

# Hamilton–Pontryagin Integrators on Lie Groups Part I: Introduction and Structure-Preserving Properties

Nawaf Bou-Rabee · Jerrold E. Marsden

Received: 5 February 2007 / Revised: 7 January 2008 / Accepted: 29 February 2008  
© SFoCM 2008

**Abstract** In this paper, structure-preserving time-integrators for rigid body-type mechanical systems are derived from a discrete Hamilton–Pontryagin variational principle. From this principle, one can derive a novel class of variational partitioned Runge–Kutta methods on Lie groups. Included among these integrators are generalizations of symplectic Euler and Störmer–Verlet integrators from flat spaces to Lie groups. Because of their variational design, these integrators preserve a discrete momentum map (in the presence of symmetry) and a symplectic form.

In a companion paper, we perform a numerical analysis of these methods and report on numerical experiments on the rigid body and chaotic dynamics of an underwater vehicle. The numerics reveal that these variational integrators possess structure-preserving properties that methods designed to preserve momentum (using the coadjoint action of the Lie group) and energy (for example, by projection) lack.

**Keywords** Variational integrators · Hamilton–Pontryagin · Lie group integrators

**Mathematics Subject Classification (2000)** 37M15 · 58E30 · 65P10 · 70EXX · 70HXX

---

Communicated by Arieh Iserles.

Research partially supported by the National Science Foundation through NSF grant DMS-0204474.

N. Bou-Rabee (✉)

Applied and Computational Mathematics, Caltech, Pasadena, CA 91125, USA

e-mail: [nawaf@acm.caltech.edu](mailto:nawaf@acm.caltech.edu)

J.E. Marsden

Control and Dynamical Systems, Caltech, Pasadena, CA 91125, USA

e-mail: [marsden@cds.caltech.edu](mailto:marsden@cds.caltech.edu)

# 1 Introduction

**Overview** This paper is concerned with efficient, structure-preserving time integrators for mechanical systems whose configuration space is a Lie group, based on the Hamilton–Pontryagin (HP) variational principle [11, 12, 20, 34, 35]. This HP principle has many attractive theoretical properties; for instance, how it handles degenerate Lagrangian systems. The present paper shows that the HP viewpoint also provides a practical way to design discrete Lagrangians, which are the cornerstone of variational integration theory. This overview explains the central idea of this paper in the context of vector spaces and shows how this idea extends to Lie groups.

The HP principle states that a mechanical system traverses a path that extremizes the following HP action integral:

$$S = \underbrace{\int_a^b L(q, v) dt}_{\text{Lagrangian}} + \underbrace{\int_a^b \langle p, \dot{q} - v \rangle dt}_{\text{kinematic constraint}}. \quad (1.1)$$

The integrand of the HP action integral consists of two terms: the Lagrangian and a kinematic constraint paired with a Lagrange multiplier (the momentum). The kinematic constraint relates the mechanical system’s velocity to a curve on the tangent bundle. In this principle, the curves  $q(t)$ ,  $v(t)$ , and  $p(t)$  are all varied independently. If  $p(t)$  is varied first, it collapses to the *usual Hamilton principle*. If on the other hand,  $v(t)$  is varied first, it defines the (negative of the) Hamiltonian as the extrema of the terms involving  $v$  and then the principle reduces to *Hamilton’s phase space principle*. This HP form of the action integral makes it amenable to discretization.

In particular, one can implement an  $s$ -stage Runge–Kutta (RK) discretization of the kinematic constraint and enforce this discretization as a constraint in a discrete action sum. The motivation is that the theory, order conditions, and implementation of such methods are mature. For this purpose, let  $[a, b]$  and  $N$  be given, and define the fixed step size  $h = (b - a)/N$  and  $t_k = hk + a$ ,  $k = 0, \dots, N$ . Let  $s$  be the number of internal stages in the RK method. In analogy with the continuous system, the discrete HP action sum takes the following form:

$$S_d = \underbrace{\sum_{k=0}^{N-1} \sum_{i=1}^s hb_i L(Q_k^i, V_k^i)}_{\text{discrete Lagrangian}} + \underbrace{\sum_{k=0}^{N-1} \sum_{i=1}^s h \left\langle p_k^i, \frac{Q_k^i - q_k}{h} - \sum_{j=1}^s a_{ij} V_k^j \right\rangle + h \left\langle p_k, \frac{q_{k+1} - q_k}{h} - \sum_{j=1}^s b_j V_k^j \right\rangle}_{\text{discrete kinematic constraint}}. \quad (1.2)$$

It consists of two parts: a weighted sum of the Lagrangian using the weights from the Butcher tableau of the RK scheme, and pairings between discrete internal and external stage Lagrange multipliers and the discretized kinematic constraint. This strategy

yields a Lagrangian analog of a well-known class of symplectic partitioned Runge–Kutta methods including the Lobatto IIIA–IIIB pair which generalize to higher-order accuracy [8, 23, 32].

In the Lie group context, one can generalize this strategy using either constrained or generalized coordinates. To use constrained coordinates, one treats the system as a holonomically constrained mechanical system. In this approach, one assumes that  $G$  can be written as the level set of some function  $g : \mathbb{R}^n \rightarrow \mathbb{R}^k$ , embeds  $G$  in a larger linear space, and uses Lagrange multipliers to enforce the constraint. The corresponding constrained action takes the following form:

$$S_d^c = \sum_{k=0}^{N-1} \sum_{i=1}^s h \left[ b_i L(Q_k^i, V_k^i) + \left\langle p_k^i, \frac{Q_k^i - q_k}{h} - \sum_{j=1}^s a_{ij} V_k^j \right\rangle + \left\langle p_{k+1}^i, \frac{q_{k+1} - q_k}{h} - \sum_{j=1}^s b_j V_k^j \right\rangle + b_i \langle \Lambda_k^i, g(Q_k^i) \rangle \right]. \quad (1.3)$$

In the present paper, a second approach based on generalized coordinates is presented. First, the paper introduces the following left-trivialized action:

$$s = \underbrace{\int_a^b \ell(g, \xi) dt}_{\text{left-trivialized Lagrangian}} + \underbrace{\int_a^b \langle \mu, g^{-1} \dot{g} - \xi \rangle dt}_{\text{reconstruction equation}}. \quad (1.4)$$

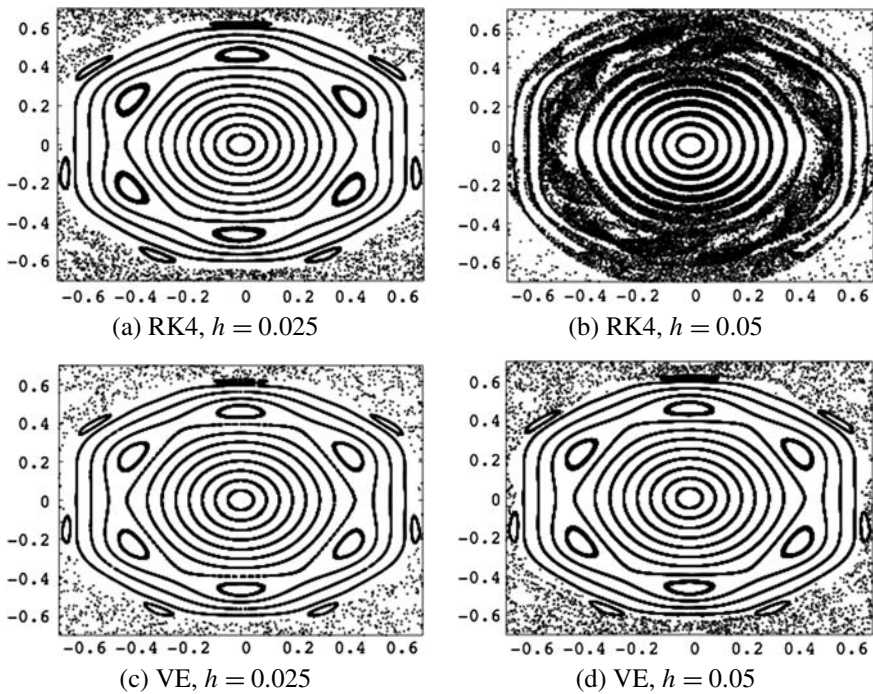
Then an equivalence is established between critical points of  $s$  and  $S$ . If the Lagrangian is left-invariant, it is shown that this principle unifies the system’s Lie–Poisson and Euler–Poincaré descriptions [6, 22]. Since the reconstruction equation is a differential equation on a Lie group, one cannot directly discretize it by an RK method. However, one can discretize it using an  $s$ -stage Runge–Kutta–Munthe-Kaas (RKMK) method [26–29]. The integral of the left-trivialized Lagrangian is approximated using a weighted sum given by the  $b$ -vector in the Butcher tableau of the RKMK scheme. This approach is shown to yield a novel class of variational partitioned Runge–Kutta (VPRK) methods on Lie groups; including generalizations of symplectic Euler and Störmer–Verlet methods on flat spaces.

## 2 Background and Setting

In the next paragraphs, we will give some background material for the reader’s convenience as well as to put the paper into context.

### 2.1 Variational Integrators

Variational integration theory derives integrators for mechanical systems from discrete variational principles. The theory includes discrete analogs of the Lagrangian, Noether’s theorem, the Euler–Lagrange equations, and the Legendre transform. Variational integrators can readily incorporate holonomic constraints (via Lagrange multipliers or the discrete null-space method [18]) and nonconservative effects (via



**Fig. 1** Underwater vehicle dynamics. This figure shows a computation of Poincaré sections using a second-order accurate variational Euler integrator (VE) as compared to fourth-order accurate Runge–Kutta (RK4). Both methods agree with the benchmark at the finer stepsize  $h = 0.025$ . However, at the coarser stepsize  $h = 0.05$ , RK4 corrupts chaotic invariant sets while the lower-order accurate VE method preserves the structure of the benchmark

their virtual work [23]), as well as discrete optimal control (see [19] and references therein). Altogether, this description of mechanics stands as a self-contained theory of mechanics akin to Hamiltonian, Lagrangian, or Newtonian mechanics.

One of the distinguishing features of variational integrators is their ability to compute statistical properties of mechanical systems, such as Poincaré sections, the instantaneous temperature of a system, etc. For example, as a consequence of their variational design, variational integrators are symplectic. A single-step integrator applied to a mechanical system is called *symplectic* if the discrete flow map it defines exactly preserves the canonical symplectic form and is otherwise called standard. Using backward error analysis, one can show that symplectic integrators applied to Hamiltonian systems nearly preserve the energy of the continuous mechanical system for exponentially long periods of time and that the modified equations are also Hamiltonian [8]. Standard integrators often introduce spurious dynamics in long-time simulations, e.g., artificially corrupt chaotic invariant sets as is well illustrated in a computation from [4], namely of a Poincaré section of an underwater vehicle in Fig. 1 using a fourth-order accurate Runge–Kutta (RK4) method and a variational Euler (VE) method designed for rigid-body type systems.

In addition to correctly computing chaotic invariant sets and long-time excellent energy behavior, evidence is mounting that variational integrators correctly compute

other statistical quantities in long-time simulations. For example, in a simulation of a coupled spring-mass lattice, Lew et al. [15] found that variational integrators accurately compute the time-averaged instantaneous temperature (mean kinetic energy over all particles) over long-time intervals, whereas RK4 exhibits a artificial drift in this statistical quantity.

## 2.2 Structure-Preserving Lie Group Integrators

For a mechanical system on a Lie group that possesses the symmetry of that Lie group, in addition to the symplectic structure, the resulting flow preserves a momentum map associated with the Lie group symmetry. In this context, there are several different strategies available to derive structure-preserving Lie group integrators; some of these are discussed here.

One strategy involves generalizations of the classical Newmark algorithms to Lie groups due to [30, 31]. It was not apparent that the proposed Lie–Newmark methods had the necessary structure-preserving properties. In fact, Simo and Wong proposed another set of algorithms which preserve momentum by using the coadjoint action on  $\mathrm{SO}(3)$  to advance the flow. Such integrators will be referred to as *coadjoint-orbit preserving methods*. Later, Austin et al. [1] showed that the midpoint rule member of the Lie–Newmark family with a Cayley reconstruction procedure was, in fact, a coadjoint-orbit preserving method for  $\mathrm{SO}(3)$ . They also numerically demonstrated the method’s good performance crediting it to third-order accuracy in the discrete approximation to the Lie–Poisson structure. In related work, McLachlan and Scovel [24] construct reduced, coadjoint-orbit preserving integrators by reducing  $G$ -equivariant integrators on  $T^*G$  obtained by embedding  $G$  in a linear space using holonomic constraints.

Coadjoint-orbit and energy preserving methods of the Simo and Wong type that further preserve the symplectic structure were developed for  $\mathrm{SO}(3)$  by Lewis and Simo [16, 17]. This was done by defining a one-parameter family of coadjoint and energy-preserving algorithms of the Simo and Wong type in which the free parameter is a functional. The function was specified so that the resulting map defined a transformation which preserves the canonical symplectic form.

Endowing coadjoint-orbit preserving methods with energy-preserving properties was also the subject of Engø and Faltinsen [7]. Specifically, they introduced integrators of the Runge–Kutta–Munthe-Kaas type that preserved coadjoint orbits and energy using the coadjoint action on  $\mathrm{SO}(3)$  and a numerical estimate of the gradient of the Hamiltonian.

Variational integration techniques have been used to derive structure-preserving integrators on Lie groups; see [2, 3, 21, 25, 33]. Moser and Veselov derived a variational integrator for the free rigid body by embedding  $\mathrm{SO}(3)$  in the linear space of  $3 \times 3$  matrices,  $\mathbb{R}^9$ , and using Lagrange multipliers to constrain the matrices to  $\mathrm{SO}(3)$ . This procedure was subsequently generalized to mechanical systems on more general configuration manifolds by the introduction of a constrained discrete Hamilton’s principle in [33].

Another approach is to use *reduction* to derive variational integrators on *reduced spaces*. Marsden et al. [21] developed a discrete analog of EP reduction theory from

which one could design *reduced* numerical algorithms. They did this by constructing a discrete Lagrangian on  $G \times G$  that inherited the  $G$ -symmetry of the continuous Lagrangian, and restricting it to the reduced space  $(G \times G)/G \sim G$ . Using this discrete reduced Lagrangian and a discrete EP (DEP) principle, they derived DEP algorithms on the discrete reduced space. They also considered using generalized coordinates to parametrize this discrete reduced space, specifically the exponential map from the Lie algebra to the Lie group. These techniques were applied to bodies with attitude-dependent potentials, discrete optimal control of rigid bodies, and extended to higher-order accuracy in [13, 14].

Bobenko and Suris [2] considered a more general case where the symmetry group is a subgroup of the Lie group  $G$  in the context of semidirect Euler–Poincaré theory (see [9]). They did this by writing down the discrete Euler–Lagrange equations for this system and left-trivializing them. For the case when the symmetry group is  $G$  itself, one recovers the DEP algorithm as pointed out in [21]. In addition, Bobenko and Suris [3] used this theory to derive an elegant, integrable discretization of the Lagrange top.

### 2.3 Organization of the Paper

In Sect. 3, continuous HP mechanics and its left-trivialization are presented. In particular, it is shown that the HP variational principle and its left-trivialization are equivalent to Hamilton’s and the EP variational principles. Moreover, the symplectic property of the resulting flow map is detailed. In Sect. 4, the discrete analog of the continuous HP theory is developed.

*Part II of this Paper* The second installment of this paper will be devoted to the numerical analysis of HP methods along with numerical experiments on a class of nonreversible mechanical systems on Lie groups as well as the chaotic dynamics of an underwater vehicle. A specific outline of that paper is given in the conclusion section of the present paper.

## 3 HP Mechanics on Lie Groups

This section develops basic mechanics on Lie groups from the Hamilton–Pontryagin perspective.

*The HP Principle* Consider a mechanical system whose configuration space is a Lie group  $G$ . Let its tangent and cotangent bundles be denoted  $TG$  and  $T^*G$ , respectively, and its Lie algebra and dual be given by  $\mathfrak{g}$  and  $\mathfrak{g}^*$ , respectively. In this paragraph, the left-trivialization of the HP principle for a Lagrangian  $L : TG \rightarrow \mathbb{R}$  will be derived.

**Definition 3.1** (Pontryagin bundle) The *Pontryagin bundle* of a manifold  $M$  is defined as  $PM = TM \oplus T^*M$ .

The HP principle unifies the Hamiltonian and Lagrangian descriptions of a mechanical system, as shown in [34, 35]. It states the following critical point condition on  $PG$ ,

$$\delta \int_a^b [L(g, v) + \langle p, \dot{g} - v \rangle] dt = 0,$$

where  $(g(t), v(t), p(t)) \in PG$  are varied arbitrarily and independently with endpoint conditions  $g(a)$  and  $g(b)$  fixed. This builds in the Legendre transformation as well as the Euler–Lagrange equations into one principle.

**Definition 3.2** Following standard conventions, the left action of  $G$  on  $TG$  or  $T^*G$  is denoted by simple concatenation. The left-trivialized Lagrangian  $\ell : G \times \mathfrak{g} \rightarrow \mathbb{R}$  is defined as:

$$\ell(g, \xi) = L(g, g\xi).$$

The HP principle for mechanical systems on Lie groups is equivalent to the left trivialized HP principle:

$$\delta \int_a^b [\ell(g, \xi) + \langle \mu, g^{-1}\dot{g} - \xi \rangle] dt = 0$$

where there are no constraints on the variations; that is, the curves  $\xi(t) \in \mathfrak{g}$ ,  $\mu(t) \in \mathfrak{g}^*$  and  $g(t) \in G$  can be varied arbitrarily and independently with endpoint conditions  $g(a)$  and  $g(b)$  fixed. To see this, we proceed as follows.

**Definition 3.3** (Path spaces) Fixing the interval  $[a, b]$  and  $g_a, g_b \in G$ , define *path space* as

$$\mathcal{C}(PG) = \{(g, v, p) \in C^\infty([a, b], PG) \mid g(a) = g_a, g(b) = g_b\}.$$

Let  $S : \mathcal{C}(PG) \rightarrow \mathbb{R}$  denote the HP action integral and defined as

$$S(g, v, p) = \int_a^b [L(g, v) + \langle p, \dot{g} - v \rangle] dt.$$

Then a simple calculation shows that

$$\begin{aligned} S(g, v, p) &= \int_a^b [L(g, g\xi) + \langle p, gg^{-1}(\dot{g} - v) \rangle] dt \\ &= \int_a^b [\ell(g, \xi) + \langle gp, g^{-1}(\dot{g} - v) \rangle] dt \\ &= \int_a^b [\ell(g, \xi) + \langle \mu, g^{-1}\dot{g} - \xi \rangle] dt \\ &= s(g, \xi, \mu) \end{aligned}$$

where  $s$  is the left-trivialized HP action integral,  $\xi(t) = g(t)^{-1}v(t) \in \mathfrak{g}$ , and  $\mu(t) = g(t)p(t) \in \mathfrak{g}^*$ . From this equality, one can derive the following key theorem.

**Theorem 3.4** *Consider a Lagrangian system on a Lie group  $G$  with Lagrangian  $L : TG \rightarrow \mathbb{R}$ . Let  $\ell : G \times \mathfrak{g} \rightarrow \mathbb{R}$  be its left-trivialization. Then the following are equivalent:*

1. *Hamilton's principle for  $L$  on  $G$*

$$\delta \int_a^b L(g, \dot{g}) dt = 0$$

*holds, for arbitrary variations  $g(t)$  with endpoint conditions  $g(a)$  and  $g(b)$  fixed.*

2. *The following variational principle holds on  $\mathfrak{g}$ :*

$$\delta \int_a^b \ell(g, \xi) dt = 0$$

*using variations of the form*

$$\delta \xi = \dot{\eta} + \text{ad}_{\xi} \eta$$

*where  $\eta(a) = \eta(b) = 0$  and  $\xi = g^{-1}\dot{g}$ ; i.e.,  $\xi = TL_{g^{-1}}\dot{g}$ .*

3. *The HP principle*

$$\delta \int_a^b [L(g, v) + \langle p, \dot{g} - v \rangle] dt = 0$$

*holds, where  $(g(t), v(t), p(t)) \in PG$ , can be varied arbitrarily and independently with endpoint conditions  $g(a)$  and  $g(b)$  fixed.*

4. *The left-trivialized HP principle*

$$\delta \int_a^b [\ell(g, \xi) + \langle \mu, g^{-1}\dot{g} - \xi \rangle] dt = 0$$

*holds, where  $(g(t), \xi(t), \mu(t)) \in G \times \mathfrak{g} \oplus \mathfrak{g}^*$  can be varied arbitrarily and independently with endpoint conditions  $g(a)$  and  $g(b)$  fixed.*

**Remark** If the Lagrangian is left-invariant, i.e.,  $L(g, v) = L(hg, hv)$  for all  $h \in G$ , then the left-trivialized Lagrangian simplifies. In particular, taking  $h = g^{-1}$ ,  $\ell(\xi) = L(g, g\xi) = L(e, \xi)$ , where  $e$  is the identity element of the group. In this case, the left-trivialized HP principle unifies the Euler–Poincaré and Lie–Poisson descriptions on  $\mathfrak{g}$  and  $\mathfrak{g}^*$ , respectively, consistent with the results of [6, 22].

**The HP flow** From the left-trivialized HP principle, the variations of  $s$  with respect to  $\xi$  and  $\mu$  give

$$\text{varying } \mu \text{ gives } \xi = g^{-1}\dot{g} \quad (\text{reconstruction equation}), \quad (3.1)$$

$$\text{varying } \xi \text{ gives } \mu = \frac{\partial \ell}{\partial \xi}(g, \xi) \quad (\text{Legendre transform}). \quad (3.2)$$



Also, setting the variation of  $s$  with respect to  $g$  equal to zero gives

$$\begin{aligned} & \int_a^b \left[ \left\langle \frac{\partial \ell}{\partial g}, \delta g \right\rangle + \langle \mu, \delta(g^{-1} \dot{g}) \rangle \right] dt \\ &= \int_a^b \left[ \left\langle g \frac{\partial \ell}{\partial g}, g^{-1} \delta g \right\rangle + \langle \mu, -g^{-1} \delta g g^{-1} \dot{g} + g^{-1} \delta \dot{g} \rangle \right] dt = 0. \end{aligned} \quad (3.3)$$

Observe that

$$\int_a^b [\langle \mu, \delta(g^{-1} \dot{g}) \rangle] dt = \int_a^b [\langle \mu, -g^{-1} \delta g g^{-1} \dot{g} + g^{-1} \delta \dot{g} \rangle] dt.$$

Let  $\eta = g^{-1} \delta g$ . Using the product rule and (3.1), we see that

$$\frac{d}{dt} \eta = -\xi \eta + g^{-1} \frac{d}{dt} \delta g, \quad \text{which implies} \quad g^{-1} \frac{d}{dt} \delta g = \frac{d}{dt} \eta + \xi \eta.$$

Substituting this relation into (3.3) gives

$$\int_a^b \left[ \left\langle g \frac{\partial \ell}{\partial g}, \eta \right\rangle + \left\langle \mu, \frac{d}{dt} \eta + \text{ad}_\xi \eta \right\rangle \right] dt = 0.$$

Integration by parts and using the boundary conditions on  $g$  yields

$$\int_a^b \left[ \left\langle -\frac{d}{dt} \mu + \text{ad}_\xi^* \mu + g \frac{\partial \ell}{\partial g}, \eta \right\rangle \right] dt = 0.$$

Since the variations are arbitrary, one arrives at

$$\frac{d}{dt} \mu = \text{ad}_\xi^* \mu + g \frac{\partial \ell}{\partial g}. \quad (3.4)$$

In sum, the left-trivialized HP equations are given by:

$$\begin{cases} \frac{d}{dt} g = g \xi, \\ \frac{d}{dt} \mu = \text{ad}_\xi^* \mu + g \frac{\partial \ell}{\partial g}, \\ \mu = \frac{\partial \ell}{\partial \xi}(g, \xi). \end{cases} \quad (3.5)$$

Assuming that the Legendre transform is invertible, (3.5) describes an IVP on the left-trivialized Pontryagin bundle  $G \times \mathfrak{g} \oplus \mathfrak{g}^*$ .

**Definition 3.5** Let  $\mathcal{I}_{\text{HP}} \subset PG$  be defined as,

$$\mathcal{I}_{\text{HP}} := \left\{ (g, v, p) \in PG \mid p = \frac{\partial L}{\partial v}(g, v) \right\}. \quad (3.6)$$

Let  $\mathcal{I}_{\text{hp}} \subset G \times \mathfrak{g} \oplus \mathfrak{g}^*$  denote the left-trivialization of  $\mathcal{I}_{\text{HP}}$  and be defined as,

$$\mathcal{I}_{\text{hp}} := \left\{ (g, \xi, \mu) \in G \times \mathfrak{g} \oplus \mathfrak{g}^* \mid \mu = \frac{\partial \ell}{\partial \xi}(g, \xi) \right\}. \quad (3.7)$$

The *natural projection* is denoted by  $\pi_{\text{HP}} : PG \rightarrow T^*G$  and defined as,

$$\pi_{\text{HP}}(g, v, p) := (g, p), \quad \pi_{\text{HP}}^{-1}(g, p) = (g, v, p), \quad (g, v) = \mathbb{F}L^{-1}(g, p)$$

where  $\mathbb{F}L$  is the Legendre transform.

Given a time-interval  $[a, b]$  and an initial  $(g(a), \xi(a), \mu(a)) \in \mathcal{I}_{\text{hp}}$ , one can solve for  $(g(b), \xi(b), \mu(b)) \in \mathcal{I}_{\text{hp}}$  by eliminating  $\xi$  using the left-trivialized Legendre transform (3.2) and solving the ODEs (3.1) and (3.4) for  $g$  and  $\mu$ . Let this map on  $\mathcal{I}_{\text{hp}}$  be called the *left-trivialized HP flow map*,  $F_{\text{hp}} : \mathcal{I}_{\text{hp}} \rightarrow \mathcal{I}_{\text{hp}}$ .

The flow map  $F_{\text{hp}}$  is equivalent to the HP flow on  $\mathcal{I}_{\text{HP}}$  through left trivialization which defines a diffeomorphism between  $PG$  and  $G \times \mathfrak{g} \oplus \mathfrak{g}^*$ , and hence, between  $\mathcal{I}_{\text{HP}}$  and  $\mathcal{I}_{\text{hp}}$ . Through  $\pi_{\text{HP}}$ , the HP flow is identical to the Hamiltonian flow for the Hamiltonian of this mechanical system on  $T^*G$  obtained via the Legendre transformation. Although  $\pi_{\text{HP}}$  is not a diffeomorphism from  $PG$  to  $T^*G$ , it is a diffeomorphism when its domain is restricted to  $\mathcal{I}_{\text{HP}}$ . Thus, the left-trivialized HP, and the HP and Hamiltonian flows of this mechanical system are all equivalent. This observation makes the subsequent proof of symplecticity seem superfluous, since this structure obviously follows from the standard theory of Hamiltonian systems with symmetry. However, this verification is still important since it serves as a model for the less obvious discrete theory.

It will be helpful to define  $\pi_{\mathcal{I}_{\text{HP}}} = \pi_{\text{HP}}|_{\mathcal{I}_{\text{HP}}}$ . The manifold  $PG$  is a presymplectic manifold with the *HP presymplectic form*,  $\Omega_{\text{HP}} = \pi_{\text{HP}}^* \Omega$ , and the manifold  $\mathcal{I}_{\text{HP}}$  is a symplectic manifold with the *HP symplectic form*,  $\Omega_{\mathcal{I}_{\text{HP}}} = \pi_{\mathcal{I}_{\text{HP}}}^* \Omega$ . Similarly, the manifold  $G \times \mathfrak{g} \oplus \mathfrak{g}^*$  is a presymplectic manifold with the presymplectic form  $\omega_{\text{HP}}$  that is obtained by pulling-back the HP presymplectic form by the left trivialization of  $PG$ ,  $\phi : G \times \mathfrak{g} \oplus \mathfrak{g}^* \rightarrow PG$ , i.e.,  $\omega_{\text{HP}} = \phi^* \Omega_{\text{HP}}$ . However, if the left-trivialization is restricted to  $\mathcal{I}_{\text{hp}}$ ,  $\phi|_{\mathcal{I}_{\text{hp}}} = \phi|_{\mathcal{I}_{\text{hp}}}$ , then  $\mathcal{I}_{\text{hp}}$  is a symplectic manifold with the symplectic form given by  $\omega_{\mathcal{I}_{\text{hp}}} = \phi_{\mathcal{I}_{\text{hp}}}^* \Omega_{\mathcal{I}_{\text{HP}}}$ .

**Symplecticity** The symplectic structure of left-trivialized HP flows is obvious from the standard theory of Hamiltonian systems with symmetry, but reviewing the proof will help since it parallels the discrete case.

Consider the restriction of the left-trivialized HP action integral to solutions of (3.5):  $\hat{s}$ . Since the space of solutions of (3.5) can be identified with  $\mathcal{I}_{\text{hp}}$ ,  $\hat{s} : \mathcal{I}_{\text{hp}} \rightarrow \mathbb{R}$ . The differential of  $\hat{s}$  can be written as,

$$\begin{aligned} d\hat{s} \cdot (\delta g(a), \delta \xi(a), \delta \mu(a)) &= \int_a^b \left[ (g^{-1} \dot{g} - \xi) \cdot \delta \mu + \left( \mu - \frac{\partial \ell}{\partial \xi} \right) \cdot \delta \xi \right] dt \\ &\quad + \int_a^b \left[ \left( -\frac{d}{dt} \mu + \text{ad}_{\xi}^* \mu + g \frac{\partial \ell}{\partial g} \right) \cdot g^{-1} \delta g \right] dt + \langle \mu, g^{-1} \delta g \rangle \Big|_a^b \\ &= \langle \mu, g^{-1} \delta g \rangle \Big|_a^b. \end{aligned}$$

Since  $\mathbf{d}^2\hat{s} = 0$ , observe that

$$\mathbf{d}^2\hat{s} = (F_{\text{hp}})^*\omega_{\mathcal{I}_{\text{hp}}} - \omega_{\mathcal{I}_{\text{hp}}} = 0.$$

And hence, as a map on  $\mathcal{I}_{\text{hp}}$ ,  $F_{\text{hp}}$  is symplectic.

**Proposition 3.6** *Left-trivialized HP flows preserve the symplectic two-form  $\omega_{\mathcal{I}_{\text{hp}}}$ .*

## 4 Lie Group VPRK Integrators

The purpose of this section is to use the general HP methodology to derive a variety of integrators of variational partitioned Runge–Kutta (VPRK) type for Lie groups. Recall the left-trivialized HP action integral is given by:

$$s = \underbrace{\int_a^b \ell(g, \xi) dt}_{\text{left-trivialized Lagrangian}} + \underbrace{\int_a^b \langle \mu, g^{-1}\dot{g} - \xi \rangle dt}_{\text{reconstruction equation}}. \quad (4.1)$$

To approximate  $s$  by an action sum the reconstruction equation is discretized using a Runge–Kutta–Munthe-Kaas (RKMK) approximant and in terms of this discretization an approximant to the integral of the left-trivialized Lagrangian is introduced. The section shows this approach leads naturally to VPRK Integrators on Lie groups. These integrators include an attractive generalization of the Störmer–Verlet method to Lie groups, symplectic Euler methods on Lie groups, and Euler–Poincaré integrators.

**Canonical Coordinates of the First Kind** To setup the discrete HP principle, we introduce a map  $\tau : \mathfrak{g} \rightarrow G$ . Let  $e \in G$  be the identity element of the group. The map  $\tau$  is assumed to be a local diffeomorphism mapping a neighborhood of zero on  $\mathfrak{g}$  to one of  $e$  on  $G$  with  $\tau(0) = e$ , assumed to be analytic in this neighborhood, and assumed to satisfy  $\tau(\xi) \cdot \tau(-\xi) = e$ . Thereby,  $\tau$  provides a local chart on the Lie group. By left translation, this map can be used to construct an atlas on  $G$ . An example of a  $\tau$  is the exponential map on  $G$ , but there are other interesting examples as well, as we shall see shortly.

**Definition 4.1** The local coordinates associated with the map  $\tau$  are called *canonical coordinates of the first kind* or just *canonical coordinates*.

For an exposition of canonical coordinates of the first and second kind, and their applications, the reader is referred to [10]. In what follows, we will prove some properties of these coordinates that will be needed shortly.

**Derivative of  $\tau$  and Its Inverse** To derive the integrator that comes from a discrete left-trivialized HP principle, we will need to differentiate  $\tau^{-1}$ . The *right trivialized tangent* of  $\tau$  and its inverse will play an important role in writing this derivative in an efficient way. The following is taken from Definition 2.19 in [10].

**Definition 4.2** Given a local diffeomorphism  $\tau : \mathfrak{g} \rightarrow G$ , we define its *right trivialized tangent* to be the function  $d\tau : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  which satisfies,

$$D\tau(\xi) \cdot \delta = TR_{\tau(\xi)} d\tau_{\xi}(\delta).$$

The function  $d\tau$  is linear in its second argument.

Figure 2 illustrates the geometry behind this definition. From this definition, the following lemma is deduced.

**Lemma 4.3** *The following identity holds:*

$$d\tau_{\xi}(\delta) = \text{Ad}_{\tau(\xi)} d\tau_{-\xi}(\delta).$$

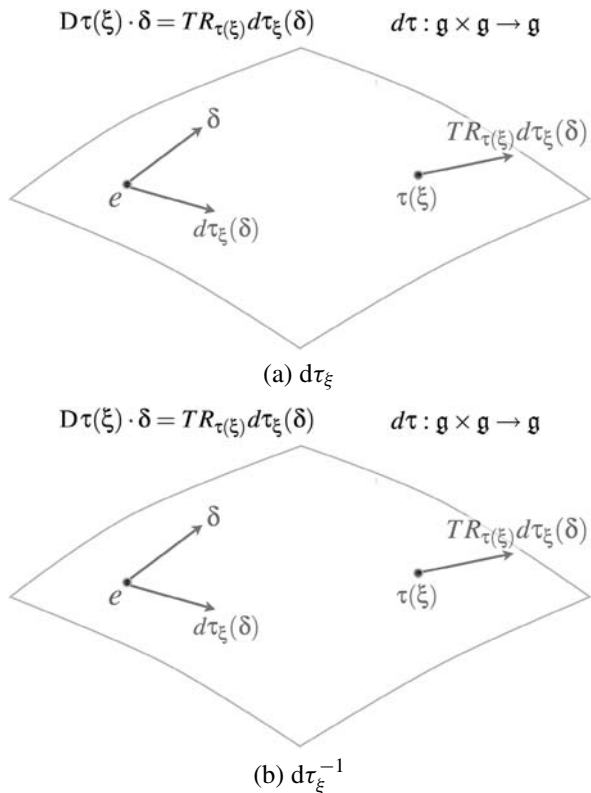
*Proof* Differentiation of  $\tau(\xi) \cdot \tau(-\xi) = e$  gives

$$D\tau(-\xi) \cdot \delta = -TL_{\tau(-\xi)} TR_{\tau(-\xi)} (D\tau(\xi) \cdot \delta).$$

While the chain rule yields

$$D\tau(-\xi) \cdot \delta = -TR_{\tau(-\xi)} d\tau_{-\xi}(\delta).$$

**Fig. 2** Derivatives of  $\tau$  and  $\tau^{-1}$ . Definition 4.2 splits the differential of  $\tau$  into a map on the Lie algebra (the right trivialized tangent of  $\tau$ ) and right multiplication to the tangent space at  $\tau(\xi)$ , while Definition 4.4 splits the differential of  $\tau^{-1}$  into right multiplication to the Lie algebra and a map on the Lie algebra (the right trivialized tangent of  $\tau^{-1}$ )



Combining these two identities and using the definition above,

$$-TR_{\tau(-\xi)} d\tau_{-\xi}(\delta) = -TL_{\tau(-\xi)} TR_{\tau(-\xi)} TR_{\tau(\xi)} d\tau_{\xi}(\delta).$$

Simplifying this expression gives

$$TL_{\tau(\xi)} d\tau_{-\xi}(\delta) = TR_{\tau(\xi)} d\tau_{\xi}(\delta),$$

which proves the identity.  $\square$

We will also need a simple expression for the differential of  $\tau^{-1}$ .

**Definition 4.4** The *inverse right trivialized tangent* of  $\tau$  is the function  $d\tau^{-1} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  which satisfies for  $g = \tau(\xi)$ ,

$$D\tau^{-1}(g) \cdot \delta = d\tau_{\xi}^{-1}(TR_{\tau(-\xi)}\delta), \quad d\tau_{\xi}^{-1}(d\tau_{\xi}(\delta)) = \delta.$$

The function  $d\tau^{-1}$  is always linear in its second argument.

Figure 2 illustrates the geometry behind this definition.

The following lemma follows from this definition and Lemma 4.3 above.

**Lemma 4.5** *The following identity holds:*

$$d\tau_{\xi}^{-1}(\delta) = d\tau_{-\xi}^{-1}(\text{Ad}_{\tau(-\xi)}\delta).$$

*Proof* This follows directly from Lemma 4.3. Let  $\delta \rightarrow d\tau_{\xi}^{-1}(\delta)$  in that identity to obtain

$$\delta = \text{Ad}_{\tau(\xi)} d\tau_{-\xi}(d\tau_{\xi}^{-1}(\delta)).$$

And now solve this equation for  $d\tau_{\xi}^{-1}(\delta)$ ,

$$d\tau_{\xi}^{-1}(\delta) = d\tau_{-\xi}^{-1}(\text{Ad}_{\tau(-\xi)}\delta). \quad \square$$

**RKMK Discretization of Reconstruction Equation** Let  $[a, b]$  and  $N$  be given, let  $h = (b - a)/N$  be a fixed integration time step and  $t_k = hk + a$  for  $k = 0, \dots, N$ . A good candidate for discretizing the reconstruction equation is given by a generalization of  $s$ -stage Runge–Kutta methods to differential equations on Lie groups, namely Runge–Kutta–Munthe-Kaas (RKMK) methods introduced in the following series of papers: [26–29]. The idea behind these papers is to use canonical coordinates on the Lie group to transform the differential equation on  $TG$ , e.g., given by,

$$\dot{g}(t) = g(t)f(t, g(t)), \quad g(a) = g_a, \quad g(t) \in G, \quad f(t, g(t)) \in \mathfrak{g}, \quad t \in [a, b], \quad (4.2)$$

to a differential equation on  $\mathfrak{g}$ . Specifically, substitute the following parametrization  $g(t) = g_0\tau(\Theta(t))$  into (4.2) to obtain,

$$\dot{g} = TL_{g_0} TR_{\tau(\Theta)} d\tau_{\Theta} \dot{\Theta} = TL_{g_0} TL_{\tau(\Theta)} f.$$

Using Lemma 4.3, this equation can be rewritten as

$$TL_{\tau(-\Theta)} TR_{\tau(\Theta)} d\tau_{\Theta} \dot{\Theta} = \text{Ad}_{\tau(-\Theta)} d\tau_{\Theta} \dot{\Theta} = d\tau_{-\Theta} \dot{\Theta} = f.$$

Solving for  $\dot{\Theta}$  gives

$$\dot{\Theta} = d\tau_{-\Theta}^{-1} f, \quad \Theta(a) = 0, \quad \Theta(t) \in \mathfrak{g}, \quad t \in [a, b]. \quad (4.3)$$

As described in the following definition, a RKMK method is obtained by applying an  $s$ -stage RK method to (4.3).

**Definition 4.6** Consider the first-order differential equation  $\dot{g}(t) = f(t, g(t))$  for the curve  $(g(t), f(t, g(t))) \in TG$ . Given coefficients  $b_i, a_{ij} \in \mathbb{R}$  ( $i, j = 1, \dots, s$ ) and set  $c_i = \sum_{j=1}^s a_{ij}$ . An  $s$ -stage Runge–Kutta–Munthe-Kaas (RKMK) approximant to (4.2) is given by

$$G_k^i = g_k \tau(h\Theta_k^i), \quad (4.4)$$

$$\Theta_k^i = h \sum_{j=1}^s a_{ij} d\tau_{-h\Theta_k^j}^{-1} f(t_k + c_j h, G_k^j), \quad i = 1, \dots, s, \quad (4.5)$$

$$g_{k+1} = g_k \tau \left( h \sum_{j=1}^s b_j d\tau_{-h\Theta_k^j}^{-1} f(t_k + c_j h, G_k^j) \right). \quad (4.6)$$

If  $a_{ij} = 0$  for  $i \leq j$  the RKMK method is called explicit, and implicit otherwise. The variables  $g_k$  and  $G_k^i$  are called external and internal stage configurations, respectively.

It follows that for given  $\tau$ , an  $s$ -stage RKMK method is determined by its  $a$ -matrix and  $b$ -vector which are typically displayed using the so-called Butcher tableau:

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array}$$

Suppose that  $\xi(t)$ ,  $t \in [a, b]$ , is a given vector field on  $\mathfrak{g}$ . From the definition above, it is clear that an  $s$ -stage RKMK method applied to  $\dot{g} = g\xi$  can be written as:

$$\begin{cases} \frac{\tau^{-1}(g_k^{-1} G_k^i)}{h} = \sum_{j=1}^s a_{ij} d\tau_{-h\Theta_k^j}^{-1} \mathcal{E}_k^j = \Theta_k^i, & i = 1, \dots, s, \\ \frac{\tau^{-1}(g_k^{-1} g_{k+1})}{h} = \sum_{j=1}^s b_j d\tau_{-h\Theta_k^j}^{-1} \mathcal{E}_k^j \end{cases} \quad (4.7)$$

where  $\mathcal{E}_k^i = \xi(t_k + c_i h)$ . In practice, one often truncates the series expansion associated to the map  $d\tau_{-h\Theta_k^j}^{-1}$ . The following theorem guides how to do this without degrading the order of accuracy of the RKMK method [8].

**Theorem 4.7** *Given an  $r$ th-order approximant to the exact exponential map:  $\tau : \mathfrak{g} \rightarrow G$ . Suppose the underlying RK method in (4.7) is of order  $p$  with  $r \geq p$ . If the truncation index of  $d\tau_{\Theta}^{-1}$  in (4.7) satisfies  $q \geq p - 2$ , then the RKMK method is of order  $p$  as well.*

**VPRK Integrators on Lie Groups** The discrete HP principle states that the discrete path the discrete system takes is one that extremizes an action sum that will be introduced shortly. To discretize the left-trivialized HP action integral (cf. (4.1)), the reconstruction equation is discretized using (4.7). This discretized reconstruction equation is (analogous to the continuous HP principle) treated as a constraint in the action sum. The integral of the left-trivialized Lagrangian is approximated by the following quadrature whose weights are determined by the  $b$ -vector in the RKMK method:

$$\int_{t_k}^{t_k+h} \ell(g, \xi) dt \approx \sum_{i=1}^s h b_i \ell(G_k^i, \Xi_k^i). \quad (4.8)$$

The truncation index of  $d\tau_{\Theta}^{-1}$  in (4.7) is chosen to be  $q = 0$ . By Theorem 4.7, the resulting RKMK method is at most second-order accurate. In the sequel to this paper, we will show if an RKMK method is of order  $p$ , then its associated VPRK scheme is of order  $p$  as well.

**Definition 4.8** Let  $[a, b]$  and  $N$  be given, let  $h = (b - a)/N$  be a fixed integration time step and  $t_k = hk + a$  for  $k = 0, \dots, N$ . Given  $g_a, g_b \in G$  and an  $s$ -stage RKMK method with  $b_j \neq 0$  for  $j = 1, \dots, s$ , define its associated *discrete path space* as

$$\mathcal{C}_d = \left\{ (g, \hat{\mu}, \{\Theta^i, \Xi^i, \mu^i\}_{i=1}^s) : \{t_k\}_{k=0}^N \rightarrow (G \times \mathfrak{g}^*) \times (\mathfrak{g} \times \mathfrak{g} \times \mathfrak{g}^*)^s \mid \right. \\ \left. g(t_0) = g_a, g(t_N) = g_b \right\},$$

and its associated *action sum*  $s_d : \mathcal{C}_d \rightarrow \mathbb{R}$  as

$$s_d = \sum_{k=0}^{N-1} \sum_{i=1}^s h \left[ b_i \ell(G_k^i, \Xi_k^i) + \left\langle \mu_k^i, \frac{\tau^{-1}(g_k^{-1} G_k^i)}{h} - \sum_{j=1}^s a_{ij} \Xi_k^j \right\rangle \right. \\ \left. + \left\langle \hat{\mu}_{k+1}, \frac{\tau^{-1}(g_k^{-1} g_{k+1})}{h} - \sum_{j=1}^s b_j \Xi_k^j \right\rangle \right]. \quad (4.9)$$

Let  $(g_k, \hat{\mu}_k, \{\Theta_k^i, \Xi_k^i, \Psi_k^i\}) = (g, \hat{\mu}, \{\Theta^i, \Xi^i, \mu^i\}_{i=1}^s)_d(t_k)$  for  $i = 1, \dots, s$  and  $k = 0, \dots, N$ . Observe that  $s_d$  is an approximation of the left-trivialized HP action integral by numerical quadrature. The definition of  $\tau$  as a map from  $\mathfrak{g}$  to  $G$  ensures that the pairings  $s_d$  are well defined. The discrete left trivialized HP principle states that

$$\delta s_d = 0$$

for arbitrary and independent variations of the external stage variables  $(g_k, \hat{\mu}_k) \in G \times \mathfrak{g}^*$  and the internal stage variables  $(\Theta_k^i, \Xi_k^i, \Psi_k^i) \in \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g}^*$  for  $i = 1, \dots, s$

and  $k = 0, \dots, N$  subject to fixed endpoint conditions on  $\{g_k\}_{k=0}^N$ . The following theorem states this discrete principle and its relation to the HP symplectic form  $\omega_{\mathcal{I}_{\text{hp}}}$ .

**Theorem 4.9** *Let  $\ell : G \times \mathfrak{g} \rightarrow \mathbb{R}$  be a smooth, left-trivialized Lagrangian. A discrete curve  $c_d \in \mathcal{C}_d$  is a critical point of the function  $s_d : \mathcal{C}_d \rightarrow \mathbb{R}$  if and only if it shadows a solution to the following VPRK scheme:*

$$\frac{\tau^{-1}(g_k^{-1}G_k^i)}{h} = \sum_{j=1}^s a_{ij} \Xi_k^j = \Theta_k^i, \quad (4.10)$$

$$\frac{\tau^{-1}(g_k^{-1}g_{k+1})}{h} = \sum_{j=1}^s b_j \Xi_k^j = \xi_{k+1}, \quad (4.11)$$

$$\begin{aligned} & (\mathrm{d}\tau_{h\xi_{k+1}}^{-1})^* M_k^i \\ &= \mu_k + h \sum_{j=1}^s \left( b_j (\mathrm{d}\tau_{h\Theta_k^j}^{-1})^* - \frac{b_j a_{ji}}{b_i} (\mathrm{d}\tau_{h\xi_{k+1}}^{-1})^* \right) (\mathrm{d}\tau_{-h\Theta_k^j})^* G_k^j \frac{\partial \ell}{\partial g}(G_k^j, \Xi_k^j), \end{aligned} \quad (4.12)$$

$$\mu_{k+1} = \mu_k + h \sum_{j=1}^s b_j (\mathrm{d}\tau_{h\Theta_k^j}^{-1})^* (\mathrm{d}\tau_{-h\Theta_k^j})^* G_k^j \frac{\partial \ell}{\partial g}(G_k^j, \Xi_k^j), \quad (4.13)$$

$$M_k^i = \frac{\partial \ell}{\partial \xi}(G_k^i, \Xi_k^i). \quad (4.14)$$

for  $i = 1, \dots, s$  and  $k = 0, \dots, N-1$ . Moreover, assuming the discrete flow,  $F_{\text{hp}}^d : \mathcal{I}_{\text{hp}} \rightarrow \mathcal{I}_{\text{hp}}$ , defined by this VPRK scheme exists; then

$$(F_{\text{hp}}^d)^* \omega_{\mathcal{I}_{\text{hp}}} = \omega_{\mathcal{I}_{\text{hp}}}.$$

That is, the discrete flow is symplectic.

**Remark** Consider the case when  $G$  is the Euclidean space  $\mathbb{R}^n$  with ordinary vector addition as the group operation. Set  $\tau(\xi) = \exp(\xi)$ , so that  $\mathrm{d}\tau_\xi \delta = \mathrm{d}\tau_\xi^{-1} \delta = \delta$ . In this case, (4.10–4.14) become the well-known symplectic partitioned Runge–Kutta methods [8, 23].

**Proof of Theorem 4.9** Set  $\eta_k = g_k^{-1} \delta g_k$  and  $H_k^i = G_k^{i-1} \delta G_k^i$ . The differential of  $s_d(c_d)$  in the direction  $z = (\{\delta g_k, \delta \hat{\mu}_k\}, \{\delta G_k^i, \delta \Xi_k^i, \delta \mu_k^i\}_{i=1}^s)$  is given by

$$\begin{aligned} \mathbf{d}s_d \cdot z &= \sum_{k=0}^{N-1} \sum_{i=1}^s h b_i G_k^i \frac{\partial \ell}{\partial g}(G_k^i, \Xi_k^i) \cdot H_k^i + h b_i \frac{\partial \ell}{\partial \xi}(G_k^i, \Xi_k^i) \cdot \delta \Xi_k^i \\ &\quad + h \left\langle \delta \mu_k^i, \frac{\tau^{-1}(g_k^{-1}G_k^i)}{h} - \sum_{j=1}^s a_{ij} \Xi_k^j \right\rangle \end{aligned}$$



$$\begin{aligned}
 & + h \left\langle \delta \hat{\mu}_{k+1}, \frac{\tau^{-1}(g_k^{-1} g_{k+1})}{h} - \sum_{j=1}^s b_j \mathcal{E}_k^j \right\rangle \\
 & + h \left\langle \mu_k^i, -\frac{d\tau_{h\Theta_k^i}^{-1} \eta_k}{h} + \frac{d\tau_{-h\Theta_k^i}^{-1} H_k^i}{h} - \sum_{j=1}^s a_{ij} \delta \mathcal{E}_k^j \right\rangle \\
 & + h \left\langle \hat{\mu}_{k+1}, -\frac{d\tau_{h\xi_{k+1}}^{-1} \eta_k}{h} + \frac{d\tau_{-h\xi_{k+1}}^{-1} \eta_{k+1}}{h} - \sum_{j=1}^s b_j \delta \mathcal{E}_k^j \right\rangle.
 \end{aligned}$$

Collecting terms with the same variations and summation by parts using the boundary conditions  $\delta g_0 = \delta g_N = 0$  gives

$$\begin{aligned}
 \mathbf{d}s_d \cdot z &= \sum_{k=1}^{N-1} \sum_{i=1}^s h \left\langle \delta \mu_k^i, \frac{\tau^{-1}(g_k^{-1} G_k^i)}{h} - \sum_{j=1}^s a_{ij} \mathcal{E}_k^j \right\rangle \\
 & + h \left\langle \delta \hat{\mu}_{k+1}, \frac{\tau^{-1}(g_k^{-1} g_{k+1})}{h} - \sum_{j=1}^s b_j \mathcal{E}_k^j \right\rangle \\
 & + h \left\langle b_i \frac{\partial \ell}{\partial \xi}(G_k^i, \mathcal{E}_k^i) - \sum_{j=1}^s a_{ji} \mu_k^j - b_i \hat{\mu}_{k+1}, \delta \mathcal{E}_k^i \right\rangle \\
 & + \left\langle (d\tau_{-h\Theta_k^i}^{-1})^* \mu_k^i + h b_i G_k^i \frac{\partial \ell}{\partial g}(G_k^i, \mathcal{E}_k^i), H_k^i \right\rangle \\
 & + \left\langle -(d\tau_{h\xi_{k+1}}^{-1})^* \hat{\mu}_{k+1} + (d\tau_{-h\xi_k}^{-1})^* \hat{\mu}_k - \sum_{j=1}^s (d\tau_{h\Theta_k^j}^{-1})^* \mu_k^j, \eta_k \right\rangle.
 \end{aligned}$$

Since  $\mathbf{d}s_d(c_d) = 0$  if and only if  $\mathbf{d}s_d \cdot z = 0$  for all  $z \in T_{c_d} \mathcal{C}_d$ , one arrives at the desired equations by elimination of  $\mu_k^i$  and introduction of the internal stage momenta  $M_k^i = \partial \ell / \partial \xi(G_k^i, \mathcal{E}_k^i)$  for  $i = 1, \dots, s$  and the external stage momenta  $\mu_k = (d\tau_{-h\xi_k}^{-1})^* \hat{\mu}_k$  for  $k = 0, \dots, N-1$ . For  $h$  sufficiently small, this change of variables is invertible by the implicit function theorem. Conversely, if  $c_d$  satisfies (4.10–4.14) under the aforementioned change of variables, then  $\mathbf{d}s_d(c_d) = 0$ .

Consider the subset of  $\mathcal{C}_d$  consisting of solutions of (4.10–4.14) under the aforementioned change of variables. Let  $\hat{s}_d$  denote the restriction of  $s_d$  to this space. Since each of these solutions is determined by a point in  $\mathcal{I}_{hp}$ , one can identify this space with  $\mathcal{I}_{hp}$ , and hence,  $\hat{s}_d : \mathcal{I}_{hp} \rightarrow \mathbb{R}$ . Since  $\hat{s}_d$  is restricted to solution space,

$$\mathbf{d}\hat{s}_d \cdot \hat{z} = \langle \mu_N, g_N^{-1} \delta g_N \rangle - \langle \mu_0, g_0^{-1} \delta g_0 \rangle.$$

Preservation of  $\omega_{\mathcal{I}_{hp}}$  follows from  $\mathbf{d}^2 \hat{s}_d = 0$ .  $\square$

The external and internal stages of (4.10–4.14) define update schemes on  $G \times \mathfrak{g}^*$  and  $(\mathfrak{g} \times \mathfrak{g} \times \mathfrak{g}^*)^s$ , respectively.

*Störmer–Verlet Integrators on Lie Groups* A generalization of the Störmer–Verlet method to Lie groups is given by evaluating (4.10)–(4.14) at the following two-stage RK tableau (implicit trapezoidal rule),

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1/2 & 1/2 \\ \hline & 1/2 & 1/2 \end{array}$$

Suppose for simplicity the left-trivialized Lagrangian is separable, i.e.,  $\ell(g, \xi) = T(\xi) - U(g)$ . Given  $h$  and  $(g_k, \mu_k)$ , the method determines  $(g_{k+1}, \mu_{k+1})$  by solving the following system of equations:

$$M_k^{1/2} = \frac{\partial T}{\partial \xi}(\xi_k^{1/2}), \quad (4.15)$$

$$(\mathrm{d}\tau_{h\xi_k^{1/2}}^{-1})^* M_k^{1/2} = \mu_k - \frac{h}{2} g_k \frac{\partial U}{\partial g}(g_k), \quad (4.16)$$

$$g_{k+1} = g_k \tau(h\xi_k^{1/2}), \quad (4.17)$$

$$\mu_{k+1} = (\mathrm{d}\tau_{-h\xi_k^{1/2}}^{-1})^* M_k^{1/2} - \frac{h}{2} g_{k+1} \frac{\partial U}{\partial g}(g_{k+1}). \quad (4.18)$$

In particular, one uses the following algorithm:

*Step 1.* Solve for  $M_k^{1/2}$  and  $\xi_k^{1/2}$  using (4.15) and (4.16). This update is implicit.

*Step 2.* Update  $g_{k+1}$  using (4.17). This update is explicit.

*Step 3.* Solve for  $\mu_{k+1}$  using (4.18). This update is explicit.

*Remark* Recall, the Störmer–Verlet integrator for separable Lagrangian systems on flat spaces is explicit, symmetric, and second-order accurate. Although this generalization of Störmer–Verlet to Lie groups is no longer explicit because a key point is that the nonlinearity in step 1 does not involve the potential force field; as is easy to confirm the integrator is also symmetric and second-order accurate.

*Variational Euler on Lie Groups* Variational Euler schemes can be derived from the following VPRK action sums:

$$\begin{aligned} s_d^f &= \sum_{k=0}^{N-1} h \left[ \ell(g_{k+1}, \xi_k) + \langle \mu_k, \tau^{-1}(g_k^{-1} g_{k+1}) / h - \xi_k \rangle \right], \\ s_d^b &= \sum_{k=0}^{N-1} h \left[ \ell(g_k, \xi_{k+1}) + \langle \mu_{k+1}, \tau^{-1}(g_k^{-1} g_{k+1}) / h - \xi_{k+1} \rangle \right]. \end{aligned}$$

Given  $h$  and  $(g_k, \mu_k)$ , the forward variational Euler method comes from stationarity of  $s_d^f$  and determines  $(g_{k+1}, \mu_{k+1})$  by solving the following system of equations:

$$\begin{cases} g_{k+1} = g_k \tau(h\xi_k), \\ (d\tau_{h\xi_{k+1}}^{-1})^* \mu_{k+1} = (d\tau_{-h\xi_k}^{-1})^* \mu_k + hg_{k+1} \frac{\partial \ell}{\partial g}(g_{k+1}, \xi_k), \\ \mu_k = \frac{\partial \ell}{\partial \xi}(g_{k+1}, \xi_k). \end{cases} \quad (4.19)$$

The backward variational Euler method comes from stationarity of  $s_d^b$  and determines  $(g_{k+1}, \mu_{k+1})$  by solving the following system of equations:

$$\begin{cases} g_{k+1} = g_k \tau(h\xi_{k+1}), \\ (d\tau_{h\xi_{k+1}}^{-1})^* \mu_{k+1} = (d\tau_{-h\xi_k}^{-1})^* \mu_k + hg_k \frac{\partial \ell}{\partial g}(g_k, \xi_{k+1}), \\ \mu_{k+1} = \frac{\partial \ell}{\partial \xi}(g_k, \xi_{k+1}). \end{cases} \quad (4.20)$$

**Euler–Poincaré Integrators** In the case when the Lagrangian is  $G$ -left-invariant, the body angular momentum updates in the above variational Euler schemes are given by:

$$(d\tau_{h\xi_{k+1}}^{-1})^* \mu_{k+1} = (d\tau_{-h\xi_k}^{-1})^* \mu_k. \quad (4.21)$$

**Examples 1** We now exhibit some Euler–Poincaré integrators by evaluating (4.21) at some specific maps  $\tau$ .

(a) *Matrix exponential.* Suppose

$$\tau = \exp(\xi), \quad \tau : \mathfrak{g} \rightarrow G,$$

which is a local diffeomorphism.

Using standard convention the right trivialized tangent of the exponential map and its inverse are denoted by  $\text{dexp} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  and  $\text{dexp}^{-1} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ , and are explicitly given by

$$\text{dexp}(x)y = \sum_{j=0}^{\infty} \frac{1}{(j+1)!} \text{ad}_x^j y, \quad \text{dexp}^{-1}(x)y = \sum_{j=0}^{\infty} \frac{B_j}{j!} \text{ad}_x^j y \quad (4.22)$$

where  $B_j$  are the *Bernoulli numbers*; see Sect. 3.4 of [8] for a detailed exposition and derivation.

Hence, (4.21) takes the form

$$(\text{dexp}^{-1}(h\xi_k))^* \mu_k = (\text{dexp}^{-1}(-h\xi_{k-1}))^* \mu_{k-1}. \quad (4.23)$$

(b) *Padé (1, 1) approximant.* Suppose

$$\tau(\xi) = \text{cay}(\xi) = (e - \xi/2)^{-1}(e + \xi/2), \quad (4.24)$$

which is the Padé (1,1) approximant to the matrix exponential and better known as the *Cayley transform*. The Cayley transform maps to the group for quadratic Lie groups ( $SO(n)$ , the symplectic group  $Sp(2n)$ , the Lorentz group  $SO(3, 1)$ ) and the special Euclidean group  $SE(3)$ .

The right-trivialized tangent of the Cayley transform and its inverse are written below

$$\text{dcay}(x)y = (e - x/2)^{-1}y(e + x/2)^{-1}, \quad \text{dcay}^{-1}(x)y = (e - x/2)y(e + x/2). \quad (4.25)$$

For a derivation and exposition, the reader is referred to Sect. 4.8.3 of [8]. Using these expressions (4.21) can be written as

$$\mu_k = \mu_{k-1} + \frac{h}{2} \text{ad}_{\xi_k}^* \mu_k + \frac{h}{2} \text{ad}_{\xi_{k-1}}^* \mu_{k-1} + \frac{h^2}{4} (\xi_k^* \mu_k \xi_k^* - \xi_{k-1}^* \mu_{k-1} \xi_{k-1}^*). \quad (4.26)$$

- (c) *Padé (1, 0) or (0, 1) approximant*. Rather than use the exact matrix exponential, one can use a Padé approximant, e.g., the Padé (1, 0) approximant

$$\exp(\xi) \approx e + \xi$$

or Padé (0, 1) approximant

$$\exp(\xi) \approx (e - \xi)^{-1}.$$

However, since a Padé approximant is not guaranteed to lie on the group, one needs to use a projector from  $GL(n)$  to  $G$ . In what follows,  $G = SO(n)$  will be considered where a natural choice of projector is given by skew symmetrization.

Suppose

$$\tau^{-1}(g) = \text{skew}(g) = \frac{g - g^*}{2},$$

which comes from a first order approximant to the matrix exponential. This map is a local diffeomorphism from a neighborhood of  $e$  to a neighborhood of 0 and its differential is the identity. Its right-trivialized tangent can be computed from its derivative:

$$D \text{skew}(g) \cdot \delta = \frac{\delta - \delta^*}{2} = \frac{(\delta g^{-1}g) - (\delta g^{-1}g)^*}{2}.$$

By definition of the right-trivialized tangent of  $\tau^{-1}$ , it then follows that

$$d \text{skew}(x)(y) = \frac{y\tau(x) - (y\tau(x))^*}{2}. \quad (4.27)$$

Cardoso and Leite [5] obtained the following proposition that explicitly determines  $\tau(\xi)$ . Moreover, they give necessary and sufficient conditions for its existence.

**Proposition 4.10** *Given  $\xi \in \mathfrak{so}(n)$ , a special orthogonal solution to the equation*

$$\xi = \frac{\tau(\xi) - \tau(\xi)^*}{2}$$

*can be written as*

$$\tau(\xi) = \xi + (\xi^2 + e)^{1/2},$$

*where  $(\xi^2 + e)^{1/2}$  is a symmetric square root.*

Hence, (4.21) can be written as

$$\begin{aligned} & \frac{\mu_k(h^2\xi_k^2 + e)^{1/2} + (h^2\xi_k^2 + e)^{1/2}\mu_k}{2} \\ &= \frac{\mu_{k-1}(h^2\xi_{k-1}^2 + e)^{1/2} + (h^2\xi_{k-1}^2 + e)^{1/2}\mu_{k-1}}{2} + \frac{h}{2} \text{ad}_{\xi_k}^* \mu_k + \frac{h}{2} \text{ad}_{\xi_{k-1}}^* \mu_{k-1}. \end{aligned} \quad (4.28)$$

## 5 Conclusion

In this paper, a left-trivialized Hamilton–Pontryagin principle is derived for mechanical systems on a Lie group  $G$ . If the Lagrangian is left-invariant with respect to the action of  $G$ , it is shown that this left-trivialized HP principle unifies the Euler–Poincaré and Lie–Poisson descriptions. In addition to its utility for implicit Lagrangian systems, the paper shows that this principle provides a practical way to design discrete Lagrangians. In particular, the paper explains how one can discretize the kinematic constraint using a Runge–Kutta–Munthe-Kaas (RKMK) method. The paper shows that this leads to a novel generalization of variational partitioned Runge–Kutta methods from flat spaces to Lie groups. In particular, one can generalize variational (or symplectic) Euler and Störmer–Verlet methods to Lie groups in this fashion. These methods inherit many of their attractive properties on flat spaces: efficiency, order of accuracy, symplecticity, symmetry, etc.

Part II of this paper will develop a basic numerical analysis of these methods and report on numerical experiments on a class of nonreversible mechanical systems on Lie groups (moving rigid body systems) and chaotic dynamics of an underwater vehicle. To be specific, the paper will:

- prove order of accuracy of the VPRK integrators presented in this paper by invoking the variational proof of order of accuracy [23];
- explain the numerics behind the Poincaré sections provided in Fig. 1;
- demonstrate the superiority of these VPRK integrators compared to symmetric rigid body integrators when applied to a nonreversible system such as a rigid body on a turntable.

## References

1. M.A. Austin, P.S. Krishnaprasad, L. Wang, Almost-Poisson integration of rigid body systems, *J. Comput. Phys.* **107**, 105–117 (1993).
2. A.I. Bobenko, Y.B. Suris, Discrete Lagrangian reduction, discrete Euler–Poincaré equations, and semi-direct products, *Lett. Math. Phys.* **49**, 79–93 (1999).
3. A.I. Bobenko, Y.B. Suris, Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top, *Commun. Math. Phys.* **204**, 147–188 (1999).
4. N. Bou-Rabee, *Hamilton–Pontryagin integrators on Lie groups*, Ph.D. thesis, California Institute of Technology (2007).
5. J.R. Cardoso, F. Leite, The Moser–Veselov equation, *Linear Algebra Appl.* **360**, 237–248 (2003).
6. H. Cendra, J.E. Marsden, S. Pekarsky, T.S. Ratiu, Variational principles for Lie–Poisson and Hamilton–Poincaré equations, *Mosc. Math. J.* **3**, 833–867 (2003).
7. K. Engø, S. Faltinsen, Numerical integration of Lie–Poisson systems while preserving coadjoint orbits and energy, *SIAM J. Numer. Anal.* **39**, 128–145 (2001).
8. E. Hairer, C. Lubich, G. Wanner, *Geometric Numerical Integration*, 2nd edn. Springer Series in Computational Mathematics, vol. 31 (Springer, Berlin, 2006).
9. D.D. Holm, J.E. Marsden, T.S. Ratiu, The Euler–Poincaré equations and semidirect products with applications to continuum theories, *Adv. Math.* **137**, 1–81 (1998).
10. A. Iserles, H.Z. Munthe-Kaas, S.P. Nørsett, A. Zanna, Lie-group methods, *Acta Numer.* **9**, 215–365 (2000).
11. L. Kharevych, Weiwei, Y. Tong, E. Kanso, J.E. Marsden, P. Schroder, M. Desbrun, Geometric, variational integrators for computer animation, in *Eurographics/ACM SIGGRAPH Symposium on Computer Animation*, 2006.
12. S. Lall, M. West, Discrete variational Hamiltonian mechanics, *J. Phys. A Math. Gen.* **39**, 5509–5519 (2006).
13. T. Lee, M. Leok, N.H. McClamroch, Lie group variational integrators for the full body problem, *Comput. Methods Appl. Mech. Eng.* **196**, 2907–2924 (2007).
14. M. Leok, N.H. McClamroch, T. Lee, A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum, in *Proceedings of IEEE Conference on Control Applications* (2005), pp. 962–967.
15. A. Lew, J.E. Marsden, M. Ortiz, M. West, Variational time integrators, *Int. J. Numer. Methods Eng.* **60**, 153–212 (2004).
16. D. Lewis, J.C. Simo, Conserving algorithms for the dynamics of Hamiltonian systems on lie groups, *J. Nonlinear Sci.* **4**, 253–299 (1994).
17. D. Lewis, J.C. Simo, Conserving algorithms for the  $N$ -dimensional rigid body, *Fields Inst. Commun.* **10**, 121–139 (1996).
18. S. Leyendecker, J.E. Marsden, M. Ortiz, Variational integrators for constrained mechanical systems, Preprint (2007).
19. S. Leyendecker, S. Ober-Blöbaum, J.E. Marsden, M. Ortiz, Discrete mechanics and optimal control for constrained multibody dynamics, in *Proceedings of the 6th International Conference on Multibody Systems, Nonlinear Dynamics, and Control. ASME International Design Engineering Technical Conferences* (Proceedings of IDETC/MSNDC), 2007, pp. 1–10.
20. G.H. Livens, On Hamilton’s principle and the modified function in analytical dynamics, *Proc. R. Soc. Edinb.* **39**, 113 (1919).
21. J.E. Marsden, S. Pekarsky, S. Shkoller, Discrete Euler–Poincaré and Lie–Poisson equations, *Nonlinearity* **12**, 1647–1662 (1998).
22. J.E. Marsden, J. Scheurle, The reduced Euler–Lagrange equations, *Fields Inst. Commun.* **1**, 139–164 (1993).
23. J.E. Marsden, M. West, Discrete mechanics and variational integrators, *Acta Numer.* **10**, 357–514 (2001).
24. R. McLachlan, C. Scovel, Equivariant constrained symplectic integration, *J. Nonlinear Sci.* **5**, 233–256 (1995).
25. J. Moser, A.P. Veselov, Discrete versions of some classical integrable systems and factorization of matrix polynomials, *Commun. Math. Phys.* **139**, 217–243 (1991).
26. H. Munthe-Kaas, Lie–Butcher theory for Runge–Kutta methods, *BIT* **43**, 572–587 (1995).
27. H. Munthe-Kaas, Runge–Kutta methods on Lie groups, *BIT* **38**, 92–111 (1998).
28. H. Munthe-Kaas, B. Owren, Computations in a free Lie algebra, *Philos. Trans. R. Soc. A* **357**, 957–982 (1999).

29. H. Munthe-Kaas, A. Zanna, *Numerical Integration of Differential Equations on Homogeneous Manifolds* (Springer, Berlin, 1997), pp. 305–315.
30. J.C. Simo, L. Vu-Quoc, On the dynamics in space of rods undergoing large motions—a geometrically exact approach, *Comput. Methods Appl. Mech. Eng.* **66**, 125–161 (1988).
31. J.C. Simo, T.S. Wong, Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum, *Int. J. Numer. Methods Eng.* **31**, 19–52 (1991).
32. Y.B. Suris, Hamiltonian methods of Runge–Kutta type and their variational interpretation, *Math. Model.* **2**, 78–87 (1990).
33. J.M. Wendlandt, J.E. Marsden, Mechanical integrators derived from a discrete variational principle, *Physica D* **106**, 223–246 (1997).
34. H. Yoshimura, J.E. Marsden, Dirac structures and Lagrangian mechanics Part I: Implicit Lagrangian systems, *J. Geom. Phys.* **57**, 133–156 (2006).
35. H. Yoshimura, J.E. Marsden, Dirac structures in Lagrangian mechanics. Part II: Variational structures, *J. Geom. Phys.* **57**, 209–250 (2006).