IMA Journal of Numerical Analysis (2009) **29**, 421–443
doi:10.1093/imanum/drn018
Advance Access publication on May 9, 2008

Stochastic variational integrators

NAWAF BOU-RABEE†
Applied and Computational Mathematics (ACM), Caltech, Pasadena, CA 91125, USA

AND

HOUMAN OWHADI‡
Applied and Computational Mathematics (ACM) and Control and Dynamical Systems (CDS), Caltech, Pasadena, CA 91125, USA

[Received on 21 October 2007; revised on 13 February 2008]

This paper presents a continuous and discrete Lagrangian theory for stochastic Hamiltonian systems on manifolds, akin to the Ornstein–Uhlenbeck theory of Brownian motion in a force field. The main result is to derive governing SDEs for such systems from a critical point of a stochastic action. Using this result, the paper derives Langevin-type equations for constrained mechanical systems and implements a stochastic analogue of Lagrangian reduction. These are easy consequences of the fact that the stochastic action is intrinsically defined. Stochastic variational integrators (SVIs) are developed using a discrete variational principle. The paper shows that the discrete flow of an SVI is almost surely symplectic and in the presence of symmetry almost surely momentum-map preserving. A first-order mean-squared convergent SVI for mechanical systems on Lie groups is introduced. As an application of the theory, SVIs are exhibited for multiple, randomly forced and torqued rigid bodies interacting via a potential.

Keywords: variational integrators; Ornstein-Uhlenbeck process; stochastic Hamiltonian systems.

1. Introduction

Since the foundational work of Bismut (1981), the field of stochastic geometric mechanics is emerging in response to the demand for tools to analyse continuous and discrete mechanical systems with uncertainty (Bismut, 1981; Liao, 1997; Liao & Wang, 2005; Milstein *et al.*, 2002, 2003; Talay, 2002; Vanden-Eijnden & Ciccotti, 2006; Lazaro-Cami & Ortega, 2007a,b; Malham & Wiese, 2007; Ciccotti *et al.*, 2008). Within this context, the goal of this paper is to develop variational integrators for the simulation of stochastic Hamiltonian systems on manifolds. For this purpose, the paper develops a Lagrangian description of stochastic Hamiltonian systems akin to the Ornstein–Uhlenbeck theory of Brownian motion in a force field. Other approaches to random mechanics include Feynmann's path integral approach to quantum mechanics (Feynmann & Hibbs, 1981) and Nelson's stochastic mechanics (Nelson, 1985). In the context of the former, there is also a generalization of Noether's theorem presented in Thieullen & Zambrini (2008).

1.1 *Variational integrators*

Variational integration theory derives integrators for mechanical systems from discrete variational principles (Veselov, 1988; MacKay, 1992; Wendlandt & Marsden, 1997; Marsden & West, 2001). The

†Email: nawaf@acm.caltech.edu
‡Email: owhadi@acm.caltech.edu

theory includes discrete analogues of the Lagrangian Noether theorem, the Euler–Lagrange equations and the Legendre transform. Variational integrators can readily incorporate holonomic constraints (via, e.g., Lagrange multipliers) and nonconservative effects (via, e.g., their virtual work) (Wendlandt & Marsden, 1997; Marsden & West, 2001). Altogether, this description of mechanics stands as a self-contained theory of mechanics comparable to Hamiltonian, Lagrangian or Newtonian mechanics. One of the distinguishing features of variational integrators is their ability to accurately compute statistics of mechanical systems, such as in computing Poincaré sections, the instantaneous temperature of a system, etc. For example, as a consequence of their variational construction, variational integrators are symplectic (de Vogelaere, 1956; Ruth, 1983; Feng, 1986). A single-step integrator applied to a mechanical system is called ‘symplectic’ if the discrete- flow map that it defines exactly preserves the canonical symplectic form; otherwise it is 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 (for a detailed exposition, see Hairer *et al.*, 2006). Standard integrators often introduce spurious dynamics in long-time simulations, e.g. artificially corrupt chaotic invariant sets as illustrated in Fig. 2.1 of Bou-Rabee & Marsden (2008). The figure compares computations of Poincaré sections of an underwater vehicle obtained using a fourth-order accurate Runge-Kuta (RK4) method and a second-order accurate variational Euler (VE) method. In particular, for a sufficiently long time-span of integration, the RK4 method is shown to corrupt chaotic invariant sets while the lower-order accurate VE method preserves such structures.

In addition to correctly computing chaotic invariant sets and long-time excellent energy behaviour, here is mounting evidence that variational integrators accurately compute other statistics of mechanical systems. For example, in a simulation of a coupled spring–mass lattice, Lew *et al.* (2004, Fig. 1) found that variational integrators accurately 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 an artificial drift in this statistical quantity. These structure-preserving properties of variational integrators are the motivation for their extension to stochastic Hamiltonian systems.

1.2 *Main results*

In his foundational work, Bismut showed that the stochastic flow of certain randomly perturbed Hamiltonian systems on flat spaces extremizes a stochastic action. He called such systems ‘stochastic Hamiltonian systems’ and used this property to prove symplecticity and extend Noether’s theorem to such systems (Bismut, 1981). Mean-squared symplectic integrators for stochastic Hamiltonian systems on flat spaces and driven by Wiener processes have been developed (Milstein *et al.*, 2002, 2003).

Bismut’s work was further enriched and generalized to manifolds by recent work (Lazaro-Cami & Ortega, 2007a,b). Lazaro-Cami & Ortega (2007b) showed that stochastic Hamiltonian systems on manifolds extremize a stochastic action defined on the space of manifold-valued semimartingales. Moreover, they performed a reduction of stochastic Hamiltonian systems on the cotangent bundle of a Lie group to obtain stochastic Lie–Poisson equations (Lazaro-Cami & Ortega, 2007a). However, as far as we can tell, the converse to Bismut’s original theorem, namely that a critical point of a stochastic action satisfies stochastic Hamilton’s equations, has not been proved. In fact, as pointed out by Lazaro-Cami & Ortega (2007b), a counterexample can be constructed to prove that the converse to Bismut’s theorem is not true ‘for a certain choice of stochastic action’.

In this paper, we restrict our attention to stochastic Hamiltonian systems driven by Wiener processes and assume that the space of admissible curves in configuration space is of class C^1. From the viewpoint of randomly perturbed mechanical systems, this latter restriction is reasonable since random effects often appears, not in the kinematic equation, but rather in the balance of momentum equation as white noise forces and torques. It should be mentioned that the ideas in this paper can be readily extended to stochastic Hamiltonian systems driven by more general semimartingales, but for the sake of clarity we restrict to Wiener processes. Within this context, the results of the paper are as follows.

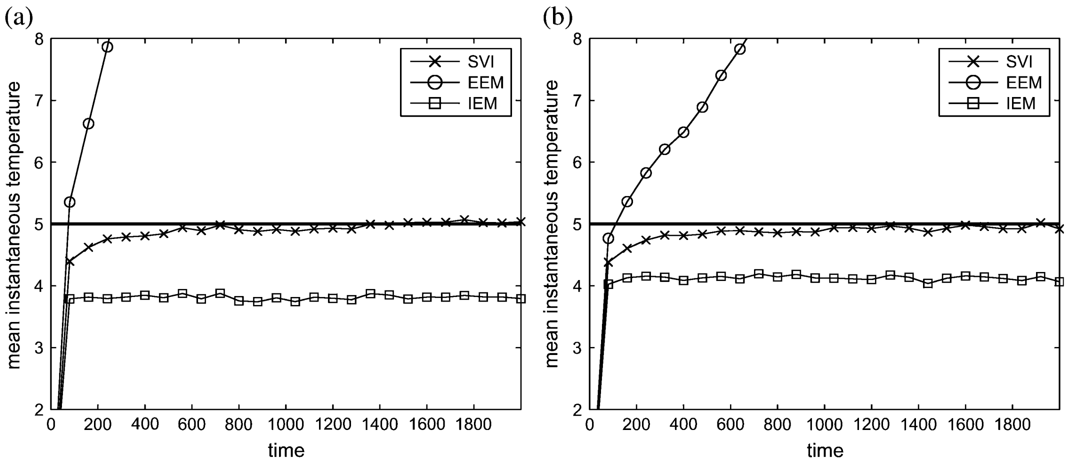
- For a class of mechanical systems whose configuration space is a paracompact manifold and which is subjected to multiplicative white noise forces and torques, the paper proves almost surely that a curve satisfies stochastic Hamilton equations if and only if it extremizes a stochastic action. This theorem is the main result of the paper.
- The paper derives governing SDEs for stochastic Hamiltonian systems with holonomic constraints using a constrained variational principle, and for stochastic Hamiltonian systems with nonconservative force in the drift vector field using a Lagrange–d'Alembert principle (for deterministic treatments, see Marsden & Ratiu, 1999). The paper performs Lagrangian reduction for stochastic Hamiltonian systems whose configuration space is a Lie group, and provides stochastic Euler–Poincaré/Lie–Poisson equations for such systems (for deterministic treatments, see Marsden & Ratiu, 1999). These are easy consequences of the fact that the stochastic action is intrinsically defined.
- The paper shows how to discretize variational principles to obtain stochastic variational integrators (SVIs), stochastic RATTLE-type integrators for constrained stochastic Hamiltonian systems and stochastic Euler–Poincaré/Lie–Poisson integrators for stochastic Hamiltonian systems on Lie groups (for deterministic treatments, see Moser & Veselov, 1991; Wendlandt & Marsden, 1997; Marsden *et al.*, 1998; Hairer *et al.*, 2006). In addition, the paper describes how to derive quasi-symplectic methods for rigid-body-type systems at uniform temperature.

1.3 *Organization of the paper*

Sufficient conditions for existence, uniqueness and almost sure differentiability of stochastic flows on manifolds are recalled in Section 2. In Section 3, we extend the Hamilton–Pontryagin (HP) principle to the stochastic setting to prove that a class of mechanical systems with multiplicative noise appearing as forces and torques possess a variational structure. It should be emphasized (and it is explained in the section) that the mechanical system could evolve on a nonlinear configuration space and involve holonomic constraints or nonconservative effects in the drift. The HP viewpoint is adopted since it unifies the Hamiltonian and Lagrangian descriptions of the system. By left trivializing this principle, we also show how to perform Lagrangian reduction in this stochastic setting for stochastic rigid-body-type systems.

In Section 4, SVIs are derived from an abstract discrete Lagrangian and the structure of the resulting discrete-flow map is analysed. In Section 5, we concretely show how to design a single-step, stochastic VE integrator for mechanical systems whose configuration space is a Lie group using a simple stochastic discrete HP principle. If the configuration space is flat, the resulting SVIs are in one-to-one correspondence with symplectic integrators for stochastic Hamiltonian systems (Bismut, 1981; Milstein *et al.*, 2002, 2003) and, with the addition of dissipation in the drift term, in one-to-one correspondence with quasi-symplectic methods for Langevin-type systems. These symplectic and quasi-symplectic integrators have been numerically tested and shown to possess excellent properties for computing energy behaviour and statistics of mechanical systems governed by Langevin-type equations; see Milstein & Tretyakov (2003, 2004).

Our own simulations confirm those findings. A sample of such results is provided in Fig. 1. It compares an SVI to standard, presumably nonvariational methods on a ballistic pendulum at uniform



$h = 0.1$ $h = 0.05$

FIG. 1. Ballistic pendulum at uniform temperature (Bou-Rabee & Owhadi, 2007). Plots of the mean instantaneous temperature (kinetic energy) of a ballistic pendulum computed using an SVI, EEM and IEM for time steps h as indicated. The correct temperature is indicated by the solid line. Observe that the EEM and IEM schemes artificially heat and cool the system, respectively. A key feature of the ballistic pendulum is that the diffusion and drift matrices associated to the momentums are degenerate, yet the system is still at uniform temperature.

temperature (Bou-Rabee & Owhadi, 2007). The figure shows that an SVI correctly computes the temperature of the system (defined as the mean of the instantaneous temperature with respect to realizations), whereas explicit Euler–Maruyama (EEM) and implicit Euler–Maruyama (IEM) schemes do not. All these methods are first-order mean-squared convergent. This computation suggests that an SVI has favourable energy behaviour, whereas EEM and IEM artificially heat and cool the system, respectively. On the other hand, this paper focuses on SVI theory and the structure-preserving properties of SVIs.

In Section 6, as an application we explain how one can add multiplicative white noise forces and torques to multiple rigid bodies in a fashion that preserves variational structure. With the addition of dissipation, these become Langevin-type equations. An SVI is provided for such systems.

It is easy to check first-order accuracy in the mean-squared (or L^2) sense using standard stochastic numerics (see, e.g. Talay, 1995, and Milstein & Tretyakov, 2004, for an expository treatment of stochastic numerics). We address this matter in Bou-Rabee & Owhadi (2008). In the references Talay (1995, 2002) and Milstein & Tretyakov (2004), one mainly considers the approximation of statistics of the law of the solution (moments at finite times, invariant measures, etc.) which is the aim of Monte Carlo or ergodic simulations. Those works do not emphasize L^2-estimates because in practice L^2-estimates are not necessary for convergence in law since the exact solution and the discretization scheme may not live on the same probability space. Furthermore, the estimates on the approximation of statistics that can be deduced from L^2-estimates are often crude and do not lead to the true convergence rates. Finally, multiple stochastic integrals cannot be simulated in a pathwise sense.

2. Stochastic flows on manifolds

Some standard results on flows of SDEs on manifolds are reviewed here for the reader's convenience. The reader is referred to the following textbooks on the subject for more detailed exposition: Elworthy

(1982), Emery (1989), Ikeda & Watanabe (1989) and Kunita (1990). This section parallels the treatment of deterministic flows on manifolds found in Chapter 4 of Abraham *et al.* (2007).

We start by introducing notation for deterministic vector fields on manifolds which are an important component of SDEs on manifolds. Let M be an n-manifold. Recall that a vector field on M is a section of the tangent bundle TM of M. The set of all C^k vector fields on M is denoted by $\mathfrak{X}^k(M)$.

The notion of a probability space is introduced in order to extend the definition of a dynamical system to incorporate noise. A stochastic dynamical system consists of a base flow on the probability space which propagates the noise and a stochastic flow on M which depends on the noise.

DEFINITION 2.1 (Stochastic dynamical system) A 'stochastic dynamical system' consists of a base flow on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a stochastic flow on a manifold M. The 'base flow' is a $\mathbb{P}$-preserving map $\theta\colon \mathbb{R} \times \Omega \to \Omega$, which satisfies

1. $\theta_0 = \mathrm{id}_\Omega\colon \Omega \to \Omega$ is the identity on Ω;
2. for all $s, t \in \mathbb{R}$, the group property: $\theta_s \circ \theta_t = \theta_{s+t}$.

Given times $0 \leqslant r \leqslant s \leqslant t$, the 'stochastic flow' on M is a map $\varphi_{t,s}\colon \Omega \times M \to M$ such that

1. for almost all $\omega \in \Omega$, the map $(s, t, \omega, x) \mapsto \varphi_{t,s}(\omega)x$ is continuous in s, t and x;
2. $\varphi_{s,s}(\omega) = \mathrm{id}_M\colon M \to M$ is the identity map on M for all $s \in \mathbb{R}$;
3. φ satisfies the cocycle property

$$\varphi_{t,s}(\theta_s(\omega)) \circ \varphi_{s,r}(\omega) = \varphi_{t,r}(\omega).$$

This paper is concerned with stochastic dynamical systems that come from 'stochastic laws of motion', i.e. ones whose stochastic flows define solutions of SDEs. The Stratonovich definition of stochastic integrals is adopted to extend SDEs from flat spaces to manifolds, the main advantage of the Stratonovich approach being that the chain rule holds for the Stratonovich differential. Consider a manifold M modelled on a Banach space E and vector fields $X_i \in \mathfrak{X}^k(M)$ for $i = 0, \dots, m$. Let $\mathcal{F}_t$ be a nondecreasing family of σ-subalgebras of $\mathcal{F}$ and let $(W_i(t, \omega), \mathcal{F}_t)$, $i = 1, \dots, m$, be independent Wiener processes for $0 \leqslant t \leqslant T$. In terms of these objects, the Stratonovich SDE that the paper considers takes the form

$$\mathrm{d}z = X_0(z)\mathrm{d}t + \sum_{i=1}^{m} X_i(z) \circ \mathrm{d}W_i, \quad z(0) = z_0. \tag{2.1}$$

X_0 is referred to as the 'drift vector field' and X_i, $i = 1, \dots, m$, are the 'diffusion vector fields'. A 'Stratonovich integral curve' of (2.1) is a C^0-map $c(\cdot, \omega)\colon [0, T] \to M$, which satisfies

$$c(t, \omega) = z_0 + \int_0^t X_0(c(s, \omega))\mathrm{d}s + \sum_{i=1}^{m} \int_0^t X_i(c(s, \omega)) \circ \mathrm{d}W_i(s, \omega),$$

for all $t \in [0, T]$. Uniqueness of solutions to (2.1) will be defined in the pathwise sense.

DEFINITION 2.2 (Pathwise uniqueness) Let c be a Stratonovich integral curve of (2.1). 'Pathwise uniqueness' of c means that if $\bar{c}\colon I \to M$ is also a solution to (2.1) on the same filtered probability space with the same Brownian motion and initial random variable, then

$$P\left(c(t, \omega) = \bar{c}(t, \omega), \forall\, t \in [0, T]\right) = 1.$$

In the following, we define the mean-squared norm on the model space E with the understanding that this notion can be extended to M using a local representative.

DEFINITION 2.3 (Mean-squared norm) The mean squared norm of $f\colon E \times \Omega \to E$ is given by

$$\| f(x, \omega)\| = (\mathbb{E}(|f(x, \omega)|^2))^{1/2}.$$

As is standard, for the rest of the paper the explicit dependence of stochastic maps on the point $\omega \in \Omega$ will usually be suppressed. With these definitions, one can state the following key, but standard, theorem (Elworthy, 1982; Emery, 1989; Ikeda & Watanabe, 1989; Kunita, 1990).

THEOREM 2.4 (Existence, uniqueness and smoothness) Let M be a manifold with model space E. Suppose that $X_i \in \mathfrak{X}^k(M)$, $i = 0, \ldots, m$ and $k \geqslant 1$, are uniformly Lipschitz and measurable with respect to $x \in M$. Let $I = [0, T]$. Then the following statements hold.

1. For each $u \in M$, there is almost surely a C^0-curve $c\colon I \to M$ such that $c(0) = u$ and c satisfies (2.1) for all $t \in I$. This curve $c\colon I \to M$ is called a 'maximal solution'.
2. c is pathwise unique.
3. There is almost surely a mapping $F\colon I \times M \to M$ such that the curve $c_u\colon I \to M$ defined by $c_u(t) = F_t(u)$ is a curve satisfying (2.1) for all $t \in I$. Moreover, almost surely F is C^k in u and C^0 in t.

3. Stochastic HP mechanics

In this section, a variational principle is introduced for a class of stochastic Hamiltonian systems on manifolds. The stochastic action presented is a sum of the classical action and several stochastic integrals. The key feature of this principle is that one can recover stochastic Hamilton equations for these systems. Roughly speaking, this is accomplished by means of taking variations of this action within the space of curves only (not the probability space) and imposing the condition that this "partial differential" of the action must be zero.

3.1 *Setting*

The setting is a paracompact, configuration manifold Q. In the context of this paper, a stochastic Hamiltonian system is specified by a Hamiltonian $H\colon T^*Q \to \mathbb{R}$ and m deterministic functions $\gamma_i\colon Q \to \mathbb{R}$ for $i = 1, \ldots, m$. Define the Lagrangian $\mathcal{L}\colon TQ \to \mathbb{R}$ to be the Legendre transform of H. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Fix an interval $[a, b] \subset \mathbb{R}$. To describe the stochastic perturbation, we introduce a probability space $(\Omega, \mathcal{F}, P)$ and $(W_i(t), \mathcal{F}_t)_{t\in[a,b]}$, for $i = 1, \ldots, m$, where $\{W_i\}_{i=1}^m$ are independent, real-valued Wiener processes and $\mathcal{F}_t$ is the filtration generated by these Wiener processes.

3.2 *Stochastic HP principle*

The paper adopts an HP viewpoint to develop a Lagrangian description of stochastic Hamiltonian systems. The HP principle unifies the Hamiltonian and Lagrangian descriptions of a mechanical system (Yoshimura & Marsden, 2006a,b; Bou-Rabee, 2007; Bou-Rabee & Marsden, 2008). The classical HP action integral will be perturbed using deterministic functions $\gamma_i\colon Q \to \mathbb{R}$ for $i = 1, \ldots, m$. Roughly speaking, in the stochastic context the HP principle states the following critical point condition on $TQ \oplus T^*Q$:

$$\delta \int_a^b \left[\mathcal{L}(q, v)\mathrm{d}t + \sum_{i=1}^m \gamma_i(q) \circ \mathrm{d}W_i + \left\langle p, \frac{\mathrm{d}q}{\mathrm{d}t} - v \right\rangle \mathrm{d}t \right] = 0,$$

where $(q(t), v(t), p(t)) \in TQ \oplus T^*Q$ are varied arbitrarily and independently, with end point conditions $q(a)$ and $q(b)$ fixed. This principle builds in a Legendre transform, stochastic Hamilton equations and stochastic Euler–Lagrange equations. The action integral in the above principle consists of two Lebesgue integrals with respect to t and m Stratonovich stochastic integrals with respect to W_i for $i = 1, \dots, m$. This action is random; i.e. for every sample point $\omega \in \Omega$ one will obtain a different, time-dependent Lagrangian system. However, each system possesses a variational structure which we will make precise in this section. For a deterministic treatment of time-dependent continuous and discrete Lagrangian systems, the reader is referred to Marsden & West (2001).

DEFINITION 3.1 (Pontryagin bundle) The 'Pontryagin bundle' of a manifold M is defined as $PM = TM \oplus T^*M$.

The Pontryagin bundle is a vector bundle over Q whose fibre at $q \in Q$ is the vector space $P_q Q = T_q Q \oplus T_q^* Q$. In terms of the Pontryagin bundle, we can define the path spaces of the stochastic Hamiltonian systems in question.

DEFINITION 3.2 (Path spaces) Fixing the interval $[a, b]$ and $q_a, q_b \in Q$, define the 'path space' as

$$\mathcal{C}(PQ) = \{(q, v, p) \in C^0([a, b], PQ) | q \in C^1([a, b], Q),\ q(a) = q_a, q(b) = q_b\}.$$

Let $\mathfrak{G}\colon \Omega \times \mathcal{C}(PQ) \to \mathbb{R}$ denote the 'stochastic HP action integral':

$$\mathfrak{G}(q, v, p) = \int_a^b \left[\mathcal{L}(q, v)\mathrm{d}t + \sum_{i=1}^m \gamma_i(q) \circ \mathrm{d}W_i(t) + \left\langle p, \frac{\mathrm{d}q}{\mathrm{d}t} - v \right\rangle \mathrm{d}t \right].$$

The HP path space is a smooth infinite-dimensional manifold. One can show that its tangent space at $c = (q, v, p) \in \mathcal{C}([a, b], q_1, q_2)$ consists of maps $w = (q, v, p, \delta q, \delta v, \delta p) \in C^0([a, b], T(PQ))$ such that $\delta q(a) = \delta q(b) = 0$ and $q, \delta q$ are of class C^1. Let $(q, v, p)(\cdot, \epsilon) \in \mathcal{C}(PQ)$ denote a one-parameter family of curves in $\mathcal{C}$ that is differentiable with respect to ϵ. Define the differential of $\mathfrak{G}$ as

$$\mathrm{d}\mathfrak{G} \cdot (\delta q, \delta v, \delta p) := \frac{\partial}{\partial \epsilon}\, \mathfrak{G}(\omega, q(t, \epsilon), v(t, \epsilon), p(t, \epsilon))|_{\epsilon=0}\,,$$

where

$$\delta q(t) = \frac{\partial}{\partial \epsilon} q(t, \epsilon)|_{\epsilon=0}, \quad \delta q(a) = \delta q(b) = 0, \quad \delta v(t) = \frac{\partial}{\partial \epsilon} v(t, \epsilon)|_{\epsilon=0}, \quad \delta p(t) = \frac{\partial}{\partial \epsilon} p(t, \epsilon)|_{\epsilon=0}.$$

In terms of this differential, one can state the following critical point condition.

THEOREM 3.3 (Stochastic variational principle of HP) Let $\mathcal{L}\colon TQ \to \mathbb{R}$ be a Lagrangian on TQ of class C^2 with respect to q and v and with globally Lipschitz first derivatives with respect to q and v. Let $\gamma_i\colon Q \to \mathbb{R}$ be of class C^2 and with globally Lipschitz first derivatives for $i = 1, \dots, m$. Then, almost surely, a curve $c = (q, v, p) \in \mathcal{C}(PQ)$ satisfies the stochastic HP equations

$$\begin{cases} \mathrm{d}q = v\,\mathrm{d}t, \\ \mathrm{d}p = \frac{\partial \mathcal{L}}{\partial q}\mathrm{d}t + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} \circ \mathrm{d}W_i, \\ p = \frac{\partial \mathcal{L}}{\partial v}, \end{cases} \tag{3.1}$$

if and only if it is a critical point of the function $\mathfrak{G}\colon \Omega \times \mathcal{C}(PQ) \to \mathbb{R}$, i.e. $\mathrm{d}\mathfrak{G}(c) = 0$.

Proof. Let us first prove almost surely that a critical point of the function satisfies (3.1). The differential of $\mathfrak{G}$ is given by

$$\begin{aligned} \mathrm{d}\mathfrak{G}(c) \cdot (\delta q, \delta v, \delta p) = \int_a^b \Bigg[& \frac{\partial \mathcal{L}}{\partial q} \cdot \delta q\, \mathrm{d}s + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} \cdot \delta q \circ \mathrm{d}W_i \\ & + \frac{\partial \mathcal{L}}{\partial v} \cdot \delta v\, \mathrm{d}s + \left\langle \delta p, \frac{\mathrm{d}q}{\mathrm{d}t} - v \right\rangle \mathrm{d}s + \left\langle p, \delta \frac{\mathrm{d}q}{\mathrm{d}t} - \delta v \right\rangle \mathrm{d}s \Bigg]. \end{aligned}$$

One can use a dominated convergence argument to show that differentiation and stochastic integration commute in the above stochastic integrals, as γ_i are of class C^2 for $i = 1, \ldots, m$, and the curves are continuous. Consider the term involving δp. Since δp is arbitrary and the integrand is continuous, the kinematic constraint holds: $\mathrm{d}q/\mathrm{d}t = v$. (This follows from the basic lemma that if $f, g \in C^0([a, b], \mathbb{R})$ and g is arbitrary, then $\int_a^b f(t)g(t)\mathrm{d}t = 0 \iff f(t) = 0\ \forall t \in [a, b]$.) Similarly, the Legendre transform is obtained from the δv term $\partial \mathcal{L}/\partial v = p$.

Collecting the variations with respect to δq in the differential gives

$$\int_a^b \left[\frac{\partial \mathcal{L}}{\partial q} \cdot \delta q\, \mathrm{d}s + \left\langle p, \delta \frac{\mathrm{d}q}{\mathrm{d}t} \right\rangle \mathrm{d}s + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} \cdot \delta q \circ \mathrm{d}W_i \right].$$

The first two terms are standard Lebesgue integrals and the last m terms are Stratonovich stochastic integrals. The following definition is introduced for notational convenience.

DEFINITION 3.4 Let $E = \mathbb{R}^n$. Given $f_1 \in C^0([0, T], E^*)$ and $f_2 \in C^1([0, T], E)$, define

$$\int_0^t \langle \mathrm{d} f_1,\ f_2 \rangle := \langle f_1,\ f_2 \rangle|_0^t - \int_0^t \left\langle f_1, \frac{\mathrm{d} f_2}{\mathrm{d}s} \right\rangle \mathrm{d}s, \quad t \in [0, T].$$

Using this definition and the boundary conditions $\delta q(a) = \delta q(b) = 0$, the following function $I\colon \mathcal{C}(q_1, q_2, [a, b]) \times C^1([a, b], TQ) \to \mathbb{R}$ is introduced:

$$I(q, v, p, f) = \int_a^b \left[\left\langle \frac{\partial \mathcal{L}}{\partial q}\mathrm{d}s + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} \circ \mathrm{d}W_i - \mathrm{d}p,\ f \right\rangle \right],$$

so that

$$I(q, v, p, \delta q) = \int_a^b \left[\frac{\partial \mathcal{L}}{\partial q} \cdot \delta q\, \mathrm{d}s + \left\langle p, \delta \frac{\mathrm{d}q}{\mathrm{d}t} \right\rangle \mathrm{d}s + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} \cdot \delta q \circ \mathrm{d}W_i \right].$$

In the following, it is shown that if $I(q, v, p, f) = 0$ for arbitrary f of class C^1, then (q, v, p) satisfy (3.1).

Let $\{U_\alpha, g_\alpha\}$ be a partition of unity on PQ. Expand I in terms of this partition of unity:

$$I = \sum_\alpha \int_a^b \left[g_\alpha(q, v, p) \left(\frac{\partial \mathcal{L}}{\partial q} \mathrm{d}t + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} \circ \mathrm{d}W_i - \mathrm{d}p \right) \cdot f \right].$$

Since the curves (q, v, p) are compactly supported, only a finite number of the g_α are nonzero. For each g_α nonzero, the terms in the integral can be expressed in local coordinates.

We will select f to single out the jth component of the covector field in I. Introduce the following function $h: \mathbb{R} \to \mathbb{R}$ for this purpose:

$$h(t) = 2\frac{t}{\epsilon} - \frac{t^2}{\epsilon^2}.$$

Observe that $h(0) = 0$, $h(\epsilon) = 1$ and $h'(\epsilon) = 0$. Let $\{e_j\}_{j=1}^n$ be a basis for the model space of Q. Now, fix j and define $f_\epsilon \in C^1([a, b], TQ)$ in local coordinates as follows:

$$f_\epsilon(s) = \begin{cases} h(s-a)e_j, & \text{if } a \leqslant s \leqslant a + \epsilon, \\ e_j, & \text{if } a + \epsilon < s < t - \epsilon, \\ h(t-s)e_j, & \text{if } t - \epsilon \leqslant s \leqslant t, \\ 0, & \text{if } t < s \leqslant b. \end{cases}$$

Introduce the following label to simplify subsequent calculations:

$$A(s) = \left(\frac{\partial \mathcal{L}}{\partial q}(q(s), v(s))\mathrm{d}s + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q}(q(s)) \circ \mathrm{d}W_i(s) - \mathrm{d}p(s) \right) \cdot e_j.$$

In terms of $A(s)$, one can write

$$I(q, v, p, f_\epsilon) = \sum_\alpha \left[\int_a^{a+\epsilon} h(s-a) g_\alpha(s) A(s) + \int_{a+\epsilon}^{t-\epsilon} g_\alpha(s) A(s) + \int_{t-\epsilon}^t h(t-s) g_\alpha(s) A(s) \right].$$

We will show in the mean-squared norm (cf. Definition 2.3)

$$\lim_{\epsilon \to 0} I(q, v, p, f_\epsilon) = \sum_\alpha \int_a^t g_\alpha A(s) =: I^*. \tag{3.2}$$

Using this result and the Borel–Cantelli lemma, one can deduce that there exists $\{\epsilon_n\}$ that converges to 0 such that $I(q, v, p, f_{\epsilon_n})$ almost surely converges to I^*. It follows that $I^* = 0$ almost surely.

We proceed to prove (3.2). Since $(a+b)^2 \leqslant 2a^2+2b^2$,

$$\begin{aligned}
&\left\| \sum_\alpha \int_a^t g_\alpha A(s) - I(q, \upsilon, p, f_\epsilon) \right\|^2 \\
&= \left\| \sum_\alpha \int_a^{a+\epsilon} (1-h(s-a)) g_\alpha A(s) + \int_{t-\epsilon}^t (1-h(t-s)) g_\alpha A(s) \right\|^2 \\
&\leqslant 2 \left\| \sum_\alpha \int_a^{a+\epsilon} (1-h(s-a)) g_\alpha A(s) \right\|^2 + 2 \left\| \sum_\alpha \int_{t-\epsilon}^t (1-h(t-s)) g_\alpha A(s) \right\|^2 .
\end{aligned}$$

We will only show how to bound the first term since bounding the second term is very similar. By continuity of (q, υ, p), one can pick ϵ small enough so that the support of (q, υ, p) lies in a single chart. On this chart, since q is differentiable, the Stratonovich–Ito conversion formula implies that

$$\int_a^{a+\epsilon} \frac{\partial \gamma_i}{\partial q} \cdot \delta q \circ \mathrm{d}W_i = \int_a^{a+\epsilon} \frac{\partial \gamma_i}{\partial q} \cdot \delta q \, \mathrm{d}W_i ,$$

for $i = 1, \ldots, m$. Therefore,

$$\begin{aligned}
&\left\| \sum_\alpha \int_a^{a+\epsilon} (1-h(s-a)) g_\alpha A(s) \right\|^2 \\
&= \left\| \int_a^{a+\epsilon} (1-h(s-a)) \left(\frac{\partial \mathcal{L}}{\partial q} \mathrm{d}s + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} \mathrm{d}W_i - \mathrm{d}p \right) \cdot e_j \right\|^2 \\
&\leqslant 3 \left\| \int_a^{a+\epsilon} (1-h(s-a)) \frac{\partial \mathcal{L}}{\partial q^j} \mathrm{d}s \right\|^2 + 3 \sum_{i=1}^m \left\| \int_a^{a+\epsilon} (1-h(s-a)) \frac{\partial \gamma_i}{\partial q^j} \mathrm{d}W_i \right\|^2 \\
&\quad +3 \left\| \int_a^{a+\epsilon} (1-h(s-a)) \mathrm{d}p \cdot e_j \right\|^2 .
\end{aligned}$$

Since $\frac{\partial \mathcal{L}}{\partial q^j}$ is continuous on $s \in [a, a+\epsilon]$, the first term can be bounded:

$$\left\| \int_a^{a+\epsilon} (1-h(s-a)) \frac{\partial \mathcal{L}}{\partial q^j} \mathrm{d}s \right\|^2 \leqslant \frac{M^2 \epsilon^2}{9} .$$

Similarly, by the Ito isometry and since $\frac{\partial \gamma_i}{\partial q^j}$ is continuous on $s \in [a, a+\epsilon]$, the second m terms can similarly be bounded; e.g. the ith Stratonovich integral can be bounded as follows:

$$\left\| \int_a^{a+\epsilon} (1-h(s-a)) \frac{\partial \gamma_i}{\partial q^j} \mathrm{d}W_i \right\|^2 = E\left(\int_a^{a+\epsilon} \left| (1-h(s-a)) \frac{\partial \gamma_i}{\partial q^j} \right|^2 \mathrm{d}s \right) \leqslant \frac{M^2 \epsilon}{5} .$$

Using Definition 3.4 and the integral mean-value theorem, the final term can be bounded as well:

$$\left\| \int_a^{a+\epsilon} (1 - h(s-a)) \mathrm{d}p \cdot e_j \right\|^2 = \left\| (1 - h(s-a)) p_j(s) \Big|_{s=a}^{s=a+\epsilon} + \int_a^{a+\epsilon} p_j(s) h'(s-a) \mathrm{d}s \right\|^2$$
$$= \| - p_j(a) + p_j(a + c_1 \epsilon) \|^2,$$

where $c_1 \in [0, 1]$ is a real constant. Since p_j is of class C^0, as $\epsilon \to 0$ this term vanishes. Since j is arbitrary, we have proved (3.2). Therefore, almost surely, if c is a critical point of $\mathfrak{G}$, then $\mathrm{d}\mathfrak{G}(c) \cdot w = 0$ for all $w \in T_c\mathcal{C}(PQ)$, and hence c satisfies the stochastic HP equations.

On the other hand, almost surely, if c satisfies (3.1), then it is a critical point of $\mathfrak{G}$. This direction is easy to confirm, since as a solution to the stochastic HP equations c is a measurable diffusion process. In fact, this direction is similar to the one that Bismut originally established, namely, that a solution of stochastic Hamilton equations extremizes an action function, although the stochastic action used by Bismut has a different domain from the stochastic action used in this proof (Bismut, 1981). □

As a consequence of the stochastic HP equations that the critical points of $\mathfrak{G}$ almost surely satisfy, it is easy to confirm that these critical points are adapted to the filtration $\mathcal{F}_t$ of the driving Wiener processes.

COROLLARY 3.5 If $z \in \mathcal{C}(PQ)$ extremizes $\mathfrak{G}$, then it is adapted to the filtration $(\mathcal{F}_t)_{t \geqslant 0}$ of the driving Wiener processes.

Equations (3.1) are a stochastic differential algebraic system of equations. Assuming that one can eliminate v using the Legendre transform, these equations can be viewed as a Cauchy problem. This paper is primarily concerned with forces or torques that appear as white noise in the balance of momentum equations, which explains the choice of $\gamma_i = \gamma_i(q)$. Observe that by the Ito–Stratonovich conversion formula, the Ito modification to the drift is equal to 0, and hence, (3.1) can be written in Ito form as

$$\mathrm{d}q = v \, \mathrm{d}t,$$
$$\mathrm{d}p = \frac{\partial \mathcal{L}}{\partial q} \mathrm{d}t + \sum_{i=1}^{m} \frac{\partial \gamma_i}{\partial q} \mathrm{d}W_i,$$
$$p = \frac{\partial \mathcal{L}}{\partial v}.$$

In what follows, structure-preserving properties of the flow map defined by the maximal solution of these equations over $[a, b]$ will be investigated. First, observe that because of the smoothness conditions assumed in Theorem 3.3, a solution almost surely exists and is pathwise unique on $[a, b]$ by the results in Section 2. When γ_i is constant for $i = 1, \ldots, m$, the reader is referred to the following texts for deterministic treatments of symplecticity, momentum map preservation and holonomically constrained mechanical systems: see Marsden & Ratiu (1999) and Marsden & West (2001).

3.3 *Symplecticity*

Assuming that one can eliminate v using the Legendre transform, the stochastic HP equations define a ‘stochastic flow’ on the symplectic manifold (T^*Q, κ), where κ is the canonical symplectic form (Marsden & Ratiu, 1999). We denote this flow by $F_t \colon T^*Q \to T^*Q$. With this assumption and in a more general context, Bismut extended the variational proof of symplecticity and Noether’s theorem to stochastic Hamiltonian systems (Bismut, 1981). In fact, one does not need to prove both directions of (3.3) to perform this extension. These proofs are repeated here in the context of stochastic Hamiltonian systems driven by Wiener processes for the reader's convenience and for completeness.

The variational proof of symplecticity will be used to show that this flow preserves κ (Marsden & Ratiu, 1999). By Theorem 2.4, assuming the Lagrangian is sufficiently smooth, $F_t\colon T^*Q \to T^*Q$ is almost surely differentiable.

The idea of the proof is to restrict $\mathfrak{G}$ to the space of pathwise unique solutions, i.e. to define $\hat{\mathfrak{G}} = \mathfrak{G}|_{\text{solutions}}$. On the same filtered probability space with the same Brownian motion, this solution space can be identified with the set of initial conditions; i.e. this restricted action can be expressed as $\hat{\mathfrak{G}}\colon T^*Q \to \mathbb{R}$. For each initial condition, by Theorem 2.4 there exists a pathwise unique solution almost surely. One then computes $\mathrm{d}\hat{\mathfrak{G}}$:

$$\mathrm{d}\hat{\mathfrak{G}}(q(a), p(a)) \cdot (\delta q(a), \delta p(a)) =$$
$$\int_a^b \left[\left(\frac{\partial \mathcal{L}}{\partial q} \mathrm{d}t + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} \circ \mathrm{d}W_i - \mathrm{d}p \right) \cdot \delta q + \delta p \cdot \left(\frac{\mathrm{d}q}{\mathrm{d}t} - v \right) \mathrm{d}s + \left(\frac{\partial \mathcal{L}}{\partial v} - p \right) \cdot \delta v \, \mathrm{d}t \right] + \langle p, \delta q \rangle|_a^b.$$

The integral in the above vanishes since $\hat{\mathfrak{G}}$ is restricted to solution space. The boundary terms define in local coordinates the canonical 1-form Θ on T^*Q. Computing $\mathrm{d}^2\hat{\mathfrak{G}}$ gives almost surely conservation of the canonical symplectic form.

THEOREM 3.6 (Conservation of stochastic symplectic form) Provided that one can eliminate v using the Legendre transform, the flow of (3.1) preserves the canonical symplectic form almost surely.

3.4 *Noether's theorem*

In what follows, we review for completeness Bismut's extension of Noether's theorem (Bismut, 1981). Let G be a Lie group with Lie algebra $\mathfrak{g}$. The 'left action' of G on Q is denoted by $\Phi\colon G \times Q \to Q$. The 'cotangent lift' of this action is likewise denoted by $\Phi^{T^*Q}\colon G \times T^*Q \to T^*Q$:

$$\Phi^{T^*Q}(h, q, p) = (\Phi(h, q), ((D_q\Phi(h, q))^{-1})^* \cdot p).$$

The corresponding 'infinitesimal generators' are $\xi^Q\colon Q \to T^*Q$ and $\xi^{T^*Q}\colon T^*Q \to T(T^*Q)$ and by definition we have

$$\xi^Q(q, p) = \frac{\mathrm{d}}{\mathrm{d}s}[\Phi(\exp(s\xi), q)]_{s=0}, \quad \xi^{T^*Q}(q, p) = \frac{\mathrm{d}}{\mathrm{d}s}[\Phi^{T^*Q}(\exp(s\xi), q, p)]_{s=0}.$$

This action gives rise to the following momentum map $J\colon T^*Q \to \mathfrak{g}^*$:

$$J(q, p) \cdot \xi = \langle p, \xi^Q(q, p) \rangle.$$

The following conservation law follows if $\mathfrak{G}$ is infinitesimally invariant with respect to the G-action. We remark in passing that infinitesimal invariance of $\mathfrak{G}$ follows from left invariance of the stochastic HP action with respect to the G-action.

THEOREM 3.7 (Stochastic Noether theorem) Let G be a Lie group. If $\mathfrak{G}$ is infinitesimally symmetric with respect to the left (or right) action of G, then the corresponding momentum map is conserved almost surely; i.e. $J = \langle p, \xi_Q(q) \rangle$ is a conserved quantity under the flow of (3.1).

Proof. This proof is terse. Consider the differential of $\hat{\mathfrak{G}}$ in the direction of ξ^{T^*Q}:

$$\mathrm{d}\hat{\mathfrak{G}}(q(a), p(a)) \cdot \xi^{T^*Q}(q(a), p(a)) = \langle p, \xi^Q(q, p)\rangle|_a^b.$$

Moreover, infinitesimal symmetry implies that

$$\mathrm{d}\hat{\mathfrak{G}} \cdot \xi^{T^*Q}(q(a), p(a)) = 0 \implies J(q(b), p(b)) \cdot \xi - J(q(a), p(a)) \cdot \xi = 0,$$

and hence J is conserved under the flow since ξ is arbitrary. □

3.5 *Holonomic constraints*

The following results will require the proof of the converse of Theorem 3.3. The setting in this part is an n-manifold Q and a stochastic Hamiltonian system with holonomic constraint. To be specific, suppose that the motion of the mechanical system is given by a constraint submanifold $S \subset Q$ defined as $S = g^{-1}(0)$, where $g\colon Q \to \mathbb{R}^k$, $k < n$, g is smooth and 0 is a regular value of g. A stochastic Hamiltonian system is specified by an unconstrained Hamiltonian $H\colon T^*Q \to \mathbb{R}$ and m deterministic functions $\gamma_j\colon Q \to \mathbb{R}$ for $j = 1, \ldots, m$. These functions $\{\gamma_j\}_{j=1}^m$ specify the structure of the noisy forces or torques. Let $\mathcal{L}\colon TQ \to \mathbb{R}$ be the Legendre transform of H. Set $\mathcal{L}^S = \mathcal{L}|_{TS}$ and $\gamma_j^S = \gamma_j|_S$ for $j = 1, \ldots, m$.

As opposed to using generalized coordinates on PS, we wish to describe the mechanical system using constrained coordinates on PQ and introduce Lagrange multipliers to enforce the constraint. However, because of the stochastic component of the action, one cannot introduce Lagrange multipliers in the standard way. Instead, we will introduce the Lagrange multiplier as a semimartingale using Definition 3.4. In particular, consider the following constrained variational principle:

$$\delta\left(\mathfrak{G} + \int_a^b \langle \mathrm{d}\lambda, g\rangle\right) = 0,$$

where using Definition 3.4 and the boundary conditions $g(q(a)) = g(q(b)) = 0$,

$$\int_a^b \langle \mathrm{d}\lambda, g\rangle = -\int_a^t \left\langle \lambda, \frac{\mathrm{d}g}{\mathrm{d}t}\right\rangle \mathrm{d}t, \quad t \in [a, b].$$

In this case, $\lambda(t)$ is a Lagrange multiplier dual to $\frac{\mathrm{d}}{\mathrm{d}t}g(q(t))$ for $t \in [a, b]$, and we assume that it is of class C^0. The corresponding equations of motion in constrained coordinates are obtained in a similar fashion to (3.1). To be precise, fix $q_a, q_b \in S$ and set $\mathcal{C}(Q) = \{q \in C^1([a, b], Q) \mid q(a) = q_a,\ q(b) = q_b\}$. Define a modified constraint function on the space of paths as $\Phi\colon \mathcal{C}(Q) \to C^0([a, b], \mathbb{R}^k)$ defined pointwise as $\Phi(q)(t) = \frac{\mathrm{d}}{\mathrm{d}t}(g(q(t)))$. In terms of these, one can prove the following equivalence.

THEOREM 3.8 (Constrained, stochastic HP principle) Given the constrained and unconstrained action integrals $\mathfrak{G}_c\colon \mathcal{C}(PS) \times \Omega \to \mathbb{R}$ and $\mathfrak{G}\colon \mathcal{C}(PQ) \times \Omega \to \mathbb{R}$ and the modified constraint function Φ : $\mathcal{C}(Q) \to C^0([0, T], \mathbb{R}^k)$, let $\langle\langle \cdot, \cdot \rangle\rangle$ denote the L_2-inner product on $C^0([0, T], \mathbb{R}^k)$. Then the following are equivalent:

(i) $z \in \mathcal{C}(PS)$ extremizes $\mathfrak{G}_c$, and hence, satisfies stochastic HP equations (cf. Theorem 3.3)

$$\begin{cases} \mathrm{d}q = v\,\mathrm{d}t, \\[2ex] \mathrm{d}p = \frac{\partial \mathcal{L}^S}{\partial q}(q, v)\mathrm{d}t + \sum_{j=1}^m \frac{\partial \gamma_j^S}{\partial q}(q) \circ \mathrm{d}W_j, \\[2ex] p = \frac{\partial \mathcal{L}^S}{\partial v}(q, v). \end{cases} \tag{3.3}$$

(ii) $z = (q, v, p) \in \mathcal{C}(PQ)$ and $\lambda \in C^0([0, T], \mathbb{R}^k)$ extremize the augmented action $\bar{\mathfrak{G}}(z, \lambda) = \mathfrak{G}(z) + \langle\langle \lambda, \Phi(q) \rangle\rangle$.

(iii) $z = (q, v, p) \in \mathcal{C}(PQ)$ and $\lambda \in C^0([0, T], \mathbb{R}^k)$ satisfy the constrained, stochastic HP equations

$$\begin{cases} \mathrm{d}q = v\,\mathrm{d}t, \\ \mathrm{d}p = \frac{\partial \mathcal{L}}{\partial q}(q, v)\mathrm{d}t + \sum_{j=1}^{m} \frac{\partial \gamma_j}{\partial q}(q) \circ \mathrm{d}W_j + \frac{\partial g}{\partial q}(q)^* \cdot \mathrm{d}\lambda, \\ p = \frac{\partial \mathcal{L}}{\partial v}(q, v), \\ g(q) = 0. \end{cases} \tag{3.4}$$

From this equivalence, it follows that the flow of (3.4) is mean-squared symplectic. For a proof of this theorem and more exposition, the reader is referred to Bou-Rabee & Owhadi (2008). With the modified constraint function, Φ, it is a standard application of the Lagrange multiplier theorem.

3.6 *Nonconservative effects*

Nonconservative effects are incorporated by considering the 'Lagrange–d'Alembert–Pontryagin principle'. In this principle, the effect of a nonconservative force appears as virtual work. Consider a force field $F\colon TQ \to T^*Q$. Then, the Lagrange–d'Alembert–Pontryagin principle is given by

$$\delta \int_a^b \left[\mathcal{L}(q, v)\mathrm{d}t + \sum_{i=1}^{m} \gamma_i(q) \circ \mathrm{d}W_i + \left\langle p, \frac{\mathrm{d}q}{\mathrm{d}t} - v \right\rangle \mathrm{d}t \right] + \int_a^b F(q, v) \cdot \delta q\, \mathrm{d}t = 0.$$

This principle provides a simple way to add dissipative effects into the drift which, e.g., appear in the standard Langevin equations for particles.

3.7 *Lagrangian reduction*

For background and exposition on Lagrangian reduction in the deterministic setting, the reader is referred to Marsden & Scheurle (1993) and Marsden & Ratiu (1999). Suppose that Q is a Lie group G with Lie algebra $\mathfrak{g}$. In this context, one can define a 'left-trivialized Lagrangian' by using the left action of G to left trivialize $\mathcal{L}$. One does this by transforming a point $(g(t), v(t)) \in TG$ to $(g(t), \xi(t)) \in G \times \mathfrak{g}$ via the relation between the velocity at $g(t) \in G$ and the 'body angular velocity' at $e \in G$ given by $\xi(t) = g(t)^{-1} v(t) \in \mathfrak{g}$. Denote by $l\colon G \times \mathfrak{g} \to \mathbb{R}$ the deterministic left-trivialized Lagrangian defined as $l(g(t), \xi(t)) = \mathcal{L}(g(t), g(t)\xi(t))$. The variational principle associated with l is the 'left-trivialized HP principle', which can be written as

$$\delta \int_a^b \left[l(g, \xi)\mathrm{d}t + \sum_{i=1}^{m} \gamma_i(g) \circ \mathrm{d}W_i + \left\langle \mu, g^{-1}\frac{\mathrm{d}g}{\mathrm{d}t} - \xi \right\rangle \mathrm{d}t \right] = 0.$$

In this principle, the Lagrange multiplier $\mu \in \mathfrak{g}^*$ is the body angular momentum. For more details on the geometry of this principle in the deterministic setting, the reader is referred to Bou-Rabee & Marsden (2008) and Bou-Rabee (2007). The resulting equations are obtained by taking arbitrary variations with fixed end points on g. For a function $U\colon G \to \mathbb{R}$, define its 'left-trivialized differential',

$U_g \in \mathfrak{g}^*$, as

$$U_g \cdot \eta := \left\langle \frac{\partial U}{\partial g}, TL_g \eta \right\rangle, \quad \eta \in \mathfrak{g}. \tag{3.5}$$

Using this definition, one can write the 'stochastic left-trivialized HP equations'

$$\frac{\mathrm{d}g}{\mathrm{d}t} = g\xi, \tag{3.6}$$

$$\mathrm{d}\mu = l_g \, \mathrm{d}t + \sum_{i=1}^{m} (\gamma_i)_g \circ \mathrm{d}W_i(\omega, t), \tag{3.7}$$

$$\mu = \frac{\partial l}{\partial \xi}. \tag{3.8}$$

By eliminating ξ using (3.8), one obtains an SDE on $G \times \mathfrak{g}^*$. The kinematic constraint in this context is referred to as the 'reconstruction equation'. We summarize this section with the following theorem.

THEOREM 3.9 (Stochastic left-trivialized HP principle) Consider a mechanical system on a Lie group G with left-trivialized Lagrangian $l\colon G \times \mathfrak{g} \to \mathbb{R}$ and deterministic functions $\gamma_i\colon G \to \mathbb{R}$ for $i = 1, \ldots, m$. Let s denote the left-trivialized action given by

$$s = \int_a^b \left[l(g, \xi)\mathrm{d}t + \sum_{i=1}^{m} \gamma_i(g) \circ \mathrm{d}W_i + \left\langle \mu, g^{-1}\frac{\mathrm{d}g}{\mathrm{d}t} - \xi \right\rangle \mathrm{d}t \right].$$

Almost surely, the stochastic HP principle on a Lie group (cf. Theorem 3.3) is equivalent to the stochastic left-trivialized HP principle:

$$\delta s = 0,$$

where the curves

$$g(t) \in G, \quad \xi(t) \in \mathfrak{g}, \quad \mu(t) \in \mathfrak{g}^*, \quad t \in [a, b],$$

can be varied arbitrarily with $\delta g(a) = \delta g(b) = 0$. Almost surely, a curve is a critical point of the left-trivialized action and only if it satisfies the left-trivialized HP equations given by (3.6–3.8).

4. Stochastic variational integrators

The standard approach of deriving variational integrators is extended to the stochastic context in this section; see, e.g. Marsden & West (2001). The cornerstone of variational integration theory is the discrete Lagrangian. In this section, we develop and analyse integrators from an abstract discrete Lagrangian that takes values on the configuration space squared. In the subsequent sections, discrete Lagrangians will be specified and the associated time integrators analysed from the HP viewpoint. Let $[a, b]$ and N be given and define the fixed step size $h = (b - a)/N$ and $t_k = hk + a$, $k = 0, \ldots, N$.

DEFINITION 4.1 Consider a mechanical system with given discrete Lagrangian $\mathcal{L}_\mathrm{d}\colon Q \times Q \to \mathbb{R}$. Let $\theta_t\colon \Omega \to \Omega$, $t \in [a, b]$, denote the base flow on the probability space (cf. Definition 2.1). Given $\omega \in \Omega$ let the approximant to the Stratonovich integrals will be denoted by

$$B_\mathrm{d}(t_k, q_k, q_{k+1}, \omega)h \approx \sum_{i=1}^{m} \int_{t_k}^{t_{k+1}} \gamma_i(q(t)) \circ \mathrm{d}W_i(t, \theta_{t_k+t}(\omega)).$$

The associated 'stochastic discrete Lagrangian' $L_{\mathrm{d}}\colon \mathbb{R} \times \Omega \times Q \times Q \to \mathbb{R}$ is defined as

$$L_{\mathrm{d}}(t_k, \omega, q_k, q_{k+1}) = \mathcal{L}(q_k, q_{k+1}) + B_{\mathrm{d}}(t_k, q_k, q_{k+1}, \omega).$$

Fixing the interval $[a, b]$ and $q_a, q_b \in Q$, define the 'discrete path space' as

$$\mathcal{C}_{\mathrm{d}}(Q) = \{q_{\mathrm{d}}\colon \{t_k\}_{k=0}^{N} \to Q | q_{\mathrm{d}}(a) = q_a, q_{\mathrm{d}}(b) = q_b\}.$$

Let $\mathfrak{G}_{\mathrm{d}}\colon \Omega \times \mathcal{C}_{\mathrm{d}}(Q) \to \mathbb{R}$ denote the 'stochastic action sum':

$$\mathfrak{G}_{\mathrm{d}}(\omega, q_{\mathrm{d}}) = \sum_{k=0}^{N-1} L_{\mathrm{d}}(t_k, \omega, q_k, q_{k+1})h.$$

The 'discrete stochastic Hamilton's principle' states that the path that the mechanical system takes in $\mathcal{C}_{\mathrm{d}}$ is one that extremizes $\mathfrak{G}_{\mathrm{d}}(\omega, \cdot)$ subject to fixed end point conditions $q_0 = q(a)$ and $q_N = q(b)$. By discrete integration by parts (reindexing),

$$\begin{aligned} \mathrm{d}\mathfrak{G}_{\mathrm{d}}(\omega, q_{\mathrm{d}}) \cdot \{\delta q_k\} &= \sum_{k=1}^{N-1} (D_3 L_{\mathrm{d}}(t_k, \omega, q_k, q_{k+1}) + D_4 L_{\mathrm{d}}(t_{k-1}, \omega, q_{k-1}, q_k)) \cdot \delta q_k \\ &\quad + D_3 L_{\mathrm{d}}(t_0, \omega, q_0, q_1) \cdot \delta q_0 + D_4 L_{\mathrm{d}}(t_{N-1}, \omega, q_{N-1}, q_N) \cdot \delta q_N. \end{aligned}$$

Using the end point conditions $\delta q_0 = \delta q_N = 0$, one obtains

$$\mathrm{d}\mathfrak{G}_{\mathrm{d}}(\omega, q_{\mathrm{d}}) \cdot \{\delta q_k\} = \sum_{k=1}^{N-1} (D_3 L_{\mathrm{d}}(t_k, \omega, q_k, q_{k+1}) + D_4 L_{\mathrm{d}}(t_{k-1}, \omega, q_{k-1}, q_k)) \cdot \delta q_k.$$

Stationarity of this action sum implies the following 'stochastic discrete Euler–Lagrange' equations:

$$D_3 L_{\mathrm{d}}(t_k, \omega, q_k, q_{k+1}) + D_4 L_{\mathrm{d}}(t_{k-1}, \omega, q_{k-1}, q_k) = 0,$$

for $k = 1, \ldots, N-1$. The resulting update scheme is not self-starting. To initialize the method, one needs to provide $(q_0, q_1) \in Q \times Q$ as opposed to a point $(q_0, v_0) \in TQ$.

4.1 *Symplecticity*

As in the continuous theory, symplecticity follows from restricting $\mathfrak{G}_{\mathrm{d}}(\omega, \cdot)$ to pathwise unique solutions of the stochastic discrete Euler–Lagrange equations, $\hat{\mathfrak{G}}_{\mathrm{d}}$. Since pathwise unique solutions can be parameterized by initial conditions, we regard the restricted action as $\hat{\mathfrak{G}}_{\mathrm{d}}\colon \Omega \times Q \times Q \to \mathbb{R}$. Taking its first variation gives

$$\begin{aligned} \mathrm{d}\hat{\mathfrak{G}}_{\mathrm{d}}(\omega, q_0, q_1) \cdot (\delta q_0, \delta q_1) &= \sum_{k=1}^{N-1} (D_3 L_{\mathrm{d}}(t_k, \omega, q_k, q_{k+1}) + D_4 L_{\mathrm{d}}(t_{k-1}, \omega, q_{k-1}, q_k)) \cdot \delta q_k \\ &\quad + D_3 L_{\mathrm{d}}(t_0, \omega, q_0, q_1) \cdot \delta q_0 + D_4 L_{\mathrm{d}}(t_{N-1}, \omega, q_{N-1}, q_N) \cdot \delta q_N. \end{aligned}$$

Because of the restriction to solution space, the sum vanishes and the boundary terms remain:

$$\mathrm{d}\hat{\mathfrak{G}}_{\mathrm{d}}(\omega, q_0, q_1) \cdot (\delta q_0, \delta q_1) = D_3 L_{\mathrm{d}}(t_0, \omega, q_0, q_1) \cdot \delta q_0 + D_4 L_{\mathrm{d}}(t_{N-1}, \omega, q_{N-1}, q_N) \cdot \delta q_N.$$

These boundary terms define left and right one forms as follows:

$$\Theta^+(t_k, \omega, q_k, q_{k+1}) \cdot (\delta q_k, \delta q_{k+1}) = D_4 L_\mathrm{d}(t_k, \omega, q_k, q_{k+1}) \cdot \delta q_{k+1},$$

$$\Theta^-(t_k, \omega, q_k, q_{k+1}) \cdot (\delta q_k, \delta q_{k+1}) = D_3 L_\mathrm{d}(t_k, \omega, q_k, q_{k+1}) \cdot \delta q_k,$$

which from $\mathrm{d}^2 L_\mathrm{d} = 0$ satisfy

$$\mathrm{d}\Theta^+ = -\mathrm{d}\Theta^- = \Omega.$$

Applying the second exterior derivative to $\hat{\mathfrak{G}}_\mathrm{d}$ implies that

$$\Omega(t_{N-1}, \omega, q_{N-1}, q_N)(\delta q^1_{N-1}, \delta q^1_N)(\delta q^2_{N-1}, \delta q^2_N) = \Omega(t_0, \omega, q_0, q_1)(\delta q^1_0, \delta q^1_1)(\delta q^2_0, \delta q^2_1),$$

since $\mathrm{d}^2 \hat{\mathfrak{G}}_\mathrm{d} = 0$. Hence, the discrete flow almost surely preserves the symplectic form Ω.

4.2 *Discrete momentum map*

Consider the left action of a Lie Group G on Q. If the stochastic discrete Lagrangian is infinitesimally symmetric, then the associated momentum map is preserved. A sufficient condition for this is that the discrete Lagrangian is invariant with respect to the left action of G. The proof is sketched out here.

Let the action on the discrete configuration manifold be denoted by $\Phi_{Q\times Q}\colon G \times Q \times Q \to Q \times Q$ and defined by

$$\Phi_{Q\times Q}(g, q_1, q_2) = (\Phi(g, q_1), \Phi(g, q_2)).$$

The associated infinitesimal generator is denoted by $\xi_{Q\times Q}\colon Q \times Q \to T(Q \times Q)$, and by definition

$$\xi_{Q\times Q}(q_1, q_2) = \frac{\mathrm{d}}{\mathrm{d}s} \Phi_{Q\times Q}(\exp(s\xi), q_1, q_2)|_{s=0}.$$

Assume that L_d is infinitesimally symmetric; i.e.

$$\mathrm{d}L_\mathrm{d} \cdot \xi_{Q\times Q} = \Theta^+ \cdot \xi_{Q\times Q} + \Theta^- \cdot \xi_{Q\times Q} = 0.$$

By this condition, the left and right discrete momentum maps, $J^+, J^-\colon Q \times Q \to \mathfrak{g}^*$, namely

$$J^+ \cdot \xi = \Theta^+ \cdot \xi_{Q\times Q},$$

$$J^- \cdot \xi = -\Theta^- \cdot \xi_{Q\times Q},$$

are equal; i.e. $J^+ = J^- = J$. Consider the restricted action sum and compute its differential in the direction of the infinitesimal generator to obtain

$$\begin{aligned}\mathrm{d}\hat{\mathfrak{G}}_\mathrm{d}(\omega, q_0, q_1) \cdot \xi_{Q\times Q}(q_0, q_1) &= \Theta^-(t_0, \omega, q_0, q_1) \cdot \xi_{Q\times Q}(q_0, q_1) \\ &\quad + \Theta^+(t_{N-1}, \omega, q_{N-1}, q_N) \cdot \xi_{Q\times Q}(q_{N-1}, q_N),\end{aligned}$$

which can be rewritten in terms of the momentum maps as

$$\mathrm{d}\hat{\mathfrak{G}}_\mathrm{d}(\omega, q_0, q_1) \cdot \xi_{Q\times Q}(q_0, q_1) = -J^-(t_0, \omega, q_0, q_1) \cdot \xi + J^+(t_{N-1}, \omega, q_{N-1}, q_N) \cdot \xi = 0.$$

Since the left and right momentum maps evaluated at the same point are equal, the momentum map J is preserved under the discrete flow.

5. Stochastic VE integrator

In the deterministic setting, the HP context provides a practical way to design discrete Lagrangians and obtain one-step methods on TQ or T^*Q as pointed out in Bou-Rabee (2007). In this section, we examine VE methods extended to the stochastic context following the continuous stochastic HP theory laid out in Section 2.

5.1 *Stochastic VE on $\mathbb{R}^n$*

To discretize the stochastic HP action integral, one needs to replace the continuous Lagrangian, stochastic integral and kinematic constraint by discrete approximants. We begin by introducing a first-order discretization of the kinematic constraint in (3.1). Let $[a, b]$ and N be given and define the fixed step size $h = (b - a)/N$ and $t_k = hk$, $k = 0, \dots, N$.

A discretization of the kinematic constraint can be obtained by introducing a discrete sequence $\{q_k\}_{k=0}^N$ taking values in Q and a finite-difference map $\varphi\colon Q \times Q \to TQ$. For example, if Q is a vector space the following backward difference map can be introduced:

$$\varphi(q_k, q_{k+1}) = \left(q_{k+1}, \frac{q_{k+1} - q_k}{h}\right).$$

Let $B_i^k \sim \mathcal{N}(0, h)$ be normally distributed random variables for $i = 1, \dots, m$ and $k = 0, \dots, N - 1$. In terms of the discretization of the kinematic constraint, the corresponding discrete HP action sum takes the following simple form:

$$\mathfrak{G}_{\mathrm{d}}^e = \sum_{k=0}^{N-1} \left[\mathcal{L}(q_k, v_k)h + \sum_{i=1}^{m} \gamma_i(q_k) B_i^k + \left\langle p_{k+1}, \frac{(q_{k+1} - q_k)}{h} - v_{k+1} \right\rangle h \right].$$

The stochastic discrete HP equations are given by

$$\begin{aligned} q_{k+1} &= q_k + h v_{k+1}, \\ p_{k+1} &= p_k + h \frac{\partial \mathcal{L}}{\partial q}(q_k, v_k) + \sum_{i=1}^{m} \frac{\partial \gamma_i}{\partial q}(q_k) B_i^k, \\ p_k &= \frac{\partial \mathcal{L}}{\partial v}(q_k, v_k). \end{aligned}$$

5.2 *Stochastic VE on Lie groups*

In the context of Lie groups, the reconstruction equation is discretized using canonical coordinates of the first kind, $\tau\colon \mathfrak{g} \to G$, as explained in Bou-Rabee (2007) and Bou-Rabee & Marsden (2008). As in the vector space case, we define a finite-difference map $\varphi\colon G \times G \to G \times \mathfrak{g}$ that provides a first-order approximant to the reconstruction equation:

$$\varphi(g_k, g_{k+1}) = \left(g_{k+1}, \frac{\tau^{-1}(g_k^{-1} g_{k+1})}{h}\right) \in G \times \mathfrak{g}.$$

A first-order approximant to the stochastic left-trivialized action integral is given by

$$s_{\mathrm{d}} = \sum_{k=0}^{N-1} \left[l(g_k, \xi_k)h + \sum_{i=1}^{m} \gamma_i(g_k) B_i^k + \left\langle \mu_{k+1}, \frac{\tau^{-1}(g_k^{-1} g_{k+1})}{h} - \xi_{k+1} \right\rangle h \right]. \tag{5.1}$$

Let $\mathrm{d}\tau^{-1} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ denote the right-trivialized tangent of τ^{-1} as defined in Bou-Rabee & Marsden (2008). The stochastic left-trivialized discrete HP equations are (cf. Definition 3.5)

$$g_{k+1} = g_k \tau(h\xi_{k+1}),$$

$$\left(\mathrm{d}\tau^{-1}_{h\xi_{k+1}}\right)^* \mu_{k+1} = \left(\mathrm{d}\tau^{-1}_{-h\xi_k}\right)^* \mu_k + h l_g(g_k, \xi_k) + \sum_{i=1}^{m} (\gamma_i)_g(g_k) B_i^k,$$

$$\mu_k = \frac{\partial l}{\partial \xi}(g_k, \xi_k).$$

5.3 *Holonomic constraints and nonconservative effects*

Holonomic constraints can be added via discrete Lagrange multipliers and nonconservative effects via discrete impulses as described below in the Lie group context. Suppose that G is an n-manifold and that the mechanical system evolves on a submanifold $S \subset G$ defined as the zero-level set of $\varphi \colon G \to \mathbb{R}^k$, where $k < n$ and $S = \varphi^{-1}(0)$. Further, suppose that there exists a force field $F \colon G \to T^*G$. These effects are appended by using the following action principle:

$$\delta \sum_{k=0}^{N-1} \left[l(g_k, \xi_k)h + \sum_{i=1}^{m} \gamma_i(g_k) B_i^k + \left\langle \mu_{k+1}, \frac{\tau^{-1}(g_k^{-1} g_{k+1})}{h} - \xi_{k+1} \right\rangle h + \langle \lambda_k, \varphi(g_k) \rangle h \right]$$
$$+ \sum_{k=0}^{N-1} F(g_k) \cdot \delta g_k h = 0.$$

The algorithm that one obtains from this principle is the stochastic analogue of constrained symplectic Euler, and the numerical analysis of this method is discussed in Bou-Rabee & Owhadi (2008).

6. Langevin-type equations for multiple rigid bodies

6.1 *Continuous description*

Consider a mechanical system consisting of K rigid bodies interacting via a potential dependent on their positions and orientations. Let $(x_i(t), v_i(t), R_i(t), \omega_i(t)) \in \mathrm{TSE}(3)$ denote the translational position, translational velocity, rotational position and spatial angular velocity of body i, where $i = 1, \ldots, K$. Let m_i and $\mathbb{I}_i$ denote the mass and diagonal inertia tensor of body i. The left-trivialized Lagrangian for the system is given by

$$l(x_i, v_i, R_i, \omega_i) = \sum_{i=1}^{K} \frac{m_i}{2} v_i^{\mathrm{T}} v_i + \frac{1}{2} \omega_i^{\mathrm{T}} R_i \mathbb{I}_i R_i^{\mathrm{T}} \omega_i - U(x_i, R_i).$$

Note that $l(x_i, v_i, R_i, \omega_i)$ is shorthand notation for $l(x_1, v_1, R_1, \omega_1, \dots, x_K, v_K, R_K, \omega_K)$. We will use this shorthand notation elsewhere to simplify the expressions. The path that the stochastic mechanical system takes on the time interval $[a, b]$ is one that extremizes the HP action:

$$s = \int_a^b \left[l(x_i, v_i, R_i, \omega_i)\mathrm{d}t + \sum_{q=1}^m \gamma_q(x_i, R_i) \circ \mathrm{d}W_q + \left\langle p_i, \frac{\mathrm{d}x_i}{\mathrm{d}t} - v_i \right\rangle \mathrm{d}t + \left\langle \widehat{\pi}_i, \frac{\mathrm{d}R_i}{\mathrm{d}t} R_i^{-1} - \widehat{\omega}_i \right\rangle \mathrm{d}t \right],$$

for arbitrary variations with fixed end points $(x_i(a), R_i(a))$ and $(x_i(b), R_i(b))$. The corresponding SDEs of motion are given by

$$\begin{aligned}
\mathrm{d}x_i &= v_i\,\mathrm{d}t \quad \text{(reconstruction equation)},\\
\mathrm{d}p_i &= -U_{x_i}\,\mathrm{d}t + \sum_{q=1}^m (\gamma_q)_{x_i} \circ \mathrm{d}W_q(t, \omega) \quad \text{(stochastic EL equations)},\\
p_i &= m_i v_i \quad \text{(Legendre transform)},\\
\mathrm{d}R_i &= \widehat{\omega}_i R_i\,\mathrm{d}t \quad \text{(reconstruction equation)},\\
\mathrm{d}\pi_i &= -U_{R_i}\,\mathrm{d}t + \sum_{q=1}^m (\gamma_q)_{R_i} \circ \mathrm{d}W_q \quad \text{(stochastic LP equations)},\\
\pi_i &= R_i \mathbb{I}_i R_i^{\mathrm{T}} \omega_i \quad \text{(reduced Legendre transform)},
\end{aligned}$$

for $i = 1, \dots, K$. The terms U_{x_i} and U_{R_i} are defined in terms of the inner product on $\mathbb{R}^3$ as

$$U_{x_i}^{\mathrm{T}} y = \left\langle \frac{\partial U}{\partial x_i}, y \right\rangle = \partial_{x_i} U(x_i, R_i) \cdot y,$$

$$U_{R_i}^{\mathrm{T}} y = \left\langle \frac{\partial U}{\partial R_i} R_i^{\mathrm{T}}, \widehat{y} \right\rangle = \partial_{R_i} U(x_i, R_i) \cdot \widehat{y} R_i,$$

where $\partial_{R_i} U \colon \mathrm{SO}(3) \to T^*_{R_i}\mathrm{SO}(3)$ and $\partial_{x_i} U \colon \mathbb{R}^3 \to T^*_{x_i}\mathbb{R}^3$. Adding dissipation so that the Gibbs distribution is invariant under the stochastic process defined by the above SDE with dissipative drift yields Langevin-type equations for rigid-body systems (see, e.g. Bou-Rabee & Owhadi, 2007).

6.2 *Stochastic VE integrator*

For the discrete description, the VE integrator provided earlier is implemented. Let $B_q^k \sim \mathcal{N}(0, h)$ be normally distributed random variables for $q = 1, \dots, m$ and $k = 0, \dots, N-1$. The action sum is given by

$$\begin{aligned}
s_{\mathrm{d}} = \sum_{k=0}^{N-1} &\left[\ell(x_i^k, v_i^k, R_i^k, \omega_i^k)h + \left\langle p_i^{k+1}, \frac{(x_i^{k+1} - x_i^k)}{h} - v_i^{k+1} \right\rangle h \right]\\
&+ \left[\left\langle \widehat{\pi_i^{k+1}}, \tau^{-1} \frac{(R_i^{k+1}(R_i^k)^{\mathrm{T}})}{h} - \widehat{\omega_i^{k+1}} \right\rangle h \right] + \sum_{q=1}^m \gamma_q(x_i^k, R_i^k) B_q^k.
\end{aligned}$$

Stationarity of this action sum implies the following discrete scheme:

$$
\begin{aligned}
x_i^{k+1} &= x_i^k + h v_i^{k+1}, \\
p_i^{k+1} &= p_i^k - hU_{x_i}(x_i^k, R_i^k) + \sum_{q=1}^{m} (\gamma_q)_{x_i}(x_i^k, R_i^k) B_q^k, \\
p_i^k &= m v_i^k, \\
R_i^{k+1} &= \tau(\widehat{\omega_i^{k+1}} h) R_i^k, \\
\left(\mathrm{d}\tau^{-1}_{h\omega_i^{k+1}}\right)^* \pi_i^{k+1} &= \left(\mathrm{d}\tau^{-1}_{h\omega_i^{k}}\right)^* \pi_i^{k} - hU_{R_i}(x_i^k, R_i^k) + \sum_{q=1}^{m} (\gamma_q)_{R_i}(x_i^k, R_i^k) B_q^k, \\
\pi_i^k &= R_i^k \mathbb{I}_i (R_i^k)^{\mathrm{T}} \omega_i^k,
\end{aligned}
$$

for $i = 1, \ldots, K$. Assuming that the Legendre transforms are invertible, this integrator has the attractive property that the translational and rotational configuration updates, and the translational momentum update, are explicit. One has only to perform an implicit solution for the discrete Lie–Poisson part. Even that computation is straightforward since the torque due to the potential is only a function of the orientation and position at the previous time step.

Acknowledgements

We wish to thank Andreu Lazaro, Jerry Marsden and Juan-Pablo Ortega for stimulating discussions.

REFERENCES

ABRAHAM, R., MARSDEN, J. E. & RATIU, T. (2007) *Manifolds, Tensors, Analysis, and Applications*. New York: Springer.

BISMUT, J. M. (1981) *Mécanique aléatoire*. Berlin: Springer.

BOU-RABEE, N. (2007) Hamilton–Pontryagin integrators on Lie groups. *Ph.D. Thesis*, Pasadena, CA: California Institute of Technology.

BOU-RABEE, N. & OWHADI, H. (2007) Ballistic transport at uniform temperature (in press).

BOU-RABEE, N. & MARSDEN, J. E. (2008) Hamilton–Pontryagin integrators on lie groups part I: introduction and structure-preserving properties. *Found. Comput. Math.* (to appear).

BOU-RABEE, N. & OWHADI, H. (2008) Symplectic-momentum Euler–Maruyama integrator for stochastic Hamiltonian systems with constraints. *SIAM J. Numer. Anal.* (to appear).

CICCOTTI, G., LELIEVRE, T. & VANDEN-EIJNDEN, E. (2008) Projections of diffusions on submanifolds: application to mean force computation. *Commun. Pure Appl. Math.*, **61**, 0001–0039.

DE VOGELAERE, R. (1956) Methods of integration which preserve the contact transformation property of the Hamiltonian equations. *Technical Report*, Notre Dama, Ind: Department of Mathematics, University of Notre Dame.

ELWORTHY, K. D. (1982) *Stochastic Differential Equations on Manifolds*. Cambridge, UK: Cambridge University Press.

EMERY, M. (1989) *Stochastic Calculus in Manifolds*. Berlin: Springer.

FENG, K. (1986) Difference schemes for Hamiltonian formalism and symplectic geometry. *J. Comput. Math.*, **4**, 279–289.

FEYNMANN, R. P. & HIBBS, A. R. (1981) *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill.

HAIRER, E., LUBICH, C. & WANNER, G. (2006) *Geometric Numerical Integration*. Springer Series in Computational Mathematics, vol. 31. Berlin: Springer.

IKEDA, N. & WATANABE, S. (1989) *Stochastic Differential Equations and Diffusion Processes*. Amsterdam: North-Holland.

KUNITA, H. (1990) *Stochastic Flows and Stochastic Differential Equations*. Berlin: Cambridge University Press.

LAZARO-CAMI, J. A. & ORTEGA, J. P. (2007a) Reduction and reconstruction of symmetric stochastic differential equations. *Rep. Math. Phys.* (in press).

LAZARO-CAMI, J. A. & ORTEGA, J. P. (2007b) Stochastic Hamiltonian dynamical systems. *Rep. Math. Phys.* (in press).

LEW, A., MARSDEN, J. E., ORTIZ, M. & WEST, M. (2004) Variational time integrators. *Int. J. Numer. Methods Eng.*, **60**, 153–212.

LIAO, M. (1997) Random motion of a rigid body. *J. Theor. Probab.*, **10**, 201–211.

LIAO, M. & WANG, L. (2005) Motion of a rigid body under random perturbation. *Electron. Commun. Probab.*, **10**, 235–243.

MACKAY, R. (1992) *Some Aspects of the Dynamics of Hamiltonian Systems*. Oxford, UK: Clarendon Press, pp. 137–193.

MALHAM, S. J. A. & WIESE, A. (2007) Stochastic Lie group integrators. *SIAM J. Sci. Comput.*, **30**, 597–617.

MARSDEN, J. E. & SCHEURLE, J. (1993) The reduced Euler–Lagrange equations. *Fields Inst. Commun.*, **1**, 139–164.

MARSDEN, J. E., PEKARSKY, S. & SHKOLLER, S. (1998) Discrete Euler–Poincaré and Lie–Poisson equations. *Nonlinearity*, **12**, 1647–1662.

MARSDEN, J. & RATIU, T. (1999) *Introduction to Mechanics and Symmetry*. Springer Texts in Applied Mathematics. New York: Springer.

MARSDEN, J. E. & WEST, M. (2001) Discrete mechanics and variational integrators. *Acta Numer.*, **10**, 357–514.

MILSTEIN, G. N., REPIN, YU. M. & TRETYAKOV, M. V. (2002) Symplectic methods for Hamiltonian systems with additive noise. *SIAM J. Numer. Anal.*, **39**, 1–9.

MILSTEIN, G. N., REPIN, YU. M. & TRETYAKOV, M. V. (2003) Symplectic methods for stochastic systems preserving symplectic structure. *SIAM J. Numer. Anal.*, **40**, 1–9.

MILSTEIN, G. N. & TRETYAKOV, M. V. (2003) Quasi-symplectic methods for Langevin-type equations. *IMA J. Numer. Anal.*, **23**, 593–626.

MILSTEIN, G. N. & TRETYAKOV, M. V. (2004) *Stochastic Numerics for Mathematical Physics*. Berlin: Springer.

MOSER, J. & VESELOV, A. P. (1991) Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Commun. Math. Phys.*, **139**, 217–243.

NELSON, E. (1985) *Quantum Fluctuations*. Princeton, NJ: Princeton University Press.

RUTH, R. D. (1983) A canonical integration technique. *IEEE Trans. Nucl. Sci.*, **30**, 2669–2671.

TALAY, D. (1995) Simulation and numerical analysis of stochastic differential systems: a review. *Probabilistic Methods in Applied Physics* (P. Krèe & W. Wedig eds), vol. 451. Berlin: Springer, pp. 54–96.

TALAY, D. (2002) Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Relat. Fields*, **8**, 1–36.

THIEULLEN, M. & ZAMBRINI, J. C. (2008) Probability and quantum symmetries. I. The theorem of Noether in Schrödinger's Euclidean quantum mechanics. *Ann. Inst. H. Poincaré Phys. Théor.*, **67**, 297–338.

VANDEN-EIJNDEN, E. & CICCOTTI, G. (2006) Second-order integrators for Langevin equations with holonomic constraints. *Chem. Phys. Lett.*, **429**, 310–316.

VESELOV, M. (1988) Integrable discrete-time systems and difference operators. *Funct. Anal. Appl.*, **22**, 83–94.

WENDLANDT, J. M. & MARSDEN, J. E. (1997) Mechanical integrators derived from a discrete variational principle. *Physica D*, **106**, 223–246.

YOSHIMURA, H. & MARSDEN, J. E. (2006a) Dirac structures and Lagrangian mechanics part I: implicit Lagrangian systems. *J. Geom. Phys.*, **57**, 133–156.

YOSHIMURA, H. & MARSDEN, J. E. (2006b) Dirac structures and Lagrangian mechanics part II: variational structures. *J. Geom. Phys.*, **57**, 209–250.