsilon \in \mathbb{R}$. Equivalently, since the Lagrangian does not depend explicitly of time, this statement says that a trajectory can be translated uniformly in time without changing the value of the action. This is called a time-translation symmetry, and there is a way to frame this symmetry in the context of Noether's theorem, which we shall not pursue here (see, e.g., Marsden and West 2001). Differentiating this last expression, we get

$$\begin{aligned} 0 &= \left.\frac{dI(\epsilon)}{d\epsilon}\right|_{\epsilon=0} \\ &= \left[\int_{0+\epsilon}^{T+\epsilon} \frac{d}{d\epsilon} L\left(\mathbf{q}(s-\epsilon), \dot{\mathbf{q}}(s-\epsilon)\right) ds\right]_{\epsilon=0} + L\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)\Big|_0^T \\ &= -\int_0^T \frac{\partial L}{\partial \dot{\mathbf{q}}} \ddot{\mathbf{q}}\, dt - \int_0^T \frac{\partial L}{\partial \mathbf{q}} \dot{\mathbf{q}}\, dt + L(\mathbf{q}, \dot{\mathbf{q}})\Big|_0^T. \end{aligned}$$

Integrating by parts the first term of the right-hand side

$$\int_0^T \frac{\partial L}{\partial \dot{\mathbf{q}}} \ddot{\mathbf{q}}\, dt = \left(\frac{\partial L}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}}\right)\Big|_0^T - \int_0^T \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) \dot{\mathbf{q}}\, dt,$$

and replacing in the above equation yields

$$\left.\frac{dI(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \int_0^T \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial L}{\partial \mathbf{q}}\right) \dot{\mathbf{q}}\, dt + \left[L(\mathbf{q}, \dot{\mathbf{q}}) - \frac{\partial L}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}}\right]_0^T = 0.$$

The term under the integral vanish identically since $\mathbf{q}(t)$ is a solution trajectory of the system. The remaining boundary term is precisely a statement of the conservation of the energy, i.e.,

$$E(T) = E(0).$$

Therefore, the energy of mechanical systems described by autonomous Lagrangians is an invariant of the dynamics, and it is a result of the invariance of the action upon time-translating a trajectory in time.

2.3.2 Symplecticity: Lagrangian Point of View

Lagrangian mechanical systems also have another important conservation property: they conserve a skew-symmetric bilinear form known[4] as the *symplectic Lagrangian*

[4]We recall that a bilinear form on a vector space V is a mapping $\boldsymbol{w} : V \times V \to \mathbb{R}$ that is linear in both arguments. It is skew-symmetric if $\boldsymbol{w}(\boldsymbol{u}, \boldsymbol{v}) = -\boldsymbol{w}(\boldsymbol{v}, \boldsymbol{u})$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$.

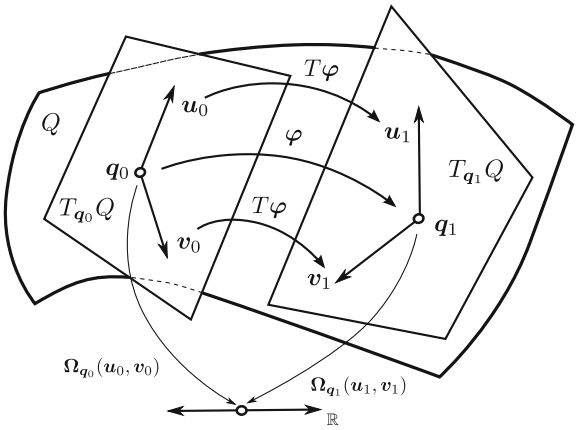


Fig. 6 Consider two points $\boldsymbol{q}_0, \boldsymbol{q}_1$ belonging to the configuration manifold Q such that $\boldsymbol{q}_1 = \varphi(\boldsymbol{q}_0)$. The tangent map $T\varphi$ relates the vectors $\boldsymbol{u}_0, \boldsymbol{v}_0 \in T_{\boldsymbol{q}_0}Q$ with $\boldsymbol{u}_1, \boldsymbol{v}_1 \in T_{\boldsymbol{q}_1}Q$ according to (25). The two-form $\boldsymbol{\Omega} : TQ \times TQ \to \mathbb{R}$ is conserved by φ if $\boldsymbol{\Omega}_{\boldsymbol{q}_0}(\boldsymbol{u}_0, \boldsymbol{v}_0) = \boldsymbol{\Omega}_{\boldsymbol{q}_1}(\boldsymbol{u}_1, \boldsymbol{v}_1)$

form along solution trajectories (Marsden and Ratiu 1999). In contrast to the conservation of energy or Noether's theorem, which are properties associated to individual trajectories, the symplectic form is associated to the behavior of nearby trajectories, namely, trajectories with very close initial conditions. We will discuss the closest to an "intuitive" explanation of the symplectic form we are aware of in a later section. We also refer the reader to Leimkuhler and Reich (2005), which contains a very approachable (the geometric concepts are progressively introduced) introduction and discussion on this topic.

Symplectic Map and Symplectic Form. We start by considering a N-dimensional configuration manifold Q with $N \in \mathbb{N}$ being an even number and a smooth enough and one-to-one mapping $\varphi : Q \to Q$. The image through the mapping φ of an arbitrary point $\boldsymbol{q}_0 \in Q$ is given by $\boldsymbol{q}_1 = \varphi(\boldsymbol{q}_0)$ which also belongs to Q. As explained in Sect. 2.1, it is possible to attach a tangent space to every point in a configuration manifold. Therefore, we construct the tangent spaces $T_{\boldsymbol{q}_0}Q$ and $T_{\boldsymbol{q}_1}Q$. See Fig. 6.

Additionally, we define *tangent map of* φ, denoted by $T\varphi$, as a mapping between elements of the tangent spaces according to

$$\begin{aligned} T\varphi : \quad TQ &\to TQ \\ (\boldsymbol{q}_0, \boldsymbol{u}_0) &\mapsto (\boldsymbol{q}_1, \boldsymbol{u}_1), \end{aligned}$$

where $\boldsymbol{q}_1 = \varphi(\boldsymbol{q}_0)$ and $\boldsymbol{u}_1 \in T_{\boldsymbol{q}_1}Q$ is obtained as

$$\boldsymbol{u}_1 := T\varphi \cdot \boldsymbol{u}_0 = \left. \frac{d\varphi}{ds}(\boldsymbol{c}(s)) \right|_{s=0}, \tag{25}$$

and $\boldsymbol{c}(s)$ is a s-parametrized curve on Q such that $\boldsymbol{c}'(0) = \boldsymbol{u}_0$. As it is schematically depicted in Fig. 6, $T\varphi$ maps elements in the tangent space of $\boldsymbol{q}_0$ to elements in the

tangent space of $\boldsymbol{q}_1$. In terms of components, we have that

$$u_1^a = \sum_{j=1}^{N} \frac{\partial \varphi^a(\boldsymbol{q}_0)}{\partial q^j} u_0^j, \qquad a = 1, ..., N.$$

Further details can be found in (Marsden and Hughes 1983).

Furthermore, assume that we are given with a nondegenerated and skew-symmetric bilinear form defined for every $\boldsymbol{q} \in Q$ according to

$$\begin{aligned} \boldsymbol{\Omega}_{\boldsymbol{q}} : T_{\boldsymbol{q}} Q \times T_{\boldsymbol{q}} Q &\to \mathbb{R} \\ (\boldsymbol{u}, \boldsymbol{v}) &\mapsto \Omega_{\boldsymbol{q}}(\boldsymbol{u}, \boldsymbol{v}). \end{aligned}$$

The explicit form of $\boldsymbol{\Omega}_{\boldsymbol{q}}$ is problem depend. Some examples are given in e.g., (Marsden and Ratiu 1999, Chap. 2). The subscript in $\boldsymbol{\Omega}_{\boldsymbol{q}}$ highlights the dependency of the two-form on the point of the manifold.

The map φ is called *symplectic* or *canonical* is it preserves $\boldsymbol{\Omega}_{\boldsymbol{q}}$ in the following sense:

$$\boldsymbol{\Omega}_{\boldsymbol{q}_0}(\boldsymbol{u}_0, \boldsymbol{v}_0) = \boldsymbol{\Omega}_{\varphi(\boldsymbol{q}_0)} \left(T\varphi \cdot \boldsymbol{u}_0, T\varphi \cdot \boldsymbol{v}_0 \right),$$

where $\boldsymbol{u}_0, \boldsymbol{v}_0 \in T_{\boldsymbol{q}_0} Q$. Note that the above expression equals to zero, unless N is an even number. If this is the case, $\boldsymbol{\Omega}_{\boldsymbol{q}}$ is called a *symplectic two-form*. See Fig. 6 for a schematic representation.

In following, we explain how the solution trajectory of a Lagrangian mechanical system implicitly defines a time-dependent symplectic transformation over the tangent space of the configuration manifold, TQ. Furthermore, we show that the Lagrangian function of the system allows to construct the matrix representation of the corresponding symplectic two-form.

Lagrangian symplecticity. Consider a mechanical system characterized by a Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}})$ evolving on a configuration manifold Q. We construct a solution trajectory over TQ by means considering the following smooth enough and one-to-one mapping,

$$\begin{aligned} \psi : TQ \times [0, T] &\to \quad TQ \\ (z, t) &\mapsto \psi(z, t) \end{aligned} \tag{26}$$

where $z = (\mathbf{q}, \dot{\mathbf{q}})$ is an arbitrary point in TQ. The mapping ψ is such that $\psi(z, 0) = z$. Therefore, given certain initial conditions $z_0 \in TQ$, we define the trajectory of the mechanical system as

$$z(t) := \psi(z_0, t),$$

for all $t \in [0, T]$. Note that TQ posses the structure of a (even-dimensional) smooth manifold and therefore, it is possible to attach a tangent space to every $z \in TQ$. See Sect. A.1 and Fig. 7.

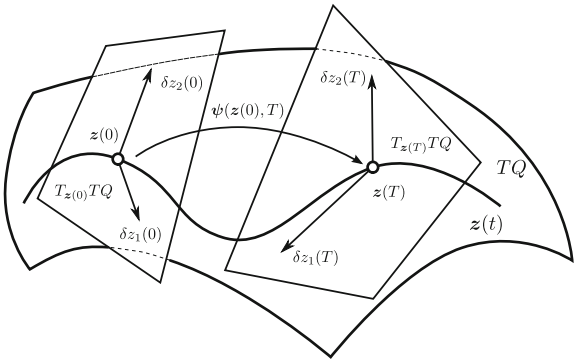


Fig. 7 Consider two points $z(0)$, $z(T)$ over the solution trajectory in the tangent space TQ. These points are related by ψ according to $z(T) = \psi(z(0), T)$. The variations $\delta z_i(0)$ and $\delta z_i(T)$, $i = 1, 2$, belong to the tangent spaces $T_{z(0)}TQ$ and $T_{z(T)}TQ$, respectively

A remarkable property of Lagrangian systems is given by the fact that $\psi(\cdot, t)$ results to be symplectic or in other words, there exist a skew-symmetric bilinear form, $\mathbf{\Omega}_L$, defined for every point over the solution trajectory $z(t)$ which enjoys the following conservation property,

$$\mathbf{\Omega}_L\big(z(0)\big)\Big(\delta z_1(0), \delta z_2(0)\Big) = \mathbf{\Omega}_L\big(z(T)\big)\Big(\delta z_1(T), \delta z_2(T)\Big), \tag{27a}$$

for arbitrary variations

$$\delta z_i(t) = \begin{bmatrix} \delta \mathbf{q}_i(t) \\ \delta \dot{\mathbf{q}}_i(t) \end{bmatrix} \in T_{z(t)}TQ, \quad i = 1, 2, \tag{27b}$$

such that

$$\delta z_i(T) = T\psi\big(z(0), T\big) \cdot \delta z_i(0), \qquad i = 1, 2. \tag{27c}$$

Moreover, the Lagrangian symplectic two-form may be represented in matrix form by

$$\mathbf{\Omega}_L(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} 0 & A_{12} & \dots & A_{1N} & B_{11} & \dots & B_{1N} \\ -A_{12} & 0 & \dots & A_{2N} & B_{21} & \dots & B_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -A_{1N} & -A_{2N} & \dots & 0 & B_{N1} & \dots & B_{NN} \\ -B_{11} & -B_{12} & \dots & -B_{1N} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -B_{N1} & -B_{N2} & \dots & -B_{NN} & 0 & \dots & 0 \end{bmatrix}, \tag{27d}$$

where

$$A_{ij}(\mathbf{q},\dot{\mathbf{q}}) = -A_{ji}(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}\left(\frac{\partial^2 L}{\partial q_j \partial \dot{q}_i}(\mathbf{q},\dot{\mathbf{q}}) - \frac{\partial^2 L}{\partial q_i \partial \dot{q}_j}(\mathbf{q},\dot{\mathbf{q}})\right)$$
$$B_{ij}(\mathbf{q},\dot{\mathbf{q}}) = \frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_i}(\mathbf{q},\dot{\mathbf{q}}). \tag{27e}$$

To show how the above result can be deduced, we follow the procedure presented in (Marsden and West 2001; Lew et al. 2004). Consider a two-parameter family of solution trajectories $(\mathbf{q}^{\epsilon,\nu}(t), \boldsymbol{v}^{\epsilon,\nu}(t)) \in TQ$ with $\epsilon, \nu \in \mathbb{R}$ and compute the following variations

$$\delta\mathbf{q}_1^{\epsilon}(t) = \left.\frac{\partial \mathbf{q}^{\epsilon,\nu}}{\partial \nu}(t)\right|_{\nu=0} \tag{28a}$$
$$\delta\mathbf{q}_2^{\nu}(t) = \left.\frac{\partial \mathbf{q}^{\epsilon,\nu}}{\partial \epsilon}(t)\right|_{\epsilon=0} \tag{28b}$$
$$\delta^2\mathbf{q}(t) = \left.\frac{\partial^2 \mathbf{q}^{\epsilon,\nu}}{\partial \epsilon \partial \nu}(t)\right|_{\epsilon,\nu=0}, \tag{28c}$$

and the same applies for $\delta\boldsymbol{v}_1^{\epsilon}(t)$, $\delta\boldsymbol{v}_2^{\nu}(t)$, and $\delta^2\boldsymbol{v}(t)$. See Fig. 8. Moreover, we write

$$\delta\mathbf{q}_1(t) = \delta\mathbf{q}_1^0(t), \quad \delta\mathbf{q}_2(t) = \delta\mathbf{q}_2^0(t) \quad \text{and} \quad \mathbf{q}^{\epsilon}(t) = \mathbf{q}^{\epsilon,0}(t). \tag{28d}$$

Example. We build the two-parameters family of curves

$$\mathbf{q}^{\epsilon,\nu}(t) = \exp\left[\epsilon\widetilde{\boldsymbol{\omega}}\right](\mathbf{q}(t) + \nu\,\mathbf{k}),$$

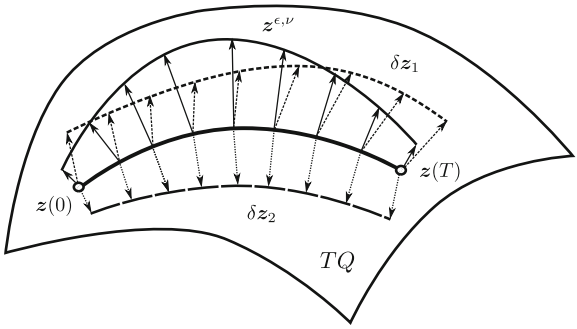


Fig. 8 Two-parameter family of solution trajectories $z(t)^{\epsilon,\nu} = (\boldsymbol{q}(t)^{\epsilon,\nu}, \dot{\mathbf{q}}(t)^{\epsilon,\nu}) \in TQ$. The corresponding variations $\delta z_1(t)$ and $\delta z_2(t)$ belong to the tangent space $T_{z(t)}TQ$

where $\widetilde{\boldsymbol{\omega}}$ is a constant and skew-symmetric tensor and $\mathbf{k} \in \mathbb{R}^3$. Then, we have

$$\begin{aligned}
\delta\mathbf{q}_1^{\epsilon}(t) &= \exp\left[\epsilon\widetilde{\boldsymbol{\omega}}\right]\mathbf{k}, \\
\delta\mathbf{q}_2^{\nu}(t) &= \widetilde{\boldsymbol{\omega}}\left(\mathbf{q}(t) + \nu\mathbf{k}\right), \\
\delta\mathbf{q}_1^{0}(t) &= \mathbf{k}, \\
\delta\mathbf{q}_2^{0}(t) &= \widetilde{\boldsymbol{\omega}}\mathbf{q}(t), \\
\delta^2\mathbf{q}(t) &= \left.\frac{d\delta\mathbf{q}_1^{\epsilon}}{d\epsilon}(t)\right|_{\epsilon=0} = \left.\frac{d\delta\mathbf{q}_2^{\nu}}{d\nu}(t)\right|_{\nu=0} = \widetilde{\boldsymbol{\omega}}\mathbf{k}.
\end{aligned}$$

□

Since we assumed that $\mathbf{q}^{\epsilon,\nu}(t)$ are solutions of the Euler–Lagrange equations, then

$$\left.\frac{\partial}{\partial\nu}\right|_{\nu=0} S\left[\mathbf{q}^{\epsilon,\nu}\right] = \sum_{i=1}^{N} \frac{\partial L}{\partial\dot{q}_i}(\mathbf{q}^{\epsilon,0}(t), \dot{q}^{\epsilon,0}(t))\ \delta\mathbf{q}_{1i}^{\epsilon}(t)\Big|_0^T \tag{29}$$

for any ϵ. We obtain the second variation of the action in the direction of these variations as (omitting arguments of functions for simplicity)

$$\begin{aligned}
\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=0}\left.\frac{\partial}{\partial\nu}\right|_{\nu=0} S\left[\mathbf{q}^{\epsilon,\nu}\right] &= \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=0}\left(\sum_{i=1}^{N}\frac{\partial L}{\partial\dot{q}_i}\,\delta\mathbf{q}_{1i}^{\epsilon}\Big|_0^T\right) \\
&= \sum_{i,j=1}^{N}\left[\frac{\partial^2 L}{\partial q_j\partial\dot{q}_i}\delta q_{1i}\delta q_{2j}\Big|_0^T + \frac{\partial^2 L}{\partial\dot{q}_j\partial\dot{q}_i}\delta q_{1i}\delta\dot{q}_{2j}\Big|_0^T\right] \\
&\quad + \sum_{i=1}^{N}\frac{\partial L}{\partial\dot{q}_i}\delta^2 q_i\Big|_0^T.
\end{aligned}$$

An equivalent expression is obtained by reversing the order of differentiation with respect to ϵ and ν since mixed partial derivatives are equal. Subtracting both expression we obtain

$$0 = \sum_{i,j=1}^{N}\left[\frac{\partial^2 L}{\partial q_j\partial\dot{q}_i}\left(\delta q_{1i}\delta q_{2j} - \delta q_{2i}\delta q_{1j}\right)\Big|_0^T + \frac{\partial^2 L}{\partial\dot{q}_j\partial\dot{q}_i}\left(\delta q_{1i}\delta\dot{q}_{2j} - \delta q_{2i}\delta\dot{q}_{1j}\right)\Big|_0^T\right]. \tag{30}$$

This identity can be rewritten to obtain an expression equivalent to (27a) with the $(2N) \times (2N)$ Lagrangian symplectic matrix given by (27d) and (27e), i.e.,

$$\left[\,\delta\mathbf{q}_1(0)\;\,\delta\dot{\mathbf{q}}_1(0)\,\right]^T\cdot\boldsymbol{\Omega}_L(\mathbf{q}(0),\dot{\mathbf{q}}(0))\cdot\begin{bmatrix}\delta\mathbf{q}_2(0)\\ \delta\dot{\mathbf{q}}_2(0)\end{bmatrix}=$$
$$\left[\,\delta\mathbf{q}_1(T)\;\,\delta\dot{\mathbf{q}}_1(T)\,\right]^T\cdot\boldsymbol{\Omega}_L(\mathbf{q}(T),\dot{\mathbf{q}}(T))\cdot\begin{bmatrix}\delta\mathbf{q}_2(T)\\ \delta\dot{\mathbf{q}}_2(T)\end{bmatrix}. \tag{31}$$

Therefore, we say that the Lagrangian symplectic two-form is exactly conserved along solution trajectories over TQ.

Particle in a hoop. Considering the Lagrangian function (4) we have that

$$\frac{\partial^2 L}{\partial\theta\partial\dot{\theta}}=0\quad\text{and}\quad\frac{\partial^2 L}{\partial\dot{\theta}^2}=mr^2,$$

and thus, the matrix representing the symplectic two-form is

$$\boldsymbol{\Omega}_L\big(\theta(t),\dot{\theta}(t)\big)=\begin{bmatrix}0 & mr^2\\ -mr^2 & 0\end{bmatrix}.$$

Consider, for example, the following bi-parametric family of solution trajectories

$$\theta^{\epsilon,\nu}(t)=e^{\epsilon}\theta(t)+\nu t$$
$$\dot{\theta}^{\epsilon,\nu}(t)=\frac{d\theta^{\epsilon,\nu}(t)}{dt}=e^{\epsilon}\dot{\theta}(t)+\nu,$$

where $\epsilon,\nu\in\mathbb{R}$. Then,

$$\delta\theta_1(t)=t,\quad\delta\theta_2(t)=\theta(t),\quad\delta\dot{\theta}_1(t)=1\quad\text{and}\quad\delta\dot{\theta}_2(t)=\dot{\theta}(t).$$

Notice that $(\delta\theta_i(t),\delta\dot{\theta}_i(t))\in T_{\theta(t),\dot{\theta}(t)}TS^1$, $i=1,2$. The conservation of the symplectic two-form along solution trajectories (31) implies that

$$\begin{bmatrix}\delta\theta_1(0)\\ \delta\dot{\theta}_1(0)\end{bmatrix}\cdot\boldsymbol{\Omega}_L\begin{bmatrix}\delta\theta_2(0)\\ \delta\dot{\theta}_2(0)\end{bmatrix}=\begin{bmatrix}\delta\theta_1(T)\\ \delta\dot{\theta}_1(T)\end{bmatrix}\cdot\boldsymbol{\Omega}_L\begin{bmatrix}\delta\theta_2(T)\\ \delta\dot{\theta}_2(T)\end{bmatrix},$$

or equivalently,

$$\begin{bmatrix}0\\1\end{bmatrix}\cdot\begin{bmatrix}0 & mr^2\\ -mr^2 & 0\end{bmatrix}\begin{bmatrix}\theta(0)\\ \dot{\theta}(0)\end{bmatrix}=\begin{bmatrix}T\\1\end{bmatrix}\cdot\begin{bmatrix}0 & mr^2\\ -mr^2 & 0\end{bmatrix}\begin{bmatrix}\theta(T)\\ \dot{\theta}(T)\end{bmatrix}.$$

Two particles joined by a rigid rod. We note that the only nonzero coefficients of the Lagrangian symplectic two-form are

$$B_{11}=\frac{\partial^2 L}{\partial\dot{x}_{CG}^2}=2m,\quad B_{22}=\frac{\partial^2 L}{\partial\dot{y}_{CG}^2}=2m,\quad B_{33}=\frac{\partial^2 L}{\partial\dot{\theta}^2}=2mL^2,$$

and therefore, the symplectic two-form is represented by

$$\boldsymbol{\Omega}_L\big(z(t)\big) = \begin{bmatrix} \mathbf{0}_{3\times 3} & \mathbf{B} \\ -\mathbf{B} & \mathbf{0}_{3\times 3} \end{bmatrix},$$

where $z(t) = (x_{CG}(t), y_{CG}(t), \theta(t), \dot{x}_{CG}(t), \dot{y}_{CG}(t), \dot{\theta}(t)) \in TQ$, $\mathbf{0}_{3\times 3}$ is a 3×3 matrix with of zeros and

$$\mathbf{B} = 2m \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & L^2 \end{bmatrix}.$$

Consider two admissible variations $\delta z_i(t) \in T_{z(t)}TQ, i = 1, 2$, computed according to (28a)–(28d). The conservation of the symplectic two-form along solution trajectories ensures that

$$\delta z_2(t) \cdot \boldsymbol{\Omega}_L\big(\mathbf{z}(t)\big)\delta z_1(t),$$

remains constant for all $t \in [0, T]$.

2.4 Hamiltonian Mechanics

Hamiltonian Mechanics reformulates Lagrange's equations of motion in generalized coordinates in a way that presents the motion of the system as a *flow* over *phase space*, as we explain next.

Given a Lagrangian L, the *conjugate momenta* are defined as

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}), \tag{32}$$

or in coordinates,

$$p_i = \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) \qquad i = 1, \ldots, d. \tag{33}$$

This defines a map $(\mathbf{q}, \dot{\mathbf{q}}) \mapsto (\mathbf{q}, \mathbf{p})$, or $(\mathbf{q}, \mathbf{p}) = \mathrm{FL}(\mathbf{q}, \dot{\mathbf{q}})$. This map is termed the *Legendre transform*. The space of all possible values of $(\mathbf{q}, \mathbf{p})$ is called *phase space* T^*Q. To be precise, T^*Q is the cotangent bundle, also a manifold, which informally speaking is defined by attaching to each point $q \in Q$ the dual space T_q^*Q to T_qQ.

For typical mechanical systems $\mathrm{FL} : TQ \to T^*Q$ is bijective (and hence invertible) and onto, because the Lagrangian is strictly convex in $\dot{\mathbf{q}}$ for each $\mathbf{q}$. This means that for these systems all possible values of the conjugate momenta are attained at each point $\mathbf{q} \in Q$, and hence $\mathrm{FL}^{-1}(\mathbf{q}, \mathbf{p})$ is defined for all $\mathbf{p} \in \mathbb{R}^d$.

The *Hamiltonian* $H\colon T^*Q \to \mathbb{R}$ is defined as

$$H(\mathbf{p}, \mathbf{q}) = \mathbf{p} \cdot \dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}), \tag{34}$$

where $\dot{\mathbf{q}} = FL^{-1}(\mathbf{q}, \mathbf{p})$.

Lagrange's equations of motion can then be written in terms of $(\mathbf{q}, \mathbf{p})$ as (in coordinates)

$$\dot{q}_i = \frac{\partial H}{\partial p_i}(\mathbf{q}, \mathbf{p}) \tag{35a}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}(\mathbf{q}, \mathbf{p}). \tag{35b}$$

These are *Hamilton's equations of motion*, and are, again, valid for any choice of generalized coordinates. These equations follow easily from using (34), namely,

$$\frac{\partial H}{\partial p_j} = \dot{q}_j + \sum_{i=1}^{d} \left[p_i \frac{\partial \dot{q}_i}{\partial p_j} - \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial p_j} \right] = \dot{q}_j$$

$$\frac{\partial H}{\partial q_j} = \sum_{i=1}^{d} \left[p_i \frac{\partial \dot{q}_i}{\partial q_j} - \frac{\partial L}{\partial q_j} - \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial q_j} \right] = -\frac{\partial L}{\partial q_j}$$

and replacing in (11).

Particle in a hoop. The Legendre transform is

$$p_\theta = \mathrm{FL}(\theta, \dot{\theta}) = mr^2\dot{\theta}$$

which is precisely the angular momentum of the particle around the origin. By solving the above equation for $\dot{\theta}$, and replacing in (34), we obtain the Hamiltonian

$$H(\theta, p_\theta) = \frac{p_\theta^2}{2mr^2} + mgr \sin\theta.$$

Hamilton's equations of motion are

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2}$$

$$\dot{p}_\theta = -\frac{\partial H}{\partial \theta} = -mgr \cos\theta.$$

General multibody system. The Legendre transform in this case is

$$\boldsymbol{\pi} = \mathbf{M}(\mathbf{q})\dot{\mathbf{q}}. \tag{36}$$

Since $\mathbf{M}(\mathbf{q})$ is positive definite for any $\mathbf{q}$, it is invertible for any $\mathbf{q}$. Hence,

$$\dot{\mathbf{q}} = \mathbf{M}^{-1}(\mathbf{q})\boldsymbol{\pi}. \tag{37}$$

The Hamiltonian then follows as

$$H(\mathbf{q}, \boldsymbol{\pi}) = \frac{1}{2}\boldsymbol{\pi} \cdot \mathbf{M}^{-1}(\mathbf{q})\boldsymbol{\pi} + U(\mathbf{q}), \tag{38}$$

and Hamilton's equations of motion as

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{M}^{-1}(\mathbf{q})\boldsymbol{\pi}, \\ \dot{\boldsymbol{\pi}} &= -\frac{1}{2}\boldsymbol{\pi} \cdot \frac{\partial \mathbf{M}^{-1}}{\partial \mathbf{q}}(\mathbf{q})\boldsymbol{\pi} - \frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}). \end{aligned}$$

□

Hamilton's equations (35a) and (35b) can be written in a more succinct form as

$$\dot{\boldsymbol{y}} = \mathbf{J}\frac{\partial H}{\partial \boldsymbol{y}}(\boldsymbol{y}) = \boldsymbol{X}_H(\boldsymbol{y}), \tag{39}$$

where $\boldsymbol{y} = (\boldsymbol{q}, \boldsymbol{p})$ is a point on the phase space T^*Q, $\boldsymbol{X}_H(\boldsymbol{y}) \in T_{\boldsymbol{y}}T^*Q$ is known as the *Hamiltonian vector field* and the skew-symmetric matrix

$$\mathbf{J} = \begin{bmatrix} \mathbf{0}_{d\times d} & \mathbf{I}_{d\times d} \\ -\mathbf{I}_{d\times d} & \mathbf{0}_{d\times d} \end{bmatrix}, \tag{40}$$

is the *canonical symplectic matrix* which represents a skew-symmetric bilinear form, the *canonical symplectic two-fom*, that is conserved along solution trajectories on the phase space. This crucial aspect will be considered in detail in the following section. In (40) $\mathbf{I}_{d\times d}$ and $\mathbf{0}_{d\times d}$, are the $d \times d$ identity and zero matrices, respectively.

The expression (39) provides an intrinsic definition for Hamiltonian systems since given a smooth enough Hamiltonian function $H : T^*Q \to \mathbb{R}$, the matrix $\mathbf{J}$ may be though as a the linear transformation that maps $\partial H(\boldsymbol{y})/\partial \boldsymbol{y}$ which belongs to $T^*_{\boldsymbol{y}}T^*Q$, to the Hamiltonian vector field $\boldsymbol{X}_H(\boldsymbol{y})$ which belongs to $T_{\boldsymbol{y}}T^*Q$. See Fig. 9.

Note that as long as a solution of the system (39) exists, it is possible to define a function

$$\begin{aligned} \boldsymbol{\Phi} : T^*Q \times \mathbb{R} &\to T^*Q \\ (\boldsymbol{y}, t) &\mapsto \boldsymbol{\Phi}(\boldsymbol{y}, t) \end{aligned} \tag{41}$$

with $\boldsymbol{\Phi}(\boldsymbol{y}, 0) = \boldsymbol{y}_0$ so that the trajectory of the mechanical system that starts at $\boldsymbol{y}_0 \in T^*Q$ is given by

$$\boldsymbol{y}(t) = \boldsymbol{\Phi}(\boldsymbol{y}_0, t). \tag{42}$$

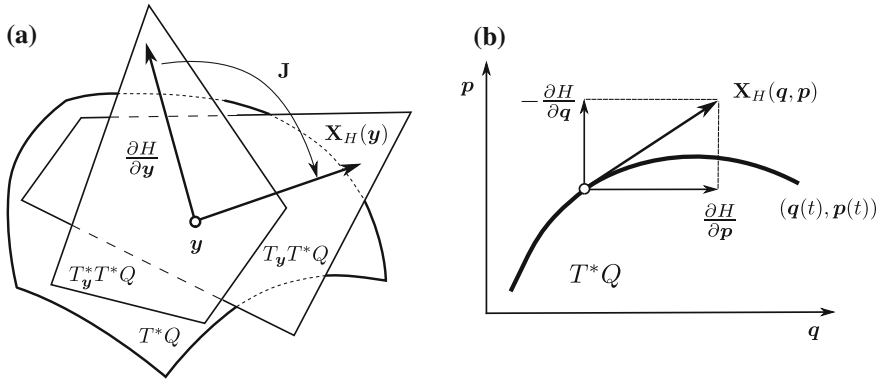


Fig. 9 Consider a point $\boldsymbol{y} = (\mathbf{q}, \mathbf{p})$ belonging to the phase field T^*Q. **a** Shows how the gradient $\partial H(\boldsymbol{y})/\partial \boldsymbol{y}$ which belongs to the linear space $T^*_{\boldsymbol{y}}T^*Q$ is related to the Hamiltonian map $\mathbf{X}_H(\boldsymbol{y}) \in T_{\boldsymbol{y}}T^*Q$ through the map $\mathbf{J}$. In (**b**) a geometric interpretation of Hamilton's equations is provided. The Hamiltonian vector field $\boldsymbol{X}_H(\boldsymbol{y}) = (\partial H/\partial \mathbf{p}, -\partial H/\partial \mathbf{q})$ is shown as a velocity vector which is tangent to the solution trajectory

The map $\boldsymbol{\Phi}$ is called the *flow map* of the Hamiltonian vector field $\mathbf{X}_H$, and it enables to consider the mechanical system as evolving collections of initial conditions, instead of only individual ones.

2.5 Conservation Properties: Hamiltonian Point of View

Since the Hamiltonian and Lagrangian points of view of mechanics are equivalent,[5] conservation properties may be stated in terms of variables on T^*Q. In this section we show how the properties described in Sect. 2.3 look when observed from the Hamiltonian point of view of mechanics. In particular, we present a comprehensive description of the connection existing between the symplectic nature of Hamiltonian flows and the conservation of volume in the phase space. This might seems at first glance, as a technicality of minor significance, however it has important consequences in the study of conservative and dissipative perturbations of nearly integrable Hamiltonian systems (Meyer et al. 2009; Maddocks and Overton 1995; Stoffer 1997, 1998; Hairer and Lubich 1999) and in the formulation of structure-preserving algorithms (Hairer et al. 2006, Chap. X, XI, XII; Channell and Scovel 1990; Meyer et al. 2011). The uninterested reader may skip this section, and continue directly to the formulation of variational integrators.

[5]There exist some exceptions in this regard such as for example, when the Legendre transform is not well defined.

2.5.1 Conservation of Energy

An autonomous Hamiltonian system conserves the energy, or in other words, its Hamiltonian function is exactly conserved along solution trajectories. This conservation property is readily verified by computing the time derivative of the Hamiltonian function (34) over a solution trajectory $(\mathbf{q}(t), \mathbf{p}(t))$ that satisfies (35a) and (35b), i.e.,

$$\dot{H}(\mathbf{q}, \mathbf{p}) = \frac{\partial H}{\partial \mathbf{p}} \cdot \dot{\mathbf{p}} + \frac{\partial H}{\partial \mathbf{q}} \cdot \dot{\mathbf{q}} = -\frac{\partial H}{\partial \mathbf{p}} \cdot \frac{\partial H}{\partial \mathbf{q}} + \frac{\partial H}{\partial \mathbf{q}} \cdot \frac{\partial H}{\partial \mathbf{p}} = 0. \tag{43}$$

2.5.2 Symplecticity of the Flow

An outstanding property of Hamiltonian systems is the *symplecticity* of their flows (41) on the phase space. In fact, the preservation of a discrete form of this property by the so-called *symplectic methods* (Yoshida 1993; Hairer et al. 2003) contributes to explain their superior performance in numerical simulations in terms of long-term stability and error propagation, when applied to Hamiltonian problems and/or to perturbations of them (Maddocks and Overton 1995; Hairer and Lubich 1999; Hairer et al. 2006, Chap. XII). A nice introduction to this subject can be consulted in (Leimkuhler and Reich 2005, Chap. 3) where a geometric interpretation of symplecticity in terms of volume preservation on the phase space is provided. A complete survey about symplectic forms on differentiable manifolds can be found elsewhere, e.g., (Arnold 1989; Marsden and Ratiu 1999).

To make a proper introduction of this geometric property of Hamiltonian flows, we closely follow (Leimkuhler and Reich 2005) and (Hairer et al. 2006). We first restrict the discussion to mappings from $\mathbb{R}^{2d}$ to itself and we show that for $d = 1$ the symplectic nature of the flow is manifested through area preservation on the phase space. The generalization to $d > 1$ allows to identify symplecticity with volume preservation. Finally, we provide some rudiments in order to extend the ideas to symplectic maps on cotangent bundles.

Symplectic two-fom. Consider two points $z_1 = (q_1^1, ..., q_1^n, p_1^1, ..., p_1^n)$ and $z_2 = (q_2^1, ..., q_2^n, p_2^1, ..., p_2^n)$ both belonging to $\mathbb{R}^{2n}$. The oriented area of the parallelogram spanned by the projections of z_1 and z_2 on the coordinate plane (q^i, p^i) is given by

$$\rho_i = p_2^i q_1^i - q_2^i p_1^i, \qquad i = 1, ...n, \tag{44}$$

see Fig. 10. The sum of all the oriented areas is

$$\omega(z_1, z_1) = \rho_1 + \cdots + \rho_n. \tag{45}$$

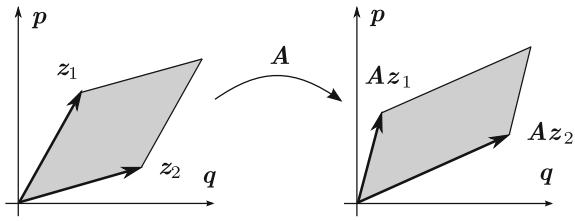


Fig. 10 Consider $z_1 = (q_1, p_1)$ and $z_2 = (q_2, p_2) \in \mathbb{R}^2$. The area of the parallelogram spanned is given by z_1 and z_2 is $\omega(z_1, z^2) = p_1 q_2 - p_2 q_1$. The linear mapping $\boldsymbol{A} : \mathbb{R}^2 \to \mathbb{R}^2$ is symplectic if $\omega(z_1, z_2) = \omega(\mathbf{A}z_1, \mathbf{A}z_2)$. This is equivalent to area conservation

This expression defines the skew-symmetric bilinear map $\boldsymbol{\omega} : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \mapsto \mathbb{R}$ that may be represented in matrix form as

$$\omega(z_1, z_2) = z_2 \cdot \mathbf{J}^{-1} z_1, \tag{46}$$

where $\mathbf{J}$ is the canonical symplectic matrix (40).

Symplectic mappings. A linear mapping $\boldsymbol{A} : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is called symplectic if

$$\boldsymbol{A}^{\mathrm{t}} \mathbf{J}^{-1} \boldsymbol{A} = \mathbf{J}^{-1}, \tag{47}$$

which is equivalent to

$$\omega(z_1, z_2) = \omega(\boldsymbol{A}z_1, \boldsymbol{A}z_2), \quad \text{for all } z_1, z_2 \in \mathbb{R}^{2n}. \tag{48}$$

Then, a symplectic linear map conserves the sum of the oriented areas of the parallelograms spanned by the projections of any two arbitrary points z_1 and z_2 on the coordinate planes $(q^i, p^i), i = 1, ..., n$. For $n = 1$ it is equivalent to area preservation in the phase space and for $n > 1$ to volume preservation. See Fig. 10.

For nonlinear maps, an analogous result holds. A differentiable map $\psi : U \subset \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is symplectic if its Jacobian matrix $\partial \psi(\boldsymbol{y}) / \partial \boldsymbol{y}$ is symplectic for all $\boldsymbol{y} \in U$, i.e.,

$$\left(\frac{\partial \psi}{\partial \boldsymbol{y}}(\boldsymbol{y}) \right)^{\mathrm{t}} \mathbf{J}^{-1} \left(\frac{\partial \psi}{\partial \boldsymbol{y}}(\boldsymbol{y}) \right) = \mathbf{J}^{-1}, \tag{49}$$

which is equivalent to (48) after replacing $\boldsymbol{A}$ by $\partial \psi(\boldsymbol{y}) / \partial \boldsymbol{y}$. A geometric interpretation of symplecticity for nonlinear maps can be consulted in for example (Hairer et al. 2006, Chap. VI.2) or (Leimkuhler and Reich 2005).

Taking into account the above definitions it is possible to state the following fundamental result due to Poincaré:

Theorem (Poincaré 1899). For each fixed time t, the flow map $\boldsymbol{\Phi}(\boldsymbol{y}_0, t)$ of a Hamiltonian system with a twice continuously differentiable function $H(\boldsymbol{y})$, defines a symplectic transformation, i.e.,

$$\left(\frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{y}_0}(\boldsymbol{y}_0, t)\right)^{\mathrm{t}} \mathbf{J}^{-1} \left(\frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{y}_0}(\boldsymbol{y}_0, t)\right) = \mathbf{J}^{-1}. \tag{50}$$

Proof See (Hairer et al. 2006, pp. 184–185) or (Arnold 1989, Chap. 8) □

Equation (50) provides an intrinsic definition for Hamiltonian systems in the sense that any continuously differentiable $\boldsymbol{f} : U \subset \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ can be locally written as

$$\boldsymbol{f} = \mathbf{J} \frac{\partial H}{\partial \boldsymbol{y}}(\boldsymbol{y}),$$

for an appropriate Hamiltonian function $H(\boldsymbol{y})$ if the flow generated by

$$\dot{\boldsymbol{y}} = \boldsymbol{f}(\boldsymbol{y}),$$

is symplectic for all $\boldsymbol{y} \in U$ and small enough t (Hairer et al. 2006).

Hamiltonian flows on manifolds. The form in which Hamiltonian systems generate symplectic flows on the phase space can be alternatively explained following the ideas presented in Sect. 2.3.2. To this end, we consider a Hamiltonian system with differentiable flow map $\boldsymbol{\Phi}(\boldsymbol{y}, t)$ where $\boldsymbol{y} \in T^*Q$ and $t > 0$. We also consider the elements $\boldsymbol{u}$ and $\boldsymbol{v}$ which belong to $T_{\boldsymbol{y}} T^*Q$. Owing to the fact that $\boldsymbol{\Phi}$ is symplectic, it preserves the canonical symplectic form $\boldsymbol{\omega}$, according to

$$\begin{aligned} \boldsymbol{\omega}_{\boldsymbol{y}}(\boldsymbol{u}, \boldsymbol{v}) &= \boldsymbol{\omega}_{\boldsymbol{\Phi}(\boldsymbol{y},t)}(T\boldsymbol{\Phi} \cdot \boldsymbol{u}, T\boldsymbol{\Phi} \cdot \boldsymbol{v}), \\ \boldsymbol{u} \cdot \mathbf{J}^{-1} \boldsymbol{v} &= (T\boldsymbol{\Phi} \cdot \boldsymbol{u}) \cdot \mathbf{J}^{-1} (T\boldsymbol{\Phi} \cdot \boldsymbol{v}) \end{aligned}$$

which, since

$$T\boldsymbol{\Phi}(\boldsymbol{y}, t) \cdot \boldsymbol{u} = \frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{y}}(\boldsymbol{y}, t)\boldsymbol{u} = \boldsymbol{\Phi}' \boldsymbol{u},$$

yields to

$$\boldsymbol{u} \cdot \mathbf{J}^{-1} \boldsymbol{v} = \boldsymbol{u} \cdot \left(\boldsymbol{\Phi}'' \mathbf{J}^{-1} \boldsymbol{\Phi}'\right) \boldsymbol{v}.$$

The condition given in (50) is recovered after noticing that $\boldsymbol{u}$ and $\boldsymbol{v}$ are arbitrary elements.

Figure 11 shows an illustration of the flow map Φ for the particle in the hoop setting $m = 1$, $r = 1$, and $g = 1$. The horizontal axis has the angle coordinate θ, and the vertical axis has the momentum, $mr^2\dot{\theta}$, which for these constants reduces to $\dot{\theta}$. We consider initial conditions belonging to the boundary of the square $[0, 4] \times [-2, 2] \in T^*Q \subset \mathbb{R}^2$ drawn in black color and to the boundary of the square $[0, 2] \times [-1, 1] \subset \mathbb{R}^2$ drawn in blue color. Then, we define the set of initial conditions as

$$\mathbf{B}^0 = \text{boundary}\big([0, 2] \times [-1, 1]\big) \cup \text{boundary}\big([0, 4] \times [-2, 2]\big). \tag{51}$$

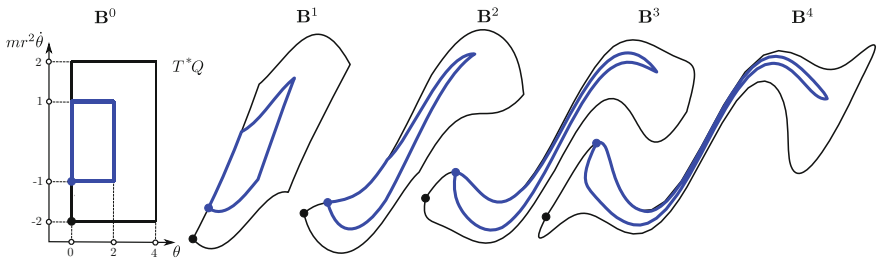


Fig. 11 Illustration of the flow map Φ for the particle in the hoop with $m = 1$, $r = 1$, and $g = 1$. The horizontal axis has the angle coordinate and the vertical axis has the momentum. For every time instant $\{t^i\}_{i=1,\ldots,4}$ the map Φ generates an updated configuration, $\mathbf{B}^i$, of the set of initial conditions (51) according to (52). Despite the fact that the shape of the set $\mathbf{B}^0$ deforms drastically in time, its area remains invariant due to the symplectic (volume-preserving) nature the flow

The area enclosed by the figure is 16. The flow map Φ deforms the shape of the set of initial conditions in such a manner that the sets

$$\mathbf{B}^i := \left\{ \boldsymbol{v} \in T^*Q \;\middle|\; \boldsymbol{v} = \Phi(\boldsymbol{u}_0, t^i),\ \boldsymbol{u}_0 \in \mathbf{B}^0 \right\}, \quad i = 1, 2, 3, 4. \tag{52}$$

represent updated configurations of $\mathbf{B}^0$ at later times $\{t^i\}_{i=1,\ldots,4}$. The same figure shows that even though the shape of the $\mathbf{B}^i$'s changes drastically, its area remains invariant due to the fact that Φ is a symplectic map.

3 Discrete Lagrangian Mechanics

This section focuses on presenting a systematic methodology to construct structure-preserving methods for finite-dimensional and autonomous, Lagrangian dynamical systems. To this end, we take advantage of some concepts in discrete Lagrangian mechanics (Marsden and West 2001). The basic idea is to consider discrete trajectories for the mechanical system, and to define the dynamics of those trajectories via a discrete version of Hamilton's principle applied to an approximate action.

3.1 Construction of Variational Integrators

We begin by considering a discretization $0 = t^0 < t^1 < \ldots < t^{N-1} < t^N = T$ of the interval $[0, T]$, for some $N \in \mathbb{N}$. In the most common case $t^n = nh$, with $h = T/N$, and for simplicity, this is the case we shall adopt for these notes. A *discrete trajectory* associated to this discretization is an element of

$$Q^{N+1} = \underbrace{Q \times \cdots \times Q}_{N \text{ times}},$$

and it is indicated by

$$\{\mathbf{q}^i\}_{i=0,\ldots,N} = \{\mathbf{q}^0, \mathbf{q}^1, \ldots, \mathbf{q}^{N-1}, \mathbf{q}^N\} \subset Q^{N+1}.$$

Let us then introduce a *discrete Lagrangian* $L_d \colon Q \times Q \times \mathbb{R} \to \mathbb{R}$ such that

$$L_d(\mathbf{q}^0, \mathbf{q}^1, h) \approx \int_0^h L(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, dt \tag{53}$$

when $\mathbf{q}^0 = \mathbf{q}(0)$ and $\mathbf{q}^1 = \mathbf{q}(h)$, where $\mathbf{q}(t)$ is the exact trajectory of the system in that time interval. We shall discuss this approximation in more length later. Notice that the tangent space TQ has been replaced by $Q \times Q \times h$. Notice as well that the discrete Lagrangian is not an approximation of the Lagrangian, but rather of the action over a time interval $[0, h]$.

The discrete Lagrangian will be used to select a discrete trajectory for the system, inasmuch the Lagrangian is used to select a trajectory in the time-continuous system.

Example We explain how the discrete Lagrangians of some of the most popular time integrators for ODE's are constructed. We also apply the method to some particular Lagrangian functions.

- **Rectangle rule** (case 1). A general guideline for the construction of discrete Lagrangians is to combine an approximation space for a trajectory of the system and a quadrature rule for the integral over $[0, h]$. In this simplest case, we approximate the exact trajectory of the system $\mathbf{q}(t)$ over $[0, T]$ with a continuous piecewise linear polynomial, i.e.,

$$\mathbf{q}(t) \approx \mathbf{q}^0 + \frac{\mathbf{q}^1 - \mathbf{q}^0}{h} t, \qquad t \in [0, h].$$

Therefore, the exact velocity, $\dot{\mathbf{q}}(t)$, is approximated by

$$h^{-1}(\mathbf{q}^1 - \mathbf{q}^0) \approx \dot{\mathbf{q}}(t), \qquad t \in [0, h],$$

see Fig. 12. The discrete Lagrangian is then constructed as

$$L_d^0(\mathbf{q}^0, \mathbf{q}^1, h) = hL\left(\mathbf{q}^0, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right) \approx \int_0^h L(\mathbf{q}(t), \dot{\mathbf{q}}(t))dt, \tag{54a}$$

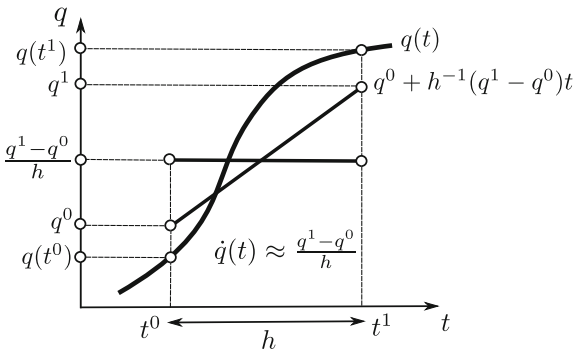


Fig. 12 Piecewise linear approximation of $\mathbf{q}(t)$, continuous across time intervals. The approximation obtained for the velocity is constant over the time interval, and hence generally discontinuous across time interval boundaries

which is equivalent to using a single quadrature point at the beginning of the interval $[0, h]$ (rectangle rule). For the particle in the hoop, this discrete Lagrangian takes the form

$$L_d^0(\theta^0, \theta^1, h) = h\left[\frac{m}{2}r^2\left(\frac{\theta^1-\theta^0}{h}\right)^2 - mgr\sin\theta^0\right]. \tag{54b}$$

For thermoelastic systems we have

$$L_d^0(\mathbf{q}^0, \mathbf{q}^1, \mathbf{\Phi}^0, \mathbf{\Phi}^1, h) = \tag{54c}$$
$$h\left[\frac{1}{2}\left(\frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right)\cdot\mathbf{M}(\mathbf{q}^0)\left(\frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right) - \mathsf{A}\left(\mathbf{q}^0, \frac{\mathbf{\Phi}^1-\mathbf{\Phi}^0}{h}\right)\right].$$

- **Rectangle rule** (case 2). Alternatively, it is possible to approximate the action functional by

$$L_d^1(\mathbf{q}^0, \mathbf{q}^1, h) = hL\left(\mathbf{q}^1, \frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right). \tag{54d}$$

- **Trapezoidal rule**. Both, (54a) and (54d) are particular cases of the discrete Lagrangian

$$L_d^\alpha(\mathbf{q}^0, \mathbf{q}^1, h) = (1-\alpha)L_d^0(\mathbf{q}^0, \mathbf{q}^1, h) + \alpha L_d^1(\mathbf{q}^0, \mathbf{q}^1, h). \tag{54e}$$

  which for $\alpha = \frac{1}{2}$ is known as the *trapezoidal rule*.
- **Midpoint rule**. The so-called *implicit midpoint method* is derived from the midpoint discrete Lagrangian

$$L_d^m(\mathbf{q}^0, \mathbf{q}^1, h) = hL\left(\frac{\mathbf{q}^1+\mathbf{q}^0}{2}, \frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right). \tag{54f}$$

For the two particles joined by the rigid rod, it is given by

$$L_d(x^0, y^0, \theta^0, x^1, y^1, \theta^1, h) = \\ m\left[\frac{(x^1_{CG} - x^0_{CG})^2}{h} + \frac{(y^1_{CG} - y^0_{CG})^2}{h} + L^2\frac{(\theta^1_{CG} - \theta^0_{CG})^2}{h}\right], \tag{54g}$$

and for thermoelastic systems,

$$L^0_d(\mathbf{q}^0, \mathbf{q}^1, \mathbf{\Phi}^0, \mathbf{\Phi}^1, h) = \\ h\left[\frac{1}{2}\left(\frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right)\cdot \mathbf{M}\left(\frac{\mathbf{q}^1 + \mathbf{q}^0}{2}\right)\left(\frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right) - \right. \\ \left. \mathsf{A}\left(\frac{\mathbf{q}^1 + \mathbf{q}^0}{2}, \frac{\mathbf{\Phi}^1 - \mathbf{\Phi}^0}{h}\right)\right]. \tag{54h}$$

Back to selecting the discrete trajectory, we construct the *discrete Action Sum* $S_d\colon Q^{N+1} \to \mathbb{R}$ as

$$S_d(\mathbf{q}^0, \ldots, \mathbf{q}^N) = \sum_{i=0}^{N-1} L_d(\mathbf{q}^i, \mathbf{q}^{i+1}, h). \tag{55}$$

The *discrete Variational Principle* is formulated mimicking the continuous case: *The trajectory* $\{\mathbf{q}_i\}_{i=0,\ldots,N}$ *is a stationary point of the discrete action sum among all variations that leave the endpoints fixed.* In other words,

$$\langle S_d(\mathbf{q}^0, \ldots, \mathbf{q}^N), (\delta\mathbf{q}^0, \ldots, \delta\mathbf{q}^N)\rangle = \sum_{i=0}^{N} \frac{\partial S_d}{\partial \mathbf{q}^i}(\mathbf{q}^0, \ldots, \mathbf{q}^N)\delta\mathbf{q}^i = 0 \tag{56}$$

for any variations that satisfy $\delta\mathbf{q}^0 = \delta\mathbf{q}^N = 0$. This happens if and only if

$$\frac{\partial S_d}{\partial \mathbf{q}^i}\left(\mathbf{q}^0, \ldots, \mathbf{q}^N\right) = 0, \qquad i = 1, \ldots, N-1. \tag{57}$$

In terms of the discrete Lagrangians, this reads

$$0 = \frac{\partial L_d}{\partial \mathbf{q}^i}(\mathbf{q}^i, \mathbf{q}^{i+1}, h) + \frac{\partial L_d}{\partial \mathbf{q}^i}(\mathbf{q}^{i-1}, \mathbf{q}^i, h), \qquad i = 1, \ldots, N-1. \tag{58}$$

These are the *discrete Euler–Lagrange equations*, or DEL equations, and they define the discrete trajectory. They also define the algorithm: If $\mathbf{q}^{i-1}, \mathbf{q}^i$ are known, these equations need to be solved to find $\mathbf{q}^{i+1}$.

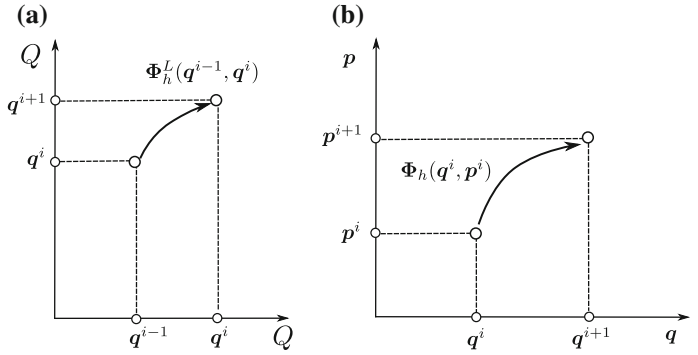


Fig. 13 Schematic representation of **a** the discrete Lagrangian map $\mathbf{\Phi}_h^L : Q \times Q \to Q \times Q$ and **b** the discrete Hamiltonian map $\mathbf{\Phi}_h : TQ^* \to TQ^*$

Thus, the discrete Euler–Lagrange equations implicitly define a map

$$\begin{aligned} \mathbf{\Phi}_h^L : \ & Q \times Q \ \to Q \times Q \\ & (\mathbf{q}^{i-1}, \mathbf{q}^i) \mapsto \mathbf{\Phi}_h^L(\mathbf{q}^{i-1}, \mathbf{q}^i) = (\mathbf{q}^i, \mathbf{q}^{i+1}) \end{aligned},$$

known as the *discrete Lagrangian map*. If $\partial L_d / \partial \mathbf{q}^i$ is invertible over $Q \times Q$, then this map is well defined. The map $\mathbf{\Phi}_h^L$ then flows the system forward over $Q \times Q$ between consecutive time-steps. See Fig. 13a.

As we shall see later, under simple conditions the discrete trajectories will approximate the exact trajectories of the Lagrangian in (53). The map $\mathbf{\Phi}_h^L$ can thus be considered a time-integrator, and because it satisfies the DEL equations of the discrete variational principle, it is called a *variational integrator*.

Example In this example, we derive the DEL equations for the discrete Lagrangians introduced so far, as applied to a general multibody system with a configuration-independent mass matrix.

- **Rectangle rules 1 and 2 and trapezoidal rule**. Consider first the discrete Lagrangian L_d^0 for such system. Then, we have that

$$\begin{aligned} \frac{\partial L_d^0}{\partial \mathbf{q}^0}(\mathbf{q}^0, \mathbf{q}^1, h) &= -\mathbf{M}\left(\frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right) - h\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}^0) \\ \frac{\partial L_d}{\partial \mathbf{q}^1}(\mathbf{q}^0, \mathbf{q}^1, h) &= \mathbf{M}\left(\frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right), \end{aligned}$$

from where the DEL equations follow as

$$\begin{aligned} 0 &= -\mathbf{M}\left(\frac{\mathbf{q}^{i+1}-\mathbf{q}^{i}}{h}\right) - h\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}^{i}) + \mathbf{M}\left(\frac{\mathbf{q}^{i}-\mathbf{q}^{i-1}}{h}\right) \\ &= -\mathbf{M}\left(\frac{\mathbf{q}^{i+1}-2\mathbf{q}^{i}+\mathbf{q}^{i-1}}{h}\right) - h\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}^{i}). \end{aligned} \tag{59}$$

These are the equations of Newmark's second-order explicit algorithm, known also as central differences or Störmer-Verlet.

Since $\mathbf{M}$ is positive definite, it is possible to solve (59) to get

$$\mathbf{q}^{i+1} = 2\mathbf{q}^{i} - \mathbf{q}^{i-1} - h^2\mathbf{M}^{-1}\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}^{i}).$$

The discrete Lagrangian map is then

$$\Phi_h^L(\mathbf{q}^{i-1}, \mathbf{q}^{i}) = \left(\mathbf{q}^{i}, 2\mathbf{q}^{i} - \mathbf{q}^{i-1} - h^2\mathbf{M}^{-1}\frac{\partial U}{\partial \mathbf{q}}(\mathbf{q}^{i})\right). \tag{60}$$

It is simple to check that the discrete Lagrangian

$$L_d^1(\mathbf{q}_0, \mathbf{q}_1, h) = hL\left(\mathbf{q}^1, \frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right),$$

gives rise to the same DEL in this case! Therefore, the *trapezoidal rule* (54e) yields the same DEL independently of the value chosen for α. This may not the case when the dependence of the Lagrangian on the velocities is more complex, such as in thermoelastic systems, see for example (Mata and Lew 2011, 2014).

- **Midpoint rule**. Using the midpoint rule, the discrete Lagrangian is given by

$$L_d^m(\mathbf{q}^0, \mathbf{q}^1, h) = h\left[\frac{1}{2}\left(\frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right)\cdot\mathbf{M}\left(\frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right) - U\left(\frac{\mathbf{q}^1+\mathbf{q}^0}{2}\right)\right], \tag{61}$$

and thus,

$$\begin{aligned} \frac{\partial L_d^m}{\partial \mathbf{q}^0}(\mathbf{q}^0, \mathbf{q}^1, h) &= -\mathbf{M}\left(\frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right) - \frac{h}{2}\frac{\partial U}{\partial \mathbf{q}}\left(\frac{\mathbf{q}^1+\mathbf{q}^0}{2}\right), \\ \frac{\partial L_d^m}{\partial \mathbf{q}^1}(\mathbf{q}^0, \mathbf{q}^1, h) &= \mathbf{M}\left(\frac{\mathbf{q}^1-\mathbf{q}^0}{h}\right) - \frac{h}{2}\frac{\partial U}{\partial \mathbf{q}}\left(\frac{\mathbf{q}^1+\mathbf{q}^0}{2}\right). \end{aligned}$$

Therefore, the DEL equations are given by

$$0 = -\mathbf{M}\left(\frac{\mathbf{q}^{i+1} - 2\mathbf{q}^i + \mathbf{q}^{i-1}}{h}\right) - \frac{h}{2}\left(\frac{\partial U}{\partial \mathbf{q}}\left(\frac{\mathbf{q}^{i+1} + \mathbf{q}^i}{2}\right) + \frac{\partial U}{\partial \mathbf{q}}\left(\frac{\mathbf{q}^i + \mathbf{q}^{i-1}}{2}\right)\right).$$

Compare with (59). Note that in this case it is not possible to provide an explicit expression for the discrete Lagrangian map $\Phi_h^L(\mathbf{q}^{i-1}, \mathbf{q}^i)$. □

More Discrete Lagrangians. The discrete Lagrangians shown so far give rise to first- or second-order variational integrators. In the following, we show some discrete Lagrangians that give rise to higher order time integrators.

- **Quadratic rule**. Consider a piecewise continuous quadratic approximation of the trajectory over $[0, h]$, i.e.,

$$\mathbf{q}^{t/h} = \mathrm{N}_1(t)\mathbf{q}^0 + \mathrm{N}_2(t)\mathbf{q}^{\frac{1}{2}} + \mathrm{N}_3(t)\mathbf{q}^1 \approx \mathbf{q}(t), \qquad t \in [0, h],$$

where $\mathbf{q}^{\frac{1}{2}} \approx \mathbf{q}(h/2)$ is a coefficient needed for the quadratic interpolation and

$$\mathrm{N}_i : [0, h] \to \mathbb{R}, \qquad i = 1, 2, 3,$$

are a set of basis functions for the set of second degree polynomials over $[0, h]$, $\mathbb{P}^2([0, h])$, which satisfy

$$\mathrm{N}_i(\tau_j) = \delta_{ij} \qquad \text{for } \tau_j \in \{0, h/2, h\},$$

see Fig. 14a. The velocity is approximated by

$$\dot{\mathbf{q}}^{t/h} = \dot{\mathrm{N}}_1(t)\mathbf{q}^0 + \dot{\mathrm{N}}_2(t)\mathbf{q}^{\frac{1}{2}} + \dot{\mathrm{N}}_3(t)\mathbf{q}^1 \approx \dot{\mathbf{q}}(t), \qquad t \in [0, h].$$

The next step to construct a discrete Lagrangian consists in providing a quadrature rule, described by a set of quadrature points over $[0, h]$ and weights

$$\{\xi_i, w_i\}_{i=1,\ldots,n_q}.$$

Then, the discrete Lagrangian is given by

$$L_d(\mathbf{q}^0, \mathbf{q}^1, h) = \inf_{\substack{\widetilde{\mathbf{q}}(t) \in \mathbb{P}^2([0,h]) \\ \widetilde{\mathbf{q}}(0)=\mathbf{q}^0,\ \widetilde{\mathbf{q}}(h)=\mathbf{q}^1}} L_d^u(\mathbf{q}^0, \mathbf{q}^{1/2}, \mathbf{q}^1, h) \tag{62a}$$

where the *unoptimized* Lagrangian is defined as

$$L_d^u(\mathbf{q}^0, \mathbf{q}^{1/2}, \mathbf{q}^1, h) = \sum_{i=1}^{n_q} w_i L(\widetilde{\mathbf{q}}(\xi_i), \dot{\widetilde{\mathbf{q}}}(\xi_i)). \tag{62b}$$

When a unique solution exists, the infimization in this last equation implicitly defines the value of $\mathbf{q}^{1/2}$ given those of $\mathbf{q}^0$ and $\mathbf{q}^1$. In practice, the infimization is imposed by requesting $\mathbf{q}^{1/2}$ to satisfy the stationarity condition

$$\frac{\partial L_d^u}{\partial \mathbf{q}^{1/2}}\left(\mathbf{q}^0, \mathbf{q}^{1/2}, \mathbf{q}^1, h\right) = \mathbf{0}. \tag{62c}$$

The DEL equations are then given by (58). In order to clarify the procedure consider the thermoelastic Lagrangian (7) with constant mass matrix along with the Gauss–Lobatto quadrature rule with points $\{0, h/2, h\}$ and weights $\{h/6, 4h/6, h/6\}$. First we define the *unoptimized* discrete Lagrangian

$$L_d^u(\mathbf{q}^0, \mathbf{\Phi}^0, \mathbf{q}^{\frac{1}{2}}, \mathbf{\Phi}^{\frac{1}{2}}, \mathbf{q}^1, \mathbf{\Phi}^{\frac{1}{2}}, h) = \frac{h}{6}\left(L^0 + 4L^{\frac{1}{2}} + L^1\right), \tag{62d}$$

where

$$\begin{aligned} L^0 &= \dot{\mathbf{q}}^0 \cdot \mathbf{M}\dot{\mathbf{q}}^0 - \mathsf{A}\left(\mathbf{q}^0, \dot{\mathbf{\Phi}}^0\right), \\ L^{\frac{1}{2}} &= \dot{\mathbf{q}}^{\frac{1}{2}} \cdot \mathbf{M}\dot{\mathbf{q}}^{\frac{1}{2}} - \mathsf{A}\left(\mathbf{q}^{\frac{1}{2}}, \dot{\mathbf{\Phi}}^{\frac{1}{2}}\right), \\ L^1 &= \dot{\mathbf{q}}^1 \cdot \mathbf{M}\dot{\mathbf{q}}^1 - \mathsf{A}\left(\mathbf{q}^1, \dot{\mathbf{\Phi}}^1\right). \end{aligned}$$

and

$$\begin{aligned} \dot{\mathbf{\Phi}}^0 &= h^{-1}\left(4\mathbf{\Phi}^{\frac{1}{2}} - 3\mathbf{\Phi}^0 - \mathbf{\Phi}^1\right), & \dot{\mathbf{\Phi}}^{\frac{1}{2}} &= h^{-1}\left(\mathbf{\Phi}^1 - \mathbf{\Phi}^0\right), \\ \dot{\mathbf{\Phi}}^1 &= h^{-1}\left(\mathbf{\Phi}^0 - 4\mathbf{\Phi}^{\frac{1}{2}} + 3\mathbf{\Phi}^1\right), & \dot{\mathbf{q}}^0 &= h^{-1}\left(4\mathbf{q}^{\frac{1}{2}} - 3\mathbf{q}^0 - \mathbf{q}^1\right), \\ \dot{\mathbf{q}}^{\frac{1}{2}} &= h^{-1}\left(\mathbf{q}^1 - \mathbf{q}^0\right), & \dot{\mathbf{q}}^1 &= h^{-1}\left(\mathbf{q}^0 - 4\mathbf{q}^{\frac{1}{2}} + 3\mathbf{q}^1\right). \end{aligned}$$

Then, the discrete Lagrangian follows as

$$L_d(\mathbf{q}^0, \mathbf{\Phi}^0, \mathbf{q}^1, \mathbf{\Phi}^1, h) = \inf_{\mathbf{q}^{\frac{1}{2}}, \mathbf{\Phi}^{\frac{1}{2}}} L_d^u(\mathbf{q}^0, \mathbf{\Phi}^0, \mathbf{q}^{\frac{1}{2}}, \mathbf{\Phi}^{\frac{1}{2}}, \mathbf{q}^1, \mathbf{\Phi}^{\frac{1}{2}}, h). \tag{62e}$$

The values of $\mathbf{q}^{\frac{1}{2}}$ and $\mathbf{\Phi}^{\frac{1}{2}}$ that infimize L_d^u for given values of $\mathbf{q}^0$, $\mathbf{\Phi}^0$, $\mathbf{q}^1$ and $\mathbf{\Phi}^1$ satisfy the equations

$$\frac{\partial L_d^u}{\partial \mathbf{q}^{\frac{1}{2}}} = \frac{4\mathbf{m}}{h^2}(\mathbf{q}^k - 2\mathbf{q}^{k+\frac{1}{2}} + \mathbf{q}^{k+1}) - \mathbf{f}^{k+\frac{1}{2}} = \mathbf{0}, \tag{62f}$$

$$\frac{\partial L_d^u}{\partial \mathbf{\Phi}^{\frac{1}{2}}} = \mathbf{\Gamma}^{k+} - \mathbf{\Gamma}^{(k+1)-} = \mathbf{0}, \tag{62g}$$

where

$$\begin{aligned}
\mathbf{f}^{k+\frac{1}{2}} &= -\frac{\partial \mathsf{A}}{\partial \mathbf{q}}\left(\mathbf{q}^{k+\frac{1}{2}}, \frac{\mathbf{\Phi}^{k+1} - \mathbf{\Phi}^{k}}{h}\right), \\
\mathbf{\Gamma}^{k+} &= -\frac{\partial \mathsf{A}}{\partial \theta}\left(\mathbf{q}^{k}, \frac{-3\mathbf{\Phi}^{k} + 4\mathbf{\Phi}^{k+\frac{1}{2}} - \mathbf{\Phi}^{k+1}}{h}\right), \\
\mathbf{\Gamma}^{k-} &= -\frac{\partial \mathsf{A}}{\partial \theta}\left(\mathbf{q}^{k}, \frac{\mathbf{\Phi}^{k-1} - 4\mathbf{\Phi}^{k-\frac{1}{2}} + 3\mathbf{\Phi}^{k}}{h}\right).
\end{aligned}$$

- **General Galerkin variational integrators**. The basic idea behind the formulation of general Galerkin VI's consists in increasing the order of the approximation polynomial along with the order of the quadrature rule used to approximate the action of the system. To this end, we consider $s+1$ control points over $[0, h]$,

$$\left\{\mathbf{q}^{0_\nu}\right\}_{\nu=0,\ldots,s},$$

which satisfy $\mathbf{q}^{0_0} = \mathbf{q}^0$ and $\mathbf{q}^{0_s} = \mathbf{q}^1$ and correspond to the values of the trajectory at control times

$$\{d_\nu h\}_{\nu=0,\ldots,s} \quad \text{with} \quad 0 = d_0 < d_1 < \ldots < d_s = 1.$$

Moreover, we assume that the exact trajectory over $[0, h]$ is approximated by a unique s-degree polynomial

$$\mathbf{q}_p(t; \mathbf{q}^{0_0}, \ldots, \mathbf{q}^{0_s}, h) \in \mathbb{P}^s([0, h])$$

such that

$$\mathbf{q}_p(d_\nu h) = \mathbf{q}^{0_\nu}, \qquad \nu = 0, 1, \ldots, s.$$

See Fig. 14b. Then, the discrete Lagrangian is obtained after providing an appropriate quadrature rule, $\{w_i, \xi_i\}_{i=1,\ldots,n_q}$, as

$$L_d(\mathbf{q}^0, \mathbf{q}^1, h) = \inf_{\substack{\mathbf{q}(t)\in\mathbb{P}^s([0,h]) \\ \mathbf{q}(0)=\mathbf{q}^0, \mathbf{q}(h)=\mathbf{q}^1}} L_d^u\left(\mathbf{q}^{0_0}, \ldots, \mathbf{q}^{0_s}, h\right), \tag{62h}$$

where

$$L_d^u\left(\mathbf{q}^{0_0}, \ldots, \mathbf{q}^{0_s}, h\right) = \sum_{i=1}^{n_q} w_i L(\mathbf{q}_p(\xi_i), \dot{\mathbf{q}}_p(\xi_i)). \tag{62i}$$

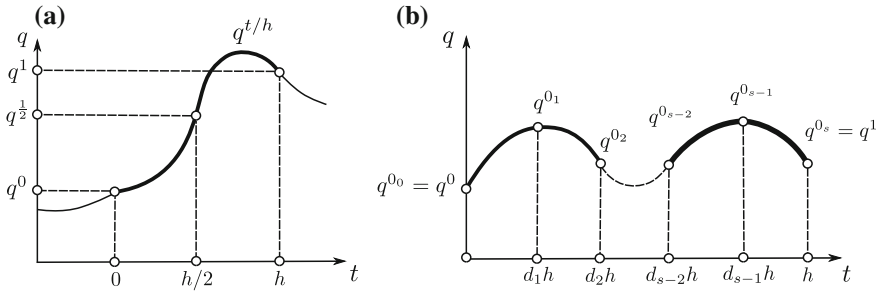


Fig. 14 **a** The polynomial $\mathbf{q}^{t/h}$ defines a continuous piecewise quadratic approximation of $\mathbf{q}(t)$ over $[0, h]$. **b** Higher order approximation of the trajectory by means of high-order polynomials

The infimization in (62h) implicitly defines the value of $\{\mathbf{q}^{0_\nu}\}_{\nu=1,\ldots,s-1}$ given those of $\mathbf{q}^0$ and $\mathbf{q}^1$. In this case the stationarity conditions are given by

$$\frac{\partial L_d^u}{\partial \mathbf{q}^{0_\nu}}\left(\mathbf{q}^{0_0}, \ldots, \mathbf{q}^{0_s}, h\right) = \mathbf{0}, \qquad \nu = 1, \ldots, s-1. \tag{63}$$

□

Example Consider the discrete Lagrangian (62e) for thermoelastic systems along with the stationarity (62f) and (62g). The DEL equations follow as

$$\frac{\mathbf{m}}{h^2}(-\mathbf{q}^{k+1} + 8\mathbf{q}^{k+\frac{1}{2}} - 14\mathbf{q}^k + 8\mathbf{q}^{k-\frac{1}{2}} - \mathbf{q}^{k-1}) = \frac{\mathbf{f}^{k+} + \mathbf{f}^{k-}}{2}, \tag{64a}$$

$$\boldsymbol{\Gamma}^{(k+1)-} - 4\boldsymbol{\Gamma}^{k+\frac{1}{2}} - 3\boldsymbol{\Gamma}^{k+} + 3\boldsymbol{\Gamma}^{k-} + 4\boldsymbol{\Gamma}^{k-\frac{1}{2}} - \boldsymbol{\Gamma}^{(k-1)+} = \mathbf{0}, \tag{64b}$$

where

$$\mathbf{f}^{k+} = -\frac{\partial \mathsf{A}}{\partial \mathbf{q}}\left(\mathbf{q}^k, \frac{-3\boldsymbol{\Phi}^k + 4\boldsymbol{\Phi}^{k+\frac{1}{2}} - \boldsymbol{\Phi}^{k+1}}{h}\right) \tag{65a}$$

$$\mathbf{f}^{k-} = -\frac{\partial \mathsf{A}}{\partial \mathbf{q}}\left(\mathbf{q}^k, \frac{\boldsymbol{\Phi}^{k-1} - 4\boldsymbol{\Phi}^{k-\frac{1}{2}} + 3\boldsymbol{\Phi}^k}{h}\right), \tag{65b}$$

$$\boldsymbol{\Gamma}^{k+\frac{1}{2}} = -\frac{\partial \mathsf{A}}{\partial \theta}\left(\mathbf{q}^{k+\frac{1}{2}}, \frac{\boldsymbol{\Phi}^{k+1} - \boldsymbol{\Phi}^k}{h}\right). \tag{65c}$$

Together, (62f), (62g), (64a) and (64b) provide enough equations to solve for $(\mathbf{q}^{k+\frac{1}{2}}, \mathbf{q}^{k+1}, \boldsymbol{\Phi}^{k+\frac{1}{2}}, \boldsymbol{\Phi}^{k+1})$ given $(\mathbf{q}^{k-\frac{1}{2}}, \mathbf{q}^k, \boldsymbol{\Phi}^{k-\frac{1}{2}}, \boldsymbol{\Phi}^k)$.

Remark Higher order methods can be constructed by means of dividing $[0, h]$ into subintervals and applying *composition methods*. The basic idea consists in combining several discrete Lagrangians together to obtain a new discrete Lagrangian with higher

order of accuracy. In (Marsden and West 2001, pp. 49–51) the methodologies to construct both *multistep* and *single step, multisubsteps* methods are presented. □

3.2 Computation of Conjugate Momenta

The discrete Lagrangian map defines an evolution over $Q \times Q$, so it does not define either velocities or conjugate momenta. To do this, it is necessary to introduce the *discrete Legendre transforms* $\mathrm{F}^{\pm}L_d \colon Q \times Q \to T^*Q$ as

$$\mathrm{F}^{+}L_d(\mathbf{q}^0, \mathbf{q}^1, h) = (\mathbf{q}^1, \mathbf{p}^1) = \left(\mathbf{q}^1, \frac{\partial L_d}{\partial \mathbf{q}^1}(\mathbf{q}^0, \mathbf{q}^1, h)\right) \tag{66}$$

$$\mathrm{F}^{-}L_d(\mathbf{q}^0, \mathbf{q}^1, h) = (\mathbf{q}^0, \mathbf{p}^0) = \left(\mathbf{q}^0, -\frac{\partial L_d}{\partial \mathbf{q}^0}(\mathbf{q}^0, \mathbf{q}^1, h)\right). \tag{67}$$

So, the discrete Legendre transforms define conjugate momenta $\mathbf{p}^0$ and $\mathbf{p}^1$ at $\mathbf{q}^0$ and at $\mathbf{q}^1$, respectively. Notice that along a discrete trajectory a single value of a conjugate momentum is defined for each $\mathbf{q}^i$, instead of possibly two, coming each of the two surrounding time-steps. This is because the discrete trajectory satisfies the DEL equations, which state that

$$\frac{\partial L_d}{\partial \mathbf{q}^i}(\mathbf{q}^{i-1}, \mathbf{q}^i, h) = -\frac{\partial L_d}{\partial \mathbf{q}^i}(\mathbf{q}^i, \mathbf{q}^{i+1}, h), \tag{68}$$

so

$$\begin{aligned}\mathrm{F}^{-}L_d(\mathbf{q}^i, \mathbf{q}^{i+1}, h) &= \left(\mathbf{q}^i, -\frac{\partial L_d}{\partial \mathbf{q}^i}(\mathbf{q}^i, \mathbf{q}^{i+1}, h)\right) \\ &= \left(\mathbf{q}^i, \frac{\partial L_d}{\partial \mathbf{q}^i}(\mathbf{q}^{i-1}, \mathbf{q}^i, h)\right) = \mathrm{F}^{+}L_d(\mathbf{q}^{i-1}, \mathbf{q}^i, h)\end{aligned}$$

and the momentum at $\mathbf{q}^i$ is uniquely defined.

The introduction of the conjugate momentum permits rewriting the DEL in the so-called position-momentum form:

$$\mathbf{p}^i = -\frac{\partial L_d}{\partial \mathbf{q}^i}(\mathbf{q}^i, \mathbf{q}^{i+1}, h) \tag{69a}$$

$$\mathbf{p}^{i+1} = \frac{\partial L_d}{\partial \mathbf{q}^{i+1}}(\mathbf{q}^i, \mathbf{q}^{i+1}, h). \tag{69b}$$

Then, given $(\mathbf{q}^i, \mathbf{p}^i)$, these equations define $(\mathbf{q}^{i+1}, \mathbf{p}^{i+1})$. The map $\boldsymbol{\Phi}_h \colon T^*Q \to T^*Q$ defined as $(\mathbf{q}^{i+1}, \mathbf{p}^{i+1}) = \boldsymbol{\Phi}_h(\mathbf{q}^i, \mathbf{p}^i)$ is the *discrete Hamiltonian map*. See Fig. 13b.

The velocities do not have an intrinsic definition in discrete Lagrangian mechanics in terms of L_d, instead the velocities are approximated by inverting the Legendre transform, namely,

$$(\mathbf{q}^i, \dot{\mathbf{q}}^i) = FL^{-1}(\mathbf{q}^i, \mathbf{p}^i) \tag{70a}$$

or

$$\dot{\mathbf{q}}^i = \frac{\partial H}{\partial \mathbf{p}}(\mathbf{q}^i, \mathbf{p}^i) \tag{70b}$$

for each i. Notice that $\dot{\mathbf{q}}^i$ defined in this way will generally be different than the approximation of the velocities that might have been used to construct the discrete Lagrangian. As we shall see later, approximating the velocity as in (70a) guarantees it will enjoy the same rate of convergence that $\mathbf{q}$ and $\mathbf{p}$ have.

Example In following we derive the position-momentum form of the DEL of some discrete Lagrangians.

- **Rectangle rules** (cases 1 and 2). The position-momentum form of the DEL corresponding to $L_d^0(\mathbf{q}^0, \mathbf{q}^1, h)$ is given by the equations

$$\begin{aligned} \mathbf{p}^0 &= -\frac{\partial}{\partial \mathbf{q}^0}\left[hL\left(\mathbf{q}^0, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right)\right] &&= \left[-h\frac{\partial L}{\partial \mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}}\right]\Bigg|_{\left(\mathbf{q}^0, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right)}, \\ \mathbf{p}^1 &= \frac{\partial}{\partial \mathbf{q}^1}\left[hL\left(\mathbf{q}^0, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right)\right] &&= \frac{\partial L}{\partial \dot{\mathbf{q}}}\left(\mathbf{q}^0, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right), \end{aligned} \tag{71}$$

  which for the case of a particle in the hoop are

$$\begin{aligned} p^0 &= mr\left(r\left(\frac{\theta^1 - \theta^0}{h}\right) + hg\cos\theta^0\right), \\ p^1 &= mr^2\left(\frac{\theta^1 - \theta^0}{h}\right), \end{aligned}$$

  therefore, the discrete Hamiltonian map is

$$(\theta^1, p^1) = \Phi_h^0(\theta^0, p^0) = \left(\theta^0 + \frac{hp^0}{mr^2} - \frac{h^2 g}{r}\cos\theta^0,\ p^0 - hmgr\cos\theta^0\right).$$

  For the case of $L_d^1(\mathbf{q}^0, \mathbf{q}^1, h)$ we have

$$\begin{aligned} \mathbf{p}^0 &= -\frac{\partial}{\partial \mathbf{q}^0}\left[hL\left(\mathbf{q}^1, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right)\right] &&= \frac{\partial L}{\partial \dot{\mathbf{q}}}\left(\mathbf{q}^1, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right), \\ \mathbf{p}^1 &= \frac{\partial}{\partial \mathbf{q}^1}\left[hL\left(\mathbf{q}^1, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right)\right] &&= \left[h\frac{\partial L}{\partial \mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}}\right]\Bigg|_{\left(\mathbf{q}^1, \frac{\mathbf{q}^1 - \mathbf{q}^0}{h}\right)}. \end{aligned} \tag{72}$$

As we mentioned earlier, both discrete Lagrangians L_d^0 and L_d^1 lead to the same DEL. However, they define different approximations to the momenta, and hence to the velocity. This is how these two discrete Lagrangians, and hence their integrators, differ.

In the case of the particle in the hoop,

$$p^0 = mr^2\left(\frac{\theta^1-\theta^0}{h}\right),$$
$$p^1 = mr\left(r\left(\frac{\theta^1-\theta^0}{h}\right) - hg\cos\theta^1\right),$$

and the discrete Hamiltonian map is given by

$$(\theta^1, p^1) = \Phi_h^1(\theta^0, p^0) = \left(\theta^0 + \frac{hp^0}{mr^2}, p^0 - hmgr\cos\theta^1\right).$$

The velocities are approximated considering (70a) as

$$\dot{\theta}^0 = FL^{-1}L(\theta^0, p^0) = \frac{p^0}{mr^2}$$
$$\dot{\theta}^1 = FL^{-1}(\theta^1, \dot{\theta}^1) = \frac{p^1}{mr^2}.$$

- **Trapezoidal rule**. If the trapezoidal rule is selected to construct the discrete Lagrangian for the particle in the hoop, $L_d^{\frac{1}{2}}(\theta^0, \theta^1, h)$, the following discrete Hamiltonian map is obtained

$$\begin{aligned}(\theta^1, p^1) = \Phi_h^{\frac{1}{2}}(\theta^0, p^0) &= \frac{1}{2}\left(\Phi_h^0(\theta^0, p^0) + \Phi_h^1(\theta^0, p^0)\right)\\ &= \left(\theta^0 + \frac{hp^0}{mr^2} - \frac{h^2 g}{2r}\cos\theta^0, p^0 - hmgr\frac{\cos\theta^1 - \cos\theta^0}{2}\right)\end{aligned}$$

- **Midpoint rule**. Consider the midpoint discrete Lagrangian (54g). The position momentum form of the DEL equations is given by

$$\mathbf{p}^0 = \begin{pmatrix} p_x^0 \\ p_y^0 \\ p_\theta^0 \end{pmatrix} = -\frac{2m}{h}\begin{pmatrix} x_{CG}^1 - x_{CG}^0 \\ y_{CG}^1 - y_{CG}^0 \\ L^2(\theta^1 - \theta^0) \end{pmatrix}, \tag{73a}$$

and

$$\mathbf{p}^1 = \begin{pmatrix} p_x^1 \\ p_y^1 \\ p_\theta^1 \end{pmatrix} = -\mathbf{p}^0. \tag{73b}$$

□

The relations (69a) and (69b) are also valid for Galerkin VI's. However, they need to be complemented with (63) in order to fulfill the condition (62h).

Example Consider the discrete Lagrangian (62e) for the thermoelastic systems. The conjugate momenta are given by

$$\mathbf{p}^k = \frac{\mathbf{m}}{3h}(8\mathbf{q}^{k+\frac{1}{2}} - 7\mathbf{q}^k - \mathbf{q}^{k+1}) - \frac{h}{6}\mathbf{f}^{k+}, \tag{74a}$$

$$\mathbf{p}^{k+1} = \frac{\mathbf{m}}{3h}(\mathbf{q}^k - 8\mathbf{q}^{k+\frac{1}{2}} + 7\mathbf{q}^{k+1}) + \frac{h}{6}\mathbf{f}^{(k+1)-}, \tag{74b}$$

$$\boldsymbol{\eta}^k = \frac{3}{6}\boldsymbol{\Gamma}^{k+} - \frac{1}{6}\boldsymbol{\Gamma}^{(k+1)-} + \frac{4}{6}\boldsymbol{\Gamma}^{k+\frac{1}{2}}, \tag{74c}$$

$$\boldsymbol{\eta}^{k+1} = \frac{3}{6}\boldsymbol{\Gamma}^{(k+1)-} - \frac{1}{6}\boldsymbol{\Gamma}^{k+} + \frac{4}{6}\boldsymbol{\Gamma}^{k+\frac{1}{2}}. \tag{74d}$$

□

3.3 Implementation of Variational Integrators

The position-momentum form of the DEL equations provide a natural way to implement a variational integrator in a computer code. We describe the general format of an implementation next.

3.3.1 Initial Conditions

Most commonly initial conditions are provided in terms of positions and velocities instead of positions and momenta as required by variational integrators in position-momentum form. Then, we take advantage of the Legendre transform to compute the initial momentum as

$$\mathbf{p}^0 = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(0), \dot{\mathbf{q}}(0)). \tag{75}$$

3.3.2 Basic Algorithm

The computer implementation of variational methods in position-momentum form follows a general structure summarized in Algorithm 1.

3.3.3 Post-processing Velocities

Consistent approximations to the velocities can be computed with the help of (35a) as

Data: Require $(\mathbf{q}^0, \mathbf{p}^0)$, h and N.
forall the $k = 0, 1, \ldots, N-1$ **do**
  Solve $\mathbf{p}^k = -\dfrac{\partial L_d}{\partial \mathbf{q}^k}(\mathbf{q}^k, \mathbf{q}^{k+1}, h)$ for $\mathbf{q}^{n+1}$.
  Set $\mathbf{p}^{k+1} = \dfrac{\partial L_d}{\partial \mathbf{q}^{k+1}}(\mathbf{q}^k, \mathbf{q}^{k+1}, h)$.
end

Algorithm 1: Basic implementation of variational integrators.

$$\dot{\mathbf{q}}^k = \frac{\partial H}{\partial \mathbf{p}}(\mathbf{q}^k, \mathbf{p}^k).$$

3.4 Approximation Properties and Convergence

In contrast to the traditional approach to constructing time integrators, which begins by approximating the equations of motion, the construction of a variational integrator departs from an approximation of the action. The order of convergence of a traditional integrator is generally assessed by the order of the consistency error. The question is then how the order of convergence of a variational integrator can be determined from the approximation of the action. This question was answered in (Marsden and West 2001, Sect. 2.3), and we describe the main ideas next.

First we define the *exact discrete Lagrangian*

$$L_d^E\left(\mathbf{q}^0, \mathbf{q}^1, h\right) = \int_0^h L\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)\, dt,$$

where $\mathbf{q}(t)$ is the solution of the E–L equations satisfying $\mathbf{q}(0) = \mathbf{q}^0$ and $\mathbf{q}(h) = \mathbf{q}^1$. In other words, L_d^E is a discrete Lagrangian that exactly matches the value of the action for the exact trajectory in the time interval $[0, h]$.

We can now define the *local variational order*. The discrete Lagrangian L_d is of order $r \geq 1$ if for any solution $\mathbf{q}$ of the E–L equations there exists $h_v > 0$ and $C_v > 0$ independent of h such that

$$\left\| L_d(\mathbf{q}(0), \mathbf{q}(h), h) - L_d^E(\mathbf{q}(0), \mathbf{q}(h), h) \right\| \leq C_v h^{r+1}, \tag{76}$$

for all $0 < h < h_v$.[6]

[6]Notice that C_v can depend on $(\mathbf{q}(0), \dot{\mathbf{q}}(0))$. For simplicity, we deliberately avoided the additional requirement that C_v should be uniformly bounded over a subset of TQ (see Marsden and West 2001). In designing a variational integrator and evaluating its order, this is a secondary condition not difficult to satisfy.

We are now ready to answer the question we started from: A fundamental result presented in (Marsden and West 2001, Theorem 2.3.1, pp. 43–44) states that *the variational integrator obtained from a discrete Lagrangian of order $r+1$ has order r*. Thus, to design a variational integrator of order r it is enough to construct a discrete Lagrangian of order $r+1$. We show examples of order calculation below.

3.4.1 Order Calculation

To compute the order of $L_d(\mathbf{q}(0), \mathbf{q}(h), h)$, we expand it in a Taylor series of h around $h = 0$ and compare the terms with those of the Taylor series expansion of the exact discrete Lagrangian. The first few terms of the latter are

$$L_d^E(\mathbf{q}(0), \mathbf{q}(h), h) = hL(\mathbf{q}(0), \dot{\mathbf{q}}(0)) + \\ + \frac{h^2}{2}\left(\frac{\partial L}{\partial \mathbf{q}}(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \cdot \dot{\mathbf{q}}(0) + \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \cdot \ddot{\mathbf{q}}(0)\right) + \mathcal{O}(h^3). \quad (77)$$

If the first r terms of the series of both discrete Lagrangian are the same, then the discrete Lagrangian is of order $r+1$.

Example Consider the discrete Lagrangian built on the trapezoidal rule,

$$L_d^\alpha(\mathbf{q}(0), \mathbf{q}(h), h) = f(h) = \\ (1-\alpha)hL\left(\mathbf{q}(0), \frac{\mathbf{q}(h) - \mathbf{q}(0)}{h}\right) + \alpha hL\left(\mathbf{q}(h), \frac{\mathbf{q}(h) - \mathbf{q}(0)}{h}\right).$$

Here we introduced the name $f(h)$ to explicitly indicate that the left-hand side is a function of h only. Then,

$$\begin{aligned} f(0) &= L_d^\alpha(\mathbf{q}(0), \mathbf{q}(0), 0) &&= 0, \\ f'(0) &= \frac{dL_d^\alpha}{dh}(\mathbf{q}(0), \mathbf{q}(0), 0) &&= L(\mathbf{q}(0), \dot{\mathbf{q}}(0)), \\ f''(0) &= \frac{d^2L_d^\alpha}{dh^2}(\mathbf{q}(0), \mathbf{q}(0), 0) &&= 2\alpha\frac{\partial}{\partial \mathbf{q}}L(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \cdot \dot{\mathbf{q}}(0) + \frac{\partial}{\partial \dot{\mathbf{q}}}L(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \cdot \ddot{\mathbf{q}}(0), \end{aligned}$$

therefore,

$$\begin{aligned} L_d^\alpha(\mathbf{q}(0), \mathbf{q}(h), h) = hL(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \\ + \frac{h^2}{2}\left[2\alpha\frac{\partial}{\partial \mathbf{q}}L(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \cdot \dot{\mathbf{q}}(0)\right. \\ \left. + \frac{\partial}{\partial \dot{\mathbf{q}}}L(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \cdot \ddot{\mathbf{q}}(0)\right] + \mathcal{O}\left(h^3\right), \end{aligned} \quad (78)$$

Comparing with (77) have that

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

and hence $r+1=3$ if and only if $\alpha=1/2$, and $r+1=2$ otherwise. This means that the two rectangle rules ($\alpha=0,1$) are only first-order integrators, while $\alpha=1/2$ gives rise to a second-order algorithm. A rather curious aspect of this last remark is that, as mentioned earlier, all the aforementioned algorithms give rise to the *same* DEL, but they differ on the discrete Legendre transform, or the definition of the discrete momenta. Thus, while the coordinate values $\mathbf{q}^i$ coincide in all three algorithms, the momenta $\mathbf{p}^i$ do not, and this is where the order difference between the algorithms comes from.

The midpoint rule gives rise to the same expansion (78) with $\alpha=1/2$, so its a second-order algorithm.

3.5 Conservation Properties: Discrete Point of View

This section focuses on the geometric properties of the flows generated by variational integrators. As we highlighted before, they correspond to symplectic flows that show an excellent long-term energy behavior along with the exact conservation of the invariants associated to the symmetries of the discrete Lagrangian. These properties along with the existence of a standard methodology to construct high-order methods for Lagrangian systems evolving on general manifolds have contributed to increase the use of VI's in both the scientific and engineering communities.

- **Symmetries and invariants of the dynamics**. A discrete Lagrangian posses a symmetry when it remains invariant under the action of a group on the configuration manifold. Moreover, each symmetry of the discrete Lagrangian leads to a quantity conserved by the dynamics (Marsden and West 2001) according to a discrete version of the celebrated Noether's theorem. This theorem reads as follows,

Discrete version of Noether's theorem. Consider a discrete Lagrangian $L_d(\mathbf{q}^k,\mathbf{q}^{k+1},h)$ and a one-parameter group of discrete curves $\{\mathbf{q}^{\epsilon,k}\}_{k=0,\dots,N}$ with $\epsilon>0$ and $\mathbf{q}^{0,k}=\mathbf{q}^k$, that leaves the discrete Lagrangian invariant in the following sense,

$$L_d\left(\mathbf{q}^{\epsilon,k},\mathbf{q}^{\epsilon,k+1},h\right)=L_d\left(\mathbf{q}^k,\mathbf{q}^{k+1},h\right),$$

for all $\epsilon>0$ and $k=0,\dots,N-1$. Moreover, consider the infinitesimal symmetry direction,

$$\zeta(\mathbf{q}^k)=\left.\frac{d}{d\epsilon}\mathbf{q}^{\epsilon,k}\right|_{\epsilon=0}.$$

Then

$$I(\mathbf{q}^k, \mathbf{p}^k) = \mathbf{p}^k \cdot \boldsymbol{\zeta}(\mathbf{q}^k) \quad \text{with} \quad \mathbf{p}^k = \frac{\partial L_d}{\partial \mathbf{q}^k}(\mathbf{q}^{k-1}, \mathbf{q}^k, h),$$

is an invariant of the dynamics for all $k = 0, ..., N$.

Proof See, e.g., (Lew et al. 2004) or (Hairer et al. 2006, Chap. VI.6) □

Example Consider the thermoelastic system described in page[7] 20. We know the Lagrangian function is invariant under rigid body translations and rigid body rotations in the physical space and under rigid body translations of the thermal displacements. Therefore, total linear momentum, the total angular momentum and the entropy of each thermoelastic spring are invariants of the dynamics.

We construct a discrete Lagrangian by means of applying the midpoint rule as

$$L_d^m\left(\mathbf{q}^k, \mathbf{q}^{k+1}, \boldsymbol{\Phi}^k, \boldsymbol{\Phi}^{k+1}, h\right) = \frac{h}{2}\left(\frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h}\right) \cdot \mathbf{M}\left(\frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h}\right) - \\ -h\mathsf{A}\left(\frac{\mathbf{q}^k + \mathbf{q}^{k+1}}{2}, \frac{\boldsymbol{\Phi}^{k+1} - \boldsymbol{\Phi}^k}{h}\right). \tag{79}$$

First, we consider the one-parameter group of discrete curves

$$\left\{\mathbf{q}_i^{\epsilon,k}(t)\right\}_{k=0,...,N_t} = \left\{\mathbf{q}_i^k(t) + \epsilon\mathbf{v}\right\}_{k=0,...,N_t}$$
$$\left\{\boldsymbol{\Phi}^{k,\epsilon}(t)\right\}_{k=0,...,N_t} = \left\{\boldsymbol{\Phi}^k(t) + \epsilon\mathbf{k}\right\}_{k=0,...,N_t},$$

where $\mathbf{v} \in \mathbb{R}^3$ and $\mathbf{k} \in \mathbb{R}^M$ are constant but otherwise arbitrary vectors. The corresponding infinitesimal directions are given by

$$\boldsymbol{\xi}^k = (\mathbf{v}_N, \mathbf{k}), \qquad k = 0, \ldots, N_t,$$

where $\mathbf{v}_N^k \in \mathbb{R}^d$. Noticing that

$$\mathbf{q}^{\epsilon,k+1} - \mathbf{q}^{\epsilon,k} = \mathbf{q}^{k+1} - \mathbf{q}^k$$
$$\boldsymbol{\Phi}^{\epsilon,k+1} - \boldsymbol{\Phi}^{\epsilon,k} = \boldsymbol{\Phi}^{k+1} - \boldsymbol{\Phi}^k,$$

and considering that the distance between masses is conserved by rigid body translations in space, i.e.,

$$l_i\left(\frac{\mathbf{q}^{\epsilon,k+1} + \mathbf{q}^{\epsilon,k}}{2}\right) = l_i\left(\frac{\mathbf{q}^{k+1} + \mathbf{q}^k}{2}\right), \qquad i = 1, \ldots, N,$$

[7] In this example N_t denotes the total number of time instants of the discrete trajectory and N is reserved for the number of masses of the thermoelastic system.

it is possible to verify that the discrete Lagrangian remains invariant and thus $\left\{\mathbf{q}^{\epsilon,k}, \Phi^{\epsilon,k}\right\}_{k=0,\ldots,N_t}$ is one of its symmetries.

The momentum vector is given by

$$\left(\mathbf{p}^k, \boldsymbol{\eta}^k\right),$$

where

$$\begin{aligned}\mathbf{p}^k &= \mathbf{M}\left(\frac{\mathbf{q}^{k+1}-\mathbf{q}^k}{h}\right) - \frac{h}{2}\frac{\partial \mathsf{A}}{\partial \mathbf{q}}\left(\frac{\mathbf{q}^{k+1}+\mathbf{q}^k}{2}, \frac{\boldsymbol{\Phi}^{k+1}-\boldsymbol{\Phi}^k}{h}\right)\\ \boldsymbol{\eta}^k &= -\frac{\partial \mathsf{A}}{\partial \boldsymbol{\theta}}\left(\frac{\mathbf{q}^{k+1}+\mathbf{q}^k}{2}, \frac{\boldsymbol{\Phi}^{k+1}-\boldsymbol{\Phi}^k}{h}\right),\end{aligned}$$

and according to the discrete version of Noether's theorem

$$\mathbf{v}_N \cdot \mathbf{p}^k + \mathbf{k} \cdot \boldsymbol{\eta}^k,$$

is conserved for all $k = 1, \ldots, N_t$ and $(\mathbf{v}_N, \mathbf{k}) \in \mathbb{R}^d \times \mathbb{R}^M$. Choosing

$$(\mathbf{v}_N, \mathbf{k}) = ((\boldsymbol{e}_i, \ldots, \boldsymbol{e}_i), \mathbf{0}_M), \quad i = 1, 2, 3$$

where $\{\boldsymbol{e}_i\}_{i=1,2,3}$ is a basis in $\mathbb{R}^3$ and $\mathbf{0}_M$ is a M-dimensional vector of zeros, allows to deduce that every component of the total linear momentum is exactly conserved. Moreover, setting

$$(\mathbf{v}_N, \mathbf{k}) = \left(\mathbf{0}_N, \mathbf{k}^j\right),$$

where $\mathbf{k}_i^j = \delta_i^j$ yields to the exact conservation of the entropy of every thermoelastic spring.

Alternatively, we can consider a one-parameter group of discrete curves corresponding to rigid body rotations in the physical space, namely

$$\left\{\mathbf{q}_i^{\epsilon,k}(t)\right\}_{k=0,\ldots,N_t} = \left\{\exp\left[\epsilon\widetilde{\boldsymbol{\omega}}\right]\mathbf{q}_i^k(t)\right\}_{k=0,\ldots,N_t} \tag{80a}$$

$$\left\{\boldsymbol{\Phi}^{k,\epsilon}(t)\right\}_{k=0,\ldots,N_t} = \left\{\boldsymbol{\Phi}^k(t)\right\}_{k=0,\ldots,N_t}, \tag{80b}$$

where $\widetilde{\boldsymbol{\omega}}$ is a constant and skew-symmetric but otherwise arbitrary tensor with axial vector $\widehat{\boldsymbol{\omega}} \in \mathbb{R}^3$. The corresponding infinitesimal symmetry direction is given by

$$\boldsymbol{\zeta}(\mathbf{q}^k) = \left\{\left(\left(\widehat{\boldsymbol{\omega}} \times \mathbf{q}_1^k, \ldots, \widehat{\boldsymbol{\omega}} \times \mathbf{q}_N^k\right), \mathbf{0}_M\right)\right\}_{k=0,\ldots,N_t}.$$

The discrete Lagrangian (79) is invariant under the transformations defined in (80a) and (80b) for all $\epsilon > 0$, since both the discrete version of the kinetic energy and the distance among the masses of the system remain unaffected by rigid body motions in the physical space.

Then, according to the discrete version of Noether's theorem

$$\mathbf{p}^k \cdot \boldsymbol{\xi}(\mathbf{q}^k),$$

is exactly conserved by the dynamics for all $k = 0, \ldots, N_t$. The above expression can be rewritten as

$$\mathbf{A}^k \cdot \widehat{\mathbf{W}} = \sum_{i=1}^{N} \left(\mathbf{q}_i^k \times \left[\mathbf{M} \left(\frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} \right) - \frac{\partial \mathsf{A}}{\partial \mathbf{q}_i^k} \left(\mathbf{q}^{k+\frac{1}{2}}, \theta^k \right) \right] \right) \cdot \widehat{\boldsymbol{\omega}},$$

where

$$\widehat{\mathbf{W}} := \underbrace{(\widehat{\boldsymbol{\omega}}, \ldots, \widehat{\boldsymbol{\omega}})}_{N \text{ times}}, \quad \mathbf{q}^{k+\frac{1}{2}} = \frac{\mathbf{q}^{k+1} + \mathbf{q}^k}{2} \quad \text{and} \quad \theta^k = \frac{\boldsymbol{\Phi}^{k+1} - \boldsymbol{\Phi}^k}{h}.$$

Therefore, since $\widehat{\boldsymbol{\omega}}$ is an arbitrary vector in $\mathbb{R}^3$, it is possible to conclude that a discrete version of the total angular momentum, $\mathbf{A}^k$, remains invariant. □

- **Discrete symplecticity**. In Sect. 3.2 we described how the discrete Legendre transform allows to define the discrete Hamiltonian map

$$\begin{aligned} \boldsymbol{\Phi}_h : \quad T^*Q \quad &\to \quad T^*Q \\ (\mathbf{q}^n, \mathbf{p}^n) &\mapsto \boldsymbol{\Phi}_h(\mathbf{q}^n, \mathbf{p}^n) = (\mathbf{q}^{n+1}, \mathbf{p}^{n+1}) \end{aligned},$$

  which can be used to construct the position-momentum form of a variational method. In this section we show that $\boldsymbol{\Phi}_h$ also defines a discrete symplectic flow on T^*Q.

Remark Given a particular discrete Lagrangian, to demonstrate that its discrete Hamiltonian map is symplectic, it is enough to verify that $\boldsymbol{\Phi}_h$ fulfils (50), i.e.,

$$\left(\frac{\partial \boldsymbol{\Phi}_h}{\partial \mathbf{y}^k}(\mathbf{y}^k) \right)^t \mathbf{J}^{-1} \left(\frac{\partial \boldsymbol{\Phi}_h}{\partial \mathbf{y}^k}(\mathbf{y}^k) \right) = \mathbf{J}^{-1},$$

where $\mathbf{y}^k = (\mathbf{q}^k, \mathbf{p}^k) \in T^*Q$. □

However, to prove this in a more general setting, we consider the following result:

Theorem *Any smooth enough and nondegenerate function $S(\mathbf{q}, \mathbf{Q})$ generates a symplectic flow $(\mathbf{q}, \mathbf{p}) \mapsto (\mathbf{P}, \mathbf{Q})$ if*

$$\mathbf{p} = -\frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \mathbf{Q}) \quad \textit{and} \quad \mathbf{P} = \frac{\partial S}{\partial \mathbf{Q}}(\mathbf{q}, \mathbf{Q}). \tag{81}$$

Proof See (Hairer et al. 2006, pp. 196–197) □

The function $S(\mathbf{q}, \mathbf{Q})$ is a particular case of the so-called *generating functions*.

Consider a sequence of points $\{\mathbf{q}^i\}_{i=0,\ldots,N}$ on Q that is the solution of the DEL equations (58) subjected to boundary conditions $\mathbf{q}^0$ and $\mathbf{q}^N$. Then the discrete action sum (55) can be regarded as a function of the initial and final configuration points, i.e.,

$$S_d\left(\mathbf{q}^0, \mathbf{q}^N\right) = \sum_{i=0}^{N-1} L_d\left(\mathbf{q}^i, \mathbf{q}^{i+1}, h\right).$$

Taking into account the discrete Legendre transformations (69a) and (69b), we have that

$$\begin{aligned}
\frac{\partial S_d}{\partial \mathbf{q}^0}\left(\mathbf{q}^0, \mathbf{q}^N\right) &= \frac{\partial L_d}{\partial \mathbf{q}^0}\left(\mathbf{q}^0, \mathbf{q}^1, h\right) = -\mathbf{p}^0, \\
\frac{\partial S_d}{\partial \mathbf{q}^N}\left(\mathbf{q}^0, \mathbf{q}^N\right) &= \frac{\partial L_d}{\partial \mathbf{q}^N}\left(\mathbf{q}^{N-1}, \mathbf{q}^N, h\right) = \mathbf{p}^N,
\end{aligned}$$

and therefore applying Theorem 3.5, the flow $(\mathbf{q}^0, \mathbf{p}^0) \mapsto (\mathbf{q}^N, \mathbf{p}^N)$ results to be symplectic. This result can be applied to an arbitrary time interval $[t^i, t^{i+1}], i = 0, \ldots, N$, which shows that the discrete flow of any variational integrator is automatically symplectic and that the corresponding generating function is the discrete Lagrangian $L_d(\mathbf{q}^i, \mathbf{q}^{i+1}, h)$.

Example (Hairer et al. 2006, pp. 190). Consider the midpoint discrete Lagrangian,

$$L_d(\mathbf{q}^0, \mathbf{q}^1, h) = \frac{1}{2h}\mathbf{M}(\mathbf{q}^1 - \mathbf{q}^0) \cdot (\mathbf{q}^1 - \mathbf{q}^0) + hU\left(\mathbf{q}^{\frac{1}{2}}\right),$$

where $\mathbf{q}^{\frac{1}{2}} = \frac{1}{2}(\mathbf{q}^1 + \mathbf{q}^0)$. The corresponding variational integrator in position-momentum form is given by

$$\begin{aligned}
\mathbf{q}^1 &= \mathbf{q}^0 + h\,\mathbf{M}^{-1}\mathbf{p}^{\frac{1}{2}} \\
\mathbf{p}^1 &= \mathbf{p}^0 - h\,\frac{\partial U}{\partial \mathbf{q}}\left(\mathbf{q}^{\frac{1}{2}}\right),
\end{aligned}$$

which can be rewritten as

$$z^1 = z^0 + h\,\mathbf{J}\frac{\partial H}{\partial z}\left(z^{\frac{1}{2}}\right),$$

where $z = (\mathbf{q}, \mathbf{p})$ and $H(z)$ is the Hamiltonian function of the problem. Differentiating the above equation yields

$$\left(\mathbf{I} - \frac{h}{2}\mathbf{J}\frac{\partial^2 H}{\partial z^2}\right)\frac{\partial z^1}{\partial z^0} = \left(\mathbf{I} + \frac{h}{2}\mathbf{J}\frac{\partial^2 H}{\partial z^2}\right),$$

from which it is clear that

$$\left(\frac{\partial z^1}{\partial z^0}\right)^t \mathbf{J}^{-1} \left(\frac{\partial z^1}{\partial z^0}\right) = \mathbf{J}^{-1}.$$

□

- **Long-term energy behavior**. As it has been explained in Sect. 2.5 the flow on the phase space of an autonomous Hamiltonian system is constrained to remain on a constant energy manifold which depends on the initial conditions. Unfortunately, the discrete flows generated by symplectic methods with constant time step cannot conserve exactly the energy of the original Hamiltonian system (Ge and Marsden 1988; Kane et al. 1999). However, in spite of this limitation, they show an excellent long-term behavior with errors in the energy that remain bounded for exponentially long periods of time. See, e.g., (Marsden and West (2001)). In following, we explain the reasons for the superior behavior of symplectic methods.

  In Fig. 15 the numerical flow of the Symplectic-Euler method (obtained from the discrete Lagrangian (54a)) is compared with the numerical flow of the Explicit-Euler method, when both methods are used to simulate the dynamics of the system described by the Lagrangian

$$L(q(t), \dot{q}(t)) = \frac{1}{2}\dot{q}(t)^2 - 5(q(t) - 1)^2,$$

  subjected to the initial conditions $q(0) = 2, \dot{q}(0) = 0$. This system is conservative and thus there should be no loss of energy over time. This figure shows that while

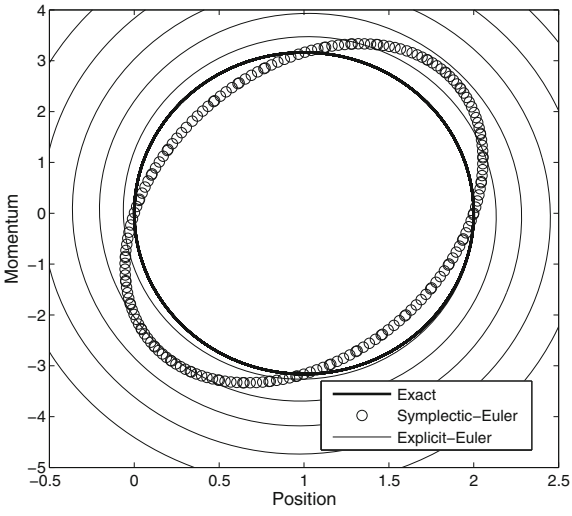


Fig. 15 Comparison between the Symplectic-Euler and the Explicit-Euler methods. Trajectories on the phase space

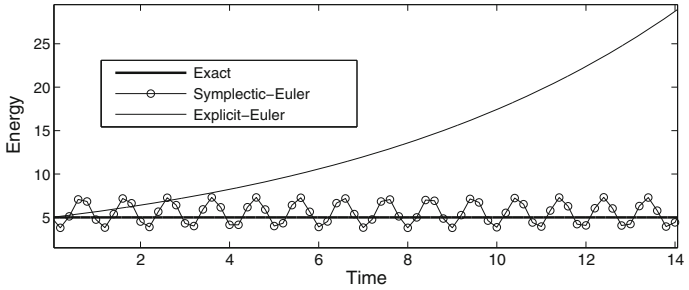


Fig. 16 Comparison between the Symplectic-Euler and the Explicit-Euler methods in terms of the numerically computed energy

the trajectory of the Explicit-Euler method departs progressively from the exact one, the trajectory of the Symplectic-Euler remains bounded and close to the exact trajectory.

In Fig. 16 a plot of the energy

$$H(q(t),\, p(t)) = \frac{1}{2} m^{-1} p(t)^2 + 5(q(t) - 1)^2,$$

evaluated on the discrete trajectories versus time is shown. The striking aspect of this graph is that while the energy associated with the Explicit-Euler method blows up due to numerical instability, for the Symplectic-Euler method the energy error remains bounded over a long period of time.

Backward error analysis. For a better understanding the above results, we may resort to use the so-called backward error analysis applied to symplectic methods (Marsden and West 2001; Lew et al. 2004; Faltinsen 2000).

Consider a numerical method, here represented by its discrete flow map $\widehat{\mathbf{\Phi}}_h : Q \to Q$, which we use to approximate the flow of the differential equation

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}), \quad \mathbf{y}(0) = \mathbf{y}^0 \in Q.$$

The basic idea of backward error analysis consist constructing *modified differential equation*

$$\widetilde{\mathbf{y}} = \mathbf{f}_h(\widetilde{\mathbf{y}}) = \mathbf{f}(\widetilde{\mathbf{y}}) + h\mathbf{f}_2(\widetilde{\mathbf{y}}) + h^2\mathbf{f}_3(\widetilde{\mathbf{y}}) + \ldots \tag{82}$$

such that its exact solution trajectory $\widetilde{\mathbf{y}}(t)$ exactly matches the discrete flow of the numerical method, i.e.,

$$\mathbf{y}^k = \widehat{\mathbf{\Phi}}_{kh}(\mathbf{y}^0) = \widetilde{\mathbf{y}}(kh), \quad k = 1, 2, \ldots, N.$$

Then, the numerical analysis focuses on studying the difference between $\mathbf{f}(\mathbf{y})$ and $\mathbf{f}_h(\mathbf{y})$ in an appropriate norm instead of studying the difference between $\mathbf{y}^k$ and $\widehat{\mathbf{\Phi}}_{kh}(\mathbf{y}^0)$ which is the focus of the more traditional *forward error analysis*. In practice the series (82) diverges and has to be truncated after a finite number of terms. Therefore, this approach allows to interpret the numerical solution of a differential equation as a *higher order approximation* of a modified system. We can now understand why variational integrators are different to standard methods. To this end we first consider the next theorem, which for sake of simplicity is restricted to Hamiltonian functions taking arguments in $\mathbb{R}^{2d}$.

Theorem *The modified equation of a symplectic method $\mathbf{\Phi}_h$ applied to a Hamiltonian system with a smooth Hamiltonian $H : \mathbb{R}^{2d} \to \mathbb{R}$ is also Hamiltonian. It means that there exist smooth functions $H_j : \mathbb{R}^{2d} \to \mathbb{R}$ for $j = 2, 3, \ldots$, such that*

$$\dot{\mathbf{y}} = \mathbf{J}\left(\frac{\partial H}{\partial \mathbf{y}}(\mathbf{y}) + h\frac{\partial H_2}{\partial \mathbf{y}}(\mathbf{y}) + h^2\frac{\partial H_3}{\partial \mathbf{y}}(\mathbf{y}) + \cdots\right).$$

Proof See, e.g., (Hairer et al. 2006). □

In other words, $\mathbf{\Phi}_h$ is a higher order approximation to the flow of the dynamical system defined by a *shadow Hamiltonian*

$$\widetilde{H}_h(\mathbf{y}) = H(\mathbf{y}) + \sum_{i=2}^{N} h^{i-1} H_i(\mathbf{y}),$$

which remains at least $\mathcal{O}(h)$ close to H.

Since every Lagrangian system admits a Hamiltonian representation, the modified differential equation of every variational integrator is Hamiltonian. This means that the discrete trajectory has all of the properties of a conservative mechanical system, such as energy conservation. This property explains the shape of the closed trajectory described by the Symplectic-Euler in Fig. 15. This also explains why the energy plots for variational integrators contain a typical oscillation about a value close to the true energy. See Fig. 16. The modified energy level set will be close to the true energy level set everywhere, but it will typically be inside it at some locations and outside it at others.

Finally, we mention a few words regarding to how much the discrete trajectory moves away from both the exact energy manifold and the constant energy manifold defined by the shadow Hamiltonian. If we apply an order p numerical method $\mathbf{\Phi}_h$ with step size h to approximate the flow of a Hamiltonian system with analytic $H : D \subset \mathbb{R}^{2d} \to \mathbb{R}$, and if the numerical solution stays in the compact set $K \subset D$, then there exist h_0 and $N = N(h)$ such that

$$\widetilde{H}(\mathbf{y}^k) = \widetilde{H}(\mathbf{y}^0) + \mathcal{O}\left(e^{h_0/2h}\right),$$
$$H(\mathbf{y}^k) = H(\mathbf{y}^0) + \mathcal{O}(h^p),$$

over exponentially long time intervals $kh \leq e^{h_0/2h}$. See (Hairer et al. 2006, Chap. IX) for details.

Therefore, it is possible to see that the discrete trajectory remains exponentially close to the constant energy manifold defined by the shadow Hamiltonian for exponentially longs periods of time. Moreover, it also remains $\mathcal{O}(h^p)$ close to the manifold of exact energy.

4 Final Examples

In this section, we formulate variational time integrators for two problems of practical interest in science and engineering. First, we develop a second-order method based on the trapezoidal rule for a free-flying body that is able to undergo arbitrarily large rotations and displacements in space. This problem has been extensively studied from both theoretical and numerical aspects (see, e.g., Meyer et al. 2009; Bauchau and Bottasso 1999; Chaturvedi et al. 2011; Lee et al. 2007; Marsden and Ratiu 1999; Simo and Wong 1991) among others) since its configuration space corresponds to a nonlinear differentiable manifold rather than a linear space. The second example corresponds to the formulation of an explicit, second-order accurate varitational integrator for finite element discretizations of geometrically exact rods. We only consider linear finite elements in space since a more general formulation can be consulted in (Mata 2015). In both examples the time interval of interest $[0, T]$ is partitioned into $N > 1$ subintervals with constant time step $\Delta t = T/N$, and we set $t^k = kT/N$, $k = 0, 1, ..., N$.

4.1 Rotating Rigid Body

Consider an inertial reference frame $\{\boldsymbol{e}_i\}_{i=1,2,3}$ in the three-dimensional space and a rigid body $\mathcal{B}$ with mass density $\rho > 0$ which has rigidly attached to its center of mass an orthogonal reference frame $\{\boldsymbol{t}_i\}_{i=1,2,3}$. See Fig. 17. The orientation of the body-fixed frame with respect to the inertial frame is specified by means of a rotation tensor $\boldsymbol{\Lambda}$ according to

$$\boldsymbol{t}_i = \boldsymbol{\Lambda}\boldsymbol{e}_i, \qquad i = 1, 2, 3.$$

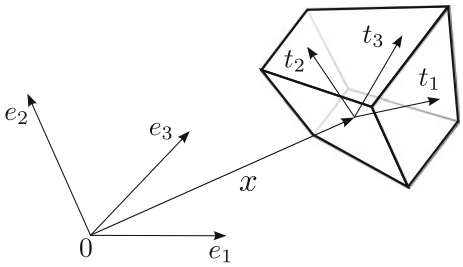


Fig. 17 Free-flying rigid body. $\{\boldsymbol{e}_i\}_{i=1,2,3}$ is a inertial reference frame and $\{\boldsymbol{t}_i\}_{i=1,2,3}$ an orthogonal body-fixed reference frame

The position vector of a material point of $\mathcal{B}$ is given by

$$\boldsymbol{y} = \boldsymbol{x} + \sum_{i=1}^{3} \xi_i \boldsymbol{t}_i = \boldsymbol{x} + \sum_{i=1}^{3} \xi_i \boldsymbol{\Lambda} \boldsymbol{e}_i,$$

where $\boldsymbol{\xi} = (\xi^1, \xi^2, \xi^3)$ is a set of coordinates with respect to $\{\boldsymbol{t}_i\}_{i=1,2,3}$ and $\boldsymbol{x} = x^1\boldsymbol{e}_1 + x^2\boldsymbol{e}_2 + x^3\boldsymbol{e}_3$ is the position vector of the center of mass of the body. The spatial position of the body is specified by the pair $(\boldsymbol{x}, \boldsymbol{\Lambda})$ which is composed of a position vector plus a rotation tensor measuring the deviation of body-fixed frame with respect to the inertial reference frame.

Note that although $\boldsymbol{x}$ belongs to $\mathbb{R}^3$ which is a linear space, $\boldsymbol{\Lambda}$ is an element of the noncommutative (Lie) group of proper rotations

$$SO(3) = \left\{ \boldsymbol{\sigma} \in \mathbb{R}^{3\times 3} \;\mid\; \boldsymbol{\sigma}^{-1} = \boldsymbol{\sigma}^t \quad \text{and} \quad \det[\boldsymbol{\sigma}] = 1 \right\},$$

which is a nonlinear manifold. A brief introduction to finite rotations is given in Sect. B. Then, the configuration the manifold is $SE(3) = \mathbb{R}^3 \times SO(3)$.

A motion of the body can be described by means of the time-dependent curve

$$\boldsymbol{\Phi} = (\boldsymbol{x}, \boldsymbol{\Lambda}) : [0, T] \to SE(3), \tag{83}$$

with velocity given by

$$\dot{\boldsymbol{\Phi}} = (\dot{\boldsymbol{x}}, \dot{\boldsymbol{\Lambda}}) \in T_{(\boldsymbol{x},\boldsymbol{\Lambda})} SE(3) = \mathbb{R}^3 \times T_{\boldsymbol{\Lambda}} SO(3).$$

For free-flying bodies the Lagrangian function $L : TSE(3) \to \mathbb{R}$ is equal to the kinetic energy

$$L(\boldsymbol{\Phi}, \dot{\boldsymbol{\Phi}}) = \frac{1}{2} \left(m\, \dot{\boldsymbol{x}} \cdot \dot{\boldsymbol{x}} + \operatorname{tr}\left[\widetilde{\boldsymbol{\Omega}} \mathbf{J}_d \widetilde{\boldsymbol{\Omega}}^t \right] \right), \tag{84}$$

where $m = \int_{\mathcal{B}} \rho \, dv$ is the total mass of the body,

$$\widetilde{\boldsymbol{\Omega}} = \boldsymbol{\Lambda}^t \dot{\boldsymbol{\Lambda}},$$

is the angular velocity tensor expressed in the body-fixed frame which belongs to the (linear) space of skew-symmetric tensors $so(3)$, and

$$\mathbf{J}_d = \int_{\mathcal{B}} \rho\, \boldsymbol{\xi} \otimes \boldsymbol{\xi} dv,$$

is a nonstandard moment of inertia tensor which is related to the standard symmetric moment of inertia tensor, $\mathbf{J}$, by

$$\mathbf{J} = \int_{\mathcal{B}} \rho\, \widetilde{\boldsymbol{\xi}}^t \widetilde{\boldsymbol{\xi}}\, dv = \mathrm{tr}\,[\mathbf{J}_d]\,\mathbf{I} - \mathbf{J}_d,$$

where $\widetilde{\boldsymbol{\xi}} = \mathrm{skew}[\boldsymbol{\xi}] \in so(3)$ is the skew-symmetric tensor obtained from $\boldsymbol{\xi} \in \mathbb{R}^3$. See Sect. B.3. More details about this relation can be found in (Lee et al. 2007).

The application of Hamilton's principle requires computing

$$\Big\langle \delta S[\boldsymbol{\Phi}(t)], \delta\boldsymbol{\Phi}(t) \Big\rangle = \frac{dL}{d\epsilon}\left(\boldsymbol{x}_\epsilon(t), \boldsymbol{\Lambda}_\epsilon(t), \dot{\boldsymbol{x}}_\epsilon(t), \dot{\boldsymbol{\Lambda}}_\epsilon(t)\right) = 0, \tag{85}$$

where $\delta\boldsymbol{\Phi}(t) = (\delta\boldsymbol{x}(t), \delta\boldsymbol{\Lambda}(t))$ represents a variation over an arbitrary element $\boldsymbol{\Phi}(t)$ belonging to the set $\mathcal{C}$ composed by all the smooth enough trajectories of the form (83) that leaves the endpoints of the trajectory fixed. This is by no means a trivial task owing to the nonlinear nature of $SE(3)$. On one hand, we have that

$$\begin{aligned}
\delta\boldsymbol{x}(t) &= \frac{d}{d\epsilon}\left(\boldsymbol{x}(t) + \epsilon\boldsymbol{u}(t)\right)\Big|_{\epsilon=0} = \boldsymbol{u}(t),\\
\delta\boldsymbol{\Lambda}(t) &= \frac{d}{d\epsilon}\left(\exp\left[\epsilon\widetilde{\boldsymbol{\Theta}}(t)\right]\boldsymbol{\Lambda}(t)\right)\Big|_{\epsilon=0} = \widetilde{\boldsymbol{\Theta}}(t)\boldsymbol{\Lambda}(t),
\end{aligned}$$

where

$$\boldsymbol{u} : [0, T] \to \mathbb{R}^3 \quad \text{and} \quad \widetilde{\boldsymbol{\Theta}} : [0, T] \to so(3).$$

Therefore, a variation over $\boldsymbol{\Phi}(\cdot)$ is given by

$$\delta\boldsymbol{\Phi} = \left(\boldsymbol{u}, \widetilde{\boldsymbol{\Theta}}\boldsymbol{\Lambda}\right) \in T_{\boldsymbol{\Phi}}\mathcal{C}.$$

The above results allow to see that

$$\delta\dot{\boldsymbol{x}} = \dot{\boldsymbol{u}} \qquad \text{and} \qquad \delta\widetilde{\boldsymbol{\Omega}} = \boldsymbol{\Lambda}^t\dot{\widetilde{\boldsymbol{\Theta}}}\boldsymbol{\Lambda}.$$

Replacing the above results in (85) yields

$$\Big\langle \delta S[\boldsymbol{\Phi}], \delta\boldsymbol{\Phi} \Big\rangle = \frac{1}{2}\int_0^T \mathrm{tr}\left[\delta\widetilde{\boldsymbol{\Theta}}\left(\frac{d}{dt}(\mathbf{j}_d\widetilde{\boldsymbol{\omega}} + \widetilde{\boldsymbol{\omega}}\mathbf{j}_d)\right)\right] dt - \int_0^T m\,\ddot{\boldsymbol{x}}\cdot\boldsymbol{u}\,dt = 0,$$

where $\widetilde{\omega} = \mathbf{\Lambda}\widetilde{\mathbf{\Omega}}\mathbf{\Lambda}^t$ and $\mathbf{j}_d = \mathbf{\Lambda}\mathbf{J}_d\mathbf{\Lambda}^t$ are the spatial forms of the angular velocity tensor and nonstandard inertia tensor, respectively. Noting that since $(\boldsymbol{u}(t), \widetilde{\mathbf{\Theta}}(t))$ are arbitrary for all $t \in [0, T]$ and that

$$\text{skew}\big[\mathbf{j}\,\boldsymbol{\omega}\big] = \mathbf{j}_d\,\widetilde{\boldsymbol{\omega}} + \widetilde{\boldsymbol{\omega}}\,\mathbf{j}_d, \qquad (\widetilde{\boldsymbol{\omega}} = \text{skew}[\boldsymbol{\omega}]),$$

the following system of Euler–Lagrange equations is obtained

$$m\ddot{\boldsymbol{x}} = 0 \qquad \text{and} \qquad \frac{d}{dt}\left(\text{skew}\left[\mathbf{j}\,\boldsymbol{\omega}\right]\right) = 0,$$

where $\boldsymbol{\omega} \in \mathbb{R}^3$ is the axial vector of $\widetilde{\boldsymbol{\omega}}$ and $\mathbf{j} = \mathbf{\Lambda}\mathbf{J}\mathbf{\Lambda}^t$. Basically, both equations are alternative statements for the conservation of the total linear momentum and total angular momentum.

4.1.1 Legendre Transforms

The momentum vectors are computed with the help of the Legendre transformation as

$$\mathbf{p} = m\dot{\boldsymbol{x}}, \qquad \text{and} \qquad \boldsymbol{\pi} = \mathbf{j}\boldsymbol{\omega},$$

which allows to define the Hamiltonian function, $H : T^*SE(3) \to \mathbb{R}$, as

$$H(\mathbf{p}, \boldsymbol{\pi}) = \frac{1}{2}\left(m^{-1}\mathbf{p}\cdot\mathbf{p} + \mathbf{j}^{-1}\boldsymbol{\pi}\cdot\boldsymbol{\pi}\right),$$

and therefore, the E–L equations may be rewritten as

$$\dot{\mathbf{p}} = 0 \qquad \text{and} \qquad \dot{\widetilde{\boldsymbol{\pi}}} = 0, \tag{86}$$

respectively.

4.1.2 Discretization

The procedure to construct a variational integrator for this problem follows some ideas presented in (Lee et al. 2007) for the full-body problem. We denote by $\boldsymbol{x}^k \approx \boldsymbol{x}(t^k)$ and $\mathbf{\Lambda}^k \approx \mathbf{\Lambda}(t^k)$, $k = 1, ..., N$ and we assume that the velocities are approximated by

$$\dot{\boldsymbol{x}}(\tau) \approx \frac{\boldsymbol{x}^1 - \boldsymbol{x}^0}{h}, \tag{87a}$$

$$\widetilde{\mathbf{\Omega}}(\tau) \approx \mathbf{\Lambda}^{0t}\frac{\mathbf{\Lambda}^1 - \mathbf{\Lambda}^0}{h}, \qquad \tau \in [t^0, t^1], \tag{87b}$$

and therefore,

$$\mathrm{tr}\left[\widetilde{\mathbf{\Omega}}\mathbf{J}_d\widetilde{\mathbf{\Omega}}^t\right] \approx \mathrm{tr}\left[\mathbf{\Lambda}^{0t}\frac{\mathbf{\Lambda}^1-\mathbf{\Lambda}^0}{h}\mathbf{J}_d\left(\mathbf{\Lambda}^{0t}\frac{\mathbf{\Lambda}^1-\mathbf{\Lambda}^0}{h}\right)^t\right], \tag{88}$$

which, taking into account that for any two matrices $\mathbf{A},\mathbf{B}\in\mathbb{R}^{n\times n}$, $\mathrm{tr}[\mathbf{AB}] = \mathrm{tr}[\mathbf{BA}] = \mathrm{tr}[\mathbf{A}^t\mathbf{B}^t]$, yields

$$\mathrm{tr}\left[\widetilde{\mathbf{\Omega}}\mathbf{J}_d\widetilde{\mathbf{\Omega}}^t\right] \approx \frac{2}{h^2}\mathrm{tr}\left[\left(\mathbf{I}-\mathbf{\Lambda}^{0t}\mathbf{\Lambda}^1\right)\mathbf{J}_d\right].$$

Then, the discrete Lagrangian based on the trapezoidal rule is given by

$$L_d(\boldsymbol{x}^0,\boldsymbol{x}^1,\mathbf{\Lambda}^0,\mathbf{\Lambda}^1) = \frac{1}{2h}m(\boldsymbol{x}^1-\boldsymbol{x}^0)\cdot(\boldsymbol{x}^1-\boldsymbol{x}^0) + \frac{1}{h}\mathrm{tr}\left[\left(\mathbf{I}-\mathbf{\Lambda}^{0t}\mathbf{\Lambda}^1\right)\mathbf{J}_d\right]. \tag{89}$$

Remark If the midpoint rule is preferred, the following approximation has to be used for the angular velocity tensor

$$\widetilde{\mathbf{\Omega}} \approx \frac{\mathbf{\Lambda}^{0t}+\mathbf{\Lambda}^{1t}}{2}\frac{\mathbf{\Lambda}^1-\mathbf{\Lambda}^0}{h}.$$

Moreover, denoting by

$$\begin{aligned}\mathbf{A} &= (\mathbf{\Lambda}^{0t}+\mathbf{\Lambda}^{1t})(\mathbf{\Lambda}^1-\mathbf{\Lambda}^0)\mathbf{J}_d\\ \mathbf{B} &= (\mathbf{\Lambda}^{1t}-\mathbf{\Lambda}^{0t})(\mathbf{\Lambda}^0+\mathbf{\Lambda}^1)\end{aligned}$$

and applying the fact that $\mathrm{tr}[\mathbf{AB}] = \mathrm{tr}\left[\mathbf{A}^t\mathbf{B}^t\right]$, it is possible to obtain

$$\mathrm{tr}\left[\widetilde{\mathbf{\Omega}}\mathbf{J}_d\widetilde{\mathbf{\Omega}}^t\right] \approx \frac{1}{2h^2}\mathrm{tr}\left[\mathbf{J}_d\left(\mathbf{I}-\mathbf{\Lambda}^{1t}\mathbf{\Lambda}^0\mathbf{\Lambda}^{1t}\mathbf{\Lambda}^0\right)\right],$$

which contains a higher power of $(\mathbf{\Lambda}^{1t}\mathbf{\Lambda}^0)$ ■

From (89) it is possible to define the discrete translational momenta as

$$\mathbf{p}^0 = -D_{\boldsymbol{x}^0}L_d = \frac{m}{h}(\boldsymbol{x}^1-\boldsymbol{x}^0), \tag{90a}$$

$$\mathbf{p}^1 = D_{\boldsymbol{x}^1}L_d = \mathbf{p}^0. \tag{90b}$$

Determining the momenta associated to the rotational part of the motion is a little bit more involved. On one hand, considering that $\delta\mathbf{\Lambda}^0 = \widetilde{\mathbf{\Theta}}^0\mathbf{\Lambda}^0$, we have that the following relation holds

$$\mathrm{tr}\left[-\left(D_{\mathbf{\Lambda}^0}L_d\right)\delta\mathbf{\Lambda}^0\right] = \mathrm{tr}\left[-\widetilde{\mathbf{\Theta}}^0\mathbf{\Lambda}^0D_{\mathbf{\Lambda}^0}L_d\right] = \mathrm{tr}\left[\widetilde{\mathbf{\Theta}}^0\left(\frac{1}{h}\mathbf{\Lambda}^1\mathbf{J}_d\mathbf{\Lambda}^{0t}\right)\right] = 0.$$

On the other hand, we know that $\widetilde{\mathbf{\Pi}}^0$, the momentum conjugated to the rotation $\mathbf{\Lambda}^0$, belongs to the linear space $so(3)^*$ which is the dual space of $so(3)$. Moreover, $so(3)^*$ is also composed by skew-symmetric tensors. See, e.g., (Wendlandt and Marsden 1997; Marsden and Ratiu 1999).

Therefore, since $\widetilde{\mathbf{\Theta}}^0$ is skew-symmetric and

$$\mathrm{tr}\left[\widetilde{\mathbf{\Theta}}^0\mathbf{B}\right] = \frac{1}{2}\mathrm{tr}\left[\widetilde{\mathbf{\Theta}}^0(\mathbf{B}-\mathbf{B}^t)\right],$$

for any 3×3 matrix $\mathbf{B}$, we have that the above equation can be rewritten as

$$\frac{1}{h}\mathrm{tr}\left[\widetilde{\mathbf{\Theta}}^0\left(\mathbf{\Lambda}^1\mathbf{J}_d\mathbf{\Lambda}^{0t}-\mathbf{\Lambda}^0\mathbf{J}_d\mathbf{\Lambda}^{1t}\right)\right] = \frac{1}{h}\mathrm{tr}\left[\widetilde{\mathbf{\Theta}}^0\widetilde{\mathbf{\Pi}}^0\right],$$

which allows to identify

$$\widetilde{\mathbf{\Pi}}^0 = \frac{1}{h}\left(\mathbf{\Lambda}^{in}\mathbf{j}_d^0-\mathbf{j}_d^0\mathbf{\Lambda}^{in(t)}\right), \tag{90c}$$

with the momentum conjugated to $\mathbf{\Lambda}^0$ after defining $\mathbf{j}_d^0=\mathbf{\Lambda}^0\mathbf{J}_d\mathbf{\Lambda}^{0t}$ and $\mathbf{\Lambda}^{in}=\mathbf{\Lambda}^1\mathbf{\Lambda}^{0t}$. On the other hand, following an analogue procedure allows to compute the skew-symmetric momentum associated to $\mathbf{\Lambda}^1$ as

$$\widetilde{\mathbf{\Pi}}^1 = \widetilde{\mathbf{\Pi}}^0. \tag{90d}$$

Note that (90b) and (90d) are the DEL equations. They also are discrete counterparts of the conservation laws $(86)^{1,2}$.

4.1.3 Solution Procedure

Since the translational and rotational momenta are exactly conserved by the algorithm, we only need to determine $(\boldsymbol{x}^1, \mathbf{\Lambda}^1)$ from $(\boldsymbol{x}^0, \mathbf{\Lambda}^0, \mathbf{p}^0, \mathbf{\Pi}^0)$. The procedure is as follows.

- The position in space is updated using (90a) as

$$\boldsymbol{x}^1 = \boldsymbol{x}^0 + \frac{h}{m}\mathbf{p}^0.$$

- To determine $\mathbf{\Lambda}^1$ we parametrize the incremental rotation tensor $\mathbf{\Lambda}^{in}$ in terms of an incremental rotation vector $\boldsymbol{\theta}\in\mathbb{R}^3$ as

$$\mathbf{\Lambda}^{in} = \exp[\widetilde{\boldsymbol{\theta}}],$$

where $\widetilde{\boldsymbol{\theta}} = \text{skew}[\boldsymbol{\theta}]$ is the skew-symmetric tensor obtained from $\boldsymbol{\theta}$ and $\exp[\bullet] : so(3) \mapsto SO(3)$ is the exponential map (see Sect. B.3). Then, replacing in (90c) yields to the following nonlinear system of equations

$$\left(\frac{\sin\theta}{\theta}\right)\mathbf{j}_d^0\,\boldsymbol{\theta} + \left(\frac{1-\cos\theta}{\theta^2}\right)\boldsymbol{\theta}\times\mathbf{j}_d^0\,\boldsymbol{\theta} = \boldsymbol{\Pi}^0, \tag{91}$$

where $\theta = \|\boldsymbol{\theta}\|$. This system is solved with the help of the Newton–Raphson scheme. Having obtained $\boldsymbol{\theta}$, we update $\boldsymbol{\Lambda}^1 = \boldsymbol{\Lambda}^{in}\boldsymbol{\Lambda}^0$.

4.2 A Model for Geometrically Exact Rods

In this example, we take advantage of some of the previous results to build an explicit time integrator for finite element discretizations of geometrically exact rods made of an isotropic, homogeneous and hyperelastic material.

4.2.1 Continuum Model

First, we briefly review some basic results of the continuum model. The reference configuration corresponds to a straight rod of length L and constant cross-section $\mathcal{A} \subset \mathbb{R}^2$. The position vector of a material point in this configuration is

$$\mathbf{X}(s, \xi_1, \xi_2) = s\mathbf{E}_1 + \xi_2\mathbf{E}_2 + \xi_3\mathbf{E}_3,$$

where $\{\mathbf{E}_i\}_{i=1,..,3}$ is an (orthogonal) inertial reference frame, $s \in [0, L]$ is an arch-length coordinate and $(\xi_2, \xi_3) \subset \mathcal{A}$ are coordinates on the cross section. The geometric place of points of the form $\mathbf{X}(s, 0, 0)$ defines a reference curve $\boldsymbol{\varphi}_0$. The current configuration of the rod is characterized by the fields

$$\boldsymbol{\Phi} = (\boldsymbol{\varphi}, \boldsymbol{\Lambda}) : [0, L] \to \mathbb{R}^3 \times SO(3), \tag{92}$$

where $\boldsymbol{\varphi}$ is obtained by adding a displacement field onto $\boldsymbol{\varphi}_0$ and $\boldsymbol{\Lambda}$ defines the orientation of the reference frame

$$\boldsymbol{t}_i(s) = \boldsymbol{\Lambda}(s)\mathbf{E}_i, \qquad s \in [0, L], \quad i = 1, 2, 3,$$

which rigidly attached to the cross-section at $s \in [0, L]$ and oriented along the principal axis of inertia by convenience. The position vector of a material point in the current configuration is given by

$$\mathbf{x}(s, \xi_2, \xi_3) = \boldsymbol{\varphi}(s) + \xi_2\mathbf{t}_2(s) + \xi_3\mathbf{t}_3(s).$$

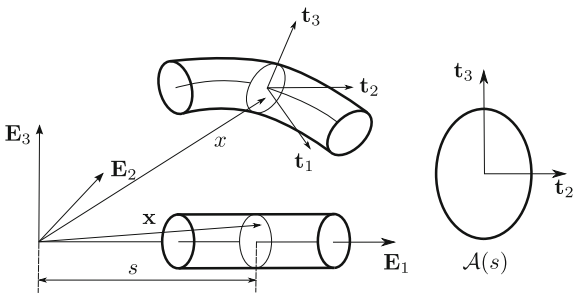


Fig. 18 Reference and current configurations of the elastic rod

See e.g., (Simo 1985; Simo and Vu-Quoc 1986, 1988; Kapania and Li 2003) and Fig. 18.

Then, the configuration manifold, $\mathcal{Q}$, is the set of all the smooth-enough fields of the form (92) subjected to the prescribed boundary conditions $\mathbf{\Phi}(0) = \mathbf{\Phi}_0$ and $\mathbf{\Phi}(L) = \mathbf{\Phi}_L$ and to the restriction $\frac{\partial \boldsymbol{\varphi}}{\partial s} \cdot \mathbf{t}_1 > 0$ (Simo et al. 1995).

Given a motion $\mathbf{\Phi}(\cdot) : [0, T] \to \mathcal{Q}$, the corresponding velocity is obtained as $\dot{\mathbf{\Phi}} = (\dot{\boldsymbol{\varphi}}, \dot{\mathbf{\Lambda}})$ which yields to the following expression for the velocity of a material point:

$$\dot{\mathbf{x}} = \dot{\boldsymbol{\varphi}} + \mathbf{\Lambda}\widetilde{\mathbf{V}}\mathbf{Z}, \tag{93}$$

where $\mathbf{Z} = \xi_2\mathbf{E}_2 + \xi_3\mathbf{E}_3$ and $\widetilde{\mathbf{V}} = \mathbf{\Lambda}^{\mathrm{t}}\dot{\mathbf{\Lambda}}$, is the material form of the angular velocity tensor.

Moreover, consider a stored energy function per unit of reference length $\psi(\mathbf{\Gamma}, \mathbf{\Omega})$ such that the material form of the cross-sectional stress resultants and stress couples are given by

$$\mathbf{n}^{\mathrm{m}} = \frac{\partial \psi}{\partial \mathbf{\Gamma}}(\mathbf{\Gamma}, \mathbf{\Omega}) = \mathbf{C}_\Gamma \mathbf{\Gamma}, \quad \text{and} \quad \mathbf{m}^{\mathrm{m}} = \frac{\partial \psi}{\partial \mathbf{\Omega}}(\mathbf{\Gamma}, \mathbf{\Omega}) = \mathbf{C}_\Omega \mathbf{\Omega}, \tag{94}$$

where

$$\mathbf{\Gamma} = \mathbf{\Lambda}^t\left(\frac{\partial \boldsymbol{\varphi}}{\partial s} - \boldsymbol{t}_1\right)$$

is the (translational) strain vector, $\mathbf{\Omega}$ is the axial vector of the curvature tensor

$$\widetilde{\mathbf{\Omega}} = \mathbf{\Lambda}^{\mathrm{t}}\frac{\partial \mathbf{\Lambda}}{\partial s}$$

and the constitutive tensors $\mathbf{C}_\Gamma$ and $\mathbf{C}_\Omega$ are given by

$$\mathbf{C}_\Gamma = \begin{bmatrix} E\mathcal{A} & 0 & \\ 0 & G\mathcal{A} & 0 \\ 0 & 0 & G\mathcal{A} \end{bmatrix} \quad \text{and} \quad \mathbf{C}_\Omega = \begin{bmatrix} GJ & 0 & 0 \\ 0 & EI_2 & 0 \\ 0 & 0 & EI_3 \end{bmatrix},$$

where E, G, J, I_2, and I_3 are an elastic modulus, a shear modulus, a torsional stiffness, and two flexural stiffnesses, respectively. We also define the spatial forms of the stress resultants and couples as

$$\mathbf{n} = \mathbf{\Lambda}\mathbf{n}^{\mathrm{m}} \quad \text{and} \quad \mathbf{m} = \mathbf{\Lambda}\mathbf{m}^{\mathrm{m}},$$

respectively.

4.2.2 Hamilton's Principle

The Lagrangian function $\mathsf{L} : T\mathcal{Q} \to \mathbb{R}$ is constructed as the kinetic minus the potential energy of the system, i.e.,

$$\mathsf{L}(\mathbf{\Phi}, \dot{\mathbf{\Phi}}) = \mathsf{K}(\dot{\mathbf{\Phi}}, \dot{\mathbf{\Phi}}) - \mathsf{U}(\mathbf{\Phi}), \tag{95a}$$

where

$$\mathsf{K}(\mathbf{\Phi}, \dot{\mathbf{\Phi}}) = \frac{1}{2}\int_0^L \left(\mathcal{A}_\rho \dot{\boldsymbol{\varphi}} \cdot \dot{\boldsymbol{\varphi}} - \mathrm{Tr}[\widetilde{\mathbf{V}}\mathbf{E}_\rho \widetilde{\mathbf{V}}]\right) ds, \tag{95b}$$

$$\mathsf{U}(\mathbf{\Phi}) = \int_0^L \psi(\mathbf{\Gamma}, \mathbf{\Omega})\, ds. \tag{95c}$$

In (95b) the mass density per unit of reference length is given by

$$\mathcal{A}_\rho = \int_{\mathcal{A}} \rho_0 \, dA$$

and

$$\mathbf{E}_\rho = \int_{\mathcal{A}} \rho_0 \mathbf{Z} \otimes \mathbf{Z} \, dA,$$

a cross-sectional nonstandard inertia tensor. Compare with (84).

Consider the set $\mathcal{C}$ composed by all the smooth enough motions $\mathbf{\Phi} : [0, T] \to \mathcal{Q}$. Hamilton's principle states that the trajectory followed by the system is a stationary point of the action under all variations in $\mathcal{C}$ that leaves fixed the end points $\mathbf{\Phi}(0) = \mathbf{\Phi}(T) = \mathbf{0}$. This principle yields to the following Euler–Lagrange equations

$$\mathcal{A}_\rho \frac{d^2\boldsymbol{\varphi}}{dt^2} = \frac{\partial \mathbf{n}}{\partial s} + \mathbf{N}, \tag{96a}$$

$$\frac{d}{dt}\left(\mathrm{skew}[\mathbf{i}_\rho \mathbf{v}]\right) = \frac{\partial \widetilde{\mathbf{m}}}{\partial s} + \frac{\partial \widetilde{\boldsymbol{\varphi}}}{\partial s}\widetilde{\mathbf{n}} - \widetilde{\mathbf{n}}\frac{\partial \widetilde{\boldsymbol{\varphi}}}{\partial s} + \widetilde{\mathbf{M}}, \tag{96b}$$

which have to be supplemented with adequate initial conditions. In the above equations $\mathbf{i}_\rho$ is the spatial form of the inertial tensor, $\mathbf{N}$ a vector of the external forces and $\widetilde{\mathbf{M}}$ a skew-symmetric tensor of external moments.

4.2.3 Hamiltonian Framework

The Momentum densities

$$\mathbf{p} = \mathcal{A}_\rho \dot{\boldsymbol{\varphi}} \qquad \text{and} \qquad \widetilde{\boldsymbol{\pi}} = \text{skew}\left[\mathbf{i}_\rho \mathbf{v}\right], \tag{97}$$

are introduced by means of the Legendre transforms. Rewriting the kinetic energy in terms of $(\mathbf{p}, \boldsymbol{\pi})$ it is possible to define the Hamiltonian function $\mathsf{H} : \mathsf{T}^*\mathcal{Q} \to \mathbb{R}$ as

$$\mathsf{H}(\mathbf{p}, \boldsymbol{\pi}, \boldsymbol{\Phi}) = \mathsf{K}(\mathbf{p}, \boldsymbol{\pi}) + \mathsf{U}(\boldsymbol{\Phi}) = \frac{1}{2}\int_0^L \left(\mathcal{A}_\rho^{-1}\mathbf{p}\cdot\mathbf{p} + \mathbf{i}_\rho^{-1}\boldsymbol{\pi}\cdot\boldsymbol{\pi}\right) ds + \mathsf{U}(\boldsymbol{\Phi}),$$

from which it is possible to obtain the balance equations in Hamiltonian form as

$$\begin{aligned}
\dot{\boldsymbol{\varphi}} &= \mathcal{A}_\rho^{-1}\mathbf{p},\\
\dot{\boldsymbol{\Lambda}} &= \text{skew}[\mathbf{i}_\rho^{-1}\boldsymbol{\pi}]\boldsymbol{\Lambda},\\
\dot{\mathbf{p}} &= \frac{\partial \mathbf{n}}{\partial s} + \mathbf{N},\\
\dot{\widetilde{\boldsymbol{\pi}}} &= \frac{\partial \widetilde{\mathbf{m}}}{\partial s} + \frac{\partial \widetilde{\boldsymbol{\varphi}}}{\partial s}\widetilde{\mathbf{n}} - \widetilde{\mathbf{n}}\frac{\partial \widetilde{\boldsymbol{\varphi}}}{\partial s} + \tilde{\mathbf{M}}.
\end{aligned}$$

Regarding to the invariants of the dynamics, we note that the Lagrangian is invariant under translations in time and under rigid body translations and rotations in space. These properties yields to the conservation of energy along with conservation of the total linear momentum, $\mathbf{L}_\text{t}$, and total angular momentum, $\mathbf{J}_\text{t}$, which are given by

$$\mathbf{L}_\text{t} = \int_0^L \mathbf{p}\, ds \qquad \text{and} \qquad \mathbf{J}_\text{t} = \int_0^L (\boldsymbol{\varphi}\times\mathbf{p} + \boldsymbol{\pi})\, ds. \tag{99}$$

4.2.4 Discretization in Space: Finite Elements

The discretization in space of the problem is carried out with the help of the finite element method. We consider a partition of $[0, L]$ in N_e linear elements with constant length h. The basic idea consists in approximating $\mathcal{Q}$ by means of a finite-dimensional subspace $\mathcal{Q}_h$. In following, calculations are performed on the basis of a generic finite element.

(i) We consider first the translational part of the motion. The current position of $\boldsymbol{\varphi}$ is approximated by

$$\boldsymbol{\varphi}_h(\zeta) = \frac{1}{2}(1-\zeta)\boldsymbol{\varphi}_1 + \frac{1}{2}(1+\zeta)\boldsymbol{\varphi}_2, \qquad \zeta \in [-1,1], \tag{100}$$

where $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2$ are the position vectors of the initial and final nodes of a generic element in the mesh (Hughes 1987). It is worth noting that $\boldsymbol{\varphi}_h(\zeta)$ belongs to $\mathbb{R}^3$ for all $\zeta \in [-1,1]$ and that

$$\frac{\partial \boldsymbol{\varphi}_h}{\partial s}(\zeta) = \frac{\boldsymbol{\varphi}_2 - \boldsymbol{\varphi}_1}{h}, \qquad \zeta \in [-1,1],$$

since $ds/d\zeta = 2/h$.

(ii) Since the rotation group is a nonlinear manifold, (100) cannot be applied. Instead, we adopt the following procedure proposed in (Sansour and Wagner 2003),

- Use the *Spurrier's algorithm* (Spurrier 1978) to represent the nodal values of the rotation tensors, $\{\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2\}$, in terms of unit quaternions[8] $\{\mathbf{q}_1, \mathbf{q}_2\}$.
- The nodal values of the rotation vectors $\{\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2\}$ are extracted from $\{\mathbf{q}_1, \mathbf{q}_2\}$ with the help of the procedure given in (Simo and Vu-Quoc 1986).
- $\boldsymbol{\Psi}_h(s)$ can be obtained by applying (100). Then, $\mathbf{q}_h(s)$ is computed from $\boldsymbol{\Psi}_h(s)$ following standard procedures (see e.g. Crisfield 1998, Chap. XVI).
- Finally, $\boldsymbol{\Lambda}_h(s)$ is obtained from $\mathbf{q}_h(s)$ by applying the classical relation between rotation tensors and quaternions.

The construction of a semi-discrete counterpart of the Lagrangian function (95a) is as follows. The semi-discrete kinetic energy $\mathsf{K}^h : \mathsf{T}\mathcal{Q}_h \times \mathbb{R} \to \mathbb{R}$ and the internal and external components of the potential energy $\mathsf{U}^{h,\text{int}}, \mathsf{U}^{h,\text{ext}} : \mathcal{Q}_h \times \mathbb{R} \to \mathbb{R}$ are given by

$$\mathsf{K}^h = \frac{h}{4}\left(\mathcal{A}_\rho(\|\dot{\boldsymbol{\varphi}}_1\|^2 + \|\dot{\boldsymbol{\varphi}}_2\|^2) - \text{Tr}\left[\widetilde{\mathbf{V}}_1 \mathbf{E}_\rho \widetilde{\mathbf{V}}_1 + \widetilde{\mathbf{V}}_2 \mathbf{E}_\rho \widetilde{\mathbf{V}}_2\right]\right., \tag{101a}$$

$$\mathsf{U}^{h,\text{int}} = h\ \psi\left(\boldsymbol{\Gamma}_h(0), \boldsymbol{\Omega}_h(0)\right), \tag{101b}$$

$$\mathsf{U}^{h,\text{ext}} = h\ \mathrm{w}_{\text{ext}}\left(\boldsymbol{\varphi}_h(0), \boldsymbol{\Lambda}_h(0)\right), \tag{101c}$$

where the explicit dependence on time has been omitted to simplify notation and

$$\boldsymbol{\Gamma}_h(0) = \boldsymbol{\Lambda}_h^t(0)\frac{\partial \boldsymbol{\varphi}_h}{\partial s}(0) - \mathbf{E}_1,$$

$$\boldsymbol{\Omega}_h(0) = \text{axial}\left[\widetilde{\boldsymbol{\Omega}}_h(0)\right] = \text{axial}\left[\boldsymbol{\Lambda}_h^t(0)\frac{\partial \boldsymbol{\Lambda}_h}{\partial s}(0)\right],$$

[8] Theoretical aspect about unit quaternions can be consulted, e.g., in (Mcrobie and Lasenby 1999; Crisfield 1998).

The semi-discrete Lagrangian is then constructed as

$$\mathsf{L}^h = \mathsf{K}^h - \mathsf{U}^{h,\text{int}} - \mathsf{U}^{h,\text{ext}}.$$

The application of Hamilton's principle allows to obtain the corresponding EL equations on the nodes of the mesh.

4.2.5 Discretization in Time: Variational Integrators

We denote by

$$(\boldsymbol{\varphi}_i^k, \boldsymbol{\Lambda}_i^k), \qquad i = 1, 2, \tag{102}$$

to the approximation of the nodal variables $\left(\boldsymbol{\varphi}_i(t^k), \boldsymbol{\Lambda}_a^k(t^k)\right)$.

The nodal value of the translational velocity is approximated by

$$\dot{\boldsymbol{\varphi}}_i(\tau) \approx \frac{\boldsymbol{\varphi}_i^{k+1} - \boldsymbol{\varphi}_i^k}{\Delta t}, \qquad i = 1, 2, \quad \tau \in (t^k, t^{k+1}), \tag{103}$$

and time derivative of the rotational tensor by

$$\dot{\boldsymbol{\Lambda}}_i(\tau) \approx \frac{\boldsymbol{\Lambda}_i^{k+1} - \boldsymbol{\Lambda}_h^k}{\Delta t}, \qquad i = 1, 2, \quad \tau \in (t^k, t^{k+1}). \tag{104}$$

The fully discrete (in space and time) counterpart of the kinetics energy is obtained by replacing $\dot{\boldsymbol{\varphi}}_i$ by (103) and $\dot{\boldsymbol{\Lambda}}_i$ by (104) in (101a) to obtain

$$\mathsf{K}_d^h\left(\boldsymbol{\Phi}_h^0, \boldsymbol{\Phi}_h^1\right) = \frac{h}{2(\Delta t)^2} \sum_{i=1}^{2} \left(\frac{\mathcal{A}_\rho}{2} \|\boldsymbol{\varphi}_i^{k+1} - \boldsymbol{\varphi}_i^k\|^2 + \text{Tr}\left[\left(\mathbf{I} - \boldsymbol{\Lambda}_i^{(k)t} \boldsymbol{\Lambda}_i^{k+1} \right) \mathbf{E}_\rho \right] \right). \tag{105}$$

Correspondingly, the discrete potential energy, $\mathsf{U}_d : \mathcal{Q}_h \to \mathbb{R}$, is given by

$$\mathsf{U}_d^h(\boldsymbol{\Phi}_h^k) = h \left(\psi\left(\boldsymbol{\Gamma}_h^k(0), \widetilde{\boldsymbol{\Omega}}_h^k(0) \right) - \mathrm{w}_{\text{ext}}\left(\boldsymbol{\varphi}_h^k(0), \boldsymbol{\Lambda}_h^k(0) \right) \right). \tag{106}$$

Consider $\alpha \in [0, 1]$ and apply the generalized trapezoidal rule (Marsden and West 2001) to construct a discrete Lagrangian,

$$\mathsf{L}_d^{h,\alpha}\left(\boldsymbol{\Phi}_h^0, \boldsymbol{\Phi}_h^1\right) = \Delta t \left(\mathsf{K}_d^h\left(\boldsymbol{\Phi}_h^0, \boldsymbol{\Phi}_h^1\right) - \alpha \mathsf{U}_d^h\left(\boldsymbol{\Phi}_h^0\right) - (1-\alpha) \mathsf{U}_d^h\left(\boldsymbol{\Phi}_h^1\right) \right). \tag{107}$$

The application of the discrete Hamilton's principle yields to the following DEL equations

$$\mathrm{m}_i \frac{\boldsymbol{\varphi}_i^{k+1} - 2\boldsymbol{\varphi}_i^k - \boldsymbol{\varphi}_i^{k-1}}{(\Delta t)^2} = \mathbf{s}_i^k + \mathbf{S}_i^k,$$
$$\frac{\mathbf{R}_i^{(k)\mathrm{t}} - \mathbf{R}_i^{(k-1)\mathrm{t}} - \mathbf{R}_i^k + \mathbf{R}_i^{(k-1)}}{(\Delta t)^2} = \widetilde{\mathbf{h}}_i^k + \widetilde{\mathbf{H}}_i^k, \qquad i = 1, 2,$$

where

$$\mathrm{m}_i = \frac{h}{2}\mathcal{A}_\rho, \tag{109a}$$

$$\mathbf{R}_i^k = \frac{h}{2}\mathbf{J}_{\rho(i)}^k \boldsymbol{\Lambda}_{\mathrm{in}(i)}^{(k)\mathrm{t}}, \tag{109b}$$

$$\mathbf{s}_i^k = (-1)^i \mathbf{n}_h^k(0), \tag{109c}$$

$$\widetilde{\mathbf{h}}_i^k = (-1)^{i+1}\widetilde{\mathbf{m}}_h^k(0) + \frac{(\widetilde{\boldsymbol{\varphi}}_2^k - \widetilde{\boldsymbol{\varphi}}_1^k)}{2}\widetilde{\mathbf{n}}_h^k(0) - \widetilde{\mathbf{n}}_h^k(0)\frac{(\widetilde{\boldsymbol{\varphi}}_2^k - \widetilde{\boldsymbol{\varphi}}_1^k)}{2}, \tag{109d}$$

$$\mathbf{S}_i^k = -\frac{h}{2}\mathbf{N}_h^k(0) \quad \text{and} \quad \widetilde{\mathbf{H}}_i^k = \frac{h}{2}\widetilde{\mathbf{M}}_h^k(0), \tag{109e}$$

with $\mathbf{J}_{\rho(i)}^k = \boldsymbol{\Lambda}_i^k \mathbf{E}_\rho \boldsymbol{\Lambda}_i^{(k)\mathrm{t}}$ and $\boldsymbol{\Lambda}_{\mathrm{in}(i)}^k = \boldsymbol{\Lambda}_i^{(k)\mathrm{t}} \boldsymbol{\Lambda}_i^{k+1}$.

For the translational part of the motion, the discrete Legendre transform allows to obtain

$$\mathbf{p}_i^k = \mathrm{m}_i \frac{\boldsymbol{\varphi}_i^{k+1} - \boldsymbol{\varphi}_i^k}{\Delta t} + \alpha\Delta t \left(\mathbf{s}_i^k + \mathbf{S}_i^k\right), \tag{110a}$$
$$\mathbf{p}_i^{k+1} = \mathrm{m}_i \frac{\boldsymbol{\varphi}_i^{k+1} - \boldsymbol{\varphi}_i^k}{\Delta t} - (1-\alpha)\Delta t \left(\mathbf{s}_i^{k+1} + \mathbf{S}_i^{k+1}\right), \quad i = 1, 2. \tag{110b}$$

For the rotational part the following relation holds,

$$\widetilde{\boldsymbol{\Pi}}_i^k = \frac{\mathbf{R}_i^{(k)\mathrm{t}} - \mathbf{R}_i^k}{\Delta t} - \alpha\Delta t \left(\widetilde{\mathbf{h}}_i^k + \widetilde{\mathbf{H}}_i^k\right), \tag{111a}$$
$$\widetilde{\boldsymbol{\Pi}}_i^{k+1} = \frac{\mathbf{R}_i^{(k)\mathrm{t}} - \mathbf{R}_i^k}{\Delta t} + (1-\alpha)\Delta t \left(\widetilde{\mathbf{h}}_i^{k+1} + \widetilde{\mathbf{H}}_i^{k+1}\right), \quad i = 1, 2. \tag{111b}$$

Subtracting (110a) from (110b) and (111a) from (111b), yields to the following relations:

$$\mathbf{p}_i^{k+1} = \mathbf{p}_i^k - \Delta t \left(\mathbf{s}_i^{k+\alpha} + \mathbf{S}_i^{k+\alpha}\right), \tag{112a}$$
$$\widetilde{\boldsymbol{\Pi}}_i^{k+1} = \widetilde{\boldsymbol{\Pi}}_i^k + \Delta t \left(\widetilde{\mathbf{h}}_i^{k+\alpha} + \widetilde{\mathbf{H}}_i^{k+\alpha}\right), \qquad i = 1, 2, \tag{112b}$$

where $(\bullet)_a^{k+\alpha} = (1-\alpha)(\bullet)_a^{k+1} + \alpha(\bullet)_a^k$. A detailed deduction of (110a), (110b), (111a), and (111b) may be found in (Mata 2015).

4.2.6 Solution Procedure

The updating procedure for the nodal values of the configuration variables is as follows.

- Since mass matrix is diagonal, it is possible to update the nodal positions explicitly according to

$$\boldsymbol{\varphi}_i^{k+1} = \boldsymbol{\varphi}_i^k + \frac{\Delta t}{\mathrm{m}_i}\left(\mathbf{p}_i^k - \alpha\Delta t\left(\mathbf{s}_i^k + \mathbf{S}_i^k\right)\right). \tag{113}$$

- Equation (111a) yields to

$$\boldsymbol{\Lambda}_{\mathrm{in}(i)}^k \mathbf{J}_{\rho(i)}^k - \mathbf{J}_{\rho(i)}^k \boldsymbol{\Lambda}_{\mathrm{in}(i)}^{k(\mathrm{t})} = \frac{2\Delta t}{h}\left(\widetilde{\boldsymbol{\Pi}}_i^k + \alpha\Delta t\left(\widetilde{\mathbf{h}}_i^k + \widetilde{\mathbf{H}}_i^k\right)\right). \tag{114}$$

To determine $\boldsymbol{\Lambda}_{\mathrm{in}(i)}^k$ we follow the method proposed in (Lee et al. 2007). We parametrize the incremental rotation tensor in terms of an incremental rotation vector as

$$\boldsymbol{\Lambda}_{\mathrm{in}(i)}^k = \exp\left[\mathrm{skew}\left(\boldsymbol{\theta}_i^k\right)\right].$$

Then, replacing the above expression in (114) the following nonlinear system of equations is obtained

$$\frac{\sin\theta_i^k}{\theta_i^k}\left(\mathbf{i}_{\rho(i)}^k\boldsymbol{\theta}_i^k\right) + \frac{1-\cos\theta_i^k}{(\theta_i^k)^2}\left(\boldsymbol{\theta}_i^k \times \mathbf{i}_{\rho(i)}^k\boldsymbol{\theta}_i^k\right) = \mathbf{Y}_i^k, \tag{115}$$

where $\theta = \|\boldsymbol{\theta}\|$ and $\mathbf{Y}_i^k$ is the axial vector of the right-hand side of (114).The system (115) is solved with the help of the Newton–Raphson scheme.

Having obtained $\boldsymbol{\theta}_i^k$, we update the nodal rotation tensors according to

$$\boldsymbol{\Lambda}_a^{k+1} = \exp[\widetilde{\boldsymbol{\theta}}_i^k]\boldsymbol{\Lambda}_i^k. \tag{116}$$

- Finally, the nodal values of the momenta are updated with the help of (112a) and (112b).

Remark We formulated an explicit method to update translational part of the configuration variables, (see (113)). However, to update the rotational part a nonlinear system of equations has to be solved iteratively in every node of the mesh. ■

4.2.7 Properties of the Resulting Scheme

The resulting integration scheme enjoys several properties which are described in following.

- The discretization in space allows to formulate an explicit time integrator since mass matrix is diagonal and positive. Moreover, the order of accuracy with the mesh size does not result affected as explained in (Mata 2015; Cohen et al. 2001).
- Fixing $h > 0$, the order of accuracy of an algorithm in position-momentum form coincides with the order of accuracy with which the discrete Lagrangian approximates the exact discrete Lagrangian (if some standard smooth conditions hold). See (Marsden and West 2001). In our case, second order of accuracy is obtained if $\alpha = \frac{1}{2}$ since the resulting method is symmetric.
- Variational methods automatically conserve a discrete analogue of the symplectic two-form (see, e.g., Lew et al. 2003; Marsden and West 2001 for a proof). Moreover, the discrete total energy

$$\mathsf{H}_d^k = \frac{1}{2}\sum_{a=1}^{n_t}\left(\mathrm{m}_a^{-1}\mathbf{p}_a^k\cdot\mathbf{p}_a^k + \mathbf{I}_{\rho(a)}^{-1}\boldsymbol{\Pi}_a^k\cdot\boldsymbol{\Pi}_a^k\right) + \mathsf{U}_d\left(\boldsymbol{\varphi}_1^k,\ldots,\boldsymbol{\varphi}_{n_d}^k,\boldsymbol{\Lambda}_1^k,\ldots,\boldsymbol{\Lambda}_{n_d}^k\right), \tag{117}$$

  remains $\mathcal{O}\left((\Delta t)^2\right)$ close to the exact value for exponentially long periods of time if a small enough $\Delta t > 0$ is provided (Hairer et al. 2006). In the above equation n_t is the total number of nodes in the mesh.
- $\mathsf{L}_d^{h,\alpha}$ results to be invariant under the action of rigid body translations and rotations. Therefore, according to a discrete version of Noether's theorem the discrete versions of the total linear momentum and total angular velocity

$$\mathbf{L}_{t,d} = \sum_{a=1}^{n_t}\mathbf{p}_a^k \quad \text{and} \quad \mathbf{J}_{t,d} = \sum_{a=1}^{n_t}(\widetilde{\boldsymbol{\Pi}}_a^k + \boldsymbol{\varphi}_a^k\times\mathbf{p}_a^k), \tag{118}$$

  are invariants of the discrete trajectories.
- Regarding to the stability limit, we follow (Lew et al. 2003) choosing Δt as a fraction of the critical time step length imposed by the Courant condition.

4.2.8 Numerical Example: Elastic Ring

In this example, we use the rod model to simulate the dynamics of the elastic ring shown in Fig. 19. The applied load are

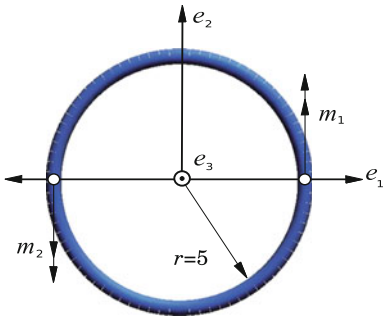
$$\boldsymbol{m}_1 t) = m(t)\mathbf{e}_2 \quad \text{and} \quad \boldsymbol{m}_2(t) = -\boldsymbol{m}_1(t),$$

where

$$m(t) = \begin{cases} (160/3)t, & 0 \le t \le 1.5 \\ 0, & t > 0 \end{cases}. \tag{119}$$

Since the applied moments are self-equilibrated, both the total linear momentum and the total angular momentum are equal to zero and exactly conserved during the free-fly phase of the motion. The mechanical properties of the cross-section are:

Fig. 19 Elastic ring



$E\mathcal{A} = G\mathcal{A} = 3 \times 10^4$, $E\mathrm{I} = 7 \times 10^3$, $\mathcal{A}_{\rho_0} = 1$ and $\mathrm{I}_{\rho_0} = 10$. The ring is discretized using 80 linear elements. The simulation is carried out during 4.5 s with a time step length $\Delta t = 0.0015$ s. The system exhibits drastic changes in the configuration during the motion. Figure 20 shows a sequence of snapshots of the motion in the $\mathbf{e}_1$-$\mathbf{e}_2$ plane. Finally, the time evolution of total energy is shown in Fig. 21 from which it is possible to verify the excellent long-term energy behavior of the algorithm.

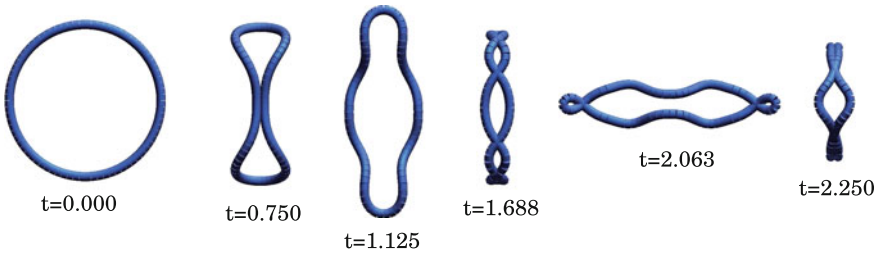


Fig. 20 Sequence of snapshots of the motion in the $\mathbf{e}_1$-$\mathbf{e}_2$ plane. Note that self-contact is not prevented in the present form of the algorithm

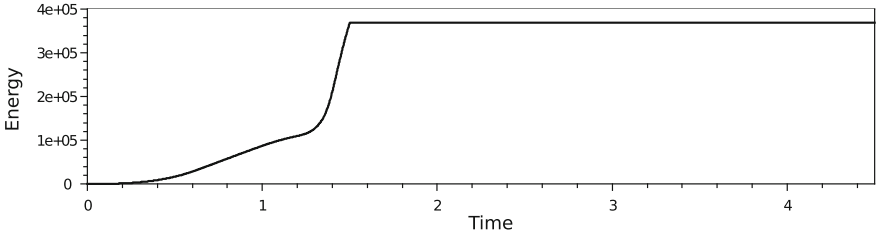


Fig. 21 Time evolution of the total energy

A (Smooth) Manifolds and Lie Groups

In this section, we briefly introduce the concepts of smooth n-dimensional manifolds and Lie groups in order to use them to make a proper introduction to finite rotations in the following section. The interested reader may consult standard textbooks such as, e.g., (Abraham et al. 1988, Arnold 1989, Mishchenko and Fomenko 1988).

A.1 Smooth n-Manifolds

A *smooth n–manifold* or manifold modeled in $\mathbb{R}^n$ is a set $\mathcal{M}$ such that:

- For each element $P \in \mathcal{M}$ there exists a subset $\mathcal{U}$ of $\mathcal{M}$ containing P and an one-to-one mapping called a *chart* or *coordinate system*, $\{x^\alpha\}_{\alpha=1,2,\ldots,n}$, from $\mathcal{U}$ onto an open set $\mathcal{V} \in \mathbb{R}^n$; x^α denote the components of this mapping ($\alpha = 1, 2, \ldots, n$).
- If x^α and $\overline{x}^\alpha$ are two of such mappings, the change of coordinate functions $\overline{x}^\alpha(x^1, \ldots, x^n)$ are C^∞ (*i.e.* it is continuously differentiable as many times as required).

Tangent Space

Let $\mathcal{M} \subset \mathbb{R}^n$ be an open set (manifold) and let $P \in \mathcal{M}$. The *tangent space* to $\mathcal{M}$ at P is simply the vector space $\mathbb{R}^n$ regarded as vectors emanating from P; this tangent space is denoted $T_P\mathcal{M}$.

A.2 Lie Groups

A *Lie group* is a smooth n–dimensional manifold $\mathcal{M}^n$ endowed with the following two smooth mappings:

(i) Multiplication:

$$\begin{aligned} \mathscr{F}_\alpha : \mathcal{M}^n \times \mathcal{M}^n &\to \mathcal{M}^n \\ (\mathbf{u}, \mathbf{v}) &\mapsto \mathscr{F}_\alpha(\mathbf{u}, \mathbf{v}) = \mathbf{u} \odot \mathbf{v}. \end{aligned}$$

where $\odot$ is used to indicate an abstract operation (multiplication) between elements of the manifold $\mathcal{M}^n$.

(ii) Construction of the inverse element:

$$\begin{aligned} \mathscr{F}_\beta : \mathcal{M}^n &\to \mathcal{M}^n \\ \mathbf{u} &\mapsto \mathscr{F}_\beta(\mathbf{u}) = \mathbf{u}^{-1}. \end{aligned}$$

Moreover, a Lie group posses a marked point $\mathbf{e} \in \mathcal{M}^n$ (the identity) which satisfies together with $\mathscr{F}_\alpha$ and $\mathscr{F}_\beta$ the following relations:

- $\boldsymbol{x}_1 \odot (\boldsymbol{x}_2 \odot \boldsymbol{x}_3) = (\boldsymbol{x}_1 \odot \boldsymbol{x}_2) \odot \boldsymbol{x}_3$, for all $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \in \mathcal{M}^n$.
- $\mathbf{e} \odot \boldsymbol{x}_1 = \boldsymbol{x}_1 \odot \mathbf{e} = \boldsymbol{x}$, $\boldsymbol{x}, \mathbf{e} \in \mathcal{M}^n$.
- $\boldsymbol{x} \odot \boldsymbol{x}^{-1} = \boldsymbol{x}^{-1} \odot \boldsymbol{x} = \mathbf{e}$, $\boldsymbol{x}, \boldsymbol{x}^{-1}, \mathbf{e} \in \mathcal{M}^n$.

Lie Algebra

The *Lie algebra*, $\mathfrak{g}$, of the Lie group G is given by its tangent vector space at the identity, $\mathfrak{g} = T_{\mathbf{e}}G$, equipped with a bilinear, skew–symmetric brackets operator $[\cdot, \cdot]$ satisfying the following relations (Dubrokin et al. 2000; Mishchenko and Fomenko 1988):

(i) Jacobi's identity:

$$[\boldsymbol{x}_a, [\boldsymbol{x}_b, \boldsymbol{x}_c]] + [\boldsymbol{x}_b, [\boldsymbol{x}_c, \boldsymbol{x}_a]] + [\boldsymbol{x}_c, [\boldsymbol{x}_a, \boldsymbol{x}_b]] = 0 \quad \text{for all} \quad \boldsymbol{x}_a, \boldsymbol{x}_b, \boldsymbol{x}_c \in \mathfrak{g}.$$

(ii) Skew-symmetry:

$$[\boldsymbol{x}_a, \boldsymbol{x}_b] = -[\boldsymbol{x}_b, \boldsymbol{x}_a] \quad \text{for all} \quad \boldsymbol{x}_a, \boldsymbol{x}_b \in \mathfrak{g},$$

where the *Lie brackets* are given by $[\boldsymbol{x}_a, \boldsymbol{x}_b] = \boldsymbol{x}_a \odot \boldsymbol{x}_b - \boldsymbol{x}_b \odot \boldsymbol{x}_a$.

B Finite Rotations

This section provides a brief introduction to finite rotations and to the rotational motion. We restrict the survey to such concepts that are used through the sections of the chapter. A more extensive review can be found, e.g., in (Argyris 1982; Argyris and Poterasu 1993; Atluri and Cazzani 1995; Bauchau and Trainelli 2003).

B.1 Noncommutative Rotations

Consider the *rotation vector*

$$\hat{\boldsymbol{\phi}} = (\phi_{xx}, \phi_{yy}, \phi_{zz}) = (\pi/2, \pi/2, \pi/2).$$

Figure 22 shows that the order in which we apply the components of the rotation vector on a rigid body (in this case a rigid box) affects its final configuration in space. Therefore, $\hat{\boldsymbol{\phi}}$ can not be used to represent uniquely a rotation in space. Or, in

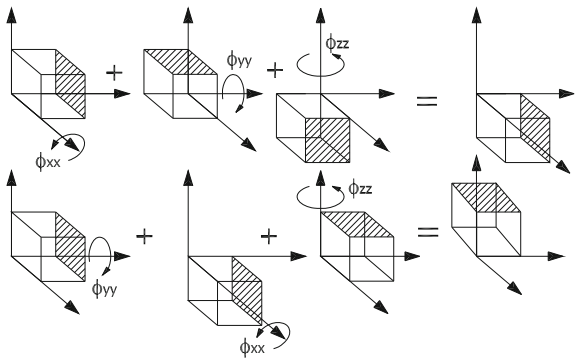


Fig. 22 Noncommutativity of the components of the *rotation vector* $\hat{\phi} = [\phi_{xx}, \phi_{yy}, \phi_{zz}]$

other words, rotations are not elements of a vector space (Jelenič and Crisfield 1999; Simo and Vu-Quoc 1986).

Alternatively, we may think of a rotation $\boldsymbol{\beta}$ as a linear application from the Euclidean vector space, $\mathbb{E}^3$, to itself. Therefore, when $\boldsymbol{\beta}$ is applied to a vector $\boldsymbol{u} \in \mathbb{E}^3$, the result is a new vector $\boldsymbol{v} = \boldsymbol{\beta u}$ conserving the original length. Consider the set

$$\mathfrak{R} = \{\boldsymbol{\beta} : \mathbb{E}^3 \to \mathbb{E}^3 \mid \boldsymbol{\beta} \text{ is a rotation}\},$$

and define the *sum of rotations* as

$$(\boldsymbol{\beta}_a \otimes \boldsymbol{\beta}_b)(\boldsymbol{x}) = \boldsymbol{\beta}_a \left(\boldsymbol{\beta}_b(\boldsymbol{x})\right) \qquad \boldsymbol{\beta}_a, \boldsymbol{\beta}_b \in \mathfrak{R}, \quad \boldsymbol{x} \in \mathbb{E}^3. \tag{120}$$

Clearly $\boldsymbol{\beta}_a \otimes \boldsymbol{\beta}_b \in \mathfrak{R}$ is a *compound rotation* applied on $\boldsymbol{x}$.

The set $\mathfrak{R}$ equipped with the operation $\otimes$ posses the algebraic structure of *noncommutative group* (Bauchau and Trainelli 2003; Mäkinen and Marjamäki 2005; Mäkinen 2004; Mishchenko and Fomenko 1988) and enjoys the following properties:

(i) Associativity:

$$\boldsymbol{\beta}_a \otimes (\boldsymbol{\beta}_b \otimes \boldsymbol{\beta}_c) = (\boldsymbol{\beta}_a \otimes \boldsymbol{\beta}_b) \otimes \boldsymbol{\beta}_c, \quad \text{for all} \quad \boldsymbol{\beta}_a, \boldsymbol{\beta}_b, \boldsymbol{\beta}_c \in \mathfrak{R}.$$

(ii) There exists a unique identity element $\boldsymbol{i} \in \mathfrak{R}$ such that

$$\boldsymbol{\beta} \otimes \boldsymbol{i} = \boldsymbol{i} \otimes \boldsymbol{\beta} = \boldsymbol{\beta}, \quad \text{for all} \quad \boldsymbol{\beta} \in \mathfrak{R}.$$

(iii) For each $\boldsymbol{\beta} \in \mathfrak{R}$ there exists a unique element belonging to $\mathfrak{R}$ called the inverse of $\boldsymbol{\beta}$ and denoted by $\boldsymbol{\beta}^{-1}$ such that

$$\boldsymbol{\beta}^{-1} \otimes \boldsymbol{\beta} = \boldsymbol{\beta} \otimes \boldsymbol{\beta}^{-1} = \boldsymbol{i}.$$

(iv) The operation $\otimes$ is, in general, noncommutative, *i.e.* ,

$$\boldsymbol{\beta}_a \otimes \boldsymbol{\beta}_b(\boldsymbol{x}) \neq \boldsymbol{\beta}_b \otimes \boldsymbol{\beta}_a(\boldsymbol{x}).$$

for all $\boldsymbol{\beta}_a, \boldsymbol{\beta}_b \in \mathfrak{R}$ and $\boldsymbol{x} \in \mathbb{E}^3$.

The following result is fundamental:

Theorem *The group $\mathfrak{R}$ is isomorphic to the set composed by all the real and orthogonal matrices of order 3, with determinant equal to 1.*

Proof See, e.g., (Pérez-Morán 2005) and references therein ■

This theorem allows to identify each finite rotation with an orthogonal *rotation tensor* belonging to the *special orthogonal group*, $SO(3)$, defined as

$$SO(3) = \left\{ \boldsymbol{\Lambda} \in \mathbf{M}^{3\times 3} \;\; | \;\; \boldsymbol{\Lambda}^t \boldsymbol{\Lambda} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^t = \mathbf{I}; \;\; \det[\boldsymbol{\Lambda}] = 1 \right\}, \tag{121}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{M}^{3\times 3}$ is the set composed by all the 3×3 matrices with real coefficients. It is not difficult to see that $SO(3)$ also has the structure of a smooth differentiable manifold (Dubrokin et al. 2000). See Sect. A.1. Moreover, under the usual matrix multiplication, it has the structure of a Lie group. See Sect. A.2.

The components of a rotation tensor depend on the reference frame adopted and thus to compose two rotations $\boldsymbol{\Lambda}_a, \boldsymbol{\Lambda}_b \in SO(3)$ two situations can happen:

(i) *Spatial description of rotations*. In this case the components of $\boldsymbol{\Lambda}_a$ and $\boldsymbol{\Lambda}_b$ are expressed in terms of a fixed reference frame (Argyris 1982). The rotation tensor representing the result of applying $\boldsymbol{\Lambda}_b$ after $\boldsymbol{\Lambda}_a$ is obtained as

$$\boldsymbol{\Lambda}_b \circ \boldsymbol{\Lambda}_a = \boldsymbol{\Lambda}_b \boldsymbol{\Lambda}_a \in SO(3),$$

Therefore, the *inverse multiplicative rule* for rotation tensors applies.

(ii) *Material description of rotations*. In the second case, $\boldsymbol{\Lambda}_a$ moves the reference frame and therefore, the components of $\boldsymbol{\Lambda}_b$ are expressed in an updated reference frame. Then, we have that

$$\boldsymbol{\Lambda}_b \circ \boldsymbol{\Lambda}_a = \boldsymbol{\Lambda}_a \boldsymbol{\Lambda}_b \in SO(3).$$

B.2 Parametrization of $SO(3)$

The rotational motion can be described by a means of a trajectory on $SO(3)$. Therefore, it can not be described trivially by using standard coordinates such as those employed for trajectories belonging to a linear space. Rotations may be parametrized using suitable charts which are inherently not global and/or singular. Over

the years, numerous techniques have been developed to cope with the description of rotational motion (Pérez-Morán 2005; Simo and Vu-Quoc 1986; Bauchau and Choi 2003; Milenkovic and Milenkovic 1997; Stuelpnagel 1964; Grassia 1998; Cottingham and Doyle 2001; Innocenti and Paganelli 2006). All these techniques show certain balance between advantages and drawbacks when compared each to other. In the following, we describe a (minimal) parametrization of rotation tensors in terms of rotation vectors. It is based on the following result:

Fundamental Theorem of Euler

The general displacement of a rigid body or vector, with one point fixed is a rotation about some axis which passes through that point.

Proof See (Crisfield 1998, Vol 2.) ■

Basically, the above theorem shows that the rotational motion is completely described by a unit vector $\hat{\mathbf{e}} \in \mathbb{R}^3$ defining an axis of rotation in space and a *rotation angle* of magnitude $\theta \in [0, 2\pi]$. Moreover, the corresponding rotation tensor is expressed according to the *Rodrigues's formula*:

$$\mathbf{\Lambda} = \mathbf{I} + \frac{\sin\theta}{\theta}\widetilde{\boldsymbol{\theta}} + \frac{(1-\cos\theta)}{\theta^2}\widetilde{\boldsymbol{\theta}}\widetilde{\boldsymbol{\theta}} = \mathbf{I} + \sin\theta\widetilde{\mathbf{e}} + (1-\cos\theta)\widetilde{\mathbf{e}}\widetilde{\mathbf{e}}. \tag{122}$$

where we use $\widetilde{\boldsymbol{u}}$ to denote the skew-symmetric tensor obtained from the vector $\boldsymbol{u} \in \mathbb{R}^3$.

B.3 Tangent Spaces

Consider $\mathbf{\Lambda} \in SO(3)$. The variations of $\mathbf{\Lambda}\mathbf{\Lambda}^t$ and $\mathbf{\Lambda}^t\mathbf{\Lambda}$ are given by

$$\begin{aligned}\delta(\mathbf{\Lambda}\mathbf{\Lambda}^t) &= \delta\mathbf{\Lambda}\mathbf{\Lambda}^t + \mathbf{\Lambda}\delta\mathbf{\Lambda}^t = \widetilde{\phi} + \widetilde{\phi}^t = 0\\ \delta(\mathbf{\Lambda}^t\mathbf{\Lambda}) &= \delta\mathbf{\Lambda}^t\mathbf{\Lambda} + \mathbf{\Lambda}^t\delta\mathbf{\Lambda} = \widetilde{\mathbf{\Phi}}^t + \widetilde{\mathbf{\Phi}} = 0,\end{aligned}$$

from which it is possible to deduce that $\widetilde{\phi}$ and $\widetilde{\mathbf{\Phi}}$ are skew-symmetric tensors. Moreover, we have that

$$\delta\mathbf{\Lambda} = \widetilde{\phi}\mathbf{\Lambda} = \mathbf{\Lambda}\widetilde{\mathbf{\Phi}}. \tag{123}$$

Clearly $\delta\mathbf{\Lambda}$ belongs to the tangent space to $SO(3)$ at $\mathbf{\Lambda}$, $T_{\mathbf{\Lambda}}SO(3)$.

The tangent space at the identity forms the Lie algebra of $SO(3)$ and is denoted by

$$so(3) = T_{\mathbf{I}}SO(3).$$

From (123) it is possible to see that $so(3)$ corresponds to the linear space of skew-symmetric tensors of the form

$$\widetilde{\boldsymbol{\theta}} = \begin{bmatrix} 0 & -\theta_3 & \theta_2 \\ \theta_3 & 0 & -\theta_1 \\ -\theta_2 & \theta_1 & 0 \end{bmatrix} = \text{skew}[\boldsymbol{\theta}]. \tag{124}$$

Since $so(3)$ is isomorphic to $\mathbb{R}^3$, every $\widetilde{\boldsymbol{\theta}} \in so(3)$ can be represented by a vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$. Moreover, $\widetilde{\boldsymbol{\Phi}} = \boldsymbol{\Lambda}^t \widetilde{\boldsymbol{\phi}} \boldsymbol{\Lambda}$ and $\boldsymbol{\Phi} = \boldsymbol{\Lambda}^t \boldsymbol{\phi}$.

The exponential map

$$\begin{aligned} \exp[\bullet] : so(3) &\to SO(3) \\ \widetilde{\boldsymbol{\theta}} &\mapsto \exp\left[\widetilde{\boldsymbol{\theta}}\right] = \boldsymbol{\Lambda}\left(\widetilde{\boldsymbol{\theta}}\right), \end{aligned} \tag{125}$$

allows to parametrize any rotation tensor in terms of an element of $so(3)$. See (122).

C Quaternions

The (minimal) vectorial parametrization of $SO(3)$ shows some limitations due to the fact that the exponential map is not a bijective application for angles greater than π (Pérez-Morán 2005; Simo and Vu-Quoc 1988; Crisfield 1998). The problem can be avoided if *unit quaternions* are used. A unit quaternion is defined using four parameters, q_0– q_3, so that:

$$\mathbf{q}_\theta = \cos(\theta/2) + \sin(\theta/2)\,\hat{\mathbf{e}} = \begin{bmatrix} \mathbf{q} \\ q_0 \end{bmatrix} = \begin{bmatrix} \sin\left(\frac{\theta}{2}\right)\hat{\mathbf{e}} \\ \cos\left(\frac{\theta}{2}\right) \end{bmatrix}, \tag{126}$$

where θ and $\hat{\mathbf{e}}$ are an angle of rotation and an axis of rotation, respectively. It is possible to see that $|\mathbf{q}_\theta| = 1$.

A rotation tensor $\boldsymbol{\Lambda} \in SO(3)$ is *uniquely* parametrized in terms of a unit quaternion $\mathbf{q}_\theta$ according to the formula:

$$\boldsymbol{\Lambda}(\mathbf{q}_\theta) = (\mathrm{q}_0^2 - \mathbf{q}\cdot\mathbf{q})\mathbf{I} + 2\mathbf{q}\otimes\mathbf{q} + 2\mathrm{q}_0\widetilde{\mathbf{q}}, \tag{127}$$

The quaternion compound rotation of the quaternions $\mathbf{q}_a = (a_0, \mathbf{a})$ and $\mathbf{q}_b = (b_0, \mathbf{b})$ is given by

$$\mathbf{q}_{ab} = \mathbf{q}_b\mathbf{q}_a = a_0 b_0 - \mathbf{a}\cdot\mathbf{b} + a_0\mathbf{b} + b_0\mathbf{a} - \mathbf{a}\times\mathbf{b}.$$

C.1 Normalized Quaternion from the Rotation Tensor

A general procedure for obtaining the rotation vector from the rotation tensor involves the computation of the Euler parameters, q_0–q_3. This can be achieved via the *Spurrier*'s algorithm (Crisfield 1998), which involves computing

$$a = \max\Big[\mathrm{Tr}[\mathbf{\Lambda}],\, \Lambda_{11},\, \Lambda_{22},\, \Lambda_{33}\Big]$$

where $\mathrm{Tr}[\bullet]$ is the trace operator and

$$\begin{aligned}
&\text{if} \quad a = \mathrm{Tr}(\mathbf{\Lambda}) \rightarrow \begin{cases} \mathrm{q}_0 = \frac{1}{2}(1+a)^{\frac{1}{2}} \\ \mathrm{q}_i = (\Lambda_{kj} - \Lambda_{jk})/4q_0;\ i = 1,3 \end{cases} \\
&\text{else if} \ \ a = \Lambda_{ii} \quad \rightarrow \begin{cases} \mathrm{q}_i = (\frac{1}{2}a + \frac{1}{4}\big[1 - \mathrm{Tr}[\mathbf{\Lambda}]\big]) \\ \mathrm{q}_0 = \frac{1}{4}(\Lambda_{kj} - \Lambda_{jk})/q_i \\ \mathrm{q}_l = \frac{1}{4}(\Lambda_{li} + \Lambda_{il})/q_i; \qquad l = j,k \end{cases}
\end{aligned}$$

where $i,\ j, k$ are a cyclic combinations of 1, 2, and 3.

References

Abraham, R., Marsden, J., & Ratiu, T. (1988). *Manifolds, Tensor Analysis, and Applications* (2nd ed., Vol. 75). Applied Mathematical Sciences. New York: Springer.

Ainsley, S., Vouga, E., Grinspun, E., & Tamstorf, R. (2012). Speculative parallel asynchronous contact mechanics. *ACM Transactions of Graphics*, *31*(6), 151:1–151:8.

Argyris, J. (1982). An excursion into large rotations. *Computer Methods in Applied Mechanics and Engineering*, *32*, 85–155.

Argyris, J., & Poterasu, V. F. (1993). Large rotations revisited application of lie algebra. *Computer Methods in Applied Mechanics and Engineering*, *103*, 11–42.

Armero, F., & Petoz, E. (1999). A new dissipative time-stepping algorithm for frictional contact problems: formulation and analysis. *Computer Methods in Applied Mechanics Engineering*, *179*, 151–178.

Armero, F., & Romero, I. (2001). On the formulation of high-frequency dissipative time-stepping algorithms for nonlinear dynamics. Part I: Low-order methods for two model problems and nonlinear elastodynamics. *Computer Methods in Applied Mechanics Engineering*, *190*, 2603–2649.

Armero, F., & Simo, J. C. (1992). A new unconditionally stable fractional step method for non-linear coupled thermomechanical problems. *International Journal for Numerical Methods in Engineering*, *35*(4), 737–766.

Arnold, V. I. (1989). *Mathematical methods of classical mechanics* (2nd ed., Vol. 60). Graduate texts in mathematics. New York: Springer.

Arnold, V. I. & Khesin, B. A. (1998). *Topological methods in hydrodynamics* (Vol. 125). Applied Mathematical Sciences. New York: Springer.

Arnold, V. I., Kozlov, V. V., & Neishtadt, A. I. (2006). *Mathematical aspects of classical and celestial mechanics* (3rd ed., Vol. 3)., Encyclopaedia of mathematical sciences. Berlin: Springer.

Atluri, S. N., & Cazzani, A. (1995). Rotations in computational solid mechanics. *Archives of Computational Methods in Engineering*, *2*(1), 49–138.

Bargmann, S., & Steinmann, P. (2008a). Modeling and simulation of first and second sound in solids. *International Journal of Solids and Structures*, *45*, 6067–6073.
Bargmann, S., & Steinmann, P. (2008b). An incremental variational formulation of dissipative and non-dissipative coupled thermoelasticity for solids. *Heat Mass Transfer*, *45*, 107–116.
Bathe, K. J. (1996). *Finite Element Procedures*. Prentice-Hall.
Bauchau, O., & Bottasso, C. (1999). On the design of energy preserving and decaying schemes for flexible, nonlinear multi-body systems. *Computer Methods in Applied Mechanics and Engineering*, *169*, 61–79.
Bauchau, O., & Trainelli, L. (2003). The vectorial parametrization of rotation. *Nonlinear Dynamics*, *32*, 71–92.
Bauchau, O. A., & Choi, J. I. (2003). The vector parameterization of motion. *Nonlinear Dynamics*, *33*, 165–188.
Bayliss, A., & Issacson, E. (1975). How to make your algorithm conservative. *American Mathematical Society*, *22*, A594–A595.
Betsch, P., & Uhlar, S. (2007). Energy-momentum conserving integration of multibody dynamics. *Multibody System Dynamics*, *17*, 243–289.
Betsch, P., Hesch, C., Sänger, N. & Uhlar, S. (2010). Variational integrators and energy-momentum schemes for flexible multibody dynamics. *Journal of Computational and Nonlinear Dynamics*, **5**(3):031001/1–11.
Bogolyubov, N. N. (1972). Part 2, chapter the model hamiltonian in superconductivity theory. *Particles and nuclei* (Vol. 1, pp. 1–52). US: Springer.
Borri, M., Bottasso, L., & Trainelli, L. (2001). Integration of elastic multibody system by invariant conserving/dissipating algorithms. I. Formulation. *Computer Methods in Applied Mechanics and Engineering*, *190*, 3669–3699.
Bou-Rabee, N., & Marsden, J. E. (2009). Hamilton-pontryagin integrators on lie groups part I: Introduction and structure-preserving properties. *Foundations of Computational Mathematics*, *9*(2), 197–219.
Bou-Rabee, N., & Owhadi, H. (2007). Stochastic variational partitioned runge-kutta integrators for constrained systems, arxiv.org/abs/0709.2222.
Bou-Rabee, N., & Owhadi, H. (2009). Stochastic variational integrators. *IMA Journal of Numerical Analysis*, *29*(2), 421–443.
Bou-Rabee, N., & Owhadi, H. (2010). Long-run accuracy of variational integrators in the stochastic context. *SIAM Journal on Numerical Analysis*, *48*(1), 278–297.
Cannarozzi, A. A., & Ubertini, F. (2001). A mixed variational method for linear coupled thermoelastic analysis. *International Journal of Solids and Structures*, *38*, 717–739.
Cano, B., & Sanz-Serna, J. M. (1988). Error growth in the numerical integration of periodic orbits by multistep methods, with application to reversible systems. *IMA Journal of Numerical Analysis*, *18*, 57–75.
Celledoni, E., & Owren, B. (2003). Lie group methods for rigid body dynamics and time integration on manifolds. *Computer Methods in Applied Mechanics and Engineering*, *192*(34), 421–438.
Celledoni, E., Marthinsen, H. & Owren, B. (2014). An introduction to lie group integrators - basics, new developments and applications. *Journal of Computational Physics*, **257**(Part B) 1040–1061.
Channell, P. J., & Scovel, C. (1990). Symplectic integration of hamiltonian systems. *Nonlinearity*, *3*, 231–259.
Chaturvedi, N. A., Lee, T., Leok, M., & McClamroch, N. H. (2011). Nonlinear Dynamics of the 3D Pendulum. *Journal of Nonlinear Science*, *21*(1), 3–32.
Chen, Y. (1990). A proof of the structure of the minimum-time control law of robotic manipulators using a hamiltonian formulation. *IEEE Robotics and Automation*, *6*(3), 388–393.
Chyba, M., Hairer, E., & Vilmart, G. (2009). The role of symplectic integrators in optimal control. *Optimal Control Applications and Methods*, *30*(4), 367–382.
Cirak, F., & West, M. (2005). Decomposition contact response (DCR) for explicit finite element dynamics. *International Journal for Numerical Methods in Engineering*, *64*(8), 1078–1110.

Clemente-Gallardo, J., & Scherpen, J. M. A. (2003). Relating lagrangian and hamiltonian formalisms of LC circuits. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, *50*(10), 1359–1363.

Cohen, G., Joly, P., Roberts, J. E., & Tordjman, N. (2001). Higher order triangular finite elements with mass lumping for the wave equation. *SIAM Journal on Numerical Analysis*, *38*(6), 2047–2078.

Coleman, B. D., & Noll, W. (1963). The thermodynamics of elastic materials with heat conduction and viscosity. *Archive for Rational Mechanics and Analysis*, *13*(3), 167–178.

Cottingham, W. N., & Doyle, D. D. (2001). The rotational dynamics of rigid bodies implemented with the cayley klein parametrization. *Molecular Physics*, *99*, 1839–1843.

Crisfield, M. A. (1998). *Non-linear finite element analysis of solids and structures* (Vol. 1, 2). Wiley.

De León, M., Marrero, J. C., & Martín De Diego, D. (2008). Some applications of semi-discrete variational integrators to classical field theories. *Qualitative Theory of Dynamical Systems*, *7*(1), 195–212.

Demoures, F., Gay-Balmaz, F., Kobilarov, M., & Ratiu, T. S. (2014). Multisymplectic lie group variational integrator for a geometrically exact beam in R^3. *Communications in Nonlinear Science and Numerical Simulation*, *19*(10), 3492–3512.

Demoures, F. M. A. (2012). *Lie group and lie algebra variational integrators for flexible beam and plate in* R^3. Ph.D thesis, École Polytechnique Fédérale de Lausanne.

Desbrun, M., Gawlik, E. S., Gay-Balmaz, F., & Zeitlin, V. (2014). Variational discretization for rotating stratified fluids. *Discrete and Continuous Dynamical Systems—Series A (DCDS-A)*, *34*(2), 477–509.

Dubrokin, B. A., Fomenko, A. T., & Nóvikov, S. P. (2000). *Geometría moderna. métodos y aplicaciones* (Vol. 1 & 2). Moscow: Mir, URSS.

Esposito, G., Marmo, G., & Sudarshan, G. (2004). *From classical to quantum mechanics: an introduction to the formalism foundations and applications*. New York: Cambridge University Press.

Faltinsen, S. (2000). Backward error analysis for Lie-group methods. *BIT Numerical Mathematics*, *40*(4), 652–670.

Feng, K., & Qin, M. (2010). *Symplectic Geometric Algorithms for Hamiltonian Systems*. Berlin: Zhejiang Publishing United Group Zhejiang Science and Technology Publishing House, Hangzhou and Springer.

Fetecau, R. C. (2003). *Variational methods for nonsmooth mechanics*. Ph.D thesis, California Institute of Technology, Pasadena, California, USA.

Fetecau, R. C., Marsden, J. E., & West, M. (2003a). Variational multisymplectic formulations of nonsmooth continuum mechanics. *Perspectives and problems in nonlinear science* (pp. 229–261). New York: Springer.

Fetecau, R. C., Marsden, J. E., Ortiz, M., & West, M. (2003b). Nonsmooth Lagrangian mechanics and variational collision integrators. *SIAM Journal on Applied Dynamical Systems*, *2*(3), 381–416.

Focardi, M., & Maria-Mariano, P. (2008). Convergence of asynchronous variational integrators in linear elastodynamics. *International Journal for Numerical Methods in Engineering*, *75*(7), 755–769.

Fong, W., Darve, E., & Lew, A. (2008). Stability of asynchronous variational integrators. *Journal of Computational Physics*, *227*, 8367–8394.

Gambar, K., & Markus, F. (1994). Hamilton-Lagrange formalism of nonequilibrium thermodynamics. *Physical Review E*, *50*(2), 1227–1231.

Gawlik, E. S., Mullen, P., Pavlov, D., Marsden, J. E., & Desbrun, M. (2011). Geometric, variational discretization of continuum theories. *Physica D: Nonlinear Phenomena*, *240*(21), 1724–1760.

Gay-Balmaz, F., Holm, D. D., & Ratiu, T. S. (2009). Variational principles for spin systems and the Kirchhoff rod. *The Journal of Geometric Mechanics (JGM)*, *1*(4), 417–444.

Ge, Z., & Marsden, J. E. (1988). Lie-poisson hamilton-jacobi theory and lie-poisson integrators. *Physics Letters A*, *133*(3), 134–139.

Gonzalez, O. (2000). Exact energy and momentum conserving algorithms for general models in nonlinear elasticity. *Computer Methods in Applied Mechanics and Engineering*, *190*, 1763–1783.

Grassia, F. S. (1998). Practical parameterization of rotations using the exponential map. *Journal of Graphics Tools*, *3*, 29–48.

Green, A., & Naghdi, P. (1991). A re-examination of the basic postulates of thermomechanics. *Proceedings: Mathematical and Physical Sciences*, *432*, 171–194.

Green, A., & Naghdi, P. (1995). A unified procedure for construction of theories of deformable media. I. classical continuum physics. *Mathematical and Physical Sciences*, *448*(1934), 335–356.

Green, A. E., & Naghdi, P. M. (1993). Thermoelasticity without energy dissipation. *Journal of Elasticity*, *31*(3), 189–208.

Grispun, E., Hirani, A., Desbrun, M. & Schröder, P. (2003). Discrete shells. In *Symposium on computer animation*, (pp. 62–67), San Diego, California.

Gross, M., & Betsch, P. (2006). An energy consistent hyprid space-time Galerkin method for nonlinear thermomechanical problems. *PAMM, Proceedings in Applied Mathematics and Mechanics*, *6*, 443–444.

Gross, M., Betsch, P. (2007). On deriving higher-order and energy-momentum-consistent time-stepping-schemes for thermo-viscoelastodynamics from a new hybrid space-time Galerkin method. In Bottasso, C. L., Masarati, P., & Trainelli, L., (Eds.), *Proceedings of the ECCOMAS Thematic Conference on Multibody Dynamics*, Milano, Italy: Politecnico di Milano.

Hairer, E., & Lubich, C. (1999). Invariant tori of dissipatively perturbed hamiltonian systems under symplectic discretization. *Applied Numerical Mathematics*, *29*(1), 57–71.

Hairer, E., & Wanner, G. (1996). *Solving ordinary differential equations II. Stiff and differential-algebraic problems* (Vol. 14). Springer Series in Computational Mathematics. Berlin: Springer.

Hairer, E., Norsett, S. P. & Wanner, G. (1993). *Solving ordinary differential equations I. Nonstiff problems*, (Vol 8). Springer Series in Computational Mathematics. Berlin: Springer.

Hairer, E., Lubich, C., & Wanner, G. (2003). Geometric numerical integration illustrated by the Störmer-Verlet method. *Acta Numerica*, *12*, 399–450.

Hairer, E., Lubich, C., & Wanner, G. (2006). *Geometric numerical integration. Structure preserving algorithms for ordinary differential equations*. Springer Series in Computational Mathematics. Springer.

Hall, J. & Leok, M. (2014a). Spectral variational integrators. *Numerische Mathematik*.

Hall, J. & Leok, M. (2014b). Spectral variational integrators. arXiv:1402.3327.

Harmon, D., Vouga, E., Smith, B., Tamstorf, R. & Grinspun, E. (2009). Asynchronous contact mechanics. In *SIGGRAPH'09 (ACM Transactions on Graphics)*, New York, USA: ACM, ISBN: 978-1-60558-726-4.

Holmes, M. H. (2007). *Introduction to numerical methods in differential equations* (Vol. 52). Texts in applied mathematics. New York: Springer.

Hughes, T. J. R. (1987). *The finite element method: linear static and dynamic finite element analysis*. Prentice Hall Inc.

Hutter, M., & Tervoort, T. A. (2007). Finite anisotropic elasticity and material frame indifference from a nonequilibrium thermodynamics perspective. *Journal of Non-Newtonian Fluid Mechanics*, *152*, 45–52.

Innocenti, C., & Paganelli, D. (2006). *Advances in robot kinematics mechanisms and motion*, chapter Determining the 3×3 rotation matrices that satisfy three linear equations in the direction cosines. Springer.

Iserles, A. (1997). *Foundations of computational mathematics*, chapter Numerical methods on (and off) manifolds, (pp. 180–189). Number 10208. Berlin: Springer.

Iserles, A., Munthe-Kaas, H. Z., Norsett, S. P., & Zanna, A. (2000). Lie-group methods. *Acta Numerica*, *9*, 215–365.

Jelenič, G. & Crisfield, M. A. (1999) Geometrically exact 3d beam theory: implementation of a strain-invariant finite element for static and dynamics. *Computer Methods in Applied Mechanics and Engineering*, *171*, 141–171.

Jiménez, F., Kobilarov, M., & Martín de Diego, M. (2013). Discrete variational optimal control. *Journal of Nonlinear Science*, *23*(3), 393–426.

Johnson, G., Leyendecker, S., & Ortiz, M. (2014). Discontinuous variational time integrators for complex multibody collisions. *International Journal for Numerical Methods in Engineering*, *100*(12), 871–913.

Kale, K. G., & Lew, A. J. (2006). Parallel asynchronous variational integrators. *International Journal for Numerical Methods in Engineering*, *70*(3), 291–321.

Kane, C., Marsden, J. E., & Ortiz, M. (1999). Symplectic-energy-momentum preserving variational integrators. *Journal of mathematical physics*, *40*(7), 3353–3371.

Kane, C., Marsden, J. E., Ortiz, M., & West, M. (2000). Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *International Journal for Numerical Methods in Engineering*, *49*, 1295–1325.

Kapania, R. K., & Li, J. (2003). On a geometrically exact curved/twisted beam theory under rigid cross-section assumption. *Computational Mechanics*, *30*, 428–443.

Kasdin, N. J., Gurfil, P., & Kolemen, E. (2005). Canonical modelling of relative spacecraft motion via epicyclic orbital elements. *Celestial Mechanics and Dynamical Astronomy*, *92*.

Kern, D., Bär, S. & Groß, M. (2014). Variational integrators for thermomechanical coupled dynamic systems with heat conduction. *Proceedings in applied mathematics and mechanics, PAMM*, *14*(1), 47–48.

Kharevych, L., Weiwei, Y., Tong, Y., Kanso, E., Marsden, J. E., Schröder, P., & Desbrun, M. (2006). Geometric, variational integrators for computer animation. In *Eurographics/ACM SIGGRAPH Symposium on Computer Animation*.

Kirwan, A. D. (2008). Quantum and ecosystem entropies. *Entropy*, *10*, 58–70.

Kobilarov, M. (2014). *Multibody dynamics*, chapter Solvability of geometric integrators for multibody systems. (Vol. 35, pp. 145–174), Computational methods in applied sciences. Switzerland: Springer International Publishing.

Koon, W. S., Lo, M. W., Marsden, J. E. & Ross, S. D. (2011). *Dynamical systems, the three-body problem and space mission design*. Marsden Books.

Kraus, M. (2013). *Variational integrators in plasma physics*. Ph.D thesis, Technische Universität München.

Kuang, J., Leung, A. Y. T., & Tan, S. (2003). Hamiltonian and chaotic attitude dynamics of an orbiting gyrostat satellite under gravity-gradient torques. *Physica D: Nonlinear Phenomena*, *186*(1–2), 1–19.

Labudde, R. A., & Greenspan, D. (1976). Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion part II. *Numerisch Mathematik*, *26*, 1–16.

Larsson, J. (1996). A new hamiltonian formulation for fluids and plasmas. part 3. multifluid electrodynamics. *Journal of Plasma Physics*, *55*(02), 279–300.

Lee, T., Leok, M., & McClamroch, N. H. (2007). Lie group variational integrators for the full body problem. *Computer Methods in Applied Mechanics and Engineering*, *196*(29–30), 2907–2924.

Lee, T., Leok, M., & McClamroch, N. H. (2009). Lagrangian mechanics and variational integrators on two-spheres. *International Journal for Numerical Methods in Engineering*, *79*(9), 1147–1174.

Leimkuhler, B., & Reich, S. (2005). *Simulating hamiltonian dynamics*. Cambridge Monographs on Applied and Computational Mathematics.

Leok, M. (2005). Generalized galerkin variational integrators. arXiv:math/0508360v1.

Leok, M., & Shingel, T. (2012a). General techniques for constructing variational integrators. *Frontiers of Mathematics in China*, *7*(2), 273–303.

Leok, M., & Shingel, T. (2012b). Prolongationcollocation variational integrators. *IMA Journal of Numerical Analysis*, *32*, 1194–1216.

Lew, A. (2003). *Variational time integrators in computational solid mechanics*. Ph.D thesis, California Institute of Technology, Pasadena, California, USA.

Lew, A., Marsden, J. E., Ortiz, M., & West, M. (2003). Asynchronous variational integrators. *Archive for Rational Mechanics and Analysis*, *2*, 85–146.

Lew, A., Marsden, J. E., Ortiz, M., & West, M. (2004). Variational time integrators. *International Journal for Numerical Methods in Engineering*, *60*, 153–212.

Leyendecker, S., Ober-Blöbaum, S., Marsden, J. E., & Ortiz, M. (2007). Discrete mechanics and optimal control for constrained multibody dynamics. In *Proceedings of the 6th International Conference on Multibody Systems, Nonlinear Dynamics, and Control, ASME* (pp. 1–10).

Leyendecker, S., Marsden, J. E., & Ortiz, M. (2008). Variational integrators for constrained dynamical systems. *ZAMM—Journal of Applied Mathematics and Mechanics*, *88*(9), 677–708.

Leyendecker, S., Hartmann, C., & Koch, M. (2012). Variational collision integrator for polymer chains. *Journal of Computational Physics*, *231*(10), 3896–3911.

Luo, M. Q., Liu, H., & Li, Y. M. (2013). Seismic wave modeling with implicit symplectic method based on spectral factorization on helix. *Chinese Journal of Geophysics*, *44*(3), 376–385.

Macchelli, A., Melchiorri, C., & Stramigioli, S. (2009). Port-based modeling and simulation of mechanical systems with rigid and flexible links. *IEEE Transactions on Robotics*, *25*(5), 1016–1029.

Maddocks, J. H. & Overton, M. L. (1995). Stability theory for dissipatively perturbed Hamiltonian systems. *Communications on pure and applied mathematics*, *XLVIII*, 583–610.

Maeda, S. (1980). Canonical structure and symmetries for discrete systems. *Mathematica Japonica*, *25*, 405–420.

Maeda, S. (1982). Lagrangian formulation of discrete systems and concept of difference space. *Mathematica Japonica*, *27*, 345–356.

Mäkinen, J. (2004). *A Formulation for flexible multibody mechanics. Lagrangian geometrically exact beam elements using constrain manifold parametrization*. Ph.D thesis, Tampere University of Technology, Institute of Applied Mechanics and Optimization.

Mäkinen, J., & Marjamäki, H. (2005). Total lagrangian parametrization of rotation manifold. In *ENOC-2005, Fifth EUROMECH Nonlinear Dynamics Conference*, (pp. 522–530).

Manning, R. S., & Maddocks, J. H. (1999). Symmetry breaking and the twisted elastic ring. *Computer Methods in Applied Mechanics and Engineering*, *170*(3–4), 313–330.

Marsden, J. E. (1988). The hamiltonian formulation of classical field theory. *Contemporary Mathematics*, *71*, 221–235.

Marsden, J. E., & Hughes, T. J. R. (1983). *Mathematical foundations of elasticity*. Prentice-Hall.

Marsden, J. E., & Ratiu, T. (1999). *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*. Springer-Verlag Gmbh.

Marsden, J. E., & Wendlandt, J. M. (1997). chapter Mechanical systems with symmetry, variational principles, and integration algorithms. *Current and future directions in applied mathematics* (pp. 219–261). Boston: Birkhuser.

Marsden, J. E., & West, W. (2001). Discrete mechanics and variational integrators. *Acta Numerica*, *10*, 357–514.

Marsden, J. E., Patrick, G. W., & Shkoller, S. (1998). Multisymplectic geometry, variational integrators, and nonlinear pdes. *Communications in Mathematical Physics*, *199*(2), 351–395.

Marsden, J. E., Pekarsky, S., & Shkoller, S. (1999). Discrete euler-poincar and lie-poisson equations. *Nonlinearity*, *12*(6), 1647–1662.

Marsden, J. E., Pekarsky, S., & Shkoller, S. (2000). Symmetry reduction of discrete lagrangian mechanics on lie groups. *Journal of Geometry and Physics*, *36*(1–2), 140–151.

Marsden, J. E., Pekarsky, S., Shkoller, S., & West, M. (2001). Variational methods, multisymplectic geometry and continuum mechanics. *Journal of Geometry and Physics*, *38*, 253–284.

Mata, P. (2015). Explicit symplectic momentum-conserving time-stepping scheme for the dynamics of geometrically exact rods. *Finite Elements in Analysis and Design*, *96*, 11–22.

Mata, P., & Lew, A. (2011). Variational time integrators for finite dimensional thermo-elastodynamics without heat conduction. *International Journal for Numerical Methods in Engineering*, *88*(1), 1–30.

Mata, P., & Lew, A. (2012). Structure-preserving time integrators for thermo-elasticity with heat conduction. Abstract in the *European Congress on Computational Methods in Applied Sciences and Engineering* Vienna, Austria, 10–14 September.

Mata, P. & Lew, A. (2014). Variational integrators for the dynamics of thermo-elastic solids with finite speed thermal waves. *Journal of Computational Physics*, *257*(Part B), 1423–1443.

Mata, P., Oller, S., & Barbat, A. H. (2008). Dynamic analysis of beam structures considering geometric and constitutive nonlinearity. *Computer Methods in Applied Mechanics and Engineering*, *197*, 857–878.

Mata, P., Barbat, A. H., Oller, S., & Boroschek, R. (2009). Non-linear seismic analysis of rc structures with energy-dissipating devices. *International Journal for Numerical Methods in Engineering*, *78*(9), 1037–1075.

Maugin, G. A. (2000). Towards an analytical mechanics of dissipative materials. *Rendiconti del Seminario Matematico. Geometry, Continua and Microstuctures. Universita e Politecnico di Torino*, Torino, *58*(2), 171–180.

Maugin, G. A., & Kalpakides, V. K. (2002). A Hamiltonian formulation for elasticity and thermoelasticity. *Journal of Physics A: Mathematical and General*, *35*, 10775–10788.

McLachlan, R. I., Perlmutter, M., & Quispel, G. R. W. (2004). On the nonlinear stability of symplectic integrators. *BIT Numerical Mathematics*, *44*, 99–117.

Mcrobie, F. A., & Lasenby, J. (1999). Simo-Vu Quoc rods using Clifford algebra. *International Journal for Numerical Methods in Engineering*, *45*, 377–398.

Meyer, K. R., Hall, G. R., & Offin, D. (2009). *introduction to hamiltonian dynamical systems and the n-body problem* (2nd ed., Vol. 90)., Applied mathematical sciences. Springer.

Meyer, K. R., Palacián, J. F., & Yanguas, P. (2011). Geometric averaging of hamiltonian systems: Periodic solutions, stability, and KAM tori. *SIAM Journal on Applied Dynamical Systems (SIADS)*, *10*(3), 817–856.

Milenkovic, V. J., & Milenkovic, V. (1997). Rational orthogonal approximations to orthogonal matrices. *Computational Geometry: Theory and Applications*, *7*, 25–32.

Mishchenko, A., & Fomenko, A. (1988). *A course of differential geometry and topology*. Moscow: Mir Publisher.

Moser, J., & Veselov, A. P. (1991). Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Communications in Mathematical Physics*, *139*(2), 217–243.

Nichols, K., & Murphey, T. D. (2008). Variational integrators for constrained cables. In *IEEE International Conference on Automation Science and Engineering, 2008. CASE 2008*, (pp. 802–807), Arlington, VA, August 2008.

Noether, E. (1918). Invariante variationsprobleme. *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys. Klasse*, *VI.6*, 235–257.

Ober-Blöbauma, S., Tao, M., Cheng, M., Owhadi, H., & Marsden, J. E. (2013). Variational integrators for electric circuits. *Journal of Computational Physics*, *242*, 498–530.

Patrick, G. W., & Cuell, C. (2009). Error analysis of variational integrators of unconstrained lagrangian systems. *Numerische Mathematik*, *113*(2), 243–264.

Pavlov, D. (2009). *Structure-preserving discretization of incompressible fluids*. Ph.D thesis, California Institute of Technology, Pasadena, California.

Pavlov, D., Mullen, P., Tong, Y., Kanso, E., Marsden, J. E., & Desbrun, M. (2011). Structure-preserving discretization of incompressible fluids. *Physica D: Nonlinear Phenomena*, *240*(6), 443–458.

Pérez-Morán, A. (2005). *Formulaciones tangente y secante en análisis no lineal de vigas de Cosserat*. Ph.D thesis, Universitat Politècnica de Catalunya, Spain.

Poincaré, H. (1899). *Les Méthodes Nouvelles de la Mécanique Céleste. Tome III*. Gauthiers-Villars.

Romero, I. (2009). Thermodynamically consistent time-stepping algorithms for non-linear thermomechanical systems. *International Journal for Numerical Methods in Engineering*, *79*, 706–732.

Romero, I. (2010). Algorithms for coupled problems that preserve symmetries and the laws of thermodynamics part I: Monolithic integrators and their application to finite strain thermoelasticity. *Computer Methods in Applied Mechanics and Engineering*, *199*, 1841–1858.

Romero, I., & Armero, F. (2002). An objective finite element approximation of the kinematics of geometrically exact rods and its use in the formulation of an energymomentum conserving scheme in dynamics. *International Journal for Numerical Methods in Engineering*, *54*, 1683–1716.

Ryckman, R. & Lew, A. (2010). Explicit asynchronous contact algorithm for elastic-rigid body interaction. In *Proceedings of the First International Conference in Computational Contact Mechanics*.

Ryckman, R., & Lew, A. (2011). *Trends in computational contact mechanics*, chapter Explicit asynchronous contact algorithm for elastic-rigid body interaction. (Vol. 58, pp. 169–191)., Lecture notes in applied and computational mechanics. Berlin: Springer.

Ryckman, R. A., & Lew, A. J. (2012). An explicit asynchronous contact algorithm for elastic bodyrigid wall interaction. *International Journal for Numerical Methods in Engineering*, *89*(7), 869–896.

Sansour, C., & Wagner, W. (2003). Multiplicative updating of the rotation tensor in the finite element analysis of rods and shells—a path independent approach. *Computational Mechanics*, *31*(1–2), 153–162.

Schmidt, B., Leyendecker, S., & Ortiz, M. (2009). $\gamma-$convergence of variational integrators for constrained systems. *Journal of Nonlinear Science*, *19*, 153–177.

Simo, J. C. (1985). A finite strain beam formulation. the three-dimensional dynamic problem. part i. *Computer Methods in Applied Mechanics and Engineering*, *49*, 55–70.

Simo, J. C., & Miehe, C. (1992). Associative coupled thermoplasticity at finite strains: Formulation, numerical analysis and implementation. *Computer Methods in Applied Mechanics and Engineering*, *98*(1), 41–104.

Simo, J. C., & Tarnow, N. (1994). A new energy and momentum conserving algorithm for the nonlinear dynamics of shells. *International Journal for Numerical Methods in Engineering*, *37*(15), 2527–2549.

Simo, J. C., & Vu-Quoc, L. (1986). A three-dimensional finite-strain rod model. part II: Computational aspects. *Computer Methods in Applied Mechanics and Engineering*, *58*, 79–116.

Simo, J. C., & Vu-Quoc, L. (1988). On the dynamics in space of rods undergoing large motions—a geometrically exact approach. *Computer Methods in Applied Mechanics and Engineering*, *66*, 125–161.

Simo, J. C., & Wong, K. K. (1991). Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum. *International Journal for Numerical Methods in Engineering*, *31*(1), 19–52.

Simo, J. C., Marsden, J. E., & Krishnaprasad, P. S. (1988). The hamiltonian structure of nonlinear elasticity: The material and convective representations of solids, rods, and plates. *Archive for Rational Mechanics and Analysis*, *104*(2), 125–183.

Simo, J. C., Tarnow, N., & Wong, K. K. (1992). Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics. *Computer Methods in Applied Mechanics and Engineering*, *100*, 63–116.

Simo, J. C., Tarnow, N., & Doblare, M. (1995). Non-linear dynamics of three-dimensional rods: Exact energy and momentum conserving algorithms. *International Journal for Numerical Methods in Engineering*, *38*(9), 1431–1473.

Spurrier, R. A. (1978). Comment on singularity-free extraction of a quaternion from a direction-cosine matrix. *Journal of Spacecraft and Rockets*, *15*(4), 255–255.

Stavros, F. (2014). *Nonlinear Hamiltonian mechanics applied to molecular dynamics. Theory and computational methods for understanding molecular spectroscopy and chemical reactions*. Springer.

Stern, A., & Grinspun, E. (2009). Implicit-explicit variational integration of highly oscillatory problems. *Multiscale Modelling and Simulation*, *7*, 1779–1794.

Stoffer, D. (1997). On the qualitative behaviour of symplectic integrators part I: Perturbed linear systems. *Numerische Mathematik*, *77*(4), 535–547.

Stoffer, D. (1998). On the qualitative behaviour of symplectic integrators. part III. Perturbed integrable systems. *Journal of Mathematical Analysis and Applications*, *217*(2), 521–545.

Stuelpnagel, J. (1964). On the parametrization of the three-dimensional rotation group. *SIAM Review*, *6*, 422–430.

Tao, M., Owhadi, H., & Marsden, J. E. (2010). Nonintrusive and structure preserving multiscale integration of stiff ODEs, SDEs, and hamiltonian systems with hidden slow dynamics via flow averaging. *Multiscale Modeling and Simulation*, *8*(4), 1269–1324.

Van Bargena, H., & Dimitroff, G. (2009). Isotropic ornstein-uhlenbeck flows. *Stochastic Processes and their Applications*, *119*(7), 2166–2197.

Veselov, A. P. (1988). Integrable discrete-time systems and difference operators. *Functional Analysis and Its Applications*, *22*(2), 83–93.

Vujanovic, B., & Djukic, D. J. (1971). On the variational principle of Hamilton's type for nonlinear heat transfer problem. *International Journal of Heat Mass Transfer*, *15*, 1111–1123.

Wang, L. (2007). *Variational integrators and generating functions for stochastic Hamiltonian systems*. Ph.D thesis, Universität Karlsruhe, Germany.

Wang, L., Hong, J., Scherer, R., & Bai, F. (2009). Dynamics and variational integrators of stochastic Hamiltonian systems. *International Journal of Numerical Analysis and Modeling*, *6*(4), 586–602.

Wendlandt, J. M., & Marsden, J. E. (1997). Mechanical integrators derived from a discrete variational principle. *Physica D: Nonlinear Phenomena*, *106*(3–4), 223–246.

West, W. (2004). *Variational integrators*. Ph.D thesis, California Institute of Technology, Pasadena, California, USA.

Wolff, S., & Bucher, C. (2013). Asynchronous collision integrators: Explicit treatment of unilateral contact with friction and nodal restraints. *International Journal for Numerical Methods in Engineering*, *95*(7), 562–586.

Yang, Q., Stainier, L., & Ortiz, M. (2006). A variational formulation of the coupled thermo-mechanical boundary-value problem for general dissipative solids. *Journal of the Mechanics and Physics of Solids*, *56*, 401–424.

Yoshida, H. (1993). Recent progress in the theory and application of symplectic integrators. *Celestial Mechanics and Dynamical Astronomy*, *56*(1–2), 27–43.