Hamilton-Pontryagin Integrators on Lie Groups Part I: 
Introduction & Structure-Preserving Properties

Nawaf Bou-Rabee*  Jerrold E. Marsden†

December, 2007

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.

Contents

1 Introduction 2
2 Background and Setting 3
3 HP Mechanics 7
4 Lie Group VPRK Integrators 11
5 Conclusion 21

*Applied & Computational Math., Caltech, Pasadena, CA 91125 (nawaf@acm.caltech.edu).
†Control & Dynamical Systems, Caltech, Pasadena, CA 91125 (marsden@cds.caltech.edu). Research partially supported by the National Science Foundation through NSF grant DMS-0204474.
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 Livens [1919]; Lall & West [2006]; Kharevych et al. [2006]; Yoshimura and Marsden [2006a,b]. 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:

$$
S_{\text{HP}} = \int_a^b L(q,v)dt + \int_a^b \langle p, \dot{q} - v \rangle dt.
$$

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 $q(t)$, $v(t)$, $p(t)$ are all varied independently. If $p$ 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 $t_k = hk + a, k = 0, ..., N$. Let $s$ be the number of stages in the RK method. In analogy with the continuous system, the discrete HP action integral takes the following form:

$$
S_{\text{HP}}^d = \sum_{k=0}^{N-1} \sum_{i=1}^s hb_i L(Q_i^k, V_i^k) \\
+ \sum_{k=0}^{N-1} \sum_{i=1}^s h \left( p_{ki} \frac{Q_i^k - q_k}{h} - \sum_{j=1}^s a_{ij} V_j^k \right) + h \left( p_k \frac{q_{k+1} - q_k}{h} - \sum_{j=1}^s b_j V_j^k \right). 
$$

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 Suris [1990]; Marsden & West [2001]; Hairer, Lubich, and Wanner [2006].

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 $g : \mathbb{R}^n \to \mathbb{R}^k$, embeds $G$ in a larger linear space, and uses Lagrange multipliers to enforce the constraint. This approach is discussed in Bou-Rabee & Owhadi [2007b]. The corresponding constrained action takes the following form:

$$
\mathcal{S}_d = \sum_{k=0}^{N-1} \sum_{i=1}^{s} h \left[ b_i L(Q^i_k, V^i_k) + \left\langle p^i_k, \frac{Q^i_k - q^i_k}{h} - \sum_{j=1}^{s} a_{ij} V^j_k \right\rangle \right. \\
+ \left. \left\langle p^{k+1}, \frac{q^{k+1} - q_k}{h} - \sum_{j=1}^{s} b_j V^j_k \right\rangle + b_i \left\langle \Lambda^i_k, g(Q^i_k) \right\rangle \right].
$$

(1.3)

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

$$
\mathcal{S}_{HP} = \int_a^b \ell(g, \xi) dt + \int_a^b \langle \mu, g^{-1} \dot{g} - \xi \rangle dt.
$$

(1.4)

Then an equivalence is established between critical points of $\mathcal{S}_{HP}$ and $\mathcal{S}_{HP}$. If the Lagrangian is left-invariant, it is shown that this principle unifies the system’s Lie-Poisson and Euler-Poincaré descriptions Marsden & Scheurle [1993]; Cendra, Marsden, Pekarsky, and Ratiu [2003]. 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 Munthe-Kaas [1995]; Munthe-Kaas & Zanna [1997]; Munthe-Kaas [1998]; Munthe-Kaas & Owren [1999]. 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.

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; Leyendecker, Marsden, and Ortiz [2007]) and non-conservative effects (via their virtual work) Marsden & West [2001], as well as discrete optimal control (see Leyendecker, Ober-Blöbaum, Marsden, and Ortiz [2007] 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 in computing 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 2-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 Hairer, Lubich, and Wanner [2006]. Standard integrators often introduce spurious dynamics in long-time simulations, e.g., artificially corrupt chaotic invariant sets is well–illustrated in a computation from Bou-Rabee [2007], namely of a Poincaré section of an underwater vehicle in Fig. 2.1 using a fourth-order accurate Runge-Kutta (RK4) method and a variational Euler (VE) method designed for rigid-body type systems.

\[\begin{array}{lll}
(a) \text{RK4} & h=0.025 & (b) \text{RK4} & h=0.05 \\
(c) \text{VE} & h=0.025 & (d) \text{VE} & h=0.05 \\
\end{array}\]

Figure 2.1: 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.

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, Marsden, Ortiz, and West [2004] found that variational integrators correctly compute the time-averaged instantaneous temperature (mean kinetic energy over all particles) over long-time intervals, whereas standard methods (even a higher-order accurate one) exhibit a artificial drift in this statistical
quantity. These structure-preserving properties of variational integrators motivated their extension to stochastic Hamiltonian systems.

**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 the so-called Lie-Newmark method due to Simo & Vu-Quoc [1988] and Simo & Wong [1991]. These methods were motivated by the need to develop conserving algorithms that efficiently simulate the structural dynamics of rods and shells. For example, the configuration space of a discrete, three-dimensional finite-strain rod model, would involve \( N \) copies of \( \mathbb{R}^3 \times \text{SO}(3) \) where \( N \) is the number of points in the discretization of the line of centroids of the rod. For each point on the line of centroids, the orientation of the rod at that point is specified by an element of \( \text{SO}(3) \). In such models the mathematical description of the rotational degrees of freedom at these points is equivalent to the EP description of a free rigid body with added nonconservative effects due to the elastic coupling between points.

It was not apparent that the proposed Lie-Newmark methods had the necessary structure-preserving properties. In fact, Simo & Wong proposed another set of algorithms which preserve momentum by using the coadjoint action on \( \text{SO}(3) \) to advance the flow. Such integrators will be referred to as *coadjoint-preserving methods*. Later, Austin et al. [1993] showed that the midpoint rule member of the Lie-Newmark family with a Cayley reconstruction procedure was, in fact, a coadjoint-preserving method for \( \text{SO}(3) \). 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 & Scovel [1995] 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 and energy preserving methods of the Simo & Wong type that further preserve the symplectic structure were developed for \( \text{SO}(3) \) by Lewis & Simo [1994, 1996]. This was done by defining a one-parameter family of coadjoint and energy-preserving algorithms of the Simo & 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 continuous symplectic form.

Endowing coadjoint methods with energy-preserving properties was also the subject of Engø & Faltinsen [2001]. Specifically, they introduced integrators of the Runge-Kutta Munthe-Kaas type that preserved coadjoint orbits and energy using the coadjoint action on \( \text{SO}(3) \) 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 Moser & Veselov [1991]; Wendlandt and Marsden [1997]; Marsden, Pekarsky, and Shkoller [1998]; Bobenko and Suris [1999a]; Bobenko & Suris [1999b]. Moser and Veselov derived a variational integrator for the free rigid body by embedding \( \text{SO}(3) \) in the linear space of \( 3 \times 3 \) matrices, \( \mathbb{R}^9 \), and using
Lagrange multipliers to constrain the matrices to SO(3). This procedure was subsequently generalized to Lagrangian systems on more general configuration manifolds by the introduction of a discrete Hamilton’s principle on the larger linear space with holonomic constraints to constrain to the configuration manifold in Wendlandt and Marsden [1997]. They also considered the specific example of deriving a variational integrator for the free rigid body on the Lie group $S^3$ by embedding $S^3$ into $\mathbb{R}^4$ and using a holonomic constraint. The constraint ensured that the configuration update remained on the space of unit quaternions (a Lie group) and was enforced using a Lagrange multiplier.

Another approach is to use reduction to derive variational integrators on reduced spaces. Marsden, Pekarsky, and Shkoller [1998] 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 to higher-order accuracy in Leok, McClamroch, and Lee [2005] and Lee, Leok, and McClamroch [2007].

Bobenko and Suris [1999a] 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 Holm et al. [1998]). 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 Marsden, Pekarsky, and Shkoller [1998]. In addition, Bobenko & Suris [1999b] used this theory to determine and analyze an elegant, integrable discretization of the Lagrange top.

The perspective in this paper on Lie group variational integrators is different. Recognizing that Euler’s equations for a rigid body are in fact decoupled from the dynamics on the Lie group, and more generally, that the EP equation is decoupled from the dynamics on the Lie group, the paper aims to develop discrete variational schemes that analogously consist of a reconstruction rule and discrete EP equations that can be solved independently of the reconstruction equation and on a lower dimensional linear space. As mentioned in the overview the central idea is to discretize the reduced HP principle.

**Organization of the Paper.** In §3 continuous HP mechanics and its reduction is presented. In particular, it is shown that the reduced and unreduced HP variational principle are equivalent to Hamilton’s and the EP variational principles. Moreover, properties of the HP flow map are verified mainly to guide the discrete theory. In §4 the reduced discrete analog of the HP theory is developed. Properties of the discrete flow map are verified including discrete momentum map and symplectic form preservation. The theory is illustrated on several specific examples. In §5 the
structure-preserving Lie group integrators relevant to this paper are presented. In §6 the free rigid body and underwater vehicle examples are presented, the structure-preserving methods from §5 specialized to these examples, and results of numerical experiments are presented.

**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

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 $L : TG \to \mathbb{R}$ will be derived.

The HP principle unifies the Hamiltonian and Lagrangian descriptions of a mechanical system, as shown in Yoshimura and Marsden [2006a,b]. It states the following critical point condition on $TG \oplus T^*G$,

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

where $(g(t), v(t), p(t)) \in TG \oplus T^*G$ 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 3.1.** Following standard conventions, the left action of $G$ on $TG$ or $T^*G$ is denoted by simple concatenation. The left-trivialized Lagrangian $\ell : G \times \mathfrak{g} \to \mathbb{R}$ 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:

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

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. To see this, we proceed as follows.
Let $S(g,v,p)$ denote the HP action functional or integral,

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

Fixing the interval $[a,b]$, we regard $S$ as a map on path space: $S : C(TG \oplus T^*G) \to \mathbb{R}$, where

$$C(TG \oplus T^*G) = \{(g,v,p) : [a,b] \to TG \oplus T^*G \mid (g,v,p) \in C^\infty([a,b], TG \oplus T^*G)\}.$$ 

Then a simple calculation shows that,

$$S(g,v,p) = \int_{a}^{b} \left[ L(g,g\xi) + \langle p, gg^{-1}(\dot{g} - v) \rangle \right] dt = \int_{a}^{b} \left[ \ell(g,\xi) + \langle gp, g^{-1}(\dot{g} - v) \rangle \right] dt = \int_{a}^{b} \left[ \ell(g,\xi) + \langle \mu, g^{-1}\dot{g} - \xi \rangle \right] dt = s(g,\xi,\mu)$$

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

**Theorem 3.2.** Consider a Lagrangian system on a Lie group $G$ with Lagrangian $L : TG \to \mathbb{R}$. Let $\ell : G \times \mathfrak{g} \to \mathbb{R}$ 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

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

holds, where $(g(t),v(t),p(t)) \in TG \oplus T^*G$, can be varied arbitrarily and independently with endpoint conditions $g(a)$ and $g(b)$ fixed;
4. the left-trivialized HP principle

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

holds, where \((g(t), \xi(t), \mu(t)) \in G \times g \times g^*\) 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 \(g\) and \(g^*\) respectively, consistent with the results of Marsden & Scheurle [1993] and Cendra, Marsden, Pekarsky, and Ratiu [2003].

The HP Flow. From the left-trivialized HP principle, the variations of \(s\) with respect to \(\xi\) and \(\mu\) give

- varying \(\mu\) gives \(\xi = g^{-1} \dot{g}\), (reconstruction equation), \((3.1)\)
- varying \(\xi\) gives \(\mu = \frac{\partial \ell}{\partial \xi}(g, \xi)\), (Legendre transform). \((3.2)\)

Also, setting the variation of \(s\) with respect to \(g\) equal to zero gives

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

Observe that

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

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

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

Substituting this relation into \((3.3)\) gives

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

Integration by parts and using the boundary conditions on \(g\) yields

\[ \int_a^b \left[ \left\langle -\frac{d}{dt} \mu + \mathrm{ad}_\xi \mu + g \frac{\partial \ell}{\partial g}, \eta \right\rangle \right] dt = 0. \]
Since the variations are arbitrary, one arrives at
\[ \frac{d}{dt} \mu = \text{ad}_\xi^* \mu + g \frac{\partial \ell}{\partial g}. \]  

(3.4)

In sum, the left-trivialized HP equations are given by:
\[
\begin{aligned}
\frac{d}{dt} g &= g \xi, \\
\frac{d}{dt} \mu &= \text{ad}_\xi^* \mu + g \frac{\partial \ell}{\partial g}, \\
\mu &= \frac{\partial \ell}{\partial \xi}(g, \xi).
\end{aligned}
\]

(3.5)

Assuming that the Legendre transform is invertible, (3.5) describes an IVP on the left-trivialized space \(G \times g \times g^*\).

**Definition 3.3.** Let \(I_{HP}\) denote the admissible space and defined as,
\[
I_{HP} := \left\{ (g, v, p) \in TG \oplus T^*G \ \bigg| \ p = \frac{\partial L}{\partial v}(g, v) \right\}.
\]

(3.6)

Let \(I_{hp}\) denote its left-trivialization and defined as the subset of \(G \times g \times g^*\) that satisfies (3.2), i.e.,
\[
I_{hp} := \left\{ (g, \xi, \mu) \in G \times g \times g^* \ \bigg| \ \mu = \frac{\partial \ell}{\partial \xi}(g, \xi) \right\}.
\]

(3.7)

The natural projection is denoted by \(\pi_{HP} : TG \oplus T^*G \to T^*G\) and defined as,
\[
\pi_{HP}(g, v, p) := (g, p), \quad \pi_{HP}^{-1}(g, v) = (g, v, p), \quad (g, v) = FL^{-1}(g, p)
\]

where \(FL\) is the Legendre transform.

Given a time-interval \([a, b]\) and an initial \((g(a), \xi(a), \mu(a)) \in I_{hp}\), one can solve for \((g(b), \xi(b), \mu(b)) \in I_{hp}\) 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 \(I_{hp}\) be called the left-trivialized HP flow map, \(F_{hp} : I_{hp} \to I_{hp}\).

The flow map \(F_{hp}\) is equivalent to the HP flow on \(I_{HP}\) through left trivialization which defines a diffeomorphism between \(TG \oplus T^*G\) and \(G \times g \times g^*\), and hence, between \(I_{HP}\) and \(I_{hp}\). Through \(\pi_{HP}\) 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 \(\pi_{HP}\) is not a diffeomorphism from \(TG \oplus T^*G\) to \(T^*G\), it is a diffeomorphism when its domain is restricted to \(I_{HP}\). Thus, the left-trivialized HP, 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 \(\pi_{I_{HP}} = \pi_{HP}|_{I_{HP}}\). The manifold \(TG \oplus T^*G\) is a presymplectic manifold with the HP presymplectic form, \(\Omega_{HP} = \pi_{HP}^* \Omega\), and the
manifold $\mathcal{I}_{HP}$ is a symplectic manifold with the $HP$ symplectic form, $\Omega_{I_{HP}} = \pi^*_{I_{HP}} \Omega$. Similarly, the manifold $G \times \mathfrak{g} \times \mathfrak{g}^*$ is a presymplectic manifold with the presymplectic form $\omega_{HP}$ that is obtained by pulling-back the $HP$ presymplectic form by the left trivialization of $TG \oplus T^*G$, $\phi: G \times \mathfrak{g} \times \mathfrak{g}^* \to TG \oplus T^*G$, i.e., $\omega_{HP} = \phi^* \Omega_{HP}$. However, if the left-trivialization is restricted to $I_{hp}$, $\phi_{I_{hp}} = \phi|_{I_{hp}}$, then $I_{hp}$ is a symplectic manifold with the symplectic form given by $\omega_{I_{hp}} = \phi^*_{I_{hp}} \Omega_{I_{HP}}$.

**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 $I_{hp}$, $\hat{s}: I_{hp} \to \mathbb{R}$. The differential of $\hat{s}$ can be written as,

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

where we have introduced the left-trivialized HP one-form, $\theta_{I_{hp}} = \phi_{I_{hp}}^* \Theta_{I_{HP}}$. Since $d^2 \hat{s} = 0$, observe that

$$d^2 \hat{s} = (F_{hp})^* \omega_{I_{hp}} - \omega_{I_{hp}} = 0.$$ 

And hence, as a map on $I_{hp}$, $F_{hp}$ is symplectic.

**Theorem 3.4.** Left-trivialized HP flows preserve the symplectic two-form $\omega_{I_{hp}}$.

### 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. After introducing the map $\tau$ which is typified by the exponential map, and its properties, we use an $s$-stage Runge-Kutta-Munthe-Kaas (RKMK) approximation to the reconstruction equation, which leads naturally to the introduction of VPRK Integrators on Lie groups. This includes the Störmer-Verlet method for Lie groups, variational 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 $\tau : \mathfrak{g} \to G$. 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 Iserles, Munthe-Kaas, Nørsett, and Zanna [2000]. 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 Iserles, Munthe-Kaas, Nørsett, and Zanna [2000].

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

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

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

Figure 4.1 illustrates the geometry behind this definition.

![Figure 4.1: Derivative of $\tau$.](image)

**Lemma 4.3.** *The following identity holds,*

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

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

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.

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 $d\tau^{-1}: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which satisfies for $g = \tau(\xi)$,

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

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

Figure 4.2 illustrates the geometry behind this definition.

$$D\tau^{-1}(g) \cdot \delta = d\tau^{-1}_{\xi}(TR_{\tau(-\xi)} \delta) \quad d\tau^{-1}: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

**Figure 4.2:** Derivative of $\tau^{-1}$. 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 $\tau^{-1}$).

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

**Lemma 4.5.** The following identity holds,

$$d\tau^{-1}_{\xi}(\delta) = d\tau^{-1}_{-\xi}(Ad_{\tau(-\xi)} \delta).$$
Proof. This follows directly from Lemma 4.3. Let \( \delta \rightarrow d\tau^{-1}_\xi(\delta) \) in that identity to obtain
\[
\delta = \text{Ad}_{\tau(\xi)} d\tau_\xi(\delta).
\]
And now solve this equation for \( d\tau^{-1}_\xi(\delta) \),
\[
d\tau^{-1}_\xi(\delta) = d\tau^{-1}_\xi \left( \text{Ad}_{\tau(-\xi)} \delta \right).
\]

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 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: Munthe-Kaas [1995]; Munthe-Kaas & Zanna [1997]; Munthe-Kaas [1998]; Munthe-Kaas & Owren [1999]. The idea behind those papers is to use canonical coordinates on the Lie group to transform the differential equation on \(TG\), e.g., given by,
\[
\dot{g} = gf(t,g), \quad g(0) = g_0, \quad g(t) \in G, \quad f(t,g(t)) \in \mathfrak{g}.
\]
to a differential equation on \(\mathfrak{g}\). Specifically, substitute the following parametrization \(g(t) = g_0\tau(\Theta(t))\) into (4.1) to obtain,
\[
\dot{g} = TL_{g_0} TR_{\tau(\Theta)} d\tau \Theta = TL_{g_0} TL_{\tau(\Theta)} f(t,g).
\]
Using Lemma 4.3 this equation can be rewritten as,
\[
TL_{\tau(-\Theta)} TR_{\tau(\Theta)} d\tau \Theta = \text{Ad}_{\tau(-\Theta)} d\tau \Theta = d\tau \Theta = f(t,g).
\]
Solving for \(\dot{\Theta}\) gives
\[
\dot{\Theta} = d\tau^{-1}_\Theta f(t,g), \quad \Theta(0) = 0, \quad \Theta(t) \in \mathfrak{g}.
\] (4.2)
As described in the following definition, the RKMK method is obtained by applying an \(s\)-stage RK method to (4.2).

Definition 4.6. Consider the first-order differential equation \(\dot{g} = f(t,g)\) for the curve \((g(t), f(t,g(t))) \in TG\). Given coefficients \(b_i, a_{ij} \in \mathbb{R} \) \((i, j = 1, \cdots, s)\) and set \(c_i = \sum_{j=1}^{s} a_{ij}\). An \(s\)-stage Runge-Kutta-Munthe-Kaas (RKMK) approximation is given by
\[
G_k^i = g_k \tau(h\Theta_k^i),
\]
\[
\Theta_k^i = h \sum_{j=1}^{s} a_{ij} d\tau^{-1}_{-h\Theta_k^i} f(t_k + c_j h, G_k^j), \quad i = 1, \cdots, s,
\]
\[
g_{k+1} = g_k \tau \left( h \sum_{j=1}^{s} b_j d\tau^{-1}_{-h\Theta_k^j} f(t_k + c_j h, G_k^j) \right).
\]
(4.3) (4.4) (4.5)
If \( a_{ij} = 0 \) for \( i \leq j \) the RKMK method is called explicit, and implicit otherwise. The vectors \( g_k \) and \( G^i_k \) 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:

\[
\begin{array}{c|cccc}
& a_{11} & \cdots & a_{1s} \\
\vdots & \vdots & \ddots & \vdots \\
& a_{s1} & \cdots & a_{ss} \\
\hline
& b_1 & \cdots & b_s
\end{array}
\]

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

\[
\begin{align*}
\tau^{-1}(g_{k+1}^- G^i_k)/h &= \sum_{j=1}^{s} a_{ij} d\tau^{-1}_{-h\Theta^j_k} \Xi^j_k = \Theta^i_k, \quad i = 1, \ldots, s, \\
\tau^{-1}(g_{k}^- g_{k+1})/h &= \sum_{j=1}^{s} b_{j} d\tau^{-1}_{-h\Theta^j_k} \Xi^j_k.
\end{align*}
\]

where \( \Xi^i_k = \xi(t_k + c_i h) \). In practice one often truncates the series expansion of \( d\tau^{-1}_{-h\Theta^j_k} \). The following theorem guides how to do this without degrading the order of accuracy Hairer, Lubich, and Wanner [2006].

**Theorem 4.7.** Given a qth order approximant to the exact exponential: \( \tau : g \to G \). If the underlying RK method is of order \( p \) and the truncation index of \( d\tau^{-1}_{-h\Theta^j_k} \) satisfies \( q \geq p - 2 \) then the RKMK method is of order \( p \).

**VPRK Integrators on Lie Groups.** The discrete HP principle states that the discrete path the discrete system takes is one that extremizes a reduced action sum that will be introduced shortly. To discretize the action integral, (4.6) is treated as a constraint in the discrete HP action, and the integral of the left-trivialized Lagrangian is approximated by the following quadrature:

\[
\int_{t_k}^{t_k+h} \ell(g, \xi) dt \approx \sum_{i=1}^{s} h b_i \ell(G^i_k, \Xi^i_k).
\]

The truncation index of \( d\tau^{-1}_{-h\Theta^j_k} \) in (4.6) is chosen to be \( q = 0 \). By theorem 4.7 one can obtain second-order accurate methods from this principle.

**Definition 4.8.** Given an \( s \)-stage RKMK method with \( b_j \neq 0 \) for \( j = 1, \ldots, s \), define the discrete VPRK path space,

\[
C_d(g_1, g_2) = \{(g, \mu, \{\Theta^i, \Xi^i, \mu^i\}_{i=1}^{s}) : \{t_k\}_{k=0}^{N} \to (G \times g^*) \times (g \times g \times g^*)^s \mid g(t_0) = g_1, \ g(t_N) = g_2\}.
\]
and the *action sum* $s_d : C_d(g_1, g_2) \to \mathbb{R}$ as

$$s_d = \sum_{k=0}^{N-1} \sum_{i=1}^s \left[ b_i \ell(G_k^i, \Xi_k^i) + \left( \mu_k, \tau^{-1}(g_k^{-1}G_k^i)/h - \sum_{j=1}^s a_{ij} \Xi_k^j \right) \right] + \left( \mu_{k+1}, \tau^{-1}(g_k^{-1}g_{k+1})/h - \sum_{j=1}^s b_j \Xi_k^j \right).$$

(4.8)

Observe that $s_d$ is an approximation of the reduced HP action integral by numerical quadrature. The definition of $\tau$ as a map from $g$ to $G$ ensures that the pairings in the above sum are well defined. The discrete left-trivialized HP principle states that

$$\delta s_d = 0$$

for arbitrary and independent variations of the external stage vectors $(g_k, \mu_k) \in G \times g^*$ and the internal stage vectors $(\Theta_k^i, \Xi_k^i, \Psi_k^i) \in g \times g \times g^*$ for $i = 1, \cdots, s$ and $k = 0, \cdots, N$ subject to fixed endpoint conditions on $(g_k)_{k=0}^N$.

**Theorem 4.9.** Let $\ell : G \times g \to \mathbb{R}$ be a smooth, left-trivialized Lagrangian. A discrete curve $c_d \in C_d(g_1, g_2)$ satisfies the following VPRK scheme:

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

(4.9)

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

(4.10)

$$(d\tau^{-1}_{\xi_{k+1}})^* M_k^i = (d\tau^{-1}_{\Theta_k})^* \mu_k$$

$$+ h \sum_{j=1}^s b_{ij} (d\tau^{-1}_{\Theta_k})^* \left( a_{ij} (d\tau^{-1}_{\Theta_k})^* + b_{ij} (d\tau^{-1}_{\Theta_k})^* \right) \left( d\tau_{h\Theta_k} \right) \left( G_k \frac{\partial \ell}{\partial g} (G_k^j, \Xi_k^j) \right),$$

(4.11)

$$(d\tau^{-1}_{\Theta_k})^* \mu_{k+1} = (d\tau^{-1}_{\Theta_k})^* \mu_k + h \sum_{j=1}^s b_{ij} (d\tau^{-1}_{\Theta_k})^* \left( G_k^j \frac{\partial \ell}{\partial g} (G_k^j, \Xi_k^j) \right),$$

(4.12)

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

(4.13)

for $i = 1, \cdots, s$ and $k = 0, \cdots, N - 1$, if and only if it is a critical point of the function $s_d : C_d(g_1, g_2) \to \mathbb{R}$, that is, $d s_d(c_d) = 0$. Moreover, the discrete flow map defined by the above scheme, $F_h : I_{hp} \to I_{hp}$, preserves the symplectic form $\omega_{I_{hp}}$.

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

\[
\mathbf{d}s_d \cdot z = \sum_{k=0}^{N-1} \sum_{i=1}^s \left( h b_i G_k^i \frac{\partial \ell}{\partial \Xi_k^i} (G_k^i, \Xi_k^i) \cdot H_k^i + h b_i \frac{\partial \ell}{\partial \Xi_k^i} (G_k^i, \Xi_k^i) \cdot \delta \Xi_k^i \right) + h \left( \delta \mu_k^i, \tau^{-1}(g_k^{-1}G_k^i)/h - \sum_{j=1}^s a_{ij} \Xi_j^i \right) + h \left( \delta \mu_{k+1}, \tau^{-1}(g_k^{-1}g_{k+1})/h - \sum_{j=1}^s b_{ij} \Xi_j^i \right) + h \left( \mu_k, -d\tau^{-1}_h\Theta_k h\eta_k/h + d\tau^{-1}_h\Theta_k h - \sum_{j=1}^s a_{ij} \delta \Xi_j^i \right) + h \left( \mu_{k+1}, -d\tau^{-1}_h\Theta_{k+1} h\eta_k/h + d\tau^{-1}_h\Theta_{k+1} h - \sum_{j=1}^s b_{ij} \delta \Xi_j^i \right)
\]

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

\[
\mathbf{d}s_d \cdot z = \sum_{k=1}^{N-1} \sum_{i=1}^s h \left( \delta \mu_k^i, \tau^{-1}(g_k^{-1}G_k^i)/h - \sum_{j=1}^s a_{ij} \Xi_j^i \right) + h \left( \delta \mu_{k+1}, \tau^{-1}(g_k^{-1}g_{k+1})/h - \sum_{j=1}^s b_{ij} \Xi_j^i \right) + h \left( b_i \frac{\partial \ell}{\partial \Xi_k^i} (G_k^i, \Xi_k^i) - \sum_{j=1}^s a_{ji} \mu_k^j - b_i \mu_{k+1}, \delta \Xi_k^i \right) + \left( (d\tau^{-1}_h\Theta_k)^i \mu_k^i + h b_i G_k^i \frac{\partial \ell}{\partial \Xi_k^i} (G_k^i, \Xi_k^i), H_k^i \right) + \left( - (d\tau^{-1}_h\Theta_{k+1})^i \mu_{k+1} + (d\tau^{-1}_h\Theta_{k+1})^i \mu_k - \sum_{j=1}^s (d\tau^{-1}_h\Theta_k)^j \mu_k^j, \eta_k \right)
\]

Since \( \mathbf{d}s_d(c_d) = 0 \) if and only if \( \mathbf{d}s_d \cdot z = 0 \) for all \( z \in T_{c_d}C_d \), one arrives at the desired equations with the elimination of \( \mu_k^i \) and the introduction of the internal stage variables \( M_k^i = \partial \ell / \partial \Xi_k^i (G_k^i, \Xi_k^i) \) for \( i = 1, \cdots, s \). Conversely, if \( c_d \) satisfies (4.9)–(4.13) then \( \mathbf{d}s_d(c_d) = 0 \).

Consider the subset of \( C_d \) given by solutions of (4.9)–(4.13). Let \( \hat{s}_d \) denote the restriction of \( s_d \) to this space. Since each of these solutions is determined by a point in \( I_{hp} \), one can identify this space with \( I_{hp} \), and hence, \( \hat{s}_d : I_{hp} \to \mathbb{R} \). Since \( \hat{s}_d \) is restricted to solution space,

\[
\mathbf{d}\hat{s}_d(g_0, \xi_0, \mu_0) \cdot (\delta g_0, \delta \xi_0, \delta \mu_0) = \left( (d\tau^{-1}_h\xi_N)^* \mu_N, g_N^{-1} \delta g_N \right) - \left( (d\tau^{-1}_h\xi_0)^* \mu_0, g_0^{-1} \delta g_0 \right).
\]

Preservation of \( \omega_{I_{hp}} \) follows from \( \mathbf{d}^2 \hat{s}_d = 0 \). 

\[ \blacksquare \]
The external and internal stages of (4.9)–(4.13) define update schemes on $G \times g^*$ and $(g \times g \times g^*)^s$, respectively.

**Störmer-Verlet Integrators on Lie Groups.** The generalization of the Störmer-Verlet method to Lie groups is given by evaluating (4.9)–(4.13) at the following two-stage RK tableau (implicit trapezoidal rule),

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

Given $h$ and $(g_k, \mu_k)$, the method determines $(g_{k+1}, \mu_{k+1})$ by solving the following system of equations:

1. $M_k^{1/2} = \frac{\partial \ell}{\partial \xi}(g_k, \Xi^1_k)$, 
   
   \hspace{1cm} (4.14)

2. $M_k^{1/2} = \frac{\partial \ell}{\partial \xi}(g_{k+1}, \Xi^2_k)$, 
   
   \hspace{1cm} (4.15)

3. $(d\tau_{-h\Xi_{k+1}}^{-1})^* M_k^{1/2} = (d\tau_{-h\Xi_k}^{-1})^* \mu_k + \frac{h}{2} g_k \frac{\partial \ell}{\partial g}(g_k, \Xi^1_k)$, 
   
   \hspace{1cm} (4.16)

4. $g_{k+1} = g_k \tau \left( h^{-1/2} (\Xi^1_k + \Xi^2_k) \right)$, 
   
   \hspace{1cm} (4.17)

5. $\mu_{k+1} = M_k^{1/2} + \frac{h}{2} (d\tau_{-h\Xi_{k+1}}^{-1})^* g_{k+1} \frac{\partial \ell}{\partial g}(g_{k+1}, \Xi^2_k)$.
   
   \hspace{1cm} (4.18)

In particular, one uses the following procedure:

- Eliminate $g_{k+1}$ in (4.15) using (4.17). Then solve for $M_k^{1/2}$, $\Xi^1_k$, and $\Xi^2_k$ using (4.14)–(4.16). This update is in general implicit.
- Update $g_{k+1}$ using (4.17). This update is explicit.
- Solve for $\mu_{k+1}$ using (4.18). This update is explicit.

Observe that if the Lagrangian is separable, then (4.16) is not implicit in the potential force term and one does not need to eliminate $g_{k+1}$ in (4.15) using (4.17).

**Variational Euler on Lie Groups.** The variational Euler schemes come from evaluating (4.9)–(4.13) with the following tableaux:

- \[ \begin{array}{c|cc}
0 & 0 & 0 \\
1 & 1 & 1 \\
\end{array} \text{ forward Euler} \]
- \[ \begin{array}{c|cc}
0 & 0 & 0 \\
1 & 1 & 1 \\
\end{array} \text{ backward Euler} \]

The corresponding VPRK action sums take the following simple forms:

- $s^e_a = \sum_{k=0}^{N-1} h \left[ \ell(g_k, \xi_k) + \langle \mu_k, \tau^{-1}(g_k^{-1} g_{k+1})/h - \xi_k \rangle \right]$
- $s^e_b = \sum_{k=0}^{N-1} h \left[ \ell(g_{k+1}, \xi_{k+1}) + \langle \mu_{k+1}, \tau^{-1}(g_k^{-1} g_{k+1})/h - \xi_{k+1} \rangle \right]
Given \( h \) and \((g_k, \mu_k)\), the forward variational Euler method determines \((g_{k+1}, \mu_{k+1})\) by solving the following system of equations:

\[
\begin{align*}
    g_{k+1} &= g_k \tau (h \xi_k), \\
    (d \tau^{-1}_{h \xi_k})^* \mu_{k+1} &= (d \tau^{-1}_{-h \xi_k})^* \mu_k + h g_{k+1} \frac{\partial \ell}{\partial g}(g_{k+1}, \xi_{k+1}), \\
    \mu_{k+1} &= \frac{\partial \ell}{\partial \xi}(g_{k+1}, \xi_{k+1}).
\end{align*}
\] (4.19)

The backward variational Euler method determines \((g_{k+1}, \mu_{k+1})\) by solving the following system of equations:

\[
\begin{align*}
    g_{k+1} &= g_k \tau (h \xi_{k+1}), \\
    (d \tau^{-1}_{h \xi_{k+1}})^* \mu_{k+1} &= (d \tau^{-1}_{-h \xi_k})^* \mu_k + h g_{k+1} \frac{\partial \ell}{\partial g}(g_k, \xi_k), \\
    \mu_{k+1} &= \frac{\partial \ell}{\partial \xi}(g_{k+1}, \xi_{k+1}).
\end{align*}
\] (4.20)

**Euler-Poincaré Integrators.** In the case when the Lagrangian is \( G \)-left-invariant, the angular momentum updates in the above methods are identical and given by:

\[(d \tau^{-1}_{h \xi_k})^* \mu_{k+1} = (d \tau^{-1}_{-h \xi_k})^* \mu_k \] (4.21)

**Examples.** We now give various examples of Euler-Poincaré integrators by making different choices of the map \( \tau \) and evaluating (4.21).

(a) **Matrix exponential.** Suppose

\[\tau = \exp(\xi), \quad \tau : g \rightarrow G,\]

which is a local diffeomorphism.

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

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

where \( B_j \) are the Bernoulli numbers; see §3.4 of Hairer, Lubich, and Wanner [2006] for a detailed exposition and derivation.

Hence, (4.21) takes the form,

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

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

\[\tau(\xi) = \text{cay}(\xi) = (e - \xi/2)^{-1}(e + \xi/2),\] (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

\[
dcay(x)y = \left( e^{-x/2} \right)^{-1} y(e + x/2)^{-1}, \quad dcay^{-1}(x)y = (e - x/2)y(e + x/2).
\]

For a derivation and exposition the reader is referred to §4.8.3 of Hairer, Lubich, and Wanner [2006]. Using these expressions (4.21) can be written as,

\[
\mu_k = \mu_{k-1} + \frac{\hbar}{2} \text{ad}_{\xi_k} \mu_k + \frac{\hbar}{2} \text{ad}_{\xi_{k-1}} \mu_{k-1} + \frac{\hbar^2}{4} \left( \xi_k \mu_k \xi_k^* - \xi_{k-1} \mu_{k-1} \xi_{k-1}^* \right). \quad (4.26)
\]

(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:

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

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

\[
dskew(x)(y) = \frac{y\tau(x) - (y\tau(x))^*}{2}. \quad (4.27)
\]

Cardoso & Leite [2003] obtained the following theorem that explicitly determines \(\tau(\xi).\) Moreover, they give necessary and sufficient conditions for its existence.
Theorem 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.

Proof. Since the skew-symmetric part of $g$ is $\xi$, one can write $g$ as a sum of $\xi$ and a symmetric matrix $S$,

$$\tau(\xi) = S + \xi.$$

Observe that $\xi$ commutes with $\tau(\xi)$ since

$$2\xi \tau(\xi) = (\tau(\xi) - \tau(\xi)^*) \tau(\xi) = \tau(\xi)^2 - e = 2\tau(\xi)\xi.$$

Moreover, $S$ satisfies an algebraic Riccati equation because,

$$\tau(\xi)^* \tau(\xi) = e \implies S^2 + S\xi - \xi S - (\xi^2 + e) = 0.$$

And since $\xi$ commutes with $S$ (because it commutes with $g$),

$$S^2 = (\xi^2 + e),$$

which completes the proof. ■

Hence, (4.21) can be written as,

$$\mu_k \left( h^2 \xi_k^2 + e \right)^{1/2} + \left( h^2 \xi_k^2 + e \right)^{1/2} \mu_k$$

$$= \frac{\mu_{k-1} \left( h^2 \xi_{k-1}^2 + e \right)^{1/2} + \left( h^2 \xi_{k-1}^2 + e \right)^{1/2} \mu_{k-1}}{2}$$

$$+ \frac{h}{2} \text{ad}_{\xi_k}^* \mu_k + \frac{h}{2} \text{ad}_{\xi_{k-1}}^* \mu_{k-1}$$

(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 Marsden & West [2001];
- explain the numerics behind the Poincaré sections provided in Figure 2.1;
- 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

Austin, M. A., P. S. Krishnaprasad, and L. Wang [1993], Almost-Poisson Integration of Rigid Body Systems. *Journal of Computational Physics*, **107**, 105–117.

Bobenko, A. I., and Y. B. Suris [1999a], Discrete Lagrangian reduction, discrete Euler-Poincaré equations, and semi-direct products. *Lett. Math. Phys.*, **49**, 79–93.

Bobenko, A. I., and Y. B. Suris [1999b], Discrete time Lagrangian Mechanics on Lie groups, with an application to the Lagrange top. *Comm. Math. Phys.*, **204**, 147–188.

Bou-Rabee, N. [2007], *Hamilton-Pontryagin integrators on Lie groups*, PhD thesis, California Institute of Technology.

Bou-Rabee, N., and H. Owhadi [2007b], Stochastic Variational Partitioned Runge-Kutta Integrators for Constrained Systems. Submitted; arXiv:0709.2222.

Cardoso, J., R., and F. Leite [2003], The Moser-Veselov Equation. *Linear Algebra and its Applications*, **360**, 237–248.

Celledoni, E., and A. Iserles [2001], Methods for the approximation of the matrix exponential in a Lie-algebraic setting. *IMA J. Num. Anal.*, **21**, 463–488.

Cendra, H., J. E. Marsden, S. Pekarsky, and T. S. Ratiu [2003], Variational principles for Lie-Poisson and Hamilton-Poincaré equations, *Mosc. Math. J.*, **3**, 833–867.

Dahlquist, G. [1975], Error analysis for a class of methods for stiff nonlinear initial boundary value problems. *Lecture Notes in Mathematics*, **506**, 60–74.

Dullweber, A., B. Leimkuhler, and R. McLachlan [1997], Symplectic splitting methods for rigid body molecular dynamics. *J. Chem. Phys.*, **107**, 5840–5851.

Engø, K., and S. Faltinsen [2001], Numerical integration of Lie-Poisson systems while preserving coadjoint orbits and energy. *SIAM J. Numer. Anal.*, **39**, 128–145.

Feng, K. [1986], Difference Schemes for Hamiltonian Formalism and Symplectic Geometry. *J. Comp. Math.*, **4**, 279–289.

Gonzalez, O., and J. C. Simo [1996], On the stability of symplectic and energy-momentum algorithms for nonlinear Hamiltonian systems with symmetry. *Comput. Methods Appl. Mech. Eng.*, **134**, 197–222.

Hairer, E., C. Lubich, and G. Wanner [2006], *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition.

Holm, D. D., J. E. Marsden, and T. S. Ratiu [1998], The Euler–Poincaré equations and semidirect products with applications to continuum theories, *Adv. in Math.*, **137**, 1–81.
REFERENCES

Holmes, P., J. Jenkins, and N. E. Leonard [1998], Dynamics of the Kirchhoff equations I: Coincident centers of gravity and buoyancy. *Phys. D*, **118**, 311–342.

Iserles, A., H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna [2000], Lie-group methods. *Acta numerica.*, **9**, 215–365.

Iserles, A. [2001], On Cayley-Transform methods for the discretization of Lie-group equations. *Found. Comp. Maths.*, **1**, 129–160.

Kane, C., J. E. Marsden, M. Ortiz, and M. West [2000], Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *Int. J. Num. Meth. Eng’g.*, **49**, 1295–1325.

Kane, C., J. E. Marsden, and M. Ortiz [1999], Symplectic-Energy-Momentum Preserving Variational Integrators. *J. Math. Phys.*, **40**, 3353–3371.

Kanso, E., J. E. Marsden, C. W. Rowley, and J. B. Melli-Huber [2005], Locomotion of articulated bodies in a perfect fluid. *J. Nonlinear Sci.*, **15**, 255–289.

Kharevych, L., Weiwei, Y. Tong, E. Kanso, J. E. Marsden, P. Schroder, and M. Desbrun [2006], Geometric, Variational Integrators for Computer Animation. *Eurographics/ACM SIGGRAPH Symposium on Computer Animation*.

Lall, S. and M. West [2006], *Discrete variational Hamiltonian mechanics*. Journal of Physics A: Mathematical and General, **39**, 5509–5519.

Leimkuhler, B., and S. Reich [2004], Simulating Hamiltonian Dynamics. *Cambridge Monographs on Applied and Computational Mathematics*, **14**.

Leok, M., N. H. McClamroch, and T. Lee [2005], A Lie Group Variational Integrator for the Attitude Dynamics of a Rigid Body with Applications to the 3D Pendulum. *Proc. IEEE Conf. on Control Applications*, 962–967.

Lee, T., M. Leok, and N. H. McClamroch [2007], Lie group variational integrators for the full body problem, *Comput. Methods Appl. Mech. Engrg*. **196**, 2907–2924.

Lew, A., J. E. Marsden, M. Ortiz, and M. West [2003], Asynchronous Variational Integrators. *Arch. Rational Mech. Anal.*, **167**, 85–146.

Lew, A., J. E. Marsden, M. Ortiz, and M. West [2004a], An Overview of Variational Integrators. *Finite Element Methods: 1970’s and Beyond*, CIMNE Barcelona, 98–115.

Lew, A., J. E. Marsden, M. Ortiz, and M. West [2004], Variational time integrators, *Internat. J. Numer. Methods Engrg*. **60**, 153–212.

Lewis, D. and J. C. Simo [1994], Conserving Algorithms for the Dynamics of Hamiltonian Systems on Lie Groups. *J. Nonlinear Science*, **4**, 253–299.

Lewis, D. and J. C. Simo [1996], Conserving Algorithms for the N-dimensional rigid body. *Fields. Inst. Comm.*, **10**, 121–139.
Leyendecker, S., S. Ober-Blöbaum, J. E. Marsden, and M. Ortiz [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 International Design Engineering Technical Conferences., volume Proceedings of IDETC/MSNDC 2007, pages 1–10.

Leyendecker, S., J. E. Marsden, and M. Ortiz [2007], Variational integrators for constrained mechanical systems, (preprint).

Livens, G. H. [1919], On Hamilton’s principle and the modified function in analytical dynamics, Proc. Roy. Soc. Edingburgh 39, 113.

MacIver, A., E. Fontaine, and J. W. Burdick [2004], Designing future underwater vehicles: Principles and mechanisms of the weakly electric fish. IEEE J. Oceanic Engineering, 29, 651–659.

MacKay, R. [1992], Some Aspects of the Dynamics of Hamiltonian Systems. The Dynamics of Numerics and the Numerics of Dynamics, Clarendon Press, 137–193.

Marsden, J. E., and J. Scheurle [1993], The reduced Euler-Lagrange Equations. Fields. Inst. Comm., 1, 139–164.

Marsden, J. E., S. Pekarsky, and S. Shkoller [1998], Discrete Euler-Poincaré and Lie-Poisson Equations. Nonlinearity, 12, 1647–1662.

Marsden, J. E., and T. Ratiu [1999], Introduction to Mechanics and Symmetry. Springer Texts in Applied Mathematics, 17.

Marsden, J. E., and M. West [2001], Discrete mechanics and variational integrators. Acta Numerica, 10, 357–514.

Mclachlan, R., and C. Scovel [1995], Equivariant constrained symplectic integration. J. Nonlinear Science, 5, 233–256.

Moser, J., and A. P. Veselov [1991], Discrete versions of some classical integrable systems and factorization of matrix polynomials. Comm. Math. Phys., 139, 217–243.

Munthe-Kaas, H. [1995], Lie-Butcher theory for Runge-Kutta methods. BIT, 572–587.

Munthe-Kaas, H. [1998], Runge-Kutta methods on lie groups. BIT, 38, 92–111.

Munthe-Kaas, H. and B. Owren [1999], Computations in a free lie algebra. Phil. Trans Royal Society A, 357, 957–982.

Munthe-Kaas, H. and A. Zanna [1997], Numerical integration of differential equations on homogeneous manifolds. Springer, 305–315.
Ruth, R. D. [1983], A canonical integration technique. *IEEE Transactions on Nuclear Science*, **30**, 2669–2671.

Suris, Y. B. [1990], Hamiltonian methods of Runge-Kutta type and their variational interpretation. *Math. Modelling*, **2**, 78–87.

Simo, J. C. and L. Vu-Quoc [1988], On the Dynamics in Space of Rods Undergoing Large Motions - A Geometrically Exact Approach. *Comp. Methods in Applied Mech. and Eng’g.*, **66**, 125–161.

Simo, J. C. and T. S. Wong [1991], Unconditionally Stable Algorithms for Rigid Body Dynamics That Exactly Preserve Energy and Momentum. *Int. J. Numer. Methods Engrg.*, **31**, 19–52.

Veselov, M. [1988], Integrable discrete-time systems and difference operators. *Func. An. and Appl.*, **22**, 83–94.

de Vogelaere, R. [1956], Methods of Integration which Preserve the Contact Transformation Property of the Hamiltonian Equations. Report No. 4, Dept. Math., Univ. of Notre Dame.

Wendlandt, J. M. and J. E. Marsden [1997], Mechanical integrators derived from a discrete variational principle, *Physica D* **106**, 223–246.

Yoshimura, H. and J. E. Marsden [2006], Dirac structures and Lagrangian Mechanics Part I: Implicit Lagrangian systems, *J. Geom. and Physics* **57**, 133–156.

Yoshimura, H. and J. E. Marsden [2006b], Dirac structures in Lagrangian mechanics Part II: Variational structures, *J. Geom. and Physics* **57**, 209–250.