Stochastic Geometric Mechanics and Symmetry

Archishman Saha

Thesis submitted to the University of Ottawa in partial fulfillment of the requirements for the degree of
Doctorate in Philosophy Mathematics and Statistics[*]

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

---

[*] The Ph.D. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

Abstract

We explore the interaction between geometry, symmetry, and randomness in stochastically perturbed mechanical systems, in Hamiltonian as well as Lagrangian (variational) formulations.

On the Hamiltonian side, we discuss a stochastic Kepler problem with a noisy angular momentum vector. We show that the radial distance and speed evolve deterministically. This allows us to regularize collisional singularities in a procedure similar to Moser's regularization. We generalize this to stochastic collective Hamiltonian systems. We show that the solutions to these systems are given by the action of a Lie group valued semimartingale on the deterministic solution and consequently, along directions transverse to the group orbits, these systems retain the same dynamics as their deterministic counterpart. Furthermore, we show that these systems can be described by coupling a deterministic Hamiltonian system to a stochastic one. We also identify conditions under which the momentum map of a stochastic Hamiltonian system evolves as a martingale on a reductive coadjoint orbit.

On the variational side, we construct fixed endpoint, local, and adapted variations of semimartingales in manifolds. These variations are used to prove a stochastic version of the Fundamental Lemma of the Calculus of Variations, which is subsequently applied to studying the stochastic Hamilton-Pontryagin principle. We also provide a stochastic analogue of Noether's theorem. We treat the corresponding global form of the stochastic Hamilton-Pontryagin principle via a novel approach to global variational principles by using Stratonovich operators. We describe the reduction by symmetry of the stochastic Hamilton-Pontryagin principle. Moreover, we also discuss stochastic collective dynamics from the variational point of view. Similar to the Hamiltonian case, we show that the critical point of the stochastic action is given by the action of a Lie group valued semimartingale on the critical point of the deterministic action. We also provide a description of coupling to a Lie group on the variational side.

Additionally, we extend the theory of reduction of Stratonovich differential equations, which arises in stochastic Hamiltonian systems, to stochastic differential equations given by Schwartz operators.

Dedications

This thesis is dedicated to all of my grandparents.

Acknowledgements

It is difficult to adequately express the immense gratitude I feel for my supervisors, Prof. Cristina Stoica and Prof. Tanya Schmah. Their exceptional guidance, insight, and intellectual depth have made a profound impact in my thesis and my journey as a Ph.D. student. They have also been an unfailing source of patience and encouragement, always ready to help me navigate any hurdle. I am deeply thankful for their genuine warmth, care, and positivity, which have extended well beyond the office and made milestones truly special. Without their unwavering support and mentorship, this work would simply not exist.

I am also very grateful to Prof. Ana Bela Cruzeiro (Instituto Superior Técnico, Lisbon), Prof. Juan Pablo Ortega (Nanyang Technological University, Singapore), Prof. Abbas Momeni (Carleton University), and Prof. Raluca Balan (University of Ottawa) for their careful reading of this work, their thoughtful questions, and their constructive feedback, which significantly improved both the clarity and depth of the final manuscript.

I am thankful to Prof. Cruzeiro as well as Prof. Jean-Claude Zambrini (University of Lisbon) for many insightful discussions and suggestions during my two-week visit to Lisbon in June 2024. I also owe a debt of gratitude to Dr. Erwin Luesink (University of Amsterdam) for many stimulating mathematical discussions during his two-week visit to the University of Ottawa in October 2025.

My thanks extend to the instructors and faculty members at both the University of Ottawa and Carleton University for broadening my mathematical horizons through the courses I have taken during my Ph.D. I am equally thankful to the administrative and office staff at the Department of Mathematics and Statistics at the University of Ottawa, as well as the library facilities and university administration, for their seamless support and assistance in making sure the journey was smooth.

Looking back further, I would like to thank the instructors at my previous university (Visva-Bharati, India) who laid the foundation for my academic career. I am also deeply grateful to my Master's supervisor, Prof. Partha Guha (Khalifa University, Abu Dhabi), for his continued encouragement, mentorship, and discussions over the past few years.

On a personal note, my deepest love and gratitude belong to my parents. Long before I entered graduate school, they encouraged me to master new subjects by teaching the concepts to them, patiently critiquing my explanations. This unique approach instilled an internal "student-teacher" dialogue that I still rely on today to refine my thinking. I also strive to embody my mother's guiding principles of discipline, determination, and dedication. Without their profound sacrifices, hard work, and unconditional care, this journey would never have been possible.

I am equally grateful to my extended family, including my grandparents, uncles, aunts, and cousins—who have been a constant wellspring of support and motivation throughout this journey. I must also send a very special thank you to my four-legged brother, Michu; his faithful companionship has been a source of immense joy and a comforting presence through it all.

The journey through graduate school would have been incomplete without an amazing circle of friends. I'm grateful to all of my friends from my previous university, especially my math colleagues, Shivam and Sameer. I have greatly benefited from many discussions with them. I am also thankful to Shivangi for her constant support and encouragement. At the University of Ottawa, my sincere thanks go to Utkarsh, Mihir, Prangya, Daniel, Ekta, Masoomeh, and Mico. I am especially grateful to my roommate, Utkarsh, for being an incredibly supportive pillar throughout this journey, and to Mihir, for countless refreshing discussions over coffee.

I am also immensely grateful to Véronique and her family, especially Fréderic, Julie, Andy, Claude, and Tin. Their immense support, boundless encouragement, and warmth made a world of difference to me during this intensive process.

Finally, I cherish the memories of many enjoyable and relaxing evenings spent with Utkarsh, Véronique, Mihir, as well as Sidhesh and Steph, which provided a much-needed escape from research and the perfect balance to the demands of graduate life.

Contents

Chapter 1

Introduction

> *Opposition brings concord. Out of discord comes the fairest harmony.*
>
> Heraclitus, translated by Philip Ellis Wheelwright, *Heraclitus* (Princeton University Press, 1959)

Geometric mechanics studies symmetries and dynamics of mechanical systems using differential geometric techniques. Not only does this facilitate the development of mechanics in a coordinate-free language, but it also provides a unifying framework for the study of many well-known mechanical systems. For instance, Euler's equation for the free rigid body and the Euler equations for an ideal fluid can both be viewed as special cases of Euler–Poincaré equations on Lie groups. Lie group actions on manifolds provide an effective setting for studying symmetries, allowing us to reduce the dimension of the phase space of a mechanical system. From a numerical point of view, variational integrators and symplectic integrators provide a way of numerically integrating mechanical systems such that the first integrals remain conserved.

In stochastic geometric mechanics, the dynamic setting changes from deterministic, smooth trajectories, to generally rough, non-differentiable paths. This is seemingly at odds with the idea of "differential" geometry, with the immediate obstacle being the loss of classical notion of the velocity of a curve.

To our knowledge, there are two ways of overcoming this in the literature. The first uses forward and backward derivatives of stochastic processes. This approach is used in Nelson [73] to describe a kinematical equation for Brownian motion that closely resembles Newton's second law. A derivation of the Schrödinger equation from Brownian motion can be found in Nelson [72]. Yasue [92] shows that the kinematical equations of Nelson can be described from a variational point of view. Applications of this approach to the (deterministic) Navier-Stokes equation can be found in Nakagomi, Yasue, and Zambrini [71], Arnaudon and Cruzeiro [9, 8] and Cipriano and Cruzeiro [21]. Arnaudon, Chen, and Cruzeiro [7] carry out Euler-Poincaré reduction in this framework and semi-direct product reduction has been considered in Chen, Cruzeiro, and Ratiu [20]. A Legendre transform as well as the Hamiltonian counterpart of this formulation can be found in Huang and Zambrini [44].

The second approach involves working at the level of vector fields, rather than velocities. The rationale is that vector fields are geometric objects, and therefore symmetry and reduction methods apply to them. In the stochastic setting, we consider a deterministic vector field X as well as multiple vector fields corresponding to noise perturbations. Using all of these vector fields we describe a Stratonovich stochastic differential equation, so that the change of variables formula can be applied.

This approach is well-suited for studying stochastic perturbations of deterministic mechanical systems. In the real world, mechanical systems and measurements performed on them are prone to noise. Mechanical systems perturbed by external noise were studied from the Hamiltonian perspective by Bismut [11] in Euclidean spaces. This was extended to Poisson manifolds by Lázaro-Camí and Ortega [55]. Reduction and reconstruction, as well as skew-product decomposition of stochastic differential equations with symmetry were studied by Lázaro-Camí and Ortega [54], wherein the deterministic theories of Poisson and symplectic reduction of Hamiltonian systems are also extended to their stochastic counterparts.

Stochastic variational principles in this perturbative setting have been studied in [55] using a stochastic version of Hamilton's principle in phase space. In Holm [43], a stochastic Clebsch variational formulation is introduced, which is now known as Stochastic Advection by Lie Transport (SALT). The SALT framework has led to several developments in data-driven models of uncertainty in geophysical fluid dynamics and oceanography. See, for instance, [43], Cotter, Gottwald and Holm [24], Gay-Balmaz and Holm [39], Cotter et al. [23], Drivas, Holm and Leahy [34] and Street and Crisan [88]. Finite-dimensional applications have been considered in Arnaudon, de Castro and Holm [3, 4] and Cruzeiro, Holm and Ratiu [29].

The approach to stochastic geometric mechanics adopted in this thesis is the second one. In this context we study symmetries in mechanical systems perturbed by external noise from both the Hamiltonian and variational principle perspective.

This thesis is divided into three parts. Part I consists of preliminaries, Part II studies stochastic Hamiltonian systems, and Part III studies stochastic variational principles.

Part I is divided into two chapters. The first chapter is a brief introduction to deterministic geometric mechanics. The topics we review are Hamiltonian vector fields, momentum maps, Hamiltonian reduction and reconstruction, variational principles, and Lagrangian reduction.

The second chapter introduces stochastic calculus in manifolds using the language of second order differential geometry. The use of second order differential geometry in stochastic calculus is due to the work Schwartz [83] and Meyer [67, 66], and is motivated by the second order nature of the Itô change of variables formula. Much of the exposition in this chapter closely follows that of Emery [36]. In Section 3.5, we provide results on reduction and reconstruction of Stratonovich differential equations due to Lázaro-Camí and Ortega [54].

We also extend their results on reduction of Stratonovich equations to reduction of more general stochastic differential equations given by Schwartz operators.

Part II addresses stochastic Hamiltonian systems, and is organized into five chapters. In Chapter 4 we present stochastic Hamiltonian systems and their symmetries, as well as associated reduction theorems, following [55, 54].

In Chapter 5 we introduce a stochastic perturbation of the deterministic Kepler problem that affects the angular momentum vector. This perturbation destroys the symmetries of the Kepler problem, and therefore stochastic symmetry reduction methods do not apply. However, along the radial direction the stochastic problem has the exact same evolution as its unperturbed counterpart. This allows us to regularize collisional solutions in the stochastic Kepler problem by adapting a regularization method due to Moser [70].

The next three chapters are all motivated by the stochastic Kepler problem. In Chapter 6, we generalize the setup of the stochastic Kepler problem to stochastic collective Hamiltonian systems. These systems describe a deterministic Hamiltonian system, invariant under the action of a Lie group G, whose associated momentum map is subjected to stochastic perturbations. The stochastic systems are generally not invariant under the action of the symmetry group and hence, symmetry reduction methods for stochastic differential equations cannot be applied in this setting. However, along the 'symmetry direction' the evolution of the stochastic system coincides with the deterministic evolution. More precisely, we show that the solutions of the stochastic problem can be expressed as the deterministic solution with an additional G-valued stochastic phase. This allows us to prove that the projection of a solution of the stochastic system onto the reduced space evolves deterministically, enabling full reconstruction of the stochastic dynamics from reduced solutions of the deterministic problem. This is done in two steps - the first step is deterministic, and it involves lifting a solution of the deterministic reduced problem to a solution of the unreduced deterministic problem. The second step is stochastic, and it corresponds to adding a stochastic phase to the deterministic solution of the unreduced problem.

In Chapter 7 we couple a G-invariant deterministic Hamiltonian system defined on a phase space P with a stochastic Hamiltonian system defined on T^*G. We show that this coupling gives rise to a stochastic collective Hamiltonian system on P. Moreover, the stochastic phase and the momentum map of the stochastic collective Hamiltonian system correspond to the dynamics of the stochastic Hamiltonian system on T^*G.

In some cases, the evolution of the momentum map is actually a Brownian motion on a coadjoint orbit. This implies that, although the momentum map is not conserved, its drift is a conserved quantity. As a generalization of this idea, in Chapter 8 we look for conditions when the momentum map of a stochastic perturbation of a Hamiltonian system is a martingale on a coadjoint orbit. We find necessary conditions for the momentum map to be a martingale on reductive coadjoint orbits equipped with the canonical affine connection. This also allows us to generalize the stochastic Noether's theorem due to Lázaro-Camí and Ortega [55].

Part III consists of three chapters and focusses on stochastic variational principles. In Chapter 9 we provide a solution to the 'adaptedness problem' in stochastic variational principles. Roughly speaking, this problem asks whether, given a (continuous) semimartingale Γ in a manifold M, a chart $U \subseteq M$ and a fixed time $T > 0$, it is possible to create a deformation $\epsilon \mapsto \Gamma_\epsilon$ of Γ so that the associated variation $\delta\Gamma$ vanishes everywhere outside U as well as at times $t = 0$ and $t = T$. Here we require that Γ_ϵ and $\delta\Gamma$ are semimartingales on M and TM respectively, so that the stochastic action integrals and their variations are well-defined. There are two main difficulties here. The first is to make sure that the deformation and variations constructed remain adapted to the underlying filtration. The second one stems from the fact that the entry and exit times for a chart are usually stopping times, and therefore are random. Thus, we need to ensure that the variations vanish not only at deterministic time $t = 0$ and $t = T$, but also at random times. Both of these difficulties are resolved in this chapter.

Having constructed variations of semimartingales, we prove a stochastic analogue of the Fundamental Lemma of Calculus of Variations. We use this to describe the local form of the stochastic Hamilton-Pontryagin principle introduced by Bou-Rabee and Owhadi [13] and recently studied by Street and Takao [89]. We also prove the variational principle side of the stochastic Noether's theorem. For the corresponding global formulation of the stochastic Hamilton-Pontryagin action principle, we show that the problem may be reformulated in terms of Stratonovich operators. These are deterministic objects which serve as generalizations of vector fields and are used to define Stratonovich stochastic differential equations. This allows us to recast the stochastic variational problem as a deterministic variational problem, which can be tackled by using the (deterministic) Fundamental Lemma of the Calculus of Variations.

Chapter 10 studies the reduction by symmetry of the stochastic Hamilton-Pontryagin principle. We assume that the stochastic action integral is invariant under the free and proper action of a Lie group G. Then, by choosing a principal connection, we show that it drops to a stochastic action integral on a reduced space. The critical points of this reduced action are described by two sets of equations in the reduced space. One set of equations take the form of stochastic Euler-Lagrange equations with a curvature-induced forcing term. These equations correspond to the evolution in the horizontal direction determined by the connection. The other set corresponds to the evolution in the vertical direction, and is similar to stochastic Euler-Poincaré equations. We use covariant Stratonovich calculus extensively to ensure that the reduced equations are coordinate independent.

In Chapter 11, we study stochastic collective motion from the Lagrangian point of view. Similar results to Chapter 6 can also be obtained in the Lagrangian formulation, that is, the critical point of the stochastic action functional may be expressed as the action of a Lie group valued semimartingale on the critical point of the deterministic action functional. We further formulate a coupling between a stochastic variational principle on the Lie group and a deterministic variational principle on the configuration manifold. This allows us to recover both the evolution of the momentum map and the stochastic phase from a single coupled variational principle. In this sense, this provides the variational counterpart to the coupling mechanism discussed in Chapter 7.

A summary of the main contributions is given below:

1. In Section 3.5, we formulate the reduction of stochastic differential equations given by Schwartz operators.
2. In Chapter 5, we describe a stochastic perturbation of the Kepler problem that only affects the angular direction. We also show that collisions in this problem are regularized in the sense of Moser. The results in this chapter also appear in [80].
3. We generalize the example of the stochastic Kepler problem to stochastic collective perturbations of deterministic Hamiltonian systems in Chapter 6. We show that these stochastic perturbations leave the symmetry direction unaffected. In Chapter 7, we show that these systems model the coupling between a deterministic and a stochastic Hamiltonian system.
4. In Chapter 8, we establish necessary conditions for the momentum map to be a martingale on a reductive coadjoint orbit.
5. In Chapter 9, we provide a detailed construction of fixed-endpoint variations of semimartingales on manifolds. This is used to prove a stochastic analogue of the Fundamental Lemma of Calculus of Variations. As an application of the variational framework we present a proof of the local form of the stochastic Hamilton-Pontryagin principle. We also provide a stochastic version of Noether's theorem on the variational principle side. Then we discuss the intrinsic form of the stochastic Hamilton-Pontryagin principle by working at the level of Stratonovich operators. These results are published in [81].
6. We develop the reduction by symmetry of the stochastic Hamilton-Pontryagin principle in Chapter 10.
7. The variational principle viewpoint of stochastic collective dynamics is investigated in Chapter 11. We provide analogous results to Chapter 6 on the Lagrangian side and describe the coupling mechanism via a coupled stochastic variational principle.

We end by outlining some future areas of research.

Part I

Preliminaries

Chapter 2

A Brief Review of Geometric Mechanics

The formulation of mechanics in the language of differential geometry serves three major purposes. The first one is that the laws of mechanics can be formulated on manifolds in a coordinate-free way. Coordinate invariance allows us to study important geometric features underlying mechanical systems, both in finite and infinite dimensions. Moreover, it makes the equations of motion invariant under coordinate change, allowing us the choice of any preferred coordinate system for tackling a specific problem. Second, the geometric ideas involved can be used to provide a unifying framework for studying many seemingly disparate examples, such as the Kepler problem, the rigid body, fluids, and quantum mechanics. Third, the formulation of mechanical systems on smooth manifolds allows us to study symmetries through Lie group actions, which greatly help in reducing the dimension of the phase space of a mechanical system. We shall provide a brief sketch of geometric mechanics following Marsden [60], Marsden and Ratiu [62], and Holm, Schmah, and Stoica [42].

In Section 2.1, we discuss the Hamiltonian formulation of mechanics. In Hamiltonian mechanics, we define a smooth function H on a manifold with some additional structure, such as a symplectic form or a Poisson bracket. The smooth function H is usually thought of as the total energy of a mechanical system. The additional structure on the manifold allows us to define a special vector field, called the Hamiltonian vector field, corresponding to H. In Cartesian coordinates, if the Hamiltonian H is the sum of kinetic and potential energies of a mechanical system, then the integral curves of the Hamiltonian vector field describe solutions to Newton's second law of motion and their associated momentum.

Following this, we shall discuss Noether's theorem, which, roughly speaking, states that conserved quantities are naturally associated with symmetries of a mechanical system. This idea will be formalized in Section 2.2, where we consider momentum maps for mechanical systems.

In Section 2.3 we consider reduction and reconstruction methods in Hamiltonian mechanics. In Hamiltonian reduction, at the geometric level, one shows that the geometric structure underlying the Hamiltonian system drops to a similar geometric structure on a reduced space. At the dynamic level, one shows that the Hamiltonian vector field drops to a vector field on the reduced space, which is Hamiltonian with respect to the reduced geometric structure.

In the next section, we focus on the variational formulation of mechanics where one constructs an action integral starting from a Lagrangian function defined on the tangent bundle of a manifold. The Lagrangian is often the difference of the kinetic and the potential energies of a mechanical system. The trajectory of the mechanical system is determined by finding the critical point of this action integral among all curves with fixed endpoints, which corresponds to solving the Euler-Lagrange equations for the Lagrangian. The connection with Hamiltonian mechanics is made using the Legendre transform.

Finally, in Section 2.5 we discuss reduction of variational principles. Roughly speaking, one shows that in the presence of symmetries, both the action integral and Euler-Lagrange equations drop to a reduced action integral and a reduced set of equations of motion respectively, and finding the critical points of the reduced action is equivalent to solving these reduced equations of motion.

2.1 Hamiltonian Mechanics

All manifolds considered here will be assumed to be connected and finite dimensional, unless otherwise mentioned.

Definition 2.1.1. A **Poisson bracket** on a smooth manifold P is a bilinear, skew-symmetric, binary operation

$$\{\cdot,\cdot\} : C^\infty(P) \times C^\infty(P) \to C^\infty(P)$$

such that for all $f, g, h \in C^\infty(P)$

1. $\{fg, h\} = f\,\{g, h\} + g\,\{f, h\}$
2. $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0.$

A **Poisson manifold** is a pair $(P, \{\cdot,\cdot\})$ where $\{\cdot,\cdot\}$ is a Poisson bracket on P. Often we will denote this pair by P if there is no ambiguity.

Given a Poisson manifold P, if $F \in C^\infty(P)$ satisfies $\{F, f\} = 0$ for all $f \in C^\infty(P)$ then we say that F is a **Casimir** of the Poisson manifold.

Morphisms between Poisson manifolds are defined as follows:

Definition 2.1.2. Given two Poisson manifolds, $(P_1, \{\cdot,\cdot\}_1)$ and $(P_2, \{\cdot,\cdot\}_2)$, a **Poisson map** $F : P_1 \to P_2$ is a smooth map that satisfies $\{f, g\}_2 \circ F = \{f \circ F, g \circ F\}_1$ for all $f, g \in C^\infty(P_2)$.

Poisson manifolds are closely linked to symplectic manifolds, which we now define.

Definition 2.1.3. A closed, non-degenerate 2-form Ω on P is called a **symplectic form** and the pair (P, Ω) is called a **symplectic manifold**. As in the case of Poisson manifolds, we will write this pair as P when there is no reason for confusion.

Morphisms between symplectic manifolds are often called canonical transformations in the context of classical mechanics. The definition follows:

Definition 2.1.4. Let (P, Ω) and (M, Ξ) be symplectic manifolds. A map $F : P \to M$ is said to be **symplectic** or **canonical** if $F^*\Xi = \Omega$, where F^* denotes the pullback of Ξ by F. F is said to be a **symplectomorphism** if F is a symplectic diffeomorphism.

It is a standard result that finite dimensional symplectic manifolds are even dimensional. The proof consists of first showing that if V is a vector space and Ω is a skew-symmetric, non-degenerate bilinear form on V, then V must be even dimensional. Next, if (P, Ω) is a symplectic manifold, then, for every $p \in P$, with $V = T_pP$ we obtain that $\dim T_pP = \dim P$ is even.

An important class of examples of symplectic manifolds are cotangent bundles of manifolds. Let Q be a smooth manifold and consider the 1-form θ on T^*Q defined at each point $(q, p) \in T^*Q$ by

$$\theta(q, p) = p \circ T_p\pi_{T^*Q},$$

where $\pi_{T^*Q} : T^*Q \to Q$ is the basepoint projection $(q, p) \in T^*Q \mapsto q \in Q$. This is called the **tautological 1-form** or the **Liouville 1-form** on T^*Q. Then, denoting the exterior derivative by $\mathbf{d}$, the 2-form $\Omega = -\mathbf{d}\theta$ is a symplectic form on T^*Q, called the **canonical symplectic form**. If (q^i, p_i) are cotangent-lifted local coordinates on T^*Q, then $\theta = p_i dq^i$ and $\Omega = dq^i \wedge dp_i$ (where we assume the summation convention). Unless otherwise mentioned, from now on we will always consider tangent-lifted coordinates on TQ and cotangent-lifted coordinates on T^*Q.

The next theorem shows that every symplectic manifold locally behaves like a cotangent bundle of a smooth manifold.

Theorem 2.1.5 (Darboux)**.** Let (P, Ω) be a symplectic manifold with $\dim P = 2n$. Given any $p \in P$, there exists a coordinate chart $(U, (x^1, \cdots, x^n, y^1, \cdots, y^n))$ such that $\Omega_p = dx^i \wedge dy^i$.

Having established the geometric setup, we now define Hamiltonian vector fields.

Definition 2.1.6. Let P be a smooth manifold and consider a smooth function $H \in C^\infty(P)$ that we will call a **Hamiltonian**.

1. Let Ω be a symplectic form on P. The **Hamiltonian vector field** of H is the vector field X_H satisfying
$$i_{X_H}\Omega = dH,$$
where i_{X_H} denotes the interior product of the vector field X_H with Ω.

2. Let $\{\cdot,\cdot\}$ be a Poisson bracket on P. The **Hamiltonian vector field** of H is the vector field X_H satisfying
$$X_H[f] = \{f, H\},$$
for all $f \in C^\infty(P)$. Here, $X_H[f]$ is the Lie derivative of f in the direction of X_H.

In case of symplectic manifolds, the finite dimensionality hypothesis assures us that the Hamiltonian vector field exists uniquely. Under a choice of Darboux coordinates (x^i, y^i) on P (i.e. coordinates that satisfy Darboux theorem), the integral curves for the Hamiltonian vector field X_H are given by the following system of differential equations:

$$\dot{x}^i = \frac{\partial H}{\partial y^i}$$
$$\dot{y}^i = -\frac{\partial H}{\partial x^i}$$

and in the case when $P = T^*Q$ with coordinates (q^i, p_i), these equations are locally given by

$$\dot{q}^i = \frac{\partial H}{\partial p_i} \tag{2.1.1}$$
$$\dot{p}_i = -\frac{\partial H}{\partial q^i}. \tag{2.1.2}$$

An important property for the flow of X_H in the symplectic case is that it preserves the symplectic form.

Proposition 2.1.7. Let $\varphi_t : P \to P$ denote the time t flow of X_H. Then φ_t is symplectic.

Every symplectic form naturally leads to the musical isomorphisms

$$\Omega^\flat : TP \to T^*P \text{ and } \Omega^\sharp : T^*P \to TP$$

given by

$$\left\langle \Omega^\flat_p(X_p), Y_p \right\rangle = \Omega_p(X_p, Y_p) \text{ for all } p \in P \text{ and } X_p, Y_p \in T_pP$$
$$\Omega^\sharp = (\Omega^\flat)^{-1}.$$

In terms of these musical isomorphisms, we have

$$\Omega^\flat(X_H) = dH \text{ and } X_H = \Omega^\sharp(dH).$$

Now, suppose P is a Poisson manifold. Then $z(t)$ is an integral curve of X_H if for every $f \in C^\infty(P)$,
$$\frac{df(z(t))}{dt} = X_H[f](z(t)) = \{f, H\}(z(t)).$$
Akin to the symplectic case, when P is a Poisson manifold, the time t flow of X_H is a Poisson map. Note that, with $f = H$, we see that H is conserved along any integral curve. Moreover, Casimirs are always conserved along integral curves as well.

The two notions of Hamiltonian vector fields are related as follows. Suppose (P, Ω) is a symplectic manifold. Define $\{\cdot, \cdot\}$ on P by $\{F, G\} = \Omega(X_F, X_G)$, for all $F, G \in C^\infty(P)$. Here X_F and X_G are the Hamiltonian vector fields of F and G on the symplectic manifold (P, Ω). Then, $\{\cdot, \cdot\}$ is a Poisson bracket on P. Moreover, given $H \in C^\infty(P)$, denote by $\bar{X}_H$, the Hamiltonian vector field of H on the Poisson manifold $(P, \{\cdot, \cdot\})$. Then, by definition, for every $f \in C^\infty(P)$,

$$\bar{X}_H[f] = \{f, H\} = \Omega(X_f, X_H) = \langle df, X_H \rangle = X_H[f].$$

Therefore, $\bar{X}_H = X_H$, that is, the two notions of Hamiltonian vector fields coincide.

Note that every symplectic manifold is also Poisson. However, the converse is false. For instance, if $P = \mathbb{R}$ and $\{\cdot, \cdot\} = 0$, then $\{\cdot, \cdot\}$ is a Poisson bracket but $\mathbb{R}$ is not a symplectic manifold since it is not even-dimensional.

An important example of a Poisson manifold is the dual of the Lie algebra $\mathfrak{g}^*$ of a Lie group G equipped with the Lie-Poisson bracket, which we now describe.

Example 2.1.8 (Lie-Poisson bracket)**.** Let G be a Lie group, $\mathfrak{g} = T_eG$ be its Lie algebra and $\mathfrak{g}^*$ be its dual Lie algebra. Let $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ be the Lie bracket on $\mathfrak{g}$. Given $f \in C^\infty(\mathfrak{g}^*)$, define $\frac{\delta f}{\delta \mu} \in \mathfrak{g}$ by $\left\langle \nu, \frac{\delta f}{\delta \mu} \right\rangle = df(\mu)(\nu)$, for every $\nu \in \mathfrak{g}^*$. The **±-Lie-Poisson** brackets on $\mathfrak{g}^*$ are defined to be

$$\{f, g\}_\pm(\mu) = \pm \left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu} \right] \right\rangle \tag{2.1.3}$$

for every $f, g \in C^\infty(\mathfrak{g}^*)$ and $\mu \in \mathfrak{g}^*$. Given $h \in C^\infty(\mathfrak{g}^*)$, the integral curves of X_h with respect to the ± Lie-Poisson brackets are given by the **Lie-Poisson equations**

$$\dot{\mu} = \mp \mathrm{ad}^*_{\frac{\delta h}{\delta \mu}} \mu,$$

where, given any $\xi \in \mathfrak{g}$, $\mathrm{ad}^*_\xi : \mathfrak{g}^* \to \mathfrak{g}^*$ is defined by

$$\left\langle \mathrm{ad}^*_\xi \mu, \eta \right\rangle = \langle \mu, \mathrm{ad}_\xi \eta \rangle = \langle \mu, [\xi, \eta] \rangle ,$$

for all $\mu \in \mathfrak{g}^*$ and $\eta \in \mathfrak{g}$.

2.2 Momentum Maps

We assume that $\phi : G \times P \to P$ is a smooth action of a Lie group G on a Poisson manifold P. Given any $g \in G$ and $p \in P$, we will let $\phi_g(p)$ or gp denote the element $\phi(g, p) \in P$.

Definition 2.2.1. We will say that the G-action on P is **canonical** or **Poisson** if for each $g \in G$, the map $p \in P \mapsto gp \in P$ is a Poisson map.

We will henceforth assume that the action of G on P is Poisson. Given $\xi \in \mathfrak{g}$, the G-action on P induces a vector field on P in the direction of ξ in the following way: let $\exp : \mathfrak{g} \to G$ denote the Lie exponential map. Then, we define the **infinitesimal vector field on** P **along** ξ or **generated by** ξ as $\xi_P(p) := \frac{d}{dt}\Big|_{t=0} (\exp(t\xi)p)$ for all $p \in P$.

Definition 2.2.2. Let $J : P \to \mathfrak{g}^*$ be a smooth map and for each $\xi \in \mathfrak{g}$, define $J^\xi : P \to \mathbb{R}$ by $J^\xi(p) = \langle J(p), \xi\rangle$ for all $p \in P$. J is called a **momentum map** if for every $\xi \in \mathfrak{g}$, $X_{J^\xi} = \xi_P$.

Remark 2.2.3. If J is a momentum map and $\mu_0 \in \mathfrak{g}^*$ is a constant then $J + \mu_0$ is also a momentum map.

Suppose $P = T^*Q$ and G has a left action ϕ on Q. Given each $g \in G$ and $p_q \in T^*_qQ$, we define the **cotangent lifted action** of G on T^*Q by $p_q \mapsto \left(T_{gq}\phi_{g^{-1}}\right)^* (p_q)$. Then, the cotangent lifted action of G on P admits a momentum map given by the **Noether's formula**:

$$J^\xi(p_q) = \langle p_q, \xi_Q(q)\rangle \, .$$

As a special case, the momentum map of cotangent lift of the action of $SO(3)$ on $\mathbb{R}^3$ is given by

$$(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto (\mathbf{q} \times \mathbf{p})\breve{} \in \mathfrak{so}(3)^*,$$

where the 'breve' map $\boldsymbol{\mu} \in \mathbb{R}^3 \mapsto \breve{\boldsymbol{\mu}} \in \mathfrak{so}(3)^*$ identifies $\mathbb{R}^3$ with $\mathfrak{so}(3)^*$. The details of this identification, as well as the identification of $\mathfrak{so}(3)$ with $\mathbb{R}^3$ via the 'hat' map $\boldsymbol{\xi} \in \mathbb{R}^3 \mapsto \hat{\boldsymbol{\xi}} \in \mathfrak{so}(3)$ can be found in [42, Examples 5.35 and 5.36].

The presence of a momentum map helps us in relating conserved quantities with symmetries via Noether's theorem, which we now state.

Theorem 2.2.4 (Noether)**.** Let H be a G-invariant Hamiltonian and suppose that the G-action on the Poisson manifold P admits a momentum map J. Then J is conserved along the flow of X_H.

Definition 2.2.5. Suppose $J : P \to \mathfrak{g}^*$ is a momentum map for the G-action on P. For $g \in G$, let $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ denote the adjoint action of g on $\mathfrak{g}$, that is, $\mathrm{Ad}_g(\xi) = T_g R_{g^{-1}} \circ T_e L_g(\xi)$, for every $\xi \in \mathfrak{g}$. Denote by $\mathrm{Ad}^*_{g^{-1}} : \mathfrak{g}^* \to \mathfrak{g}^*$ the dual map of $\mathrm{Ad}_{g^{-1}}$. We say that J is:

1. **Equivariant** or **coadjoint equivariant** if the following diagram commutes for all $g \in G$:

$$\begin{tikzcd} P \arrow[r, "J"] \arrow[d, "\phi_g"] & \mathfrak{g}^* \arrow[d, "Ad^*_{g^{-1}}"] \\ P \arrow[r, "J"] & \mathfrak{g}^* \end{tikzcd}$$

2. **Infinitesimally equivariant** if $J^{[\xi,\eta]} = \{J^\xi, J^\eta\}$, for every $\xi, \eta \in \mathfrak{g}$.

The momentum maps corresponding to cotangent lifted group actions are equivariant. In general, equivariant momentum maps are infinitesimally equivariant. The converse is true when G is connected, but not true in general. This is proven in [62, Theorem 12.3.2].

An important property of infinitesimally equivariant momentum maps is that they are Poisson maps. More precisely, we have the following proposition:

Proposition 2.2.6. Suppose a Lie group G acts canonically on a Poisson manifold P and the action admits an infinitesimally equivariant momentum map $J : P \to \mathfrak{g}^*$. Then, given any $f, g \in C^\infty(\mathfrak{g}^*)$:

$$\{f, g\}_+ \circ J = \{f \circ J, g \circ J\}\,.$$

If one considers a right action instead of a left action then the (+) Lie-Poisson bracket is replaced with a (−) Lie Poisson bracket.

An important class of Hamiltonians are those which are smooth functions of momentum maps.

Definition 2.2.7. Let $f \in C^\infty(\mathfrak{g}^*)$ and $J : P \to \mathfrak{g}^*$ be a momentum map for the action of a Lie group G on P. Then $f \circ J : P \to \mathbb{R}$ is called a **collective Hamiltonian**.

The next theorem describes the Hamiltonian vector field corresponding to a collective Hamiltonian.

Theorem 2.2.8 (Collective Hamiltonian Theorem)**.** Let $f \in C^\infty(\mathfrak{g}^*)$ and $H = f \circ J$. Then, given any $p \in P$

$$X_H(p) = \left(\frac{\delta f}{\delta \mu}\Big|_{\mu = J(p)} \right)_P (p).$$

2.2.1 The Action of a Lie Group on its Cotangent Bundle

An important special case is the action of Lie group G on its cotangent bundle T^*G. For every $g \in G$, let $L_g : G \to G$ and $R_g : G \to G$ denote the left and right translation maps $h \in G \mapsto L_g(h) = gh \in G$ and $h \in G \mapsto R_g(h) = hg \in G$. Then, this induces corresponding cotangent lifted left and right actions of G on T^*G given by,

$$\begin{aligned} g \cdot \alpha_h &= (T_{gh} L_{g^{-1}})^* \alpha_h =: \phi_g^L(\alpha_h) \\ \alpha_h \cdot g &= (T_{hg} R_{g^{-1}})^* \alpha_h =: \phi_g^R(\alpha_h). \end{aligned}$$

Let J_L and J_R denote momentum maps for the cotangent lifts of the left and right action of G on itself, respectively. Then, given any $\alpha_g \in T_g^*G$ and $\xi \in \mathfrak{g}$, by Noether's formula, we have

$$\begin{aligned} J_L^\xi(\alpha_g) &= \langle \alpha_g, T_e R_g(\xi) \rangle = \langle T_e^* R_g(\alpha_g), \xi \rangle \\ J_R^\xi(\alpha_g) &= \langle \alpha_g, T_e L_g(\xi) \rangle = \langle T_e^* L_g(\alpha_g), \xi \rangle \end{aligned}$$

since $\xi_G(g) = T_e R_g(\xi)$ for the left action of G on itself and $\xi_G(g) = T_e L_g(\xi)$ for the right action of G on itself. As a result, $J_L(\alpha_g) = \alpha_g \cdot g^{-1}$ and $J_R(\alpha_g) = g^{-1} \cdot \alpha_g$. Given any $\alpha_g \in T_g G$ and $\xi \in \mathfrak{g}$, we have

$$\begin{aligned} \left\langle \mathrm{Ad}_{g^{-1}}^* J_R(\alpha_g), \xi \right\rangle &= \langle (T_e^* L_g) \alpha_g, T_g L_{g^{-1}} \circ T_e R_g(\xi) \rangle \\ &= \langle \alpha_g, T_e R_g(\xi) \rangle \end{aligned}$$

$$\begin{aligned} &= \langle T_e^* R_g(\alpha_g), \xi \rangle \\ &= \langle J_L(\alpha_g), \xi \rangle \end{aligned}$$

which shows that $J_R(\alpha_g) = \mathrm{Ad}^*_{g^{-1}} J_L(\alpha_g)$. Given any $h \in G$, we also have

$$\begin{aligned} \langle J_R(\alpha_g \cdot h), \xi \rangle &= \left\langle (gh)^{-1} \cdot (\alpha_g \cdot h), \xi \right\rangle \\ &= \langle (T_e L_{gh})^* \circ (T_{gh} R_{h^{-1}})^* (\alpha_g), \xi \rangle \\ &= \langle (T_{gh} R_{h^{-1}} \circ T_e L_{gh})^* (\alpha_g), \xi \rangle \\ &= \langle \alpha_g, (T_{gh} R_{h^{-1}} \circ T_e L_{gh}) \xi \rangle \\ &= \langle \alpha_g, T_e L_g (\mathrm{Ad}_h(\xi)) \rangle \\ &= \langle \mathrm{Ad}^*_h (T_e^* L_g(\alpha_g)), \xi \rangle \\ &= \langle \mathrm{Ad}^*_h (J_R(\alpha_g)), \xi \rangle \,. \end{aligned}$$

Hence, $J_R(\alpha_g \cdot h) = \mathrm{Ad}^*_h J_R(\alpha_g)$. This is still coadjoint equivariant, once we consider the coadjoint action $\mu \in \mathfrak{g}^* \mapsto \mathrm{Ad}^*_g \mu \in \mathfrak{g}^*$ for every $g \in G$ and $\mu \in \mathfrak{g}^*$. This means that following diagram commutes for all $g \in G$:

$$\begin{tikzcd} T^*G \arrow[r, "J_R"] \arrow[d, "\phi^R_g"] & \mathfrak{g}^* \arrow[d, "Ad^*_g"] \\ T^*G \arrow[r, "J_R"] & \mathfrak{g}^*. \end{tikzcd}$$

For the case of infinitesimal equivariance, one uses the $(-)$ Lie-Poisson bracket for J_R and the $(+)$ Lie-Poisson bracket for J_L. Note that, by Noether's theorem, given a Hamiltonian $H \in C^\infty(T^*G)$, J_L is conserved if H is invariant under the cotangent lift of left translation and J_R is conserved if H is invariant under cotangent lift of right translations.

2.3 Hamiltonian Reduction and Reconstruction

We now describe how symmetries reduce the phase space of a Hamiltonian system, and how we can reconstruct the trajectories of the original system from the trajectories of the reduced system. First, we will state the Poisson reduction theorem and the Lie-Poisson reduction theorem, which is the Poisson Reduction Theorem applied to cotangent bundles of Lie groups, and then we will state the symplectic reduction theorem.

Theorem 2.3.1 (Poisson Reduction Theorem)**.** Let $(P, \{\cdot,\cdot\})$ be a Poisson manifold and assume that the Lie group G acts on P by Poisson maps. Assume that P/G is a smooth manifold and the projection $\pi : P \to P/G$ is a submersion (this is true if the action of G on P is free and proper). There exists a unique Poisson bracket $\{\cdot,\cdot\}_{\text{red}}$ on P/G defined by $\{f, g\}_{\text{red}} \circ \pi = \{f \circ \pi, g \circ \pi\}$, for all $f, g \in C^\infty(P/G)$, for which π is a Poisson map.

Further, suppose H is a G-invariant Hamiltonian on P and define the **reduced Hamiltonian** $h \in C^\infty(P/G)$ as the unique Hamiltonian on P/G satisfying $H = h \circ \pi$. Let φ and

φ^{red} denote the Hamiltonian flows of H and h, respectively. Then, for all t, $\varphi_t^{\mathrm{red}} \circ \pi = \pi \circ \varphi_t$. If $z(t)$ is an integral curve of X_H, then $\pi \circ z(t)$ is an integral curve of X_h.

As a special case, with $P = T^*G$, we obtain **Lie-Poisson reduction**. We provide some more details in this case.

If $P = T^*G$ then G acts on P via cotangent lifts of left, as well as right actions, as described in the previous section. Corresponding to each of these actions, one obtains diffeomorphisms

$$\begin{aligned} \lambda_L &: T^*G \to G \times \mathfrak{g}^*, \quad \alpha_g \in T^*G \mapsto (g, g^{-1} \cdot \alpha_g) \in G \times \mathfrak{g}^* \\ \lambda_R &: T^*G \to G \times \mathfrak{g}^*, \quad \alpha_g \in T^*G \mapsto (g, \alpha_g \cdot g^{-1}) \in G \times \mathfrak{g}^*. \end{aligned}$$

The maps λ_L and λ_R are called the **left** and **right trivialization maps** respectively. Then the canonical Poisson bracket $\{\cdot,\cdot\}$ on T^*G can be pushed forward to a Poisson bracket on $G \times \mathfrak{g}^*$ under each of these trivializations. Let $\pi_{\mathfrak{g}^*} : G \times \mathfrak{g}^* \to \mathfrak{g}^*$ denote the projection onto the $\mathfrak{g}^*$-direction. Then $J_R = \pi_{\mathfrak{g}^*} \circ \lambda_L$ and $J_L = \pi_{\mathfrak{g}^*} \circ \lambda_R$, where J_R and J_L are the momentum maps for the cotangent lifts of the right and left translation actions respectively. By equivariance of J_R and J_L, these maps are Poisson maps from $(T^*G, \{\cdot,\cdot\})$ onto $\mathfrak{g}^*_{\mp} := (\mathfrak{g}^*, \{\cdot,\cdot\}_{\mp})$ respectively, where $\{\cdot,\cdot\}_{\mp}$ are the $(-)$ and $(+)$ Lie-Poisson brackets on $\mathfrak{g}^*$.

Now let $H \in C^\infty(T^*G)$ be a left G-invariant Hamiltonian, that is, it is invariant under cotangent lifts of left translations . Then, by G-invariance, for all $g \in G$, we obtain a reduced Hamiltonian h_L in $C^\infty(\mathfrak{g}^*_-)$ that satisfies $H = h_L \circ J_R$. Explicitly, we have, $h_L(\mu) = H(\lambda_L^{-1}(e, \mu))$. The Lie-Poisson equations for h_L is given by

$$\dot{\mu}_L = \mathrm{ad}^*_{\frac{\delta h_L}{\delta \mu_L}} \mu_L.$$

Note that μ_L describes the evolution of J_R along integral curves of X_H. Thus, for left Lie-Poisson reduction, J_R acts as the dynamic variable for the reduced system, and J_L acts as the conserved momentum. For right Lie-Poisson reduction, one reverses the roles of J_L and J_R and uses the $(+)$ Lie-Poisson bracket on $\mathfrak{g}^*$.

Moreover, given an integral curve $\mu_L(t)$ of the Lie-Poisson equations for h_L, with $\mu_L(0) = \mu_0$, an integral curve for X_H starting at $\alpha_{g_0} \in T^*_{g_0}G$ can be easily reconstructed. Let $z(t)$ be the required integral curve, and suppose $z(t)$ projects to a curve $g(t)$ on G. Then

$$J_L(z(t)) = \mathrm{Ad}^*_{g(t)^{-1}} J_R(z(t)) = \mathrm{Ad}^*_{g(t)^{-1}} \mu_L(t).$$

Differentiating this relation, and using the fact that J_L is conserved, one obtains the following result:

Proposition 2.3.2 (Reconstruction Theorem for Lie-Poisson Reduction)**.** Let $\mu_L(t)$ be an integral curve of the Lie-Poisson equation

$$\dot{\mu}_L = \mathrm{ad}^*_{\frac{\delta h_L}{\delta \mu_L}} \mu_L$$

with $\mu_L(0) = J_R(\alpha_{g_0})$ for some $g_0 \in G$ and $\alpha_{g_0} \in T^*_{g_0}G$. Let $g(t)$ be the curve in G satisfying

$$\dot{g}(t) = T_e L_{g(t)} \left(\frac{\delta h_L}{\delta \mu_L} \right), \quad g(0) = g_0.$$

Then $z(t) = g(t) \cdot \mu(t)$ is an integral curve of X_H with $z(0) = \alpha_{g_0}$.

If P is a symplectic manifold and the G-action on P admits a coadjoint equivariant momentum map, then one can use symplectic reduction to obtain dimensionally smaller reduced spaces in general.

Theorem 2.3.3 (Symplectic Reduction Theorem, Marsden and Weinstein [58])**.** Let (P, Ω) be a symplectic manifold and G be a Lie group acting canonically on P. Assume that the G-action on P is free and proper. Let J be a coadjoint equivariant momentum map for this action. Let μ be a regular value of J and suppose G_μ is the isotropy subgroup of μ under the coadjoint action of G on $\mathfrak{g}^*$. Additionally, assume that G_μ acts freely and properly on $J^{-1}(\mu)$, so that $P_\mu := J^{-1}(\mu)/G_\mu$ is a smooth manifold. Then, there exists a unique symplectic form Ω_μ on P_μ satisfying $i^*_\mu \Omega = \pi^*_\mu \Omega_\mu$, where $i_\mu : J^{-1}(\mu) \hookrightarrow P$ denotes the inclusion and $\pi_\mu : J^{-1}(\mu) \to P_\mu$ denotes the projection.

Moreover, let $H \in C^\infty(P)$ be a G-invariant Hamiltonian and define the reduced Hamiltonian $h_\mu \in C^\infty(P_\mu)$ as the unique Hamiltonian on P_μ satisfying $h_\mu \circ \pi_\mu = H \circ i_\mu$. Then, the integral curves of X_H project to those of X_{h_μ}.

We remark that applying the symplectic reduction theorem to the case $P = T^*G$ equipped with the canonical symplectic form, with G acting on it via cotangent lifted left (respectively, right) action, yields the coadjoint orbit $\mathcal{O}_\mu$ through a regular value μ_R (respectively, μ_L) of J_L (respectively, J_R) as the reduced space. The reduced symplectic form on $\mathcal{O}_\mu$ is consistent with the $(+)$ (respectively $(-)$) Lie-Poisson bracket on $\mathfrak{g}^*$. The symplectic forms on $\mathcal{O}_\mu$ thus obtained are called the **orbit symplectic forms** or the **Kirillov-Kostant-Souriau** (KKS) symplectic forms.

Closely related to the symplectic reduction theorem is the orbit reduction theorem due to Marle [57] and Kazhdan, Kostant, and Sternberg [45].

Theorem 2.3.4 (Orbit Reduction Theorem)**.** Let (P, Ω) be a symplectic manifold and G be a Lie group acting canonically on P. Assume that the G-action on P is free and proper. Let J be a coadjoint equivariant momentum map for this action and H be a G-invariant Hamiltonian on P. Given a regular value μ of J, suppose $\mathcal{O}_\mu$ is the coadjoint orbit through μ. Then $P_{\mathcal{O}_\mu} := J^{-1}(\mathcal{O}_\mu)/G$ is a regular symplectic quotient manifold with the symplectic form $\Omega_{\mathcal{O}_\mu}$ defined as follows:
Let $J_{\mathcal{O}_\mu}$ denote the restriction of J to $J^{-1}(\mathcal{O}_\mu)$ and $\omega^+_{\mathcal{O}_\mu}$ denote the orbit symplectic form corresponding to the $(+)$-Lie-Poisson bracket. Let $\pi_{\mathcal{O}_\mu} : J^{-1}(\mathcal{O}_\mu) \to P_{\mathcal{O}_\mu}$ and $i_{\mathcal{O}_\mu} : J^{-1}(\mathcal{O}_\mu) \hookrightarrow P$ denote the natural projection and inclusion, respectively. Then $\Omega_{\mathcal{O}_\mu}$ is uniquely characterized by

$$i^*_{\mathcal{O}_\mu} \Omega = \pi^*_{\mathcal{O}_\mu} \Omega_{\mathcal{O}_\mu} + J^*_{\mathcal{O}_\mu} \omega^+_{\mathcal{O}_\mu}.$$

Let H be a G-invariant Hamiltonian and define the reduced Hamiltonian $h_{\mathcal{O}_\mu} \in C^\infty(P_{\mathcal{O}_\mu})$ by

$$h_{\mathcal{O}_\mu} \circ \pi_{\mathcal{O}_\mu} = H \circ i_{\mathcal{O}_\mu}.$$

Moreover, let μ denote the conserved value of J along the flow φ_t of X_H. Then φ_t leaves the connected components of $J^{-1}(\mathcal{O}_\mu)$ invariant and commutes with the G-action. The reduced flow $\varphi_t^{\mathcal{O}_\mu}$ on $P_{\mathcal{O}_\mu}$ uniquely determined by

$$\pi_{\mathcal{O}_\mu} \circ \varphi_t \circ i_{\mathcal{O}_\mu} = \varphi_t^{\mathcal{O}_\mu} \circ \pi_{\mathcal{O}_\mu}$$

is Hamiltonian, and corresponds to the flow of $X_{h_{\mathcal{O}_\mu}}$ on $P_{\mathcal{O}_\mu}$.

It is shown in [78, Theorem 6.4.1] that the orbit reduced spaces $P_{\mathcal{O}_\mu}$ are symplectomorphic to the symplectic reduced spaces P_μ. The relation between symplectic reduction, orbit reduction, and Poisson reduction is shown in the following diagram:

$$\begin{tikzcd}
J^{-1}(\mu) \arrow[d, "\text{symplectic reduction}"'] & \subseteq & J^{-1}(\mathcal{O}_\mu) \arrow[d, "\text{orbit reduction}"'] & \subseteq & P \arrow[d, "\text{Poisson reduction}"'] \\
J^{-1}(\mu)/G_\mu & \cong & J^{-1}(\mathcal{O}_\mu)/G & \subseteq & P/G
\end{tikzcd}$$

2.3.1 Reconstruction of Dynamics

Complementary to the process of reduction is the process of reconstruction. While reduction removes symmetries and provides a reduced Hamiltonian vector field on a dimensionally smaller space, reconstruction allows us to lift the solution of the reduced Hamiltonian vector field to a solution of the original Hamiltonian vector field in the unreduced space. We have already seen this in the case of Lie-Poisson reduction. In general, this process of 'unreduction' requires extra structure on the manifold, namely, a principal connection. We will carry this out in the symplectic reduction case. The process is similar for Poisson and orbit reductions.

Definition 2.3.5. Let P be a smooth manifold and G be a Lie group acting freely and properly on P. A **principal connection on** $\pi : P \to P/G$ is a $\mathfrak{g}$-valued 1-form $A : TP \to \mathfrak{g}$, such that the following properties are satisfied:

1. For every $\xi \in \mathfrak{g}$, $A \circ \xi_P = \xi$, where ξ_P is the infinitesimal vector field generated by ξ.
2. A is equivariant with respect to the tangent lifted action of G on TP and the Ad action of G on $\mathfrak{g}$, that is, for all $p \in P$, $v_p \in T_pP$ and $g \in G$, we have $A(T_p\phi_g(v_p)) = \mathrm{Ad}_g(A(v_p))$.

Principal connections allow us to decompose the tangent bundle of P into a 'symmetry' direction and a 'transverse' direction. More concretely, for any point $p \in P$, we define

$$\mathrm{Hor}\, T_pP := \{v_p \in T_pP \mid A_p(v_p) = 0\},$$

$$\mathrm{Ver}\, T_pP := \{v_p \in T_pP \mid T_p\pi(v_p) = 0\}.$$

Then $T_pP = \mathrm{Hor}\, T_pP \oplus \mathrm{Ver}\, T_pP$. We say that $\mathrm{Hor}\, T_pP$ is the **horizontal space** at $p \in M$ and $\mathrm{Ver}\, T_pP$ is the **vertical space** at $p \in P$. We can define the **horizontal** and **vertical bundles** of P by

$$\mathrm{Hor}\, TP := \bigcup_{p\in P} \mathrm{Hor}\, T_pP,$$
$$\mathrm{Ver}\, TP := \bigcup_{p\in P} \mathrm{Ver}\, T_pP.$$

Then $TM = \mathrm{Hor}\, TP \oplus \mathrm{Ver}\, TP$, where $\oplus$ is the Whitney sum. Given a curve $z(t)$ in P/G, denote by $z^h(t)$ the **horizontal lift** of $z(t)$ starting at $p \in P$, that is, $z^h(t)$ satisfies $z^h(0) = p$, $A(\dot{z}^h(t)) = 0$ and $\pi(z^h(t)) = z(t)$. Then, the reconstruction theorem for symplectic reduction may be stated as follows:

Theorem 2.3.6 (Reconstruction Theorem, Marsden [60, Theorem 6.1])**.** Let P be a symplectic manifold, $H \in C^\infty(P)$ be a G-invariant Hamiltonian, and $p \in P$. Suppose A is a principal connection on the principal bundle $\pi_\mu : J^{-1}(\mu) \to P_\mu$ where μ is a regular value of J. Let $z(t)$ be an integral curve of the Hamiltonian vector field X_{h_μ} corresponding to the reduced Hamiltonian h_μ on P_μ. Then, the integral curve $\gamma(t)$ of X_h with $\gamma(0) = p$ may be constructed as follows:

1. Horizontally lift $z(t)$ to a curve $z^h(t)$ with $z^h(0) = p$.
2. Consider the $\mathfrak{g}$-valued curve $\xi(t) = A \circ X_h(z^h(t))$.
3. Let $g(t)$ be the solution of $\dot{g}(t) = T_eL_{g(t)}\xi(t)$, with $g(0)$ being the identity element in G. Then $\gamma(t) = g(t)z^h(t)$.

2.4 Variational Principles in Mechanics

The foundation of the Lagrangian formulation or variational formulation of mechanics is **Hamilton's principle of stationary action**. This states that if $\mathcal{L} \in C^\infty(TQ)$ is the Lagrangian function for a mechanical system then the evolution of the system between fixed times t_1 and t_2 is along a trajectory $q(t)$ where the functional $\mathcal{S}(q(\cdot)) := \int_{t_1}^{t_2} \mathcal{L}(q(t), \dot{q}(t))dt$ from the space of curves with fixed endpoints in Q at $t = t_1$ and $t = t_2$ to $\mathbb{R}$ has a critical point. It is well known from the theory of calculus of variations that finding such a curve is equivalent to solving the Euler-Lagrange equations for $\mathcal{L}$.

Theorem 2.4.1. The curve $q(t)$ is a critical point for $\mathcal{S}$ for all deformations $\epsilon \mapsto q_\epsilon(t)$ satisfying $q_\epsilon(t_1) = a$ and $q_\epsilon(t_2) = b$, for some $a, b \in Q$, if and only if $q(t)$ solves the **Euler-Lagrange equations**

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = 0,$$

that is,

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}^i} - \frac{\partial \mathcal{L}}{\partial q^i} = 0$$

in every coordinate system.

The passage from the Lagrangian description to the Hamiltonian description occurs through the Legendre transformation.

Definition 2.4.2. Let $\mathcal{L} \in C^\infty(TQ)$. The **Legendre transformation** of $\mathcal{L}$ is the map $F\mathcal{L} : TQ \to T^*Q$ defined by

$$\langle F\mathcal{L}(v_q), w_q \rangle = \frac{d}{ds}\bigg|_{s=0} \mathcal{L}(v_q + s w_q)$$

for every $q \in Q$ and $v_q, w_q \in T_qQ$. $\mathcal{L}$ is said to be **regular** if $F\mathcal{L}$ is locally invertible, and **hyperregular** if $F\mathcal{L}$ is a diffeomorphism.

In local coordinates (q, v) on TQ, one has $F\mathcal{L}(q, v) = \frac{\partial \mathcal{L}}{\partial v}(q, v)$. Note that if $\mathcal{L}$ is hyperregular then it is regular. For regular Lagrangians, the Euler-Lagrange equations may be cast into a similar form as Hamilton's equations on T^*Q. To see this, we pull back the Liouville 1-form θ and the symplectic form Ω on T^*Q via $F\mathcal{L}$ to obtain a 1-form $\theta_{\mathcal{L}}$ and a 2-form $\Omega_{\mathcal{L}}$ on TQ. Then, since the exterior derivative commutes with the pullback, one gets $\Omega_{\mathcal{L}} = -\mathbf{d}\theta_{\mathcal{L}}$. Define the associated **energy corresponding to $\mathcal{L}$** by

$$E_{\mathcal{L}}(v_q) = \langle F\mathcal{L}(v_q), v_q \rangle - \mathcal{L}(v_q)$$

for all $q \in Q$ and $v_q \in T_qQ$. By regularity of $\mathcal{L}$, it can be shown that $\Omega_{\mathcal{L}}$ is non-degenerate, so there exists a unique vector field $Z_{\mathcal{L}}$ on TQ, such that

$$i_{Z_{\mathcal{L}}} \Omega_{\mathcal{L}} = dE_{\mathcal{L}}.$$

The vector field $Z_{\mathcal{L}}$ is called a Lagrangian vector field. Marsden and Ratiu [62, Theorem 7.3.3] show that regularity of $\mathcal{L}$ implies that $Z_{\mathcal{L}}$ satisfies $T\pi_{TQ} \circ Z_{\mathcal{L}}(v) = v$, for all $v \in TQ$ where $\pi_{TQ} : TQ \to Q$ is the basepoint projection. Moreover, the integral curves of $Z_{\mathcal{L}}$ are obtained by solving the Euler-Lagrange equations. Thus, for regular Lagrangians, Hamilton's principle is equivalent to the problem of finding the vector field $Z_{\mathcal{L}}$ on TQ. Let us also mention that $E_{\mathcal{L}}$ is conserved along solutions of Euler-Lagrange equations, even if $\mathcal{L}$ is not regular.

Now, suppose $\mathcal{L}$ is hyperregular. Then, one has the following theorem relating Lagrangian and Hamiltonian mechanics:

Theorem 2.4.3. Suppose $\mathcal{L}$ is a hyperregular Lagrangian on TQ and define $H = E_{\mathcal{L}} \circ F\mathcal{L}^{-1}$. Then

$$(F\mathcal{L})^* X_H = Z_{\mathcal{L}}.$$

Further, let $c(t)$ be an integral curve of $Z_{\mathcal{L}}$ and $d(t)$ be an integral curve of X_H with $F\mathcal{L}(c(0)) = d(0)$. Then $F\mathcal{L}(c(t)) = d(t)$ and $c(t)$ and $d(t)$ project to the same base curve on Q.

There is also a version of Noether's theorem on the Lagrangian side.

Theorem 2.4.4 (Noether's theorem, Lagrangian version)**.** Suppose $\varphi : \mathbb{R} \times Q \to Q$ is a flow on Q and $\mathcal{L}$ is invariant with respect to the tangent lift of φ, that is, for each $t \in \mathbb{R}$, $\mathcal{L} \circ T\varphi_t = \mathcal{L}$. Then $\left\langle \theta_{\mathcal{L}}, \frac{d}{ds}\Big|_{s=0} \varphi_s(\cdot) \right\rangle$ is conserved along trajectories of the Euler-Lagrange equations for $\mathcal{L}$.

2.4.1 The Hamilton-Pontryagin Principle

Hamilton's principle is formulated on the tangent bundle of a manifold, and it yields Euler-Lagrange equations. Alternatively, one can obtain Hamilton's equations via a variational principle on the cotangent bundle, called Hamilton's principle in phase space. Fixing a Hamiltonian $H \in C^\infty(T^*Q)$, the action integral is locally given by

$$\mathcal{S}(q(\cdot), p(\cdot)) = \int_{t_1}^{t_2} [\langle p(t), \dot{q}(t)\rangle - H(q(t), p(t))]dt$$

where we search for critical points under fixed endpoint conditions $q(t_1) = a$, $q(t_2) = b$ for some $a, b \in Q$ and arbitrary variations of $p(t)$. Finding the critical point of this action integral is equivalent to solving Hamilton's equations.

The Hamilton-Pontryagin principle, studied by Yoshimura and Marsden [93, 94], is formulated on the Pontryagin bundle $\mathcal{P}Q = TQ \oplus T^*Q$ via a constrained variational principle, and can be seen as a 'mix' of Hamilton's principle on TQ and Hamilton's principle in phase space on T^*Q. We shall follow the exposition in [94].

Let $\mathcal{L} \in C^\infty(TQ)$ be a Lagrangian function. In local coordinates (q, v, p) on $\mathcal{P}Q$, the action is given by

$$\mathcal{S}(q(t), v(t), p(t)) = \int_{t_1}^{t_2} [\mathcal{L}(q(t), v(t)) + \langle p(t), \dot{q}(t) - v(t)\rangle]dt.$$

The following theorem concerns the characterization of critical points of $\mathcal{S}$ under fixed endpoint variations of $q(t)$:

Theorem 2.4.5. A curve $(q(t), v(t), p(t))$ in $\mathcal{P}Q$ is a critical point for $\mathcal{S}$ among all curves with fixed endpoints in Q at $t = t_1$ and $t = t_2$, if and only if $(q(t), v(t), p(t))$ satisfies the **implicit Euler-Lagrange equations** given by

$$\dot{q} = v, \quad \dot{p} = \frac{\partial \mathcal{L}}{\partial q}, \quad p = \frac{\partial \mathcal{L}}{\partial v}. \tag{2.4.1}$$

We note that the implicit Euler-Lagrange equations naturally incorporate the Legendre transform. The Hamilton-Pontryagin principle is related to the theory of Dirac structures and finds applications in constrained systems. For an account of this, we refer to [93, 94].

Let us also describe the Hamilton-Pontryagin principle intrinsically. Let $G : \mathcal{P}Q \to \mathbb{R}$ denote the fibrewise pairing map between TQ and T^*Q, that is, $G(q, v, p) = \langle p, v\rangle$. Denote by

$$\begin{aligned} \mathrm{Pr}_{\mathcal{P}Q} &: T\mathcal{P}Q \to \mathcal{P}Q \\ \mathrm{Pr}_{T\mathcal{P}Q} &: TT\mathcal{P}Q \to T\mathcal{P}Q \\ \mathrm{pr}_Q &: \mathcal{P}Q \to Q \\ \mathrm{pr}_{TQ} &: \mathcal{P}Q \to TQ \\ \mathrm{pr}_{T^*Q} &: \mathcal{P}Q \to T^*Q \end{aligned} \tag{2.4.2}$$

the corresponding projection maps. We define the map $\rho_{TT^*Q} : TT^*Q \to \mathcal{P}Q$ in local coordinates by setting $\rho_{TT^*Q}(q, p, v_q, v_p) = (q, v_q, p)$. As shown in [93], this map is independent of the choice of local coordinates. Let $\mathcal{G}$ denote the 1-form on $\mathcal{P}Q$ given by $\mathcal{G} = G \circ \rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}$. In local coordinates, if $(u_q, u_v, u_p) \in T_{(q,v,p)}\mathcal{P}Q$ then

$$\mathcal{G}(q, v, p)(u_q, u_v, u_p) = G(q, u_q, p) = \langle p, u_q\rangle\,. \tag{2.4.3}$$

Define the **generalized energy for $\mathcal{L}$** by

$$E_{\mathcal{L}} : \mathcal{P}Q \to \mathbb{R}, \quad E_{\mathcal{L}} = G - \mathcal{L} \circ \mathrm{pr}_{TQ}.$$

In coordinates, $E_{\mathcal{L}}(q, v, p) = \langle p, v\rangle - L(q, v)$. Then, given a curve $x(t)$ in $\mathcal{P}Q$

$$\mathcal{S}(x(t)) = \int_{t_1}^{t_2} [\mathcal{G}(x(t)) - E_{\mathcal{L}}(x(t))]dt. \tag{2.4.4}$$

The next theorem describes the intrinsic form of the implicit Euler-Lagrange equations.

Theorem 2.4.6. Let θ_{T^*Q} and $\theta_{T^*T^*Q}$ denote the Liouville 1-forms on T^*Q and T^*T^*Q respectively, Ω denote the symplectic form on T^*Q, and $\chi = (\Omega^\flat)^*(\theta_{T^*T^*Q})$. A curve $x(t)$ in $\mathcal{P}Q$, $t_1 \leq t \leq t_2$ joining $\mathrm{pr}_Q(x(t_1)) = q_1$ and $\mathrm{pr}_Q(x(t_2)) = q_2$ satisfies

$$(T\mathrm{pr}_{T^*Q})^*\chi(x(t), \dot{x}(t)) = (\mathrm{Pr}_{\mathcal{P}Q})^*\mathbf{d}E_{\mathcal{L}}(x(t), \dot{x}(t))$$

if and only if it is a critical point of $\mathcal{S}$ among all curves $x(t)$ such that $\mathrm{pr}_Q(x(t_1))$ and $\mathrm{pr}_Q(x(t_2))$ are fixed.

2.5 Implicit Lagrange-Poincaré Reduction

In this section, following Yoshimura and Marsden [95], we carry out reduction by symmetry for the Hamilton-Pontryagin principle, called implicit Lagrange-Poincaré reduction. For reduction by symmetry of Hamilton's principle, we refer to Cendra, Marsden, and Ratiu [19] and for Hamilton's principle in phase space we refer to Cendra, Marsden, Pekarsky, and Ratiu [18]. Implicit Lagrange-Poincaré reduction may be thought of as the variational equivalent of Poisson reduction. The analogue of symplectic reduction is called Routh reduction and can be found in Marsden and Scheurle [63]. Note that, unlike implicit Lagrange-Poincaré reduction, Routh reduction occurs at a fixed value of the momentum map.

Following [95] and [19], we first describe the geometric background. Let $\phi : G \times Q \to Q$ be a free and proper action of a Lie group G on Q. If $v_q \in T_qQ$ and $p_q \in T_q^*Q$, we let $gv_q := T_q\phi_g(v_q)$ and $gp_q = T^*_{gq}\phi_{g^{-1}}(p_q)$ denote the image of v_q and p_q under the tangent lifted action of G on TQ and the cotangent lifted action of G on T^*Q respectively . Let $A : TQ \to \mathfrak{g}$ denote a principal connection on the principal bundle $\pi : Q \to Q/G$. For each $q \in Q$, the equivalence class $\pi(q)$ will be denoted by $[q]$ or $[q]_G$. The horizontal and vertical subspaces at q are denoted by Hor_q and Ver_q respectively, whereas $\mathrm{Hor}\, TQ$ and $\mathrm{Ver}\, TQ$ will denote the corresponding horizontal and vertical bundles.

Lemma 2.5.1. Let $q(t)$ be a curve in Q, $q^h(t)$ the horizontal lift of $\pi(q(t))$ starting at $q_0 = q(0)$, and $g^q(t)$ a G-valued curve satisfying $q(t) = g^q(t)q^h(t)$. Then $A(q(t), \dot{q}(t)) = \dot{g}^q(t)g^q(t)^{-1}$.

The curvature of A is the $\mathfrak{g}$-valued 2-form B defined by $B(X, Y) = \mathbf{d}A(\mathrm{Hor}\, X, \mathrm{Hor}\, Y)$, where $\mathbf{d}A$ is the exterior derivative of A and X and Y are vector fields on Q. This satisfies

$$B(X, Y) = -A([\mathrm{Hor}\, X, \mathrm{Hor}\, Y]) = \mathbf{d}A(X, Y) - [A(X), A(Y)]. \tag{2.5.1}$$

Let G act on $Q \times \mathfrak{g}$ by $(q, \xi) \mapsto (gq, \mathrm{Ad}_g\xi)$ for all $g \in G$ and $\tilde{\mathfrak{g}} = (Q \times \mathfrak{g})/G$. The equivalence class of (q, ζ) will be denoted by $[q, \zeta]_G$. We consider the associated vector bundle $\pi_{\tilde{\mathfrak{g}}} : (Q \times \mathfrak{g})/G \to Q/G$ with fibre $\mathfrak{g}$. There is a natural Lie algebra structure on the fibres of $\tilde{\mathfrak{g}}$ given by $[[q, \eta]_G, [q, \zeta]_G] = [q, [\eta, \zeta]]_G$. By using Ad-invariance of the Lie bracket, one can show that this is well defined. The covariant derivative of a curve $[q(t), \zeta(t)]_G$ in $\tilde{\mathfrak{g}}$ is given

$$\frac{D}{Dt}[q(t), \zeta(t)]_G = [q(t), [\zeta(t), A(q(t), \dot{q}(t))] + \dot{\zeta}(t)]_G. \tag{2.5.2}$$

The corresponding connection on $\tilde{\mathfrak{g}}$ is denoted by $\nabla^{\tilde{\mathfrak{g}}}$. Moreover, let $\bar{\mu}(t)$ and $\bar{\zeta}(t)$ be curves in $\tilde{\mathfrak{g}}^*$ and $\tilde{\mathfrak{g}}$ respectively such that they project to the same curve in Q/G. We define the covariant derivative of $\bar{\mu}(t)$ via the following relation:

$$\frac{d}{dt}\left\langle \bar{\mu}(t), \bar{\zeta}(t) \right\rangle = \left\langle \bar{\mu}(t), \frac{D\bar{\zeta}(t)}{Dt} \right\rangle + \left\langle \frac{D\bar{\mu}(t)}{Dt}, \bar{\zeta}(t) \right\rangle. \tag{2.5.3}$$

The covariant derivative of curves $\bar{\mu}(t)$ in $\tilde{\mathfrak{g}}^*$ defines a unique connection on $\tilde{\mathfrak{g}}^*$ which we denote by $\nabla^{\tilde{\mathfrak{g}}^*}$.

We have a well-defined isomorphism of fibre bundles $\Psi_A : TQ/G \to T(Q/G) \oplus \tilde{\mathfrak{g}}$ given by

$$[v_q]_G \in TQ/G \mapsto (T_q\pi(v_q), [q, A_q(v_q)]_G) \in T(Q/G) \oplus \tilde{\mathfrak{g}}.$$

Dually, one has an isomorphism $(\Psi_A^{-1})^* : T^*Q/G \to T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$ given by

$$\left\langle (\Psi_A^{-1})^*([\alpha_q]), (u_{[q]}, [q, \eta]_G) \right\rangle = \left\langle (\alpha_q)_q^{h^*}, u_{[q]} \right\rangle + \left\langle J(\alpha_q), \eta \right\rangle,$$

where $(u_{[q]}, [q,\eta]_G) \in T(Q/G) \oplus \tilde{\mathfrak{g}}$, $J : T^*Q \to \mathfrak{g}^*$ is the momentum map of the cotangent lifted G-action on T^*Q, $(\cdot)^h_q : T_{[q]}(Q/G) \to T_qQ$ is the horizontal lift map that sends $v_{[q]} \in T_{[q]}(Q/G)$ to $(T_q\pi|_{\mathrm{Hor}_q})^{-1}(v_{[q]})$ and $(\cdot)^{h^*}_q$ is the dual of $(\cdot)^h_q$. Using Ψ_A and $(\Psi_A^{-1})^*$ we obtain an isomorphism

$$\tilde{\Psi}_A : \mathcal{P}Q/G \to T(Q/G) \oplus T^*(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^* = \mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*.$$

Remark 2.5.2. The bundle TQ/G over Q/G is called the **Atiyah quotient** (see Mackenzie [56]) while $T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$ is called the **Weinstein space** (see Ortega and Ratiu [78]).

Corresponding to the curvature 2-form B, we can define a $\tilde{\mathfrak{g}}$-valued curvature 2-form on Q/G given by

$$\left\langle \bar{\mu}, \tilde{B}(\pi(q))(u_{[q]}, v_{[q]}) \right\rangle = \left\langle \bar{\mu}, [q, B(q)(u_q, v_q)]_G \right\rangle, \tag{2.5.4}$$

where $q \in Q$, $\mu \in \mathfrak{g}^*$, $\bar{\mu} = [q,\mu]_G \in \tilde{\mathfrak{g}}^*$, and $u_{[q]}$ and $v_{[q]}$ are elements of $T_{\pi(q)}(Q/G)$ with $u_{[q]} = T_q\pi(u_q)$ and $v_{[q]} = T_q\pi(v_q)$. Equivariance of B can be used to show that $\tilde{B}$ is well-defined.

2.5.1 The Structure of Variations

Let $q(t)$ be a curve in Q and $\epsilon \mapsto q_\epsilon(t)$ be a deformation of $q(t)$. The corresponding variation, $\frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0} q_\epsilon(t)$ is denoted by δq. Yoshimura and Marsden [95] prove the following results:

Theorem 2.5.3. Let $\xi(t) = A(q(t), \dot{q}(t))$. Then the following holds:

1. If δq is vertical then
$$\delta\xi = \dot{\eta} + [\eta, \xi], \tag{2.5.5}$$
where $\eta(t) = A(q(t), \delta q(t))$. If $\delta q(t)$ vanishes at $t = t_1$ and $t = t_2$ then the same holds for $\eta(t)$.
2. If δq is horizontal then
$$\delta\xi = B(q)(\delta q, \dot{q}). \tag{2.5.6}$$

Let $\bar{\xi}(t) = [q(t), \xi(t)]_G$. Given a deformation $\epsilon \mapsto q_\epsilon(t)$, we define the covariant variation of $\bar{\xi}$ by

$$\delta^A\bar{\xi} = \frac{D}{D\epsilon}\Big|_{\epsilon=0} [q_\epsilon(t), \xi_\epsilon(t)]_G. \tag{2.5.7}$$

As a consequence of Theorem 2.5.3 and the definition of the covariant derivative on the adjoint bundle, it can be shown that

$$\delta^A\bar{\xi} = \frac{D\bar{\eta}}{Dt} + [\bar{\xi}, \bar{\eta}] + \tilde{B}(\delta x, \dot{x}), \tag{2.5.8}$$

where $\bar{\eta} = [q, A(q, \delta q)]_G$. If $\delta q(t)$ vanishes at $t = t_1$ and $t = t_2$ then so does $\bar{\eta}(t)$.

Let $(q(t), v(t), p(t))$ be a curve in $\mathcal{P}Q$ and consider the projection $[q(t), v(t), p(t)]_G$ on the reduced Pontryagin bundle $\mathcal{P}Q/G$. We let

$$\tilde{\Psi}_A([q(t), v(t), p(t)]_G) = (x(t), u(t), y(t), \bar{\zeta}(t), \bar{\mu}(t)).$$

Here $x(t) = \pi(q(t))$, $u(t) = T_{q(t)}\pi(v(t))$, $y(t) = (p(t))_q^{h^*}$, $\bar{\zeta}(t) = [q(t), A(q(t), v(t))]_G$ and $\bar{\mu}(t) = [q(t), J(q(t), p(t))]_G$. If $(\delta q(t), \delta v(t), \delta p(t))$ is a general variation of $(q(t), v(t), p(t))$ then the variation of the curve $(x(t), u(t), y(t), \bar{\zeta}(t), \bar{\mu}(t))$ satisfies

$$(\delta x(t), \delta u(t), \delta y(t), \delta\bar{\zeta}(t), \delta\bar{\mu}(t)) \in T_{x(t)}(Q/G) \oplus T_{u(t)}T(Q/G) \oplus T_{y(t)}T^*(Q/G) \oplus T_{\bar{\zeta}(t)}\tilde{\mathfrak{g}} \oplus T_{\bar{\mu}(t)}\tilde{\mathfrak{g}}^*.$$

Remark 2.5.4. As explained in [95] and [19], given a curve $(q(t), \dot{q}(t), p(t))$ with $\bar{\xi} := [q(t), A(q(t), \dot{q}(t))]_G$ we will only consider deformations $\epsilon \mapsto \bar{\xi}_\epsilon(t)$ such that $\pi_{\tilde{\mathfrak{g}}}(\bar{\xi}_\epsilon(t)) := x_\epsilon(t)$ does not depend on ϵ. The variations corresponding to such deformations are called $\tilde{\mathfrak{g}}$-fiber variations, and $\delta\bar{\xi}(t)$ can be identified with an element in $\tilde{\mathfrak{g}}$ instead of $T_{\bar{\xi}(t)}\tilde{\mathfrak{g}}$. The variation $\delta^A\bar{\xi}(t) = \frac{D}{D\epsilon}\Big|_{\epsilon=0}\bar{\xi}_\epsilon(t)$ is an instance of a $\tilde{\mathfrak{g}}$-fiber variation and is an element of the fiber $\tilde{\mathfrak{g}}_{x(t)} \cong \mathfrak{g}$. If $x_\epsilon(t) \oplus \bar{\xi}_\epsilon(t)$ is a family of curves in $(Q/G) \oplus \tilde{\mathfrak{g}}$ depending smoothly on ϵ then its **covariant variation** is given by

$$\delta x(t) \oplus \delta^A\bar{\xi}(t) := \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0} x_\epsilon(t) \oplus \frac{D}{D\epsilon}\Big|_{\epsilon=0}\bar{\xi}_\epsilon(t).$$

In the context of implicit Lagrange-Poincaré reduction, we are interested in the following example of covariant variation. Suppose $q(t)$ is a curve in Q and $\epsilon \mapsto q_\epsilon(t)$ is a deformation of $q(t)$. This induces a variation of the curve $x(t) \oplus \bar{\xi}(t) := \pi(q(t)) \oplus [q(t), A(q(t), \dot{q}(t))]_G$ given by

$$\epsilon \mapsto x_\epsilon(t) \oplus \bar{\xi}_\epsilon(t) := \pi(q_\epsilon(t)) \oplus [q_\epsilon(t), A(q_\epsilon(t), \dot{q}_\epsilon(t))]_G.$$

The covariant variation in this case is given by $\delta x \oplus \delta^A\bar{\xi}$, where

$$\delta^A\bar{\xi} = \frac{D\bar{\eta}}{Dt} + [\bar{\xi}, \bar{\eta}] + \tilde{B}(\delta x, \dot{x})$$

with $\bar{\eta} = [q, A(q, \delta q)]_G$. It is under these constrained variations that we will consider stationary points of the reduced action principle.

2.5.2 The Implicit Lagrange-Poincaré Reduction Theorem

Let $\mathrm{Pr}_{TQ/G} : TQ \to TQ/G$, $\mathrm{Pr}_{T^*Q/G} : T^*Q \to T^*Q/G$ and $\mathrm{Pr}_{\mathcal{P}Q/G} : \mathcal{P}Q \to \mathcal{P}Q/G$ denote the corresponding projections onto equivalence classes. Suppose $\mathcal{L} \in C^\infty(TQ)$ is invariant under the tangent lifted action of G on TQ. The isomorphism Ψ_A between TQ/G and $T(Q/G) \oplus \tilde{\mathfrak{g}}$ yields a reduced Lagrangian $\ell \in C^\infty(T(Q/G) \oplus \tilde{\mathfrak{g}})$ such that $\mathcal{L} = \ell \circ \Psi_A \circ \mathrm{Pr}_{TQ/G}$. Then, we have a reduced action functional corresponding to $\mathcal{S}$ on $\mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*$ given by

$$\mathcal{S}^{red}(x(t), u(t), y(t), \bar{\zeta}(t), \bar{\mu}(t)) = \int_0^T \left[\ell(x(t), u(t), \bar{\zeta}(t))dt + \langle y(t), \dot{x}(t) - u(t)\rangle\right.$$

$$+ \left\langle \bar{\mu}(t), \bar{\xi}(t) - \bar{\zeta}(t) \right\rangle \Big] . \tag{2.5.9}$$

The next theorem relates the critical points of $\mathcal{S}$ and $\mathcal{S}^{red}$. The proof of it can be found in [95].

Theorem 2.5.5 (Implicit Lagrange-Poincaré Reduction Theorem, [95, Theorem 4.1])**.** The following are equivalent:

1. The curve $(q(t), v(t), p(t))$ is a critical point of the Hamilton-Pontryagin action $\mathcal{S}$ for all variations δq, δv and δp with $\delta q(t_1) = \delta q(t_2) = 0$.
2. The curve $(q(t), v(t), p(t))$ satisfies the implicit Euler-Lagrange equations (2.4.1).
3. Let $[q(t), v(t), p(t)]_G$ denote the reduced curve and $(x(t), u(t), y(t), \bar{\zeta}(t), \bar{\mu}(t))$ denote the curve $\tilde{\Psi}_A([q(t), v(t), p(t)]_G)$. Then $(x(t), u(t), y(t), \bar{\zeta}(t), \bar{\mu}(t))$ is a critical point of the reduced action $\mathcal{S}^{red}$ for arbitrary variations δu, $\delta\bar{\zeta}$, δy and $\delta\bar{\mu}$ and for variations $\delta x \oplus \delta^A \bar{\xi}$ such that $\delta x(t_1) = \delta x(t_2) = 0$ and
$$\delta^A \bar{\xi} = \frac{D\bar{\eta}}{Dt} + [\bar{\xi}, \bar{\eta}] + \tilde{B}(\delta x, \dot{x}),$$
where $\bar{\eta}(t)$ is an arbitrary curve in $\tilde{\mathfrak{g}}$ with $\bar{\eta}(t_1) = \bar{\eta}(t_2) = 0$.
4. The curve $(x(t), u(t), y(t), \bar{\zeta}(t), \bar{\mu}(t))$ satisfies the **horizontal Lagrange-Poincaré equations**
$$\frac{Dy}{Dt} = \frac{\partial \ell}{\partial x} - \left\langle \bar{\mu}, i_{\dot{x}} \tilde{B} \right\rangle, \ \ \dot{x} = u, \ \ y = \frac{\partial \ell}{\partial u}, \tag{2.5.10}$$
and the **vertical Lagrange-Poincaré equations**
$$\frac{D\bar{\mu}}{Dt} = \mathrm{ad}^*_{\bar{\xi}} \bar{\mu}, \ \ \bar{\xi} = \bar{\zeta}, \ \ \bar{\mu} = \frac{\partial \ell}{\partial \bar{\zeta}}. \tag{2.5.11}$$

Remark 2.5.6. The covariant derivative $\frac{Dy}{Dt}$ can be defined after fixing an affine connection $\nabla^{T(Q/G)}$ on the manifold Q/G. Then the definition of $\frac{Dy}{Dt}$ is analogous to that of $\frac{D\bar{\mu}}{Dt}$ (see Equation (2.5.3)).

Remark 2.5.7. In the previous theorem, the partial derivatives of ℓ are interpreted as follows: the partial derivatives $\frac{\partial \ell}{\partial u}$ and $\frac{\partial \ell}{\partial \bar{\zeta}}$ are defined by:

$$\begin{aligned} \left\langle \frac{\partial \ell}{\partial u}, v \right\rangle &= \frac{d}{ds}\Big|_{s=0} \ell(x, u + sv, \bar{\zeta}) \\ \left\langle \frac{\partial \ell}{\partial \bar{\zeta}}, \bar{\eta} \right\rangle &= \frac{d}{ds}\Big|_{s=0} \ell(x, u, \bar{\zeta} + s\bar{\eta}), \end{aligned} \tag{2.5.12}$$

where v and $\bar{\eta}$ are arbitrary elements of $T_x(Q/G)$ and $\pi_{\tilde{\mathfrak{g}}}^{-1}(x)$ respectively. To define the partial derivative $\frac{\partial l}{\partial x}$, we use the connection $\nabla^{T(Q/G)}$ on the Q/G. Then $\nabla^{T(Q/G)} \oplus \nabla^{\tilde{\mathfrak{g}}}$ defines an affine connection on the bundle $T(Q/G) \oplus \tilde{\mathfrak{g}}$. Let $x(t)$ be a curve in Q/G with $x(0) = x_0$ and $(x(t), u^h(t), \bar{\zeta}^h(t))$ denote the horizontal lift of $x(t)$ with respect to the connection $\nabla^{T(Q/G)} \oplus \nabla^{\tilde{\mathfrak{g}}}$ with $(x(0), u^h(0), \bar{\zeta}^h(0)) = (x_0, u_0, \bar{\zeta}_0) \in T(Q/G) \oplus \tilde{\mathfrak{g}}$. We define

$$\left\langle \frac{\partial \ell}{\partial x}\Big|_{(x_0, u_0, \bar{\zeta}_0)}, (x(0), \dot{x}(0)) \right\rangle = \frac{d}{dt}\Big|_{t=0} \ell(x(t), u^h(t), \bar{\zeta}^h(t)). \tag{2.5.13}$$

2.5.3 Implicit Euler-Poincaré Reduction

Consider the special case where $Q = G$. In this case, $Q/G = \{e\}$, where $\{e\}$ is the identity element. So Q/G is trivial and the reduced curvature 2-form vanishes, as does the horizontal Lagrange-Poincaré equations. Then $\mathcal{P}Q/G$ can be naturally identified with $\mathfrak{g}\oplus\mathfrak{g}^*$. Moreover, the covariant derivatives $\frac{D}{Dt}$ can be replaced by the usual time derivative $\frac{d}{dt}$. One then has the following theorem:

Theorem 2.5.8. Suppose $\mathcal{L} \in C^\infty(TG)$ is left-invariant under the tangent lifts of left translations. Then the following are equivalent:

1. The curve $(g(t), v(t), p(t))$ is a critical point of the Hamilton-Pontryagin action

   $$\mathcal{S}(g(t), v(t), p(t)) = \int_{t_1}^{t_2} [\mathcal{L}(g(t), v(t)) + \langle p(t), \dot{g}(t) - v(t)\rangle]dt.$$

   for all variations δg, δv and δp with $\delta g(t_1) = \delta g(t_2) = 0$.

2. The curve $(g(t), v(t), p(t))$ satisfies the implicit Euler-Lagrange equations

   $$\dot{g} = v, \quad \dot{p} = \frac{\partial \mathcal{L}}{\partial g}, \quad p = \frac{\partial \mathcal{L}}{\partial v}.$$

3. Let $[g(t), v(t), p(t)]_G$ denote the reduced curve which we identify with a curve $(\zeta(t), \mu(t))$ in $\mathfrak{g} \oplus \mathfrak{g}^*$. Also, let $\xi(t) = T_{g(t)}L_{g(t)^{-1}}(\dot{g}(t))$. Let $\ell : \mathfrak{g} \to \mathbb{R}$ denote the reduced Lagrangian, explicitly defined by

   $$\ell(\zeta) = \mathcal{L}(e, \zeta).$$

   Then $(\zeta(t), \mu(t))$ is a critical point of the reduced action

   $$\mathcal{S}^{\mathrm{red}}(\zeta(t), \mu(t)) = \int_{t_1}^{t_2} [\ell(\zeta) + \langle \mu(t), \xi(t) - \zeta(t)\rangle]dt$$

   for all variations of the form

   $$\delta\xi = \dot{\eta} + [\xi, \eta]$$

   where $\eta(t)$ is an arbitrary curve in $\mathfrak{g}$ with $\eta(t_1) = \eta(t_2) = 0$.

4. The curve $(\zeta(t), \mu(t))$ satisfies the **implicit Euler-Poincaré equations** given by

   $$\dot{\mu} = \mathrm{ad}^*_\zeta \mu, \quad \xi = \zeta, \quad \mu = \frac{\delta \ell}{\delta \xi}. \tag{2.5.14}$$

From the implicit Euler-Poincaré equations we obtain

$$\frac{d}{dt}\left(\frac{\delta \ell}{\delta \xi}\right) = \mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi},$$

which is called the **Euler-Poincaré equation**.

Remark 2.5.9. If we consider the tangent lift of the right action of G on TG then the Euler-Poincaré equation is given by

$$\frac{d}{dt}\left(\frac{\delta \ell}{\delta \xi}\right) = -\mathrm{ad}^*_{\xi}\frac{\delta \ell}{\delta \xi}.$$

Using ξ, we can reconstruct a solution for the implicit Euler-Lagrange equations. To do this, we let $g(t)$ solve

$$\dot{g}(t) = T_e L_{g(t)}\xi(t), \quad g(t_1) = g_0.$$

Then $(g(t), \dot{g}(t), \frac{\partial \mathcal{L}}{\partial \dot{g}}(t))$ solves the implicit Euler-Lagrange equations on $TG \oplus T^*G$ with $g(t_1) = g_0$.

The Euler-Poincaré equations are related to the Lie-Poisson equations by the following theorem. The proof of it can be found in Holm, Schmah and Stoica [42]:

Theorem 2.5.10. Suppose $\mathcal{L} \in C^\infty(TG)$ is hyperregular and invariant under the tangent lift of left (respectively, right) translations. Let H denote the Hamiltonian corresponding to $\mathcal{L}$. Denote by $\ell \in C^\infty(\mathfrak{g})$ the reduced Lagrangian and the **reduced Legendre transform** $f\ell : \mathfrak{g} \to \mathfrak{g}^*$ by

$$\langle f\ell(\zeta), \eta\rangle = \frac{d}{ds}\bigg|_{s=0} \ell(\zeta + s\eta)$$

for all $\zeta, \eta \in \mathfrak{g}$. Then, the following holds:

1. $F\mathcal{L}$ is G-equivariant with respect to the cotangent lift of left (respectively, right) translations.

2. H is invariant with respect to the cotangent lift of left (respectively, right) translations.

3. Denote by $e_\ell : \mathfrak{g} \to \mathbb{R}$ the map $e_\ell(\zeta) = E_{\mathcal{L}}(e, \zeta)$. Then $e_\ell(\zeta) = \langle f\ell(\zeta), \zeta\rangle - \ell(\zeta)$, for all $\zeta \in \mathfrak{g}$.

4. Let h^{red} be the reduced Hamiltonian corresponding to H. Then $h^{\mathrm{red}} = e_\ell \circ f\ell^{-1}$.

5. Let $\xi(t)$ be a solution of the Euler-Poincaré equations

   $$\frac{d}{dt}\left(\frac{\delta \ell}{\delta \xi}\right) = \mathrm{ad}^*_{\xi}\frac{\delta \ell}{\delta \xi} \left(\text{respectively } \frac{d}{dt}\left(\frac{\delta \ell}{\delta \xi}\right) = -\mathrm{ad}^*_{\xi}\frac{\delta \ell}{\delta \xi}\right),$$

   and $\mu(t)$ solve the Lie-Poisson equations

   $$\dot{\mu} = \mathrm{ad}^*_{\frac{\delta h^{\mathrm{red}}}{\delta \mu}}\mu \left(\text{respectively } \dot{\mu} = -\mathrm{ad}^*_{\frac{\delta h^{\mathrm{red}}}{\delta \mu}}\mu\right)$$

   with $\mu(0) = f\ell(\xi(0))$. Then $\mu(t) = f\ell(\xi(t))$.

Chapter 3

Stochastic Calculus on Manifolds

This chapter is meant as a brief introduction to stochastic calculus on manifolds, which will allow us to transition from geometric mechanics to stochastic geometric mechanics. Stochastic calculus on Euclidean spaces can be extended to the manifold setting by employing second order differential geometry. This idea was formulated in the works of Schwartz [83] and Meyer [66]. For most parts of this chapter, we shall refer to Emery [36].

In Section 3.1, we introduce concepts and notations from stochastic analysis in $\mathbb{R}$ that will be used in the thesis. In particular, two important concepts that we will review are stochastic integrals, and convergence of semimartingales. Stochastic integrals will play an important role throughout the thesis, and the notion of semimartingale convergence will be used in case of stochastic variational principles.

In the following sections we turn our attention to stochastic processes on (smooth) manifolds. We introduce second order differential geometry in Section 3.2. This plays an important role in the formulation of stochastic processes on manifolds due to the inherent second order nature of the change of variables formula for Itô integrals.

Having outlined the geometric formalism, in Sections 3.3 and 3.4 we extend many of the notions introduced in Section 3.1 to the manifold setting. Particularly, in Section 3.4, we introduce a general stochastic integral on manifolds, which yields the Itô and Stratonovich integrals as special cases. The Stratonovich integral, due to its convenient change of variables formula, will be heavily used in the formulation of stochastic geometric mechanics.

In Section 3.5 we turn our attention to stochastic differential equations on manifolds. Particularly important in this regard are Stratonovich stochastic differential equations, which geometrically behave very similar to ordinary differential equations. For instance, as shown by Lázaro-Camí and Ortega [54], reduction and reconstruction methods for ordinary differential equations also extend to Stratonovich stochastic differential equations. Here we also extend their work to show that reduction methods apply to stochastic differential equations in general.

3.1 Stochastic Analysis

In this section we provide a brief review of stochastic analysis in $\mathbb{R}$. A detailed exposition may be found in Protter [79] and Karatzas and Shreve [47].

Definition 3.1.1. Let $\mathcal{T}$ be an index set and $(\boldsymbol{\Omega}, \mathfrak{F})$ and $(\boldsymbol{\Sigma}, \mathfrak{G})$ be measurable spaces. A **stochastic process indexed by** $\mathcal{T}$ is a collection $X = (X_t)_{t \in \mathcal{T}}$ of $\boldsymbol{\Sigma}$-valued random variables on $(\boldsymbol{\Omega}, \mathfrak{F})$. Given any $\omega \in \boldsymbol{\Omega}$, the map $t \in \mathcal{T} \mapsto X_t(\omega) \in \boldsymbol{\Sigma}$ is called a **path** of the stochastic process X.

Note that we can think of the stochastic process as a map $X : \boldsymbol{\Omega} \times \mathcal{T} \to \boldsymbol{\Sigma}$. Often $\mathcal{T}$ is a partially ordered set corresponding to the flow of time. In case of continuous-time stochastic processes we take $\mathcal{T} = [0, \infty)$, and in case of discrete-time stochastic processes we take $\mathcal{T}$ to be the set of non-negative integers. In this thesis we will restrict ourselves to the case $\mathcal{T} = [0, \infty)$ or any subinterval of it. The codomain $(\boldsymbol{\Sigma}, \mathfrak{G})$ will usually be $\mathbb{R}$, $\mathbb{R}^k$ or a smooth manifold.

Definition 3.1.2. Let X be a stochastic process on $(\boldsymbol{\Omega}, \mathfrak{F})$.

1. Suppose X takes values in a topological space $\boldsymbol{\Sigma}$. We will say that X is **continuous** if almost all of its paths are continuous.

2. Suppose X takes values in $\mathbb{R}$. We will say that X is **right continuous** (respectively, **increasing**) if almost all of its paths are right continuous (respectively, increasing).

Unless otherwise mentioned, all stochastic processes considered in this thesis will be assumed to have continuous paths. Often this assumption can be relaxed to processes with right continuous paths with finite left limits.

Definition 3.1.3. A **filtration** of $\mathfrak{F}$ is a collection $\{\mathfrak{F}_t\}_{t \geq 0}$ of sub-σ-algebras of $\mathfrak{F}$ such that $\mathfrak{F}_s \subseteq \mathfrak{F}_t$ whenever $s < t$. Given a probability measure $\boldsymbol{\mathcal{P}}$ on $(\boldsymbol{\Omega}, \mathfrak{F})$, we say that the filtration $\{\mathfrak{F}_t\}_{t \geq 0}$ satisfies the **usual hypotheses** if $\mathfrak{F}_0$ contains all sets of probability measure 0 in $\mathfrak{F}$ and $\mathfrak{F}_t = \bigcap_{s > t} \mathfrak{F}_s$.

We will always assume that filtrations satisfy the usual hypotheses. If $\{\mathfrak{F}_t\}_{t \geq 0}$ is a filtration on a probability space $(\boldsymbol{\Omega}, \mathfrak{F}, \boldsymbol{\mathcal{P}})$ then we say that $(\boldsymbol{\Omega}, \mathfrak{F}, \{\mathfrak{F}_t\}_{t \geq 0})$ is a **filtered space** and $(\boldsymbol{\Omega}, \mathfrak{F}, \{\mathfrak{F}_t\}_{t \geq 0}, \boldsymbol{\mathcal{P}})$ is a **filtered probability space**. We will drop the subscript $t \geq 0$ from the filtration when the time interval $[0, \infty)$ is understood from context.

Definition 3.1.4. Let X be a stochastic process defined on a filtered probability space $(\boldsymbol{\Omega}, \mathfrak{F}, \{\mathfrak{F}_t\}, \boldsymbol{\mathcal{P}})$:

1. We say that X is **adapted to the filtration** $\{\mathfrak{F}_t\}$ if for all s, X_s is $\mathfrak{F}_s$-measurable.

2. The smallest σ-algebra on $\boldsymbol{\Omega} \times [0, \infty)$ that makes all (real-valued) continuous and adapted stochastic processes measurable is called the **predictable σ-algebra** on $\boldsymbol{\Omega} \times [0, \infty)$. Any stochastic process that is measurable with respect to the predictable σ-algebra on $\boldsymbol{\Omega} \times [0, \infty)$ is called **predictable**.

Note that a stochastic process is always adapted to the filtration it generates. From now on, unless otherwise mentioned, we will always assume that we have a filtered probability space.

Definition 3.1.5. A random variable $\tau : \mathbf{\Omega} \to [0, \infty]$ is called a **stopping time** if the process

$$(\mathbf{1}_{[[\tau,\infty[[})_t(\omega) := \begin{cases} 1, & \text{if}\, t \geq \tau(\omega) \\ 0, & \text{if}\, t < \tau(\omega) \end{cases}$$

is adapted. The σ-algebra

$$\mathfrak{F}_\tau := \{A \in \mathfrak{F} \mid A \cap \{\tau \leq t\} \in \mathfrak{F}_t, \text{ for all}\, t \geq 0\},$$

is called the **stopping time σ-algebra**.

Remark 3.1.6. We would like to point out some notational conventions adopted in this thesis. Both X and X_t denote the same stochastic process $X = \{X_t\}_{t \geq 0}$. For deterministic processes, and in particular, smooth curves, we will write the time variable in parentheses. For instance, if (q, v) denotes local coordinates on the tangent bundle of a manifold Q then $(q(t), v(t))$ will denote a TQ-valued smooth curve and (q_t, v_t) will denote a TQ-valued process. If the symbol used for a stochastic process also appears as a coordinate variable or as a generic element of the codomain, then we will explicitly include the time variable when referring to the process. If τ is a stopping time then X_τ will denote the random variable $\omega \mapsto X_{\tau(\omega)}(\omega)$. In particular, if τ is constant $T \geq 0$ then X_τ is the random variable $\omega \mapsto X_T(\omega)$.

Definition 3.1.7. Let τ be a stopping time such that there is an increasing sequence of stopping times τ_n with $\tau_n < \tau$ whenever $\tau > 0$, and $\tau_n \uparrow \tau$ on $\{\tau > 0\}$. Then τ_n is said to **announce** τ and τ is said to be a **predictable stopping time**.

By Dellacherie and Meyer [32, Theorem 77], this is equivalent to saying that $\mathbf{1}_{[[\tau,\infty[[}$ is predictable.

Let τ_1 and τ_2 be stopping times (respectively, predictable stopping times). Then the following are stopping times (respectively, predictable stopping times):

1. $\tau_1 \wedge \tau_2 := \min(\tau_1, \tau_2)$.
2. $\tau_1 \vee \tau_2 := \max(\tau_1, \tau_2)$.
3. $\tau_1 + \tau_2$.
4. $a\tau_1$ provided $a > 0$.

Given a set A and a stochastic process X, we define

$$\tau_A^h(X)(\omega) = \inf\{t \mid X_t(\omega) \in A\}$$

and

$$\tau^e_A(X)(\omega) = \inf\{t \mid X_t(\omega) \notin A\}.$$

We say that $\tau^h_A(X)$ and $\tau^e_A(X)(\omega)$ are the **first hitting time of** A and **first exit time from** A of X, respectively. If A is Borel (respectively, closed) then these are stopping times (respectively, predictable stopping times). We will drop X or A from the notation whenever they are understood from context. We refer to Bass [10] for a proof of this result.

Definition 3.1.8. Let τ_1 and τ_2 be stopping times. We define the **stochastic intervals**

$$\begin{aligned}
[[\tau_1, \tau_2]] &:= \{(\omega, t) \in \mathbf{\Omega} \times [0, \infty) \mid \tau_1(\omega) \le t \le \tau_2(\omega)\} \\
[[\tau_1, \tau_2[[&:= \{(\omega, t) \in \mathbf{\Omega} \times [0, \infty) \mid \tau_1(\omega) \le t < \tau_2(\omega)\} \\
]]\tau_1, \tau_2]] &:= \{(\omega, t) \in \mathbf{\Omega} \times [0, \infty) \mid \tau_1(\omega) < t \le \tau_2(\omega)\} \\
]]\tau_1, \tau_2[[&:= \{(\omega, t) \in \mathbf{\Omega} \times [0, \infty) \mid \tau_1(\omega) < t < \tau_2(\omega)\}
\end{aligned}$$

As suggested by the name, stopping times are used to stop processes at a particular random time.

Definition 3.1.9. Let X be a stochastic process and τ be a stopping time. The process $X^{|\tau}$ defined by $X^{|\tau}_t = X_{t\wedge\tau}$ is called the process X **stopped at** τ.

We now define the following notions of equality between stochastic processes:

Definition 3.1.10. Let X, Y be stochastic processes taking values in $(\mathbf{\Sigma}, \mathfrak{G})$. We say that

1. X is a **modification of** Y or a **version of** Y if $X_t = Y_t$ a.s., for all $t \in \mathcal{T}$.
2. X is said to be **indistinguishable from** Y if

$$\mathcal{P}(\{\omega \in \mathbf{\Omega} \mid X_t(\omega) = Y_t(\omega) \text{ for all } t \in \mathcal{T}\}) = 1.$$

For continuous (or, more generally, right continuous with finite left limits) stochastic processes taking values in $\mathbb{R}^k$ equipped with its Borel σ-algebra, these notions coincide. In general, indistinguishability implies that the processes are modifications of each other, but the converse is not true.

An important example of a continuous stochastic process is a Brownian motion, which we now define.

Definition 3.1.11. A **(standard) Brownian motion** is a continuous, adapted stochastic process B defined on some filtered probability space $(\mathbf{\Omega}, \mathfrak{F}, \{\mathfrak{F}_t\}, \mathcal{P})$ and taking values in $\mathbb{R}$ such that:

1. $B_0 = 0$ a.s.
2. If $0 = t_0 < t_1 < \cdots < t_n$ then $\{B_{t_i} - B_{t_{i-1}}\}_{i=1}^n$ are independent.
3. If $s < t$ then $B_t - B_s$ is normally distributed with mean 0 and variance $t - s$

Two important properties of Brownian motion are as follows:

1. Almost every sample path of a Brownian motion is nowhere differentiable a.s.
2. Almost every sample path of a Brownian motion has unbounded variation.

The second statement implies that Brownian motion cannot serve as an integrator in the Lebesgue-Stieltjes integral. This necessitates the development of a separate theory of stochastic integration.

3.1.1 Change of Time

Definition 3.1.12. A **change of time** or **time change** is a positive, continuous, strictly increasing random process A such that each A_t is a stopping time.

Given a stopping time τ, the processes $t \mapsto t \wedge \tau$ and $t \mapsto t + \tau$ are two important examples of changes of time. The first change of time converts a process X to $X^{|\tau}$ and the second one forgets all information up to time τ.

3.1.2 Martingales, Local Martingales, and Semimartingales

Definition 3.1.13. Let $(\boldsymbol{\Omega}, \mathfrak{F}, \{\mathfrak{F}_t\}, \boldsymbol{\mathcal{P}})$ be a filtered probability space.

1. An adapted stochastic process X is said to be a **martingale** if X_t is integrable for each t and $X_s = \mathbb{E}[X_t|\mathfrak{F}_s]$ whenever $s \leq t$.
2. An adapted stochastic process X is said to be a local martingale if there exists an increasing sequence $\{\tau_n\}$ of stopping times with $\lim_{n\to\infty} \tau_n = \infty$ a.s. such that $X^{|\tau_n}\mathbf{1}_{\{\tau_n>0\}}$ is a martingale. If X is continuous then we say that X is a **continuous local martingale**.
3. A continuous, adapted process A is said to have **finite variation** if each path of it has bounded variation on compact subintervals of $[0, \infty)$.
4. A **semimartingale** is any process which can be written as a sum of a continuous local martingale and a finite variation process. The set of all real-valued semimartingales will be denoted by $\mathscr{S}(\mathbb{R})$ or $\mathscr{S}$.

Remark 3.1.14. In the definition of a local martingale, the localization is in time, as opposed to localization in space.

Remark 3.1.15. The decomposition of a semimartingale X as $X = X_0 + M + A$, where X_0 is its initial value, M is a local martingale, A is a finite variation process and $M_0 = A_0 = 0$ a.s. is unique. The uniqueness is in an almost sure sense, and it follows from the Doob-Meyer decomposition for submartingales.

Remark 3.1.16. Note that, by our definition, semimartingales are continuous. In stochastic analysis one also studies discontinuous semimartingales, but we won't use them in this thesis.

Let us mention two important stability properties of semimartingales. The first one is stability under change of probability measure. If $\boldsymbol{\mathcal{P}}_1$ and $\boldsymbol{\mathcal{P}}_2$ are two probability measures on the filtered space $(\boldsymbol{\Omega}, \boldsymbol{\mathfrak{F}}, \{\boldsymbol{\mathfrak{F}}_t\})$ such that $\boldsymbol{\mathcal{P}}_2$ is absolutely continuous with respect to $\boldsymbol{\mathcal{P}}_1$, then a semimartingale with respect to $(\boldsymbol{\Omega}, \boldsymbol{\mathfrak{F}}, \{\boldsymbol{\mathfrak{F}}_t\}, \boldsymbol{\mathcal{P}}_1)$ is a modification of a semimartingale with respect to $(\boldsymbol{\Omega}, \boldsymbol{\mathfrak{F}}, \{\boldsymbol{\mathfrak{F}}_t\}, \boldsymbol{\mathcal{P}}_2)$. The second one is stability under changes of time. Let A be a change of time. If X is a semimartingale then X_{A_t} is a semimartingale for the filtration $\{\boldsymbol{\mathfrak{F}}_{A_t}\}$.

Another important fact is that being a semimartingale is a temporally local property: X is a semimartingale if and only if there exists an increasing sequence $\{\tau_n\}$ of stopping times with $\lim_{n\to\infty} \tau_n = \infty$ a.s. such that $X^{|\tau_n} \mathbf{1}_{\{\tau_n > 0\}}$ is a semimartingale.

3.1.3 Stochastic Integration

First, let us define the notion of uniform convergence on compact sets in probability.

Definition 3.1.17. A sequence $\{X^n\}$ of stochastic processes **converges uniformly on compact sets in probability** (ucp) to a stochastic process X if the sequence

$$\left\{ \sup_{0\le s\le t} |X^n_s - X_s| \right\}_n$$

converges to 0 in probability for every $t > 0$.

Denote by $\mathbb{D}$ the space of all adapted processes having right continuous paths with finite left limits. Convergence in ucp induces a complete metrizable topology on $\mathbb{D}$, defined via the metric

$$d_c(X, Y) = \sum_n 2^{-n} \mathbb{E}\left[1 \wedge \sup_{0\le t\le n} |X_t - Y_t|\right].$$

This also works for continuous, adapted processes taking values in $\mathbb{R}^k$; one simply replaces the absolute value by the Euclidean norm (see Arnaudon and Thalmaier [5]). The space of continuous, adapted process on $\mathbb{R}^k$ endowed with the ucp topology will be denoted by $\mathbb{D}_c(\mathbb{R}^k)$.

A useful fact is that continuous functions preserve ucp convergence. More precisely, we have the following proposition:

Proposition 3.1.18 ([5, Proposition 2.6 (1)])**.** Let $\{X_n\}$ be a sequence in $\mathbb{D}_c(\mathbb{R}^k)$ that converges to X in ucp. For every continuous function $h : \mathbb{R}^k \to \mathbb{R}$, we have $h(X_n) \xrightarrow{ucp} h(X)$.

Corollary 3.1.19. Let $\{X_n\}$ and $\{Y_n\}$ be two sequences of semimartingales such that $X_n \xrightarrow{ucp} X$ and $Y_n \xrightarrow{ucp} Y$, where X and Y are semimartingales. Then $X_n Y_n \xrightarrow{ucp} XY$.

Definition 3.1.20. A stochastic process H is said to be **locally bounded** if there exists an increasing sequence $\{\tau_n\}$ of stopping times with $\lim_{n\to\infty}\tau_n = \infty$ a.s. such that $H^{|\tau_n}\mathbf{1}_{\{\tau_n>0\}}$ is bounded.

Now we define the stochastic integral of predictable, locally bounded processes with respect to semimartingales. Following [36, Chapter 1], we will define the stochastic integral as a map from predictable, locally bounded processes to semimartingales satisfying some additional properties.

Definition 3.1.21. Given a semimartingale X and a locally bounded and predictable process H the **stochastic integral** or **Itô integral of** H **with respect to** X is a semimartingale $\int HdX$ (its value at time T is denoted by $\int_0^T HdX$) satisfying the following properties:

1. $\int_0^0 HdX = 0$.
2. It is linear in H and X.
3. If τ is a stopping time and h is $\mathfrak{F}_\tau$-measurable then $\int h\mathbf{1}_{]]\tau,\infty[[}dX = h(X - X^{|\tau})$.
4. Let $\{\tau_n\}$ be an increasing sequence of stopping times such that $\lim_{n\to\infty}\tau_n = \infty$ a.s., $\tau_0 = 0$ a.s. and $|\sigma| := \sup_n |\tau_n - \tau_{n-1}|$ is uniformly bounded. If $|\sigma|$ goes to zero as $n \to \infty$ then
$$\sum_n H_{\tau_n}(X^{|\tau_{n+1}} - X^{|\tau_n}) \xrightarrow{ucp} \int HdX.$$
5. (Module associativity property) If H_1 and H_2 are locally bounded and predictable processes then
$$\int H_1 d\left(\int H_2 dX\right) = \int H_1 H_2 dX.$$
6. (Dominated Convergence Theorem) If $\{H_n\}$ is a sequence of predictable processes converging pointwise to H (a.s.) on $[[0,\infty[[$ and $\{H_n\}$ is dominated by a locally bounded process K then
$$\int H_n dX \xrightarrow{ucp} \int HdX.$$
7. Let $\mathcal{Q}$ be a probability measure that is absolutely continuous with respect to $\mathcal{P}$. Then the stochastic integral computed on $(\boldsymbol{\Omega}, \mathfrak{F}, \{\mathfrak{F}_t\}, \mathcal{Q})$ is a modification of the stochastic integral computed on $(\boldsymbol{\Omega}, \mathfrak{F}, \{\mathfrak{F}_t\}, \mathcal{P})$.
8. Let A be a change of time. Set $\bar{H}_t = H_{A_t}$ and $\bar{X}_t = X_{A_t}$. Then
$$\int_{A_0}^{A_T} H_t dX_t = \int \mathbf{1}_{]]A_0,A_T]]}(t) H_t dX_t = \int_0^T \bar{H}_t d\bar{X}_t,$$
for all $T \geq 0$.

An important class of integrands are continuous semimartingales themselves, since they are locally bounded and predictable. Moreover, if X is a local martingale, then $\int HdX$ is a local martingale.

In [79], the stochastic integral with respect to a fixed semimartingale is first defined as a continuous linear map from the space of simple, predictable processes to the space of right continuous adapted processes with finite left limits, where both are equipped with the ucp topologies. We recall that a simple predictable process is given by

$$H = H_0 \mathbf{1}_{\{0\}} + \sum_{i=0}^{n} H_i \mathbf{1}_{]]\tau_i, \tau_{i+1}]]},$$

where $0 = \tau_0 \leq \tau_1 \leq \cdots \tau_{n+1} < \infty$ is a finite sequence of stopping times and $H_i \in \mathfrak{F}_{\tau_i}$ with $|H_i| < \infty$ a.s. for all $0 \leq i \leq n$. Then, it is shown that the space of simple predictable processes is dense in the space of adapted processes with left continuous paths and finite right limits, so the definition of the stochastic integral extends to this space as a continuous linear map via a density argument. Since the stochastic integrals of semimartingales are semimartingales, we obtain the following result:

Proposition 3.1.22. Let $\{X_n\}$ be a sequence of semimartingales converging to X in ucp, and let Y be a fixed semimartingale. Then,

$$\int X_n dY \xrightarrow{ucp} \int X dY.$$

3.1.4 Quadratic Variation

Definition 3.1.23. Let X and Y be semimartingales. The **joint quadratic variation** of X and Y is defined as

$$[X, Y] := XY - X_0 Y_0 - \int X dY - \int Y dX.$$

If $X = Y$ then we say that $[X, X]$ is the **quadratic variation** of X.

This is not necessarily 0 in general, but it is a process of bounded variation. Moreover, if X or Y has finite variation then $[X, Y] = 0$. Thus, the joint quadratic variation of two semimartingales corresponds to the joint quadratic variation of their local martingale parts. As an example, we mention that a continuous local martingale B with $B_0 = 0$ is a Brownian motion if and only if $[B, B]_t = t$.

From the definition of the quadratic variation, we see that the product rule for calculus does not extend to stochastic integrals in general. One has

$$\int X dY + \int Y dX + [X, Y] = XY - X_0 Y_0.$$

One way to extend the product rule to stochastic integration is to note that $[X,Y]=[Y,X]$ and rewrite the left side as follows:

$$\begin{aligned}\int XdY + \int YdX + [X,Y] &= \int XdY + \int YdX + \frac{1}{2}([X,Y]+[Y,X]) \\ &= \left(\int XdY + \frac{1}{2}[X,Y]\right) + \left(\int YdX + \frac{1}{2}[Y,X]\right).\end{aligned}$$

We define the **Stratonovich integral** of X with respect to Y by

$$\int X \bullet dY = \int XdY + \frac{1}{2}[X,Y].$$

Then, we have

$$\int X \bullet dY + \int Y \bullet dX = XY - X_0 Y_0.$$

The Stratonovich integral lacks the dominated convergence theorem of Itô integrals. However, unlike the Itô integral, the Stratonovich integral satisfies the usual chain rule of calculus, as the next proposition shows. This result can be found in:

Proposition 3.1.24. Let X be a semimartingale in $\mathbb{R}^k$.

1. If $f \in C^2(\mathbb{R}^k)$ then

$$f(X) - f(X_0) = \sum_{i=1}^{k} \int \frac{\partial f}{\partial X^i}(X) dX^i + \frac{1}{2} \sum_{i,j=1}^{k} \int \frac{\partial^2 f}{\partial X^i \partial X^j} d[X^i, X^j].$$

2. If $f \in C^3(\mathbb{R}^k)$ then

$$f(X) - f(X_0) = \sum_{i=1}^{k} \int \frac{\partial f}{\partial X^i}(X) \bullet dX^i.$$

Since the processes on the right are semimartingales, this shows that smooth functions of semimartingales are semimartingales. Semimartingales will play an important role in the rest of the thesis, mainly due to this property as well as their role as integrators in stochastic integration. We remark that smooth functions do not preserve the martingale or local martingale property. As an example, the Brownian motion B is a martingale, but B^2 is not even a local martingale.

3.1.5 The Semimartingale Topology

We have mentioned that a dominated convergence theorem for Stratonovich integrals cannot be formulated in the ucp topology. On the other hand, the semimartingale topology allows us to establish continuity of the Stratonovich integral with respect to each of its arguments.

Definition 3.1.25. Let $\{X_n\}$ be a sequence of semimartingales. We say that $\{X_n\}$ **converges to X in the semimartingale topology** if $X_n \xrightarrow{ucp} X$ and $\int H dX_n \xrightarrow{ucp} \int H dX$ for all bounded and predictable processes H.

This topology is induced by the metric $d_S(X, Y) = \sup_{|H|\leq 1} d_c(\int H dX, \int H dY)$ where the supremum is over all predictable processes bounded by 1. Under the semimartingale topology $\mathscr{S}$ is a complete topological vector space.

Now let X_n be a sequence of semimartingales converging to X in the semimartingale topology and suppose Y is a semimartingale. We have already seen that $\int X_n dY \xrightarrow{ucp} \int X dY$ by continuity of the stochastic integral with respect to the integrand. We wish to look at convergence with respect to the integrator. For this, we will use the next result, which is a special case of Proposition 2.6 in Arnaudon and Thalmaier [5]:

Proposition 3.1.26. Let $\{X_n\}$ and $\{Y_n\}$ be two sequences of semimartingales converging to X and Y respectively, in the semimartingale topology. Additionally, let $h \in C^k(\mathbb{R}^2)$, where $k \geq 2$. Then:

1. $h(X_n, Y_n)$ converges to $h(X, Y)$ in the semimartingale topology.
2. $[X_n, Y_n]$ converges to $[X, Y]$ in the semimartingale topology.

As a result, we get

$$\begin{aligned} \int Y dX_n &= (X_n Y - (X_n)_0 Y_0) - \int X_n dY - [Y, X_n] \\ &\xrightarrow{ucp} (XY - X_0 Y_0) - \int X dY - [Y, X] \\ &= \int Y dX. \end{aligned}$$

Moreover, by the same kind of reasoning, $\int X_n \bullet dY = \int X_n dY + \frac{1}{2}[X_n, Y]$ and $\int Y \bullet dX_n = \int Y dX_n + \frac{1}{2}[Y, X_n]$ also converge to $\int X \bullet dY$ and $\int Y \bullet dX$, respectively, in ucp. We rewrite this result with a slight modification in the following proposition:

Proposition 3.1.27. Let I be a closed interval containing 0 and $\epsilon \in I \mapsto X_\epsilon \in \mathscr{S}$ be a map from I to $\mathscr{S}$. Suppose $X_\epsilon \to X$ in the semimartingale topology as $\epsilon \to 0$. Then $\int X_\epsilon \bullet dY \xrightarrow[ucp]{\epsilon\to 0} \int X \bullet dY$ and $\int Y \bullet dX_\epsilon \xrightarrow[ucp]{\epsilon\to 0} \int Y \bullet dX$.

3.2 Second Order Differential Geometry

Unless otherwise mentioned, manifolds will be assumed to be smooth, connected, second countable, and finite-dimensional.

Definition 3.2.1. Let M be a manifold.

1. A **second order tangent vector** at $p \in M$ is a differential operator of order up to 2 at p with no constant coefficients.
2. The **second order tangent space** at $p \in M$ is the space of all second order tangent vectors at $p \in M$. The **second order tangent bundle** of M is
$$\tau M = \bigcup_{p \in M} \tau_p M.$$
We denote the projection $L_p \in \tau M \mapsto p \in M$ onto the basepoint of a second order tangent vector by $\pi_{\tau M} : \tau M \to M$.
3. A **second order vector field** on M is a smooth map $A : M \to \tau M$ such that $\pi_{\tau M} \circ A$ is the identity map on M. The space of second order vector fields on M will be denoted by $\mathfrak{X}^{(2)}(M)$.

Note that every (first order) tangent vector is a second order tangent vector. Locally, a second order tangent vector L_p at $p \in M$ can be written as

$$L_p = a^i \frac{\partial}{\partial x^i} + \frac{1}{2} a^{ij} \frac{\partial^2}{\partial x^i \partial x^j},$$

where $a^{ij} = a^{ji}$ and we have assumed the summation convention. The next result shows that a symmetric 2-tensor can be attached to every second order tangent vector. This result is a special case of Proposition 1.3 in Loos [53].

Theorem 3.2.2. Let $\mathrm{Sym}^2(TM)$ denote the space of all symmetric 2-tensors on TM and $v_1 \odot v_2$ denote the symmetric tensor product of tangent vectors v_1 and v_2. The map $P_2 : \tau M \to \mathrm{Sym}^2(TM)$ defined by

$$P_2\left(\left.\frac{\partial^2}{\partial x^i \partial x^j}\right|_p\right) = \left.\frac{\partial}{\partial x^i}\right|_p \odot \left.\frac{\partial}{\partial x^j}\right|_p$$
$$P_2\left(\left.\frac{\partial}{\partial x^i}\right|_p\right) = 0.$$

is well-defined (independent of the choice of coordinates) with kernel TM.

Thus, we have a short exact sequence

$$0 \to TM \hookrightarrow \tau M \xrightarrow{P_2} \mathrm{Sym}^2(TM) \to 0.$$

Definition 3.2.3. Let $F : M \to N$ be a smooth map between manifolds M and N.

1. The **second order tangent map of** F **at** $p \in M$ is the linear map $\tau_p F : \tau_p M \to \tau_{F(p)} N$ defined by $\tau_p F(L)[f] = L[f \circ F]$, for every $f \in C^\infty(N)$ and $L \in \tau_p M$. The **second order tangent map of** F $\tau F : \tau M \to \tau N$ is obtained by gathering all of the maps $\tau_p F$ for each $p \in M$.

2. Suppose $\gamma : I \to M$ is a curve in M where I is an interval in $\mathbb{R}$. Then the second order tangent vector $\tau_t\gamma\left(\frac{d^2}{dt^2}\right)$ is called the **acceleration of γ at t** and denoted by $\ddot{\gamma}(t)$.

3. Suppose $f \in C^\infty(M)$. Then τf is called the **second order differential of f** and is denoted by $d_2 f$.

While every tangent vector can be expressed as a velocity vector of a curve, every second order tangent vector can be expressed as a linear combination of accelerations of curves.

Second order tangent maps also satisfy the chain rule of first order calculus. Indeed, if $F : M \to N$ and $G : P \to M$, then given any $f \in C^\infty(N)$ and $L \in \tau_p P$ we have

$$\begin{aligned}\tau_{G(p)}F(\tau_p G(L))[f] &= \tau_p G(L)[f \circ F]\\ &= L[f \circ (F \circ G)]\\ &= \tau_p(F \circ G)(L)[f],\end{aligned}$$

as required.

3.2.1 Connections

The first order part of a second order tangent vector cannot be defined intrinsically, that is, if

$$L_p = a^i \frac{\partial}{\partial x^i} + \frac{1}{2} a^{ij} \frac{\partial^2}{\partial x^i \partial x^j}$$

is a second order tangent vector at $p \in M$, then the coefficients a^i do not transform covariantly under a change of coordinates. To obtain the first order part of a second order tangent vector, we need an extra structure on a manifold, namely a connection.

Definition 3.2.4. Assume that ∇ is a torsion-free, affine connection on M. The **Hessian** of a smooth function $f \in C^\infty(M)$ with respect to ∇ is a smooth map $\mathrm{Hess}_\nabla(f) : M \to T^*M \otimes T^*M$ defined by

$$\mathrm{Hess}_\nabla(f)(A, B) = A[B[f]] - \nabla_A B[f].$$

Note that, by definition, for every smooth function $f \in C^\infty(M)$ and vector fields A, B,

$$\begin{aligned}\mathrm{Hess}_\nabla(f^2)(A, B) &= A[B[f^2]] - \nabla_A B[f^2]\\ &= A[2fB[f]] - 2f\nabla_A B[f]\\ &= 2fA[B[f]] - 2f\nabla_A B[f] + 2A[f]B[f]\\ &= (2f\mathrm{Hess}_\nabla(f) + 2(df \otimes df))(A, B)\end{aligned}$$

Conversely, any assignment $f \in C^\infty(M) \to \mathrm{Hess}(f) \in \Gamma(T^*M \otimes T^*M)$ such that $\mathrm{Hess}(f)$ is symmetric and

$$\mathrm{Hess}(f^2) = 2f\mathrm{Hess}(f) + 2(df \otimes df),$$

one can define an affine, torsion-free connection ∇ via

$$\nabla_A B[f] = A[B[f]] - \mathrm{Hess}(f)(A, B)$$

such that $\mathrm{Hess}(f) = \mathrm{Hess}_\nabla(f)$ for all $f \in C^\infty(M)$.

Theorem 3.2.5. There is a bijective correspondence $\nabla \mapsto F_\nabla$ between the space of affine, torsion-free, connections on a manifold M and the space of $C^\infty(M)$-linear maps $F_\nabla : \mathfrak{X}^{(2)}M \to \mathfrak{X}^\infty(M)$ satisfying

$$\begin{aligned} F_\nabla(A) &= A, \text{ if } A \in \mathfrak{X}^\infty(M), \\ F_\nabla(L)[f] &= Lf - \langle \mathrm{Hess}_\nabla(f), P_2(L)\rangle \text{ for all } L \in \mathfrak{X}^{(2)}M, \\ F_\nabla(AB)[f] &= A[B[f]] - \mathrm{Hess}_\nabla(f)(A, B) = \nabla_A B[f], \text{ for all } A, B \in \mathfrak{X}^\infty(M). \end{aligned}$$

3.2.2 Second Order Cotangent Vectors

Now we move on to the dual space of the second order tangent space.

Definition 3.2.6. Let M be a manifold.

1. Given $p \in M$, the dual space of $\tau_p M$, denoted by $\tau_p^* M$ is called the **second order cotangent space at** $p \in M$.

2. The **second order cotangent bundle of** M is defined by

$$\tau^* M := \bigcup_{p \in M} \tau_p^* M.$$

   The projection $\alpha_p \in \tau^* M \mapsto p \in M$ of a second order cotangent vector onto its basepoint is denoted by $\pi_{\tau^* M}$.

3. A **second order 1-form** is a smooth assignment $\alpha : M \to \tau^* M$ such that $\pi_{\tau^* M} \circ \alpha$ is the identity map on M.

Example 3.2.7. Let us mention some examples of second order 1-forms:

1. If $f \in C^\infty(M)$ then $d_2 f$ is a second order 1-form. We have $\langle d_2 f, L\rangle = L[f]$, for all $L \in \tau M$.

2. If $f, g \in C^\infty(M)$ then $\Gamma(f, g)$ defined by

$$\langle \Gamma(f, g), L\rangle = L[fg] - fL[g] - gL[f]$$

   for all $L \in \tau M$ is a second order 1-form.

3. Denote the dual of $P_2 : \tau M \to \mathrm{Sym}^2(TM)$ by P_2^*. If $b : M \to \mathrm{Sym}^2(T^* M)$ is a symmetric bilinear form then $P_2^*(b)$ is a second order 1-form.

Note that each second order 1-form α induces a (first order) 1-form $\mathrm{Res}(\alpha)$ defined by restricting α_p to T_pM, for each $p \in M$. More precisely, $\mathrm{Res}(\alpha)_p = \alpha_p|_{T_pM}$ for all $p \in M$. Thus, we have a map $\mathrm{Res} : \tau^*M \to T^*M$. Note that $\mathrm{Res} \circ P_2^* : \mathrm{Sym}^2(T^*M) \to T^*M$ is always 0, since for any symmetric bilinear form b and any tangent vector X, if $i : TM \hookrightarrow \tau M$ denotes the inclusion, then

$$\begin{aligned}\langle (\mathrm{Res} \circ P_2^*(b)), X\rangle &= \langle (P_2^*(b)), i(X)\rangle \\ &= \langle b, P_2(i(X))\rangle \\ &= 0\end{aligned}$$

since TM is the kernel of P_2. Thus, we get a short exact sequence

$$0 \to \mathrm{Sym}^2(T^*M) \xrightarrow{P_2^*} \tau^*M \xrightarrow{\mathrm{Res}} T^*M \to 0.$$

If (x^i) is a system of local coordinates around a point $x \in M$ then a second order cotangent vector $\alpha \in \tau_x^*M$ can be written locally as

$$\alpha(x) = \alpha_i d_2 x^i + \frac{1}{2}\alpha_{ij} P_2^*(dx^i \odot dx^j),$$

where $\alpha_{ij} = \alpha_{ji}$.

The next result shows that first order 1-forms can yield second order 1-forms in two different ways.

Theorem 3.2.8. There exists a unique $\mathbb{R}$-linear map $\boldsymbol{d}$ from (first order) 1-forms to second order 1-form $\boldsymbol{d}\alpha$ such that for all 1-forms α and smooth functions $f \in C^\infty(M)$:

1. If $\alpha = df$ then $\boldsymbol{d}\alpha = d_2 f$.
2. $\boldsymbol{d}(f\alpha) = P_2^*(df \odot \alpha) + f\boldsymbol{d}\alpha$.
3. $(\mathrm{Res} \circ \boldsymbol{d})(\alpha) = \alpha$.
4. For all vector fields A, B

$$\begin{aligned}\langle \boldsymbol{d}\alpha, [A, B]\rangle &= \langle \alpha, [A, B]\rangle \\ \langle \boldsymbol{d}\alpha, \{A, B\}\rangle &= A[\langle \alpha, B\rangle] + B[\langle \alpha, A\rangle],\end{aligned}$$

where $\{A, B\} = \frac{1}{2}(AB + BA)$.

We have mentioned that endowing M with an affine, torsion-free connection ∇ is equivalent to having a $C^\infty(M)$-linear map $F_\nabla : \mathfrak{X}^{(2)}(M) \to \mathfrak{X}^\infty(M)$ which restricts to the identity on $\mathfrak{X}^\infty(M)$ and satisfies some additional properties. Observe that $\boldsymbol{d}$ also induces a map $\boldsymbol{d}^* : \mathfrak{X}^{(2)}(M) \to \mathfrak{X}^\infty(M)$ by duality, which restricts to the identity on $\mathfrak{X}^\infty(M)$. However, this map $\boldsymbol{d}^*$, does not correspond to a connection, since this is $\mathbb{R}$-linear and not $C^\infty(M)$-linear. On the other hand, corresponding to F_∇, one can define, again by duality, a map that converts first order 1-forms to second order 1-forms. More precisely, one has the following result:

Theorem 3.2.9. There is a bijective correspondence $\nabla \mapsto G_\nabla$ between the space of affine, torsion-free, connections on a manifold M and the space of $C^\infty(M)$-linear maps from first order 1-forms to second order 1-forms satisfying the following properties:

1. For every $f \in C^\infty(M)$,
$$G_\nabla(df) = d_2 f - P_2^*(\mathrm{Hess}_\nabla(f)).$$

2. For every $f \in C^\infty(M)$ and vector fields A, B,
$$\mathrm{Hess}_\nabla(f)(A,B) = A[B[f]] - \langle G_\nabla(df), AB\rangle .$$

Moreover, for all second order vector fields L and first order 1-form α,

$$\langle \alpha, F_\nabla(L)\rangle = \langle G_\nabla(\alpha), L\rangle .$$

3.3 Semimartingales on Manifolds

We now assume that $(\boldsymbol{\Omega}, \mathfrak{F}, \{\mathfrak{F}_t\}, \boldsymbol{\mathcal{P}})$ is a filtered probability space.

Definition 3.3.1. Let M be a manifold. A (continuous) semimartingale on a manifold M is a (continuous) stochastic process Γ such that $f(\Gamma)$ is a real-valued semimartingale for all $f \in C^\infty(M)$. The space of continuous semimartingales on a manifold M will be denoted by $\mathscr{S}(M)$.

We will always assume that semimartingales are continuous. The property of being a semimartingale is preserved under smooth maps: if $F : N \to M$ is a smooth map between manifold N and M, and Γ is a semimartingale on N, then $F(\Gamma)$ is a semimartingale on M. Another important fact is that the property of being a semimartingale is both a temporally and spatially local property, as the next two results show:

Proposition 3.3.2. Suppose Γ is a stochastic process in M and $\{\tau_n\}$ is an increasing sequence of stopping times with $\lim_{n\to\infty} \tau_n = \infty$ a.s. The following are equivalent:

1. Γ is a semimartingale in M.

2. For every n, $\Gamma^{|\tau_n}$ is a semimartingale in M.

3. For every n, the process $Y_t^n := \Gamma_{(\tau_n+t)\wedge\tau_{n+1}}$ is a semimartingale in M for the filtration $\mathfrak{G}_t^n := \mathfrak{F}_{\tau_n+t}$ and the probability measure $\boldsymbol{\mathcal{P}}_n := \boldsymbol{\mathcal{P}}[\,\cdot\,|\{\tau_{n+1} > \tau_n\}]$.

Proposition 3.3.3. Suppose $\{U_n\}$ is a countable open covering of M. If Γ is any continuous, adapted process in M, then there exists a predictable, increasing sequence of stopping times $\{\tau_k\}$, with $\tau_0 = 0$ a.s., such that $\lim_{k\to\infty} \tau_k = \infty$ a.s. and on each of the events

$$[[\tau_k, \tau_{k+1}]] \bigcap \{\tau_k < \tau_{k+1}\}$$

Γ takes its value in one of the U_n's.

3.3.0.1 The Ucp and Semimartingale Convergence on Manifolds

Definition 3.3.4. Let M be a manifold:

1. Suppose $\{\Gamma_n\}$ is a sequence of continuous, adapted processes in M. We say the $\{\Gamma_n\}$ **converges to Γ in the topology of uniform convergence on compacts in probability (ucp)** if for every smooth function f, $f(\Gamma_n) \xrightarrow{ucp} f(\Gamma)$.
2. Suppose $\{\Gamma_n\}$ is a sequence of semimartingales in M. We say the $\{\Gamma_n\}$ **converges to Γ in the semimartingale topology** if for every smooth function f, $f(\Gamma_n) \to f(\Gamma)$ in the semimartingale topology on $\mathscr{S}(\mathbb{R})$.

Note that, by definition, the limit process Γ is a continuous, adapted process (respectively, semimartingale) when $\{\Gamma_n\}$ is a sequence of continuous, adapted processes (respectively, semimartingales) converging to Γ in the ucp (respectively, semimartingale) topology.

We record the following important proposition for future reference:

Proposition 3.3.5. Suppose $f \in C^\infty(M)$, I is an interval in $\mathbb{R}$ containing 0 as an interior point, and $\epsilon \in I \mapsto \Gamma_\epsilon \in \mathscr{S}(M)$ is a map such that $\Gamma_\epsilon \to \Gamma \in \mathscr{S}(M)$ in the semimartingale topology as $\epsilon \to 0$. Additionally, suppose $\epsilon \in I \mapsto X_\epsilon \in \mathscr{S}(\mathbb{R})$ is a map such that $X_\epsilon \to X \in \mathscr{S}(\mathbb{R})$ in the semimartingale topology as $\epsilon \to 0$. Then

$$\int f(\Gamma_\epsilon) dX_\epsilon \xrightarrow[ucp]{\epsilon \to 0} \int f(\Gamma) dX$$

and

$$\int f(\Gamma_\epsilon) \bullet dX_\epsilon \xrightarrow[ucp]{\epsilon \to 0} \int f(\Gamma) \bullet dX.$$

Proof. First suppose f is compactly supported. Then, there exists $K > 0$ with $|f| < K$. Let us write

$$\int f(\Gamma_\epsilon) dX_\epsilon = \int f(\Gamma_\epsilon) dX + \int f(\Gamma_\epsilon) d(X_\epsilon - X).$$

The first term converges in the ucp topology to $\int f(\Gamma) dX$ by the Dominated Convergence Theorem since $f(\Gamma_\epsilon)$ is uniformly bounded. For the second term, by definition of the metric in semimartingale topology, for each ϵ, we have

$$\begin{aligned} d_S(X_\epsilon, X) &= \sup_{|H| \le 1} d_c\left(\int H dX_\epsilon, \int H dX\right) \\ &\ge d_c\left(\int \frac{f(\Gamma_\epsilon)}{K} dX_\epsilon, \int \frac{f(\Gamma_\epsilon)}{K} dX\right). \end{aligned}$$

Since this holds for each ϵ and the left side goes to zero as ϵ goes to 0, it follows that $d_c\left(\int \frac{f(\Gamma_\epsilon)}{K} dX_\epsilon, \int \frac{f(\Gamma_\epsilon)}{K} dX\right)$ converges to zero. Hence, $\int f(\Gamma_\epsilon) d(X_\epsilon - X)$ converges to 0 in ucp.

For the more general case, let $\{K_n\}$ be a sequence of compact sets with $M = \bigcup_n K_n$ and $K_n \subseteq K_{n+1}$. Suppose τ_n is the exit time of Γ from K_n and $\tau_0 = 0$. Then $\{\tau_n\}$ is a sequence of predictable stopping times increasing to infinity. For each n, $f(\Gamma_\epsilon^{|\tau_n})$ converges to $f(\Gamma^{|\tau_n})$ in the semimartingale topology, by [5, Proposition 2.6 (3)]. On $[[0, \tau_n[[$, replace f by a smooth function F such that $F = f$ on K_n and F is supported in K_{n+1}. Then, we have

$$\begin{aligned}\left(\int f(\Gamma_\epsilon) dX_\epsilon\right)^{|\tau_n} &= \int F(\Gamma_\epsilon^{|\tau_n}) dX_\epsilon^{|\tau_n} \\ &\xrightarrow[ucp]{\epsilon\to 0} \int F(\Gamma^{|\tau_n}) dX^{|\tau_n} \\ &= \left(\int f(\Gamma) dX\right)^{|\tau_n}.\end{aligned}$$

Since the convergence in ucp holds on each of the stochastic intervals $[[0, \tau_n[[$, it follows that

$$\int f(\Gamma_\epsilon) dX_\epsilon \xrightarrow[ucp]{\epsilon\to 0} \int f(\Gamma) dX.$$

To prove that

$$\int f(\Gamma_\epsilon) \bullet dX_\epsilon \xrightarrow[ucp]{\epsilon\to 0} \int f(\Gamma) \bullet dX,$$

we simply note that

$$\begin{aligned}\int f(\Gamma_\epsilon) \bullet dX_\epsilon &= \int f(\Gamma_\epsilon) dX_\epsilon + \frac{1}{2}[f(\Gamma_\epsilon), X_\epsilon] \\ &\xrightarrow[ucp]{\epsilon\to 0} \int f(\Gamma) dX + \frac{1}{2}[f(\Gamma), X] \\ &= \int f(\Gamma) \bullet dX.\end{aligned}$$

This completes the proof.

∎

An important notion is that of differentiability with respect to the semimartingale topology. We will use this when we discuss variations of semimartingales in manifolds.

Definition 3.3.6. Let $I \subseteq \mathbb{R}$ be an interval and $\epsilon \in I \mapsto \Gamma_\epsilon \in \mathscr{S}(M)$ be a map from I to $\mathscr{S}(M)$ with $\Gamma = \Gamma_{\epsilon_0}$, where $\epsilon_0 \in I$. We say that $\epsilon \mapsto \Gamma_\epsilon$ is **differentiable in the semimartingale topology at** $\epsilon = \epsilon_0$ if there exists a $Y \in \mathscr{S}(TM)$ such that:

1. Let $\pi_{TM} : TM \to M$ denote the projection onto the basepoint. Then $\pi_{TM}(Y) = \Gamma$.
2. For every $f \in C^\infty(M)$,
$$\lim_{\epsilon\to\epsilon_0} \frac{f(\Gamma_\epsilon) - f(\Gamma)}{\epsilon - \epsilon_0} = \langle df, Y\rangle (\Gamma)$$
in the semimartingale topology.

As mentioned in Arnaudon and Thalmaier [5], pathwise smoothness does not imply differentiability in the semimartingale topology, that is, even when the maps $\epsilon \mapsto \Gamma_\epsilon(\cdot)$ are smooth, $\epsilon \mapsto \Gamma_\epsilon$ may not be differentiable in the semimartingale topology. On the other hand, this is true if the convergence in Definition 3.3.6 is replaced by ucp convergence, in which case, we say that Γ_ϵ is **differentiable in the ucp topology at** $\epsilon = \epsilon_0 \in I$.

3.4 Stochastic Integration on Manifolds

Given a smooth function $f \in C^\infty(\mathbb{R}^k)$ and a semimartingale X in $\mathbb{R}^k$, recall the change of variables formula for semimartingales:

$$f(X) - f(X_0) = \sum_{i=1}^{k} \int \frac{\partial f}{\partial X^i}(X) dX^i + \frac{1}{2} \sum_{i,j=1}^{k} \int \frac{\partial^2 f}{\partial X^i \partial X^j} d[X^i, X^j].$$

Since the left side does not depend on coordinates, we expect the right side to be coordinate independent as well. The stochastic integral on a manifold expresses this coordinate independence by using second order differential geometry.

Definition 3.4.1. Let M be a manifold.

1. A **locally bounded** process in M is a stochastic process H taking values in M such that there exists an increasing sequence of stopping times $\{\tau_n\}$ with $\lim_{n\to\infty} \tau_n = \infty$ a.s. and $H^{|\tau_n} \mathbf{1}_{\{\tau_n > 0\}}$ takes values in a relatively compact subset of M.
2. Let (E, π, M) be a fibre bundle, Γ be any stochastic process in M and H be any stochastic process in E. If $\pi(H) = \Gamma$ then we say that H **is over** Γ.

The next theorem introduces the stochastic integral.

Theorem 3.4.2. Let $\Gamma \in \mathscr{S}(M)$. There exists a unique linear map $H \mapsto \int H d\Gamma$ from the set of all predictable, locally bounded τ^*M-valued processes over Γ to $\mathscr{S}(\mathbb{R})$ such that, for every $f \in C^\infty(M)$, every locally bounded and predictable real-valued process K, and every locally bounded and predictable τ^*M-valued process H over Γ, we have

$$f(\Gamma) - f(\Gamma_0) = \int d_2 f(\Gamma) d\Gamma$$

$$\int KH d\Gamma = \int K d\left(\int H d\Gamma\right).$$

We will call $\int H d\Gamma$ the **stochastic integral of** H **with respect to** Γ. If α is a second order 1-form then we will denote $\int \alpha(\Gamma) d\Gamma$ by $\int \alpha \, d\Gamma$ and call it the **stochastic integral of** α **with respect to** Γ.

We mention some properties of the stochastic integral in the following proposition.

Proposition 3.4.3 (Properties of the stochastic integral)**.** Suppose $\Gamma \in \mathscr{S}(M)$.

1. Suppose $\boldsymbol{\mathcal{Q}}$ is a probability measure that is absolutely continuous with respect to $\boldsymbol{\mathcal{P}}$. Then the stochastic integral computed with respect to $\boldsymbol{\mathcal{Q}}$ is a version of the one computed with respect to $\boldsymbol{\mathcal{P}}$.
2. If $t \mapsto A_t$ is a continuous change of time and H is a locally bounded and predictable process over Γ, then
$$\int_0^T H_{A_t} d\Gamma_{A_t} = \int_{A_0}^{A_T} H d\Gamma.$$
In particular,
$$\int H d\Gamma^{|\tau} = \int H d\Gamma^{|\tau}$$
for any stopping time τ.
3. If $F : M \to N$ is a smooth map between manifolds M and N, and H is a locally bounded and predictable $\tau^* N$-valued process over $F(\Gamma)$ then
$$\int H dF(\Gamma) = \int F^* H d\Gamma,$$
where F^*H is the τ^*M-valued process defined by $(F^*H)_t(\omega)(L) = H_t(\omega)(\tau_{\Gamma_t(\omega)} F(L))$, for any $L \in \tau_{\Gamma_t(\omega)} M$ and all $t \geq 0$ and $\omega \in \boldsymbol{\Omega}$.
4. Let $b \in \mathrm{Sym}^2(T^*M)$ be a symmetric bilinear form. Denote by $\int b(d\Gamma, d\Gamma)$ the semimartingale $2\int P_2^*(b) d\Gamma$. Then for all $f, g \in C^\infty(M)$
$$\int (fb)(d\Gamma, d\Gamma) = \int f(\Gamma) d\left(\int b(d\Gamma, d\Gamma)\right)$$
$$\int (df \odot dg)(d\Gamma, d\Gamma) = [f(\Gamma), g(\Gamma)]$$
5. Suppose H_1 and H_2 are locally bounded and predictable processes over Γ. Then
$$\left[\int H_1 d\Gamma, \int H_2 d\Gamma\right] = 2 \int P_2^*(\mathrm{Res}(H_1) \odot \mathrm{Res}(H_2)) d\Gamma.$$
6. If Γ is a smooth curve and α is a second order 1-form then
$$\int \alpha \, d\Gamma = \int \left\langle \mathrm{Res}(\alpha), \dot{\Gamma} \right\rangle dt.$$
7. If a second order 1-form α is written locally as
$$\alpha(x) = \alpha^i(x) d_2 x^i + \frac{1}{2} \alpha_{ij}(x) P_2^*(dx^i \odot dx^j)$$
and $\Gamma = (\Gamma^i)$ locally, then
$$\int \alpha \, d\Gamma = \int \alpha_i(\Gamma) d\Gamma^i + \frac{1}{2} \int \alpha_{ij}(\Gamma) d[\Gamma^i, \Gamma^j].$$

The Stratonovich and Itô integrals on a manifold are special cases of the stochastic integral.

Definition 3.4.4. Let $\Gamma \in \mathscr{S}(M)$ and α be any 1-form.

1. The **Stratonovich integral of** α **with respect to** Γ is defined to be the stochastic integral of $\boldsymbol{d}\alpha$ with respect to Γ. We denote the Stratonovich integral by $\int \alpha \bullet d\Gamma$
2. Suppose M is endowed with an affine, torsion-free connection ∇. The **Itô integral of** α **with respect to** Γ is defined to be the stochastic integral of $G_\nabla(\alpha)$ with respect to Γ. We denote the Itô integral by $\int \alpha \, d^\nabla \Gamma$

The Stratonovich and Itô integrals can also be extended to locally bounded and predictable processes over Γ. The change of variables formula for Stratonovich and Itô integrals are given below:

Theorem 3.4.5. Suppose $\Gamma \in \mathscr{S}(M)$.

1. (Change of variables formula for Stratonovich integrals)
For every $f \in C^\infty(M)$
$$f(\Gamma) - f(\Gamma_0) = \int df \bullet d\Gamma.$$
2. (Change of variables formula for Itô integrals)
Let M be endowed with an affine, torsion-free connection ∇. For every $f \in C^\infty(M)$
$$f(\Gamma) - f(\Gamma_0) = \int df \, d^\nabla \Gamma + \frac{1}{2} \mathrm{Hess}_\nabla(f)(d\Gamma, d\Gamma).$$

More generally, for Stratonovich integrals, the following property is very useful:

Proposition 3.4.6. Let $F : M \to N$ be a smooth map between manifolds M and N, α_N be a 1-form on N and $\Gamma \in \mathscr{S}(M)$. Let $\alpha_M := F^*\alpha_N$, that is, α_M satisfies $\langle \alpha_M(p), v_p \rangle = \langle \alpha_N(F(p)), T_pF(v_p) \rangle$ for all $p \in M$ and $v_p \in T_pM$. Then
$$\int \alpha_N \bullet dF(\Gamma) = \int \alpha_M \bullet d\Gamma.$$

Differentiability results in the semimartingale topology for Stratonovich and Itô integrals for manifolds can be found in Arnaudon and Thalmaier [5][Corollary 3.18]:

Proposition 3.4.7. Suppose α is a 1-form on M. Let I be an interval in $\mathbb{R}$ and $\epsilon \in I \mapsto \Gamma_\epsilon \in \mathscr{S}(M)$ be differentiable in the semimartingale topology. Then the maps
$$\epsilon \in I \mapsto \int \alpha \bullet d\Gamma_\epsilon \in \mathscr{S}(\mathbb{R})$$
and
$$\epsilon \in I \mapsto \int \alpha \, d^\nabla \Gamma_\epsilon \in \mathscr{S}(\mathbb{R})$$
are differentiable in the semimartingale topology.

In case of real-valued processes, a semimartingale is a martingale if and only if all of its Itô integrals are local martingales. A similar result holds on manifolds.

Theorem 3.4.8. Given a manifold M endowed with a connection ∇ and a semimartingale $\Gamma \in \mathscr{S}(M)$, the following are equivalent:

1. For every smooth function $f \in C^\infty(M)$
$$f(\Gamma) - f(\Gamma_0) - \frac{1}{2}\mathrm{Hess}_\nabla(f)(d\Gamma, d\Gamma)$$
is a local martingale.

2. There exists an adapted, locally bounded, τM-valued process L over Γ with $F_\nabla L = 0$ and an increasing, adapted, and continuous process C such that for every smooth function $f \in C^\infty(M)$
$$f(\Gamma) - f(\Gamma_0) - \int LfdC$$
is a local martingale.

If Γ satisfies either of these conditions then we say that Γ is a **∇-martingale** or simply, a **martingale** on M.

3.5 Stochastic Differential Equations on Manifolds

In the deterministic case, an ordinary differential equation can be viewed as the problem of finding an integral curve of a vector field on a manifold. In the stochastic case, we generally do not encounter differentiable objects (with respect to time), so stochastic differential equations need to be formalized via the stochastic integral, which is a second order object. Moreover, in practice, one is often interested in multiple sources of noise. For instance, one can have a deterministic vector field V on a manifold M perturbed by k sources of independent Brownian noise $B^1, \cdots, B^k$. But the vector $\mathbf{B} = (B^1, \cdots, B^k)$ lives in $\mathbb{R}^k$, so we need a recipe to generate noise in the manifold M from noise in $\mathbb{R}^k$. This is provided by a Schwartz operator.

Definition 3.5.1. Let M and N be manifolds.

1. A **Schwartz morphism** from N to M is a fibrewise linear map $F : \tau N \to \tau M$ such that the following properties hold:
   (a) The image of every tangent vector in N is a tangent vector in M.
   (b) Let $F|_{TN}$ denote the restriction of F to TN. For every second order tangent vector L in M, $P_2(F(L)) = (F|_{TN} \odot F|_{TN})(P_2(L))$.

2. A **Schwartz operator** from N to M is a family $F = \{F(x,y)\}_{(x,y)\in N\times M}$ of Schwartz morphisms from N to M, smoothly depending on x and y.

3. Given a Schwartz operator F from N to M, a **solution to the stochastic differential equation** (SDE)
$$d\Gamma = F(X,\Gamma)dX$$
is a semimartingale $\Gamma \in \mathscr{S}(M)$ such that for all second order 1-forms α
$$\int \alpha \, d\Gamma = \int F^*(X,\Gamma)(\alpha)dX$$
where $F^*(x,y) : \tau_y^* M \to \tau_x^* N$ is the dual map of $F(x,y) : \tau_x N \to \tau_y M$.

The next theorem establishes existence and uniqueness for solutions to stochastic differential equations:

Theorem 3.5.2. Let $X \in \mathscr{S}(N)$, F be a Schwartz operator from N to M, and Z be a random variable in M. There exists a predictable stopping time $\tau > 0$ (possibly infinite) and a semimartingale $\Gamma \in \mathscr{S}(M)$ such that Γ solves
$$d\Gamma = F(X,\Gamma)dX, \quad \Gamma_0 = Z$$
on $[[0,\tau[[$ and explodes on $\{\tau < \infty\}$. Moreover, if $\bar{\Gamma}$ is a semimartingale and $\bar{\tau}$ is a predictable stopping time such that $\bar{\Gamma}$ also solves the same SDE with $\bar{\Gamma}_0 = Z$ on $[[0,\bar{\tau}[[$, then $\bar{\tau} \leq \tau$ and $\bar{\Gamma} = \Gamma$ on $[[0,\bar{\tau}[[$.

While stochastic differential equations are second order objects, their first order counterparts are called Stratonovich equations. We will frequently use them in this thesis.

Definition 3.5.3. Let M and N be manifolds.

1. A **Stratonovich operator** from N to M is a family of linear maps
$$S = \{S(x,y) : T_x N \to T_y M\}_{(x,y)\in N\times M}$$
depending smoothly on x and y. The collection of all Stratonovich operators from N to M will be denoted by $\mathrm{Strat}(N,M)$.

2. Given a Stratonovich operator $S \in \mathrm{Strat}(N,M)$, a **solution to the Stratonovich stochastic differential equation** (Stratonovich SDE)
$$\bullet d\Gamma = S(X,\Gamma) \bullet dX$$
is a semimartingale $\Gamma \in \mathscr{S}(M)$ such that for all 1-forms α
$$\int \alpha \bullet d\Gamma = \int S^*(X,\Gamma)(\alpha) \bullet dX$$
where $S^*(x,y) : T_y^* M \to T_x^* N$ is the dual map of $S(x,y) : T_x N \to T_y M$.

To every Stratonovich operator, one can assign a Schwartz operator uniquely. This is the content of the next lemma. Note that this also establishes existence and uniqueness of solutions for Stratonovich SDEs, by passing to the SDE defined by the corresponding Schwartz operator.

Lemma 3.5.4. Let $S \in \mathrm{Strat}(N, M)$. There exists a unique Schwartz operator F_S from N to M such that if the curve $(x(t), y(t))$ in $N \times M$ satisfies $\dot{y}(t) = S(x(t), y(t))(\dot{x}(t))$ then $\ddot{y}(t) = F_S(x(t), y(t))(\ddot{x}(t))$. Moreover, for every $X \in \mathscr{S}(N)$, a semimartingale Γ in M solves $\bullet d\Gamma = S(X, \Gamma) \bullet dX$ if and only if it solves $d\Gamma = F_S(X, \Gamma) dX$ with the same initial condition.

We remark that Stratonovich and Schwartz operators are deterministic objects and the Schwartz operator corresponding to a Stratonovich operator can be determined by purely geometric arguments. The main reason Stratonovich SDEs feature prominently in this thesis is due to the Malliavin Transfer Principle, which roughly states that geometric properties of ordinary differential equations carry over to Stratonovich SDEs.

3.5.1 Reduction and Reconstruction of Stratonovich SDEs

If an ordinary differential equation (ODE) is symmetric under a Lie group action, then often we have a simpler ODE on a reduced space, and the solution of the original ODE can be reconstructed from the solution of the reduced ODE. Lázaro-Camí and Ortega [54] have shown that these results also carry over to Stratonovich SDEs, and we will refer to their exposition here. We will assume that $\phi : G \times M \to M$ is a smooth action of a Lie group G on M and for all $g \in G$, $\phi_g : M \to M$ is the map $p \in M \mapsto \phi(g, p)$. The Lie algebra of G, will be denoted by $\mathfrak{g}$.

Definition 3.5.5. Let $S \in \mathrm{Strat}(N, M)$. We say that S is G**-invariant** or G**-equivariant** if

$$S(x, \phi_g(y)) = T_y\phi_g \circ S(x, y)$$

for all $(x, y) \in N \times M$ and $g \in G$. The collection of all G-invariant Stratonovich operators from N to M will be denoted by $\mathrm{Strat}^G(N, M)$.

G-invariance of Stratonovich operators lead to symmetries in the solutions of the Stratonovich SDEs they describe.

Proposition 3.5.6. Let $S \in \mathrm{Strat}^G(N, M)$ and $X \in \mathscr{S}(N)$. If Γ is a solution of

$$\bullet d\Gamma = S(X, \Gamma) \bullet dX$$

then, for all $g \in G$, $\bar{\Gamma} := \phi_g(\Gamma)$ satisfies

$$\bullet d\bar{\Gamma} = S(X, \bar{\Gamma}) \bullet dX.$$

Associated to G-invariant Stratonovich operators are invariant submanifolds for the corresponding differential equation.

Definition 3.5.7. Suppose $S \in \mathrm{Strat}(N, M)$ and $X \in \mathscr{S}(N)$. Let P be an immersed submanifold in M. P is said to be an **invariant submanifold** for the Stratonovich differential equation

$$\bullet d\Gamma = S(X, \Gamma) \bullet dX$$

if for every $\mathfrak{F}_0$-measurable random variable Z taking values in P, $\Gamma_0 = Z$ implies that Γ takes values in P in $[[0, \tau[[$, where τ is the explosion time of Γ.

Assume that the G-action on M is proper. Given any isotropy subgroup I of G, we define the **isotropy type submanifold** M_I of M by $M_I := \{z \in M \mid G_z = I\}$, where G_z is the isotropy subgroup of z. By properness of the action, I is compact, and the connected components of M_I are embedded submanifolds in M. This is shown in Ortega and Ratiu [78, Proposition 2.4.7]. The next proposition states that the isotropy type submanifolds are also invariant submanifolds.

Proposition 3.5.8. Assume the G-action on M is proper and let $S \in \mathrm{Strat}(N, M)$ be G-invariant. Then, given any $X \in \mathscr{S}(N)$, the isotropy type submanifolds are invariant manifolds for the Stratonovich stochastic differential equation $\bullet d\Gamma = S(X, \Gamma) \bullet dX$.

This proposition is called the **law of conservation of the isotropy**. The next theorem is the **reduction theorem** for Stratonovich operators.

Theorem 3.5.9. Let $X \in \mathscr{S}(N)$ be any semimartingale and $S \in \mathrm{Strat}(N, M)$ be G-invariant. Let I be any isotropy subgroup in G, M_I be its isotropy type submanifold, $N(I)$ be its normalizer and $L_I := N(I)/I$. Then L_I acts freely and properly on M_I. Moreover, let $\pi_I : M \to M_I/L_I$ denote the projection and define the **reduced Stratonovich operator** $S_I \in \mathrm{Strat}(N, M_I/L_I)$ by $S_I(x, \pi_I(y)) = T_y\pi_I \circ S(x, y)$, for all $(x, y) \in N \times M_I$. If Γ solves $\bullet d\Gamma = S(X, \Gamma) \bullet dX$ and Γ_0 takes values in M_I, then $\Gamma^{\mathrm{red}} := \pi_I(\Gamma)$ satisfies $\bullet d\Gamma^{\mathrm{red}} = S_I(X, \Gamma^{\mathrm{red}}) \bullet dX$ with initial condition $\pi_I(\Gamma_0)$.

Now we describe the process of reconstruction of solutions. We will carry this out with the assumption that G acts freely and properly on M, so that we have a principal bundle $\pi : M \to M/G$. The extension to the more general setup in the reduction theorem is similar - one replaces M by M_I and G by L_I.

Let A be a principal connection on $\pi : M \to M/G$. If $\{\mathbf{e}_i\}_{i=1}^{\dim \mathfrak{g}}$ is a basis of $\mathfrak{g}$ then A can be written as $\sum_{i=1}^{\dim \mathfrak{g}} \alpha^i \mathbf{e}_i$, for some 1-forms α^i on M. Then, given any $\Gamma \in \mathscr{S}(M)$, the Stratonovich integral $\int A \bullet d\Gamma$ is defined to be equal to the $\mathfrak{g}$-valued semimartingale $\sum_{i=1}^{\dim \mathfrak{g}} \left(\int \alpha^i \bullet d\Gamma \right) \mathbf{e}_i$. Note that by linearity of the Stratonovich integral, this definition is basis independent.

Given a semimartingale $\Gamma^{M/G}$ in M/G and a $\mathfrak{F}_0$-measurable random variable Z in M with $\pi(Z) = \Gamma_0$, we define the **horizontal lift of** Γ **to** M **starting at** Z as the semimartingale Γ^h in M with $\pi(\Gamma^h) = \Gamma^{M/G}$, $\int A \bullet d\Gamma^h = 0$ and $\Gamma^h_0 = Z$. The existence and uniqueness of the horizontal lift is shown in Shigekawa [84, Theorem 2.1] and Catuogno [15, Theorem 1]. The reconstruction theorem uses the horizontal lift of the reduced solution to yield a solution of the original Stratonovich operator.

Theorem 3.5.10. Let $X \in \mathscr{S}(N)$ be any semimartingale and $S \in \mathrm{Strat}(N, M)$ be a G-invariant Stratonovich operator. Suppose $S^{\mathrm{red}} \in \mathrm{Strat}(N, M/G)$ is the reduced Stratonovich operator on M/G defined by $S^{\mathrm{red}}(x, \pi(y)) = S(x, y)$ for all $(x, y) \in N \times M$. Let $\Gamma^{M/G} \in \mathscr{S}(M/G)$ solve $\bullet d\Gamma = S^{\mathrm{red}}(X, \Gamma) \bullet dX$ with initial condition $\Gamma_0 = Z_{M/G}$, where $Z_{M/G}$ is an $\mathfrak{F}_0$-measurable M/G-valued random variable. Fix a $\mathfrak{F}_0$-measurable random variable Z in M with $\pi(Z) = Z_{M/G}$ and define a semimartingale $\Gamma \in \mathscr{S}(M)$ as follows:

1. Consider the horizontal lift Γ^h of Γ starting at Z.
2. Let ξ be the $\mathfrak{g}$-valued semimartingale defined by $\xi = \int A \circ S(X, \Gamma^h) \bullet dX$.
3. Define the Stratonovich operator $L \in \mathrm{Strat}(\mathfrak{g}, G)$ by $L(\eta, g) = T_e L_g(\eta)$ for all $(\eta, g) \in \mathfrak{g} \times G$ and let Ξ solve the Stratonovich stochastic differential equation

$$\bullet d\Xi = L(\xi, \Xi) \bullet d\xi, \quad \Xi_0 = e.$$

Set $\Gamma = \phi_{\Xi}(\Gamma^h)$

Then Γ satisfies

$$\bullet d\Gamma = S(X, \Gamma) \bullet dX, \quad \Gamma_0 = Z$$

and $\pi(\Gamma) = \Gamma^{M/G}$.

3.5.2 Reduction of Schwartz Operators

In the previous section, we discussed reduction of Stratonovich operators. This naturally leads to the question of reduction of Schwartz operators. For instance, let S be a G-invariant Stratonovich operator from N to M. Then one can obtain a reduced Stratonovich operator S^{red} on a reduced space corresponding to S. By Lemma 3.5.4, there are Schwartz operators F_S and $F_{S^{\mathrm{red}}}$ corresponding to S and S^{red} respectively. We would like to show that F and $F_{S^{\mathrm{red}}}$ are G-invariant and they are connected by reduction as well.

More generally, we derive a general reduction theorem for stochastic differential equations defined by Schwartz operators. We do not assume that the Schwartz operators are derived from Stratonovich operators by Lemma 3.5.4, although every Schwartz operator naturally defines a Stratonovich operator. If F is a Schwartz operator from N to M, then one can obtain a Stratonovich operator S from N to M by restricting F to the tangent bundles on N, namely, for each $(x, y) \in N \times M$, one defines $S(x, y) = F(x, y)|_{T_x N}$. However, F may not be equal to F_S - the Schwartz operator obtained from S via Lemma 3.5.4. We provide a concrete counterexample here:

Example 3.5.11. Define a Schwartz operator F from $\mathbb{R}^2$ to $\mathbb{R}^2$ as follows:

$$F(x, y)\left(\frac{\partial}{\partial x}\right) = \frac{\partial}{\partial x}$$

$$F(x, y)\left(\frac{\partial}{\partial y}\right) = 0$$

$$\begin{aligned}
F(x,y)\left(\frac{\partial^2}{\partial x^2}\right) &= \frac{\partial^2}{\partial x^2} \\
F(x,y)\left(\frac{\partial^2}{\partial y^2}\right) &= \frac{\partial}{\partial y} \\
F(x,y)\left(\frac{\partial^2}{\partial x \partial y}\right) &= 0.
\end{aligned}$$

We extend $F(x,y)$ to all of $\tau_{(x,y)}\mathbb{R}^2$ by linearity. Note that $F(x,y)$ satisfies the conditions for a Schwartz morphism, so F indeed is a well-defined Schwartz operator. The Stratonovich operator S obtained by restriction is given by

$$S(x,y)\left(a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}\right) = a\frac{\partial}{\partial x}$$

for all $a, b \in \mathbb{R}$. We shall now determine F_S. Suppose $(\mathbf{x}(t), \mathbf{y}(t))$ is a curve in $\mathbb{R}^2 \times \mathbb{R}^2$ such that

$$\dot{\mathbf{y}}(t) = S(\mathbf{x}(t), \mathbf{y}(t))(\dot{\mathbf{x}}(t)).$$

Let $\mathbf{x}(t) = (x^1(t), x^2(t))$ and $\mathbf{y}(t) = (y^1(t), y^2(t))$. Then

$$\dot{y}^1(t)\frac{\partial}{\partial x} + \dot{y}^2(t)\frac{\partial}{\partial y} = S(\mathbf{x}(t), \mathbf{y}(t))(\dot{\mathbf{x}}(t)) = \dot{x}^1(t)\frac{\partial}{\partial x}.$$

Therefore, $\dot{y}^1(t) = \dot{x}^1(t)$ and $\dot{y}^2(t) = 0$. Next, for any $f \in C^\infty(\mathbb{R}^2)$

$$\begin{aligned}
\frac{d^2}{dt^2}(f \circ \mathbf{y}(t)) &= \frac{d}{dt}\left[\dot{y}^1(t)\frac{\partial f}{\partial x}(\mathbf{y}(t)) + \dot{y}^2(t)\frac{\partial f}{\partial y}(\mathbf{y}(t))\right] \\
&= \ddot{y}^1(t)\frac{\partial f}{\partial x}(\mathbf{y}(t)) + \ddot{y}^2(t)\frac{\partial f}{\partial y}(\mathbf{y}(t)) \\
&+ \dot{y}^1(t)^2\frac{\partial^2 f}{\partial x^2} + \dot{y}^2(t)^2\frac{\partial^2 f}{\partial y^2} + 2\dot{y}^1(t)\dot{y}^2(t)\frac{\partial^2 f}{\partial x \partial y}.
\end{aligned}$$

Since $\dot{y}^1(t) = \dot{x}^1(t)$, $\dot{y}^2(t) = 0$ and f is arbitrary, we conclude that

$$\ddot{\mathbf{y}}(t) = \ddot{x}^1(t)\frac{\partial}{\partial x} + \dot{x}^1(t)^2\frac{\partial^2}{\partial x^2}.$$

Moreover, by a similar computation,

$$\begin{aligned}
\ddot{\mathbf{x}}(t) = \ddot{x}^1(t)\frac{\partial}{\partial x} &+ \ddot{x}^2(t)\frac{\partial}{\partial y} \\
&+ \dot{x}^1(t)^2\frac{\partial^2}{\partial x^2} + \dot{x}^2(t)^2\frac{\partial^2}{\partial y^2} \\
&+ 2\dot{x}^1(t)\dot{x}^2(t)\frac{\partial^2}{\partial x \partial y}.
\end{aligned}$$

Hence, by linearity of F_S

$$F_S(\mathbf{x}(t), \mathbf{y}(t))(\ddot{\mathbf{x}}(t)) = \ddot{x}^1(t)F_S\left(\frac{\partial}{\partial x}\right) + \ddot{x}^2(t)F_S\left(\frac{\partial}{\partial y}\right)$$

$$\begin{aligned}&+\dot{x}^1(t)^2F_S\left(\frac{\partial^2}{\partial x^2}\right)+\dot{x}^2(t)^2F_S\left(\frac{\partial^2}{\partial y^2}\right)\\&+2\dot{x}^1(t)\dot{x}^2(t)F_S\left(\frac{\partial^2}{\partial x\partial y}\right).\end{aligned}$$

Since the left side equals $\ddot{\mathbf{y}}(t)$ and the curve $\mathbf{x}(t)$ is arbitrary, this implies that

$$\begin{aligned}F_S(x,y)\left(\frac{\partial}{\partial x}\right)&=\frac{\partial}{\partial x}\\F_S(x,y)\left(\frac{\partial}{\partial y}\right)&=0\\F_S(x,y)\left(\frac{\partial^2}{\partial x^2}\right)&=\frac{\partial^2}{\partial x^2}\\F_S(x,y)\left(\frac{\partial^2}{\partial y^2}\right)&=0\\F_S(x,y)\left(\frac{\partial^2}{\partial x\partial y}\right)&=0.\end{aligned}$$

But note that $F_S(x,y)\left(\frac{\partial^2}{\partial y^2}\right)\neq F(x,y)\left(\frac{\partial^2}{\partial y^2}\right)$.

Given a Lie group G acting on M, the notion of a G-invariant Schwartz operator is defined very similarly to that of G-invariant Stratonovich operator. We shall continue to use the notations of the previous section for group actions:

Definition 3.5.12. Let F be a Schwartz operator from N to M. We say that F is **G-invariant** or **G-equivariant** if for every $(x,y)\in N\times M$ and $g\in G$

$$F(x,\phi_g(y))=\tau_y\phi_g\circ F(x,y).$$

G-invariant Schwartz operators yield symmetries in terms of the solutions of the SDEs they define. More precisely, we have the following proposition:

Proposition 3.5.13. Let F be a G-invariant Schwartz operator and $X\in\mathscr{S}(N)$. If Γ solves

$$d\Gamma=F(X,\Gamma)dX$$

then $\bar{\Gamma}:=\phi_g(\Gamma)$ solves

$$d\bar{\Gamma}=F(X,\bar{\Gamma})dX.$$

Proof. For every second order 1-form α

$$\begin{aligned}\int\alpha\,d\bar{\Gamma}&=\int\phi_g^*(\alpha)\,d\Gamma\\&=\int F^*(X,\Gamma)\phi_g^*(\alpha)\,dX\end{aligned}$$

$$= \int F^*(X, \bar{\Gamma})(\alpha)\, dX.$$

This shows that $\bar{\Gamma}$ solves

$$d\bar{\Gamma} = F(X, \bar{\Gamma})\, dX.$$

∎

G-invariant Schwartz operators are related naturally to G-invariant Stratonovich operators. The Stratonovich operator obtained by restricting a G-invariant Schwartz operator to the (first order) tangent spaces is G-invariant. Conversely, the Schwartz operator induced by a G-invariant Stratonovich operator is G-invariant. We prove both of these statements in the next proposition.

Proposition 3.5.14. Let $S \in \mathrm{Strat}(N, M)$ and F be a Schwartz operator from N to M.

1. Suppose F is G-invariant and for all $(x, y) \in N \times M$, $S(x, y) = F(x, y)|_{T_x N}$. Then S is G-invariant.
2. Suppose S is G-invariant. If $F = F_S$ then F is G-invariant.

Proof.

1. Let $(x, y) \in N \times M$. Then, for all $g \in G$

$$\begin{aligned} S(x, \phi_g(y)) &= F(x, \phi_g(y))|_{T_x N} \\ &= \tau_y \phi_g \circ F(x, y)|_{T_x N} \\ &= T_y \phi_g \circ F(x, y)|_{T_x N} \\ &= T_y \phi_g \circ S(x, y), \end{aligned}$$

   where the second last equality holds since the restriction of the second order tangent map $\tau_y \phi_g$ to the tangent space $T_y M$ is the usual (first order) tangent map $T_y \phi_g$.

2. Let $(x, y) \in N \times M$ and $g \in G$. Suppose $(x(t), y(t))$ is a curve in $N \times M$ with $(x(0), y(0)) = (x, y)$ and

$$\dot{y}(t) = S(x(t), y(t))(\dot{x}(t)).$$

   Then, by definition,

$$\ddot{y}(t) = F_S(x(t), y(t))(\ddot{x}(t)).$$

   By G-invariance of S, $\overline{y}(t) := \phi_g(y(t))$ solves

$$\dot{\overline{y}}(t) = S(x(t), \overline{y}(t))(\dot{x}(t)), \quad \overline{y}(0) = \phi_g(y).$$

   This implies that

$$\ddot{\overline{y}}(t) = F_S(x(t), \overline{y}(t))(\ddot{x}(t)).$$

   As a result

$$F_S(x(t), \phi_g(y(t)))(\ddot{x}(t)) = \frac{d^2}{dt^2} \phi_g(y(t))$$

$$
\begin{aligned}
&= \tau_{y(t)}\phi_g \left(\frac{d^2}{dt^2} y(t) \right) \\
&= \tau_{y(t)}\phi_g \circ F_S(x(t), y(t))(\ddot{x}(t)).
\end{aligned}
$$

Evaluating this at $t = 0$ gives

$$F_S(x, \phi_g(y))(\ddot{x}(0)) = \tau_y\phi_g \circ F_S(x, y)(\ddot{x}(0)).$$

Since the accelerations $\ddot{x}(0)$ of curve $x(t)$ with $x(0) = x$ linearly span $\tau_x N$, we conclude that

$$F_S(x, \phi_g(y)) = \tau_y\phi_g \circ F_S(x, y)$$

for all $(x, y) \in N \times M$ and $g \in G$. This completes the proof.

∎

Similar to the case of Stratonovich equations, we can define invariant manifolds for SDEs defined by Schwartz operators.

Definition 3.5.15. Suppose F is a Schwartz operator from N to M and $X \in \mathscr{S}(N)$. Let P be an immersed submanifold in M. P is said to be an **invariant submanifold** for the stochastic differential equation

$$d\Gamma = F(X, \Gamma)dX$$

if for every $\mathfrak{F}_0$-measurable random variable Z taking values in P, $\Gamma_0 = Z$ implies that Γ takes values in P in $[[0, \tau[[$, where τ is the explosion time of Γ.

The next result, proven in Emery [36], allows us to prove that a submanifold P of M is invariant.

Proposition 3.5.16 ([36, Corollary 6.44])**.** Let F be a Schwartz operator from N to M and $X \in \mathscr{S}(N)$. Suppose that for every $(x, y) \in N \times P$ and $L \in \tau_x N$, $F(x, y)(L) \in \tau_y P$. Then P is an invariant submanifold for the stochastic differential equation

$$d\Gamma = F(X, \Gamma)dX.$$

We will use this proposition to prove the analogue of Proposition 3.5.8 for Schwartz operators. Namely, we will show that the isotropy type submanifolds are invariant submanifolds for G-invariant Schwartz operators.

Proposition 3.5.17 (Law of Conservation of the Isotropy)**.** Assume the G-action on M is proper and let F be a G-invariant Schwartz operator from N to M. Then, given any $X \in \mathscr{S}(N)$, the isotropy type submanifolds are invariant manifolds for the stochastic differential equation $d\Gamma = F(X, \Gamma)dX$.

To prove this, we need the following lemma:

Lemma 3.5.18. Let F be a Schwartz operator from N to M. Define a Stratonovich operator $S \in \mathrm{Strat}(N, M)$ by

$$S(x, y) = F(x, y)|_{T_xN}, \text{for all } (x, y) \in N \times M.$$

Let F_S be the Schwartz operator from N to M obtained from S via Lemma 3.5.4. Then, for every $(x, y) \in N \times M$ and $L \in \tau_x N$,

$$(F(x, y) - F_S(x, y))(L) \in T_y M.$$

Proof. Let $(x, y) \in N \times M$. Then

$$(F(x, y) - F_S(x, y))|_{T_xN} = 0.$$

Hence, given any $L \in \tau_x N$,

$$\begin{aligned}&P_2((F(x, y) - F_S(x, y))(L))\\&= ((F(x, y) - F_S(x, y))|_{T_xN} \odot (F(x, y) - F_S(x, y))|_{T_xN})(P_2(L))\\&= 0.\end{aligned}$$

Since the kernel of $P_2|_{\tau_y M} = T_y M$, the result follows. ∎

Proof of Proposition 3.5.17. Let I be an isotropy subgroup of G. By Ortega and Ratiu [78, Proposition 2.4.7], for all $y \in M_I$

$$T_y M_I = \{v \in T_y M \mid T_y\phi_g(v) = v \text{ for all } g \in I.\}$$

Define a Stratonovich operator $S \in \mathrm{Strat}(N, M)$ by

$$S(x, y) = F(x, y)|_{T_xN}, \text{for all } (x, y) \in N \times M,$$

and let F_S be the Schwartz operator from N to M obtained via Lemma 3.5.4. Then S, F_S and F are all G-invariant. For all $(x, y) \in N \times M_I$, let

$$V(x, y) = F(x, y) - F_S(x, y).$$

Then $F(x, y) = F_S(x, y) + V(x, y)$. For every $g \in I$, since the image of $V(x, y)$ is a first order tangent vector and $\phi_g(y) = y$

$$\begin{aligned}T_y\phi_g \circ V(x, y) &= \tau_y\phi_g \circ V(x, y)\\&= \tau_y\phi_g \circ F(x, y) - \tau_y\phi_g \circ F_S(x, y)\\&= F(x, \phi_g(y)) - F_S(x, \phi_g(y))\\&= V(x, y).\end{aligned}$$

Therefore, $V(x, y)$ takes values in $T_y M_I$. Next, we show that $F_S(x, y)$ takes values in $\tau_y M_I$. Let $(x(t), y(t))$ be a curve in $N \times M$ with $(x(0), y(0)) = (x, y) \in N \times M_I$ and

$$\dot{y}(t) = S(x(t), y(t))(\dot{x}(t)).$$

Applying Proposition 3.5.8 with $X_t = x(t)$ shows that M_I is an invariant submanifold for this ordinary differential equation. Hence the curve $y(t)$ takes its values in M_I. Consequently $\ddot{y}(t) \in \tau_{y(t)}M_I$. On the other hand, by definition of F_S,

$$\ddot{y}(t) = F_S(x(t), y(t))(\ddot{x}(t)).$$

Evaluating this at $t = 0$ shows that $F_S(x, y)(\ddot{x}(0)) \in \tau_y M_I$. Since the accelerations $\ddot{x}(0)$ of curve $x(t)$ with $x(0) = x$ linearly span $\tau_x N$, this implies that $F_S(x, y)$ takes values in $\tau_y M_I$.

Since, for every $(x, y) \in N \times M_I$, $F(x, y) = F_S(x, y) + V(x, y) \in \tau_y M_I$, by Proposition 3.5.16 it follows that M_I is an invariant submanifold for $d\Gamma = F(X, \Gamma)\, dX$ for every $X \in \mathscr{S}(N)$. ∎

Now we are ready to prove the reduction theorem for SDEs defined by Schwartz operators.

Theorem 3.5.19 (Reduction Theorem for Schwartz Operators)**.** Let $X \in \mathscr{S}(N)$ be any semimartingale and F be a G-invariant Schwartz operator from N to M. Let I be any isotropy subgroup in G, M_I be its isotropy type submanifold, $N(I)$ be its normalizer and $L_I := N(I)/I$. Then L_I acts freely and properly on M_I. Moreover, let $\pi_I : M \to M_I/L_I$ denote the projection and F_I be a Schwartz operator from N to M_I/L_I defined by $F_I(x, \pi_I(y)) = \tau_y \pi_I \circ F(x, y)$, for all $(x, y) \in N \times M_I$. If Γ solves $d\Gamma = F(X, \Gamma)dX$ and Γ_0 takes values in M_I, then $\Gamma^{\mathrm{red}} := \pi_I(\Gamma)$ satisfies $d\Gamma^{\mathrm{red}} = F_I(X, \Gamma^{\mathrm{red}})dX$ with initial condition $\pi_I(\Gamma_0)$.

Proof. We remark that the proof is very similar to the proof of the reduction theorem in Lázaro-Camí and Ortega [54]. Since L_I acts freely and properly on M_I, M_I/L_I is a regular quotient manifold. This is proven, for instance, in Duistermaat and Kolk [35].

Next we show that the Schwartz operator F_I is well-defined. Suppose $(x, y) \in N \times M_I$ and $\bar{y} = \phi_g(y)$, with $g \in L_I$. Then

$$\begin{aligned} F_I(x, \bar{y}) &= \tau_{\bar{y}} \pi_I \circ F(x, \bar{y}) \\ &= \tau_{\phi_g(y)} \pi_I \circ \tau_y \phi_g \circ F(x, y) \\ &= \tau_y (\pi_I \circ \phi_g) \circ F(x, y) \\ &= \tau_y \pi_I \circ F(x, y) \\ &= F_I(x, y). \end{aligned}$$

Suppose Γ solves $d\Gamma = F(X, \Gamma)\, dX$ and Γ_0 takes values in M_I. Since M_I is an invariant submanifold it follows that Γ takes its values in M_I. Therefore $\Gamma^{\mathrm{red}} = \pi_I(\Gamma)$ is well-defined at $\Gamma_0^{\mathrm{red}} = \pi_I(\Gamma_0)$. Moreover, for all second order 1-forms α on M_I/L_I

$$\begin{aligned} \int \alpha\, d\Gamma^{\mathrm{red}} &= \int \pi_I^*(\alpha)\, d\Gamma \\ &= \int F^*(X, \Gamma)(\pi_I^*(\alpha))\, dX \end{aligned}$$

$$
\begin{aligned}
&= \int F_I^*(X, \pi_I(\Gamma))(\alpha)\, dX \\
&= \int F_I^*(X, \Gamma^{\text{red}})(\alpha)\, dX.
\end{aligned}
$$

This shows that Γ^{red} solves the SDE $d\Gamma^{\text{red}} = F_I(X, \Gamma^{\text{red}})dX$. ∎

Finally, we show that the Schwartz operators corresponding to a G-invariant Stratonovich operator S and its reduced counterpart S^{red} are connected by reduction.

Theorem 3.5.20. Let $S \in \text{Strat}(N, M)$ be G-invariant, I denote an isotropy subgroup of G and $L_I = N(I)/I$. Let $S_I \in \text{Strat}(N, M_I/L_I)$ be the reduced Stratonovich operator as constructed in Theorem 3.5.9. Suppose F_S (respectively, F_{S_I}) is the Schwartz operator from N to M (respectively, M_I/L_I) corresponding to S (respectively, S_I) as obtained via Lemma 3.5.4. Then, $F_{S_I}(x, \pi_I(y)) = \tau_y \pi_I \circ F_S(x, y)$ for all $(x, y) \in N \times M_I$.

Proof. Let $(x, y) \in N \times M_I$ and suppose $(x(t), y(t))$ is a curve in $N \times M$ with $(x(0), y(0)) = (x, y)$ and

$$\dot{y}(t) = S(x(t), y(t))(\dot{x}(t)).$$

Since M_I is an invariant submanifold for this differential equation, it follows that $y(t)$ takes its values in M_I. Moreover, by the reduction theorem for Stratonovich operators, $y^{\text{red}}(t) := \pi_I(y(t))$ solves

$$\dot{y}^{\text{red}}(t) = S_I(x(t), y^{\text{red}}(t))(\dot{x}(t)), \quad y^{\text{red}}(0) = \pi_I(y).$$

On the other hand, we have $\ddot{y}(t) = F_S(x(t), y(t))(\ddot{x}(t))$ and $\ddot{y}^{\text{red}}(t) = F_{S_I}(x(t), y^{\text{red}}(t))(\ddot{x}(t))$. Since $\ddot{y}^{\text{red}}(t) = \tau_{y(t)}\pi_I(\ddot{y}(t))$, it follows that

$$F_{S_I}(x(t), y^{\text{red}}(t))(\ddot{x}(t)) = \tau_{y(t)}\pi_I \circ F_S(x(t), y(t))(\ddot{x}(t)).$$

Evaluating this at $t = 0$ gives

$$F_{S_I}(x, \pi_I(y))(\ddot{x}(0)) = \tau_y \pi_I \circ F_S(x, y)(\ddot{x}(0)),$$

which implies that

$$F_{S_I}(x, \pi_I(y)) = \tau_y \pi_I \circ F_S(x, y).$$

This completes the proof. ∎

Remark 3.5.21. We do not pursue the reconstruction for Schwartz operators here. This is left as a prospect for future research.

Remark 3.5.22. Huang and Zambrini [44] has developed a Hamiltonian formulation of stochastic mechanics using generalized derivatives and second order differential geometry. In their setting, the canonical symplectic form on T^*Q is extended to a 2-form-type object ω on the second-order cotangent bundle τ^*Q. For a smooth Hamiltonian $H \in C^\infty(\tau^*Q)$, they define a second-order vector field A_H on τ^*Q via the relation

$$i_{A_H}\omega = d_2 H.$$

The uniqueness of A_H is established under additional hypotheses. Given such a second-order field and a choice of Riemannian metric on Q, they then construct an associated Itô equation via a martingale problem formulation.

As shown in Emery [37], an Itô equation on a manifold canonically determines a Schwartz operator. If the metric on Q is G-invariant, the complete lift of the corresponding Levi-Civita connection described by Yano and Patterson [91] yields a G-invariant connection on T^*Q. If the Itô equation resulting from the martingale problem in [44] is G-invariant, its associated Schwartz operator can also be shown to be G-invariant. Consequently, the reduction of G-invariant Schwartz operators carried out in this section also applies in this setting.

Additionally, symmetries of stochastic differential equations under simultaneous transformations of the phase space variables and the time parameter are considered in [44], by employing the prolongation of SDEs to an appropriate stochastic tangent bundle. In their setup, the driving noise is of the form $(t, B^1, \ldots, B^k)$, where each B^i is a Brownian motion. In contrast, the present work focuses on symmetries acting purely on the phase space with driving noises which are arbitrary semimartingales X taking values in a manifold N. Moreover, every symmetry transformations considered in this section are also symmetry transformations in the sense of [44], but the converse is not necessarily true.

Part II

Stochastic Hamiltonian Systems

Chapter 4

Stratonovich Formulation of Stochastic Hamiltonian Systems

In this chapter we will look at the Hamiltonian formulation of stochastic geometric mechanics using Stratonovich differential equations. Stratonovich differential equations allow us to transfer many of the geometric properties of Hamiltonian systems to the stochastic case.

As in the case of deterministic geometric mechanics, the phase space for a stochastic Hamiltonian system is a Poisson manifold. In the deterministic case, one starts with a Hamiltonian h on a Poisson manifold and the central object of study is the Hamiltonian vector field. In the stochastic case, the goal is to study stochastic perturbations of X_h by noise vector fields. This is done by replacing X_h by a Stratonovich operator. In order to ensure that the stochastic flows are Hamiltonian in nature, we consider Stratonovich operators given by a sum of Hamiltonian vector fields. In this sense, stochastic geometric mechanics can be thought of as the study of vector-valued Hamiltonians on a Poisson manifold.

Stochastic perturbations of Hamiltonian systems via Stratonovich differential equations was studied by Bismut [11] on Euclidean spaces and generalized by Lázaro-Camí and Ortega [55, 54] to Poisson manifolds. Following [55, 54], we introduce the general framework of stochastic Hamiltonian systems and discuss the analogues of Poisson, symplectic, and Lie-Poisson reduction for stochastic Hamiltonian systems.

4.1 Definitions and Symmetries of Stochastic Hamiltonian Systems

Definition 4.1.1. Let $(P, \{\cdot, \cdot\})$ be a Poisson manifold, and V be a vector space with basis $\{e_1, \cdots, e_k\}$. Suppose $Y \in \mathscr{S}(V)$ and $H : P \to V^*$ is a smooth function with component functions $(h_1, \cdots, h_k)$ with respect to a basis $\{\epsilon^1, \cdots, \epsilon^k\}$ for V^*. Define a Stratonovich operator $S_H \in \operatorname{Strat}(V, P)$ by

$$S_H(v,z)(u) = \sum_{i=1}^{k} \left\langle \epsilon^i, u \right\rangle X_{h_i}(z), \quad v \in V,\ z \in P,\ u \in T_vV. \tag{4.1.1}$$

The **Hamilton equations with stochastic component** Y, and **Hamiltonian function** H are the Stratonovich differential equations associated with S_H:

$$\bullet d\Gamma = S_H(Y, \Gamma) \bullet dY. \tag{4.1.2}$$

The 4-tuple $(P, \{\cdot,\cdot\}, H, Y)$ will be called a **stochastic Hamiltonian system**.

For the sake of brevity, we will write Equation (4.1.2) as

$$\bullet d\Gamma = \sum_{i=1}^{k} X_{h_i}(\Gamma) \bullet dY^i. \tag{4.1.3}$$

Often, we think of this equation as stochastically perturbing a Hamiltonian vector field X_h by noisy Hamiltonian vector fields. More concretely, let $V = \mathbb{R}^{k+1}$ with its standard basis, $Y = (t, Y^1, \cdots, Y^k)$ where $Y^1, \cdots, Y^k$ are real-valued semimartingales and $H = (h, h_1, \cdots, h_k) : P \to (\mathbb{R}^{k+1})^*$. Then the corresponding stochastic Hamiltonian system is written as

$$\bullet d\Gamma = X_h(\Gamma) dt + \sum_{i=1}^{k} X_{h_i}(\Gamma) \bullet dY^i.$$

As a consequence of the definition of a stochastic Hamiltonian system we have the following proposition:

Theorem 4.1.2. Let α be a 1-form on P, Γ solve Equation (4.1.3) and τ be the explosion time of Γ. Then, on $[[0, \tau[[$

$$\int \alpha \bullet d\Gamma = \sum_{i=1}^{k} \int \left\langle \alpha, X_{h_i} \right\rangle (\Gamma) \bullet dY^i. \tag{4.1.4}$$

Let $\alpha = df$, where $f \in C^\infty(P)$. Since $\langle df, X_{h_i} \rangle = X_{h_i}[f] = \{f, h_i\}$, we obtain

$$f(\Gamma) - f(\Gamma_0) = \sum_{i=1}^{k} \int \{f, h_i\} (\Gamma) \bullet dY^i. \tag{4.1.5}$$

The next proposition describes the Schwartz operator for Equation (4.1.3). We will, however, restrict ourselves to the Stratonovich description in most applications.

Proposition 4.1.3 ([55, Proposition 2.3])**.** The Schwartz operator $F_{S_H}(v,p) : \tau_vV \to \tau_pP$ associated to S_H is described as follows: Given $L \in \tau_vV$ and $f \in C^\infty(M)$

$$F_{S_H}(v,p)(L)(f) = \left\langle \sum_{i,j=1}^{k} \{f, h_j\}(p)\epsilon^j + \{\{f, h_j\}, h_i\}(p) P_2^*(\epsilon^i \odot \epsilon^j), L \right\rangle.$$

Now we define conserved quantities for stochastic Hamiltonian systems.

Definition 4.1.4. Let $f \in C^\infty(P)$.

1. We say that f is a **conserved quantity** for the stochastic Hamiltonian system (4.1.3) if for any solution Γ, $f(\Gamma) - f(\Gamma_0) = 0$ almost surely.
2. We say that f is a **weakly conserved quantity** for the stochastic Hamiltonian system (4.1.3) if for any solution Γ, $\mathbb{E}[f(\Gamma)] - \mathbb{E}[f(\Gamma_0)] = 0$.

From Equation (4.1.5) we see that f is conserved along Γ if $\{f, h_i\} = 0$ for all $i = 1, \cdots, k$. In particular, Casimirs are always conserved along solutions of stochastic Hamiltonian systems.

On the other hand, conservation of $f \in C^\infty(P)$ along solutions does not automatically imply that $\{f, h_i\} = 0$ for all $i = 1, \cdots, k$. A partial converse is provided by the following proposition:

Proposition 4.1.5 ([54, Proposition 2.6])**.** Let f be conserved along a solution Γ of Equation (4.1.3). Suppose there exists $j \in \{1, \cdots, k\}$ with the following property: there exists $A \in \mathfrak{F}$ and $\delta > 0$ such that $\boldsymbol{\mathcal{P}}(A) > 0$ and $[Y^j, Y^j]_t(\omega) > [Y^j, Y^j]_0(\omega)$ for all $t < \delta$ and $\omega \in A$. Then $\{f, h_j\} = 0$.

4.1.1 Reduction of Symmetric Stochastic Hamiltonian Systems

Now we turn to the problem of reduction of symmetric stochastic Hamiltonian system. Suppose G is a Lie group acting canonically on P. Lázaro-Camí and Ortega [54] prove the following theorem:

Theorem 4.1.6. Let G be a Lie group acting canonically on P. If h_i is G-invariant for all $i = 1, \cdots, k$ then the Stratonovich operator S_H introduced in Definition 4.1.1 is G-invariant.

Suppose the G-action on P admits a momentum map $J : P \to \mathfrak{g}^*$. Then Noether's theorem also carries over to the stochastic Hamiltonian case.

Theorem 4.1.7 (Stochastic Noether's Theorem, [54, Proposition 6.4])**.** The level sets of J are invariant submanifolds of Equation (4.1.3). Moreover, its components are conserved quantities.

Next, we discuss stochastic analogues of Poisson, symplectic, and Lie-Poisson reduction. We will use the notations of deterministic Hamiltonian reduction introduced in Section 2.3. The proofs may be found in [54].

Theorem 4.1.8. Let V be a vector space with basis $\{e_1, \cdots, e_k\}$. Suppose $Y \in \mathscr{S}(V)$ and $H : P \to V^*$ is a smooth function with component functions $(h_1, \cdots, h_k)$ with respect to a basis $\{\epsilon^1, \cdots, \epsilon^k\}$ for V^*. Suppose G is a Lie group acting freely, properly, and canonically on P and h_i is G-invariant for all $i = 1, \cdots, k$.

1. **Stochastic Poisson reduction**: Let $\pi : P \to P/G$ denote the projection and $\{\cdot,\cdot\}^{\mathrm{red}}$ denote the reduced Poisson bracket on P/G. For all $i = 1, \cdots, k$, denote by $h_i^{\mathrm{red}} \in C^\infty(P/G)$ the reduced Hamiltonian corresponding to h_i, that is, $h_i^{\mathrm{red}} \circ \pi = h_i$ and define $H^{\mathrm{red}} : P/G \to V^*$ by $H^{\mathrm{red}} = \sum_{i=1}^k h_i^{\mathrm{red}} \epsilon^i$. Let $S_{H^{\mathrm{red}}} \in \mathrm{Strat}(V, P/G)$ denote the Stratonovich operator given by

$$S_{H^{\mathrm{red}}}(v,z)(u) = \sum_{i=1}^k \left\langle \epsilon^i, u \right\rangle X^{P/G}_{h_i^{\mathrm{red}}}(z),$$

   for all $(v,z) \in V \times P/G$ and $u \in T_vV$. Here $X^{P/G}_{h_i^{\mathrm{red}}}(z)$ is the Hamiltonian vector field of h_i^{red} on P/G.

   Then, $S_{H^{\mathrm{red}}}$ is the reduced Stratonovich operator corresponding to S_H in the sense of Theorem 3.5.9. Moreover, if Γ is a solution of Equation (4.1.2) with initial condition Γ_0 then $\Gamma^{\mathrm{red}} := \pi(\Gamma)$ is a solution of $\bullet d\Gamma^{\mathrm{red}} = S_{H^{\mathrm{red}}}(Y, \Gamma^{\mathrm{red}}) \bullet dY$ with initial condition $\pi(\Gamma_0)$.

2. **Stochastic Symplectic reduction**: Assume that (P, Ω) is a symplectic manifold, and the G-action on P admits a coadjoint equivariant momentum map $J : P \to \mathfrak{g}^*$. Suppose μ is a regular value of J and let G_μ denote the coadjoint isotropy subgroup through μ. Let P_μ denote the symplectic quotient $J^{-1}(\mu)/G_\mu$ with reduced symplectic form Ω_μ. Denote by $i_\mu : J^{-1}(\mu) \hookrightarrow P$ the inclusion and $\pi_\mu : J^{-1}(\mu) \to P_\mu$ denote the projection. Define $h_i^{\mathrm{red}} \in C^\infty(P_\mu)$ via $h_i^{\mathrm{red}} \circ \pi_\mu = h_i \circ i_\mu$ for all $i = 1, \cdots, k$. Let $H^{\mathrm{red}} : P_\mu \to V^*$ by $H^{\mathrm{red}} = \sum_{i=1}^k h_i^{\mathrm{red}} \epsilon^i$. Let $S_{H^{\mathrm{red}}} \in \mathrm{Strat}(V, P_\mu)$ denote the Stratonovich operator given by

$$S_{H^{\mathrm{red}}}(v,z)(u) = \sum_{i=1}^k \left\langle \epsilon^i, u \right\rangle X^{P_\mu}_{h_i^{\mathrm{red}}}(z),$$

   for all $(v,z) \in V \times P_\mu$ and $u \in T_vV$. Here $X^{P_\mu}_{h_i^{\mathrm{red}}}(z)$ is the Hamiltonian vector field of h_i^{red} on P_μ.

   Then, $S_{H^{\mathrm{red}}}$ is the reduced Stratonovich operator corresponding to S_H in the sense of Theorem 3.5.9. Moreover, if Γ is a solution of Equation (4.1.2) with initial condition Γ_0 such that $J(\Gamma_0) = \mu$, then $\Gamma^{\mathrm{red}} := \pi_\mu(\Gamma)$ is a solution of $\bullet d\Gamma^{\mathrm{red}} = S_{H^{\mathrm{red}}}(Y, \Gamma^{\mathrm{red}}) \bullet dY$ with initial condition $\pi_\mu(\Gamma_0)$.

3. **Stochastic Lie-Poisson Reduction**: Let $P = T^*G$ with its canonical Poisson bracket. Assume that h_i is invariant with respect to the cotangent lifts of left translations. Denote by $J_R : T^*G \to \mathfrak{g}^*$ the momentum map for the cotangent lifted right action of G on T^*G. Let $h_i^{\mathrm{red}} \in C^\infty(\mathfrak{g}^*_-)$ be defined by $h_i^{\mathrm{red}} \circ J_R = h_i$. Define $H^{\mathrm{red}} : \mathfrak{g}^* \to V^*$ by $H^{\mathrm{red}} = \sum_{i=1}^k h_i^{\mathrm{red}} \epsilon^i$. Let $S_{H^{\mathrm{red}}} \in \mathrm{Strat}(V, \mathfrak{g}^*)$ denote the Stratonovich operator given by

$$S_{H^{\mathrm{red}}}(v,\mu)(u) = \sum_{i=1}^k \left\langle \epsilon^i, u \right\rangle \mathrm{ad}^*_{\frac{\delta h_i^{\mathrm{red}}}{\delta \mu}} \mu,$$

for all $(v, \mu) \in V \times \mathfrak{g}^*$ and $u \in T_v V$.

Then, $S_{H^{\mathrm{red}}}$ is the reduced Stratonovich operator corresponding to S_H in the sense of Theorem 3.5.9. Moreover, if Γ is a solution of Equation (4.1.2) with initial condition Γ_0 then $\Gamma^{\mathrm{red}} := J_R(\Gamma)$ is a solution of $\bullet d\Gamma^{\mathrm{red}} = S_{H^{\mathrm{red}}}(Y, \Gamma^{\mathrm{red}}) \bullet dY$ with initial condition $J_R(\Gamma_0)$.

Remark 4.1.9. The case of Lie-Poisson reduction for the cotangent lifts of right actions is similar. The definition of $S_{H^{\mathrm{red}}}$ must be modified to

$$S_{H^{\mathrm{red}}}(v, \mu)(u) = -\sum_{i=1}^{k} \left\langle \epsilon^i, u \right\rangle \mathrm{ad}^*_{\frac{\delta h_i^{\mathrm{red}}}{\delta \mu}} \mu.$$

The stochastic reconstruction theorem 3.5.10 can be used to reconstruct the solution of S_H starting from a solution of $S_{H^{\mathrm{red}}}$.

Chapter 5

Kepler Problem with Perturbed Angular Momentum

In this chapter we consider a concrete example of a stochastically perturbed Kepler problem with noise affecting the angular momentum vector. Specifically, the Hamiltonian vector field of the Kepler problem is perturbed by noisy Hamiltonian vector fields arising out of the components of the angular momentum vector. While stochastic mechanical systems with symmetries have been well-studied (see, for instance, Lázaro-Camí and Ortega [55, 54], Arnaudon, de Castro, and Holm [3, 4], and Street and Takao [89]), this perturbation destroys the symmetries of the Kepler problem (the angular momentum and the Laplace Runge-Lenz vector). Therefore, the reduction and reconstruction methods mentioned in Section 4.1 do not apply.

On the other hand, even though the symmetries of the Kepler problem are no longer conserved, the radial distance of the particle from the gravitational source as well as its speed behaves deterministically. Another feature is that the momentum vector is always at a constant distance away from a stochastic process on a sphere, providing an analogue of the deterministic Hamilton's theorem (see Milnor [69]).

Given the deterministic behaviour of the distance and speed, collisions are well-defined. In case of the deterministic Kepler problem, Moser regularization transforms Keplerian orbits at a fixed negative energy level to orbits of the geodesic flow on the 3-sphere (see Moser [70], Cushman and Bates [30], and Heckmann and de Laat [41]). In our problem, we are able to carry out a transformation similar to Moser's, whereby orbits for a fixed negative energy level are transformed to the geodesic flow on the 3-sphere. Moreover, the transformation is 'noise-preserving' in the sense that the structure of the noise perturbation remains unchanged after applying the Moser map.

For other stochastic versions of the Kepler problem, we refer to Albeverio, Blanchard, and Høegh-Krohn [1, 2], Nottale [76], Nottale, Schumacher, and Gay [77], Cresson [25], Cresson and Sébastien [26], Cresson, Pierret, and Puig [28], Cresson, Nottale and Lehner [27].

5.1 The Equations of Motion

Let B^1, B^2 and B^3 be independent Brownian motions. Let $h : T^*(\mathbb{R}^3 \setminus \{0\}) \to \mathbb{R}$ denote the Hamiltonian of the Kepler problem

$$h = \frac{||\mathbf{p}||^2}{2} - \frac{1}{||\mathbf{q}||}$$

and

$$\mathbf{J} = (J^1, J^2, J^3) = (q^2 p_3 - q^3 p_2, q^3 p_1 - q^1 p_3, q^1 p_2 - q^2 p_1)$$

denote the angular momentum vector. We also assume positive constants ν_1, ν_2 and ν_3 to take into account the intensity of perturbations along different angular momentum components. We consider the stochastic Hamiltonian system:

$$\begin{aligned} \bullet d\Gamma &= X_h(\Gamma)dt + \sum_{i=1}^{3} X_{\nu_i J^i}(\Gamma) \bullet dB^i \\ &= X_h(\Gamma)dt + \sum_{i=1}^{3} X_{J^i}(\Gamma) \bullet d(\nu_i B^i). \end{aligned} \tag{5.1.1}$$

Since h commutes with the components of the angular momentum, we note that h is a conserved quantity for this problem. Let $\mathbf{B}$ denote the process $(\nu_1 B^1, \nu_2 B^2, \nu_3 B^3)$ in $\mathbb{R}^3$. Let ϵ_{ijk} denote the Levi-Civita tensor. Suppose $\mathbf{Z} = (Z^1, Z^2, Z^3)$ satisfies $\{Z^i, J^j\} = \sum_{k=1}^{3} \epsilon_{ijk} Z^k$ and $\{Z^i, h\} = 0$ for all $i, j \in \{1, 2, 3\}$. Then

$$\begin{aligned} Z^1(\Gamma) - Z^1(\Gamma_0) &= \nu_2 \int_0^t \{Z^1, J^2\}(\Gamma) \bullet dB^2 + \nu_3 \int_0^t \{Z^1, J^3\}(\Gamma) \bullet ddB^3 \\ &= \nu_2 \int_0^t Z^3(\Gamma) \bullet dB^2 - \nu_3 \int_0^t Z^2(\Gamma) \bullet dB^3 \end{aligned}$$

which implies that

$$\bullet dZ^1(\Gamma) = \nu_2 Z^3(\Gamma) \bullet dB^2 - \nu_3 Z^2(\Gamma) \bullet dB^3.$$

Using a similar calculation for $Z^2(\Gamma)$ and $Z^3(\Gamma)$, we obtain

$$\bullet d\mathbf{Z}(\Gamma) = -\mathbf{Z}(\Gamma) \times \bullet d\mathbf{B}\,.$$

Therefore

$$\bullet d||\mathbf{Z}(\Gamma_t)||^2 = 2\mathbf{Z}(\Gamma_t) \cdot \bullet d\mathbf{Z}(\Gamma_t) = 0\,, \tag{5.1.2}$$

which shows that $||\mathbf{Z}||$ is conserved along solutions and hence $\mathbf{Z}(\Gamma)$ lies on a sphere. Note that $\mathbf{Z}$ is not conserved in general. Letting $\mathbf{Z} = \mathbf{J}$ and $\mathbf{Z} = \mathbf{A}$, where $\mathbf{A} = \mathbf{p} \times \mathbf{J} - \frac{\mathbf{q}}{||\mathbf{q}||}$ is the Runge-Lenz vector, we see that $||\mathbf{J}||$ and $||\mathbf{A}||$ are conserved along solutions.

The Stratonovich equations for $\mathbf{q}(\Gamma)$ and $\mathbf{p}(\Gamma)$ are given by

$$\bullet d\mathbf{q}(\Gamma) = \mathbf{p}(\Gamma)dt - \mathbf{q}(\Gamma) \times \bullet d\mathbf{B}\,, \tag{5.1.3}$$

$$\bullet d\mathbf{p}(\Gamma) = -\frac{1}{||\mathbf{q}(\Gamma)||^3}\mathbf{q}(\Gamma)dt - \mathbf{p}(\Gamma) \times \bullet d\mathbf{B}\,. \tag{5.1.4}$$

The evolution of $||\mathbf{q}||(\Gamma)$ is given by

$$||\mathbf{q}||(\Gamma) - ||\mathbf{q}||(\Gamma_0) = \int_0^t \frac{(\mathbf{q}\cdot\mathbf{p})}{||\mathbf{q}||}(\Gamma_s)ds\,.$$

By the Fundamental Theorem of Calculus, the left side of this equation is differentiable almost surely and we have,

$$\frac{d}{dt}||\mathbf{q}||(\Gamma) = \frac{(\mathbf{q}\cdot\mathbf{p})}{||\mathbf{q}||}(\Gamma)\,. \tag{5.1.5}$$

On the other hand, using $\mathbf{q}(\Gamma)\cdot(\mathbf{p}(\Gamma)\times\bullet d\mathbf{B}) = -\mathbf{p}(\Gamma)\cdot(\mathbf{q}(\Gamma)\times\bullet d\mathbf{B})$, we have

$$\begin{aligned}\bullet d(\mathbf{q}\cdot\mathbf{p})(\Gamma) &= \mathbf{p}(\Gamma)\cdot\bullet d\mathbf{q}(\Gamma) + \bullet d\mathbf{p}(\Gamma_t)\cdot\mathbf{q}(\Gamma_t)\\ &= \left(||\mathbf{p}||^2(\Gamma) - \frac{1}{||\mathbf{q}||(\Gamma)}\right)dt\end{aligned}$$

which shows that $(\mathbf{q}\cdot\mathbf{p})(\Gamma)$ is differentiable almost surely. This means that the right side of Equation (5.1.5) is differentiable. Using Lagrange's identity $||\mathbf{J}||^2 = ||\mathbf{q}||^2||\mathbf{p}||^2 - (\mathbf{q}\cdot\mathbf{p})^2$ and setting $||\mathbf{J}|| = J =$ constant, we get

$$\begin{aligned}\frac{d^2}{dt^2}||\mathbf{q}||(\Gamma_t) &= -\left(\frac{1}{||\mathbf{q}||^3}(\mathbf{q}\cdot\mathbf{p})^2\right)(\Gamma_t) + \frac{1}{||\mathbf{q}||}\left(||\mathbf{p}||^2 - \frac{1}{||\mathbf{q}||}\right)(\Gamma_t)\\ &= \frac{J^2}{||\mathbf{q}||^3(\Gamma_t)} - \frac{1}{||\mathbf{q}||^2(\Gamma_t)}.\end{aligned} \tag{5.1.6}$$

A similar calculation with the deterministic Kepler problem $\dot{\gamma}(t) = X_h(\gamma(t))$ shows that $||\mathbf{q}||(\gamma(t))$ also satisfies Equation (5.1.6). Consequently, for almost every $\omega \in \mathbf{\Omega}$,

$$||\mathbf{q}||(\Gamma_t)(\omega) = ||\mathbf{q}||(\gamma(t))\,. \tag{5.1.7}$$

In conclusion, we obtain that $||\mathbf{q}||(\Gamma)$ behaves deterministically. Also, since h is conserved along solutions we conclude that $||\mathbf{p}||(\Gamma)$ behaves deterministically as well.

5.1.1 Hamilton's velocity vector circle theorem

Consider now

$$\mathbf{X} := \frac{\mathbf{A}\times\mathbf{J}}{||\mathbf{J}||^2}\,.$$

In case of the deterministic Kepler problem, by direct computation we can show that

$$\frac{d}{dt}||\mathbf{p}(\gamma(t)) - \mathbf{X}(\gamma(t))|| = 0\,.$$

along any solution $\gamma(t)$ and so $\mathbf{X}$ is the center of the momentum hodographs (see Milnor [69]). In case of our stochastic Kepler problem we have

$$\begin{aligned}
\bullet d\mathbf{X}(\Gamma) &= \frac{1}{J^2}\left(\bullet d\mathbf{A}(\Gamma)\times\mathbf{J}(\Gamma)+\mathbf{A}(\Gamma)\times\bullet d\mathbf{J}(\Gamma)\right)\\
&= -\frac{1}{J^2}\left((\mathbf{A}(\Gamma)\times\bullet d\mathbf{B})\times\mathbf{J}(\Gamma)+\mathbf{A}(\Gamma)\times(\mathbf{J}(\Gamma)\times\bullet d\mathbf{B})\right)\\
&= -\frac{1}{J^2}\left((\mathbf{A}(\Gamma)\times\bullet d\mathbf{B})\times\mathbf{J}(\Gamma)-(\mathbf{J}(\Gamma)\times\bullet d\mathbf{B})\times\mathbf{A}(\Gamma)\right)\\
&= -\frac{1}{J^2}\left((\mathbf{A}(\Gamma)\times\bullet d\mathbf{B})\times\mathbf{J}(\Gamma)+(\bullet d\mathbf{B}\times\mathbf{A}(\Gamma))\times\mathbf{J}(\Gamma)+(\mathbf{A}(\Gamma)\times\mathbf{J}(\Gamma))\times\bullet d\mathbf{B}\right)\\
&= -\mathbf{X}(\Gamma)\times\bullet d\mathbf{B}.
\end{aligned}$$

Hence $||\mathbf{X}||(\Gamma)$ is constant and $\mathbf{X}(\Gamma)$ evolves on a sphere. Next we calculate

$$\begin{aligned}
\mathbf{q}\cdot\mathbf{X} &= \frac{1}{||\mathbf{J}||^2}[\mathbf{q}\cdot(\mathbf{A}\times\mathbf{J})]\\
&= -\frac{1}{||\mathbf{J}||^2}[\mathbf{A}\cdot(\mathbf{q}\times\mathbf{J})]\\
&= \frac{1}{||\mathbf{J}||^2}[\mathbf{A}\cdot(||\mathbf{q}||^2\mathbf{p}-(\mathbf{q}\cdot\mathbf{p})\mathbf{q})]\\
&= \frac{1}{||\mathbf{J}||^2}[||\mathbf{q}||^2(\mathbf{A}\cdot\mathbf{p})-(\mathbf{A}\cdot\mathbf{q})(\mathbf{q}\cdot\mathbf{p})]\\
&= \frac{1}{||\mathbf{J}||^2}[-||\mathbf{q}||(\mathbf{q}\cdot\mathbf{p})-||\mathbf{J}||^2(\mathbf{q}\cdot\mathbf{p})+||\mathbf{q}||(\mathbf{q}\cdot\mathbf{p})]\\
&= \mathbf{q}\cdot\mathbf{p}.
\end{aligned}$$

This implies

$$\begin{aligned}
\bullet d||\mathbf{p}(\Gamma)-\mathbf{X}(\Gamma)||^2 &= \bullet d||\mathbf{p}||^2(\Gamma)-2\mathbf{X}(\Gamma)\cdot\bullet d\mathbf{p}(\Gamma)-2\mathbf{p}(\Gamma)\cdot\bullet d\mathbf{X}(\Gamma)\\
&= 2\frac{(\mathbf{q}\cdot\mathbf{p})(\Gamma)}{||\mathbf{q}||^3(\Gamma)}dt+2\mathbf{X}(\Gamma)\cdot\frac{\mathbf{q}}{||\mathbf{q}||}(\Gamma)\\
&\quad+2\mathbf{X}(\Gamma)\cdot(\mathbf{p}(\Gamma)\times\bullet d\mathbf{B})+2\mathbf{p}(\Gamma)\cdot(\mathbf{X}\times\bullet d\mathbf{B})\\
&= \frac{2}{||\mathbf{q}||^3(\Gamma)}[(\mathbf{q}\cdot\mathbf{p})(\Gamma)+(\mathbf{X}\cdot\mathbf{q})(\Gamma)]dt\\
&= 0.
\end{aligned}$$

Consequently, the momentum vector $\mathbf{p}(\Gamma)$ is always a constant distance away from a stochastic process on a sphere. This serves as the analogue of the deterministic Hamilton's theorem (see Milnor [69]) for the stochastic Kepler problem.

5.2 Collisions and the Moser Regularization

In the previous section we showed that if Γ is a solution of (5.1.1) then $||\mathbf{q}||(\Gamma)$ behaves exactly as in the deterministic case. Therefore, akin to the deterministic case, we define a collision solution as one for which $||\mathbf{q}||_t := ||\mathbf{q}||(\Gamma_t)$ approaches 0 in finite time. Any such solution must have $||\mathbf{J}|| = 0$ throughout since $||\mathbf{J}||$ is conserved. This, in turn, implies that $\mathbf{J} = 0$ along the solution. The converse is true as well, since any solution with $||\mathbf{J}|| = 0$, or equivalently, $\mathbf{J} = 0$ must behave exactly the same as the deterministic Kepler problem by replacing in $J^i = 0$ in equation (5.1.1). This means that $||\mathbf{q}|| \to 0$ in finite time as in the deterministic Kepler problem. In summary, *the solutions ending up in a collision are precisely the ones with angular momentum equal to zero.*

This allows us to use the Moser map to regularize the collisions. The procedure is similar to the deterministic Kepler problem, for which we refer to Cushman and Bates [30] and Heckmann and de Laat [41]. We define

$$F(\mathbf{q},\mathbf{p}) : T^*(\mathbb{R}^3 \setminus \{0\}) \to \mathbb{R}$$

to be the Hamiltonian

$$F(\mathbf{q},\mathbf{p}) = \frac{1}{2}(||\mathbf{q}||)\left(||\mathbf{p}||^2 + 1\right).$$

Let $(\mathbf{Q}(t), \mathbf{P}(t))$ denote a solution of the deterministic Kepler problem and define a new time parameter s by $\frac{ds}{dt} = \frac{1}{||\mathbf{Q}(t)||}$. Then the integral curves of X_F on $F^{-1}(1)$ with respect to the time parameter s are the integral curves of X_h on $h^{-1}(-\frac{1}{2})$ with respect to time t.

Next, let Γ denote a solution to the stochastic Kepler problem. Let $s(t) = \int_0^t \frac{dt}{||\mathbf{q}||_t}$, and note that since $||\mathbf{q}||_t = ||\mathbf{Q}(t)||$, $s(t)$ is identical to the time parameter s in the previous paragraph. Since on $h^{-1}(-\frac{1}{2}) = F^{-1}(1)$, we have $X_h(\Gamma_t)||\mathbf{q}||_t = X_F(\Gamma_t)$, it follows that the equation

$$\bullet d\tilde{\Gamma}_t = X_F(\tilde{\Gamma})ds(t) + \sum_{i=1}^{3} X_{J^i}(\tilde{\Gamma}_t) \bullet d(\nu_i B_t^i). \tag{5.2.1}$$

has the same solutions on $F^{-1}(1)$ as (5.1.1) on $h^{-1}(-\frac{1}{2})$. It is important to note that the "true" time parameter in this equation is t and we are not considering the noise to be parametrized by s.

Define $K : T^*(\mathbf{R}^3 \setminus \{0\}) \to \mathbb{R}$ by $K(\mathbf{q},\mathbf{p}) = \frac{1}{2}F^2(\mathbf{q},\mathbf{p})$. Then, on $K^{-1}\left(\frac{1}{2}\right) = F^{-1}(1)$, X_K and X_F are equal, and consequently, solutions to (5.2.1) also satisfy

$$\bullet d\tilde{\Gamma} = X_K(\tilde{\Gamma}_t)ds + \sum_{i=1}^{3} X_{J^i}(\tilde{\Gamma}_t) \bullet d(\nu_i B^i). \tag{5.2.2}$$

The Moser map is the map

$$\phi_M : T^*(\mathbf{R}^3 \setminus \{0\}) \to \{(\mathbf{u},\mathbf{v}) \in T^*S^3 |\ \mathbf{u} \neq (0,0,0,1), \mathbf{v} \neq 0\} =: T^\times S^3_{np}$$

given by

$$\phi_M(\mathbf{q},\mathbf{p}) = \left(\left(\frac{2\mathbf{p}}{||\mathbf{p}||^2+1}, \frac{||\mathbf{p}||^2-1}{||\mathbf{p}||^2+1}\right), \left(-(||\mathbf{p}||^2+1)\frac{\mathbf{q}}{2} + (\mathbf{q}\cdot\mathbf{p})\mathbf{p}, -\mathbf{q}\cdot\mathbf{p}\right)\right).$$

The Moser map is a symplectomorphism, and in particular, restricts to a diffeomorphism between $h^{-1}(-\frac{1}{2})$ and $T_1^\times S^3_{np} := \{(\mathbf{u},\mathbf{v}) \in T^\times S^3_{np}| \ ||\mathbf{v}||^2 = 1\}$. On $T_1^\times S^3_{np}$, consider the Hamiltonian obtained by pushing forward K on $K^{-1}\left(\frac{1}{2}\right) = h^{-1}(-\frac{1}{2})$ by ϕ_M, namely, we let

$$G(\mathbf{u},\mathbf{v}) := K \circ \phi_M^{-1} = \frac{1}{2}||\mathbf{v}||^2.$$

G can be extended to a smooth function on $T^*S^3 \bigcap G^{-1}(\frac{1}{2})$, which we also denote by G.

The Hamiltonian G is symmetric under the cotangent-lifted $SO(4)$ action on $G^{-1}\left(\frac{1}{2}\right)$. In particular, each of the components of

$$\mathbf{\Lambda}(\mathbf{u},\mathbf{v}) = (\lambda^1(\mathbf{u},\mathbf{v}), \lambda^2(\mathbf{u},\mathbf{v}), \lambda^3(\mathbf{u},\mathbf{v})) := (u^2v_3 - u^3v_2, -u^1v_3 + u^3v_1, u^1v_2 - u^2v_1)$$

are conserved. The Moser map pulls back λ^i to the angular momentum components J^i, for each $i = 1, 2, 3$, restricted to $h^{-1}(-\frac{1}{2})$.

We now turn our attention back to the stochastic Kepler problem (5.1.1). Using the Moser map, and noting the fact that the Moser map pushes forward the components of J^i to λ^i, we obtain the following stochastic perturbation of the geodesic flow on $G^{-1}(\frac{1}{2})$:

$$\bullet d\tilde{\Gamma} = X_G(\tilde{\Gamma}_t)ds + \sum_{i=1}^{3} X_{\lambda^i}(\tilde{\Gamma}) \bullet d(\nu_i B^i). \tag{5.2.3}$$

Note that λ^i is not a conserved quantity for Equation (5.2.3). However, arguing similar to $||\mathbf{J}||$ as in the previous section, we can show that $||\mathbf{\Lambda}||$ is conserved for Equation (5.2.3). The next theorem shows that the collision solutions of the stochastic Kepler problem with $h = -\frac{1}{2}$ map to great circles passing through the north pole $(0, 0, 0, 1)$ of S^3, exactly as in Moser regularization of the deterministic Kepler problem.

Theorem 5.2.1. A solution $\tilde{\Gamma}$ of (5.2.3) has $||\mathbf{\Lambda}||(\tilde{\Gamma}) = 0$ if and only if it is a collision solution of the stochastic Kepler problem (5.1.1).

Moreover, the set $||\mathbf{\Lambda}||^{-1}(0)$ is the set of all solutions of equation (5.2.3) that pass through the set

$$C := \{(\mathbf{u},\mathbf{v}) \in T^\times S^3_{np}| \ \mathbf{u} = (0, 0, 0, 1)\}.$$

If $\tilde{\Gamma}_t$ is such a solution, then the projection of $\tilde{\Gamma}_t$ on S^3 is a great circle passing through $(0, 0, 0, 1)$.

Proof: Let $\tilde{\Gamma}$ be a solution of (5.1.1) on $G^{-1}(\frac{1}{2})$. Since the Moser map pulls back the components of $\mathbf{\Lambda}$ to the corresponding components of $\mathbf{J}$, it follows that has $\mathbf{\Lambda}(\tilde{\Gamma}) = 0$ if and only if the corresponding process Γ obtained by pulling back $\tilde{\Gamma}$ by the Moser map is a solution of the stochastic Kepler problem (5.1.1) on $h^{-1}(-\frac{1}{2})$ with $\mathbf{J}(\Gamma) = 0$. This is possible if and only if Γ is a collision solution.

First, note that C is a subset of $||\mathbf{\Lambda}||^{-1}(0)$ since the first three components of $\mathbf{u}$ are 0. So we need to prove the reverse inclusion. If $||\mathbf{\Lambda}||(\tilde{\Gamma}) = 0$, then $\tilde{\Gamma}$ satisfies

$$\bullet d\tilde{\Gamma} = X_G(\tilde{\Gamma}_t)ds.$$

Recalling $s(t) = \frac{dt}{||\mathbf{q}||_t}$ and recalling that $||\mathbf{q}||_t$ is deterministic and differentiable, the previous equation is the following ODE:

$$\frac{d\gamma}{dt} = X_G(\gamma(t))\frac{1}{||\mathbf{q}||_t}. \tag{5.2.4}$$

Now we change the time parameter to s in the above ODE. This yields,

$$\frac{d\gamma}{ds} = X_G(\gamma(s)). \tag{5.2.5}$$

The remainder of the proof proceeds similarly to the proof of (4.10) in Cushman and Bates [30]. The projection of $\gamma(s)$ on S^3 is a great circle. Such a great circle always intersects the equator of S^3 given by

$$\{\mathbf{u} = (u^1, u^2, u^3, u^4) \in S^3 |\ u^4 = 0\}.$$

Let $(\tilde{\mathbf{u}}, 0, \tilde{\mathbf{v}}, v_4)$ be the point of intersection. Since $\mathbf{\Lambda} = 0$ and $\tilde{\mathbf{u}} \neq 0$, we have $\tilde{\mathbf{v}} = 0$ or $\tilde{\mathbf{u}} = \mu\tilde{\mathbf{v}}$ for some non-zero $\mu \in \mathbb{R}$. The latter implies that $(\tilde{\mathbf{u}}, 0) \cdot (\tilde{\mathbf{v}}, v_4) = \mu$, and since $(\mathbf{u}, \mathbf{v}) \in T^*S^3$, it follows that $\mu = 0$. But then $\tilde{\mathbf{u}} = 0$, which is impossible. Hence, $\tilde{\mathbf{v}} = 0$. Moreover, since equation (5.2.3) is defined on $G^{-1}\left(\frac{1}{2}\right)$, we obtain $v_4^2 = 1$, or $v_4 = \pm 1$.

Now, we solve the ODE (5.2.5) with initial conditions $(\tilde{\mathbf{u}}, 0, 0, v_4)$. Keeping in mind that $G = \frac{1}{2}$, we have

$$\gamma(s) = (\cos(s)\tilde{\mathbf{u}}, \sin(s)v_4, -\sin(s)\tilde{\mathbf{u}}, \cos(s)v_4)$$

At time $s = \pi - v_4\frac{\pi}{2}$, $\cos(s) = 0$ and $\sin(s) = v_4$, so $\gamma(s)$ reaches C. This completes the proof.

This procedure can be extended to any fixed energy level $k < 0$ by rescaling $(\mathbf{q}, \mathbf{p}) \in T^*(\mathbf{R}^3 \setminus \{0\})$ and the noise parameters ν_1, ν_2 and ν_3. Indeed, given $a \in (0, \infty)$, rescale time by $t \mapsto a^3 t$, $(\mathbf{q}, \mathbf{p})$ by $(\mathbf{q}, \mathbf{p}) \mapsto \left(a^2\mathbf{q}, \frac{1}{a}\mathbf{p}\right)$ and the noise intensities ν_1, ν_2, ν_3 by $\nu_i \mapsto \tilde{\nu}_i = \nu_i a^{-\frac{3}{2}}$. The transformation of the phase space variables yields a symplectomorphism $F : T^*(\mathbf{R}^3 \setminus \{0\}) \to T^*(\mathbf{R}^3 \setminus \{0\})$. The usual symplectic form Ω on $T^*(\mathbf{R}^3 \setminus \{0\})$ is pushed forward by F to $\tilde{\Omega} = F_*\Omega = a\Omega$. Moreover, h induces a Hamiltonian $\tilde{h} = h \circ F^{-1}$ on $(T^*(\mathbf{R}^3 \setminus \{0\}), \tilde{\Omega})$. We have, $\tilde{h}\,(\mathbf{q}, \mathbf{p}) = a^2 h(\mathbf{q}, \mathbf{p})$. Similarly, defining $\tilde{\mathbf{J}} = \mathbf{J} \circ F^{-1}$, we obtain $\tilde{\mathbf{J}}(\mathbf{q}, \mathbf{p}) = \frac{1}{a}\mathbf{J}(\mathbf{q}, \mathbf{p})$. This implies that the Hamiltonian vector fields of h on $(T^*(\mathbf{R}^3 \setminus \{0\}), \Omega)$, denoted by X_h^{Ω}, and $\tilde{h}$ on $(T^*(\mathbf{R}^3 \setminus \{0\}), \tilde{\Omega})$, denoted by $X_{\tilde{h}}^{\tilde{\Omega}}$ are related by $X_{\tilde{h}}^{\tilde{\Omega}} = \frac{1}{a^3}X_h^{\Omega}$. Similarly, we have, $X_{\tilde{J}^i}^{\tilde{\Omega}} = X_{J^i}^{\Omega}$. Next, using the scaling property of Brownian motion, that is, for any positive constant c, $B_{ct} = \sqrt{c}B_t$, we find that $B^i_{a^3 t} = a^{\frac{3}{2}}B^i_t$. Therefore, letting $\tilde{t} = a^3 t$, the solutions of

$$\bullet d\Gamma_{\tilde{t}} = X_{\tilde{h}}^{\tilde{\Omega}}(\Gamma_{\tilde{t}})d\tilde{t} + \sum_{i=1}^{3} X_{\tilde{J}^i}^{\tilde{\Omega}}(\Gamma_{\tilde{t}}) \bullet d(\tilde{\nu}_i B^i_{\tilde{t}}). \tag{5.2.6}$$

are identical to the solutions of Equation (5.1.1). We can then choose $a = \sqrt{-2k}$ to transform the stochastic Kepler problem with energy k to the stochastic Kepler problem with energy $-\frac{1}{2}$.

Chapter 6

Stochastic Collective Hamiltonian Systems

This chapter generalizes the example of the stochastic Kepler problem presented in the previous chapter. We will study a class of stochastic Hamiltonian systems which exhibit deterministic behaviour along 'symmetry directions'.

In Lázaro-Camí and Ortega [54, Section 7.1], the authors consider symmetric Hamiltonian systems perturbed by noisy Hamiltonian vector fields arising out of coadjoint invariant functions of a coadjoint equivariant momentum map. The symplectic reduction of such systems agree with the symplectic reduction of their deterministic counterparts. This provides a method of stochastically perturbing a mechanical system while respecting its symmetries and keeping the deterministic behaviour of the reduced system intact.

Although the stochastic Kepler problem does not directly fit in this category (since it is not symmetric), the stochastic perturbations considered in that case are the components, and hence smooth functions, of the angular momentum. This situation is generalized in Section 6.1 where we define stochastic collective Hamiltonian systems. These involve a deterministic Hamiltonian vector field X_h, invariant under an action of a Lie group G, which is perturbed by noise Hamiltonian vector fields arising out of arbitrary smooth functions of a (coadjoint equivariant) momentum map. In this case, the solution can be expressed as the action of a G-valued semimartingale, which we call the stochastic phase, on an integral curve of X_h. As a direct consequence, we show that the relative equilibria of the stochastic system coincide with the relative equilibria of its deterministic counterpart. The conservation of the momentum map no longer holds, however, the momentum map still evolves on a coadjoint orbit. This directly leads to an associated coadjoint motion, which can be expressed as Lie-Poisson equation. In this connection, Cruzeiro, Holm, and Ratiu [29] also obtain coadjoint motion via a stochastic Clebsch variational principle.

Even if the deterministic Hamiltonian is invariant under a free, proper, and canonical action of a Lie group, the Stratonovich differential equations governing the stochastic dynamics are generally not invariant and hence reduction and reconstruction methods mentioned in Section 4.1 no longer apply. Nevertheless, in Section 6.2, we show that the projection of the stochastic solution on the reduced space agrees with the projection of the deterministic solution. This means that the evolution of every G-invariant smooth function along the stochastic solution coincides with its evolution along the deterministic trajectory.

This also allows us to reconstruct a solution of the stochastic Hamiltonian system from a solution of the deterministic reduced system. Once a solution of the deterministic system is known in the reduced space, the solution of the stochastic system involves a two-step reconstruction process. In the first step, the solution of the deterministic reduced system is lifted to a solution of the unreduced deterministic system by the choice of a principal connection. Then, a stochastic phase is added by solving a Lie group valued stochastic differential equation. This second reconstruction step is independent of the choice of a connection. We also show that in the reconstruction process, while the geometric phase is unaffected, the dynamic phase must be modified.

In Section 6.3 we apply this framework to Lie-Poisson reduction. Here, due to the symmetry between left and right actions, we can view a stochastic collective Hamiltonian system as either a symmetric deterministic system with a collective stochastic perturbation, or a collective deterministic system with a symmetric stochastic perturbation. In the first case, the solution is given by an action of a stochastic phase on a deterministic curve, while in the second case, the solution can be written as an action of a deterministic phase on a semimartingale. As an example, we consider the free rigid body with a perturbed spatial angular momentum, where the body angular momentum evolves deterministically.

In the final section, we consider systems with a semi-direct product structure. Semi-direct products arise due to symmetry breaking, when there is an advected quantity associated with the mechanical system, for instance, a heavy top. Here, we show that the stochastic phase affects both the evolution of the mechanical system and the advected quantity. As a specific example, we consider the heavy top with a perturbed spatial angular momentum.

In the next chapter we shall explain stochastic collective Hamiltonian systems from the point of view of coupling deterministic and stochastic dynamics. Chapter 11 revisits stochastic collective dynamics and the coupling mechanism from the variational viewpoint.

6.1 Collective Perturbations of Stochastic Hamiltonian Systems

Let $(P, \{\cdot,\cdot\})$ be a Poisson manifold. Throughout we assume that G is a Lie group acting freely, properly, and canonically on P and this action has a coadjoint equivariant momentum map $J : P \to \mathfrak{g}^*$, where $\mathfrak{g}^*$ is the dual Lie algebra of G. The action is denoted by $\phi : G \times P \to P$ and given $g \in G$ and $p \in P$, we will write $\phi_g(p)$ or $g \cdot p$ for the element $\phi(g, p) \in P$. Let $h \in C^\infty(P)$ be G-invariant and for each $i = 1, \cdots, k$, let $h_i = f_i \circ J$, where $f_i \in C^\infty(\mathfrak{g}^*)$.

Our main subject of study is the following stochastic Hamiltonian system:

$$\bullet d\Gamma = X_h(\Gamma)dt + \sum_{i=1}^{k} X_{h_i}(\Gamma) \bullet dY^i, \tag{6.1.1}$$

where $Y = (Y^1, \cdots, Y^k)$ is an arbitrary semimartingale in $\mathbb{R}^k$. We call the terms $h_i := f_i \circ J$ **collective perturbations**, following the terminology in the deterministic case, where functions of this form are collective Hamiltonians. For any solution Γ of (6.1.1), we have

$$F(\Gamma) - F(\Gamma_0) = \int \{F, h\}(\Gamma)dt + \sum_{i=1}^{k} \int \{F, h_i\}(\Gamma) \bullet dY^i. \tag{6.1.2}$$

In Lázaro-Camí and Ortega [54], the specific case of collective perturbations in which the functions f_i are invariant under the coadjoint action are discussed. Note that here we relax this requirement.

The following observations (Lemma 6.1.1 and Proposition 6.1.2) appear in [54].

Lemma 6.1.1. Any collective function Poisson commutes with any G-invariant function. That is, for any $f \in C^\infty(\mathfrak{g}^*)$, and any G-invariant $F \in C^\infty(P)$, we have $\{f \circ J, F\} = 0$.

Proof. By the Collective Hamiltonian Theorem (Theorem 2.2.8)

$$X_{f\circ J}(z) = \left(\frac{\delta f}{\delta \mu}\right)_P (z),$$

for every $z \in P$, so

$$\{F, f \circ J\}(z) = dF(z) \cdot \left(\frac{\delta f}{\delta \mu}\right)_P (z) = 0,$$

since F is G-invariant. ∎

Proposition 6.1.2. If $f_i \in C^\infty(\mathfrak{g}^*)$ is Ad*-invariant, for $i = 1, \cdots, k$, then any collective function is a conserved quantity of (6.1.1).

Proof. Let $K \in C^\infty(\mathfrak{g}^*)$. Since h is assumed G-invariant, the lemma implies $\{K \circ J, h\} = 0$. For all i, the composition $h_i = f_i \circ J$ is G-invariant, since J is equivariant and f_i is invariant. Thus the lemma implies that $\{K \circ J, h_i\} = 0$ for all i. From (6.1.2), we see that $K \circ J$ is a conserved quantity of (6.1.1). ∎

Corollary 6.1.3. If $f_i \in C^\infty(\mathfrak{g}^*)$ is Ad*-invariant, for $i = 1, \cdots, k$, then J is conserved by the flow of (6.1.1).

This is a corollary of the preceding proposition, and also a consequence of the stochastic version of Noether's theorem in [54], applied to the G-invariant function $H := h + \sum_i h_i$.

Remark 6.1.4. If at least one of the f_i's is *not* Ad*-invariant, then h_i is not G-invariant, and the system (6.1.1) need not conserve momentum, as we will see in examples.

We now consider general functions $f_i \in C^\infty(\mathfrak{g}^*)$, not necessarily coadjoint invariant.

Proposition 6.1.5. If $K \in C^\infty(\mathfrak{g}^*)$ is Ad^*-invariant, then the collective function $K \circ J$ is a conserved quantity of (6.1.1).

Proof. Since h is assumed G-invariant, Lemma 6.1.1 implies $\{K \circ J, h\} = 0$. The composition $K \circ J$ is G-invariant, since J is equivariant and K is invariant, so the same lemma also implies that $\{f_i \circ J, K \circ J\} = 0$ for all i. From (6.1.2), we see that $K \circ J$ is a conserved quantity of (6.1.1). ∎

Remark 6.1.6. $K \in C^\infty(\mathfrak{g}^*)$ is Ad^*-invariant if and only if K is a Casimir of the Lie-Poisson bracket.

Remark 6.1.7. In the perturbed Kepler problem, the spatial angular momentum is not conserved. However, its length *is* conserved. This is a consequence of the preceding proposition with $K(\mathbf{x}) = ||\mathbf{x}||$, which is rotationally invariant. Note that the coadjoint action of $SO(3)$ on $\mathfrak{so}(3)^*$ corresponds to the rotational action on $\mathbb{R}^3$.

Remark 6.1.8. If h is G-invariant and the functions f_i are invariant under the coadjoint action of G on $\mathfrak{g}^*$, then Equation (6.1.1) is G-invariant and the solutions remain on level sets of the momentum map. Additionally, if P is a symplectic manifold then it is shown in [54] that symplectic reduction of Equation (6.1.1) agrees with the deterministic Hamiltonian system on the reduced space $J^{-1}(\mu)/G_\mu$ obtained from the deterministic symplectic reduction of h. However, when f_i is not assumed to be Ad^*_G-invariant then one cannot apply the reduction and reconstruction techniques for Stratonovich equations described in [54], since Equation (6.1.1) is not G-invariant in general.

Generally, if h is a G-invariant Hamiltonian then the solution of Equation (6.1.1) can be constructed from an integral curve of X_h and a solution of a Stratonovich equation in G. Before stating the result in the stochastic case, we state and prove a deterministic analogue, to illustrate the geometric structure of the argument. The notation z^L in this proposition is used because it is an integral curve of the *left*-invariant vector field X_h (notation that will be used again in Section 6.3).

Proposition 6.1.9. Let h be a G-invariant Hamiltonian on P, and let $f \in C^\infty(\mathfrak{g}^*)$. Let $z_0 \in P$ and $\mu_0 = J(z_0)$, and suppose that $g(t)$ solves the following initial value problem in G:

$$\dot{g} = T_e R_g \left(\left. \frac{\delta f}{\delta \mu} \right|_{\mu = \mathrm{Ad}^*_{g^{-1}} \mu_0} \right), \quad g(0) = e, \tag{6.1.3}$$

where $R_g : G \to G$ is right multiplication by g. If $z^L(t)$ is the integral curve of X_h starting at z_0 then $z(t) := g(t) \cdot z^L(t)$ is a solution to

$$\dot{z} = X_h(z) + X_{f \circ J}(z), \quad z(0) = z_0.$$

Proof. Applying the product rule for group actions to $z(t) := g(t) \cdot z^L(t)$, using the multiplicative notation $\dot{g}g^{-1} := (T_e R_g)^{-1} \dot{g} \in \mathfrak{g}$, gives

$$\begin{aligned} \dot{z} &= \left(\dot{g}g^{-1}\right)_P \left(g \cdot z^L\right) + g \cdot \dot{z}^L \\ &= \left(\frac{\delta f}{\delta \mu}\Big|_{\mu = \mathrm{Ad}^*_{g^{-1}} \mu_0} \right)_P \left(g \cdot z^L\right) + g \cdot X_h \left(z^L\right). \end{aligned}$$

Since X_h is G-invariant, the second term equals $X_h\left(g \cdot z^L\right)$. Since J is conserved along the integral curve z^L, we have $J(z^L(t)) = J(z_0) = \mu_0$ for all t, so $\mathrm{Ad}^*_{g^{-1}}\mu_0 = \mathrm{Ad}^*_{g^{-1}} J(z^L) = J(g \cdot z^L)$, by the equivariance of J. By the Collective Hamiltonian theorem,

$$\left(\frac{\delta f}{\delta \mu}\Big|_{\mu = J(g \cdot z^L)} \right)_P (g \cdot z^L) = X_{f \circ J}(g \cdot z^L).$$

Therefore

$$\dot{z} = X_{f \circ J}\left(g \cdot z^L\right) + X_h\left(g \cdot z^L\right).$$

∎

In the following theorem, $z^{det}(t)$ is the same integral curve as $z^L(t)$ above, with the name changed to indicate that it is the deterministic part of the solution. The semimartingale g^S in the theorem, called the **stochastic phase**, corresponds to the path $g(t)$ in the above proposition.

Theorem 6.1.10. Suppose h is a G-invariant Hamiltonian, $\mu_0 = J(z_0)$, and g^S solves the following Stratonovich equation in G:

$$\bullet dg = \sum_{i=1}^{k} T_e R_g \left(\frac{\delta f_i}{\delta \mu}\big|_{\mu = \mathrm{Ad}^*_{g^{-1}} \mu_0} \right) \bullet dY^i, \quad g_0 = e. \tag{6.1.4}$$

If z^{det} is the integral curve of X_h starting at $\Gamma_0 = z_0 \in P$ then $\Gamma = g^S \cdot z^{\mathrm{det}}$ solves Equation (6.1.1).

The first step of the proof requires a stochastic version of the product rule for group actions. This is given by the following lemma in Shigekawa [84]:

Lemma 6.1.11. Let M, N and L be smooth manifolds, X be a semimartingale on M, Y be a semimartingale on N and $\Phi : M \times N \to L$ be a smooth map. Suppose $Z = \Phi(X, Y)$. Given any 1-form α on L

$$\int \alpha \bullet dZ = \int \Phi^*_X \alpha \bullet dY + \int \Phi^*_Y \alpha \bullet dX.$$

Proof of Theorem 6.1.10. Using the previous lemma, for all 1-forms α on P

$$\int \alpha \bullet d\Gamma = \int \phi^*_{g^S}\alpha \bullet dz^{\det} + \int \phi^*_{z^{\det}}\alpha \bullet dg^S.$$

Since $\mu_0 = J(z_0) = J\left(z^{\det}\right)$ and J is coadjoint equivariant, we have $J(\Gamma) = \mathrm{Ad}^*_{(g^S)^{-1}} J\left(z^{\det}\right) = \mathrm{Ad}^*_{(g^S)^{-1}}(\mu_0)$. Applying G-invariance of X_h, we have

$$\begin{aligned}
&\int \phi^*_{g^S}\alpha \bullet dz^{\det} + \int \phi^*_{z^{\det}}\alpha \bullet dg^S \\
&= \int \phi^*_{g^S}\alpha\left(X_h\left(z^{\det}\right)\right)dt + \sum_{i=1}^k \int \phi^*_{z^{\det}}\alpha\left(T_eR_{g^S}\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=\mathrm{Ad}^*_{(g^S)^{-1}}\mu_0}\right)\right) \bullet dY^i \\
&= \int \phi^*_{g^S}\alpha\left(X_h\left(z^{\det}\right)\right)dt + \sum_{i=1}^k \int \phi^*_{z^{\det}}\alpha\left(T_eR_{g^S}\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right)\right) \bullet dY^i \\
&= \int \alpha\left(X_h\left(g^S\cdot z^{\det}\right)\right)dt + \sum_{i=1}^k \int \phi^*_{z^{\det}}\alpha\left(T_eR_{g^S}\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right)\right) \bullet dY^i \\
&= \int \alpha(X_h(\Gamma))dt + \sum_{i=1}^k \int \phi^*_{z^{\det}}\alpha\left(T_eR_{g^S}\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right)\right) \bullet dY^i.
\end{aligned}$$

Given $g \in G$, $p \in P$ and $v_g \in T_gG$, denote by v_gp the derivative of the map $\varphi_p(g) = g\cdot p$ evaluated at v_g. For all $v_{g^S} \in T_{g^S}G$,

$$\begin{aligned}
\phi^*_{z^{\det}}\alpha(v_{g^S}) &= \alpha\left(v_{g^S}z^{\det}\right) \\
&= \alpha(T_e\varphi_{g^S\cdot z^{\det}} \circ T_{g^S}R_{(g^S)^{-1}}(v_{g^S})) \\
&= \alpha((T_{g^S}R_{(g^S)^{-1}}v_{g^S})_P(g^S\cdot z^{\det})) \\
&= \alpha((T_{g^S}R_{(g^S)^{-1}}v_{g^S})_P(\Gamma)).
\end{aligned}$$

The collective Hamiltonian Theorem then yields

$$\begin{aligned}
&\int \phi^*_{z^{\det}}\alpha\left(T_eR_{g^S}\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right)\right) \bullet dY^i \\
&= \int \alpha\left(\left(T_{g^S}R_{(g^S)^{-1}} \circ T_eR_{g^S}\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right)\right)_P(\Gamma)\right) \bullet dY^i \\
&= \int \alpha\left(\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right)_P(\Gamma)\right) \bullet dY^i \\
&= \int \alpha(X_{h_i}(\Gamma)) \bullet dY^i.
\end{aligned}$$

As a result, we have

$$\int \alpha \bullet d\Gamma = \int \alpha(X_h(\Gamma))dt + \sum_{i=1}^k \int \alpha(X_{h_i}(\Gamma)) \bullet dY^i,$$

which, along with the condition $\Gamma_0 = g_0^S\cdot z^{\det}(0) = z_0$ shows that Γ solves Equation (6.1.1). ∎

Remark 6.1.12. Note that the basepoint $\mu = \mathrm{Ad}^*_{g^{-1}}\mu_0$ is the semimartingale $J(\Gamma)$, where Γ solves Equation (6.1.1).

Remark 6.1.13. Suppose J is not equivariant, but P is symplectic and connected, G is connected, and σ is a one-cocycle of the G-action on P. Then, as shown in Marsden and Ratiu [62, Section 12.3], J is equivariant with respect to the action

$$g \cdot \mu = \mathrm{Ad}^*_{g^{-1}}(\mu) + \sigma(g).$$

Then, equivariance of J and the definition of σ gives

$$J(g \cdot p) = \mathrm{Ad}^*_{g^{-1}} J(p) + \sigma(g), \tag{6.1.5}$$

for every $p \in P$. In this case, Equation (6.1.4) is given by

$$\bullet dg = \sum_{i=1}^{k} T_e R_g \left(\frac{\delta f_i}{\delta \mu}\Big|_{\mu=\mathrm{Ad}^*_{g^{-1}}\mu_0+\sigma(g)} \right) \bullet dY^i, \quad g_0 = e. \tag{6.1.6}$$

Remark 6.1.14. Suppose $N \subseteq P$ is an invariant manifold under the flow of z^{det}. Assume that the G-action on P restricts to a G-action on N. Then N is an invariant manifold for the flow of the stochastic collective Hamiltonian system (6.1.1) as well.

The next theorem generalizes Theorem 6.1.10. The proof of it is a minor modification of the proof of Theorem 6.1.10.

Theorem 6.1.15. *Suppose $Y^1, \cdots, Y^k, Z^1, \cdots, Z^\ell$ are semimartingales. Let $H_0, H_1, \cdots, H_\ell \in C^\infty(P)$ be G-invariant Hamiltonians and $h_i = f_i \circ J$, for $i = 1, \cdots, k$, where $f_i \in C^\infty(\mathfrak{g}^*)$. Consider the stochastic Hamiltonian system*

$$\bullet d\Gamma = X_{H_0}(\Gamma)\, dt + \sum_{j=1}^{\ell} X_{H_j}(\Gamma) \bullet dZ^j + \sum_{i=1}^{k} X_{h_i}(\Gamma) \bullet dY^i, \tag{6.1.7}$$

with $\Gamma_0 = z_0 \in P$. Let $\tilde{\Gamma}$ denote the solution of

$$\bullet d\tilde{\Gamma} = X_{H_0}(\tilde{\Gamma})dt + \sum_{j=1}^{\ell} X_{H_j}(\tilde{\Gamma}) \bullet dZ^j, \quad \tilde{\Gamma}_0 = z_0, \tag{6.1.8}$$

and g^S solve the Stratonovich equation

$$\bullet dg = \sum_{i=1}^{k} T_e R_g \left(\frac{\delta f_i}{\delta \mu}\Big|_{\mu=\mathrm{Ad}^*_{g^{-1}}\mu_0} \right) \bullet dY^i, \quad g_0 = e. \tag{6.1.9}$$

where $\mu_0 = J(z_0)$. Then $\Gamma = g^S \cdot \tilde{\Gamma}$ solves Equation (6.1.7).

The next result is a useful corollary of Theorem 6.1.15.

Corollary 6.1.16. Let $\mu_0 = J(\Gamma_0)$. If Γ is a solution of Equation (6.1.1) or Equation (6.1.7) then $J(\Gamma)$ lies on the coadjoint orbit $\mathcal{O}_{\mu_0}$ through μ_0.

Proof. We only prove this for solutions of Equation (6.1.1). The extension to the case of Equation (6.1.7) has a similar proof by using the stochastic Noether's theorem. By coadjoint equivariance of J and Noether's theorem, we have

$$J(\Gamma) = J\left(g^S \cdot z^{\det}\right) = \mathrm{Ad}^*_{(g^S)^{-1}} J\left(z^{\det}\right) = \mathrm{Ad}^*_{(g^S)^{-1}}\mu_0.$$

∎

Remark 6.1.17. A consequence of this corollary is that for any Ad^*-invariant function $K \in C^\infty(\mathfrak{g}^*)$, the collective function $K \circ J$ is conserved by Γ. This is an alternative proof of Proposition 6.1.5.

As an important special case, we mention that if $\frac{\delta f_i}{\delta\mu} = \xi_i$, where $\xi_i \in \mathfrak{g}$ is fixed, then the stochastic phase g^S satisfies

$$\bullet dg = \sum_{i=1}^{k} T_e R_g(\xi_i) \bullet dY^i, \quad g_0 = e. \tag{6.1.10}$$

Additionally, suppose $\mathfrak{g}$ is unimodular, that is, the trace of ad_ξ vanishes for all $\xi \in \mathfrak{g}$. This is always true if G is compact. Assume that $Y = (B^1, \cdots, B^k)$, where the B^i's are independent Brownian motions, and $\{\xi_i\}_{i=1}^k$ is an orthonormal basis with respect to an inner product $\langle\cdot,\cdot\rangle$ on $\mathfrak{g}$. Then g^S is a Brownian motion in G with respect to the right invariant metric on G induced by $\langle\cdot,\cdot\rangle$ (see Liao [52, Proposition 3.23]). Moreover, an application of Theorem 3.25 in [52, Theorem 3.26] shows that the right invariant metric $\langle\cdot,\cdot\rangle$ induces a G-invariant metric $\langle\cdot,\cdot\rangle_{\mathcal{O}_{\mu_0}}$ for which

$$J(\Gamma) = \mathrm{Ad}^*_{(g^S)^{-1}}\mu_0$$

is a Brownian motion on the coadjoint orbit $\mathcal{O}_{\mu_0}$ through μ_0. We will return to this observation in Chapter 8, where we discuss a weak formulation of the stochastic Noether's theorem.

The next theorem associates a stochastic Lie-Poisson system to the stochastic collective Hamiltonian system (6.1.1).

Theorem 6.1.18. Suppose J is coadjoint equivariant. Identify $G \times \mathfrak{g}^*$ with T^*G via the right trivialization map $\alpha_g \in T^*G \mapsto (g, (T_e R_g)^*\alpha_g) \in G \times \mathfrak{g}^*$. Let $\Gamma = g^S \cdot z^{\det}$ denote a solution of Equation (6.1.1) where g^S is a solution of Equation (6.1.4) and $z^{\det}$ is an integral curve of X_h. Then $(g^S, J(\Gamma))$ is a solution of the following stochastic Hamiltonian system in $G \times \mathfrak{g}^*$:

$$\begin{aligned} \bullet d\mu &= -\sum_{i=1}^{k} \mathrm{ad}^*_{\frac{\delta f_i}{\delta\mu}}\mu \bullet dY^i \\ \bullet dg &= T_e R_g\left(\frac{\delta f_i}{\delta\mu}\right) \bullet dY^i \end{aligned} \tag{6.1.11}$$

with initial condition $(g_0, \mu_0) = (e, J(z^{\det}(0)))$.

Proof. Let $\mu = J(\Gamma)$ and $F \in C^\infty(\mathfrak{g}^*)$. Denote by $\{\cdot,\cdot\}_+$ the $+$-Lie-Poisson bracket on $C^\infty(\mathfrak{g}^*)$. Using the fact that h is G-invariant and $J : (P, \{\cdot,\cdot\}) \to (\mathfrak{g}^*, \{\cdot,\cdot\}_+)$ is a Poisson map (see Proposition 9.23 in Holm, Stoica, and Schmah [42]), we obtain

$$\begin{aligned}
F(\mu) - F(\mu_0) &= F \circ J(\Gamma) - F \circ J(\Gamma_0)\\
&= \int \{F \circ J, h\}(\Gamma) dt + \sum_{i=1}^{k} \int \{F \circ J, h_i\}(\Gamma) \bullet dY^i\\
&= \sum_{i=1}^{k} \int \{F \circ J, f_i \circ J\}(\Gamma) \bullet dY^i\\
&= \sum_{i=1}^{k} \int \{F, f_i\}_+ \circ J(\Gamma) \bullet dY^i\\
&= \sum_{i=1}^{k} \int \{F, f_i\}_+(\mu) \bullet dY^i
\end{aligned}$$

This shows that $\bullet d\mu = -\sum_{i=1}^{k} \operatorname{ad}^*_{\frac{\delta f_i}{\delta \mu}} \mu \bullet dY^i$. We also have $\mu_0 = J(\Gamma_0) = J(z^{\det}(0))$.

Next, by equivariance of J we obtain $J(\Gamma) = \operatorname{Ad}^*_{(g^S)^{-1}} J(z^{\det}) = \operatorname{Ad}^*_{(g^S)^{-1}} \mu_0$. Since g^S satisfies Equation (6.1.4), it follows that

$$\bullet dg^S = \sum_{i=1}^{k} T_e R_{g^S} \left(\frac{\delta f_i}{\delta \mu}\bigg|_{\mu = J(\Gamma)} \right) \bullet dY^i.$$

Consequently, $(g^S, J(\Gamma))$ solves the stochastic Lie-Poisson system (6.1.11). ∎

Remark 6.1.19. Solving the system (6.1.11) is not sufficient to determine the full solution $z = g^S \cdot z^{det}$. For that we would also need to solve the usual deterministic reconstruction equation, see Section 6.2.

Corollary 6.1.20. In the context of the previous theorem, suppose that T^*G is identified with $G \times \mathfrak{g}^*$ via the left trivialization map $\alpha_g \in T^*G \mapsto (g, (T_e L_g)^* \alpha_g) \in G \times \mathfrak{g}^*$. Let $F_i = -f_i \in C^\infty(\mathfrak{g}^*)$. Then $((g^S)^{-1}, J(\Gamma))$ is a solution of the following stochastic Hamiltonian system in $G \times \mathfrak{g}^*$:

$$\begin{aligned}
\bullet d\mu &= \sum_{i=1}^{k} \operatorname{ad}^*_{\frac{\delta F_i}{\delta \mu}} \mu \bullet dY^i\\
\bullet dg &= T_e L_g \left(\frac{\delta F_i}{\delta \mu} \right) \bullet dY^i
\end{aligned} \tag{6.1.12}$$

with initial condition $(g_0, \mu_0) = (e, J(z^{\det}(0)))$.

Proof. By an application of the Stratonovich product rule, we have

$$0 = \bullet d(g^S (g^S)^{-1}) = T_{(g^S)^{-1}} L_{g^S} \bullet d(g^S)^{-1} + T_{g^S} R_{(g^S)^{-1}} \bullet dg^S.$$

As a result,

$$\begin{aligned}\bullet d(g^S)^{-1} &= -(T_{(g^S)^{-1}}L_{g^S})^{-1} \circ T_{g^S}R_{(g^S)^{-1}} \bullet dg^S \\ &= -\sum_{i=1}^{k} T_eL_{(g^S)^{-1}} \circ T_{g^S}R_{(g^S)^{-1}} \circ T_eR_{g^S}\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right) \bullet dY^i \\ &= -\sum_{i=1}^{k} T_eL_{(g^S)^{-1}}\left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right) \bullet dY^i \\ &= \sum_{i=1}^{k} T_eL_{(g^S)^{-1}}\left(\frac{\delta F_i}{\delta\mu}\Big|_{\mu=J(\Gamma)}\right) \bullet dY^i.\end{aligned}$$

The remaining assertions follow directly from Theorem 6.1.18. ∎

6.1.1 Relative equilibria

Definition 6.1.21. Let $h_0, \cdots, h_k \in C^\infty(P)$ and $Y = (Y^0, \cdots, Y^k)$ be a semimartingale in $\mathbb{R}^{k+1}$. A point $z_e \in P$ is called a **relative equilibrium** of the stochastic Hamiltonian system

$$\bullet d\Gamma = \sum_{i=0}^{k} X_{h_i}(\Gamma) \bullet dY^i$$

if $X_{h_i}(z_e) \in T_{z_e}(G \cdot z_e)$ for all $i = 0, \cdots, k$, where $G \cdot z_e$ is the orbit of z_e.

This definition is equivalent to the definition of relative equilibria in stochastic Hamiltonian systems given in Fernández-Saiz et al. [38]. We refer to their paper for more details on relative equilibria of stochastic Hamiltonian systems.

Theorem 6.1.22. A point z_e is a relative equilibrium of the stochastic Hamiltonian system (6.1.1) if and only if z_e is a relative equilibrium of the deterministic Hamiltonian h.

Proof. For all μ in the image of J, denote by $\frac{\delta f_i}{\delta\mu}$ the unique element of $\mathfrak{g}$ that satisfies $\left\langle \tilde{\mu}, \frac{\delta f_i}{\delta\mu} \right\rangle = Df(\mu)(\tilde{\mu})$ for all $\tilde{\mu} \in \mathfrak{g}^*$. By the collective Hamiltonian theorem (see Marsden and Ratiu [62]),

$$X_{h_i}(z_e) = \left(\frac{\delta f_i}{\delta\mu}\Big|_{\mu=J(z_e)}\right)_P (z_e) \in T_{z_e}(G \cdot z_e).$$

It follows that z_e is a relative equilibrium of (6.1.1) if and only if z_e is a relative equilibrium of X_h. ∎

6.2 Reduction and Reconstruction

6.2.1 Reduction

Let $\pi : P \to P/G$ be the canonical projection on P/G and $\{\cdot,\cdot\}^{\text{red}}$ denote the reduced Poisson bracket on P/G such that π is a Poisson map. Denote by $h^{\text{red}} \in C^\infty(P/G)$ the reduced Hamiltonian on P/G given by $h^{\text{red}} \circ \pi = h$ with $X^{P/G}_{h^{\text{red}}}$ being the corresponding Hamiltonian vector field on P/G. By the Poisson reduction theorem $\pi\left(z^{\det}\right)$ is the integral curve of $X^{P/G}_{h^{\text{red}}}$ starting at $\pi(z_0)$. Since $\pi(\Gamma) = \pi\left(g^S \cdot z^{\det}\right) = \pi\left(z^{\det}\right)$, it follows that $\pi(\Gamma)$ is (almost surely) the integral curve of $X^{P/G}_{h^{\text{red}}}$ starting at $\pi(z_0)$. Thus, we have shown the following theorem:

Theorem 6.2.1. Suppose h is G-invariant. If Γ solves Equation (6.1.1) then $\pi(\Gamma)$ almost surely solves

$$\dot{z} = X^{P/G}_{h^{\text{red}}}(z), \quad z(0) = \pi(\Gamma_0). \tag{6.2.1}$$

As an immediate corollary, we note that the evolution of G-invariant functions along Γ is deterministic.

Corollary 6.2.2. Let $F \in C^\infty(P)$ be G-invariant and suppose $F = f \circ \pi$ for some $f \in C^\infty(P/G)$. If Γ solves (6.1.1) and z^{red} is an integral curve of $X^{P/G}_{h^{\text{red}}}$ starting at $\pi(z_0)$ then $F(\Gamma_t) = f\left(z^{\text{red}}(t)\right)$ almost surely.

Assume now that P is a symplectic manifold with symplectic form Ω. If Γ solves Equation (6.1.1) then $J(\Gamma)$ lies on the coadjoint orbit $\mathcal{O}_\mu$ where μ is the conserved value of J along $z^{\det}$. Therefore, it is natural to look at orbit reduction. Let $i_{\mathcal{O}_\mu} : J^{-1}(\mathcal{O}_\mu) \hookrightarrow P$ denote the inclusion, $\pi_{\mathcal{O}_\mu} : J^{-1}(\mathcal{O}_\mu) \to J^{-1}(\mathcal{O}_\mu)/G =: P_{\mathcal{O}_\mu}$ denote the projection, $\omega_{\mathcal{O}_\mu}$ denote the Kirillov-Kostant-Souriau on the coadjoint orbit $\mathcal{O}_\mu$ and $J_{\mathcal{O}_\mu} := J|_{J^{-1}(\mathcal{O}_\mu)}$. By the orbit reduction theorem, $P_{\mathcal{O}_\mu}$ is a symplectic leaf in P/G and the symplectic form $\Omega_{\mathcal{O}_\mu}$ on $P_{\mathcal{O}_\mu}$ is uniquely determined by

$$i^*_{\mathcal{O}_\mu}\Omega = \pi^*_{\mathcal{O}_\mu}\Omega_{\mathcal{O}_\mu} + J^*_{\mathcal{O}_\mu}\omega_{\mathcal{O}_\mu}. \tag{6.2.2}$$

Denote by h^{red} the reduced Hamiltonian on $P_{\mathcal{O}_\mu}$. This is uniquely determined by $h^{\text{red}} \circ \pi_{\mathcal{O}_\mu} = h \circ i_{\mathcal{O}_\mu}$. Then $\pi_{\mathcal{O}_\mu}\left(z^{\det}\right)$ is the integral curve of the Hamiltonian vector field $X^{P_{\mathcal{O}_\mu}}_{h^{\text{red}}}$ of h^{red} starting at $\pi_{\mathcal{O}_\mu}(z_0)$. Since

$$\pi_{\mathcal{O}_\mu}(\Gamma) = \pi_{\mathcal{O}_\mu}\left(g^S \cdot z^{\det}\right) = \pi_{\mathcal{O}_\mu}\left(z^{\det}\right),$$

we see that $\pi_{\mathcal{O}_\mu}(\Gamma)$ is (almost surely) the integral curve of $X^{P_{\mathcal{O}_\mu}}_{h^{\text{red}}}$ starting at $\pi_{\mathcal{O}_\mu}(z_0)$. Hence, we have the following theorem:

Theorem 6.2.3. If Γ solves Equation (6.1.1) then $\pi_{\mathcal{O}_\mu}(\Gamma)$ almost surely solves

$$\dot{z}^{\text{red}} = X^{P_{\mathcal{O}_\mu}}_{h^{\text{red}}}(z^{\text{red}}), \quad z^{\text{red}}(0) = \pi_{\mathcal{O}_\mu}(\Gamma_0), \tag{6.2.3}$$

where $\mathcal{O}_\mu$ is the coadjoint orbit through $\mu = J(z_0)$.

The next corollary is analogous to Corollary 6.2.2.

Corollary 6.2.4. Let Γ be a solution of Equation (6.1.1) and suppose Γ lies on $J^{-1}(\mathcal{O}_\mu)$, where $\mathcal{O}_\mu$ is the coadjoint orbit through $\mu = J(\Gamma_0)$. Suppose $F \in C^\infty(J^{-1}(\mathcal{O}_\mu))$ is G-invariant and $F = f \circ \pi_{\mathcal{O}_\mu}$ for some $f \in C^\infty(P_{\mathcal{O}_\mu})$. If z^{red} is an integral curve of $X^{P_{\mathcal{O}_\mu}}_{h^{\mathrm{red}}}$ starting at $\pi_{\mathcal{O}_\mu}(z_0)$ then $F(\Gamma_t) = f\left(z^{\mathrm{red}}(t)\right)$ almost surely.

Remark 6.2.5. As mentioned in Section 2.3, in the deterministic case orbit reduction provides an alternative characterization of symplectic reduction since the reduced space $P_{\mathcal{O}_\mu}$ is symplectomorphic to $P_\mu := J^{-1}(\mu)/G_\mu$ where $\mu \in \mathfrak{g}^*$ is a regular value of J and G_μ is the coadjoint isotropy subgroup of μ. The reduced spaces P_μ and $P_{\mathcal{O}_\mu}$ are symplectomorphic via a symplectomorphism $L_\mu : P_\mu \to P_{\mathcal{O}_\mu}$. Let z_μ be an integral curve of the reduced Hamiltonian vector field $X^{P_\mu}_{h_\mu}$ of the reduced Hamiltonian h_μ on P_μ. Then $z = L_\mu(z_\mu)$ is an integral curve of $X^{P_{\mathcal{O}_\mu}}_{h^{\mathrm{red}}}$. Thus, even though a solution Γ of Equation (6.1.1) does not lie on a level set of J, an integral curve of the deterministic symplectic reduction of X_h can be identified with the projection $\pi_{\mathcal{O}_\mu}(\Gamma)$.

6.2.2 Reconstruction

We now turn to the problem of reconstructing a solution of Equation (6.1.1) from an integral curve of the reduced Hamiltonian vector field. We will restrict ourselves to Poisson reduction, since the method of reconstruction carries over *mutatis mutandis* to orbit reduction.

We assume that we have a principal connection A on $P \xrightarrow{\pi} P/G$. Let h be G-invariant and z^{red} be a deterministic integral curve of the reduced Hamiltonian vector field $X^{P/G}_{h^{\mathrm{red}}}$ starting at $c_0 = \pi(z_0)$ in P/G. Denote by z^{Hor} the horizontal lift of z^{red} via the connection A starting at $z_0 \in \pi^{-1}(c_0)$ and let $g^{\mathrm{det}}(t)$ be a deterministic curve in G that solves

$$\dot{g}(t) = T_e L_{g(t)}(A(X_h(z^{\mathrm{Hor}}(t))))$$

with $g^{\mathrm{det}}(0) = e$. Then, by deterministic reconstruction (see Marsden [60]) $z^{\mathrm{det}}(t) = g^{\mathrm{det}}(t) \cdot z^{\mathrm{Hor}}(t)$ is an integral curve of X_h starting at $\Gamma_0 = z_0 \in P$ that projects to z^{red}. Using Theorem 6.1.10, the solution of Equation (6.1.1) can then be written as $\Gamma = g^S \cdot z^{\mathrm{det}}$, where g^S is the G-valued semimartingale solving Equation (6.1.4).

Remark 6.2.6. If P is a symplectic manifold then a solution of Equation (6.1.1) can be reconstructed from a deterministic solution of the orbit reduction of the h. In this case one replaces P by $J^{-1}(\mathcal{O}_\mu)$ where $\mathcal{O}_\mu$ is the coadjoint orbit through $\mu = J(z_0)$ in the preceding argument.

Remark 6.2.7. Note that the reconstruction procedure occurs in two stages. The first stage involves reconstructing a solution z^{det} of X_h from an integral curve z^{red} of the reduced Hamiltonian vector field. This procedure is entirely deterministic. The next stage involves solving the stochastic phase equation (6.1.4). The randomness in the solution Γ arises out of this second stage of the reconstruction process.

Remark 6.2.8. We note that the stochastic phase in Equation (6.1.4) is independent of the choice of the connection A.

6.2.3 Geometric Phases in Symplectic Reduction

Let us briefly recall reconstruction in the case of symplectic reduction. Suppose (P, Ω) is a symplectic manifold, μ is a regular value of J and $P_\mu = J^{-1}(\mu)/G_\mu$ where G_μ is the coadjoint isotropy subgroup of μ. Let $X^{P_\mu}_{h^{\mathrm{red}}}$ denote the Hamiltonian vector field of the reduced Hamiltonian h^{red} on P_μ. For reconstruction of solutions, suppose A_μ is a $\mathfrak{g}_\mu$-valued principal connection on $J^{-1}(\mu) \xrightarrow{\pi_\mu} P_\mu$. Let z^{red} denote an integral curve of $X^{P_\mu}_{h^{\mathrm{red}}}$ starting at $\pi_\mu(z_0) = \pi_\mu(\Gamma_0)$ and horizontally lift z^{red} to a curve z^{Hor} starting at z_0. Let $\xi(t) = A_\mu(X_h(z^{\mathrm{Hor}}))$ and $g_\mu(t)$ denote the solution of

$$\dot{g} = T_e L_{g(t)} \xi(t) \quad g(0) = e. \tag{6.2.4}$$

Then $z^{\mathrm{det}}(t) = g_\mu(t) \cdot z^{\mathrm{Hor}}(t)$ is the integral curve of X_h starting at z_0.

Suppose z^{red} is a closed curve with period T. Then there exists $\hat{g}, h \in G$ such that $z^{\mathrm{Hor}}(0) = \hat{g} \cdot z^{\mathrm{Hor}}(T)$ and $z^{\mathrm{det}}(T) = h \cdot z^{\mathrm{det}}(0)$. Note that $\hat{g}$ is the holonomy of z^{red} with respect to A_μ. We have, $h = g_\mu(T)\hat{g}$. For deterministic systems, we call $\hat{g}$ or $\log(\hat{g}) \in \mathfrak{g}_\mu$, the **geometric phase** and $g_\mu(T)$ or $\log(g_\mu(T))$, the **dynamic phase**.

By Theorem 6.1.10, the solution of the stochastic collective Hamiltonian system (6.1.1) is given by $\Gamma_t = (g^S_t g_\mu(t)) \cdot z^{\mathrm{Hor}}(t)$, where g^S solves Equation (6.1.4). Since the equation $z^{\mathrm{Hor}}(0) = \hat{g} \cdot z^{\mathrm{Hor}}(T)$ is purely deterministic, the geometric phase does not change under the stochastic perturbation. On the other hand at $t = T$ we have

$$\Gamma_T = (g^S_T g_\mu(T)) \cdot (\hat{g} \cdot z^{\mathrm{Hor}}(0)) = (g^S_T g_\mu(T)) \cdot (\hat{g} \cdot z^{det}(0)).$$

Therefore the element $g_\mu(T) \in G_\mu$ is replaced by the G-valued random variable $g^S_T g_\mu(T)$. On the other hand, at the Lie algebra level $\log(g_\mu(T)) \in \mathfrak{g}_\mu$ is replaced by the $\mathfrak{g}$-valued random variable $\log(g^S_T g_\mu(T))$. We will see that in this stochastic context it is important to differentiate whether we treat $g^S_T g_\mu(T)$ or $\log(g^S_T g_\mu(T))$ as the dynamic phase.

Suppose G is a compact connected Lie group. Recall that a bi-invariant metric on G corresponds to the choice of an Ad-invariant inner product. The associated Levi-Civita connection is given by $\nabla_{X^L} Y^L = \frac{1}{2}([X, Y])^L$, where $X, Y \in \mathfrak{g}$ and X^L, Y^L are the corresponding left invariant vector fields. Note that ∇ is a bi-invariant connection. Let ε^L and ε^R denote the left and right stochastic exponentials on G (see Hakim-Dowek and Lépingle[40] and Arnaudon [6]). If Y is a $\mathfrak{g}$-valued semimartingale then $\varepsilon^L(Y)$ and $\varepsilon^R(Y)$ are the solutions of

$$\bullet dg = T_e L_g \bullet dY, \quad g_0 = e, \tag{6.2.5}$$

and

$$\bullet dg = T_e R_g \bullet dY, \quad g_0 = e, \tag{6.2.6}$$

respectively. Following Catuogno and Ruffino [17] let $\mathcal{L}$ denote the left stochastic logarithm on G, defined by $\mathcal{L}(g) = \int \omega_L \bullet dg$, where g_t is a G-valued semimartingale and ω_L is the left Maurer-Cartan 1-form. Then, given any semimartingale Y in $\mathfrak{g}$, $\mathcal{L}(\varepsilon^L(Y)) = Y$ and $\mathcal{L}(\varepsilon^R(Y)) = \int \mathrm{Ad}_{(\varepsilon^R(Y))^{-1}} \bullet dY$ for a semimartingale Y in $\mathfrak{g}$.

Theorem 6.2.9. In the setup of Theorem 6.1.10, let G be a compact connected Lie group and denote by $\{\xi_i\}_{i=1}^k$ an orthonormal basis of $\mathfrak{g}$. Let $Y = (B^1, B^2, \dots B^k)$ where the B^i's are independent Brownian motions. Then

$$\mathbb{E}[\mathcal{L}(g^S g_\mu)] = \mathcal{L}(g_\mu) = \log(g_\mu). \tag{6.2.7}$$

Proof. Let $f_i \in C^\infty(\mathfrak{g}^*)$ given by $f_i(\mu) = \langle \mu, \xi_i \rangle$. Then the solution of Equation (6.1.4) is a G-valued Brownian motion and equals $\varepsilon^R(\beta)$, where $\beta = \sum_{i=1}^k \xi_i B^i$ is a Brownian motion in $\mathfrak{g}$.

Let $B = \varepsilon^R(\beta)$ and

$$\hat{\beta} = \mathcal{L}(B) = \int \mathrm{Ad}_{B^{-1}} \bullet d\beta = \sum_{i=1}^k \mathrm{Ad}_{B^{-1}} \xi_i \bullet dB^i.$$

Denote by $\Delta_\mathfrak{g}$ the Laplacian on $\mathfrak{g}$. Then, for any $f \in C^\infty(\mathfrak{g})$

$$\Delta_\mathfrak{g}(f) = \mathrm{Tr}\,\mathrm{Hess} f = \sum_{i=1}^k \mathrm{Hess} f(\xi_i, \xi_i) = \sum_{i=1}^k \xi_i\, \xi_i[f].$$

Since the inner product on $\mathfrak{g}$ is Ad-invariant, given any $t \geq 0$ $\{\mathrm{Ad}_{B_t^{-1}} \xi_i\}_{i=1}^k$ is also an orthonormal basis. As a result, given any $f \in C^\infty(\mathfrak{g})$ and $t \geq 0$, we have

$$\begin{aligned}
\Delta_\mathfrak{g}(f) &= \sum_{i=1}^k \mathrm{Hess} f(\xi_i, \xi_i) \\
&= \sum_{i=1}^k \mathrm{Hess} f(\mathrm{Ad}_{B_t^{-1}} \xi_i, \mathrm{Ad}_{B_t^{-1}} \xi_i) \\
&= \sum_{i=1}^k \mathrm{Ad}_{B_t^{-1}} \xi_i\, \mathrm{Ad}_{B_t^{-1}} \xi_i[f].
\end{aligned}$$

Hence,

$$\begin{aligned}
&f(\hat{\beta}_t) - f(\hat{\beta}_0) \\
&= \sum_{i=1}^k \int \langle df, \mathrm{Ad}_{B^{-1}} \xi_i \rangle\, dB^i + \frac{1}{2} \sum_{i=1}^k \int \mathrm{Ad}_{B^{-1}} \xi_i\, \mathrm{Ad}_{B^{-1}} \xi_i[f] dt \\
&= \sum_{i=1}^k \int \langle df, \mathrm{Ad}_{B^{-1}} \xi_i \rangle\, dB^i + \frac{1}{2} \sum_{i=1}^k \int \mathrm{Ad}_{B^{-1}} \xi_i\, \mathrm{Ad}_{B^{-1}} \xi_i[f] dt \\
&= \sum_{i=1}^k \int \langle df, \mathrm{Ad}_{B^{-1}} \xi_i \rangle\, dB^i + \frac{1}{2} \int \Delta_g(f) dt.
\end{aligned}$$

Since the first term on the right is a local martingale, therefore $\hat{\beta}$ is a $\mathfrak{g}$-valued Brownian motion. We apply now the more general formula (8) in [6], which states that $\mathcal{L}(YB) = \int \mathrm{Ad}_{B^{-1}} d\mathcal{L}(Y) + \mathcal{L}(B)$

where Y is a martingale and B and finite variation process, with $Y = g^S$ and $B = g_\mu$. We get

$$\begin{aligned}\mathcal{L}(g^S g_\mu) &= \int \mathrm{Ad}_{g_\mu^{-1}} d\mathcal{L}(g^S) + \mathcal{L}(g_\mu)\\ &= \int \mathrm{Ad}_{g_\mu^{-1}} d\hat{\beta} + \mathcal{L}(g_\mu).\end{aligned}$$

Taking expectation in $\mathfrak{g}$, the first term vanishes since $\hat{\beta}$ is a $\mathfrak{g}$-valued Brownian motion. Hence,

$$\mathbb{E}[\mathcal{L}(g^S g_\mu)] = \mathcal{L}(g_\mu) = \log(g_\mu). \tag{6.2.8}$$

∎

On the other hand, let $\rho : G \to GL(V)$ be a representation of G on a real or complex vector space V. Denote the Casimir element of G by C. This is defined by $C = \sum_{i=1}^{k} \xi_i \, \xi_i \in \mathfrak{U}(\mathfrak{g})$, where $\mathfrak{U}(\mathfrak{g})$ is the universal enveloping algebra of $\mathfrak{g}$. If $\mathfrak{U}(\mathfrak{g})$ is identified with the left invariant differential operators on G, then C coincides with the Laplace operator. Diez and Miaskiwskyi [33] have shown that $\mathbb{E}[\rho(B)] = \exp\left(\frac{t}{2}\rho(C)\right)$, where B is a G-valued Brownian motion. Consequently, we obtain the following theorem:

Theorem 6.2.10. In the setup of Theorem 6.1.10, let G be a compact connected Lie group and denote by $\{\xi_i\}_{i=1}^k$ is an orthonormal basis of $\mathfrak{g}$. Let $Y = (B^1, B^2, \dots B^k)$ where the B^i's are independent Brownian motions. Then

$$\mathbb{E}[\rho(g^S g_\mu)] = \exp\left(\frac{t}{2}\rho(C)\right)\rho(g_\mu). \tag{6.2.9}$$

Remark 6.2.11. We observe that

$$\rho\left(\exp\left(\mathbb{E}[\mathcal{L}(g^S g_\mu)]\right)\right) \neq \mathbb{E}[\rho(g^S g_\mu)]$$

.

Remark 6.2.12. If z^{red} is a closed curve with period T then $\mathbb{E}[\mathcal{L}(g^S g_\mu)] = \log(g_\mu)(T)$. Thus, on average, the logarithm of the dynamic phase is unchanged. On the other hand, $\mathbb{E}[\rho(g^S_T g_\mu(T))]$ is different from $\rho(g_\mu(T))$ by a factor of $\exp\left(\frac{T}{2}\rho(C)\right)$.

6.2.4 Central Force Motion

Let $\rho : \mathbb{R}^3 \to \mathbb{R}$ be defined by $\rho(\mathbf{x}) = ||\mathbf{x}||$, where $||\cdot||$ is the Euclidean norm. Suppose V is a real-valued function defined on an open subset $U \subseteq \mathbb{R}$. We consider the Hamiltonian

$$h(\mathbf{q}, \mathbf{p}) = \frac{1}{2}||\mathbf{p}||^2 + V(||\mathbf{q}||)$$

defined for $(\mathbf{q},\mathbf{p}) \in Q \times \mathbb{R}^3 \cong T^*Q$, where $Q = \rho^{-1}(U)$ and $(\mathbb{R}^3)^*$ is identified with $\mathbb{R}^3$ via the standard inner product. Note that h is $SO(3)$-invariant and is a conserved quantity for Equation (6.2.10). Let $\mathbf{J} = \mathbf{q} \times \mathbf{p} = (J^1, J^2, J^3)$ denote the angular momentum. We consider the stochastic Hamiltonian system

$$\bullet d\Gamma = X_h(\Gamma)dt + \sum_{i=1}^{3} X_{J^i}(\Gamma) \bullet dB^i, \quad \Gamma_0 = (\mathbf{q}_0, \mathbf{p}_0), \tag{6.2.10}$$

where the B^i's are independent Brownian motions. We assume that $||\mathbf{q}_0 \times \mathbf{p}_0|| \neq 0$. Identify the coadjoint orbit through $J(\Gamma_0)$ with the sphere $\mathbb{S}_c^2$ of radius $c = ||\mathbf{q}_0 \times \mathbf{p}_0||$ centered at the origin. Let P_c denote the reduced space $J^{-1}\left(\mathbb{S}_c^2\right)/SO(3)$. Then, by Theorem 6.2.3, $\pi(\Gamma) = \pi\left(z^{\text{det}}\right)$, where z^{det} is an integral curve of X_h. The reduced Hamiltonian on P_c is the 'radial' Hamiltonian given by

$$h^{\text{red}}(r, p_r) = \frac{p_r^2}{2} + V_{\text{eff}}(r)$$

where $V_{\text{eff}}(r)$ is the effective potential given by $V_{\text{eff}}(r) = \frac{c^2}{2r^2} + V(r)$. Note that this implies that the radial distance $r = ||\mathbf{q}||$ and the radial velocity $v_r = \dot{r}$ exhibits the same behaviour as the deterministic problem. Since $||\mathbf{q}||$ has deterministic dynamics and h is a conserved quantity for Equation (6.2.10), it follows that $||\mathbf{p}||$ also has the same dynamics as the deterministic system.

Remark 6.2.13. Since $||\mathbf{J}||$ is constant, Equation (6.2.10) describes a central force problem in which the normal to the orbital plane is a stochastic process on a sphere in $\mathbb{R}^3$. The deterministic behaviour of the reduced equations suggests that the motion of the particle on the orbital plane is deterministic while the orbital plane experiences stochastic changes in orientation.

Collisional Solutions

Suppose $V(r) = -\frac{k}{r^\alpha}$ where $\alpha < 2$ and $k > 0$. Setting $h^{\text{red}} = E$ and using the fact that $r^2 \frac{p_r^2}{2} \geq 0$, we obtain

$$Er^2 + kr^{2-\alpha} \geq \frac{c^2}{2}.$$

Since $2 - \alpha > 0$, it follows that $r \to 0$ in finite time if and only if $c = 0$. On the other hand, on $||\mathbf{J}||^{-1}(0)$, the solutions of Equation (6.2.10) are the integral curves of X_h. Hence we have the following theorem:

Theorem 6.2.14. Let $V(r) = -\frac{k}{r^\alpha}$ where $\alpha < 2$. A solution Γ of Equation (6.2.10) is a collisional solution if and only if $\Gamma_t = z^{\text{det}}(t)$ almost surely, where $z^{\text{det}}(t)$ is a collisional solution of the deterministic problem $\dot{z} = X_h(z)$.

6.3 Lie-Poisson Reduction and Reconstruction

Let us recall the general framework of Lie-Poisson reduction from Section 2.3. Let G be a Lie group and let $\{\cdot,\cdot\}$ be the canonical Poisson bracket on T^*G. Consider the two actions: left translation, $L_g(h) := gh$, and right translation $R_g(h) := hg$, and their cotangent lifted actions on T^*G. Both actions have Ad^*-equivariant momentum maps: $J_L : T^*G \to \mathfrak{g}^*$ the momentum map of the left action, given by $J_L(g,\alpha) = T^*R_g(\alpha) =: \alpha \cdot g^{-1}$, i.e. the right translation of α to $\mathfrak{g}^*$; and $J_R : T^*G \to \mathfrak{g}^*$ the momentum map of the right action, given by $J_R(g,\alpha) = T^*L_g(\alpha) =: g^{-1}\cdot\alpha$. These are related by $J_L(g,\alpha) = \mathrm{Ad}^*_{g^{-1}} J_R(g,\alpha)$.

On $\mathfrak{g}^*$, the $(\pm)$ Lie-Poisson bracket $\{\cdot,\cdot\}_L$ on $\mathfrak{g}^*$ is defined by

$$\{f,g\}_{\pm}(\mu) = \pm\left\langle \mu, \left[\frac{\delta f}{\delta\mu}, \frac{\delta g}{\delta\mu}\right]\right\rangle \tag{6.3.1}$$

By the Lie-Poisson reduction theorem we have that the left (respectively, right) Poisson reduced space T^*G/G is Poisson isomorphic to $\mathfrak{g}^*_{\pm} := (\mathfrak{g}^*, \{\cdot,\cdot\}_{\pm})$ via the isomorphism $\varphi : [(g,\alpha)] \to J_R(g,\alpha)$ (respectively, $\varphi : [(g,\alpha)] \to J_L(g,\alpha)$). The two momentum maps form a dual pair,

$$\begin{tikzcd}
& (T^*G, \{,\}) \arrow[dl, "J_R"'] \arrow[dr, "J_L"] & \\
\left(\mathfrak{g}^*, \{f,g\}^{\mathrm{red}}_{-}\right) & & \left(\mathfrak{g}^*, \{f,g\}^{\mathrm{red}}_{+}\right),
\end{tikzcd}$$

in particular, every momentum level set of J_L is mapped by J_R to a symplectic leaf, and vice versa (see Marsden and Weinstein [59]).

If $h \in C^\infty(T^*G)$ is left G-invariant, then h factors through J_R as $h = h^{\mathrm{red}}_L \circ J_R$. Similarly, if h is right G-invariant, then h factors through J_L as $h = h^{\mathrm{red}}_R \circ J_L$. Consequently, the left invariant Hamiltonians on T^*G are of the form $f^L \circ J_R$ and the right invariant Hamiltonians on T^*G are of the form $f^R \circ J_L$ with $f^L, f^R \in C^\infty(\mathfrak{g}^*)$. The Hamiltonian vector field of f^L on $\mathfrak{g}^*$ with respect to the left Lie-Poisson bracket is given by $X^{\mathfrak{g}^*}_{f^L}(\mu) = \mathrm{ad}^*_{\frac{\delta f^L}{\delta\mu}}\mu$. The Hamiltonian vector field $X^{\mathfrak{g}^*}_{f^R}$ with respect to the right Lie-Poisson bracket on $\mathfrak{g}^*$ is given by $X^{\mathfrak{g}^*}_{f^R}(\mu) = -\mathrm{ad}^*_{\frac{\delta f^R}{\delta\mu}}\mu$.

We now consider the stochastic Hamiltonian system

$$\bullet d\Gamma = X_{f^L\circ J_R}(\Gamma)dt + \sum_{i=1}^{k} X_{f^R_i\circ J_L}(\Gamma)\bullet dY^i, \tag{6.3.2}$$

where $f^L, f^R_1, \cdots, f^R_k \in C^\infty(\mathfrak{g}^*)$. Suppose μ_R and μ_L are semimartingales on $\mathfrak{g}^*$ given by $\mu_R = J_L(\Gamma)$ and $\mu_L = J_R(\Gamma)$, where Γ is a solution of Equation (6.3.2). By Theorem 6.1.18 and 6.2.2, we have

$$\dot\mu_L = \mathrm{ad}^*_{\frac{\delta f^L}{\delta\mu}\big|_{\mu=\mu_L}}\mu_L \tag{6.3.3}$$

$$\bullet d\mu_R = -\sum_{i=1}^{k} \mathrm{ad}^*_{\frac{\delta f_i^R}{\delta \mu}\Big|_{\mu=\mu_R}} \mu_R \bullet dY^i. \tag{6.3.4}$$

Let g_R and g_L denote the solutions of the following equations:

$$\begin{aligned}
\bullet dg_R &= \sum_{i=1}^{k} T_e R_{g_R} \left(\frac{\delta f_i^R}{\delta \mu}\Big|_{\mu=\mu_R} \right) \bullet dY^i \\
&= \sum_{i=1}^{k} T_e R_{g_R} \left(\frac{\delta f_i^R}{\delta \mu}\Big|_{\mu=\mathrm{Ad}^*_{g_R{}^{-1}}\mu_L} \right) \bullet dY^i \\
\dot{g}_L &= T_e L_{g_L} \left(\frac{\delta f^L}{\delta \mu}\Big|_{\mu=\mu_L} \right) = T_e L_{g_L} \left(\frac{\delta f^L}{\delta \mu}\Big|_{\mu=\mathrm{Ad}^*_{g_L}\mu_R} \right).
\end{aligned}$$

From these, an integral curve of $X_{f^L \circ J_R}$ is given by $\alpha_L = g_L \cdot \mu_L = (T_{g_L} L_{g_L^{-1}})^*(\mu_L)$ and a solution of the stochastic Hamiltonian system

$$\bullet d\Gamma = \sum_{i=1}^{k} X_{f_i^L \circ J_L}(\Gamma) \bullet dY^i \tag{6.3.5}$$

is given by $\alpha_R = \mu_R \cdot g_R = (T_{g_R} R_{g_R^{-1}})^*(\mu_R)$.

As an immediate consequence of Theorem 6.1.10, by using the cotangent lift of the left action as the group action, we obtain the following theorem:

Theorem 6.3.1. The solution Γ of Equation (6.3.2) is given by $\Gamma = g_R \cdot \alpha_L = (g_R g_L) \cdot \mu_L$.

On the other hand, if the cotangent lift of the right action is considered as the group action, then the Hamiltonian $f^L \circ J_R$ plays the role of the 'collective' part and the Hamiltonians $f^R \circ J_L$ are invariant. A careful inspection of the proof of Theorem 6.1.10 shows that g_L plays the role of the stochastic phase in this case. As a result, we obtain the following theorem:

Theorem 6.3.2. The solution Γ of Equation (6.3.2) is given by $\Gamma = \alpha_R \cdot g_L = \mu_R \cdot (g_R g_L)$.

In both of the previous theorems, the total phase is given by $g_T := g_R g_L$.

Proposition 6.3.3. The G-valued semimartingale g_T solves the following stochastic differential equation:

$$\bullet dg = T_e L_g \left(\frac{\delta f^L}{\delta \mu}\Big|_{\mu=\mu_L} \right) dt + \sum_{i=1}^{k} T_e R_g \left(\frac{\delta f_i^R}{\delta \mu}\Big|_{\mu=\mu_R} \right) \bullet dY^i. \tag{6.3.6}$$

Proof. Since $g_T = g_R g_L$, we have

$$\bullet d(g_R g_L) = T_{g_L} L_{g_R} \dot{g}^L dt + T_{g_R} R_{g_L} \bullet dg_R$$

$$= T_{g_L} L_{g_R} \circ T_e L_{g_L} \left(\frac{\delta f^L}{\delta \mu}\Big|_{\mu=\mu_L} \right) dt + \sum_{i=1}^{k} T_{g_R} R_{g_L} \circ T_e R_{g_R} \left(\frac{\delta f_i^R}{\delta \mu}\Big|_{\mu=\mu_R} \right) \bullet dY^i$$

$$= T_e L_{g_R g_L} \left(\frac{\delta f^L}{\delta \mu}\Big|_{\mu=\mu_L} \right) dt + \sum_{i=1}^{k} T_e R_{g_R g_L} \left(\frac{\delta f_i^R}{\delta \mu}\Big|_{\mu=\mu_R} \right) \bullet dY^i,$$

as required. ∎

The next theorem expresses Γ in terms of μ_R, μ_L and g_T:

Theorem 6.3.4. The solution Γ of Equation (6.3.2) is given by

$$\Gamma = \frac{1}{2}[g_T \cdot \mu_L + \mu_R \cdot g_T] = \frac{1}{2}[(T_{g_T} L_{g_T^{-1}})^* \mu_L + (T_{g_T} R_{g_T^{-1}})^* \mu_R],$$

where the operations of addition and scalar multiplication are taken fibre-wise in the fibre over g_T.

Proof. Indeed, from Theorems 6.3.1 and 6.3.2, we obtain

$$\Gamma = \frac{1}{2}(2\Gamma) = \frac{1}{2}[(g_R g_L) \cdot \mu_L + \mu_R \cdot (g_R g_L)].$$

∎

6.3.1 Free Rigid Body with a Collective Perturbation

Consider $G = SO(3)$, acting on itself by left multiplication, which is a special case of the setting of the previous section. The momentum map is defined by $J_L(\alpha) := T^* R_\Theta(\alpha)$ for all $\alpha \in T^*_\Theta SO(3)$, where R_Θ denotes right multiplication by a matrix Θ. We make the standard identifications of $\mathfrak{so}(3)$ with $\mathbb{R}^3$ via the 'hat' map and $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ via the 'breve' map (see [42]). Then $J_L(\alpha_\Theta) = \breve{\boldsymbol{\pi}}$ for a variable vector $\boldsymbol{\pi} = (\pi^1, \pi^2, \pi^3) \in \mathbb{R}^3$ called the *spatial angular momentum.* The *body angular momentum* is $\mathbf{\Pi} := \Theta^{-1}\boldsymbol{\pi} \in \mathbb{R}^3$, corresponding to $\breve{\mathbf{\Pi}} = \Theta^{-1}\breve{\boldsymbol{\pi}}\Theta = T^* L_\Theta(\alpha)$. We left-trivialize $T^*SO(3)$ via the map $\alpha \in T^*_\Theta SO(3) \mapsto (\Theta, \mathbf{\Pi}) \in SO(3) \times \mathbb{R}^3$, with $\mathbf{\Pi}$ as above.

The Hamiltonian of a free rigid body with moment of inertia matrix $\mathbb{I}$ is

$$h(\Theta, \mathbf{\Pi}) = \frac{1}{2}\mathbf{\Pi}^T \mathbb{I}^{-1} \mathbf{\Pi}. \tag{6.3.7}$$

This is left $SO(3)$-invariant and has solution:

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbb{I}^{-1}\mathbf{\Pi}, \quad \textit{(Euler equation in Lie–Poisson form)} \tag{6.3.8}$$

$$\dot{\Theta} = \Theta(\widehat{\mathbb{I}^{-1}\mathbf{\Pi}}) \quad \textit{(reconstruction equation)} \tag{6.3.9}$$

We now consider perturbations of the free rigid body, beginning with a deterministic one.

Rigid body with constant spatial torque

Let $\boldsymbol{\tau} \in \mathbb{R}^3$ be any constant vector, and define the Hamiltonian

$$H(\Theta, \boldsymbol{\Pi}) = \frac{1}{2}\boldsymbol{\Pi}^T \mathbb{I}^{-1}\boldsymbol{\Pi} + \boldsymbol{\tau} \cdot \boldsymbol{\pi}. \tag{6.3.10}$$

This Hamiltonian is no longer left $SO(3)$-invariant. Its equations of motion are

$$\dot{\boldsymbol{\Pi}} = \boldsymbol{\Pi} \times \mathbb{I}^{-1}\boldsymbol{\Pi} + \Theta^{-1}\boldsymbol{\tau}, \quad \text{(forced Euler equation)} \tag{6.3.11}$$

$$\dot{\Theta} = \Theta(\widehat{\mathbb{I}^{-1}\boldsymbol{\Pi}}), \quad \text{(same as in unperturbed rigid body)} \tag{6.3.12}$$

and the motion satisfies $\dot{\boldsymbol{\pi}} = \boldsymbol{\tau}$, i.e. the body is subject to a constant spatial torque $\boldsymbol{\tau}$. Such a torque is realisable in practice by a control moment gyro cluster such as is used in spacecraft attitude control. Note that though the second equation is identical to the reconstruction equation of the free rigid body, the equations no longer decouple.

We next consider stochastic perturbations of the same form, with the constant vector $\boldsymbol{\tau}$ replaced by Brownian motions.

Rigid body with stochastic spatial torque

Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be the canonical basis vectors for $\mathbb{R}^3$ and let $E_i = \frac{1}{\sqrt{2}}\hat{\boldsymbol{e}}_i$ for $i = 1, 2, 3$. The set $\{E_1, E_2, E_3\}$ is an orthonormal basis with respect to the trace product defined by $\langle X, Y\rangle_{\mathfrak{so}(3)} := \mathrm{tr}(XY^T)$. Consider the stochastic Hamiltonian system:

$$\bullet d\Gamma = X_h(\Gamma)dt + \sum_{i=1}^{3} X_{\pi^i}(\Gamma) \bullet dB^i, \quad \Gamma_0 = (\mathbf{I}, \boldsymbol{\Pi}_0) \tag{6.3.13}$$

where B^1, B^2 and B^3 are standard Brownian motions, h is Hamiltonian for the free rigid body and $\mathbf{I}$ is the identity matrix. Note that this equation is not invariant under the $SO(3)$ action. However, since the equations on $\mathfrak{so}(3)^*$ are deterministic, the reduced equation (6.3.8) still holds. We note that the body angular momentum of the rigid body behaves deterministically. The stochastic phase Θ^S is determined by solving the reconstruction equation

$$\bullet d\Theta = \sum_{i=1}^{3} E_i \Theta \bullet dB^i, \quad \Theta_0 = \mathbf{I}. \tag{6.3.14}$$

It follows that Θ^S is a Brownian motion in $SO(3)$ with respect to the right invariant metric induced by $\langle\cdot,\cdot\rangle_{\mathfrak{so}(3)}$. Moreover, identifying the 2-sphere $\mathbb{S}^2$ with the coadjoint orbit through $\boldsymbol{\pi}_0 = J_L(\Gamma_0)$, we see $\boldsymbol{\pi} = J_L(\Gamma)$ is a Brownian motion on $\mathbb{S}^2$. From Theorem 6.1.18, the Stratonovich equation for $\boldsymbol{\pi}$ is given by

$$\bullet d\boldsymbol{\pi} = -\sum_{i=1}^{3} (\boldsymbol{\pi} \times \mathbf{e}_i) \bullet dB^i, \tag{6.3.15}$$

where $\hat{\mathbf{e}}_i = E_i$ for all $i = 1, 2, 3$.

Let $\Theta^{\text{det}}(t)$ be the curve in $SO(3)$ that solves the reconstruction equation for the deterministic rigid body, that is,

$$\dot{\Theta}^{\text{det}}(t) = \Theta^{\text{det}}(t)\widehat{\mathbb{I}^{-1}\mathbf{\Pi}}. \tag{6.3.16}$$

Then $\Gamma = (\Theta^S\Theta^{\text{det}}, \Theta^S\Theta^{\text{det}}\breve{\mathbf{\Pi}})$ in left trivialized coordinates and $\Gamma = (\Theta^S\Theta^{\text{det}}, \breve{\boldsymbol{\pi}}\Theta^S\Theta^{\text{det}})$ in right trivialized coordinates. Let $\Theta_T := \Theta^S\Theta^{\text{det}}$. By Theorem 6.3.4

$$\Gamma = \frac{1}{2}[\Theta_T\breve{\mathbf{\Pi}} + \breve{\boldsymbol{\pi}}\Theta_T],$$

where the sum is taken fibre-wise over the semimartingale Θ_T. Note that Θ_T determines the orientation in space for the stochastic rigid body.

Remark 6.3.5. Since the body angular momentum in the above example behaves deterministically, the stochastic perturbations do not affect the internal rotational dynamics of the rigid body, such its relative equilibria. Instead, the noise affects the spatial orientation of the rigid body. As a result, while the axis of rotation undergoes stochastic perturbations, the motion of the rigid body about the axis remains deterministic.

Remark 6.3.6. We refer the reader to Lázaro-Camí and Ortega [54] and Arnaudon, De Castro and Holm [4] for the case where the perturbation involves the components of the body angular momentum. In that case, the reduced equations are stochastic as well.

Stochastic Rigid Body Phases

Suppose $T > 0$ satisfies $\mathbf{\Pi}(T) = \mathbf{\Pi}(0)$. We want to find out how much has the rigid body rotated *on average* in space. Let $G_\mu \cong S^1$ be generated by $\hat{\zeta}$, where $\zeta = \frac{\mu}{||\mu||}$. In case of the deterministic free rigid body the rotation in space is given by $\Theta^{\text{det}}(T) = \exp[\theta\hat{\zeta}]$, where $\theta = -\Lambda + 2\frac{h_\mu^{\text{red}}T}{||\mu||}$ (see Marsden [60]). Here h_μ^{red} is the conserved value of the reduced Hamiltonian corresponding to h and $\Lambda(\text{mod} 2\pi) = \frac{\text{area}\, D}{||\mu||^2}$, where D is one of the spherical caps on S^2 enclosed by $\mathbf{\Pi}(t)$. The Casimir element of $SO(3)$ is given by $C = E_1^2 + E_2^2 + E_3^2 = -\mathbf{I}$. Hence, the expectation of the stochastic phase Θ^S is $e^{-t/2}\mathbf{I}$. We recall that the stochastic phase leaves the holonomy unchanged. Moreover, it follows from Theorem 6.2.10 that on average, the rotation of the rigid body is given by $e^{-T/2}\Theta^{\text{det}}(T) = e^{-T/2}\exp[\theta\hat{\zeta}]$.

6.4 Semi-Direct Products

Let G be a Lie group, V be a vector space and $\Phi : G \to \text{Aut}(V)$ be a left representation of G on V. Let $S = G \circledS V$ be the semi-direct product of a G with V. This is a Lie group with $G \times V$ being the underlying manifold and the group operation being given by

$$(g_1, v_1) \star (g_2, v_2) = (g_1g_2, v_1 + g_1v_2) \tag{6.4.1}$$

where $(g_1, v_1), (g_2, v_2) \in G \times V$ and $g_1 v_2 = \Phi(g_1)(v_2)$. The left and right cotangent lifted actions of G on T^*G are denoted by $(\alpha_g, g) \mapsto g \cdot \alpha_g$ and $(\alpha_g, g) \mapsto \alpha_g \cdot g$ respectively and denote by $\mathrm{pr}_G : T^*G \to G$ the canonical projection. Let $\mathfrak{s}$ denote the Lie algebra of S and $\mathfrak{s}^*$ denote the dual of $\mathfrak{s}$ equipped with the left or minus Lie-Poisson bracket. Let $\xi v := \Phi'(\xi)(v)$ and following Holm, Schmah and Stoica [42], we define the diamond operator by

$$\langle v \diamond a, \xi \rangle = \langle a, \xi v \rangle \,,$$

for every $v \in V$, $a \in V^*$ and $\xi \in \mathfrak{g}$. We denote the momentum maps under the left and right cotangent lifted action of S on T^*S by J_L and J_R respectively. Let $P_L : T^*S \to T^*G \times V^*$ be the canonical map $(\alpha_g, v, a) \mapsto (\alpha_g, a)$. As shown in Marsden, Ratiu and Weinstein [61], J_R factors through P_L as $J_R = \tilde{J}_R \circ P_L$, where $\tilde{J}_R : T^*G \times V^* \to \mathfrak{s}^*$ is given by

$$\tilde{J}_R(\alpha_g, a) = (g^{-1} \cdot \alpha_g, g^{-1}a) := (T_e^* L_g(\alpha_g), \Phi(g)^*(a)).$$

A similar situation holds for J_L as well - we let $P_R : T^*S \to T^*G \times V^*$ denote the map $(\alpha_g, v, a) \mapsto (\alpha_g + (v \diamond a) \cdot g, g^{-1}a)$. Then J_L factors through P_R as $J_L = \tilde{J}_L \circ P_R$ where $\tilde{J}_L$ is given by

$$\tilde{J}_L(\alpha_g, a) = (\alpha_g \cdot g^{-1}, ga) := (T_e^* R_g(\alpha_g), \Phi(g^{-1})^*(a)).$$

Suppose that the Poisson bracket of two functions on $T^*G \times V^*$ is given by their Poisson bracket on T^*G. Let $h \in C^\infty(T^*G \times V^*)$ be a Hamiltonian in $T^*G \times V^*$ such that for all $a \in V^*$, the Hamiltonian $h_a \in C^\infty(T^*G)$ defined by $h_a(\alpha_g) = h(\alpha_g, a)$ is invariant under the cotangent lift of the stabilizer G_a of a. Then, there is a reduced Hamiltonian $h^{\mathrm{red}} \in C^\infty(\mathfrak{s}^*)$ satisfying $h^{\mathrm{red}} \circ \tilde{J}_R = h$. Define $H \in C^\infty(T^*S)$ by $H = h \circ P_L$, so that $h^{\mathrm{red}} \circ J_R = h \circ P_L = H$. Following [42], the Lie-Poisson equations for h^{red} are given by

$$\begin{aligned} \dot{\mu} &= \mathrm{ad}^*_{\frac{\delta h^{\mathrm{red}}}{\delta \mu}} \mu - \frac{\delta h^{\mathrm{red}}}{\delta \mu} \diamond a \\ \dot{a} &= -\frac{\delta h^{\mathrm{red}}}{\delta \mu} a. \end{aligned} \tag{6.4.2}$$

We now focus on the problem of reconstructing an integral curve of X_H from a solution of the reduced equations on $\mathfrak{s}^*$. Suppose $\mathscr{L}_{(g,v)}$ and $\mathscr{R}_{(g,v)}$ denote left and right translations by (g, v) in S respectively. Let $(\alpha_{g_0}, v_0, a_0) \in T^*S$, where $\mathrm{pr}_G(\alpha_{g_0}) = g_0$. Let $(\mu(t), a(t)$ denote the solution of Equation (6.4.2) with $(\mu(0), a(0)) = (T_e^* L_{g_0} \alpha_{g_0}, \Phi(g_0)^*(a_0))$. From Lie-Poisson reconstruction, the integral curve of X_H starting at (α_{g_0}, v_0, a_0) is given by

$$z^{\mathrm{det}}(t) = T^*_{(g^{\mathrm{det}}(t), v^{\mathrm{det}}(t))} \mathscr{L}_{(g^{\mathrm{det}}(t), v^{\mathrm{det}}(t))^{-1}} (\mu(t), a(t)),$$

where $(g^{\mathrm{det}}(t), v^{\mathrm{det}}(t))$ solves

$$(\dot{g}(t), \dot{v}(t)) = T_{(e,0)} \mathscr{L}_{(g(t),v(t))} \left(\frac{\delta h^{\mathrm{red}}}{\delta \mu}, \frac{\delta h^{\mathrm{red}}}{\delta a} \right), \quad (g(0), v(0)) = (g_0, v_0).$$

Consequently, $g^{\mathrm{det}}(t)$ is a solution of

$$\dot{g} = T_e L_g \frac{\delta h^{\mathrm{red}}}{\delta \mu}, \quad g(0) = g_0, \tag{6.4.3}$$

$v^{\mathrm{det}}(t) = v_0 + \int_0^t \Phi(g^{\mathrm{det}}(s))\left(\frac{\delta h^{\mathrm{red}}}{\delta a}\right) ds$ and

$$z^{\mathrm{det}}(t) = \left(T^*_{g^{\mathrm{det}}(t)} L_{g^{\mathrm{det}}(t)^{-1}} \mu(t), v^{\mathrm{det}}(t), \Phi\left(\left(g^{\mathrm{det}}(t)\right)^{-1}\right)^* a(t)\right).$$

Now we perturb X_H stochastically by collective Hamiltonians. Let $F_i \in C^\infty(\mathfrak{s}^*)$ and consider the stochastic Hamiltonian system on T^*S given by

$$\bullet d\Gamma = X_H(\Gamma)dt + \sum_{i=1}^k X_{h_i}(\Gamma) \bullet dY^i, \quad \Gamma_0 = (\alpha_{g_0}, v_0, a_0). \tag{6.4.4}$$

where Y^i is a semimartingale and $h_i = F_i \circ J_L$. By deterministic reconstruction we already have a solution z^{det} of X_H on T^*S, so we only need to determine the stochastic phase. Setting $(\nu_0, m_0) = J_L(\Gamma_0)$, the stochastic phase equation (6.1.4) yields

$$\bullet d(g,v) = \sum_{i=1}^k T_e \mathscr{R}_{(g,v)} \left(\frac{\delta F_i}{\delta \mu}, \frac{\delta F_i}{\delta a}\right) \Big|_{(\mu,a)=\mathrm{Ad}^*_{(g,v)^{-1}}(\nu_0,m_0)} \bullet dY^i, \tag{6.4.5}$$

with initial condition $(g_0, v_0) = (e, 0)$. From Holm, Schmah and Stoica [42] we have $(g,v)^{-1} = (g^{-1}, -\Phi(g^{-1})v)$ and

$$\mathrm{Ad}^*_{(g,v)^{-1}}(\mu, a) = (\mathrm{Ad}^*_{g^{-1}}\mu + v \diamond (ga), ga).$$

Thus, we obtain

$$\mathrm{Ad}^*_{(g^{-1}, -\Phi(g^{-1})v)}(\nu_0, m_0) = (\mathrm{Ad}^*_{g^{-1}}\nu_0 - \Phi(g^{-1})v \diamond (g^{-1}m_0), g^{-1}m_0).$$

Let $\boldsymbol{T} : G \times V \to \mathfrak{g}^*$ be given by $\boldsymbol{T}(g,v) = \mathrm{Ad}^*_{g^{-1}}\nu_0 - g^{-1}v \diamond (g^{-1}m_0)$. Then, Equation (6.4.5) becomes

$$\begin{aligned} \bullet dg &= \sum_{i=1}^k T_e R_g \left(\frac{\delta F_i}{\delta \mu}\right) \Big|_{\mu = \boldsymbol{T}(g,v)} \bullet dY^i \\ \bullet dv &= \sum_{i=1}^k \left[\left(\frac{\delta F_i}{\delta a}\right) \Big|_{a = g^{-1}m_0} + \left(\left(\frac{\delta F_i}{\delta \mu}\right) \Big|_{\mu = \boldsymbol{T}(g,v)}\right) v\right] \bullet dY^i \end{aligned} \tag{6.4.6}$$

with initial condition $(g_0, v_0) = (e, 0)$. Let (g^S, v^S) denote the solution of the above equations. Then the solution of Equation (6.4.4) is given by

$$\begin{aligned} \Gamma_t &= \left(g^S_t, v^S_t\right) \cdot z^{\mathrm{det}}(t) \\ &= \left(g^S_t, v^S_t\right) \cdot (\alpha_g(t), v(t), a(t)) \\ &= \left(T^*_{g^S_t g^{\mathrm{det}}(t)} L_{(g^S_t g^{\mathrm{det}}(t))^{-1}} \mu(t), v^S_t + \Phi(g^S_t) v^{\mathrm{det}}(t), \Phi\left(\left(g^S_t g^{\mathrm{det}}(t)\right)^{-1}\right)^* a(t)\right). \end{aligned}$$

Then $P_L(\Gamma_t) = \left(T^*_{g^S_t g^{\mathrm{det}}(t)} L_{(g^S_t g^{\mathrm{det}}(t))^{-1}} \mu(t), \Phi\left(\left(g^S_t g^{\mathrm{det}}(t)\right)^{-1}\right)^* a(t)\right)$.

6.4.1 The Heavy Top

Let $G = SO(3)$, $V = \mathbb{R}^3$ and $h \in C^\infty(T^*G \times V^*)$ denote the Hamiltonian of the heavy top given by

$$h(\Theta, \mathbf{\Pi}, \mathbf{v}) = \frac{1}{2}\mathbf{\Pi}^T \mathbb{I}^{-1}\mathbf{\Pi} - mg\left\langle \mathbf{v}, \Theta\boldsymbol{\chi}\right\rangle \tag{6.4.7}$$

where g is the acceleration due to gravity, m is the mass of the rigid body and $\boldsymbol{\chi}$ is the vector from the body's point of support to the centre of mass. As in the case of the free rigid body, we consider the E_1, E_2, E_3 of $\mathfrak{so}(3)$ and the collective Hamiltonians are taken to be the components of the body's spatial angular momentum $\boldsymbol{\pi} = (\pi^1, \pi^2, \pi^3) \in \mathbb{R}^3$. Letting $H = h \circ P_L$ and $\mathbf{k}$ denote the direction of gravity, we consider the stochastic Hamiltonian system:

$$\bullet d\Gamma = X_H(\Gamma)dt + \sum_{i=1}^{3} X_{\pi^i}(\Gamma) \bullet dB^i, \quad \Gamma_0 = (\mathbf{I}, \mathbf{v}_0, \mathbf{\Pi}_0, \mathbf{k}). \tag{6.4.8}$$

For physical considerations, we are interested in the solution on $T^*G \times V^*$ rather than the full solution on T^*S. The projection of Γ on $\mathfrak{so}(3)^* \times (\mathbb{R}^3)^*$ satisfies the equations:

$$\begin{aligned}\dot{\mathbf{\Pi}} &= \mathbf{\Pi} \times \mathbb{I}^{-1}\mathbf{\Pi} + mg\mathbf{a} \times \boldsymbol{\chi}\\ \dot{\mathbf{a}} &= \mathbf{a} \times \mathbb{I}^{-1}\mathbf{\Pi}.\end{aligned}$$

Since the collectives are linear and we are interested in the evolution $P_L(\Gamma)$ on $T^*G \times V^*$, hence we may ignore the v-equation in Equation (6.4.6). Then the stochastic phase equation is same as the case of the free rigid body, that is,

$$\bullet d\Theta = \sum_{i=1}^{3} E_i \Theta \bullet dB^i, \quad \Theta_0 = \mathbf{I}.$$

Let Θ^S denote the solution of the above equation. If $\Theta^{\det}(t)$ solves $\dot{\Theta}^{\det} = \Theta\breve{\mathbf{\Pi}}(t)$ with $\Theta(0) = \mathbf{I}$ and $\mathbf{\Pi}(0) = \mathbf{\Pi}_0$ then the integral curve of X_h on $T^*G \times V^*$ starting at $(\mathbf{I}, \mathbf{\Pi}_0, \mathbf{k})$ is given by $(\Theta(t), \Theta(t)\breve{\mathbf{\Pi}}(t), \Theta(t)^{-1}\mathbf{k})$ in left-trivialized coordinates. Then

$$P_L(\Gamma_t) = (\Theta^S_t\Theta^{\det}(t), \Theta^S_t\Theta^{\det}(t)\breve{\mathbf{\Pi}}(t), (\Theta^S_t\Theta^{\det}(t))^{-1}\mathbf{k}),$$

in left-trivialized coordinates.

Chapter 7

Coupling to a Lie Group

In the previous chapter, we have seen that stochastic collective Hamiltonian systems exhibit three important features:

1. The solution involves an action of a G-valued semimartingale on the deterministic solution.
2. There is a natural Lie-Poisson system associated with the stochastic collective Hamiltonian system.
3. While these systems are not G-invariant in general, if h is G-invariant one can still carry out a two-step reconstruction process starting from a reduced solution of the deterministic system. The first one involves reconstructing a solution of the deterministic system from P/G to P and introduces a deterministic phase. The second one involves solving a stochastic differential equation in G to add a stochastic phase to the deterministic solution.

We will try to explain these three features by considering an uncoupled, G-invariant stochastic Hamiltonian system on an enlarged phase space $T^*G \times P$. The Hamiltonian system on P is purely deterministic, whereas the one on T^*G is stochastic. We will first separately consider reduction via the cotangent lifted left action G on T^*G and the action of G on P. Then, we will reduce the system by considering the action of G on both of the factors. We will see that the reduced space is isomorphic to $\mathfrak{g}^* \times P$. The Poisson bracket on this space is given by the sum of the (minus) Lie-Poisson bracket on $\mathfrak{g}^*$, the Poisson bracket on P, and additional coupling or interaction terms between the two spaces. Under appropriate initial conditions, we show that the evolution of the momentum map on P can be identified with that of the momentum map for the cotangent lifted right action J_R on T^*G. Moreover, in this case, the projection of the reduced solution onto P solves a stochastic collective Hamiltonian system.

In Chapter 11 we will revisit coupling to a Lie group from a variational point of view. Therein, the momentum map J is identified with J_L, the momentum map for the cotangent lifted left action of G on T^*G. We remark that this difference arises from the fact that in this chapter we use the left action of G on T^*G, whereas in Chapter 11 we use the right action of G on T^*G.

7.1 Two Reductions and Two Geometric Phases

Suppose G is a Lie group. Assume that G acts on T^*G by cotangent lifts of left translations on T^*G and on P by a free, proper, and canonical action $\phi : G \times P \to P$. Let $J_R : T^*G \to \mathfrak{g}^*$ denote the momentum map for the cotangent lifted right action of G on T^*G. Let $f_1, \cdots, f_k \in C^\infty(\mathfrak{g}^*)$ and $h \in C^\infty(P)$ be G-invariant. Consider the stochastic Hamiltonian system:

$$\begin{aligned} \bullet d\alpha &= -\sum_{i=1}^{k} X_{f_i \circ J_R}(\alpha) \bullet dY^i \\ \dot{z}(t) &= X_h(z). \end{aligned} \tag{7.1.1}$$

We first suppose that G acts on $T^*G \times P$ by only acting on the first component, that is, given $(\alpha_{g_1}, z) \in T^*_{g_1}G \times P$ and $g_2 \in G$, we have $(\alpha_{g_1}, z) \mapsto ((T_{g_2} L_{g_2^{-1}})^* \alpha_{g_1}, z)$. Then, since $f_i \circ J_R$ is G-invariant for all i, Equation (7.1.1) is invariant under this action. Consequently, on the Poisson reduced space $\mathfrak{g}^* \times P$, the reduced system is given by

$$\begin{aligned} \bullet d\mu &= -\sum_{i=1}^{k} \mathrm{ad}^*_{\frac{\delta f_i}{\delta \mu}} \mu \bullet dY^i \\ \dot{z} &= X_h(z), \end{aligned} \tag{7.1.2}$$

where $\mathfrak{g}^*$ is equipped with the left Lie-Poisson bracket. For reconstruction to $T^*G \times P$, we need to solve the G-valued Stratonovich equation.

$$\bullet dg = -\sum_{i=1}^{k} T_e L_g \left(\frac{\delta f_i}{\delta \mu} \right) \bullet dY^i, \quad g_0 = e. \tag{7.1.3}$$

A direct calculation yields that $\bar{g}_t := g_t^{-1}$ solves the equation

$$\bullet d\bar{g} = \sum_{i=1}^{k} T_e R_{\bar{g}} \left(\frac{\delta f_i}{\delta \mu} \right) \bullet dY^i, \quad \bar{g}_0 = e. \tag{7.1.4}$$

Remark 7.1.1. We will see later that $\bar{g}$ is indeed the stochastic phase g^S introduced in the previous chapter. The key difference so far is that μ is not yet related to the momentum map of a collective Hamiltonian system. This will be achieved by the coupling mechanism.

Now consider the action of G only on the P factor of $\mathfrak{g}^* \times P$. Since h is G-invariant, so is Equation (7.1.2), and hence, we obtain a reduced Hamiltonian system on $\mathfrak{g}^* \times (P/G)$. Let h^{red} denote the reduced Hamiltonian on P/G corresponding to h, with $X^{P/G}_{h^{\mathrm{red}}}$ denoting its Hamiltonian vector field. Then, Poisson reduction of the stochastic Hamiltonian system (7.1.2) yields the following system on $\mathfrak{g}^* \times (P/G)$:

$$\begin{aligned} \bullet d\mu &= -\sum_{i=1}^{k} \mathrm{ad}^*_{\frac{\delta f_i}{\delta\mu}} \mu \bullet dY^i \\ \dot{z}^{\mathrm{red}} &= X_{h^{\mathrm{red}}}(z^{\mathrm{red}}). \end{aligned} \tag{7.1.5}$$

Next we describe the reconstruction procedure. Let $\pi : P \to P/G$ denote the projection. Let $z^{\mathrm{red}}(t)$ be an integral curve of $X_{h^{\mathrm{red}}}$ and A be a principal connection on $P \xrightarrow{\pi} P/G$. Let z^{Hor} denote the horizontal lift of $z^{\mathrm{red}}(t)$ starting at some point x_0 in the fiber of $z^{\mathrm{red}}(0)$ and $\xi(t) = A \circ X_h(z^{\mathrm{Hor}}(t))$. The phase introduced in this reconstruction process is given by solving

$$\dot{g}^{\mathrm{det}} = T_e L_{g^{\mathrm{det}}}(\xi), \quad g^{\mathrm{det}}(0) = e.$$

Remark 7.1.2. Note that this g^{det} is the same as the first phase introduced in the reconstruction from P/G to P.

The next proposition summarizes what we have obtained so far:

Proposition 7.1.3. *Let $(\mu, z^{\mathrm{red}}(t)$ be the solution of Equation (7.1.5). Let A be a principal connection on $P \xrightarrow{\pi} P/G$, $z^{\mathrm{Hor}}(t)$ be the horizontal lift of $z^{\mathrm{red}}(t)$ starting at some $x_0 \in \pi^{-1}(z^{\mathrm{red}}(0))$, $\xi(t)$ denote the $\mathfrak{g}$-valued curve $A \circ X_h(z^{\mathrm{Hor}}(t))$ and $g^{\mathrm{det}}(t)$ denote the solution of*

$$\dot{g}^{\mathrm{det}} = T_e L_{g^{\mathrm{det}}}(\xi), \quad g^{\mathrm{det}}(0) = e.$$

Furthermore, let g_t solve the G-valued Stratonovich equation

$$\bullet dg = -\sum_{i=1}^{k} T_e L_g \left(\frac{\delta f_i}{\delta\mu}\right) \bullet dY^i, \quad g_0 = e.$$

Then, the solution of Equation (7.1.1) starting at $(\mu_0, x_0) \in \mathfrak{g}^* \times P = T_e^*G \times P$ is given by $(g_t \cdot \mu_t, g^{\mathrm{det}}(t) \cdot z^{\mathrm{Hor}}(t))$.

Remark 7.1.4. We mention that these two reductions commute, namely we can also reduce first by the action of G on P, and then by the action of G on T^*G. This is because we are considering G-action on each of the factors separately.

7.2 Coupling Deterministic and Stochastic Dynamics

We assume that G acts jointly on both the factors of $T^*G \times P$. Then, since the action is free, proper, and canonical, the reduced space $(T^*G \times P)/G$ carries a Poisson structure. The next proposition characterizes this reduced space:

Proposition 7.2.1 ([46])**.** Suppose $\Phi : T^*G \times P \to \mathfrak{g}^* \times P$ is the map defined by $\Phi(\alpha_g, z) = (J_R(\alpha_g), g^{-1} \cdot z)$. Consider the Poisson bracket $\{\cdot,\cdot\}_c$ on $\mathfrak{g}^* \times P$ defined by

$$\{F, K\}_c(\mu, z) = \{F, K\}_-(\mu, z) + \{F, K\}(\mu, z) - \left\langle d_z F, \left(\frac{\delta K}{\delta \mu}\right)_P \right\rangle + \left\langle d_z K, \left(\frac{\delta F}{\delta \mu}\right)_P \right\rangle.$$

Here the Lie-Poisson bracket $\{\cdot,\cdot\}_-$ is computed by treating the variable $z \in P$ as a constant and the Poisson bracket $\{\cdot,\cdot\}$ on P is computed by treating the variable $\mu \in \mathfrak{g}^*$ as a constant. Then

1. Φ is a Poisson map.
2. Let G act on $\mathfrak{g}^* \times P$ by $g \cdot (\mu, z) \in \mathfrak{g}^* \times P \mapsto (\mathrm{Ad}^*_{g^{-1}}\mu, x)$, for all $g \in G$. Then Φ is equivariant with respect to this action, and therefore induces a Poisson diffeomorphism between $(T^*G \times P)/G$ and $\mathfrak{g}^* \times P$.

Remark 7.2.2. We note that the Poisson bracket $\{\cdot,\cdot\}_c$ on $\mathfrak{g}^* \times P$ involves the sum of the $(-)$ Lie-Poisson bracket on $\mathfrak{g}^*$, the Poisson bracket on P, as well as additional coupling terms given by $-\left\langle d_z F, \left(\frac{\delta K}{\delta \mu}\right)_P \right\rangle + \left\langle d_z K, \left(\frac{\delta F}{\delta \mu}\right)_P \right\rangle$. Such coupled Poisson brackets arise in the study of coupling flexible attachments to rigid bodies; see Krishnaprasad and Marsden [46] and Simo, Posbergh, and Marsden [85].

Let $(\alpha_g, z^{\mathrm{det}})$ denote the solution of Equation (7.1.1) and suppose g_t solves Equation (7.1.3). Then, g_t is the basepoint of α_g, which implies that

$$\Phi(\alpha_{g_t}, z^{\mathrm{det}}(t)) = (J_R(\alpha_{g_t}), g_t^{-1} \cdot z^{\mathrm{det}}) = (J_R(\alpha_{g_t}), \bar{g}_t \cdot z^{\mathrm{det}}).$$

Proposition 7.2.3. Let $\pi_P : \mathfrak{g}^* \times P \to P$ and $\pi_{\mathfrak{g}^*} : \mathfrak{g}^* \times P \to \mathfrak{g}^*$ denote the projections onto the P and $\mathfrak{g}^*$ factor respectively. On $(\mathfrak{g}^* \times P, \{\cdot,\cdot\}_c)$, $(\mu_t, \Gamma_t) = (J_R(\alpha_{g_t}), \bar{g}_t \cdot z^{\mathrm{det}})$ solves the following stochastic Hamiltonian system:

$$\bullet d(\mu, \Gamma) = X^c_{h \circ \pi_P}(\mu, \Gamma) dt + \sum_{i=1}^{k} X^c_{f_i \circ \pi_{\mathfrak{g}^*}}(\mu, \Gamma) \bullet dY^i, \tag{7.2.1}$$

where the superscript c indicates that these vector fields are computed with respect to the Poisson bracket $\{\cdot,\cdot\}_c$. Moreover, given any $f \in C^\infty(P)$, we have

$$f(\Gamma) - f(\Gamma_0) = \int \{f, h\}(\Gamma) dt + \sum_{i=1}^{k} \int \left\langle d_\Gamma f, \left(\frac{\delta f_i}{\delta \mu}\Big|_\mu\right)_P \right\rangle \bullet dY^i.$$

Proof. The fact that (μ, Γ) solves Equation (7.2.1) follows from the Poisson reduction theorem and noting that Φ induces a Poisson diffeomorphism of the Poisson reduced space $(T^*G \times P)/G$ with $\mathfrak{g}^* \times P$.

To prove the second part of the proposition, let $F = f \circ \pi_P \in C^\infty(\mathfrak{g}^* \times P)$. Then

$$F(\mu, \Gamma) - F(\mu_0, \Gamma_0) = \int dF \bullet d(\mu, \Gamma)$$

$$= \int \{F, h \circ \pi_P\}_c(\mu, \Gamma)dt - \sum_{i=1}^{k} \{F, f_i \circ \pi_{\mathfrak{g}^*}\}_c(\mu, \mathfrak{g}^*) \bullet dY^i.$$

Since $f_i \in C^\infty(\mathfrak{g}^*)$, by definition of $\{\cdot, \cdot\}_c$,

$$\begin{aligned} \{F, f_i \circ \pi_{\mathfrak{g}^*}\}_c &= \{F, f_i\}_- - \left\langle d_z F, \left(\frac{\delta f_i}{\delta \mu}\right)_P \right\rangle \\ &= -\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta f_i}{\delta \mu}\right] \right\rangle - \left\langle d_z f, \left(\frac{\delta f_i}{\delta \mu}\right)_P \right\rangle \\ &= -\left\langle d_z f, \left(\frac{\delta f_i}{\delta \mu}\right)_P \right\rangle, \end{aligned}$$

since $\frac{\delta F}{\delta \mu} = 0$. Similarly,

$$\left\langle d_z h, \left(\frac{\delta F}{\delta \mu}\right)_P \right\rangle = 0.$$

Since the projection π_P is a Poisson map, we get,

$$\{F, h \circ \pi_P\}_c = \{f, h\} \circ \pi_P.$$

Consequently, we obtain

$$\begin{aligned} F(\mu, \Gamma) - F(\mu_0, \Gamma_0) &= \int dF \bullet d(\mu, \Gamma) \\ &= \int \{F, h \circ \pi_P\}_c(\mu, \Gamma)dt - \sum_{i=1}^{k} \{F, f_i \circ \pi_{\mathfrak{g}^*}\}_c(\mu, \mathfrak{g}^*) \bullet dY^i \\ &= \int \{f, h\}(\Gamma)dt + \sum_{i=1}^{k} \int \left\langle d_\Gamma f, \left(\left.\frac{\delta f_i}{\delta \mu}\right|_\mu\right)_P \right\rangle \bullet dY^i. \end{aligned}$$

∎

The next theorem shows characterizes the evolution of the semimartingale $\mu - J(\Gamma)$.

Theorem 7.2.4. Suppose (μ, Γ) solves the equation (7.2.1) and let $\nu = \mu - J(\Gamma)$. Then

$$\bullet d\nu = \sum_{i=1}^{k} \mathrm{ad}^*_{\frac{\delta f_i}{\delta \mu}} \nu \bullet dY^i. \tag{7.2.2}$$

Proof. Let $f \in C^\infty(\mathfrak{g}^*)$ and $F = f \circ \pi_{\mathfrak{g}^*}$. Since h is G-invariant we have

$$F(\mu, \Gamma) - F(\mu_0, \Gamma_0) = -\sum_{i=1}^{k} \int \{f, f_i\}_-(\mu) \bullet dY^i.$$

Hence,

$$\bullet d\mu = -\sum_{i=1}^{k} \mathrm{ad}^*_{\frac{\delta f_i}{\delta \mu}} \mu \bullet dY^i$$

On the other hand, G-invariance of h again implies that

$$\begin{aligned} f \circ J(\Gamma) - f \circ J(\Gamma_0) &= \sum_{i=1}^{k} \left\langle d_\Gamma(f \circ J), \left(\frac{\delta f_i}{\delta \mu}\right)_P \right\rangle \bullet dY^i \\ &= -\left\langle d_{J(\Gamma)} f, \mathrm{ad}^*_{\frac{\delta f_i}{\delta \mu}} J(\Gamma) \right\rangle \bullet dY^i. \end{aligned}$$

Consequently,

$$\bullet dJ(\Gamma) = \sum_{i=1}^{k} \mathrm{ad}^*_{\frac{\delta f_i}{\delta \mu}} J(\Gamma) \bullet dY^i.$$

As a result, with $\nu = \mu - J(\Gamma)$

$$\bullet d\nu = -\sum_{i=1}^{k} \mathrm{ad}^*_{\frac{\delta f_i}{\delta \mu}} \nu \bullet dY^i$$

∎

Now we will identify the momentum map J with J_R. To do this, we solve Equation (7.2.2) with the initial condition $\nu_0 = 0$. This corresponds to setting $\mu_0 = J(\Gamma_0)$. Recall that, $\Gamma = \varphi \circ \Phi(\alpha_g, \bar{g} \cdot z^{\mathrm{det}})$, where z^{det} is an integral curve of X_h, $\bar{g}_t = g_t^{-1}$, and g_t solves

$$\bullet dg = -T_e L_g \left(\frac{\delta f_i}{\delta \mu}\bigg|_{\mu = J_R(\alpha_g)} \right) \bullet dY^i, \quad g_0 = e.$$

Therefore, setting $\nu_0 = 0$ implies that $\mu_0 = J(\Gamma_0) = J(z^{\mathrm{det}}(0))$. In this case, Equation (7.2.2) has the solution $\nu = 0$ a.s., which, by existence and uniqueness of solutions, is also unique. Therefore, we have $\mu = J(\Gamma)$ almost surely. *Thus, the evolution of the momentum map J on P can be identified with that of the momentum map J_R on T^*G.*

Let us look at the evolution of Γ. Given any $f \in C^\infty(P)$, Proposition 7.2.3 yields

$$\begin{aligned} f(\Gamma) - f(\Gamma_0) &= \int \{f, h\}(\Gamma) dt + \sum_{i=1}^{k} \int \left\langle d_\Gamma f, \left(\frac{\delta f_i}{\delta \mu}\bigg|_{\mu = J(\Gamma)} \right)_P \right\rangle \bullet dY^i \\ &= \int \{f, h\}(\Gamma) dt + \sum_{i=1}^{k} \int \{f, f_i \circ J\}(\Gamma) \bullet dY^i, \end{aligned}$$

where the last step follows by applying the Collective Hamiltonian Theorem to $X_{f_i \circ J}$. As a result Γ solves the stochastic collective Hamiltonian system

$$\bullet d\Gamma = X_h(\Gamma) dt + \sum_{i=1}^{k} X_{f_i \circ J}(\Gamma) \bullet dY^i.$$

Finally, note that by Equation (7.1.4), $\bar{g}$ solves the G-valued stochastic differential equation

$$\bullet dg = \sum_{i=1}^{k} T_e R_g \left(\frac{\delta f_i}{\delta \mu}\bigg|_{\mu = J(\Gamma)} \right) = \sum_{i=1}^{k} T_e R_g \left(\frac{\delta f_i}{\delta \mu}\bigg|_{\mu = \mathrm{Ad}^*_{g^{-1}} \mu_0} \right),$$

where $\mu_0 = J(z^{\det}(0))$. Hence, $\bar{g} = g^S$, where g^S is the stochastic phase for the stochastic collective Hamiltonian system.

Chapter 8

Martingale Evolution of Momentum Maps

Let G be a Lie group acting on a Poisson manifold P. We have seen that if G is unimodular, $\{\xi_1, \cdots, \xi_k\}$ is an orthonormal basis of $\mathfrak{g}$ with respect to a right invariant Riemannian metric on G, and $f_i(\mu) = \langle \mu, \xi_i \rangle$ then the stochastic phase g^S associated to the stochastic collective Hamiltonian system

$$\bullet d\Gamma = X_h(\Gamma)dt + \sum_{i=1}^{k} X_{f_i \circ J}(\Gamma) \bullet dB^i$$

is a Brownian motion in the Lie group, and $J(\Gamma)$ is Brownian motion in the coadjoint orbit through μ. Given a semimartingale Z on an arbitrary manifold M with a connection ∇ and associated Hessian Hess_∇, following Arnaudon, Chen and Cruzeiro [7], we define the **generalized derivative of** Z as follows: if for every $f \in C^\infty(M)$, there exists a TM-valued process A over Z such that

$$N^f = f(Z) - f(Z_0) - \frac{1}{2}\int \mathrm{Hess}_\nabla f(dZ, dZ) - \int \langle df(Z), A \rangle \, dt$$

is a real-valued local martingale, then we define

$$\frac{D^\nabla Z}{Dt} = A.$$

Z is a martingale with respect to the connection ∇, that is, for all $f \in C^\infty(M)$

$$f(Z) - f(Z_0) - \frac{1}{2}\int \mathrm{Hess}_\nabla f(dZ, dZ)$$

is a local martingale, if and only if $\frac{D^\nabla Z}{Dt} = 0$.

Using this definition, we see that a stochastic conservation law arises in this case. Since $J(\Gamma)$ is a Brownian motion on $\mathcal{O}_{\mu_0}$, in particular, it is a martingale with respect to the

Levi-Civita connection ∇^{μ_0} of $\langle\cdot,\cdot\rangle_{\mathcal{O}_{\mu_0}}$. By definition of the generalized derivative, we arrive at the conservation law

$$\frac{D^{\nabla} J(\Gamma)}{Dt} = 0.$$

We wish to point out that this conservation law is purely stochastic in nature. In particular, it does not hold in general for the deterministic collective Hamiltonian system

$$\dot{z} = X_h(z) + X_{f\circ J}(z)$$

where $f \in C^{\infty}(\mathfrak{g}^*)$ and h is G-invariant. This is because for differentiable curves, the generalized derivative coincides with the ordinary derivative. But we have already seen that $J(z)$ is not conserved along the integral curves of the above equation since $X_{f\circ J}$ is not G-invariant in general.

In this chapter, we describe a general perturbation in the vertical direction of deterministic G-invariant Hamiltonian systems such that the momentum map evolves as a martingale on the coadjoint orbit. These perturbations are introduced in the next section and they correspond to the stochastic version of adding an infinitesimally general vector field to the Hamiltonian vector field. Similar to the case of stochastic collective Hamiltonian systems, the resulting solution can be written as the action of a semimartingale in the Lie group on the deterministic integral curve. Then, following Stelmastchuk [86], we describe the general theory of martingales on reductive homogeneous spaces in Section 8.2. In Section 8.3, we will specialize this to the case of reductive coadjoint orbits. This will be used to establish conditions in Theorem 8.3.3 under which the momentum map is a martingale.

We remark that Theorem 8.3.3 generalizes the stochastic Noether's theorem of Lázaro-Camí and Ortega stated in Theorem 4.1.7. In the setting of the latter, the momentum map is conserved in the classical sense; namely, it is almost surely constant along solutions. In the present framework, this requirement is relaxed by requiring only that its generalized derivative vanishes, thereby yielding a weaker notion of conservation. For related results in stochastic calculus of variations formulated in terms of semimartingale drifts, we refer the reader to Thieullen and Zambrini [90] and Huang and Zambrini [44]. In these works, the authors study stochastic conservation by considering transformations preserving a stochastic action functional and derive necessary conditions under which associated conserved quantities arise as real-valued martingales. In this sense, our results are a manifold-level counterpart of this approach to stochastic conservation. We emphasize, however, that since smooth functions of martingales are not, in general, martingales, it follows that arbitrary smooth functions of a martingale on a coadjoint orbit need not define stochastic conserved quantities in the sense of [90].

8.1 Vertical Perturbations of Hamiltonian Systems

Let G be a Lie group and P be a Poisson manifold. Suppose $\phi : G \times P \to P$ is a smooth action of G on P. Following Hakim-Dowek and Lépingle [40] and Arnaudon [6], we recall the definitions of the left and right stochastic exponentials on G, denoted by $\varepsilon^L(Y)$ and $\varepsilon^R(Y)$, respectively. If Y is a $\mathfrak{g}$-valued semimartingale then $\varepsilon^L(Y)$ and $\varepsilon^R(Y)$ are the solutions of

$$\bullet dg = T_eL_g \bullet dY =: g(\bullet dY), \quad g_0 = e, \tag{8.1.1}$$

and

$$\bullet dg = T_eR_g \bullet dY =: (\bullet dY)g, \quad g_0 = e, \tag{8.1.2}$$

respectively. We note that if $g = \varepsilon^R(Y)$ then $g^{-1} = \varepsilon^L(-Y)$ since $g_0^{-1} = e$ and

$$\bullet dg^{-1} = -(T_{g^{-1}}L_g)^{-1} \circ T_gR_{g^{-1}} \bullet dg = T_eL_{g^{-1}} \bullet d(-Y)$$

Given any semimartingale Γ in P, we define the Stratonovich differential $(\bullet dY)_P(\Gamma)$ by setting

$$\int \alpha(\bullet dY)_P(\Gamma) = \int (\phi_\Gamma^*\alpha)_e \bullet dY. \tag{8.1.3}$$

It will be more convenient for us to think of this in terms of a basis $\{\mathbf{e}_i\}_{i=1}^k$ of $\mathfrak{g}$. Let $Y = \sum_{i=1}^k Y^i\mathbf{e}_i$, where Y^i is a real valued semimartingale. Then

$$\begin{aligned}\int \alpha(\bullet dY)_P(\Gamma) &= \int (\phi_\Gamma^*\alpha)_e \bullet dY \\ &= \sum_{i=1}^k \int (\phi_\Gamma^*\alpha)_e(\mathbf{e}_i) \bullet dY^i \\ &= \sum_{i=1}^k \int \langle \alpha, T_e\varphi_\Gamma(\mathbf{e}_i)\rangle \bullet dY^i \\ &= \sum_{i=1}^k \int \langle \alpha, (\mathbf{e}_i)_P(\Gamma)\rangle \bullet dY^i,\end{aligned}$$

where given any $p \in P$ and $g \in G$, $\varphi_p(g) = g \cdot p$. As in the proof of the stochastic collective Hamiltonian theorem, we denote the derivative of φ_p in the direction of $v_g \in T_gG$ by v_gp. We also note that the Stratonovich equations for the left and right stochastic exponentials are given by

$$\begin{aligned}\bullet dg &= \sum_{i=1}^k T_eL_g(\mathbf{e}_i) \bullet dY^i, \quad g_0 = e \\ \bullet dg &= \sum_{i=1}^k T_eR_g(\mathbf{e}_i) \bullet dY^i, \quad g_0 = e,\end{aligned}$$

respectively. The next theorem is similar to the corresponding theorem for stochastic collective Hamiltonian systems:

Theorem 8.1.1. Let $h \in C^\infty(P)$ be G-invariant and $z^{\det}$ be an integral curve of X_h. Suppose Y is a $\mathfrak{g}$-valued semimartingale and $g^S = \varepsilon^R(Y)$. Then $\Gamma := g^S \cdot z^{\det}$ solves the Stratonovich differential equation

$$\bullet d\Gamma = X_h(\Gamma)dt + (\bullet dY)_P(\Gamma). \tag{8.1.4}$$

Proof. Let $\Gamma = g^S \cdot z^{\det}$, where $g^S = \varepsilon^R(Y)$. We use the Stratonovich differential notation to prove this theorem, leaving the details to the proof of Theorem 6.1.10. With $\Gamma = g^S \cdot z^{\det}$, we have

$$\begin{aligned}
\bullet d\Gamma &= \bullet d(g^S \cdot z^{\det}) \\
&= g^S \cdot \dot{z}^{\det}(t)dt + \bullet dg^S \cdot z^{\det} \\
&= g^S \cdot X_h(z^{\det}(t))dt + \sum_{i=1}^{k} (\mathbf{e}_i g^S) z^{\det} \bullet dY^i \\
&= g^S \cdot X_h(z^{\det}(t))dt + \sum_{i=1}^{k} \left(\mathbf{e}_i (g^S)^{-1}\right)_P (g^S \cdot z^{\det}) \bullet dY^i \\
&= X_h(g^S \cdot z^{\det}(t))dt + \sum_{i=1}^{k} \left(\mathbf{e}_i (g^S)^{-1}\right)_P (g^S \cdot z^{\det}) \bullet dY^i \\
&= X_h(\Gamma)dt + (\bullet dY)_P (\Gamma),
\end{aligned}$$

as required. ∎

Remark 8.1.2. With $\bar{g} = \varepsilon^L(-Y)$, we have $g^S = \bar{g}^{-1}$.

8.2 Martingales on Reductive Homogeneous Spaces

Following Stelmastchuk [86], first we describe martingales on reductive homogeneous spaces. Let K be a closed subgroup of the Lie group G. We say that the homogeneous space G/K is **reductive** if there is an $\mathrm{Ad}(K)$ invariant subspace $\mathfrak{p}$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Here $\mathfrak{k}$ is the Lie algebra of K. We recall that if K is compact then G/K is always reductive. We let $\pi_{G/K} : G \to G/K$, $\pi_{\mathfrak{p}} : \mathfrak{g} \to \mathfrak{p}$, and $\pi_{\mathfrak{k}} : \mathfrak{g} \to \mathfrak{k}$ denote the corresponding projections. We also let $o = \pi_{G/K}(e)$.

Let $\beta_{G/K} : \mathfrak{p} \times \mathfrak{p} \to \mathfrak{p}$ be an $\mathrm{ad}(\mathfrak{k})$-invariant bilinear map. Nomizu [74, Theorem 8.1] shows that this is equivalent to choosing a G-invariant connection $\nabla^{G/K}$ on G/K. Also assume that $\nabla^{G/K}$ is torsion-free, which, by [74, Equation 9.1], implies that $\beta_{G/K}(\xi, \eta) - \beta_{G/K}(\eta, \xi) = \pi_{\mathfrak{p}} \circ \mathrm{ad}_\xi \eta$ for all $\xi, \eta \in \mathfrak{p}$. Suppose we have a bilinear map $\beta_G : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying $\pi_{\mathfrak{p}} \circ \beta_G(\xi, \eta) = \beta_{G/K}(\pi_{\mathfrak{p}}(\xi), \pi_{\mathfrak{p}}(\eta))$ and $\beta_G(\xi, \eta) - \beta_G(\eta, \xi) = \mathrm{ad}_\xi \eta$ for all $\xi, \eta \in \mathfrak{g}$. Extend β_G to a left G-invariant, torsion-free connection ∇^G on G by setting $\nabla^G_{\xi^L}(\eta^L) = \beta_G(\xi, \eta)^L$, where the superscript L denotes the left invariant vector field corresponding to a Lie algebra element. Then ∇^G reduces to $\nabla^{G/K}$ in the following sense: let ω be the left Maurer-Cartan 1-form on G. Then $\omega^{\mathfrak{k}} := \pi_{\mathfrak{k}} \circ \omega : TG \to \mathfrak{k}$ is a principal connection on the bundle $G \xrightarrow{\pi_{G/K}} G/K$. Let $(\cdot)^h$ denote the horizontal lift of a vector field on G/K to G. Then, given vector fields $X, Y \in \mathfrak{X}^\infty(G/K)$, we have $\mathbf{h}(\nabla^G_{X^h} Y^h) = (\nabla^{G/K}_X Y)^h$, where $\mathbf{h} : TG \to \mathrm{Hor}(TG)$ is the projection onto the horizontal subspace $\{T_e L_g(\mathfrak{p}) \mid g \in G\}$ defined by $\omega^{\mathfrak{k}}$.

Given this setup, the following theorem characterizes the $\nabla^{G/K}$-martingales.

Theorem 8.2.1. Let Z be a semimartingale in G/K with $Z_0 = o$ and Z^h denote the horizontal lift of Z via the connection $\omega^{\mathfrak{k}}$ starting at e, that is, $\pi_{G/K}(Z^h) = Z$, $Z^h_0 = e$ and $\int \omega^{\mathfrak{k}} \bullet dZ^h = 0$. The following are equivalent:

1. Z is a $\nabla^{G/K}$-martingale.
2. The Itô integral $\int \omega^{\mathfrak{k}} d^{\nabla^G} Z^h$ equals 0 a.s. In this case, Z^h is said to be a **horizontal martingale**.
3. The Itô integral $\int \omega d^{\nabla^G} Z^h$ is a local martingale in $\mathfrak{p}$.
4. There is a $\mathfrak{p}$-valued local martingale Y such that if $\bar{g}$ satisfies the Itô differential equation
$$d^{\nabla^G} \bar{g} = T_e L_{\bar{g}} d^{\nabla^{\mathfrak{g}}} Y, \quad \bar{g}_0 = e,$$
then $Z = \pi_{G/K}(\bar{g})$. Here $\nabla^{\mathfrak{g}}_\xi \eta = \nabla^G_\xi \eta(e)$ for all $\xi, \eta \in \mathfrak{g}$.

If ∇^G is the canonical affine connection on G given by $\nabla^G_{\xi^L} \eta^L = \frac{1}{2}[\xi, \eta]^L$, for all $\xi, \eta \in \mathfrak{g}$, then each of the Itô integrals can be replaced by their corresponding Stratonovich integrals. The corresponding connection on G/K is induced by the bilinear map $(\xi, \eta) \in \mathfrak{p} \times \mathfrak{p} \mapsto \frac{1}{2}\pi_{\mathfrak{p}}([\xi, \eta])$. Moreover, the martingales in G are the stochastic exponentials of the local martingales in $\mathfrak{g}$ (see, for instance Stelmastchuk [87, Example 5.1]). In this case, we have the following corollary:

Corollary 8.2.2. Endow G/K with the connection arising from the bilinear map $(\xi, \eta) \in \mathfrak{p} \times \mathfrak{p} \mapsto \pi_{\mathfrak{p}}([\xi, \eta])$ and define the connection ∇^G on G by $\nabla^G_{\xi^L} \eta^L = \frac{1}{2}[\xi, \eta]^L$, for all $\xi, \eta \in \mathfrak{g}$. Let Z be a semimartingale in G/K with $Z_0 = o$ and Z^h denote the horizontal lift of Z via the connection $\omega^{\mathfrak{k}}$ starting at e. Then the following are equivalent:

1. Z is a $\nabla^{G/K}$-martingale.
2. $\int \omega^{\mathfrak{k}} \bullet dZ^h$ equals 0 a.s.
3. $\int \omega \bullet dZ^h$ is a local martingale in $\mathfrak{p}$.
4. There is a $\mathfrak{p}$-valued local martingale Y such that $Z = \pi_{G/K}(\varepsilon^L(Y))$.

8.3 Martingales on Coadjoint Orbits

Now consider the case when K is the coadjoint isotropy subgroup G_μ of μ. To connect the evolution of the momentum map along solutions of Equation (8.1.4), it will be convenient to use the canonical affine connection. Suppose the coadjoint orbit $\mathcal{O}_{\mu_0} \cong G/G_{\mu_0}$ is reductive, that is, there exists an $\mathrm{Ad}(G_{\mu_0})$-invariant subspace $\mathfrak{p}$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{g}_{\mu_0} \oplus \mathfrak{p}$. Given a semimartingale $\mu_t = \mathrm{Ad}^*_{g_t^{-1}} \mu_0$ in $\mathcal{O}_{\mu_0}$, its horizontal lift with respect to the canonical affine connection is a semimartingale g^h_t satisfying the following properties:

1. $\pi_{\mathcal{O}_{\mu_0}}(g_t^h) = \mu_t$, where $\pi_{\mathcal{O}_{\mu_0}} : G \to \mathcal{O}_{\mu_0}$ is the projection onto $\mathcal{O}_{\mu_0}$.
2. $g_0^h = e$.
3. $\int \omega^{\mathfrak{g}_{\mu_0}} \bullet dg_t^h = 0$.

Thus, we obtain that $k_t := (g_t^h)^{-1} g_t$ is a semimartingale in G_{μ_0}. Reductivity and horizontality imply that

$$\bullet dk_t = T_e L_{k_t} \circ \omega^{\mathfrak{g}_{\mu_0}} \bullet dg_t, \quad k_0 = e,$$

that is, $k_t = \varepsilon^L(\int \omega^{\mathfrak{g}_{\mu_0}} \bullet dg_t)$. This allows us to determine k_t from g_t, from which g_t^h can be easily determined. As a result of Corollary 8.2.2, we obtain the following theorem:

Theorem 8.3.1. Suppose that $\mathcal{O}_{\mu_0}$ is a reductive coadjoint orbit. Endow $\mathcal{O}_{\mu_0}$ with the connection ∇^{μ_0} arising from the bilinear map $(\xi, \eta) \in \mathfrak{p} \times \mathfrak{p} \mapsto \pi_{\mathfrak{p}}([\xi, \eta])$ and define the connection ∇^G on G by $\nabla^G_{\xi^L} \eta^L = \frac{1}{2}[\xi, \eta]^L$, for all $\xi, \eta \in \mathfrak{g}$. Let $\mu_t = \mathrm{Ad}^*_{g_t^{-1}} \mu_0$ be a semimartingale in $\mathcal{O}_{\mu_0}$, and define the semimartingales $k_t = \varepsilon^L(\int \omega^{\mathfrak{g}_{\mu_0}} \bullet dg_t)$ and $g_t^h = g_t k_t^{-1}$. Then the following are equivalent:

1. μ_t is a ∇^{μ_0}-martingale.
2. $\int \omega^{\mathfrak{g}_{\mu_0}} \bullet dg^h$ equals 0 a.s.
3. $\int \omega \bullet dg^h$ is a local martingale in $\mathfrak{p}$.
4. There is a $\mathfrak{p}$-valued local martingale Y_t such that $\mu_t = \pi_{\mathcal{O}_{\mu_0}}(\varepsilon^L(Y_t))$.

As a consequence of this theorem, we record the following lemma:

Lemma 8.3.2. In the setup of Theorem 8.3.1 μ_t is a martingale in $\mathcal{O}_{\mu_0}$ if and only if $\int \mathrm{Ad}_{k_t^{-1}} \circ \pi_{\mathfrak{p}} \bullet dY_t$ is a local martingale, where $\pi_{\mathfrak{p}} : \mathfrak{g} = \mathfrak{g}_{\mu_0} \oplus \mathfrak{p} \to \mathfrak{p}$ is the projection onto the second factor.

Proof. We have

$$k_t = \varepsilon^L(\pi_{\mathfrak{g}_{\mu_0}}(-Y_t)) = \varepsilon^R(\pi_{\mathfrak{g}_{\mu_0}}(Y_t)),$$

where $\pi_{\mathfrak{g}_{\mu_0}} : \mathfrak{g} = \mathfrak{g}_{\mu_0} \oplus \mathfrak{p} \to \mathfrak{g}_{\mu_0}$ is the projection onto the first factor. Then

$$\begin{aligned}
\bullet dg_t^h &= (\bullet dg_t) k_t^{-1} - g_t k_t^{-1} (\bullet dk_t) k_t^{-1} \\
&= -g_t (\bullet dY_t) k_t^{-1} + g_t k_t^{-1} k_t \pi_{\mathfrak{g}_{\mu_0}} (\bullet dY_t) k_t^{-1} \\
&= -g_t (\pi_{\mathfrak{p}} (\bullet dY_t)) k_t^{-1} \\
&= -g_t^h (\mathrm{Ad}_{k_t^{-1}} \circ \pi_{\mathfrak{p}} (\bullet dY_t)) \\
&= g_t^h \bullet d \left(-\int \mathrm{Ad}_{k_t^{-1}} \circ \pi_{\mathfrak{p}} \bullet dY_t \right).
\end{aligned}$$

Note that, since k_t lies in G_{μ_0}, $-\int \mathrm{Ad}_{k_t^{-1}} \pi_{\mathfrak{p}} \bullet dY_t$ takes values in $\mathfrak{p}$ by the reductivity hypothesis. Thus, by Theorem 8.3.1, μ_t is a martingale if and only if $-\int \mathrm{Ad}_{k_t^{-1}} \pi_{\mathfrak{p}} \bullet dY_t$ is a local martingale, or equivalently, $\int \mathrm{Ad}_{k_t^{-1}} \pi_{\mathfrak{p}} \bullet dY_t$ is a local martingale. ∎

Now we state the main theorem of this section:

Theorem 8.3.3. Let $\Gamma = g^S \cdot z^{\det}$ solve Equation (8.1.4), where $z^{\det}$ is an integral curve of X_h and $g^S = \varepsilon^R(Y)$. Assume that the G-action on P admits a coadjoint equivariant momentum map $J : P \to \mathfrak{g}^*$ and set $\mu_0 = J(z^{\det}(0))$. Suppose that $\mathcal{O}_{\mu_0}$ is a reductive coadjoint orbit, ∇^G is the canonical affine connection on G and ∇^{μ_0} is the reduced connection on $\mathcal{O}_{\mu_0}$ obtained from ∇^G. The following are equivalent:

1. $J(\Gamma)$ is a martingale.
2. $\frac{D^{\nabla^{\mu_0}} J(\Gamma)}{Dt} = 0$.
3. Let $k_t = \varepsilon^R(\pi_{\mathfrak{g}_{\mu_0}}(Y_t))$. Then $\int \mathrm{Ad}_{k_t^{-1}} \pi_{\mathfrak{p}} \bullet dY_t$ is a local martingale.

In particular, if Y is a local martingale in $\mathfrak{p}$ then $J(\Gamma)$ is a martingale in $\mathcal{O}_{\mu_0}$.

Proof. By definition of the generalized derivative, the equivalence between (1) and (2) follows immediately. For the equivalence between (1) and (3), we note that since h is G-invariant and J is coadjoint equivariant, we have

$$J(\Gamma) = J(g^S \cdot z^{\det}) = \mathrm{Ad}^*_{(g^S)^{-1}} J(z^{\det}) = \mathrm{Ad}^*_{(g^S)^{-1}} \mu_0.$$

Then, by the previous lemma $J(\Gamma)$ is a martingale if and only if $\int \mathrm{Ad}_{k_t^{-1}} \pi_{\mathfrak{p}} \bullet dY_t$ is a local martingale.

For the second part of the theorem, suppose that Y is a local martingale in $\mathfrak{p}$. Then k_t equals the identity element e and hence, $\int \mathrm{Ad}_{k_t^{-1}} \pi_{\mathfrak{p}} \bullet dY_t = Y_t$ is a local martingale in $\mathfrak{p}$. This implies that $J(\Gamma)$ is a martingale. ∎

The next corollary is an application of this theorem to stochastic collective Hamiltonian systems with linear collectives.

Corollary 8.3.4. Let $f_i \in C^\infty(\mathfrak{g}^*)$ be defined by $f_i(\mu) = \langle \mu, \xi_i \rangle$, where $\xi_i \in \mathfrak{p}$ for all $i = 1, \cdots, k$. Let $J : P \to \mathfrak{g}^*$ be a coadjoint equivariant momentum map, $h_i = f_i \circ J$, and $Y = \sum_{i=1}^k \xi_i Y^i$, where $Y^1, \cdots, Y^k$ are real-valued semimartingales. Let ∇^G and ∇^{μ_0} be defined as in Theorem 8.3.3. Suppose $\Gamma = g^S \cdot z^{\det}$ solves

$$\bullet d\Gamma = X_h(\Gamma) dt + \sum_{i=1}^k X_{h_i}(\Gamma) \bullet dY^i,$$

where $z^{\det}$ is an integral curve of X_h and $g^S = \varepsilon^R(Y)$. Then $\frac{D^{\nabla^{\mu_0}} J(\Gamma)}{Dt} = 0$ if Y is a local martingale.

Proof. Note that Y is a semimartingale in $\mathfrak{p}$. By the Collective Hamiltonian Theorem,

$$X_h(\Gamma) dt + \sum_{i=1}^k X_{h_i}(\Gamma) \bullet dY^i$$

$$
\begin{aligned}
&= X_h(\Gamma)dt + \sum_{i=1}^{k} \left(\frac{\delta f_i}{\delta \mu}\Big|_{\mu=J(\Gamma)} \right)_P (\Gamma) \bullet dY^i \\
&= X_h(\Gamma)dt + \sum_{i=1}^{k} (\xi_i)_P(\Gamma) \bullet dY^i \\
&= X_h(\Gamma)dt + (\bullet dY)_P(\Gamma).
\end{aligned}
$$

If Y is a local martingale then by the previous theorem, $J(\Gamma)$ is a martingale. ∎

Remark 8.3.5. We mention that if $h, h_1, \cdots, h_k$ are G-invariant Hamiltonians, $Y^1, \cdots, Y^k$ are real-valued semimartingales and ξ is a semimartingale in $\mathfrak{g}$, then proceeding as in the proof of Theorem 8.1.1 the solution of the stochastic differential equation

$$\bullet d\Gamma = X_h(\Gamma)dt + \sum_{i=1}^{k} X_{h_i}(\Gamma) \bullet dY^i + (\bullet d\xi)_P(\Gamma)$$

is given by $\Gamma = g^S \cdot \tilde{\Gamma}$, where $\tilde{\Gamma}$ solves the stochastic Hamiltonian system

$$\bullet d\tilde{\Gamma} = X_h(\tilde{\Gamma})dt + \sum_{i=1}^{k} X_{h_i}(\tilde{\Gamma}) \bullet dY^i$$

and g^S solves

$$\bullet dg = T_e R_g \bullet d\xi, \quad g_0 = e.$$

By the stochastic Noether's theorem (Theorem 4.1.7), $J(\tilde{\Gamma}) = \mu_0$, where μ_0 is a constant. Hence $J(\Gamma) = \mathrm{Ad}^*_{(g^S)^{-1}} \mu_0$. In particular, if ξ is a local martingale in $\mathfrak{p}$ then

$$\frac{D^{\nabla^{\mu_0}} J(\Gamma)}{Dt} = 0.$$

In particular, with $\xi = 0$, we note that the hypotheses of the stochastic Noether's theorem 4.1.7 also imply Theorem 8.3.3.

Part III

Stochastic Variational Principles

Chapter 9

Stochastic Hamilton-Pontryagin Principle

In Lagrangian mechanics, one starts with a (smooth) curve $q(t)$ on a manifold Q and a Lagrangian function $\mathcal{L} : TQ \to \mathbb{R}$ that models the dynamics of the system. Typically, $\mathcal{L}$ is chosen to be the kinetic energy minus the potential energy. Hamilton's principle states that the trajectory of a mechanical system between time t_0 and t_1 is determined by finding the stationary point of the action $\int_{t_0}^{t_1} \mathcal{L}(q(t), \dot{q}(t))dt$ among all curves that fixes the endpoints, that is, $q(t_0) = a$ and $q(t_1) = b$ for some $a, b \in Q$. In other words, one looks at deformations $\epsilon \mapsto q_\epsilon(t)$ of $q(t)$, for $\epsilon \in (-s, s)$ with $s > 0$, such that:

1. $q_\epsilon(t)$ is smooth and $q_0(t) = q(t)$,
2. $\delta q(t) := \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0} q_\epsilon(t)$ satisfies $\delta q(t_0) = \delta q(t_1) = 0$.

Then, the stationary point is determined from $\int_{t_0}^{t_1} d\mathcal{L}(q(t), \dot{q}(t))(\delta q(t), \delta \dot{q}(t))dt = 0$.

Finding the stationary point of the action is equivalent to solving the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}}\right) = 0.$$

A standard argument to show this equivalence is to proceed in local coordinates, by subdividing the curve $q(t)$ into a finite number of segments, each of which lies in a chart. Then, in local coordinates, one writes

$$d\mathcal{L}(q(t), \dot{q}(t))(\delta q(t), \delta \dot{q}(t)) = \left\langle \frac{\partial \mathcal{L}}{\partial q}, \delta q \right\rangle + \left\langle \frac{\partial \mathcal{L}}{\partial \dot{q}}, \delta \dot{q} \right\rangle.$$

Integration by parts and an application of the fundamental lemma of the calculus of variations then yield the desired equivalence.

There are several problems that arise while trying to generalize this to the stochastic case. We list them here:

1. While taking variations of the action, the equality

$$\frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \int_{t_0}^{t_1} \mathcal{L}(q_\epsilon(t), \dot{q}_\epsilon(t))dt = \int_{t_0}^{t_1} d\mathcal{L}(q(t), \dot{q}(t))(\delta q(t), \delta \dot{q}(t))dt = 0$$

   proceeds by applying the dominated convergence theorem. However, in stochastic geometric mechanics one typically works with Stratonovich integrals, which lack the dominated convergence theorem for Itô integrals with respect to the ucp topology. In this context, the problem is in which topology should we interpret the variation δq when q is a semimartingale, and in which topology should we interpret the variation of the action itself, in order to obtain an analogue of the dominated convergence theorem for the Stratonovich case.

2. The introduction of partial derivatives of $\mathcal{L}$ involves localizing the problem in coordinate charts. In the deterministic case, when one localizes the problem to coordinate charts, the entry times and exit times for a curve are fixed, deterministic times. Thus, from a global fixed-endpoint problem, one can obtain a local fixed-endpoint problem. In the stochastic case, the entry time and exit times for a stochastic process in entering and leaving a chart are random variables. This leads to the problem of constructing variations that are fixed not only at the deterministic initial and final times but also at these random entry and exit times. Moreover, to ensure that the resulting Stratonovich integrals are well-defined, the constructed variations must also be semimartingales, and in particular, adapted to the underlying filtration.

3. The implication that a curve is a stationary point of the action only if it satisfies the Euler-Lagrange equations proceeds by applying the fundamental lemma of the calculus of variations. When $Q = \mathbb{R}$, this lemma states that if g is a continuous function defined on $(a, b) \supseteq (t_0, t_1)$ and $\int_{t_0}^{t_1} g(t)h(t)dt = 0$ for all smooth functions h compactly supported on (t_0, t_1), then $g \equiv 0$ in (t_0, t_1). Note that given a curve $\gamma : [t_0, t_1] \to \mathbb{R}$ every smooth function h which is compactly supported on (t_0, t_1) provides a fixed-endpoint deformation of γ given by $\epsilon \mapsto \gamma(t) + \epsilon h(t)$. Hence, the classical fundamental lemma operates entirely within the class of fixed-endpoint deformations used in the action principle. In the stochastic case, how does one formulate an analogous fundamental lemma for the type of fixed endpoint deformations discussed in the previous point?

4. Stochastic processes generally lack differentiability. How do we express the action integral $\int_{t_0}^{t_1} \mathcal{L}(q(t), \dot{q}(t))dt$ in terms of well-defined stochastic integrals? Moreover, in the stochastic case, how do we ensure that the action integral remains coordinate independent?

5. In [62], it is shown that under some mild hypotheses on $\mathcal{L}$ (namely, that it is a regular Lagrangian), the Euler-Lagrange equations can be viewed as determining integral curves of a specific second order vector field, called the Lagrangian vector field. Can we have a similar interpretation in the stochastic case as well?

Section 9.1 aims to answer the first three questions, and Section 9.2 addresses the final two questions.

In Section 9.1, we introduce variations and in particular, fixed endpoint variations, of semimartingales. We prove a stochastic analogue of the fundamental lemma of the calculus of variations for Stratonovich integrals, especially taking into account variations that vanish at the first hitting and exit times for a chart. In Section 9.2, we focus on the stochastic Hamilton-Pontryagin principle. The stochastic Hamilton-Pontryagin principle was formulated by Bou-Rabee and Owhadi [13] and studied more recently in Street and Takao [89]. As an application of the variational framework developed in Section 9.1, we present a proof of the local form of the stochastic Hamilton-Pontryagin principle. This generalizes the Hamilton-Pontryagin principle formulated in [89] to arbitrary noise semimartingales. We also discuss a stochastic version of Noether's theorem on the variational principle side. We then describe a novel method for working with variational principles globally on manifolds by using Stratonovich operators. Here we exploit the fact that Stratonovich equations on manifolds are determined by Stratonovich operators and these are deterministic generalizations of vector fields.

The present approach addresses several technical difficulties that have not been treated in the existing literature. In particular, Bismut [11, 12] develops stochastic variational principles for mechanical systems from the point of view stochastic optimal control. In that setting, the variations are taken to be free, while endpoint constraints are incorporated through additional terms in the action functional. The resulting backward stochastic differential equations are shown to provide sufficient conditions for criticality. However, neither a stochastic analogue of the fundamental lemma of the calculus of variations nor a proof of necessity of these equations for critical points is established therein. Variational formulations for semimartingales have also been studied in the Lie group setting by Arnaudon, Chen, and Cruzeiro [7], and, using techniques from Malliavin calculus, by Cruzeiro, Holm, and Ratiu [29] and Huang and Zambrini [44]. These works, however, do not address the localization via stopping times required to impose fixed endpoint conditions in local coordinates. Related difficulties also arise in the approaches of Crisan and Street [88] and Street and Takao [89], which are formulated using semimartingales compatible with a prescribed driving noise. By contrast, the framework developed here allows for a local-coordinate treatment while preserving both adaptedness and fixed endpoint constraints. The stochastic fundamental lemma established in Section 9.1.2 yields both necessity and sufficiency of the implicit Euler-Lagrange equations, while the intrinsic formulation in terms of solutions to Stratonovich stochastic differential equations allows us to convert stochastic variational problems to deterministic variational problems.

9.1 Variations of a Semimartingale

In this section we describe variations of a semimartingale Γ in a smooth manifold M.

Definition 9.1.1. Let Γ be a semimartingale in a smooth manifold M. A **deformation** of Γ is a map $[-s,s] \to \mathscr{S}(M)$ denoted by $\epsilon \mapsto \Gamma_\epsilon$, where $\epsilon \in [-s,s]$ for some $s > 0$, such that:

- $\Gamma_{0,t} = \Gamma_t$.

- The map $\epsilon \in [-s,s] \mapsto \Gamma_\epsilon \in \mathscr{S}(M)$ is continuous at 0 with respect to the semimartingale topology on $\mathscr{S}(M)$. Additionally, there exists a TM-valued semimartingale $\delta\Gamma$ such that for every $f \in C^\infty(M)$, $\frac{f(\Gamma_\epsilon)-f(\Gamma)}{\epsilon} \to df(\delta\Gamma)$ with respect to the semimartingale topology on $\mathscr{S}(\mathbb{R})$ as $\epsilon \to 0$. The semimartingale $\delta\Gamma$ will be called a **variation** of Γ.

Remark 9.1.2. It is assumed implicitly that the lifetimes of the semimartingales Γ_ϵ are at least as large as the lifetime of Γ.

Remark 9.1.3. By definition of differentiability in the semimartingale topology on M, Definition 9.1.1 means that the map $\epsilon \mapsto \Gamma_\epsilon$ is differentiable at $\epsilon = 0$ with respect to the semimartingale topology on $\mathscr{S}(M)$, and its derivative is $\delta\Gamma$.

Definition 9.1.4. Let M be a smooth manifold and Γ be a semimartingale in M. We say that Γ is **admissible** if, for every semimartingale Y in TM over Γ, there exists a deformation $\epsilon \mapsto \Gamma_\epsilon$ of Γ with $\delta\Gamma = Y$.

Theorem 9.1.5. Assume that Γ is a semimartingale in a Riemannian manifold M and exp denotes the exponential map on M. If $\exp_{\Gamma_t(\omega)}$ has domain $T_{\Gamma_t(\omega)}M$ for all $t \geq 0$ and $\omega \in \Omega$ then Γ is admissible.

Proof. This follows directly from Corollary 4.3 in Arnaudon and Thalmaier [5]. We remark that the hypothesis ensures that the lifetimes of Γ_ϵ are at least as large as the lifetime of Γ. ∎

Remark 9.1.6. Using the Hopf-Rinow theorem we conclude that if M is connected and M is a compact manifold or a geodesically complete manifold then every semimartingale Γ on M is admissible.

Definition 9.1.7. Let $\Gamma \in \mathscr{S}(M)$, $X \in \mathscr{S}(\mathbb{R})$, $f \in C^\infty(M)$ and α be a 1-form on M. Given a deformation $\epsilon \mapsto \Gamma_\epsilon$, we define:

1. $\mathcal{D}\int f(\Gamma) \bullet dX = \lim_{\substack{\text{ucp} \\ \epsilon\to 0}} \int \frac{f(\Gamma_\epsilon)-f(\Gamma)}{\epsilon} \bullet dX$.
2. $\mathcal{D}\int \alpha \bullet d\Gamma = \lim_{\substack{\text{ucp} \\ \epsilon\to 0}} \frac{1}{\epsilon}\left(\int \alpha \bullet d\Gamma_\epsilon - \int \alpha \bullet d\Gamma\right)$.

Remark 9.1.8. The notation $\mathcal{D}$ is used as opposed to δ in order to distinguish between ucp convergence and semimartingale convergence.

The next lemma prescribes a method for computing variations of Stratonovich integrals.

Lemma 9.1.9. Let Γ be a semimartingale in a manifold M and $\epsilon \mapsto \Gamma_\epsilon$ be a deformation of Γ.

1. For every real semimartingale X and $f \in C^\infty(M)$

$$\mathcal{D}\int f(\Gamma) \bullet dX = \int df(\delta\Gamma) \bullet dX \tag{9.1.1}$$

2. For every 1-form α on M

$$\mathcal{D}\int \alpha(\Gamma)\bullet d\Gamma = \int i_{\delta\Gamma}\mathbf{d}\alpha \bullet d\Gamma + \langle \alpha(\Gamma), \delta\Gamma\rangle - \langle \alpha(\Gamma_0), \delta\Gamma_0\rangle, \tag{9.1.2}$$

where $\mathbf{d}\alpha$ denotes the exterior derivative of α.

Proof. The first statement follows by applying Proposition 3.1.27 to $Z_\epsilon := \frac{f(\Gamma_\epsilon)-f(\Gamma)}{\epsilon}$.

The proof of the second statement is similar to the proof of Proposition 4.3 in Lázaro-Camí and Ortega [55]. By the Whitney Embedding Theorem, there exists a positive integer ℓ such that M can be smoothly embedded as a submanifold in $\mathbb{R}^\ell$. In the embedded picture, we can write $\alpha(p) = \sum_{i=1}^{\ell} \alpha_i(p) df_i(p)$ where $\alpha_1, \cdots, \alpha_\ell, f_1, \cdots, f_\ell \in C^\infty(\mathbb{R}^\ell)$. Then

$$\begin{aligned}
&\frac{1}{\epsilon}\left[\int \alpha \bullet d\Gamma_\epsilon - \int \alpha \bullet d\Gamma\right] \\
&= \sum_{i=1}^{\ell} \frac{1}{\epsilon}\left[\int \alpha_i(\Gamma_\epsilon)\bullet df_i(\Gamma_\epsilon) - \int \alpha_i(\Gamma)\bullet df_i(\Gamma)\right] \\
&= \sum_{i=1}^{\ell}\left[\frac{\int \alpha_i(\Gamma_\epsilon)\bullet df_i(\Gamma_\epsilon) - \int \alpha_i(\Gamma_\epsilon)\bullet df_i(\Gamma)}{\epsilon}\right. \\
&\left. + \frac{\int \alpha_i(\Gamma_\epsilon)\bullet df_i(\Gamma) - \int \alpha_i(\Gamma)\bullet df_i(\Gamma)}{\epsilon}\right]
\end{aligned}$$

Given $i \in \{1, \cdots, k\}$, first consider the term

$$\frac{\int \alpha_i(\Gamma_\epsilon)\bullet df_i(\Gamma_\epsilon) - \int \alpha_i(\Gamma_\epsilon)\bullet df_i(\Gamma)}{\epsilon} = \int \alpha_i(\Gamma_\epsilon)\bullet d\left(\frac{f_i(\Gamma_\epsilon)-f_i(\Gamma)}{\epsilon}\right).$$

We claim that this term converges to $\int \alpha_i(\Gamma)\bullet d\langle df_i(\Gamma), \delta\Gamma\rangle$ in ucp as ϵ goes to 0. To prove this, first note that as $\epsilon \to 0$, $\alpha_i(\Gamma_\epsilon)$ converges to $\alpha_i(\Gamma)$ and $\frac{f_i(\Gamma_\epsilon)-f_i(\Gamma)}{\epsilon}$ converges to $\langle df_i(\Gamma), \delta\Gamma\rangle$ in the semimartingale topology. Consequently, by Proposition 3.3.5, the claim holds.

Using the product rule for Stratonovich integrals, we write

$$\begin{aligned}
&\langle \alpha_i(\Gamma) df_i(\Gamma), \delta\Gamma\rangle - \langle \alpha_i(\Gamma) df_i(\Gamma_0), \delta\Gamma_0\rangle \\
&= \int \alpha_i(\Gamma)\bullet d\langle df_i(\Gamma), \delta\Gamma\rangle + \int \langle df_i, \delta\Gamma\rangle \bullet d\alpha_i(\Gamma).
\end{aligned}$$

Now we discuss the second term, that is,

$$\frac{\int \alpha_i(\Gamma_\epsilon)\bullet df_i(\Gamma) - \int \alpha_i(\Gamma)\bullet df_i(\Gamma)}{\epsilon}.$$

By definition of $\delta\Gamma$ this converges in ucp to

$$\int \langle d\alpha_i(\Gamma), \delta\Gamma\rangle \bullet df_i(\Gamma)$$

as $\epsilon \to 0$. Consequently, for the i-th term

$$\begin{aligned}
&\left[\frac{\int \alpha_i(\Gamma_\epsilon) \bullet df_i(\Gamma_\epsilon) - \int \alpha_i(\Gamma_\epsilon) \bullet df_i(\Gamma)}{\epsilon}\right. \\
&\left.+\frac{\int \alpha_i(\Gamma_\epsilon) \bullet df_i(\Gamma) - \int \alpha_i(\Gamma) \bullet df_i(\Gamma)}{\epsilon}\right] \\
&\xrightarrow[ucp]{\epsilon\to 0} \langle \alpha_i(\Gamma) df_i(\Gamma), \delta\Gamma\rangle - \langle \alpha_i(\Gamma) df_i(\Gamma_0), \delta\Gamma_0\rangle \\
&- \int \langle df_i, \delta\Gamma\rangle \bullet d\alpha_i(\Gamma) + \int \langle d\alpha_i(\Gamma), \delta\Gamma\rangle \bullet df_i(\Gamma) \\
&= \int i_{\delta\Gamma}(d\alpha^i \wedge df_i) \bullet d\Gamma + \langle \alpha_i(\Gamma) df_i(\Gamma), \delta\Gamma\rangle - \langle \alpha_i(\Gamma) df_i(\Gamma_0), \delta\Gamma_0\rangle \,.
\end{aligned}$$

The result follows by summing over i from 1 to ℓ. ∎

The next result may be obtained either by purely deterministic arguments or as a corollary of the previous lemma:

Corollary 9.1.10. Let $\gamma(t)$ be a smooth curve in M and α be a 1-form on M. Suppose $\epsilon \mapsto \gamma_\epsilon(t)$ is a variation of $\gamma(t)$. Then

$$\delta \langle \alpha(\gamma(t)), \dot\gamma(t)\rangle = i_{\delta\gamma}\mathbf{d}\alpha(\dot\gamma(t)) + \frac{d}{dt}\langle \alpha(\gamma(t)), \delta\gamma\rangle \,. \tag{9.1.3}$$

9.1.1 Fixed Endpoint Variations

We will assume that Γ is an admissible semimartingale in M. Let $T > 0$ be fixed. Suppose $g \in C^\infty(\mathbb{R})$ is supported on $(0,T)$ and $X \in \mathfrak{X}^\infty(M)$. Then $Y_t = g(t)X(\Gamma_t)$ is a semimartingale in TM over Γ (that is, the projection of Y on M is Γ) that vanishes at $t = 0$ and $t = T$. Then there exists a deformation $\epsilon \mapsto \Gamma_\epsilon$ of Γ such that $\delta\Gamma = Y$.

A second way to construct variations that vanish at $t = 0$ and $t = T$ is inspired by the works of Arnaudon, Chen and Cruzeiro [7] and Huang and Zambrini [44]. Assume M is equipped with a connection, $\Gamma_0 = a$ for some $a \in M$ and let $\|_{0\to t}^{\Gamma} v$ denote the parallel transport of a vector $v \in T_aM$ along Γ. Let $v(t)$ be a deterministic curve in T_aM such that $v(0) = v(T) = 0$. Then $Y_t := \|_{0\to t}^{\Gamma} v(t)$ is the TM-valued semimartingale over Γ such that $Y_0 = Y_T = 0$. The admissibility hypothesis ensures that there exists a deformation $\epsilon \mapsto \Gamma_\epsilon$ of Γ with $\delta\Gamma = Y$.

Now, we describe variations that vanish locally. Given a closed subset $K \subseteq M$ we describe how to construct variations of the portion of Γ that lies in K. Recall that if τ_K^h is the hitting time for K and if $\tau_K^{(h,e)} := \tau_K^e\left(\Gamma_{t+\tau_K^h(\Gamma)}\right)$ then $\Gamma|_{[[\tau_K^h,\tau_K^h+\tau_K^{(h,e)}]]}$ is the portion of Γ that lies in K. Let $f \in C^\infty(M)$ be supported on the interior int K of K and $X \in \mathfrak{X}^\infty(M)$. Then $f\cdot X$ vanishes outside int K. It follows that $\tilde Y = f\cdot X(\Gamma)$ vanishes on $[[0,\infty[[\backslash]]\tau_K^h, \tau_K^h + \tau_K^{(h,e)}[[= [[0,\tau_K^h]]\bigcup[[\tau_K^h+\tau_K^{(h,e)},\infty[[$. Let $g \in C^\infty(\mathbb{R})$ be supported on $(0,T)$. Then $Y_t := g(t)\tilde Y_t$

is a TM valued semimartingale that not only vanishes on $[[0, \tau_K^h]] \bigcup [[\tau_K^h + \tau_K^{(h,e)}, \infty[[$, but also for all $t \geq T$. The admissibility hypothesis shows that there exists a deformation $\epsilon \mapsto \Gamma_\epsilon$ of Γ with $\delta\Gamma = Y$.

Definition 9.1.11. Let $K \subseteq M$ be a closed subset and Γ be an admissible semimartingale in M.

1. A K**-deformation** of Γ is a deformation $\epsilon \mapsto \Gamma_\epsilon$ of Γ such that $\delta\Gamma$ vanishes outside $]]\tau_K^h, \tau_K^h + \tau_K^{(h,e)}[[$. The corresponding variation will be called a K**-variation**.
2. Given $T > 0$, a (K,T)**-deformation** of Γ is a K-deformation $\epsilon \mapsto \Gamma_\epsilon$ of Γ such that $\delta\Gamma$ also vanishes on $[[\left(\tau_K^h + \tau_K^{(h,e)}\right) \wedge T, \infty[[$. The associated variation $\delta\Gamma$ will be called a (K,T)**-variation**.

Lemma 9.1.12. Let $\epsilon \mapsto \Gamma_\epsilon$ be a K-deformation of Γ, where $K \subset M$ is closed. Then:

1. For every $f \in C^\infty(M)$

$$\mathcal{D} \int_0^T f(\Gamma) \bullet dX = \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} df(\delta\Gamma^{|T}) \bullet dX$$

.

2. For every 1-form α on M

$$\mathcal{D} \int_0^T \alpha \bullet d\Gamma = \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} i_{\delta\Gamma^{|T}} \mathbf{d}\alpha \bullet d\Gamma^{|T} + \langle \alpha(\Gamma_T), \delta\Gamma_T \rangle - \langle \alpha(\Gamma_0), \delta\Gamma_0 \rangle .$$

Proof. We only prove (1) since the proof of (2) is similar. It follows from the definition that $\delta\Gamma$ vanishes outside $]]\tau_K^h, \tau_K^h + \tau_K^{(h,e)}[[$. Let $\nVdash_{(\cdot)}$ denote the indicator function. Using Proposition 5.3 in Lázaro-Camí and Ortega [55], we have

$$\begin{aligned} \mathcal{D} \int_0^T f(\Gamma) \bullet dX &= \int_0^T df(\delta\Gamma) \bullet dX \\ &= \int \nVdash_{[0,T]} \nVdash_{[[\tau_K^h, \tau_K^h + \tau_K^{(h,e)}]]} df(\delta\Gamma) \bullet dX \\ &= \int \nVdash_{[[\tau_K^h, \tau_K^h + \tau_K^{(h,e)}]]} df(\delta\Gamma^{|T}) \bullet dX \\ &= \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} df(\delta\Gamma^{|T}) \bullet dX. \end{aligned}$$

∎

9.1.2 A Stochastic Analogue of the Fundamental Lemma of the Calculus of Variations

We will formulate a stochastic analogue of the fundamental lemma of the calculus of variations in coordinate charts. Let $\langle\cdot,\cdot\rangle$ denote the standard Euclidean inner product and $(e_1,\cdots,e_n)$ denote the standard basis of $\mathbb{R}^n$.

Lemma 9.1.13. Let M be a smooth n-manifold and $U \subseteq M$ be a coordinate chart. We identify U with an open subset of $\mathbb{R}^n$, also denoted by U. Let $\Gamma \in \mathscr{S}(M)$ be admissible and $\Xi : \mathscr{S}(M) \to \mathscr{S}(\mathbb{R}^n)$ be such that $\Xi(\Gamma)_{A_t} = \Xi(\Gamma_{A_t})$ for any continuous change of time $t \mapsto A_t$. If for every $(\bar{U},T)$-deformation $\epsilon \mapsto \Gamma_\epsilon$ we have

$$\int_{\tau_{\bar{U}}^h}^{\tau_{\bar{U}}^h+\tau_{\bar{U}}^{(h,e)}} \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle = 0,$$

then $\bullet d\Xi(\Gamma^{|T}) = 0$ in $]]\tau_U^h, \tau_U^h + \tau_U^{(h,e)}[[$. Here $\bullet d\Xi(\Gamma^{|T}) = 0$ means that $\Xi(\Gamma^{|T}) - \Xi(\Gamma^{|T})_{\tau_U^h} = 0$ a.s. in $]]\tau_U^h, \tau_U^h + \tau_U^{(h,e)}[[$.

Proof. First suppose U is a precompact coordinate ball and identify U with the open ball $B_r(0)$ of radius r and centered at 0. Given $s > 0$ let (g_n) be a sequence in $C^\infty(\mathbb{R})$ such that g_n is supported in $(0, s+1)$ for every n and $g_n \to \nVdash_{(0,s]}$ pointwise. Then $g_n(t)\delta\Gamma_t^{|T}$ is a $(\bar{U},T)$ variation of Γ. Using the fact that g_n is of bounded variation, the Itô dominated convergence theorem and Proposition 5.3 of Lázaro-Camí and Ortega [55], we obtain

$$\begin{aligned}
0 &= \int_{\tau_{\bar{U}}^h}^{\tau_{\bar{U}}^h+\tau_{\bar{U}}^{(h,e)}} \left\langle g_n(t)\delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle \\
&= \int_{\tau_{\bar{U}}^h}^{\tau_{\bar{U}}^h+\tau_{\bar{U}}^{(h,e)}} g_n(t) \bullet d\left(\int \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle\right) \\
&= \int_{\tau_{\bar{U}}^h}^{\tau_{\bar{U}}^h+\tau_{\bar{U}}^{(h,e)}} g_n(t)\, d\left(\int \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle\right) \\
&\xrightarrow[n\to\infty]{ucp} \int_{\tau_{\bar{U}}^h}^{\tau_{\bar{U}}^h+\tau_{\bar{U}}^{(h,e)}} \mathbf{1}_{(0,s]}\, d\left(\int \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle\right) \\
&= \int \mathbf{1}_{(0,s]} \mathbf{1}_{[[\tau_{\bar{U}}^h, \tau_{\bar{U}}^h+\tau_{\bar{U}}^{(h,e)}]]}\, d\left(\int \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle\right) \\
&= \int_0^s \mathbf{1}_{[[\tau_{\bar{U}}^h, \tau_{\bar{U}}^h+\tau_{\bar{U}}^{(h,e)}]]} \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle.
\end{aligned}$$

Since this holds for all $s > 0$ we conclude that

$$\int \mathbf{1}_{[[\tau_{\bar{U}}^h, \tau_{\bar{U}}^h+\tau_{\bar{U}}^{(h,e)}]]} \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle = 0.$$

Moreover, since $\delta\Gamma^{|T}$ vanishes outside $[[\tau^h_{\bar U}, \tau^h_{\bar U}+\tau^{(h,e)}_{\bar U}]]$, we have

$$\int \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle = 0$$

in $[[\tau^h_{\bar U}, \tau^h_{\bar U}+\tau^{(h,e)}_{\bar U}]]$.

Let $0<\eta<r$ and $h\in C^\infty(\mathbb{R}^n)$ be supported on $B_r(0)$ with $h|_{\bar B_\eta(0)}=1$, where $\bar B_\eta(0)$ denotes the closed ball of radius η centered at 0. For $j=1,\cdots,n$, let $\tilde X$ denote the vector field on $\mathbb{R}$ defined by $\tilde X_j = h(x)e_j$. Then $\tilde X_j$ vanishes on the boundary $\partial\bar B_r(0)$ of $\bar B_r(0)=\bar U$. Extending $\tilde X_j|_{\bar U}$ to a vector field X on M by setting $X=0$ outside U and letting $g\in C^\infty(M)$ be supported on $(0,T)$, we can construct a $(\bar U, T)$-deformation $\epsilon\mapsto\Gamma_\epsilon$ with variation $g(t)X(\Gamma_t)$ by the admissibility hypothesis.

Since $\int\left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle = 0$ in $[[\tau^h_{\bar U}, \tau^h_{\bar U}+\tau^{(h,e)}_{\bar U}]]$ and $[[\tau^h_{\bar B_\eta(0)}, \tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}]] \subseteq [[\tau^h_{\bar U}, \tau^h_{\bar U}+\tau^{(h,e)}_{\bar U}]]$, it follows that $\int\left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T})\right\rangle = 0$ in $[[\tau^h_{\bar B_\eta(0)}, \tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}]]$. Let Z_j denote the jth component of $\Xi(\Gamma^{|T})$. Then the previous equality and the fact that $X(\Gamma^{|T})=g(t)e_j$ on $[[\tau^h_{\bar B_\eta(0)}, \tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}]]$ implies that

$$\int \mathbf{1}_{[[\tau^h_{\bar B_\eta(0)}, \tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}]]} g(t)\bullet dZ_j = 0$$

for all $g\in C^\infty(\mathbb{R})$ supported on $(0,T)$.

Pick an arbitrary $s\in(0,T)$. Replacing g by $\tilde g_n$ where $(\tilde g_n)$ is a sequence in $C^\infty(\mathbb{R})$ that is supported on $(0,T)$ and $\tilde g_n\to\mathbf{1}_{(0,s]}$ pointwise, we get

$$\begin{aligned}
0 &= \int \mathbf{1}_{[[\tau^h_{\bar B_\eta(0)}, \tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}]]} \tilde g_n(t)\bullet dZ_{j_t}\\
&= \int \mathbf{1}_{[[\tau^h_{\bar B_\eta(0)}, \tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}]]} \tilde g_n(t)\, dZ_{j_t}\\
&\xrightarrow[n\to\infty]{ucp} \int \mathbf{1}_{[[\tau^h_{\bar B_\eta(0)}, \tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}]]} \mathbf{1}_{(0,s]}\, dZ_{j_t}\\
&= \int_{\tau^h_{\bar B_\eta(0)}}^{\tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}} \mathbf{1}_{(0,s]}\, dZ_j\\
&= \int_{\tau^h_{\bar B_\eta(0)}}^{\tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}} dZ_j^{|s}\\
&= \left(Z^{|s}_{j_{\left(\tau^h_{\bar B_\eta(0)}+\tau^{(h,e)}_{\bar B_\eta(0)}\right)}} - Z^{|s}_{j_{\tau^h_{\bar B_\eta(0)}}} \right)
\end{aligned}$$

and we used the fact that $\tilde g_n$ is of bounded variation, the Itô dominated convergence theorem and Proposition 5.3 in [55]. Since this is true for all $0<s<T$, it follows that $\bullet dZ_j = 0$

in $[[\tau^h_{\bar{B}_\eta(0)}, \tau^h_{\bar{B}_\eta(0)} + \tau^{(h,e)}_{\bar{B}_\eta(0)}]]$. Since this holds for all $0 < \eta < r$, we conclude that $\bullet dZ_j = 0$ in $]]\tau^h_{B_r(0)}, \tau^h_{B_r(0)} + \tau^{(h,e)}_{B_r(0)}[[=]]\tau^h_U, \tau^h_U + \tau^{(h,e)}_U[[$. Consequently $\bullet d\Xi(\Gamma^{|T}) = 0$ in $]]\tau^h_U, \tau^h_U + \tau^{(h,e)}_U[[$.

Now suppose $U \subseteq M$ is a coordinate chart in M. For every precompact coordinate ball U_0 in U, note that a $(\bar{U}_0, T)$-deformation of Γ is also a $(\bar{U}, T)$-deformation of Γ. By our hypothesis, for every $(\bar{U}_0, T)$-deformation of Γ, we have

$$\int_{\tau^h_{\bar{U}}}^{\tau^h_{\bar{U}} + \tau^{(h,e)}_{\bar{U}}} \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T}) \right\rangle = \int_{\tau^h_{\bar{U}_0}}^{\tau^h_{\bar{U}_0} + \tau^{(h,e)}_{\bar{U}_0}} \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T}) \right\rangle = 0.$$

This implies that $\bullet d\Xi(\Gamma^{|T}) = 0$ in $]]\tau^h_{U_0}, \tau^h_{U_0} + \tau^{(h,e)}_{U_0}[[$. Since this holds for all precompact coordinate balls $U_0 \subseteq U$, we have $\bullet d\Xi(\Gamma^{|T}) = 0$ in $]]\tau^h_U, \tau^h_U + \tau^{(h,e)}_U[[$. ∎

9.2 The Stochastic Hamilton-Pontryagin Principle

A stochastic extension of the deterministic Hamilton-Pontryagin principle was first introduced by Bou-Rabee and Owhadi [13] and has been generalized more recently by Street and Takao [89]. We will provide a proof of the local form of the stochastic Hamilton-Pontryagin principle as an application of the variational framework developed in the last section to stochastic geometric mechanics. Then we will develop a stochastic version of Noether's theorem. This will be followed by a discussion of the intrinsic form of the stochastic Hamilton-Pontryagin principle, where we will use Stratonovich operators to provide a global description.

Let us briefly recall the deterministic Hamilton-Pontryagin principle. Given a configuration manifold Q let $\mathcal{P}Q := TQ \oplus T^*Q$ denote its Pontryagin bundle. Local coordinates on $\mathcal{P}Q$ will be denoted by (q, v, p). Let $\mathcal{L} \in C^\infty(TQ)$ be a Lagrangian. The deterministic Hamilton-Pontryagin principle states that a $\mathcal{P}Q$-valued curve $(q(t), v(t), p(t))$ is a critical point of the action

$$\int_{t_0}^{t_1} [\mathcal{L}(q(t), v(t)) + \langle p(t), \dot{q}(t) - v(t) \rangle] dt$$

amongst all curves such that $q(t_0)$ and $q(t_1)$ are fixed (and variations in $v(t)$ and $p(t)$ are arbitrary), if and only if $(q(t), v(t), p(t))$ satisfies the implicit Euler-Lagrange equations given by

$$\dot{q} = v, \quad p = \frac{\partial \mathcal{L}}{\partial v}, \quad \dot{p} = \frac{\partial \mathcal{L}}{\partial q}.$$

We refer to [94] for more details on the deterministic Hamilton-Pontryagin principle and its application to constrained systems.

Definition 9.2.1. Let $X = (X^0, \cdots, X^k) \in \mathscr{S}\left(\mathbb{R}^{k+1}\right)$ and $\mathcal{L} \in C^\infty(TQ)$ be a Lagrangian. Suppose we have $L_1, \cdots, L_k \in C^\infty(Q)$ and vector fields $V_1, \cdots, V_k$ on Q. Given an admissible

$\mathcal{P}Q$-valued semimartingale $\Gamma_t = (q_t, v_t, p_t)$ we define the **stochastic Hamilton-Pontryagin action integral** as

$$\mathcal{S}_X(\Gamma) = \int_0^T \Bigg(\mathcal{L}(q_t, v_t) \bullet dX_t^0 + \sum_{i=1}^k L_i(q_t) \bullet dX_t^i \\ + \Bigg\langle p_t, \bullet dq_t - v_t \bullet dX_t^0 - \sum_{i=1}^k V_i(q_t) \bullet dX_t^i \Bigg\rangle \Bigg) . \tag{9.2.1}$$

In Bou-Rabee and Owhadi [13] the authors consider $X = (t, B_t^1, \cdots, B_t^k)$, where B^i is a Brownian motion, and $V_i = 0$. Street and Takao [89] have generalized this to the case where X is a driving semimartingale and Γ is compatible with X. The reader is referred to [89] as well as Street and Crisan [88] for further details on driving semimartingales and the compatibility hypothesis, as well as a different stochastic analogue of the fundamental lemma of the calculus of variations under these assumptions. We will only assume that $X \in \mathscr{S}\left(\mathbb{R}^{k+1}\right)$ and in particular, we will forego the assumption that $X^0 = t$.

Let us also recall the intrinsic description of the action functional. Let $G : \mathcal{P}Q \to \mathbb{R}$ denote the fibrewise pairing map between TQ and T^*Q, that is, $G(q,v,p) = \langle p, v\rangle$. The following maps

$$\begin{aligned} \mathrm{Pr}_{\mathcal{P}Q} &: T\mathcal{P}Q \to \mathcal{P}Q \\ \mathrm{Pr}_{T\mathcal{P}Q} &: TT\mathcal{P}Q \to T\mathcal{P}Q \\ \mathrm{pr}_Q &: \mathcal{P}Q \to Q \\ \mathrm{pr}_{TQ} &: \mathcal{P}Q \to TQ \\ \mathrm{pr}_{T^*Q} &: \mathcal{P}Q \to T^*Q \end{aligned} \tag{9.2.2}$$

are the corresponding projections. We also have a map $\rho_{TT^*Q} : TT^*Q \to \mathcal{P}Q$, defined in local coordinates by setting $\rho_{TT^*Q}(q, p, v_q, v_p) = (q, v_q, p)$. Then we obtain a 1-form $\mathcal{G}$ on $\mathcal{P}Q$ given by $\mathcal{G} = G \circ \rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}$. In local coordinates, if $(u_q, u_v, u_p) \in T_{(q,v,p)}\mathcal{P}Q$ then

$$\mathcal{G}(q,v,p)(u_q,u_v,u_p) = G(q,u_q,p) = \langle p, u_q\rangle . \tag{9.2.3}$$

Consequently, if $\Gamma_t = (q_t, v_t, p_t)$ then $\int \mathcal{G} \bullet d\Gamma = \int \langle p_t, \bullet dq_t\rangle$.

Given a vector field $V \in \mathfrak{X}^\infty(Q)$ define $\tilde{V} : \mathcal{P}Q \to \mathcal{P}Q$ by $\tilde{V}(x) = (V \circ \mathrm{pr}_Q(x)) \oplus \mathrm{pr}_{T^*Q}(x) \in \mathcal{P}Q$. Written in local coordinates this reads $\tilde{V}(q,v,p) = (q, V(q), p)$. For every $j \in \{0, \cdots, k\}$ define the **generalized energy** $E_j : \mathcal{P}Q \to \mathbb{R}$ by

$$E_j = \begin{cases} G - \mathcal{L} \circ \mathrm{pr}_{TQ}, & \text{if } j = 0, \\ G \circ \tilde{V}_j - L_j \circ \mathrm{pr}_Q, & \text{if } j = 1, \ldots, k. \end{cases}$$

In coordinates, $E_0(q,v,p) = \langle p, v\rangle - \mathcal{L}(q,v)$ and $E_i(q,v,p) = \langle p, V_i(q)\rangle - L_i(q)$, for $i = 1, \cdots, k$. The generalized energies E_i for $i = 1, \cdots, k$ also appear in Street and Takao [89].

We note that if $\Gamma_t = (q_t, v_t, p_t)$ in local coordinates then

$$E_j(\Gamma_t) = E_j(q_t, v_t, p_t) = \begin{cases} \langle p_t, v_t \rangle - \mathcal{L}(q_t, v_t), & \text{if } j = 0, \\ \langle p_t, V_j(q_t) \rangle - L_j(q_t), & \text{if } j = 1, \ldots, k. \end{cases}$$

Hence

$$\mathcal{S}_X(\Gamma) = \int_0^T \mathcal{G} \bullet d\Gamma - \sum_{j=0}^{k} \int_0^T E_j(\Gamma) \bullet dX_j. \tag{9.2.4}$$

9.2.1 The Local Form of the Stochastic Hamilton-Pontryagin Principle

First we describe variations in local coordinates of the terms in the Hamilton-Pontryagin action integral.

Lemma 9.2.2. Let Γ be an admissible semimartingale on $\mathcal{P}Q$ and let $\Gamma_t = (q_t, v_t, p_t)$ in local coordinates. Suppose $\epsilon \mapsto \Gamma_{\epsilon,t} = (q_{\epsilon,t}, v_{\epsilon,t}, p_{\epsilon,t})$ be a deformation of Γ. Then

$$\begin{aligned} \mathcal{D} \int_0^T \mathcal{G} \bullet d\Gamma &= \mathcal{D} \int_0^T \langle p_t, \bullet dq_t \rangle \\ &= \int_0^T \left\langle \delta p_t^{|T}, \bullet dq_t^{|T} \right\rangle - \int_0^T \left\langle \bullet dp_t^{|T}, \delta q_t^{|T} \right\rangle \\ &\quad + \langle p_T, \delta q_T \rangle - \langle p_0, \delta q_0 \rangle . \end{aligned}$$

Proof. In local coordinates $\langle \mathcal{G}(\Gamma_t), \delta\Gamma_t \rangle = \langle p_t, \delta q_t \rangle$. Then, by Lemma 9.1.9

$$\begin{aligned} \mathcal{D} \int_0^T \mathcal{G} \bullet d\Gamma &= \mathcal{D} \int_0^T \mathcal{G} \bullet d\Gamma^{|T} \\ &= \int_0^T i_{\delta\Gamma^{|T}} \mathbf{d}\mathcal{G} \bullet d\Gamma^{|T} + \langle \mathcal{G}(\Gamma_T), \delta\Gamma_T \rangle - \langle \mathcal{G}(\Gamma_0), \delta\Gamma_0 \rangle \\ &= \int_0^T i_{\delta\Gamma^{|T}} \mathbf{d}\mathcal{G} \bullet d\Gamma^{|T} + \langle p_T, \delta q_T \rangle - \langle p_0, \delta q_0 \rangle . \end{aligned}$$

Let $(q, v, p) \in \mathcal{P}Q$. Suppose $(\dot{q}, \dot{v}, \dot{p}), (w_q, w_v, w_p) \in T_{(q,v,p)}\mathcal{P}Q$. From the local coordinate expression of $\mathcal{G}$ in Eq. (9.2.3), it follows that $\mathcal{G}(q, v, p) = pdq$. Hence

$$\mathbf{d}\mathcal{G}(q, v, p) = \sum_{i=1}^{\dim Q} dp_i \wedge dq^i,$$

which yields

$$i_{(w_q, w_v, w_p)} \mathbf{d}\mathcal{G}(\dot{q}, \dot{v}, \dot{p}) = \mathbf{d}\mathcal{G}(q, v, p)((w_q, w_v, w_p), (\dot{q}, \dot{v}, \dot{p})) = \langle w_p, \dot{q} \rangle - \langle w_q, \dot{p} \rangle .$$

Consequently, since the tangent vectors considered are arbitrary, we get

$$\int_0^T i_{\delta\Gamma^{|T}} \mathbf{d}\mathcal{G} \bullet d\Gamma^{|T} = \int_0^T \left\langle \delta p_t^{|T}, \bullet dq_t^{|T} \right\rangle - \int_0^T \left\langle \bullet dp_t^{|T}, \delta q_t^{|T} \right\rangle .$$

This concludes the proof. ∎

Remark 9.2.3. The product rule is often used to prove the above lemma. Mimicking the product rule we write

$$\mathcal{D} \int_0^T \langle p_t^{|T}, \bullet dq_t^{|T} \rangle = \int_0^T \left\langle \delta p_t^{|T}, \bullet dq_t^{|T} \right\rangle + \int_0^T \left\langle p_t^{|T}, \delta \left(\bullet dq_t^{|T} \right) \right\rangle .$$

But the term $\delta \left(\bullet dq_t^{|T} \right)$ is not defined since $\bullet dq_t^{|T}$ is not a stochastic process. To define this, we recall that the $\delta q_t^{|T}$ is assumed to be a semimartingale by definition. Hence we can set

$$\int_0^T \left\langle p_t^{|T}, \delta \left(\bullet dq_t^{|T} \right) \right\rangle = \int_0^T \left\langle p_t^{|T}, \bullet d \left(\delta q_t^{|T} \right) \right\rangle .$$

The Stratonovich product rule gives us

$$\bullet d \left\langle p_t^{|T}, \delta q_t^{|T} \right\rangle = \left\langle \bullet dp_t^{|T}, \delta q_t^{|T} \right\rangle + \left\langle p_t^{|T}, \bullet d \left(\delta q_t^{|T} \right) \right\rangle$$

which implies

$$\langle p_T, \delta q_T \rangle - \langle p_0, \delta q_0 \rangle = \int_0^T \left\langle \bullet dp_t^{|T}, \delta q_t^{|T} \right\rangle + \int_0^T \left\langle p_t^{|T}, \bullet d \left(\delta q_t^{|T} \right) \right\rangle .$$

Therefore

$$\begin{aligned} \mathcal{D} \int_0^T \langle p_t, \bullet dq_t \rangle = \int_0^T \left\langle \delta p_t^{|T}, \bullet dq_t^{|T} \right\rangle - \int_0^T \left\langle \bullet dp_t^{|T}, \delta q_t^{|T} \right\rangle \\ + \langle p_T, \delta q_T \rangle - \langle p_0, \delta q_0 \rangle . \end{aligned}$$

Lemma 9.2.4. Let Γ be an admissible semimartingale on $\mathcal{P}Q$ and let $\Gamma_t = (q_t, v_t, p_t)$ in local coordinates. Suppose $\epsilon \mapsto \Gamma_{\epsilon,t} = (q_{\epsilon,t}, v_{\epsilon,t}, p_{\epsilon,t})$ be a deformation of Γ. Then

$$\begin{aligned}
& - \mathcal{D} \sum_{j=0}^k \int_0^T E_j(\Gamma) \bullet dX^j \\
& = \mathcal{D} \left[\int_0^T \left((\mathcal{L}(q_t, v_t) - \langle p_t, v_t \rangle) \bullet dX_t^0 + \sum_{i=1}^k (L_i(q_t) - \langle p_t, V_i(q_t) \rangle) \bullet dX_t^i \right) \right] \\
& = \int_0^T \frac{\partial}{\partial q_t^{|T}} \left\langle \left(\mathcal{L} \bullet dX_t^0 + \sum_{i=1}^k \left(L_i - \left\langle p_t^{|T}, V_i(q_t^{|T}) \right\rangle \right) \bullet dX_t^i \right), \delta q_t^{|T} \right\rangle \\
& + \int_0^T \left\langle \left(p_t^{|T} - \frac{\partial \mathcal{L}}{\partial v_t^{|T}} \right) \bullet dX_t^0, \delta v_t^{|T} \right\rangle \\
& - \int_0^T \left\langle \delta p_t^{|T}, v_t^{|T} \bullet dX_t^0 + \sum_{i=1}^k V_i(q_t^{|T}) \bullet dX_t^i \right\rangle .
\end{aligned}$$

Proof. By Lemma 9.1.9 we have

$$
\begin{aligned}
&-\mathcal{D}\sum_{j=0}^{k}\int_0^T E_j(\Gamma)\bullet dX^j\\
&= -\mathcal{D}\sum_{j=0}^{k}\int_0^T \left(\left\langle \frac{\partial E_j}{\partial q_t^{|T}}, \delta q_t^{|T}\right\rangle + \left\langle \frac{\partial E_j}{\partial v_t^{|T}}, \delta v_t^{|T}\right\rangle + \left\langle \frac{\partial E_j}{\partial p_t^{|T}}, \delta p_t^{|T}\right\rangle\right)\bullet dX_t^j\\
&= \int_0^T \left(\left\langle \frac{\partial \mathcal{L}}{\partial q_t^{|T}}, \delta q_t^{|T}\right\rangle + \left\langle \frac{\partial \mathcal{L}}{\partial v_t^{|T}}, \delta v_t^{|T}\right\rangle - \left\langle \delta p_t^{|T}, v_t^{|T}\right\rangle - \left\langle p_t^{|T}, \delta v_t^{|T}\right\rangle\right)\bullet dX_t^0\\
&+\sum_{i=1}^{k}\int_0^T \left(\left\langle \frac{\partial}{\partial q_t^{|T}}\left(L_i(q_t^{|T}) - \left\langle p_t^{|T}, V_i(q_t^{|T})\right\rangle\right), \delta q_t^{|T}\right\rangle - \left\langle \delta p_t^{|T}, V_i(q_t^{|T})\right\rangle\right)\bullet dX_t^i\\
&= \int_0^T \frac{\partial}{\partial q_t^{|T}}\left\langle\left(\mathcal{L}\bullet dX_t^0 + \sum_{i=1}^{k}\left(L_i - \left\langle p_t^{|T}, V_i(q_t^{|T})\right\rangle\right)\bullet dX_t^i\right), \delta q_t^{|T}\right\rangle\\
&+\int_0^T \left\langle\left(p_t^{|T} - \frac{\partial \mathcal{L}}{\partial v_t^{|T}}\right)\bullet dX_t^0, \delta v_t^{|T}\right\rangle\\
&-\int_0^T \left\langle \delta p_t^{|T}, v_t^{|T}\bullet dX_t^0 + \sum_{i=1}^{k} V_i(q_t^{|T})\bullet dX_t^i\right\rangle.
\end{aligned}
$$

∎

The local form of the stochastic Hamilton-Pontryagin principle is given by the following theorem:

Theorem 9.2.5. For every semimartingale $X = (X^0, \cdots, X^k)$ on $\mathbb{R}^{k+1}$, if $\Gamma_t = (q_t, v_t, p_t) \in \mathscr{S}(\mathcal{P}Q)$ is admissible then $\mathcal{D}\mathcal{S}_X(\Gamma) = 0$ for all deformations $\epsilon \mapsto \Gamma_\epsilon$ such that $\delta q_t = T\mathrm{pr}_Q(\Gamma_t)$ vanishes at $t = 0$ and $t = T$ if and only if $\Gamma^{|T} = \left(q_t^{|T}, v_t^{|T}, p_t^{|T}\right)$ satisfies the **stochastic implicit Euler-Lagrange equations** given by

$$
\begin{aligned}
\bullet dq_t &= v_t \bullet dX_t^0 + \sum_{i=1}^{k} V_i(q_t)\bullet dX_t^i\\
\bullet dp_t &= \frac{\partial}{\partial q_t}\left(\mathcal{L}\bullet dX_t^0 + \sum_{i=1}^{k}\left(L_i - \langle p_t, V_i(q_t)\rangle\right)\bullet dX_t^i\right)\\
\left(p_t - \frac{\partial \mathcal{L}}{\partial v_t}\right)&\bullet dX_t^0 = 0. \qquad (9.2.5)
\end{aligned}
$$

Proof. Let $\epsilon \mapsto \Gamma_{\epsilon,t} = (q_{\epsilon,t}, v_{\epsilon,t}, p_{\epsilon,t})$ be a deformation of Γ such that $\delta q_t = 0$ at $t = 0$ and $t = T$. Consequently $\langle p_t, \delta q_t\rangle = \langle \mathcal{G}(\Gamma_t), \delta\Gamma_t\rangle$ vanishes at $t = 0$ and $t = T$. Using Lemma 9.1.9

$$\mathcal{D}\mathcal{S}_X(\Gamma) = \mathcal{D}\left[\int_0^T \mathcal{G}\bullet d\Gamma - \sum_{j=0}^{k}\int_0^T E_j(\Gamma)\bullet dX_j\right]$$

$$
\begin{aligned}
&= \int_0^T i_{\delta\Gamma^{|T}}\mathbf{d}\mathcal{G}\bullet d\Gamma^{|T} - \sum_{j=0}^k \int_0^T dE_j(\delta\Gamma^{|T})\bullet dX_j\\
&+ \langle \mathcal{G}(\Gamma_T), \delta\Gamma_T\rangle - \langle \mathcal{G}(\Gamma_0), \delta\Gamma_0\rangle\\
&= \int_0^T i_{\delta\Gamma^{|T}}\mathbf{d}\mathcal{G}\bullet d\Gamma^{|T} - \sum_{j=0}^k \int_0^T dE_j(\delta\Gamma^{|T})\bullet dX_j.
\end{aligned}
$$

Suppose Γ solves the stochastic implicit Euler-Lagrange equations. We use the previous two lemmas to show that the above expression vanishes in local coordinates. We have

$$
\begin{aligned}
&\int_0^T i_{\delta\Gamma^{|T}}\mathbf{d}\mathcal{G}\bullet d\Gamma^{|T} - \sum_{j=0}^k \int_0^T dE_j(\delta\Gamma^{|T})\bullet dX_j\\
&= \int_0^T \Bigg\langle \frac{\partial}{\partial q_t^{|T}}\Big(\mathcal{L}\bullet dX_t^0\\
&+ \sum_{i=1}^k \Big(L_i - \Big\langle p_t^{|T}, V_i(q_t^{|T})\Big\rangle\Big)\bullet dX_t^i\Bigg) - \bullet dp_t^{|T}, \delta q_t^{|T}\Bigg\rangle\\
&+ \int_0^T \Bigg\langle \Bigg(p_t^{|T} - \frac{\partial\mathcal{L}}{\partial v_t^{|T}}\Bigg)\bullet dX_t^0, \delta v_t^{|T}\Bigg\rangle\\
&+ \int_0^T \Bigg\langle \delta p_t^{|T}, \bullet dq_t^{|T} - v_t^{|T}\bullet dX_t^0 - \sum_{i=1}^k V_i(q_t^{|T})\bullet dX_t^i\Bigg\rangle\\
&= \int_0^T \Bigg\langle \Big(\delta q_t^{|T}, \delta v_t^{|T}, \delta p_t^{|T}\Big), \bullet d\Bigg(\Bigg(\int\Bigg(\frac{\partial}{\partial q_t^{|T}}\Big(\mathcal{L}\bullet dX_t^0 +\\
&\sum_{i=1}^k \Big(L_i - \Big\langle p_t^{|T}, V_i(q_t^{|T})\Big\rangle\Big)\bullet dX_t^i\Bigg) - \bullet dp_t^{|T}\Bigg), \int\Bigg(p_t^{|T} - \frac{\partial\mathcal{L}}{\partial v_t^{|T}}\Bigg)\bullet dX_t^0,\\
&\int\Bigg(\bullet dq_t^{|T} - v_t^{|T}\bullet dX_t^0 - \sum_{i=1}^k V_i(q_t^{|T})\bullet dX_t^i\Bigg)\Bigg)\Bigg\rangle\\
&= \int_0^T \Bigg\langle \delta\Gamma_t^{|T}, \bullet d\Bigg(\Bigg(\int\Bigg(\frac{\partial}{\partial q_t^{|T}}\Big(\mathcal{L}\bullet dX_t^0 +\\
&\sum_{i=1}^k \Big(L_i - \Big\langle p_t^{|T}, V_i(q_t^{|T})\Big\rangle\Big)\bullet dX_t^i\Bigg) - \bullet dp_t^{|T}\Bigg), \int\Bigg(p_t^{|T} - \frac{\partial\mathcal{L}}{\partial v_t^{|T}}\Bigg)\bullet dX_t^0,\\
&\int\Bigg(\bullet dq_t^{|T} - v_t^{|T}\bullet dX_t^0 - \sum_{i=1}^k V_i(q_t^{|T})\bullet dX_t^i\Bigg)\Bigg)\Bigg\rangle \qquad (9.2.6)\\
&= 0.
\end{aligned}
$$

We now prove the converse. Let $U \subseteq Q$ be open. Then $K^0 := U \times \mathbb{R}^n \times \mathbb{R}^n$ is an arbitrary chart on $\mathcal{P}Q$. It suffices to show that the stochastic implicit Euler-Lagrange equations are satisfied by $\Gamma_t^{|T} = (q_t^{|T}, v_t^{|T}, p_t^{|T})$ in $]]\tau_{K^0}^h, \tau_{K^0}^h + \tau_{K^0}^{(h,e)}[[$. Let K be the closure of K^0 and $\epsilon \mapsto \Gamma_\epsilon$ be an arbitrary (K,T)-deformation of Γ. Then $\delta q_t = 0$ at $t = 0$ and $t = T$. Given a semimartingale $\Gamma_t = (q_t, v_t, p_t)$ on $\mathcal{P}Q$, define the $\mathbb{R}^{3(\dim Q)}$-valued semimartingale $\Xi(\Gamma)$ in local coordinates by

$$\Xi(q_t, v_t, p_t) = \left(\int \left(\frac{\partial}{\partial q_t^{|T}} \left(\mathcal{L} \bullet dX_t^0 + \sum_{i=1}^k \left(L_i - \left\langle p_t^{|T}, V_i(q_t^{|T}) \right\rangle \right) \bullet dX_t^i \right) - \bullet dp_t^{|T} \right), \right.$$
$$\left. \int \left(p_t^{|T} - \frac{\partial \mathcal{L}}{\partial v_t^{|T}} \right) \bullet dX_t^0, \int \left(\bullet dq_t^{|T} - v_t^{|T} \bullet dX_t^0 - \sum_{i=1}^k V_i(q_t^{|T}) \bullet dX_t^i \right) \right).$$

Since the Stratonovich integral commutes with time changes we have $\Xi(\Gamma_{A_t}) = \Xi(\Gamma)_{A_t}$ for any continuous time change $t \mapsto A_t$. An application of Lemma 9.1.12 shows that the integral from 0 to T in (9.2.6) can be replaced by an integral from τ_K^h to $\tau_K^h + \tau_K^{(h,e)}$. As a result, we obtain

$$\int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} \left\langle \delta\Gamma^{|T}, \bullet d\Xi(\Gamma^{|T}) \right\rangle = 0$$

for every (K,T)-deformation $\epsilon \mapsto \Gamma_\epsilon$ of Γ. By Lemma 9.1.13 $\bullet d\Xi(\Gamma^{|T}) = 0$ in $]]\tau_{K^0}^h, \tau_{K^0}^h + \tau_{K^0}^{(h,e)}[[$. This implies that

$$\bullet dq_t^{|T} = v_t^{|T} \bullet dX_t^0 + \sum_{i=1}^k V_i(q_t^{|T}) \bullet dX_t^i$$

$$\bullet dp_t^{|T} = \frac{\partial}{\partial q_t^{|T}} \left(\mathcal{L} \bullet dX_t^0 + \sum_{i=1}^k \left(L_i - \left\langle p_t^{|T}, V_i(q_t^{|T}) \right\rangle \right) \bullet dX_t^i \right)$$

$$\left(p_t^{|T} - \frac{\partial \mathcal{L}}{\partial v_t^{|T}} \right) \bullet dX_t^0 = 0$$

in $]]\tau_{K^0}^h, \tau_{K^0}^h + \tau_{K^0}^{(h,e)}[[$. This completes the proof. ∎

Remark 9.2.6. If $X^0 = t$ then the second equation relates p_t and v_t via the Legendre transform $p_t = \frac{\partial \mathcal{L}}{\partial v_t}$. Also note that if $X^i = 0$ for all $i = 1, \cdots, k$ then the stochastic implicit Euler-Lagrange equations reduce to the deterministic implicit Euler-Lagrange equations.

Remark 9.2.7. Suppose $\mathcal{L}$ is a hyperregular Lagrangian, $F\mathcal{L} : TQ \to TQ$ is the Legendre transform of $\mathcal{L}$ and $E_{\mathcal{L}}(v_q) := \langle F\mathcal{L}(v_q), v_q \rangle - \mathcal{L}(v_q)$ is the energy of $\mathcal{L}$. Define $\mathcal{H} \in C^\infty(T^*Q)$ by $\mathcal{H} = E_{\mathcal{L}} \circ F\mathcal{L}^{-1}$ and $H_i \in C^\infty(T^*Q)$ by $H_i(q,p) = \langle p, V(q) \rangle - L_i(q)$. Then, Street and Takao [89] show that, Γ solves the stochastic Hamiltonian system

$$\bullet d\Gamma = X_{\mathcal{H}}(\Gamma) \bullet dX^0 + \sum_{i=1}^k X_{H_i}(\Gamma) \bullet dX^i.$$

Remark 9.2.8. In Lázaro-Camí and Ortega [48], the authors develop a stochastic Hamilton-Jacobi equation for stochastic Hamiltonian systems using a stochastic version of Hamilton's principle in phase space. When $\mathcal{L}$ is a hyperregular Lagrangian we can pass to the Hamiltonian side as mentioned in the previous remark, and use the stochastic Hamilton-Jacobi formalism as proposed in [48]. However, a general treatment of stochastic Hamilton-Jacobi equations on the Pontryagin bundle, similar to the deterministic development in Leok, Ohsawa and Sosa [49], is left as a future prospect for research.

9.2.2 Stochastic Noether's Theorem: Variational Principle Version

Let $\Phi : Q \times \mathbb{R} \to Q$ be a deterministic smooth flow. Given $\epsilon \in \mathbb{R}$, let $\Phi_\epsilon : Q \to Q$ be the map $q \mapsto \Phi(q, \epsilon)$. Suppose $\mathcal{L}$ is invariant under the tangent lifted flow of Φ, L_i and V_i are invariant under Φ for all $i = 1, \cdots, k$ and $\epsilon \in \mathbb{R}$. For every $\epsilon \in \mathbb{R}$, define $\Psi_\epsilon : \mathcal{P}Q \to \mathcal{P}Q$ by $\Psi_\epsilon = T\Phi_\epsilon \oplus T^*\Phi_\epsilon^{-1}$. Then $\mathrm{pr}_Q \circ \Psi_\epsilon = \Phi_\epsilon$.

Let Γ be a $\mathcal{P}Q$-valued admissible semimartingale, written in coordinates as $\Gamma_t = (q_t, v_t, p_t)$ and set $\Gamma_{\epsilon,t} = \Psi_\epsilon(\Gamma_t) = (q_{\epsilon,t}, v_{\epsilon,t}, p_{\epsilon,t})$. We note that

$$\begin{aligned}\int \mathcal{G} \bullet d\Gamma_\epsilon &= \int \langle p_{\epsilon,t}, \bullet dq_{\epsilon,t} \rangle \\ &= \int \left\langle T^*\Phi_\epsilon \left(T^*\Phi_\epsilon^{-1}(p_t) \right), \bullet dq_t \right\rangle \\ &= \int \langle p_t, \bullet dq_t \rangle \\ &= \int \mathcal{G} \bullet d\Gamma\end{aligned}$$

and

$$\begin{aligned}E_j(\Gamma_{\epsilon,t}) = E_j(q_{\epsilon,t}, v_{\epsilon,t}, p_{\epsilon,t}) &= \begin{cases} \langle p_{\epsilon,t}, v_{\epsilon,t} \rangle - \mathcal{L}(q_{\epsilon,t}, v_{\epsilon,t}), & \text{if } j = 0, \\ \langle p_{\epsilon,t}, V_j(q_{\epsilon,t}) \rangle - L_j(q_{\epsilon,t}), & \text{if } j = 1, \ldots, k. \end{cases} \\ &= E_j(\Gamma_t).\end{aligned}$$

Consequently, $\mathcal{S}_X(\Gamma_\epsilon) = \mathcal{S}_X(\Gamma)$ and hence $\mathcal{D}\mathcal{S}_X(\Gamma) = 0$. Following the calculations done in the proof of Theorem 9.2.5, for $t \geq 0$ we have

$$\int_0^t \Bigg\langle \delta\Gamma_s^{|t}, \Bigg(\bullet d \int \Bigg(\frac{\partial}{\partial q_s^{|t}} \big(\mathcal{L} \bullet dX_s^0 + \\ \sum_{i=1}^k \big(L_i - \langle p_s^{|t}, V_i(q_s^{|t}) \rangle \big) \bullet dX_s^i \Bigg) - \bullet dp_s^{|t} \Bigg), \bullet d \int \left(p_s^{|t} - \frac{\partial \mathcal{L}}{\partial v_s^{|t}} \right) \bullet dX_s^0, \\ \bullet d \int \left(\bullet dq_s^{|t} - v_s^{|t} \bullet dX_s^0 - \sum_{i=1}^k V_i(q_s^{|t}) \bullet dX_s^i \right) \Bigg\rangle$$

$$+ \langle p_t, \delta q_t \rangle - \langle p_0, \delta q_0 \rangle = 0.$$

Suppose Γ solves the stochastic implicit Euler-Lagrange equations up to a maximal stopping time τ. Then for all $t \leq \tau$, $\langle p_t, \delta q_t \rangle = \langle p_0, \delta q_0 \rangle$. Hence $\langle p_t, \delta q_t \rangle$ is conserved along the solution Γ. Since $p_t = \frac{\partial \mathcal{L}}{\partial v_t}$, it follows that $\langle p_t, \delta q_t \rangle = \langle \theta_{\mathcal{L}}, \delta q \rangle$, where $\theta_{\mathcal{L}}$ is the pullback to TQ of the Liouville 1-form on T^*Q by $F\mathcal{L}$. Thus we have proven the following theorem:

Theorem 9.2.9. Suppose $\Phi : Q \times \mathbb{R} \to Q$ is a smooth deterministic flow such that $\mathcal{L}$ is invariant under the tangent lift of Φ, L_i is invariant under Φ and V_i is symmetric under Φ for all $i = 1, \cdots, k$. Let $\Gamma_t = (q_t, v_t, p_t)$ solves the stochastic implicit Euler-Lagrange equations and $q_{\epsilon,t} := \Phi_\epsilon(q_t) = \Phi_\epsilon(\mathrm{pr}_Q(\Gamma_t))$. Then $\langle \theta_{\mathcal{L}}, \delta q \rangle$ is conserved along Γ.

9.2.3 The Intrinsic Form of the Stochastic Hamilton-Pontryagin Principle

We will now develop the intrinsic form of the stochastic Hamilton-Pontryagin principle. We introduce an additional assumption, namely that our semimartingales are obtained as solutions of Stratonovich equations on manifolds. To motivate this, recall that in case of the deterministic Hamilton's principle, given a regular Lagrangian $\mathcal{L} \in C^\infty(TQ)$ there exists a second order vector field $Z_{\mathcal{L}}$ such that the integral curves of $Z_{\mathcal{L}}$ project to solutions of the Euler-Lagrange equations. Thus, for regular Lagrangians, finding a critical point of the action functional corresponds to selecting a particular vector field in $\mathfrak{X}^\infty(TQ)$. We will extend this idea to the stochastic case as well, namely, we will show that under the assumption that a semimartingale solves a Stratonovich equation, finding a critical point of $\mathcal{S}_X$ corresponds to selecting a particular Stratonovich operator.

Let $(e_0, \cdots, e_k)$ be a basis of $\mathbb{R}^{k+1}$. Let $\mathcal{M}$ be any regular submanifold of $\mathcal{P}Q$, $\mathrm{Pr}_{T\mathcal{M}}$ and $\mathrm{Pr}_{\mathcal{M}}$ denote the same projections as $\mathrm{Pr}_{T\mathcal{P}Q}$ and $\mathrm{Pr}_{\mathcal{P}Q}$ respectively with $\mathcal{P}Q$ replaced by $\mathcal{M}$ and we restrict the other maps in (9.2.2) to $\mathcal{M}$. The generalized energies are now defined on $\mathcal{M}$ as opposed to $\mathcal{P}Q$. Given any Stratonovich operator $S \in \mathrm{Strat}(\mathbb{R}^{k+1}, \mathcal{M})$, any semimartingale $X = (X^0, \cdots, X^k)$ in $\mathbb{R}^{k+1}$ and a solution Γ_X of $\bullet d\Gamma = S(X, \Gamma) \bullet dX$, from Eq. (9.2.4) we have

$$\mathcal{S}_X(\Gamma_X) = \int_0^T \mathcal{G} \bullet d\Gamma_X - \sum_{j=0}^{k} \int_0^T E_j(\Gamma_X) \bullet dX^j.$$

The dual of $S(x, y)$ will be denoted by $S^\vee(x, y)$ in this section to distinguish it from the pullback. Moreover, we adopt the following notation: Suppose $x \in \mathbb{R}^{k+1}, v \in T_x\mathbb{R}^{k+1} \cong \mathbb{R}^{k+1}$, and $S \in \mathrm{Strat}(\mathbb{R}^{k+1}, M)$. We denote the vector field $y \in \mathcal{M} \mapsto S(x, y)(v) \in T_y\mathcal{M}$ by $S^{x,v}$.

Given $x \in \mathbb{R}^{k+1}$, $y \in \mathcal{M}$ and $j \in \{0, \cdots, k\}$, suppose $z_j \in TT\mathcal{M}$, $\mathrm{Pr}_{T\mathcal{M}} = S^{x,e_j}(y)$ and $T\mathrm{Pr}_{\mathcal{M}} = w_y$, for some $w_y \in T\mathcal{M}$. Then

$$dE_j(w_y) = dE_j \circ T\mathrm{Pr}_{\mathcal{M}}(z_j) = (\mathrm{Pr}_{\mathcal{M}})^* dE_j(S^{x,e_j})(z_j).$$

Now let $\epsilon \mapsto \Gamma_{X_\epsilon}$ be any deformation of Γ_X and Z^j be any $TT\mathcal{M}$-valued semimartingale such that $T\mathrm{Pr}_{\mathcal{M}}(Z^j) = \delta\Gamma_X$ and $\mathrm{Pr}_{T\mathcal{M}}(Z^j) = S^{x,e_j}(\Gamma_X)$. Then

$$dE_j(\delta\Gamma_X) = (\mathrm{Pr}_{\mathcal{M}})^* dE_j(S^{X,e_j}(\Gamma_X))(Z^j)$$

and hence

$$\mathcal{D}\left(\sum_{j=0}^{k}\int_0^T E_j(\Gamma_X)\bullet dX^j\right) = \sum_{j=0}^{k}\int_0^T (\mathrm{Pr}_{\mathcal{M}})^* dE_j(S^{X,e_j}(\Gamma_X))(Z^j)\bullet dX^j. \tag{9.2.7}$$

Next, by Lemma 9.1.9

$$\begin{aligned}\mathcal{D}\int_0^T \mathcal{G}\bullet d\Gamma_X &= \int_0^T i_{\delta\Gamma_X}\mathbf{d}\mathcal{G}\bullet d\Gamma_X + \langle \mathcal{G}(\Gamma_{X_T}), \delta\Gamma_{X_T}\rangle - \langle \mathcal{G}(\Gamma_{X_0}), \delta\Gamma_{X_0}\rangle \\ &= \int_0^T S^\vee(X,\Gamma_X) i_{\delta\Gamma_X}\mathbf{d}\mathcal{G}\bullet dX + \langle \mathcal{G}(\Gamma_{X_T}), \delta\Gamma_{X_T}\rangle - \langle \mathcal{G}(\Gamma_{X_0}), \delta\Gamma_{X_0}\rangle .\end{aligned}$$

The calculation of $\int_0^T S^\vee(X,\Gamma_X) i_{\delta\Gamma_X}\mathbf{d}\mathcal{G}\bullet dX$ is related to the proof of Proposition 3.2 in Yoshimura and Marsden [94]. Let θ_{T^*Q} denote the Liouville 1-form on T^*Q and $\Omega_{T^*Q} = -d\theta_{T^*Q}$. Denote by

$$\Omega^\flat_{T^*Q} : TT^*Q \to T^*T^*Q$$

the bundle map associated with Ω_{T^*Q}. Also let $\theta_{T^*T^*Q}$ be the Liouville 1-form on T^*T^*Q and set $\chi = (\Omega^\flat)^*\theta_{T^*T^*Q}$. Note that χ is a 1-form on TT^*Q. We will show that, given any $x \in \mathbb{R}^{k+1}$, $y = (q,v,p) \in \mathcal{M}$, $w_y = (q,v,p,w_q,w_v,w_p) \in T\mathcal{M}$ and $z_j \in TT\mathcal{M}$ such that $T\mathrm{Pr}_{\mathcal{M}}(z_j) = w_y$ and $\mathrm{Pr}_{T\mathcal{M}}(z_j) = S^{x,e_j}(y)$, we have

$$\left\langle S^\vee(x,y) i_{w_y}\mathbf{d}\mathcal{G}, e_j\right\rangle = (T\mathrm{pr}_{T^*Q})^*\chi(S^{x,e_j}(y))(z_j). \tag{9.2.8}$$

Let

$$\begin{aligned} u_y &= S^{x,e_j}(y) = (q,v,p,u_q,u_v,u_p) \in T\mathcal{M}, \\ z_j &= (q,v,p,u_q,u_v,u_p,w_q,w_v,w_p,\tilde{w}_q,\tilde{w}_v,\tilde{w}_p).\end{aligned}$$

From the proof of Lemma 9.2.2 we get

$$\left\langle S^\vee(x,y) i_{w_y}\mathbf{d}\mathcal{G}, e_j\right\rangle = \mathbf{d}\mathcal{G}(w_y,u_y) = \langle w_p, u_q\rangle - \langle w_q, u_p\rangle .$$

On the other hand

$$\begin{aligned}\theta_{T^*T^*Q}(\Omega^\flat_{T^*Q}\circ T\mathrm{pr}_{T^*Q}(q,v,p,u_q,u_v,u_p)) &= \theta_{T^*T^*Q}(\Omega^\flat_{T^*Q}(q,p,u_q,u_p)) \\ &= \theta_{T^*T^*Q}(q,p,-u_p,u_q) \\ &= -u_p dq + u_q dp.\end{aligned}$$

Since $T_{u_y}T\mathrm{pr}_{T^*Q}(z_j) = (q,p,u_q,u_p,w_q,w_p,\tilde{w}_q,\tilde{w}_p)$, it follows that

$$T\Omega^\flat_{T^*Q}(T_{u_y}T\mathrm{pr}_{T^*Q})(z_j) = (q,p,-u_p,u_q,w_q,w_p,-\tilde{w}_p,\tilde{w}_q)$$

Note that

$$(T\mathrm{pr}_{T^*Q})^*\chi(S^{x,e_j}(y))(z_j) = \theta_{T^*T^*Q}(\Omega^\flat_{T^*Q} \circ T\mathrm{pr}_{T^*Q}(u_y)) \cdot T\Omega^\flat_{T^*Q}(T_{u_y}T\mathrm{pr}_{T^*Q}(z_j)),$$

and

$$\begin{aligned}\theta_{T^*T^*Q}(\Omega^\flat_{T^*Q} \circ T\mathrm{pr}_{T^*Q}(u_y)) \cdot T\Omega^\flat_{T^*Q}(T_{u_y}T\mathrm{pr}_{T^*Q}(z_j)) &= -\langle w_q, u_p\rangle + \langle w_p, u_q\rangle \\ &= \left\langle S^\vee(x,y) i_{w_y}\mathbf{d}\mathcal{G}, e_j \right\rangle.\end{aligned}$$

This proves our claim.

Consequently, given any $a = (a^0, \cdots, a^k) \in \mathbb{R}^{k+1}$

$$\begin{aligned}\langle S^\vee(x,y) i_{w_y}\mathbf{d}\mathcal{G}, a\rangle &= \sum_{j=0}^k a^j \langle S^\vee(x,y) i_{w_y}\mathbf{d}\mathcal{G}, e_j\rangle \\ &= \sum_{j=0}^k a^j (T\mathrm{pr}_{T^*Q})^*\chi(S^{x,e_j}(y))(z_j)\end{aligned}$$

where $z_j \in TT\mathcal{M}$, $\mathrm{Pr}_{T\mathcal{M}}(z_j) = S^{x,e_j}(y)$ and $T\mathrm{Pr}_{\mathcal{M}}(z_j) = w_y$.

Therefore, given any deformation $\epsilon \mapsto \Gamma_{X_\epsilon}$ and semimartingales $Z^0, \cdots, Z^k$ in $TT\mathcal{M}$ over $S^{X,e_j}(\Gamma_X)$ such that $T\mathrm{Pr}_{\mathcal{M}}(Z^j) = \delta\Gamma_X$ we have

$$\begin{aligned}\mathcal{D}\int_0^T \mathcal{G} \bullet d\Gamma_X &= \int_0^T S^\vee(X,\Gamma_X) i_{\delta\Gamma_X}\mathbf{d}\mathcal{G} \bullet dX + \langle \mathcal{G}(\Gamma_{X_T}), \delta\Gamma_{X_T}\rangle - \langle \mathcal{G}(\Gamma_{X_0}), \delta\Gamma_{X_0}\rangle \\ &= \sum_{j=0}^k \int_0^T (T\mathrm{pr}_{T^*Q})^*\chi(S^{X,e_j}(\Gamma_X))(Z^j) \bullet dX^j \\ &\quad + \langle \mathcal{G}(\Gamma_{X_T}), \delta\Gamma_{X_T}\rangle - \langle \mathcal{G}(\Gamma_{X_0}), \delta\Gamma_{X_0}\rangle.\end{aligned}$$

We have $\langle \mathcal{G}(\Gamma_t), \delta\Gamma_t\rangle = \langle p_t, \delta q_t\rangle$, where $\Gamma_t = (q_t, v_t, p_t)$ in coordinates. Assuming that $T\mathrm{pr}_Q(\delta\Gamma) = 0$ at $t = 0, T$, we have

$$\langle \mathcal{G}(\Gamma_T), \delta\Gamma_T\rangle = 0 = \langle \mathcal{G}(\Gamma_0), \delta\Gamma_0\rangle.$$

In that case

$$\mathcal{D}\int_0^T \mathcal{G} \bullet d\Gamma_X = \sum_{j=0}^k \int_0^T (T\mathrm{pr}_{T^*Q})^*\chi(S^{X,e_j}(\Gamma_X))(Z^j) \bullet dX^j.$$

We summarize this discussion in the following lemma:

Lemma 9.2.10. Let $S \in \mathrm{Strat}(\mathbb{R}^{k+1}, \mathcal{M})$ where $\mathcal{M}$ is a regular submanifold of $\mathcal{P}Q$. For every semimartingale $X = (X^0, \cdots, X^k) \in \mathscr{S}\left(\mathbb{R}^{k+1}\right)$, if Γ_X solves $\bullet d\Gamma = S(X,\Gamma) \bullet dX$ and Γ_X is admissible, then

$$\mathcal{D}\mathcal{S}_X(\Gamma_X) = \mathcal{D}\left[\int_0^T \mathcal{G} \bullet d\Gamma_X - \sum_{j=0}^k \int_0^T E_j(\Gamma_X) \bullet dX^j\right] = 0$$

for all variations $\epsilon \mapsto \Gamma_{X_\epsilon}$ with $T\mathrm{pr}_Q(\delta\Gamma_{X_0}) = T\mathrm{pr}_Q(\delta\Gamma_{X_T}) = 0$, if and only if

$$\sum_{j=0}^{k} \int_0^T \left[(T\mathrm{pr}_{T^*Q})^*\chi(S^{X,e_j}(\Gamma_X)) - (\mathrm{Pr}_{\mathcal{M}})^* dE_j(S^{X,e_j}(\Gamma_X))\right](Z^j) \bullet dX^j = 0 \tag{9.2.9}$$

for arbitrary $TT\mathcal{M}$-valued semimartingales $Z^0, \cdots, Z^k$ over $S^{X,e_j}(\Gamma_X)$ such that

$$T\left(\mathrm{pr}_Q \circ \mathrm{Pr}_{\mathcal{M}}\right)(Z^j) = 0$$

at $t = 0$ and $t = T$.

Definition 9.2.11. Given $y_0 \in \mathcal{M}$ and $j = 0, \cdots, k$ let $y_j(t)$ be a curve in $\mathcal{M}$ satisfying

$$(T\mathrm{pr}_{T^*Q})^*\chi(y_j(t), \dot{y}_j(t)) = (\mathrm{Pr}_{\mathcal{M}})^* dE_j(y_j(t), \dot{y}_j(t)) \tag{9.2.10}$$

with $y_j(0) = y_0$. Let $S_{HP} \in \mathrm{Strat}(\mathbb{R}^{k+1}, \mathcal{M})$ be defined by

$$S_{HP}(x_0, y_0)(a) = \sum_{j=0}^{k} a^j \dot{y}_j(0) \in T_{y_0}\mathcal{M}$$

for every $x_0 \in \mathbb{R}^{k+1}$ and $a = (a^0, \cdots, a^k) \in \mathbb{R}^{k+1} \cong T_{x_0}\mathbb{R}^{k+1}$. We will call S_{HP} a **Hamilton-Pontryagin Stratonovich operator**.

Remark 9.2.12. From Yoshimura and Marsden [94], if $j = 0$ then Eq. (9.2.10) is just the deterministic implicit Euler-Lagrange equations for $\mathcal{L}$. Since this means that the Legendre transform holds, it follows that S_{HP}^{x,e_0}, and hence S_{HP}, is well-defined if and only if $\mathcal{M}$ is the submanifold $\mathcal{K} = TQ \oplus F\mathcal{L}(TQ)$. For $j = 1, \cdots, k$, in local coordinates Eq. (9.2.10) reads

$$\dot{q} = V_j(q), \quad \dot{p} = \frac{\partial}{\partial q}\left(L_j(q) - \langle p, V_j(q)\rangle\right). \tag{9.2.11}$$

We assume enough regularity on $\mathcal{L}$ such that $\mathcal{K}$ is a well-defined submanifold. In particular, assuming that $F\mathcal{L}$ has constant rank suffices.

By definition of S_{HP}, if $\mathcal{M} = \mathcal{K}$ and $S = S_{HP}$ then Eq. (9.2.9) is satisfied, so $\mathcal{DS}_X(\Gamma_X) = 0$. We now prove the converse.

First we make an important observation. The Stratonovich operator S is a deterministic object that is defined independently of the semimartingale X. Since the equivalence in Lemma 9.2.10 is true for every semimartingale $X \in \mathscr{S}\left(\mathbb{R}^{k+1}\right)$ and a solution Γ_X of $\bullet d\Gamma = S(X, \Gamma) \bullet dX$ that is admissible, it must be true for a deterministic semimartingale of the form $X_t = X_0(t) := x_0 + te_0$, where $x_0 \in \mathbb{R}^{k+1}$ is arbitrary. In this case the solution of $\bullet d\Gamma = S(X, \Gamma) \bullet dX = S^{X_0(t),e_0}(\Gamma)dt$ is a deterministic smooth curve in $\mathcal{M}$ that we denote by $\gamma_0(t)$. Given $y_0 \in \mathcal{M}$, suppose that $\gamma_0(0) = y_0$. Note that $\gamma_0(t)$ solves

$$\dot{\gamma}_0(t) = S^{X_0(t),e_0}(\gamma_0(t)) = S(x_0 + te_0, \gamma_0(t))(e_0).$$

By Lemma 9.2.10 we have

$$\int_0^T \left[(T\mathrm{pr}_{T^*Q})^* \chi(S^{x_0+te_0,e_0}(\gamma_0(t))) - (\mathrm{Pr}_{\mathcal{M}})^* dE_0(S^{x_0+te_0,e_0}(\gamma_j(t))) \right] (z^0(t))dt = 0$$

for all smooth curves $z_0(t)$ in $TT\mathcal{M}$ such that $\mathrm{Pr}_{T\mathcal{M}}(z_0(t)) = S^{x_0+te_0,e_0}(\gamma_0(t)) = \dot{\gamma}_0(t)$ and $T\left(\mathrm{pr}_Q \circ \mathrm{Pr}_{\mathcal{M}}\right)(z_0(t)) = 0$ at $t = 0$ and $t = T$. By the (deterministic) fundamental theorem of the calculus of variations we have

$$(T\mathrm{pr}_{T^*Q})^* \chi(S^{x_0+te_0,e_0}(\gamma_0(t))) = (\mathrm{Pr}_{\mathcal{M}})^* dE_0(S^{x_0+te_0,e_0}(\gamma_0(t))).$$

Since $\dot{\gamma}_0(t) = S^{X_0(t),e_0}(\gamma_0(t))$ we have

$$(T\mathrm{pr}_{T^*Q})^* \chi(\gamma_0(t), \dot{\gamma}_0(t)) = (\mathrm{Pr}_{\mathcal{M}})^* dE_0(\gamma_0(t), \dot{\gamma}_0(t)).$$

This equation has a solution provided $\mathcal{M} = \mathcal{K}$. Moreover, by definition of S_{HP}

$$S_{HP}(x_0, y_0)(e_0) = \dot{\gamma}_0(0) = S(x_0, y_0)(e_0).$$

A similar argument with e_0 replaced by e_j shows that $S(x_0, y_0)(e_j) = S_{HP}(x_0, y_0)(e_j)$ for all $j = 0, \cdots, k$. Thus $S = S_{HP}$.

Theorem 9.2.13. Let $S \in \mathrm{Strat}(\mathbb{R}^{k+1}, \mathcal{M})$ where $\mathcal{M}$ is a regular submanifold of $\mathcal{P}Q$. For every semimartingale $X = (X^0, \cdots, X^k) \in \mathscr{S}\left(\mathbb{R}^{k+1}\right)$, if Γ_X solves $\bullet d\Gamma = S(X, \Gamma) \bullet dX$ and Γ_X is admissible then $\mathcal{DS}_X(\Gamma_X) = 0$ for all deformations $\epsilon \mapsto \Gamma_{X_\epsilon}$ satisfying $T\mathrm{pr}_Q(\delta\Gamma_{X_0}) = T\mathrm{pr}_Q(\delta\Gamma_{X_T}) = 0$ if and only if $\mathcal{M} = \mathcal{K}$ and $S = S_{HP}$.

Moreover, the stochastic implicit Euler-Lagrange equations are given by

$$\bullet d(\mathrm{pr}_{T^*Q}(\Gamma^{|T})) = T\mathrm{pr}_{T^*Q}\left(S_{HP}(X, \Gamma^{|T})\right) \bullet dX$$

on the submanifold $\mathcal{K}$.

Proof. It only remains to show that the stochastic implicit Euler-Lagrange equations are given by

$$\bullet d(\mathrm{pr}_{T^*Q}(\Gamma^{|T})) = T\mathrm{pr}_{T^*Q}\left(S_{HP}(X, \Gamma^{|T})\right) \bullet dX$$

on $\mathcal{K}$. Let $\Gamma_t = (q_t, v_t, p_t)$ in terms of local coordinates on $\mathcal{P}Q$. The restriction to $\mathcal{K}$ implies that $\left(p_t^{|T} - \frac{\partial \mathcal{L}}{\partial v_t^{|T}}\right) \bullet dX_t^0 = 0$. Given any $x \in \mathbb{R}^{k+1}$ and $a = (a^0, \cdots, a^k) \in \mathbb{R}^{k+1} \cong T_x\mathbb{R}^{k+1}$, suppose that in local coordinates, we have

$$S_{HP}(x, y)(a) = \sum_{j=0}^{k} a^j S_{HP}^{x,e_j}(q, v, p) = \sum_{j=0}^{k} a^j (u_{q_j}, u_{v_j}, u_{p_j}).$$

Using the local form of the deterministic implicit Euler-Lagrange equations for $j = 0$ and using Eq. (9.2.11) for $j = 1, \cdots, k$, we have

$$T\mathrm{pr}_{T^*Q}\left(S_{HP}(x, y)(a)\right) = \sum_{j=0}^{k} a^j T\mathrm{pr}_{T^*Q}(u_{q_j}, u_{v_j}, u_{p_j})$$

$$
\begin{aligned}
&= \sum_{j=0}^{k} a^j (u_{q_j}, u_{p_j}) \\
&= a^0 \left(v, \frac{\partial \mathcal{L}}{\partial q} \right) + \sum_{j=1}^{k} a^j \left(V_j(q), \frac{\partial}{\partial q} \left(L_j(q) - \langle p, V_j(q) \rangle \right) \right).
\end{aligned}
$$

Thus the local form of the equation $\bullet d(\mathrm{pr}_{T^*Q}(\Gamma^{|T})) = T\mathrm{pr}_{T^*Q}\left(S_{HP}(X, \Gamma^{|T})\right) \bullet dX$ is

$$
\begin{aligned}
\bullet dq_t^{|T} &= v_t^{|T} \bullet dX_t^0 + \sum_{i=1}^{k} V_i(q_t^{|T}) \bullet dX_t^i \\
\bullet dp_t^{|T} &= \frac{\partial}{\partial q_t^{|T}} \left(\mathcal{L} \bullet dX_t^0 + \sum_{i=1}^{k} \left(L_i(q^{|T}) - \left\langle p_t^{|T}, V_i(q_t^{|T}) \right\rangle \right) \bullet dX_t^i \right),
\end{aligned}
$$

which, together with the equation $\left(p_t^{|T} - \frac{\partial \mathcal{L}}{\partial v_t^{|T}} \right) \bullet dX_t^0 = 0$ gives the stochastic implicit Euler-Lagrange equations. ∎

Remark 9.2.14. This method of reformulating the problem of determining a critical point of a stochastic action to determining a Stratonovich operator can be applied to other stochastic action principles as well, for instance, the stochastic Hamilton's principle in phase space described in Lázaro-Camí and Ortega [55].

Chapter 10

Stochastic Lagrange-Poincaré Reduction

In this chapter we study symmetries of the stochastic Hamilton-Pontryagin principle, and extend the general framework of implicit Lagrange-Poincaré reduction carried out by Yoshimura and Marsden [95] and outlined in Section 2.5 to the stochastic case. We also generalize stochastic Euler-Poincaré reduction carried out by Bou-Rabee and Owhadi [14], Arnaudon, de Castro and Holm [3], and Street and Takao [89] to the case where the configuration manifold is an arbitrary smooth manifold equipped with the free and proper action of a Lie group.

To give a motivating finite-dimensional example, consider the problem of a satellite with a rotor attached to it. Suppose that the body angular momentum of the satellite and the angular momentum of the rotor can be observed, but that these observations are inherently noisy due to sensor error or environmental disturbances. These angular momenta are naturally expressed as variables on a reduced space obtained by exploiting the rotational symmetry of the system. We can then ask if these noisy equations in the reduced space can be obtained via reducing a stochastic perturbation of the entire system. This also leads to the problem of understanding when stochastic forcing preserves the symmetries underlying the reduction and when it breaks them.

In the classical setting, let G be a Lie group acting freely and properly on a configuration manifold Q and via tangent and cotangent lifts on TQ and T^*Q respectively. From Section 2.5, we recall that in implicit Lagrange-Poincaré reduction, the reduced space is the Whitney sum of the Pontryagin bundle of the shape space, Q/G, the associated adjoint bundle obtained by fixing a connection on $Q \to Q/G$ and the dual of the associated adjoint bundle. Corresponding to a G-invariant Lagrangian function, the reduced equations may be divided into a set of horizontal equations, which resemble Euler-Lagrange equations with an extra force arising from a reduced curvature 2-form, and vertical equations, which resemble Euler-Poincaré equations.

We show that a similar behaviour occurs in case of the stochastic Hamilton-Pontryagin principle. We assume that $\mathcal{L} \in C^\infty(TQ)$, $L_i \in C^\infty(Q)$ and $V_i \in \mathfrak{X}^\infty(Q)$ are all G-invariant for all $i = 1, \cdots, k$. Then we obtain a reduced stochastic action integral on the reduced space. Similar to the deterministic case, under certain constrained variations we obtain reduced horizontal implicit stochastic Lagrange-Poincaré equations and vertical implicit stochastic Lagrange-Poincaré equations. The vertical equations resemble stochastic Euler-Poincaré equations. When Q is a Lie group these equations are the stochastic Euler-Poincaré equations.

In order to express the reduced equations in a coordinate-free way, we need the notion of stochastic covariant derivatives. This is described by Norris [75] and Catuogno, Ledesma and Ruffino [16] and reviewed in Section 10.1. In the next section, we describe the reduction process and state a stochastic analogue of the implicit Lagrange-Poincaré reduction theorem. In Section 10.3, we discuss two examples, namely, the stochastic rigid body with a rotor, and a stochastic charged particle in a magnetic field.

For the convenience of the reader, below we recall some of the notations introduced in Section 2.5 in the context of deterministic implicit Lagrange-Poincaré reduction.

Notation	**Description**
$\phi : G \times Q \to Q$	Free and proper action of a Lie group G on Q. G acts on TQ via tangent lifts and on T^*Q via cotangent lifts.
gq, gv, gp	The images of $q \in Q$, $v \in TQ$, and $p \in T^*Q$ under the action of an element $g \in G$.
A, B	$A : TQ \to \mathfrak{g}$ is a principal connection on $\pi : Q \to Q/G$, and B is its curvature 2–form.
$[q]$	$\pi(q)$, where $q \in Q$.
$\tilde{\mathfrak{g}}, \tilde{\mathfrak{g}}^*$	$\tilde{\mathfrak{g}} = (Q \times \mathfrak{g})/G \xrightarrow{\pi_{\tilde{\mathfrak{g}}}} Q/G$, $\tilde{\mathfrak{g}}^* = ((Q \times \mathfrak{g})/G)^* \xrightarrow{\pi_{\tilde{\mathfrak{g}}^*}} Q/G$, the adjoint bundle and its dual bundle.
ξ	$\xi = A(q, \dot{q})$, where $q(t)$ is a curve in Q.
$\bar{\zeta} = [q, \zeta]_G,\ \bar{\mu} = [q, \mu]_G$	Elements of $\tilde{\mathfrak{g}}$ and $\tilde{\mathfrak{g}}^*$ respectively, where $q \in Q$, $\zeta \in \mathfrak{g}$, and $\mu \in \mathfrak{g}^*$.
$\nabla^{\tilde{\mathfrak{g}}},\ \nabla^{\tilde{\mathfrak{g}}^*}$	Affine connections on $\tilde{\mathfrak{g}}$ and $\tilde{\mathfrak{g}}^*$ respectively.
$\Psi_A,\ (\Psi_A^{-1})^*,\ \tilde{\Psi}_A$	Connection–induced bundle isomorphisms $$\begin{aligned} \Psi_A &: TQ/G \to T(Q/G) \oplus \tilde{\mathfrak{g}}, \\ (\Psi_A^{-1})^* &: T^*Q/G \to T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*, \\ \tilde{\Psi}_A &: \mathcal{P}Q/G \to \mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*. \end{aligned}$$
$\tilde{B}$	Reduced curvature 2–form on Q/G induced by B.

Table 10.1: List of notations from Section 2.5.

10.1 Connectors and the Stochastic Covariant Derivative

In the stochastic case, the covariant derivatives in the vertical and horizontal Lagrange-Poincaré equations are replaced by stochastic covariant derivatives. On a vector bundle with a connection, this is defined in terms of the connector or the connection map.

10.1.1 The Connector on a Vector Bundle

Let $E \xrightarrow{\pi_E} M$ be a vector bundle with a connection ∇ and let $GL(E)$ denote the frame bundle of E. From the connection ∇, we obtain a projection $\Phi_E : TE \to \operatorname{Ver} E$ to the vertical bundle $\operatorname{Ver} E$ of TE. Let $\operatorname{vlft}^E : E \times_M E \to \operatorname{Ver} E$ denote the vertical lift map given by

$$\operatorname{vlft}^E_m(v_{m_1}, v_{m_2}) = \frac{d}{ds}\Big|_{s=0}(v_{m_1} + s v_{m_2}), \tag{10.1.1}$$

for all $m \in M$, $v_{m_1}, v_{m_2} \in \pi_E^{-1}(m)$. For the proof of the next lemma, we refer the reader to Michor [68].

Lemma 10.1.1. The vertical lift map vlft^E satisfies the following properties:

1. vlft^E is a vector bundle isomorphism.
2. Let Pr_i denote the projection onto the i-th factor of $E \times_M E$. Then $\operatorname{Pr}_1 \circ (\operatorname{vlft}^E)^{-1} = \operatorname{pr}_{TE}|_{\operatorname{Ver} E}$, where $\operatorname{pr}_{TE} : TE \to E$ is the projection onto the basepoint.

Definition 10.1.2. The map $\operatorname{vpr}_E := \operatorname{Pr}_2 \circ (\operatorname{vlft}^E)^{-1} : \operatorname{Ver} E \to E$ is called the **vertical projection**. The map $K_\nabla := \operatorname{vpr}_E \circ \Phi_E : TE \to E$ is called the **connector** associated to ∇.

Let $F : E \to \mathbb{R}$ and $v_0 \in E$ be an arbitrary point with $\pi_E(v_0) = m_0$. Let $u_{m_0} \in T_{m_0}M$. We define

$$\left\langle \frac{\partial F}{\partial m}\Big|_{v_0}, (m_0, u_{m_0}) \right\rangle = \frac{d}{dt}\Big|_{t=0} F(m^h(t)) \tag{10.1.2}$$

where $m^h(t)$ is the horizontal lift of a curve $m(t)$ in M, with $m(0) = m_0$, $\dot{m}(0) = u_{m_0}$ and $m^h(0) = v_0$.

Remark 10.1.3. This is independent of the choice of the curve $m(t)$ in M, since the tangent vector $\dot{m}^h(0)$ depends only on the horizontal lift of u_{m_0}.

We also define

$$\left\langle \frac{\partial F}{\partial v}\Big|_{v_0}, v \right\rangle = \frac{d}{ds}\Big|_{s=0} F(v_0 + sv) = \left\langle dF, \operatorname{vlft}^E_{m_0}(v_0, v) \right\rangle, \tag{10.1.3}$$

for all $v \in \pi_E^{-1}(m_0)$. Note that the definition of $\frac{\partial F}{\partial v}$ is independent of ∇.

Theorem 10.1.4. Suppose $\epsilon \mapsto v_{m,\epsilon}(t)$ is a deformation of $v_m(t)$. Then

$$\langle dF, \delta v_m(t)\rangle = \left\langle \frac{\partial F}{\partial m}, \delta m(t)\right\rangle + \left\langle \frac{\partial F}{\partial v}, K_\nabla(\delta v_m(t))\right\rangle . \tag{10.1.4}$$

Proof. We have

$$\langle dF, \delta v_m(t)\rangle = \langle dF|_{\mathrm{Hor}\,E}, \mathrm{Hor}\,\delta v_m(t)\rangle + \langle dF|_{\mathrm{Ver}\,E}, \mathrm{Ver}\,\delta v_m(t)\rangle$$

where $\mathrm{Hor}\,\delta v_m(t)$ and $\mathrm{Ver}\,\delta v_m(t)$ are the horizontal and vertical components of $\delta v_m(t)$ respectively. For the first term on the right, horizontally lift $m_\epsilon(t) := \pi_E(v_{m,\epsilon}(t))$ to a curve $m^h_\epsilon(t)$ starting at $v_{m,\epsilon}(0)$. Then

$$\langle dF|_{\mathrm{Hor}\,E}, \mathrm{Hor}\,\delta v_m(t)\rangle = \left\langle dF, \delta m^h(t)\right\rangle = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} F(m^h_\epsilon(t)) = \left\langle \frac{\partial F}{\partial m}, \delta m(t)\right\rangle .$$

For the second term on the right, let $w(t) = (\mathrm{vlft}^E)^{-1} \circ \Phi_E(\delta v_m(t))$. Then, by Lemma 10.1.1 $\mathrm{Pr}_1(w(t)) = v_m(t)$, and by definition of K_∇, $\mathrm{Pr}_2(w(t)) = K_\nabla(\delta v_m(t))$. Therefore, we have

$$\begin{aligned}
\langle dF|_{\mathrm{Ver}\,E}, \mathrm{Ver}\,\delta v_m(t)\rangle &= \langle dF|_{\mathrm{Ver}\,E}, \Phi_E(\delta v_m(t))\rangle \\
&= \left\langle dF|_{\mathrm{Ver}\,E}, (\mathrm{vlft}^E)(\mathrm{Pr}_1(w(t)), \mathrm{Pr}_2(w(t)))\right\rangle \\
&= \left\langle dF|_{\mathrm{Ver}\,E}, (\mathrm{vlft}^E)(v_m(t), K_\nabla(\delta v_m(t)))\right\rangle \\
&= \left\langle dF|_{\mathrm{Ver}\,E}, \frac{d}{ds}\bigg|_{s=0} (v_m(t) + K_\nabla(\delta v_m(t)))\right\rangle \\
&= \frac{d}{ds}\bigg|_{s=0} F(v_m(t) + K_\nabla(\delta v_m(t))) \\
&= \left\langle \frac{\partial F}{\partial v}, K_\nabla(\delta v_m(t))\right\rangle .
\end{aligned}$$

This completes the proof. ∎

10.1.2 The Stochastic Covariant Derivative

Let $E \xrightarrow{\pi_E} M$ be a vector bundle with a connection ∇ and let E^* denote the dual bundle. Let K_∇ denote the associated connector. Suppose Γ is a semimartingale in E. We adopt Proposition 1 in Catuogno, Ledesma, and Ruffino [16] as the definition of the (Stratonovich) stochastic covariant derivative of Γ:

Definition 10.1.5. For every section $\alpha : M \to E^*$ of the dual bundle E^*, we define the **stochastic covariant derivative** $\int \alpha \bullet D\Gamma$ by

$$\int \langle \alpha, \bullet D\Gamma\rangle := \int \alpha(\pi_E(\Gamma)) \circ K_\nabla \bullet d\Gamma. \tag{10.1.5}$$

We denote Equation (10.1.5) in differential notation by

$$\bullet D\Gamma := K_\nabla \bullet d\Gamma.$$

Either by purely deterministic arguments or as a special case of this definition, we obtain

$$\frac{Dv(t)}{Dt} = K_\nabla(\dot{v}(t)). \tag{10.1.6}$$

It follows from the expression of the covariant derivative on the associated bundle in Equation (2.5.2) that given a semimartingale $[q_t, \bar{\zeta}_t]$ in $\tilde{\mathfrak{g}}$, we have

$$\bullet D[q_t, \bar{\zeta}_t]_G = [q_t, [\zeta_t, A(q_t, \bullet dq_t)] + \bullet d\zeta_t]_G. \tag{10.1.7}$$

Note that the connection ∇ also defines a connection on the dual bundle. This connection can be specified by specifying the covariant derivative corresponding to it. Let $\alpha(t)$ be a curve in E^* and $v(t)$ be a curve in E such that both $\alpha(t)$ and $v(t)$ project onto the same curve on M. Then, we define the covariant derivative of $\alpha(t)$ via the product rule

$$\frac{d}{dt}\langle \alpha(t), v(t)\rangle = \left\langle \frac{D\alpha(t)}{dt}, v(t)\right\rangle + \left\langle \alpha(t), \frac{Dv(t)}{Dt}\right\rangle \tag{10.1.8}$$

Then, if $K_{\nabla_{\tilde{\mathfrak{g}}}}$ denotes the connector on $\tilde{\mathfrak{g}}$ and we have a connection $\nabla_{\tilde{\mathfrak{g}}^*}$ on the dual bundle $\tilde{\mathfrak{g}}^*$ defined by Equation (10.1.8) then

$$\frac{d}{dt}\left\langle \bar{\mu}(t), \bar{\zeta}(t)\right\rangle = \left\langle \bar{\mu}(t), K_{\nabla^{\tilde{\mathfrak{g}}}}\dot{\bar{\zeta}}(t)\right\rangle + \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}}\dot{\bar{\mu}}(t), \bar{\zeta}(t)\right\rangle, \tag{10.1.9}$$

where $\bar{\zeta}(t)$ and $\bar{\mu}(t)$ are curves in $\tilde{\mathfrak{g}}$ and $\tilde{\mathfrak{g}}^*$ respectively that project to the same curve in Q/G. For semimartingales $\bar{\zeta} \in \mathscr{S}(\tilde{\mathfrak{g}})$ and $\bar{\mu} \in \mathscr{S}(\tilde{\mathfrak{g}}^*)$ that have the same projection to Q/G, we define $\int \left\langle \bullet D\bar{\mu}, \bar{\zeta}\right\rangle$ by

$$\left\langle \bar{\mu}, \bar{\zeta}\right\rangle = \int \left\langle \bullet D\bar{\mu}, \bar{\zeta}\right\rangle + \int \left\langle \bar{\mu}, \bullet D\bar{\zeta}\right\rangle. \tag{10.1.10}$$

10.2 Stochastic Implicit Lagrange-Poincaré Reduction

10.2.1 The Reduced Action Functional

We now assume that $\mathcal{L} \in C^\infty(TQ)$, $L_i \in C^\infty(Q)$ and $V_i \in \mathfrak{X}^\infty(Q)$ are G-invariant. Let $\ell \in C^\infty(T(Q/G) \oplus \tilde{\mathfrak{g}})$ denote the reduced Lagrangian corresponding to $\mathcal{L}$. By G-invariance of L_i, there exists $l_i \in C^\infty(Q/G)$ such that $l_i \circ \pi = L_i$. Similarly, G-invariance of V_i implies that there exists sections $\beta_i : Q/G \to \tilde{\mathfrak{g}}$ and vector fields $\Theta_i \in \mathfrak{X}^\infty(Q/G)$ such that $\beta_i([q]) = [q, A(V_i(q))]_G$ and $\Theta_i \circ \pi = T\pi \circ V_i$. Note that G-invariance of V_i and equivariance of A implies that β_i is well-defined. Similar to the deterministic reduced action functional $\mathcal{S}^{red}$ in Equation (2.5.9), we shall describe a reduced stochastic action $\mathcal{S}_X^{red}$ corresponding to the stochastic Hamilton-Pontryagin action $\mathcal{S}_X$ in Equation (9.2.1).

Let us briefly recall the intrinsic description of the stochastic Hamilton-Pontryagin principle. Denote by $G_{\mathcal{P}Q} : \mathcal{P}Q \to \mathbb{R}$ the pairing map $(u_q, p_q) \in \mathcal{P}Q \mapsto \langle p_q, u_q \rangle$. Let $G_{\mathcal{P}(Q/G)}$ denote the corresponding map for $\mathcal{P}(Q/G)$. Let $\rho_{TT^*Q} : TT^*Q \to \mathcal{P}Q$ denote the map $(q, p, u_q, u_p) \in TT^*Q \mapsto (q, u_q, p) \in \mathcal{P}Q$. Let

$$\begin{aligned} \mathrm{pr}_Q &: \mathcal{P}Q \to Q \\ \mathrm{pr}_{TQ} &: \mathcal{P}Q \to TQ \\ \mathrm{pr}_{T^*Q} &: \mathcal{P}Q \to T^*Q \end{aligned}$$

denote the projections $(q, v, p) \in \mathcal{P}Q \mapsto q \in Q$, $(q, v, p) \in \mathcal{P}Q \mapsto (q, v) \in TQ$ and $(q, v, p) \in \mathcal{P}Q \mapsto (q, p) \in T^*Q$, respectively. The corresponding projections in $\mathcal{P}(Q/G)$ are denoted by $\mathrm{pr}_{Q/G}$, $\mathrm{pr}_{T(Q/G)}$, and $\mathrm{pr}_{T^*(Q/G)}$, respectively. Consider the 1-form $\mathcal{G}_{\mathcal{P}Q}$ on $\mathcal{P}Q$ defined by the composition $\mathcal{G}_{\mathcal{P}Q} = G_{\mathcal{P}Q} \circ \rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}$. The corresponding 1-form on $\mathcal{P}(Q/G)$ is denoted by $\mathcal{G}_{\mathcal{P}(Q/G)}$.

Given a vector field $V \in \mathfrak{X}^\infty(Q)$, let $\bar{V} : \mathcal{P}Q \to \mathcal{P}Q$ denote the map defined by $\bar{V}(q, v, p) = (q, V(q), p) = (V \circ \mathrm{pr}_Q) \oplus \mathrm{pr}_{T^*Q}$. If V is a vector field on Q/G instead, then we will slightly abuse notation to still denote by $\bar{V} : \mathcal{P}(Q/G) \to \mathcal{P}(Q/G)$ the map $\bar{V}(x, u, y) = (x, V(x), y)$, where $(x, u, y) \in \mathcal{P}(Q/G)$. We define generalized energies $E_j \in C^\infty(\mathcal{P}Q)$ by

$$E_j = \begin{cases} G_{\mathcal{P}Q} - \mathcal{L} \circ \mathrm{pr}_{TQ}, \text{ if } j = 0, \\ G_{\mathcal{P}Q} \circ \bar{V}_j - L_j \circ \mathrm{pr}_Q, \text{ if } j = 1, \cdots, k. \end{cases}$$

In local coordinates, we have

$$E_j(q, v, p) = \begin{cases} \langle p_q, v_q \rangle - \mathcal{L}(q, v_q), \text{ if } j = 0, \\ \langle p_q, V_j(q) \rangle - L_j(q), \text{ if } j = 1, \cdots, k. \end{cases}$$

We have seen that the stochastic Hamilton-Pontryagin action functional can be written as

$$\mathcal{S}_X(\Gamma) = \int_0^T \mathcal{G}_{\mathcal{P}Q} \bullet d\Gamma - \sum_{j=0}^{k} \int_0^T E_j(\Gamma) \bullet dX^j. \tag{10.2.1}$$

Now suppose $\mathcal{L}, L_1, \cdots, L_k$ and $V_1, \cdots, V_k$ are all G-invariant. Then, for all $g \in G$, $\mathcal{S}_X(g\Gamma) = \mathcal{S}_X(\Gamma)$. Therefore $\mathcal{S}_X(\cdot)$ drops to an action $\mathcal{S}_X^{\mathrm{red}}(\cdot)$ on $\mathcal{P}Q/G$ defined by

$$\mathcal{S}_X^{\mathrm{red}}([\Gamma]) = \mathcal{S}_X(\Gamma).$$

Via the isomorphism $\mathcal{P}Q/G \cong \mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*$, we can identify $[\Gamma]$ with a semimartingale $\Gamma^{\mathcal{P}(Q/G)} \oplus \bar{\zeta} \oplus \bar{\mu}$, where $\Gamma^{\mathcal{P}(Q/G)} \in \mathscr{S}(\mathcal{P}(Q/G))$, $\bar{\zeta} \in \mathscr{S}(\tilde{\mathfrak{g}})$ and $\bar{\mu} \in \mathscr{S}(\tilde{\mathfrak{g}}^*)$.

We define $\mathcal{E}_j \in C^\infty(\mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*)$ by

$$\mathcal{E}_j(x, u, y, \bar{\zeta}, \bar{\mu}) = \begin{cases} \langle y, u \rangle + \langle \bar{\mu}, \bar{\zeta} \rangle - \ell(x, u, \bar{\zeta}), \text{ if } j = 0, \\ \langle y, V_j^{\mathrm{red}}(x) \rangle + \langle \bar{\mu}, \bar{\beta}_j(x) \rangle - l_j(x), \text{ if } j = 1, \cdots, k. \end{cases}$$

Denote by $G_{\tilde{\mathfrak{g}}} : \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^* \to \mathbb{R}$ the fibrewise pairing map $(\bar{\zeta}_x, \bar{\mu}_x) \mapsto \left\langle \bar{\zeta}_x, \bar{\mu}_x \right\rangle$, where $\bar{\zeta}_x, \bar{\mu}_x$ are elements in the fibre over $x \in Q/G$ respectively. If $\bar{\mu}$ is a semimartingale in $\tilde{\mathfrak{g}}^*$ and q_t is a semimartingale in Q with $\bar{\mu}_t$ projecting onto $\pi(q_t)$, we consider the semimartingale $\Psi_{\bar{\mu}} \in T^*Q$ over q_t defined by

$$\Psi_{\bar{\mu}_t}(v_{q_t}) = G_{\tilde{\mathfrak{g}}}(\bar{\mu}_t, [q_t, A(q_t, v_{q_t})]_G) \tag{10.2.2}$$

for all $v_{q_t} \in T_{q_t}Q$.

Given a semimartingale q_t in Q, we let $\bar{\xi}_t = [q_t, \xi_t]_G$, where $\xi_t = \int A \bullet dq_t$. Let us define

$$\int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle = \int \Psi_{\bar{\mu}_t} \bullet dq_t. \tag{10.2.3}$$

Remark 10.2.1. If Γ is a semimartingale in a manifold M, we usually think of $\bullet d\Gamma$ as a formal tangent vector in TM. However, in this case, we will think of $\bullet d\bar{\xi}$ formally as an element of $\tilde{\mathfrak{g}}$, rather than $T\tilde{\mathfrak{g}}$. To motivate this interpretation, let us note that if $\xi_t = \int A(q_t) \bullet dq_t$ then $\bullet d\xi_t = A(q_t) \bullet dq_t$ is a formal element in $\mathfrak{g}$. When $q_t = q(t)$ is a deterministic curve, we have $\dot{\xi}(t) = A(q(t), \dot{q}(t))$. The curve $A(q(t), \dot{q}(t))$ is referred to as $\xi(t)$ in Yoshimura and Marsden [95], and they set $\bar{\xi}(t) = [q(t), \xi(t)]_G$. In our case, we replace $\xi(t)$ by $\dot{\xi}(t)$ so that $\bar{\xi}(t) = [q(t), \dot{\xi}(t)]_G$. But when q_t is a semimartingale, we replace $\dot{\xi}(t) = A(q(t), \dot{q}(t))$ by $\bullet d\xi_t = A(q_t) \bullet dq_t$. Then, the right side is the formal Stratonovich differential $[q_t, A(q_t) \bullet dq_t]_G$, and hence we must replace $\bar{\xi}(t)$ by a Stratonovich differential $\bullet d\bar{\xi}_t$. The definition of $\int \left\langle \bar{\mu}, \bullet d\bar{\xi}_t \right\rangle$ interprets the pairing of $\bullet d\bar{\xi}_t$ with a semimartingale in $\tilde{\mathfrak{g}}^*$ as a well-defined Stratonovich integral in Q.

Theorem 10.2.2. The reduced action functional $\mathcal{S}_X^{\mathrm{red}}(\cdot)$ is given by

$$\begin{aligned}\mathcal{S}_X^{\mathrm{red}}(\Gamma^{\mathcal{P}(Q/G)} \oplus \bar{\zeta} \oplus \bar{\mu}) = &\int_0^T \mathcal{G}_{\mathcal{P}(Q/G)} \bullet d\Gamma^{\mathcal{P}(Q/G)} + \int_0^T \left\langle \bar{\mu}, \bullet d\bar{\xi} \right\rangle \\ &- \sum_{j=0}^{k} \mathcal{E}_j(\Gamma^{\mathcal{P}(Q/G)}, \bar{\zeta}, \bar{\mu}) \bullet dX^j.\end{aligned} \tag{10.2.4}$$

Proof. Let $(q, v, p) \in \mathcal{P}Q$, $x = \pi(q)$, $u = T_q\pi(v)$, $y = (p)_q^{h^*}$, $\bar{\zeta} = [q, A(q, v)]_G$ and $\bar{\mu} = [q, J(q, p)]_G$. Then, we have,

$$\langle p, v \rangle = \langle p, \mathrm{Hor}\,(v) \rangle + \langle p, \mathrm{Ver}\,(v) \rangle = \langle y, u \rangle + \left\langle \bar{\mu}, \bar{\zeta} \right\rangle,$$

and similarly,

$$\langle p, V_i(q) \rangle = \langle p, \mathrm{Hor}\, V_i(q) \rangle + \langle p, \mathrm{Ver}\, V_i(q) \rangle = \left\langle y, V_i^{\mathrm{red}}(x) \right\rangle + \left\langle \bar{\mu}, \bar{\beta}_i(x) \right\rangle.$$

As a result,

$$\begin{aligned}E_j(q, v, p) &= \begin{cases} \langle p_q, v_q \rangle - \mathcal{L}(q, v_q), \text{ if } j = 0, \\ \langle p_q, V_j(q) \rangle - L_j(q), \text{ if } j = 1, \cdots, k. \end{cases} \\ &= \begin{cases} \langle y, u \rangle + \left\langle \bar{\mu}, \bar{\zeta} \right\rangle - \ell(x, u, \bar{\zeta}), \text{ if } j = 0, \\ \left\langle y, V_j^{\mathrm{red}}(x) \right\rangle + \left\langle \bar{\mu}, \bar{\beta}_j(x) \right\rangle - l_j(x), \text{ if } j = 1, \cdots, k. \end{cases}\end{aligned}$$

$$= \mathcal{E}_j(x, u, y, \bar{\zeta}, \bar{\mu}).$$

Next suppose $(\dot{q}, \dot{v}, \dot{p}) \in T_{(q,v,p)}\mathcal{P}Q$. Let $\pi_{\mathcal{P}(Q/G)} : \mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^* \to \mathcal{P}(Q/G)$ denote the projection onto the $\mathcal{P}(Q/G)$ factor. Then, we obtain a map $\Pi_{\mathcal{P}(Q/G)} : \mathcal{P}Q \to \mathcal{P}(Q/G)$ by $\Pi_{\mathcal{P}(Q/G)} = \pi_{\mathcal{P}(Q/G)} \circ \tilde{\Psi}_A \circ \mathrm{Pr}_{\mathcal{P}Q/G}$. Let $(\dot{x}, \dot{u}, \dot{y}) = T_{(q,v,p)}\Pi_{\mathcal{P}(Q/G)}(\dot{q}, \dot{v}, \dot{p}) \in T_{(x,u,y)}\mathcal{P}(Q/G)$. Then

$$\begin{aligned}
\mathcal{G}_{\mathcal{P}Q}(q, v, p)(\dot{q}, \dot{v}, \dot{p}) &= \langle p, \dot{q} \rangle \\
&= \langle p, \mathrm{Hor}\, \dot{q} \rangle + \langle p, \mathrm{Ver}\, \dot{q} \rangle \\
&= \langle y, \dot{x} \rangle + \langle \bar{\mu}, [q, A(q, \dot{q})]_G \rangle \\
&= \mathcal{G}_{\mathcal{P}(Q/G)}(x, u, y)(\dot{x}, \dot{u}, \dot{y}) + \Psi_{\bar{\mu}}(\dot{q}).
\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}
\mathcal{S}_X^{\mathrm{red}}(\Gamma^{\mathcal{P}(Q/G)} \oplus \bar{\zeta} \oplus \bar{\mu}) = \int_0^T \mathcal{G}_{\mathcal{P}(Q/G)} \bullet d\Gamma^{\mathcal{P}(Q/G)} + \int_0^T \left\langle \bar{\mu}, \bullet d\bar{\xi} \right\rangle \\
- \sum_{j=0}^{k} \mathcal{E}_j(\Gamma^{\mathcal{P}(Q/G)} \oplus \bar{\zeta} \oplus \bar{\mu}) \bullet dX^j.
\end{aligned}$$

∎

If $\Gamma_t = (q_t, v_t, p_t)$ in local coordinates and $\tilde{\Psi}_A \circ \mathrm{Pr}_{\mathcal{P}Q}(\Gamma_t) = (x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t)$, then the local coordinate expression of $\mathcal{S}_X^{\mathrm{red}}(\cdot)$ is given by

$$\begin{aligned}
\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t) = \int_0^T \Bigg[& \ell(x_t, u_t, \bar{\zeta}_t) \bullet dX_t^0 + \sum_{i=1}^{k} l_i(x_t) \bullet dX_t^i && (10.2.5) \\
& + \left\langle y_t, \bullet dx_t - u_t \bullet dX_t^0 - \sum_{i=1}^{k} V_i^{\mathrm{red}}(x_t) \bullet dX_t^i \right\rangle \\
& + \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t - \bar{\zeta}_t \bullet dX_t^0 - \sum_{i=1}^{k} \bar{\beta}_i(x_t) \bullet dX_t^i \right\rangle \Bigg]. && (10.2.6)
\end{aligned}$$

10.2.2 Vertical and Horizontal Variations

Let us consider a semimartingale q_t in Q. Define

$$\xi_t = \int A(q_t) \bullet dq_t \in \mathscr{S}(\mathfrak{g}).$$

Theorem 10.2.3. Suppose $\epsilon \mapsto q_{t,\epsilon}$ is a deformation of q_t.

1. If δq_t is vertical then

$$\mathcal{D}\xi_t = \int \mathrm{ad}_{\eta_t} \bullet d\xi_t + \eta_t - \eta_0 \tag{10.2.7}$$

   where $\eta_t = A(q_t, \delta q_t)$.

2. If δq is horizontal then

$$\mathcal{D}\xi_t = \int B(q_t)(\delta q_t, \bullet dq_t) := \int i_{\delta q_t} B \bullet dq_t. \tag{10.2.8}$$

Proof.

1. Since Hor $\delta q_t = 0$ it follows from (2.5.1) that $B(q_t)(\delta q_t, \cdot) = 0$. Hence, by Lemma 9.1.9

$$\begin{aligned}\mathcal{D}\xi_t &= \int i_{\delta q_t} dA \bullet dq_t + \langle A(q_t), \delta q_t\rangle - \langle A(q_0), \delta q_0\rangle \\ &= \int \mathrm{ad}_{\eta_t} A(q_t) \bullet dq_t + \eta_t - \eta_0 \\ &= \int \mathrm{ad}_{\eta_t} \bullet d\int A(q_t) \bullet dq_t + \eta_t - \eta_0 \\ &= \int \mathrm{ad}_{\eta_t} \bullet d\xi_t + \eta_t - \eta_0.\end{aligned}$$

2. Since $A(q_t, \delta q_t) = 0$, we have

$$\begin{aligned}\mathcal{D}\xi_t &= \int i_{\delta q_t} dA \bullet dq_t + \langle A(q_t), \delta q_t\rangle - \langle A(q_0), \delta q_0\rangle \\ &= \int i_{\delta q_t} B \bullet dq_t \\ &= \int B(q_t)(\delta q_t, \bullet dq_t).\end{aligned}$$

∎

10.2.3 Covariant Variations on the Adjoint Bundle

In this section, we discuss variation of the term $\int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle$ in the reduced action functional.

Theorem 10.2.4. Let q_t be a semimartingale in Q, $\bar{\mu}_t$ be a semimartingale in $\tilde{\mathfrak{g}}^*$ that projects to $x_t := \pi(q_t)$ in Q/G and $\bar{\xi}_t = [q_t, \xi_t]_G$ with $\xi_t = \int A(q_t) \bullet dq_t$. Suppose $\epsilon \mapsto q_{\epsilon,t}$ and $\epsilon \mapsto \bar{\mu}_{\epsilon,t}$ are deformations of q_t and $\bar{\mu}_t$ respectively. Set $x_{\epsilon,t} = \pi(q_{\epsilon,t})$, $\eta_t = A(q_t, \delta q_t)$ and $\bar{\eta}_t = [q_t, \eta_t]_G$. Then

$$\begin{aligned}\mathcal{D}\int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle = \int \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t - \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi} \right\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t) \right\rangle \bullet dx_t - \int \left\langle \bullet D\bar{\mu}_t, \bar{\eta}_t \right\rangle \\ + \left\langle \bar{\mu}_t, \bar{\eta}_t \right\rangle - \left\langle \bar{\mu}_0, \bar{\eta}_0 \right\rangle. \end{aligned} \tag{10.2.9}$$

Proof. To make our calculations easier, we begin by considering the 1-form $\bar{\Psi}$ on $\tilde{\mathfrak{g}}^* \times_{Q/G} Q$ defined by $\bar{\Psi}_{(\bar{\mu},q)}(v_{\bar{\mu}}, v_q) = \Psi_{\bar{\mu}}(v_q) = \langle \bar{\mu}, [q, A(q, v_q)]_G\rangle$, where $(\bar{\mu}, q) \in \tilde{\mathfrak{g}}^* \times_{Q/G} Q$ and $(v_{\bar{\mu}}, v_q) \in T_{(\bar{\mu},q)}(\tilde{\mathfrak{g}}^* \times_{Q/G} Q)$. Then, we have,

$$\int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle = \int \bar{\Psi} \bullet d(\bar{\mu}_t, q_t). \tag{10.2.10}$$

This implies, by Lemma 9.1.9, that

$$\mathcal{D}\int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle = \int i_{(\delta\bar{\mu}_t,\delta q_t)}\mathbf{d}\bar{\Psi} \bullet d(\bar{\mu}_t, q_t) + \left\langle \bar{\Psi}(\bar{\mu}_t, q_t), (\delta\bar{\mu}_t, \delta q_t)\right\rangle - \left\langle \bar{\Psi}(\bar{\mu}_0, q_0), (\delta\bar{\mu}_0, \delta q_0)\right\rangle .$$

We now compute $i_{(\delta\bar{\mu}_t,\delta q_t)}\mathbf{d}\bar{\Psi}$ by Corollary 9.1.10. Consider tangent vectors $(v_{\bar{\mu}}, v_q), (w_{\bar{\mu}}, w_q) \in T_{(\bar{\mu},q)}(\tilde{\mathfrak{g}}^* \times_{Q/G} Q)$. Let $q_\epsilon(t)$ and $\bar{\mu}_\epsilon(t)$ be curves in Q and $\tilde{\mathfrak{g}}^*$ respectively that project on the same curve $x_\epsilon(t)$ and set $q_0(t) = q(t)$, $\bar{\mu}_0(t) = \bar{\mu}(t)$ and $x_0(t) = x(t)$. We assume that $q(0) = q$, $\bar{\mu}(0) = \bar{\mu}$, $\dot{q}(0) = v_q$, $\delta q(0) = w_q$, $\dot{\bar{\mu}}(0) = v_{\bar{\mu}}$ and $\delta\bar{\mu}(0) = w_{\bar{\mu}}$. Further let $x(0) = x$, $v_x = \dot{x}(0)$ and $w_x = \delta x(0)$. We have

$$\begin{aligned} i_{(\delta\bar{\mu}(t),\delta q(t))}\mathbf{d}\bar{\Psi}(\dot{\bar{\mu}}, \dot{q}) &= \delta \left\langle \bar{\Psi}(\bar{\mu}(t), q(t)), (\dot{\bar{\mu}}(t), \dot{q}(t))\right\rangle - \frac{d}{dt}\left\langle \bar{\Psi}(\bar{\mu}(t), q(t)), (\delta\bar{\mu}(t), \delta q(t))\right\rangle \\ &= \delta \left\langle \bar{\mu}(t), [q(t), A(q(t), \dot{q}(t))]_G \right\rangle - \frac{d}{dt}\left\langle \bar{\mu}(t), [q(t), A(q(t), \delta q(t))]_G \right\rangle . \end{aligned}$$

We now use Equation (10.1.8) to obtain

$$\begin{aligned} & i_{(\delta\bar{\mu}(t),\delta q(t))}\mathbf{d}\bar{\Psi}(\dot{\bar{\mu}}, \dot{q}) \\ &= \delta \left\langle \bar{\mu}(t), [q(t), A(q(t), \dot{q}(t))]_G \right\rangle - \frac{d}{dt}\left\langle \bar{\mu}(t), [q(t), A(q(t), \delta q(t))]_G \right\rangle \\ &= \left\langle \frac{D}{D\epsilon}\Big|_{\epsilon=0}\bar{\mu}_\epsilon(t), [q(t), A(q(t), \dot{q}(t))]_G \right\rangle + \left\langle \bar{\mu}, \frac{D}{D\epsilon}\Big|_{\epsilon=0}[q(t), A(q(t), \dot{q}(t))]_G \right\rangle \\ &- \left\langle \frac{D}{Dt}\bar{\mu}(t), [q(t), A(q(t), \delta q(t))]_G \right\rangle - \left\langle \bar{\mu}, \frac{D}{Dt}[q(t), A(q(t), \delta q(t))]_G \right\rangle \\ &= \left\langle \frac{D}{D\epsilon}\Big|_{\epsilon=0}\bar{\mu}_\epsilon(t), [q(t), A(q(t), \dot{q}(t))]_G \right\rangle - \left\langle \bar{\mu}(t), [q(t), \mathrm{ad}_{A(q(t),\delta q(t))}A(q(t), \dot{q}(t))]_G \right\rangle \\ &+ \left\langle \bar{\mu}, \frac{D}{Dt}[q(t), A(q(t), \delta q(t))]_G \right\rangle + \left\langle \bar{\mu}(t), [q(t), B(q(t))(\delta q(t), \dot{q}(t))]_G \right\rangle \\ &- \left\langle \frac{D}{Dt}\bar{\mu}(t), [q(t), A(q(t), \delta q(t))]_G \right\rangle - \left\langle \bar{\mu}, \frac{D}{Dt}[q(t), A(q(t), \delta q(t))]_G \right\rangle . \end{aligned}$$

In the last step, we have used Equation (2.5.8) to write

$$\begin{aligned} & \frac{D}{D\epsilon}\Big|_{\epsilon=0}[q_\epsilon(t), A(q_\epsilon(t), \dot{q}_\epsilon(t))]_G \\ &= \frac{D}{Dt}[q(t), A(q(t), \delta q(t))]_G - [q(t), \mathrm{ad}_{A(q(t),\delta q(t))}A(q(t), \dot{q}(t))]_G \\ &+ [q(t), B(q(t))(\delta q(t), \dot{q}(t))]_G . \end{aligned}$$

We observe that the term $\left\langle \bar{\mu}, \frac{D}{Dt}[q(t), A(q(t), \delta q(t))]_G \right\rangle$ cancels out. The covariant derivatives can be expressed by the connector via Equation (10.1.6) to yield

$$\begin{aligned} & i_{(\delta\bar{\mu}(t),\delta q(t))}\mathbf{d}\bar{\Psi}(\dot{\bar{\mu}}, \dot{q}) \\ &= \left\langle \frac{D}{D\epsilon}\Big|_{\epsilon=0}\bar{\mu}_\epsilon(t), [q(t), A(q(t), \dot{q}(t))]_G \right\rangle - \left\langle \bar{\mu}(t), [q(t), \mathrm{ad}_{A(q(t),\delta q(t))}A(q(t), \dot{q}(t))]_G \right\rangle \end{aligned}$$

$$
\begin{aligned}
&+ \langle \bar{\mu}(t), [q(t), B(q(t))(\delta q(t), \dot{q}(t))]_G \rangle - \left\langle \frac{D}{Dt}\bar{\mu}(t), [q(t), A(q(t), \delta q(t))]_G \right\rangle \\
&= \langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \bar{\mu}_\epsilon(t), [q(t), A(q(t), \dot{q}(t))]_G \rangle - \left\langle \mathrm{ad}^*_{[q(t), A(q(t), \delta q(t))]_G} \bar{\mu}(t), [q(t), A(q(t), \dot{q}(t))]_G \right\rangle \\
&+ \left\langle \left\langle \bar{\mu}(t), i_{\delta x(t)} \tilde{B}(x(t)) \right\rangle, \dot{x}(t) \right\rangle - \langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \bar{\mu}(t), [q(t), A(q(t), \delta q(t))]_G \rangle
\end{aligned}
$$

Evaluating at $t = 0$ we obtain

$$
\begin{aligned}
i_{(w_{\bar{\mu}}, w_q)} \mathbf{d}\bar{\Psi}(v_{\bar{\mu}}, v_q) &= \langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} w_{\bar{\mu}}, [q, A(q, v_q)]_G \rangle - \left\langle \mathrm{ad}^*_{[q, A(q, w_q)]_G} \bar{\mu}, [q, A(q, v_q)]_G \right\rangle \\
&\quad + \left\langle \left\langle \bar{\mu}, i_{w_x} \tilde{B}(x) \right\rangle, v_x \right\rangle - \langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} v_{\bar{\mu}}, [q, A(q, w_q)]_G \rangle .
\end{aligned}
$$

Since all of the tangent vectors chosen are arbitrary, we obtain

$$
\begin{aligned}
&\int i_{(\delta \bar{\mu}_t, \delta q_t)} \mathbf{d}\bar{\Psi} \bullet d(\bar{\mu}_t, q_t) \\
&= \int \Psi_{K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t - \mathrm{ad}^*_{[q, A(q, \delta q_t)]_G} \bar{\mu}_t} \bullet dq_t + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t) \right\rangle \bullet dx_t - \int \langle \bullet D\bar{\mu}_t, [q_t, A(q_t, \delta q_t)]_G \rangle \\
&= \int \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t - \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi} \right\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t) \right\rangle \bullet dx_t - \int \langle \bullet D\bar{\mu}_t, \bar{\eta}_t \rangle ,
\end{aligned}
$$

which yields

$$
\begin{aligned}
\mathcal{D} \int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle &= \int \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t - \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi} \right\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t) \right\rangle \bullet dx_t - \int \langle \bullet D\bar{\mu}_t, \bar{\eta}_t \rangle \\
&\quad + \left\langle \bar{\Psi}(\bar{\mu}_t, q_t), (\delta \bar{\mu}_t, \delta q_t) \right\rangle - \left\langle \bar{\Psi}(\bar{\mu}_0, q_0), (\delta \bar{\mu}_0, \delta q_0) \right\rangle \\
&= \int \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t - \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi} \right\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t) \right\rangle \bullet dx_t - \int \langle \bullet D\bar{\mu}_t, \bar{\eta}_t \rangle \\
&\quad + \langle \bar{\mu}_t, \bar{\eta}_t \rangle - \langle \bar{\mu}_0, \bar{\eta}_0 \rangle .
\end{aligned}
$$

This completes the proof. ∎

Let us now show that the above formula can also be obtained by a formal computation by considering covariant variations of the semimartingale $\bar{\xi}_t = [q_t, \xi_t]_G$. Since $\epsilon \mapsto \bar{\xi}_{\epsilon,t}$ is pathwise smooth, the covariant derivative $\frac{D}{D\epsilon}\Big|_{\epsilon=0} \bar{\xi}_{\epsilon,t}$ with respect to ϵ in the ucp topology agrees with the pathwise computed covariant $\frac{D}{D\epsilon}\Big|_{\epsilon=0} \bar{\xi}_t(\cdot)$. This is a consequence of Arnaudon and Thalmaier [5, Proposition 2.8]. Using the covariant derivative on the adjoint bundle and Proposition 10.2.3, we obtain

$$
\begin{aligned}
\delta^A \bar{\xi}_t &:= \frac{D}{D\epsilon}\Big|_{\epsilon=0} \bar{\xi}_{\epsilon,t} \\
&= \left[q_t, -\mathrm{ad}_{\eta_t} \int A(q_t) \bullet dq_t + \mathcal{D}\xi_t \right]_G \\
&= \left[q_t, -\mathrm{ad}_{\eta_t} \int A(q_t) \bullet dq_t \right]_G + \left[q_t, \int \bullet d\eta_t + \int \mathrm{ad}_{\eta_t} \bullet d\xi_t \right]_G
\end{aligned}
$$

$$+\left[q_t, \int i_{\delta q_t} B \bullet dq_t\right]_G .$$

We now formally set

$$\int \left\langle \bar{\mu}_t, \bullet d(\delta^A \bar{\xi}_t)\right\rangle = \int \left\langle \bar{\mu}_t, [q_t, -\mathrm{ad}_{\eta_t} A(q_t) \bullet dq_t]_G \right\rangle + \int \left\langle \bar{\mu}_t, [q_t, \bullet d\eta_t + \mathrm{ad}_{\eta_t} \bullet d\xi_t]_G \right\rangle$$
$$+ \int \left\langle \bar{\mu}_t, [q_t, i_{\delta q_t} B \bullet dq_t]_G \right\rangle .$$

To obtain each of the integrals on the right as Stratonovich integrals, we formally treat the Stratonovich differentials as tangent vectors in each of the pairings involved. For the first term, we have

$$\begin{aligned} \left\langle \bar{\mu}_t, [q_t, -\mathrm{ad}_{\eta_t} A(q_t) \bullet dq_t]_G \right\rangle &= - \left\langle \bar{\mu}_t, \mathrm{ad}_{\bar{\eta}_t} [q_t, A(q_t) \bullet dq_t]_G \right\rangle \\ &= - \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, [q_t, A(q_t) \bullet dq_t]_G \right\rangle \end{aligned}$$

which shows that $\int \left\langle \bar{\mu}_t, [q_t, -\mathrm{ad}_{\eta_t} A(q_t) \bullet dq_t]_G \right\rangle$ corresponds to $-\int \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle$. For the next term, we note that the stochastic covariant derivative of $\bar{\eta}_t$ on the adjoint bundle is given by

$$\bullet D\bar{\eta}_t = [q_t, [\eta_t, A(q_t) \bullet dq_t] + \bullet d\eta_t]_G.$$

Hence, $\int \left\langle \bar{\mu}_t, [q_t, \bullet d\eta_t + \mathrm{ad}_{\eta_t} \bullet d\xi_t]_G \right\rangle$ corresponds to $\int \langle \bar{\mu}, \bullet D\bar{\eta}\rangle$. For the final term on the right side, we have

$$\begin{aligned} \langle \bar{\mu}_t, [q_t, i_{\delta q_t} B \bullet dq_t]\rangle &= \left\langle \bar{\mu}_t, \tilde{B}(x_t)(\delta x_t, \bullet dx_t)\right\rangle \\ &= \left\langle \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t)\right\rangle, \bullet dx_t \right\rangle \end{aligned}$$

and hence, we can write

$$\int \langle \bar{\mu}_t, [q_t, i_{\delta q_t} B \bullet dq_t]\rangle = \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t)\right\rangle \bullet dx_t.$$

Thus, we obtain

$$\int \left\langle \bar{\mu}_t, \bullet d(\delta^A \bar{\xi}_t)\right\rangle = -\int \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle + \int \langle \bar{\mu}_t, \bullet D\bar{\eta}_t\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t)\right\rangle \bullet dx_t.$$

which we will write as

$$\bullet d(\delta^A \bar{\xi}_t) = -\mathrm{ad}_{\bar{\eta}_t} \bullet d\bar{\xi}_t + \bullet D\bar{\eta}_t + i_{\delta x_t} \tilde{B}(x_t) \bullet dx_t.$$

to emphasize the similarity with the deterministic case (see Equation (2.5.8)). By Equation (10.1.10), we have

$$\langle \bar{\mu}, \bar{\eta}\rangle - \langle \bar{\mu}_0, \bar{\eta}_0\rangle = \int \langle D\bar{\mu}, \bar{\eta}\rangle + \int \langle \bar{\mu}, \bullet D\bar{\eta}\rangle .$$

Putting all of this together, we have

$$\int \left\langle \bar{\mu}_t, \bullet d(\delta^A \bar{\xi}_t)\right\rangle = -\int \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle - \int \langle D\bar{\mu}, \bar{\eta}\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t)\right\rangle \bullet dx_t$$

$$+ \left\langle \bar{\mu}_t, \bar{\eta}_t \right\rangle - \left\langle \bar{\mu}_0, \bar{\eta}_0 \right\rangle . \tag{10.2.11}$$

Note that the terms on the right are well-defined Stratonovich integrals. By formally considering $\frac{D}{D\epsilon}\Big|_{\epsilon=0}(\bullet d\bar{\xi}_{\epsilon,t}) = \bullet d(\delta^A \bar{\xi}_t)$, we obtain

$$\mathcal{D}\int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle = \int \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle + \int \left\langle \bar{\mu}_t, \bullet d(\delta^A \bar{\xi}_t) \right\rangle .$$

Replacing $\int \left\langle \bar{\mu}_t, \bullet d(\delta^A \bar{\xi}_t) \right\rangle$ by the expression in Equation (10.2.11) gives us the expression for $\delta \int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle$ obtained in Equation (10.2.9).

Remark 10.2.5. As in the deterministic case, for the semimartingale $\bar{\xi}_t = [q_t, \int A(q_t) \bullet dq_t]_G$, we will only consider deformations $\epsilon \mapsto \bar{\xi}_{\epsilon,t}$ such that their projections on Q/G is $x_t = \pi(q_t)$. The corresponding variations can be identified with a semimartingale in $\tilde{\mathfrak{g}}$, rather than one in $T\tilde{\mathfrak{g}}$. The variation $\delta^A \bar{\xi}$ is an instance of $\tilde{\mathfrak{g}}$-fiber variation.

We have already seen that if q_t is a semimartingale such that $x_t = \pi(q_t)$ and $\bar{\xi}_t = [q_t, \int A(q_t) \bullet dq_t]_G$ then, given a deformation $\epsilon \mapsto q_{\epsilon,t}$, we (formally) have

$$\bullet d(\delta^A \bar{\xi}_t) = -\mathrm{ad}_{\bar{\eta}_t} \bullet d\bar{\xi}_t + \bullet D\bar{\eta}_t + i_{\delta x_t} \tilde{B}(x_t) \bullet dx_t.$$

Moreover, we also have

$$\mathcal{D}\int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle - \int \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle = \int \left\langle \bar{\mu}_t, \bullet d(\delta^A \bar{\xi}_t) \right\rangle$$

where

$$\begin{aligned} \int \left\langle \bar{\mu}_t, \bullet d(\delta^A \bar{\xi}_t) \right\rangle = &- \int \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi} \right\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t) \right\rangle \bullet dx_t - \int \left\langle \bullet D\bar{\mu}_t, \bar{\eta}_t \right\rangle \\ &+ \left\langle \bar{\mu}_t, \bar{\eta}_t \right\rangle - \left\langle \bar{\mu}_0, \bar{\eta}_0 \right\rangle . \end{aligned}$$

Thus, in the stochastic case, the variations $\delta^A \bar{\xi}_t$ are those $\tilde{\mathfrak{g}}$-fiber variations for which

$$\begin{aligned} &\mathcal{D}\int \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle - \int \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t, \bullet d\bar{\xi}_t \right\rangle \\ &= \int \left\langle \bar{\mu}_t, \bullet D\bar{\eta}_t \right\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t) \right\rangle \bullet dx_t - \int \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi} \right\rangle \\ &= - \int \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t, \bullet d\bar{\xi} \right\rangle + \int \left\langle \bar{\mu}_t, i_{\delta x_t} \tilde{B}(x_t) \right\rangle \bullet dx_t - \int \left\langle \bullet D\bar{\mu}_t, \bar{\eta}_t \right\rangle \\ &+ \left\langle \bar{\mu}_t, \bar{\eta}_t \right\rangle - \left\langle \bar{\mu}_0, \bar{\eta}_0 \right\rangle , \end{aligned}$$

where $\bar{\eta}_t = [q_t, A(q_t, \delta q_t)]_G$. The variations of the form $\delta x_t \oplus \delta^A \bar{\xi}_t$ are the stochastic analogues of the deterministic covariant variations. The reduced variational principle $\mathcal{S}_X^{\mathrm{red}}$ will be considered under constrained variations of this form.

10.2.4 The Stochastic Implicit Lagrange-Poincaré Reduction Theorem

In this section we will address the critical points of S_X^{red} given by Equation (10.2.5) and prove the stochastic version of the implicit Lagrange-Poincaré reduction theorem.

Fix a connection $\nabla^{T(Q/G)}$ on $T(Q/G)$, and let $\nabla^{T^*(Q/G)}$ denote the connection on $T^*(Q/G)$ obtained from $\nabla^{T(Q/G)}$. Then $\nabla^{T(Q/G)} \oplus \nabla^{\tilde{\mathfrak{g}}}$ defines a connection on $T(Q/G) \oplus \tilde{\mathfrak{g}}$ with associated connector $K_{\nabla^{T(Q/G)}} \oplus K_{\nabla^{\tilde{\mathfrak{g}}}}$, where $K_{\nabla^{T(Q/G)}}$ and $K_{\nabla^{\tilde{\mathfrak{g}}}}$ are the connectors corresponding to $\nabla^{T(Q/G)}$ and $\nabla^{\tilde{\mathfrak{g}}}$ respectively.

Theorem 10.2.6. Let $\Gamma_t = (x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t)$ be an admissible semimartingale in $\mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*$. Suppose $\epsilon \mapsto \Gamma_{\epsilon,t}^{|T} = (x_{\epsilon,t}^{|T}, u_{\epsilon,t}^{|T}, y_{\epsilon,t}^{|T}, \bar{\zeta}_{\epsilon,t}^{|T}, \bar{\mu}_{\epsilon,t}^{|T})$ is a deformation of $\Gamma_t^{|T} = (x_t^{|T}, u_t^{|T}, y_t^{|T}, \bar{\zeta}_t^{|T}, \bar{\mu}_t^{|T})$ such that the variations $\delta u_t^{|T}, \delta\bar{\zeta}_t^{|T}, \delta y_t^{|T}$ and $\delta\bar{\mu}_t^{|T}$ are arbitrary and the variations $\delta x_t^{|T} \oplus \delta^A(\bar{\xi}_t^{|T})$ satisfy $\delta x_t^{|T} = 0$ at $t = 0$ and $t = T$ and

$$\bullet d(\delta^A \bar{\xi}_t^{|T}) = -\mathrm{ad}_{\bar{\eta}_t} \bullet d\bar{\xi}_t^{|T} + \bullet D\bar{\eta}_t + i_{\delta x_t^{|T}} \tilde{B}(x_t^{|T}) \bullet dx_t^{|T}$$

in the sense of Remark 10.2.5. Here $\bar{\eta}$ is an arbitrary semimartingale in $\tilde{\mathfrak{g}}$ that vanishes at $t = 0$ and $t = T$. Then

$$\begin{aligned}
&\mathcal{D}\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t) \\
&= \int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left(\ell \bullet dX_t^0 + \sum_{i=1}^k \left(l_i - \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \right\rangle - \left\langle \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle \right) \bullet dX_t^i \right) \right. \\
&\quad \left. - \left\langle \bar{\mu}_t, i_{\bullet dx_t} \tilde{B}(x_t) \right\rangle - \bullet D y_t^{|T}, \delta x_t^{|T} \right\rangle + \int_0^T \left\langle \frac{\partial \ell}{\partial u_t^{|T}} - y_t^{|T}, K_{\nabla^{T(Q/G)}} \delta u_t^{|T} \right\rangle \bullet dX_t^0 \\
&\quad + \int_0^T \left\langle \frac{\partial \ell}{\partial \bar{\zeta}_t^{|T}} - \bar{\mu}_t^{|T}, K_{\nabla^{\tilde{\mathfrak{g}}}} \delta \bar{\zeta}_t \right\rangle \bullet dX_t^0 \\
&\quad + \int_0^T \langle K_{\nabla^{T^*(Q/G)}} \delta y_t^{|T}, \bullet dx_t^{|T} - u_t^{|T} \bullet dX_t^0 - \sum_{i=1}^k V_i^{\mathrm{red}}(x_t^{|T}) \bullet dX_t^i \rangle \\
&\quad + \int_0^T \langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} - \bar{\zeta}_t^{|T} \bullet dX_t^0 - \sum_{i=1}^k \bar{\beta}_i(x_t^{|T}) \bullet dX_t^i \rangle \\
&\quad + \int_0^T \langle - \bullet D\bar{\mu}_t^{|T} + \mathrm{ad}^*_{\bullet d\bar{\xi}_t^{|T}} \bar{\mu}_t^{|T}, \bar{\eta}_t \rangle
\end{aligned} \tag{10.2.12}$$

Consequently, $\mathcal{D}\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t) = 0$ holds if and only if the **horizontal stochastic Lagrange-Poincaré equations**

$$\begin{aligned}
\bullet D y_t = \frac{\partial}{\partial x_t} &\left(\ell \bullet dX_t^0 + \sum_{i=1}^k \left(l_i - \langle y_t, V_i^{\mathrm{red}}(x_t) \rangle - \langle \bar{\mu}_t, \bar{\beta}_i(x_t) \rangle \right) \bullet dX_t^i \right) \\
&- \left\langle \bar{\mu}_t, i_{\bullet dx_t} \tilde{B}(x_t) \right\rangle
\end{aligned}$$

$$\bullet dx_t = u_t \bullet dX_t^0 + \sum_{i=1}^{k} V_i^{\mathrm{red}}(x_t) \bullet dX_t^i$$

$$\left(y_t - \frac{\partial \ell}{\partial u_t}\right) \bullet dX_t^0 = 0$$

and the **vertical stochastic implicit Lagrange-Poincaré equations**

$$\begin{aligned}
\bullet D\bar{\mu}_t &= \mathrm{ad}^*_{\bullet d\bar{\xi}_t}\bar{\mu}_t \\
\bullet d\bar{\xi}_t &= \bar{\zeta}_t \bullet dX_t^0 + \sum_{i=1}^{k} \bar{\beta}_i(x_t) \bullet dX_t^i \\
\left(\bar{\mu}_t - \frac{\partial \ell}{\partial \bar{\zeta}_t}\right) \bullet dX_t^0 &= 0.
\end{aligned}$$

are satisfied by $(x_t^{|T}, u_t^{|T}, y_t^{|T}, \bar{\zeta}_t^{|T}, \bar{\mu}_t^{|T})$.

Proof.

We have

$$\begin{aligned}
&\mathcal{D}\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t) \\
&= \mathcal{D}\int_0^T \ell(x_t, u_t, \bar{\zeta}_t) \bullet dX_t^0 + \sum_{i=1}^{k} \mathcal{D}\int_0^T l_i(x_t) \bullet dX_t^i + \mathcal{D}\int_0^T \langle y_t, \bullet dx_t\rangle \\
&- \mathcal{D}\int_0^T \langle y_t, u_t\rangle \bullet dX_t^0 - \sum_{i=1}^{k} \int_0^T \langle y_t, V_i^{\mathrm{red}}(x_t)\rangle \bullet dX_t^i + \mathcal{D}\int_0^T \langle \bar{\mu}_t, \bullet d\bar{\xi}_t\rangle \\
&- \mathcal{D}\int_0^T \left\langle \bar{\mu}_t, \bar{\zeta}_t dt + \sum_{i=1}^{k} \bar{\beta}_i(x_t) \bullet dX_t^i \right\rangle. \\
&= \mathcal{D}\int_0^T \ell(x_t^{|T}, u_t^{|T}, \bar{\zeta}_t^{|T}) \bullet dX_t^0 + \sum_{i=1}^{k} \mathcal{D}\int_0^T l_i(x_t^{|T}) \bullet dX_t^i + \mathcal{D}\int_0^T \langle y_t^{|T}, \bullet dx_t^{|T}\rangle \\
&- \mathcal{D}\int_0^T \langle y_t^{|T}, u_t^{|T}\rangle \bullet dX_t^0 - \sum_{i=1}^{k} \int_0^T \langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T})\rangle \bullet dX_t^i + \mathcal{D}\int_0^T \langle \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T}\rangle \\
&- \mathcal{D}\int_0^T \left\langle \bar{\mu}_t^{|T}, \bar{\zeta}_t^{|T} dt + \sum_{i=1}^{k} \bar{\beta}_i(x_t^{|T}) \bullet dX_t^i \right\rangle. \\
&= \mathcal{D}\int_0^T \ell(x_t^{|T}, u_t^{|T}, \bar{\zeta}_t^{|T}) \bullet dX_t^0 + \sum_{i=1}^{k} \mathcal{D}\int_0^T l_i(x_t^{|T}) \bullet dX_t^i \\
&+ \mathcal{D}\int_0^T \langle y_t^{|T}, \bullet dx_t^{|T}\rangle - \Bigg(\mathcal{D}\int_0^T \langle y_t^{|T}, u_t^{|T}\rangle \bullet dX_t^0 \\
&+ \sum_{i=1}^{k} \mathcal{D}\int_0^T \langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T})\rangle \bullet dX_t^i \Bigg) + \mathcal{D}\int_0^T \langle \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T}\rangle
\end{aligned}$$

$$- \mathcal{D} \int_0^T \left\langle \bar{\mu}_t^{|T}, \bar{\zeta}_t^{|T} \bullet dX_t^0 + \sum_{i=1}^k \bar{\beta}_i(x_t^{|T}) \bullet dX_t^i \right\rangle . \tag{10.2.13}$$

We break down the calculation of variations of the terms in several steps:

1. **The terms** $\mathcal{D} \int_0^T \ell(x_t^{|T}, u_t^{|T}, \bar{\zeta}_t^{|T}) \bullet dX_t^0 + \sum_{i=1}^k \mathcal{D} \int_0^T l_i(x_t^{|T}) \bullet dX_t^i$

   By Theorem 10.1.4, we have

$$\begin{aligned}
&\mathcal{D} \int_0^T \ell(x_t^{|T}, u_t^{|T}, \bar{\zeta}_t^{|T}) \bullet dX_t^0 + \sum_{i=1}^k \mathcal{D} \int_0^T l_i(x_t^{|T}) \bullet dX_t^i \\
&= \int_0^T \left\langle d\ell(x_t^{|T}, u_t^{|T}, \bar{\zeta}_t^{|T}), (\delta x_t^{|T}, \delta u_t^{|T}, \delta \bar{\zeta}_t^{|T}) \right\rangle \bullet dX_t^0 + \sum_{i=1}^k \int_0^T \left\langle dl_i(x_t^{|T}), \delta x_t^{|T} \right\rangle \bullet dX_t^i \\
&= \int_0^T \left(\left\langle \frac{\partial \ell}{\partial x_t^{|T}}, \delta x_t^{|T} \right\rangle + \left\langle \frac{\partial \ell}{\partial u_t^{|T}}, K_{\nabla^{T(Q/G)}} \delta u_t^{|T} \right\rangle \right. \\
&\left. + \left\langle \frac{\partial \ell}{\partial \bar{\zeta}_t^{|T}}, K_{\nabla^{\tilde{\mathfrak{g}}}} \delta \bar{\zeta}_t^{|T} \right\rangle \right) \bullet dX_t^0 + \sum_{i=1}^k \int_0^T \left\langle \frac{\partial l_i}{\partial x_t^{|T}}, \delta x_t^{|T} \right\rangle \bullet dX_t^i .
\end{aligned} \tag{10.2.14}$$

2. **The term** $\mathcal{D} \int_0^T \langle y_t^{|T}, \bullet dx_t^{|T} \rangle$

   We recognize this term as $\mathcal{D} \int_0^T \mathcal{G}_{\mathcal{P}(Q/G)} \bullet d(x_t^{|T}, u_t^{|T}, y_t^{|T})$. Since $\mathcal{G}_{\mathcal{P}(Q/G)}(\delta x^{|T}, \delta u_t^{|T}, \delta y^{|T}) = \left\langle y_t^{|T}, \delta x_t^{|T} \right\rangle$ and $\delta x_t^{|T}$ vanishes at $t = 0$ and $t = T$, by Lemma 9.1.9, we have

$$\begin{aligned}
&\mathcal{D} \int_0^T \langle y_t^{|T}, \bullet dx_t^{|T} \rangle \\
&= \int_0^T i_{(\delta x_t^{|T}, \delta u_t^{|T}, \delta y_t^{|T})} \mathbf{d} \mathcal{G}_{\mathcal{P}(Q/G)} \bullet d(x_t^{|T}, u_t^{|T}, y_t^{|T}) \\
&+ \left\langle \mathcal{G}_{\mathcal{P}(Q/G)}(x_T^{|T}, u_T^{|T}, y_T^{|T}), (\delta x_T^{|T}, \delta u_T^{|T}, \delta y_T^{|T}) \right\rangle - \left\langle \mathcal{G}_{\mathcal{P}(Q/G)}(x_0^{|T}, u_0^{|T}, y_0^{|T}), (\delta x_0^{|T}, \delta u_0^{|T}, \delta y_0^{|T}) \right\rangle \\
&= \int_0^T i_{(\delta x_t^{|T}, \delta u_t^{|T}, \delta y_t^{|T})} \mathbf{d} \mathcal{G}_{\mathcal{P}(Q/G)} \bullet d(x_t^{|T}, u_t^{|T}, y_t^{|T}).
\end{aligned}$$

   Thus, we need to determine $i_{(\delta x_t^{|T}, \delta u_t^{|T}, \delta y_t^{|T})} \mathbf{d} \mathcal{G}_{\mathcal{P}(Q/G)}$. To do this, we will use Corollary 9.1.10.

   Let (x, u, y) be a point in $\mathcal{P}(Q/G)$. By definition, $\mathcal{G}_{\mathcal{P}(Q/G)}(x, u, y) = \langle y, dx \rangle$, so that $\mathbf{d}\mathcal{G}_{\mathcal{P}(Q/G)} = dy \wedge dx$. Suppose (w_x, w_u, w_y) and $(\tilde{w}_x, \tilde{w}_u, \tilde{w}_y)$ are tangent vectors to $\mathcal{P}(Q/G)$ at (x, u, y). Let $\gamma(t) = (x(t), u(t), y(t))$ be a curve in $\mathcal{P}(Q/G)$ such that $\gamma(0) = (x, u, y)$ and $\dot{\gamma}(0) = (w_x, w_u, w_y)$. Let $\gamma_\epsilon(t) = (x_\epsilon(t), u_\epsilon(t), y_\epsilon(t))$ be a deformation of $\gamma(t)$ with $\delta\gamma(0) = (\tilde{w}_x, \tilde{w}_u, \tilde{w}_y)$. By Corollary 9.1.10

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \langle \mathcal{G}_{\mathcal{P}(Q/G)}(\gamma_\epsilon(t)), \dot{\gamma}_\epsilon(t) \rangle = \langle i_{\delta\gamma(t)} \mathbf{d}\mathcal{G}_{\mathcal{P}(Q/G)}, \dot{\gamma}(t) \rangle + \frac{d}{dt} \langle \mathcal{G}_{\mathcal{P}(Q/G)}(\gamma(t)), \delta\gamma(t) \rangle .$$

which shows that,

$$\langle i_{\delta\gamma(t)}\mathbf{d}\mathcal{G}_{\mathcal{P}(Q/G)}, \dot{\gamma}(t)\rangle = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\langle \mathcal{G}_{\mathcal{P}(Q/G)}(\gamma_\epsilon(t)), \dot{\gamma}(t)\rangle - \frac{d}{dt}\langle \mathcal{G}_{\mathcal{P}(Q/G)}(\gamma(t)), \delta\gamma(t)\rangle$$

Expressing the right side in terms of coordinates yields,

$$\langle i_{\delta\gamma(t)}\mathbf{d}\mathcal{G}_{\mathcal{P}(Q/G)}, \dot{\gamma}(t)\rangle = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\langle y_\epsilon(t), \dot{x}_\epsilon(t)\rangle - \frac{d}{dt}\langle y(t), \delta x(t)\rangle.$$

In terms of the covariant derivatives on $T(Q/G)$ and $T^*(Q/G)$, we have

$$\begin{aligned}
&\frac{d}{d\epsilon}\Big|_{\epsilon=0}\langle y_\epsilon(t), \dot{x}_\epsilon(t)\rangle - \frac{d}{dt}\langle y(t), \delta x(t)\rangle \\
&= \left\langle \frac{D}{D\epsilon}\Big|_{\epsilon=0} y_\epsilon(t), \dot{x}(t)\right\rangle + \left\langle y(t), \frac{D}{D\epsilon}\Big|_{\epsilon=0}\dot{x}_\epsilon(t)\right\rangle \\
&- \left\langle \frac{D}{Dt}y(t), \delta x(t)\right\rangle - \left\langle y(t), \frac{D}{Dt}\delta x(t)\right\rangle
\end{aligned}$$

But by equality of mixed partial derivatives

$$\begin{aligned}
&\left\langle y(t), \frac{D}{D\epsilon}\dot{x}(t)\right\rangle - \left\langle y(t), \frac{D}{Dt}\frac{d}{d\epsilon}x(t)\right\rangle \\
&= \left\langle y_\epsilon(t), K_{\nabla^{T(Q/G)}}\frac{d^2x(t)}{d\epsilon dt}\right\rangle - \left\langle y(t), K_{\nabla^{T(Q/G)}}\frac{d^2x(t)}{dt d\epsilon}\right\rangle \\
&= 0,
\end{aligned}$$

which implies

$$\begin{aligned}
&\frac{d}{d\epsilon}\Big|_{\epsilon=0}\langle y(t), \dot{x}_\epsilon(t)\rangle - \frac{d}{dt}\langle y(t), \delta x(t)\rangle \\
&= \left\langle \frac{D}{D\epsilon}\Big|_{\epsilon=0} y_\epsilon(t), \dot{x}(t)\right\rangle - \left\langle \frac{D}{Dt}y(t), \delta x(t)\right\rangle \\
&= \left\langle K_{\nabla^{T^*(Q/G)}}\delta y(t), \dot{x}(t)\right\rangle - \left\langle K_{\nabla^{T^*(Q/G)}}\dot{y}(t), \delta x(t)\right\rangle.
\end{aligned}$$

Evaluating at $t = 0$ gives

$$\begin{aligned}
&i_{(\tilde{w}_x,\tilde{w}_u,\tilde{w}_y)}\mathbf{d}\mathcal{G}_{\mathcal{P}(Q/G)}(w_x, w_u, w_y) \\
&= \left(\frac{d}{d\epsilon}\Big|_{\epsilon=0}\langle y(t), \dot{x}_\epsilon(t)\rangle - \frac{d}{dt}\langle y(t), \delta x(t)\rangle\right)_{t=0} \\
&= \left(\langle K_{\nabla^{T^*(Q/G)}}\delta y(t), \dot{x}(t)\rangle - \langle K_{\nabla^{T^*(Q/G)}}\dot{y}(t), \delta x(t)\rangle\right)_{t=0} \\
&= \left\langle K_{\nabla^{T^*(Q/G)}}\tilde{w}_y, w_x\right\rangle - \left\langle K_{\nabla^{T^*(Q/G)}}w_y, \tilde{w}_x\right\rangle.
\end{aligned}$$

Since this holds for arbitrary tangent vectors (w_x, w_u, w_y) and $(\tilde{w}_x, \tilde{w}_u, \tilde{w}_y)$ at (x, u, y), we have

$$i_{(\delta x_t^{|T}, \delta u_t^{|T}, \delta y_t^{|T})}\mathbf{d}\mathcal{G}_{\mathcal{P}(Q/G)} \bullet d(x_t^{|T}, u_t^{|T}, y_t^{|T})$$

$$\begin{aligned}
&= \left\langle K_{\nabla T^*(Q/G)} \delta y_t^{|T}, \bullet dx_t^{|T} \right\rangle - \left\langle K_{\nabla T^*(Q/G)} \bullet dy_t^{|T}, \delta x_t^{|T} \right\rangle \\
&= \left\langle K_{\nabla T^*(Q/G)} \delta y_t^{|T}, \bullet dx_t^{|T} \right\rangle - \left\langle \bullet D y_t^{|T}, \delta x_t^{|T} \right\rangle .
\end{aligned}$$

As a result,

$$\mathcal{D} \int_0^T \langle y_t^{|T}, \bullet dx_t^{|T} \rangle = \int_0^T \left\langle K_{\nabla T^*(Q/G)} \delta y_t^{|T}, \bullet dx_t^{|T} \right\rangle - \int_0^T \left\langle \bullet D y_t^{|T}, \delta x_t^{|T} \right\rangle . \tag{10.2.15}$$

3. **The terms** $\mathcal{D} \int_0^T \langle y_t^{|T}, u_t^{|T} \rangle \bullet dX_t^0 + \sum_{i=1}^k \mathcal{D} \int_0^T \langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \rangle \bullet dX_t^i$

We use the connectors corresponding to the covariant derivatives on $T(Q/G)$ and $T^*(Q/G)$ to obtain

$$\begin{aligned}
&\mathcal{D} \int_0^T \langle y_t^{|T}, u_t^{|T} \rangle \bullet dX_t^0 \\
&= \int_0^T \left\langle K_{\nabla T^*(Q/G)} \delta y_t^{|T}, u_t^{|T} \right\rangle \bullet dX_t^0 + \int_0^T \left\langle y_t^{|T}, K_{\nabla T(Q/G)} \delta u_t^{|T} \right\rangle \bullet dX_t^0
\end{aligned}$$

To evaluate $\mathcal{D} \int_0^T \langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \rangle \bullet dX_t^i$, we apply Theorem 10.1.4 to the maps $(x, y) \in T^*(Q/G) \mapsto \left\langle y, V_i^{\mathrm{red}}(x) \right\rangle \in \mathbb{R}$. This gives,

$$\begin{aligned}
&\mathcal{D} \int_0^T \langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \rangle \bullet dX_t^i \\
&= \int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle \bullet dX_t^i \\
&+ \int_0^T \left\langle \frac{\partial}{\partial y_t^{|T}} \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \right\rangle, K_{\nabla T^*(Q/G)} \delta y_t^{|T} \right\rangle \bullet dX_t^i \\
&= \int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle \bullet dX_t^i \\
&+ \int_0^T \left\langle K_{\nabla T^*(Q/G)} \delta y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \right\rangle \bullet dX_t^i .
\end{aligned}$$

Therefore

$$\begin{aligned}
&\mathcal{D} \int_0^T \langle y_t^{|T}, u_t^{|T} \rangle \bullet dX_t^0 + \sum_{i=1}^k \mathcal{D} \int_0^T \langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \rangle \bullet dX_t^i \\
&= \int_0^T \left\langle K_{\nabla T^*(Q/G)} \delta y_t^{|T}, u_t^{|T} \bullet dX_t^0 + \sum_{i=1}^k V_i^{\mathrm{red}}(x_t^{|T}) \bullet dX_t^i \right\rangle \\
&+ \int_0^T \left\langle y_t^{|T}, K_{\nabla T(Q/G)} \delta u_t^{|T} \right\rangle \bullet dX_t^0 + \sum_{i=1}^k \int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle \bullet dX_t^i .
\end{aligned} \tag{10.2.16}$$

4. **The term** $\mathcal{D}\int_0^T \langle \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \rangle$

   Using the computations carried out in Section 10.2.3, we have

$$\begin{aligned}\mathcal{D}\int_0^T \langle \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \rangle &= \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \right\rangle - \int_0^T \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t^{|T}, \bullet d\bar{\xi}^{|T} \right\rangle \\ &\quad + \int_0^T \left\langle \bar{\mu}_t^{|T}, i_{\delta x_t^{|T}} \tilde{B}(x_t^{|T}) \right\rangle \bullet dx_t^{|T} - \int_0^T \left\langle \bullet D\bar{\mu}_t^{|T}, \bar{\eta}_t \right\rangle \\ &\quad + \left\langle \bar{\mu}_T^{|T}, \bar{\eta}_T \right\rangle - \left\langle \bar{\mu}_0^{|T}, \bar{\eta}_0 \right\rangle .\end{aligned}$$

   Since $\bar{\eta}$ vanishes at $t = 0$ and $t = T$, it follows that $\left\langle \bar{\mu}_T^{|T}, \bar{\eta}_T \right\rangle = \left\langle \bar{\mu}_0^{|T}, \bar{\eta}_0 \right\rangle = 0$. Hence,

$$\begin{aligned}\mathcal{D}\int_0^T \langle \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \rangle &= \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \right\rangle - \int_0^T \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t^{|T}, \bullet d\bar{\xi}^{|T} \right\rangle \\ &\quad + \int_0^T \left\langle \bar{\mu}_t^{|T}, i_{\delta x_t^{|T}} \tilde{B}(x_t^{|T}) \right\rangle \bullet dx_t^{|T} - \int_0^T \left\langle \bullet D\bar{\mu}_t^{|T}, \bar{\eta}_t \right\rangle \\ &= \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \right\rangle + \int_0^T \left\langle \mathrm{ad}^*_{\bullet d\bar{\xi}_t^{|T}} \bar{\mu}_t^{|T}, \bar{\eta}_t \right\rangle \\ &\quad - \int_0^T \left\langle \left\langle \bar{\mu}_t^{|T}, i_{\bullet dx_t^{|T}} \tilde{B}(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle - \int_0^T \left\langle \bullet D\bar{\mu}_t^{|T}, \bar{\eta}_t \right\rangle \\ &= \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \right\rangle - \int_0^T \left\langle \bullet D\bar{\mu}_t^{|T} - \mathrm{ad}^*_{\bullet d\bar{\xi}_t^{|T}} \bar{\mu}_t^{|T}, \bar{\eta}_t \right\rangle \\ &\quad - \int_0^T \left\langle \left\langle \bar{\mu}_t^{|T}, i_{\bullet dx_t^{|T}} \tilde{B}(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle , \end{aligned} \tag{10.2.17}$$

   where

$$\int \left\langle \mathrm{ad}^*_{\bullet d\bar{\xi}_t^{|T}} \bar{\mu}_t^{|T}, \bar{\eta}_t \right\rangle = - \int \left\langle \mathrm{ad}^*_{\bar{\eta}_t} \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \right\rangle$$

   and

$$\int \left\langle \left\langle \bar{\mu}_t^{|T}, i_{\bullet dx_t^{|T}} \tilde{B}(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle = - \int \left\langle \bar{\mu}_t^{|T}, i_{\delta x_t^{|T}} \tilde{B}(x_t^{|T}) \right\rangle \bullet dx_t^{|T}.$$

5. **The terms** $\mathcal{D}\int_0^T \left\langle \bar{\mu}_t^{|T}, \bar{\zeta}_t^{|T} \bullet dX_t^0 + \sum_{i=1}^k \bar{\beta}_i(x_t^{|T}) \bullet dX_t^i \right\rangle$

   We proceed by applying Equation (10.1.9) to yield

$$\begin{aligned}\mathcal{D}\int_0^T \left\langle \bar{\mu}_t^{|T}, \bar{\zeta}_t^{|T} \bullet dX_t^0 \right\rangle &= \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta\bar{\mu}_t^{|T}, \bar{\zeta}_t^{|T} \right\rangle \bullet dX_t^0 \\ &\quad + \int_0^T \left\langle \bar{\mu}_t^{|T}, K_{\nabla^{\tilde{\mathfrak{g}}}} \delta\bar{\zeta}_t^{|T} \right\rangle \bullet dX_t^0.\end{aligned}$$

   Next, by applying Theorem 10.1.4 to the maps $\bar{\mu}_x \in \tilde{\mathfrak{g}}^* \mapsto \langle \bar{\mu}_x, \bar{\beta}_i(x) \rangle$, we obtain

$$\mathcal{D}\int_0^T \left\langle \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle \bullet dX_t^i$$

$$\begin{aligned}
&= \int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left\langle \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle + \int_0^T \left\langle \frac{\partial}{\partial y_t^{|T}} \left\langle \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle, K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T} \right\rangle \\
&= \int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left\langle \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle + \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle .
\end{aligned}$$

Consequently,

$$\mathcal{D} \int_0^T \left\langle \bar{\mu}_t^{|T}, \bar{\zeta}_t^{|T} \bullet dX_t^0 + \sum_{i=1}^k \bar{\beta}_i(x_t^{|T}) \bullet dX_t^i \right\rangle \tag{10.2.18}$$

$$\begin{aligned}
&= \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T}, \bar{\zeta}_t^{|T} \right\rangle \bullet dX_t^0 \\
&+ \int_0^T \left\langle \bar{\mu}_t^{|T}, K_{\nabla^{\tilde{\mathfrak{g}}}} \delta \bar{\zeta}_t^{|T} \right\rangle \bullet dX_t^0 \\
&+ \left(\sum_{i=1}^k \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle \bullet dX_t^i \right. \\
&+ \left. \int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left\langle \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle \right) .
\end{aligned} \tag{10.2.19}$$

From Equations (10.2.14)-(10.2.18), we have

$$\begin{aligned}
&\mathcal{D}\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t) \\
&= \int_0^T \left(\left\langle \frac{\partial \ell}{\partial x_t^{|T}}, \delta x_t^{|T} \right\rangle + \left\langle \frac{\partial \ell}{\partial u_t^{|T}}, K_{\nabla^{T(Q/G)}} \delta u_t^{|T} \right\rangle \right. \\
&+ \left. \left\langle \frac{\partial \ell}{\partial \bar{\zeta}_t^{|T}}, K_{\nabla^{T(Q/G)}} \delta \bar{\zeta}_t^{|T} \right\rangle \right) \bullet dX_t^0 + \sum_{i=1}^k \int_0^T \left\langle \frac{\partial l_i}{\partial x_t^{|T}}, \delta x_t^{|T} \right\rangle \bullet dX_t^i \\
&+ \int_0^T \left\langle K_{\nabla^{T^*(Q/G)}} \delta y_t^{|T}, \bullet dx_t^{|T} \right\rangle - \int_0^T \left\langle \bullet D y_t^{|T}, \delta x_t^{|T} \right\rangle \\
&- \int_0^T \left\langle K_{\nabla^{T^*(Q/G)}} \delta y_t^{|T}, u_t^{|T} \bullet dX_t^0 + \sum_{i=1}^k V_i^{\mathrm{red}}(x_t^{|T}) \bullet dX_t^i \right\rangle \\
&- \int_0^T \left\langle y_t^{|T}, K_{\nabla^{T(Q/G)}} \delta u_t^{|T} \right\rangle \bullet dX_t^0 - \sum_{i=1}^k \int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle \bullet dX_t^i \\
&+ \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} \right\rangle - \int_0^T \left\langle \bullet D \bar{\mu}_t^{|T} - \mathrm{ad}^*_{\bullet d\bar{\xi}_t^{|T}} \bar{\mu}_t^{|T}, \bar{\eta}_t \right\rangle \\
&- \int_0^T \left\langle \left\langle \bar{\mu}_t^{|T}, i_{\bullet dx_t^{|T}} \tilde{B}(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle - \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T}, \bar{\zeta}_t^{|T} \right\rangle \bullet dX_t^0 \\
&- \int_0^T \left\langle \bar{\mu}_t^{|T}, K_{\nabla^{\tilde{\mathfrak{g}}}} \delta \bar{\zeta}_t^{|T} \right\rangle \bullet dX_t^0 - \left(\sum_{i=1}^k \int_0^T \left\langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle \bullet dX_t^i \right.
\end{aligned}$$

$$+\int_0^T \left\langle \frac{\partial}{\partial x_t^{|T}} \left\langle \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle, \delta x_t^{|T} \right\rangle \Bigg) .$$

Then the expression in (10.2.12) follows by appropriately grouping terms.

To prove the second part of the theorem, first note that if $(x_t^{|T}, u_t^{|T}, y_t^{|T}, \bar{\zeta}_t^{|T}, \bar{\mu}_t^{|T})$ satisfies the horizontal and vertical stochastic implicit Lagrange-Poincaré equations then

$$\mathcal{D}\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t) = 0,$$

so we only need to prove the converse. Take an arbitrary coordinate ball K in $\mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*$ and restrict to (K,T)-deformations of $(x_t^{|T}, u_t^{|T}, y_t^{|T}, \bar{\zeta}_t^{|T}, \bar{\mu}_t^{|T})$, and semimartingales $\bar{\eta}_t$ that vanish outside $]]\tau_K^h, (\tau_K^h + \tau_K^{(h,e)}) \wedge T[[$. Let τ_K^h denote the hitting time for K. Since $\bar{\eta}_t$ and the variations corresponding to (K,T)-deformations vanish outside $]]\tau_K^h, \tau_K^h + \tau_K^{(h,e)}[[$, the integral from 0 to T in the expression (10.2.12) for $\mathcal{D}\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t)$ can be replaced by an integral from τ_K^h to $\tau_K^h + \tau_K^{(h,e)}$, that is,

$$\begin{aligned}
&\mathcal{D}\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t) \\
&= \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} \left\langle \frac{\partial}{\partial x_t^{|T}} \left(\ell \bullet dX_t^0 + \sum_{i=1}^k \left(l_i - \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t^{|T}) \right\rangle - \left\langle \bar{\mu}_t^{|T}, \bar{\beta}_i(x_t^{|T}) \right\rangle \right) \bullet dX_t^i \right) \right. \\
&\left. - \left\langle \bar{\mu}_t, i_{\bullet dx_t} \tilde{B}(x_t) \right\rangle - \bullet D y_t^{|T}, \delta x_t^{|T} \right\rangle + \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} \left\langle \frac{\partial \ell}{\partial u_t^{|T}} - y_t^{|T}, K_{\nabla^{T(Q/G)}} \delta u_t^{|T} \right\rangle \bullet dX_t^0 \\
&+ \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} \left\langle \frac{\partial \ell}{\partial \bar{\zeta}_t^{|T}} - \bar{\mu}_t^{|T}, K_{\nabla^{\tilde{\mathfrak{g}}}} \delta \bar{\zeta}_t \right\rangle \bullet dX_t^0 \\
&+ \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} \langle K_{\nabla^{T^*(Q/G)}} \delta y_t^{|T}, \bullet dx_t^{|T} - u_t^{|T} \bullet dX_t^0 - \sum_{i=1}^k V_i^{\mathrm{red}}(x_t^{|T}) \bullet dX_t^i \rangle \\
&+ \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} \langle K_{\nabla^{\tilde{\mathfrak{g}}^*}} \delta \bar{\mu}_t^{|T}, \bullet d\bar{\xi}_t^{|T} - \bar{\zeta}_t^{|T} \bullet dX_t^0 - \sum_{i=1}^k \bar{\beta}_i(x_t^{|T}) \bullet dX_t^i \rangle \\
&+ \int_{\tau_K^h}^{\tau_K^h + \tau_K^{(h,e)}} \langle - \bullet D\bar{\mu}_t^{|T} + \mathrm{ad}^*_{\bullet d\bar{\xi}_t^{|T}} \bar{\mu}_t^{|T}, \bar{\eta}_t \rangle
\end{aligned}$$

By the stochastic version of the fundamental lemma of the calculus of variations, Lemma 9.1.13, this implies that the semimartingale $(x_t^{|T}, u_t^{|T}, y_t^{|T}, \bar{\zeta}_t^{|T}, \bar{\mu}_t^{|T})$ satisfies the horizontal and vertical stochastic Lagrange-Poincaré equations in $]]\tau_K^h, \tau_K^h + \tau_K^{(h,e)}[[$. Since K is an arbitrary regular coordinate ball in $\mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*$, it follows that $(x_t^{|T}, u_t^{|T}, y_t^{|T}, \bar{\zeta}_t^{|T}, \bar{\mu}_t^{|T})$ satisfies the horizontal and vertical Lagrange-Poincaré equations in $\mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*$. This completes the proof. ∎

We now summarize what we have obtained so far in the next theorem. This serves as the stochastic analogue of the deterministic implicit Lagrange-Poincaré reduction theorem (Theorem 2.5.5):

Theorem 10.2.7 (Stochastic Implicit Lagrange-Poincaré Reduction Theorem)**.** The following are equivalent:

1. An admissible semimartingale $\Gamma_t = (q_t, v_t, p_t)$ in $\mathcal{P}Q$ is a critical point of the stochastic action integral
$$\begin{aligned}\mathcal{S}_X(\Gamma) = \int_0^T \Bigg(& \mathcal{L}(q_t, v_t) \bullet dX_t^0 + \sum_{i=1}^k L_i(q_t) \bullet dX_t^i \\ & + \left\langle p_t, \bullet dq_t - v_t \bullet dX_t^0 - \sum_{i=1}^k V_i(q_t) \bullet dX_t^i \right\rangle \Bigg) . \end{aligned} \tag{10.2.20}$$
for all deformations $\epsilon \mapsto \Gamma_\epsilon$ such that δq_t vanishes at $t = 0$ and $t = T$.

2. The stochastic implicit Euler-Lagrange equations
$$\begin{aligned}\bullet dq_t &= v_t \bullet dX_t^0 + \sum_{i=1}^k V_i(q_t) \bullet dX_t^i \\ \bullet dp_t &= \frac{\partial}{\partial q_t}\left(\mathcal{L} \bullet dX_t^0 + \sum_{i=1}^k \left(L_i - \langle p_t, V_i(q_t)\rangle\right) \bullet dX_t^i \right) \\ \left(p_t - \frac{\partial \mathcal{L}}{\partial v_t}\right) & \bullet dX_t^0 = 0. \end{aligned} \tag{10.2.21}$$
are satisfied by $\Gamma_t^{|T} = (q_t^{|T}, v_t^{|T}, p_t^{|T})$.

3. The reduced semimartingale
$$[q_t, v_t, p_t]_G \cong (x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t)$$
in the reduced Pontryagin bundle $\mathcal{P}Q/G \cong \mathcal{P}(Q/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*$ is a critical point of the reduced action integral
$$\begin{aligned}&\mathcal{S}_X^{\mathrm{red}}(x_t, u_t, y_t, \bar{\zeta}_t, \bar{\mu}_t) \\ &= \int \ell(x_t, u_t, \bar{\zeta}_t) \bullet dX_t^0 + \sum_{i=1}^k l_i(x_t) \bullet dX_t^i + \left\langle y_t, \bullet dx_t - u_t \bullet dX_t^0 - \sum_{i=1}^k V_i^{\mathrm{red}}(x_t) \bullet dX_t^i \right\rangle \\ &+ \left\langle \bar{\mu}_t, \bullet d\bar{\xi}_t - \bar{\zeta}_t \bullet dX_t^0 - \sum_{i=1}^k \bar{\beta}_i(x_t) \bullet dX_t^i \right\rangle \end{aligned} \tag{10.2.22}$$
for variations arbitrary variations $\delta u_t, \delta y_t, \delta\bar{\zeta}_t$ and $\delta\bar{\mu}_t$ and for variations $\delta x_t^{|T} \oplus \delta^A(\bar{\xi}_t^{|T})$ such that $\delta x^{|T} = 0$ at $t = 0$ and $t = T$ and
$$\bullet d(\delta^A \bar{\xi}_t^{|T}) = -\mathrm{ad}_{\bar{\eta}_t} \bullet d\bar{\xi}_t^{|T} + \bullet D\bar{\eta}_t + i_{\delta x_t^{|T}} \tilde{B}(x_t^{|T}) \bullet dx_t^{|T}$$
where $\bar{\eta}$ is an arbitrary semimartingale in $\tilde{\mathfrak{g}}$ that vanishes at $t = 0$ and $t = T$.

4. The horizontal stochastic implicit Lagrange-Poincaré equations

$$\begin{aligned}\bullet Dy_t &= \frac{\partial}{\partial x_t}\left(\ell \bullet dX_t^0 + \sum_{i=1}^{k}\left(l_i - \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t)\right\rangle - \left\langle \bar{\mu}_t, \bar{\beta}_i(x_t)\right\rangle\right) \bullet dX_t^i\right)\\ &\quad - \left\langle \bar{\mu}_t, i_{\bullet dx_t}\tilde{B}(x_t)\right\rangle\\ \bullet dx_t &= u_t \bullet dX_t^0 + \sum_{i=1}^{k} V_i^{\mathrm{red}}(x_t) \bullet dX_t^i\\ \left(y_t - \frac{\partial \ell}{\partial u_t}\right) &\bullet dX_t^0 = 0\end{aligned}$$

and the vertical stochastic implicit Lagrange-Poincaré equations

$$\begin{aligned}\bullet D\bar{\mu}_t &= \mathrm{ad}^*_{\bullet d\bar{\xi}_t}\bar{\mu}_t\\ \bullet d\bar{\xi}_t &= \bar{\zeta}_t \bullet dX_t^0 + \sum_{i=1}^{k}\bar{\beta}_i(x_t) \bullet dX_t^i\\ \left(\bar{\mu}_t - \frac{\partial \ell}{\partial \bar{\zeta}_t}\right) &\bullet dX_t^0 = 0\end{aligned}$$

are satisfied by $(x_t^{|T}, u_t^{|T}, y_t^{|T}, \bar{\zeta}_t^{|T}, \bar{\mu}_t^{|T})$.

Remark 10.2.8. In the deterministic case, that is, $X^0 = t$, $X^i = 0$, all semimartingales may be replaced by smooth curves. We also replace $\bullet d\bar{\xi}_t$ by $[q(t), A(q(t), \dot{q}(t))]_G dt$. Under these replacements, the stochastic implicit horizontal and vertical Lagrange-Poincaré equations agree with their deterministic counterparts.

Coordinate expressions

Suppose Q is an n-manifold and Q/G has dimension r. We choose a trivialization of Q as $U \times G$, where $U \subseteq \mathbb{R}^r$ is an open set, and G acts only on the second factor by left multiplication. Let $(x, g) = (x^\alpha, g^a)$ be an element of $U \times G$. The principal connection A acts on a tangent vector $(\dot{x}, \dot{g})$ by

$$A(\dot{x}, \dot{g}) = \mathrm{Ad}_g(A_{(x,e)}(\dot{x}, g^{-1}\dot{g})) = \mathrm{Ad}_g(A_e(x)\dot{x} + g^{-1}\dot{g}) = \mathrm{Ad}_g(A_e(x)\dot{x}) + \dot{g}g^{-1},$$

where $A_e(x)\dot{x} = A_{(x,e)}(\dot{x}, 0)$. Hence, $A(x, g) = \mathrm{Ad}_g(A_e(x)dx + g^{-1}dg)$. Let $\xi = g^{-1}\dot{g}$. Then

$$\bar{\xi} := [(x, g), A(x, g, \dot{x}, \dot{g})]_G = [(x, e), A_e(x)\dot{x} + \xi]_G.$$

Trivializing the adjoint bundle as $(U \times G \times \mathfrak{g})/G \cong U \times \mathfrak{g}$ identifies $\bar{\xi}$ with $(x, A_e(x)\dot{x} + \xi)$. Let us write $A_e(x)$ locally as $A^a_\alpha(x)$. Then, with identify $\bar{\xi}$ with its second component and simply write $\bar{\xi}^a = \xi^a + A^a_\alpha \dot{x}^\alpha$. If we now use Stratonovich differentials, then this reads $\bullet d\bar{\xi}^a = \bullet d\xi^a + A^a_\alpha \bullet dx^\alpha$, where $\bullet d\bar{\xi}^a$, as a Stratonovich differential in $(U \times G \times \mathfrak{g})/G \cong U \times \mathfrak{g}$ is given by $[(x_t, e), A_e(x_t) \bullet dx_t + \bullet d\xi_t]$ and $\bullet d\xi_t = g_t^{-1} \bullet dg_t$.

Let C^b_{cd} denote the structure constants of the Lie algebra. Then the $\mathfrak{g}$-valued curvature 2-form B is given locally by

$$B^b_{\alpha\beta} = \frac{\partial A^b_\beta}{\partial x^\alpha} - \frac{\partial A^b_\alpha}{\partial x^\beta} - C^b_{cd} A^c_\alpha A^d_\beta.$$

Let the G-invariant vector fields V_i be written locally as $V_i(x,g) = V_i^\alpha(x)\frac{\partial}{\partial x^\alpha} + g\beta^a(x)$, where $\beta^a : U \to \mathfrak{g}$ is a smooth map. Then the reduced vector fields V_i^{red} on U given locally by $V_i^{\mathrm{red}}(x) = V_i^\alpha(x)\frac{\partial}{\partial x^\alpha}$. The sections $\bar{\beta}_i$ of the associated bundle are given by

$$\bar{\beta}_i(x) = [(x,e), A(V_i(x,e))]_G.$$

We think of these as maps from $U \to \mathfrak{g}$ given by $\bar{\beta}^a_i(x) = A^a_\alpha(x) V_i^\alpha(x) + \beta^a_i(x)$.

Following the local coordinate calculations done in Cendra, Marsden and Ratiu [19], the local form of the horizontal stochastic implicit Lagrange-Poincaré equations is given by

$$\begin{aligned}
\bullet dy_{\alpha_t} &= \frac{\partial}{\partial x^\alpha_t}\left(\ell \bullet dX^0_t + \sum_{i=1}^k l_i \bullet dX^i_t\right) - \sum_{i=1}^k \left(y_{\beta_t}\frac{\partial V_i^\beta(x_t)}{\partial x^\alpha_t} + \right.\\
&\quad \left. + \bar{\mu}_a\left(C^a_{bd}\bar{\beta}^b_i(x_t)A^d_\alpha + \frac{\partial \beta^a_i(x_t)}{\partial x^\alpha_t}\right)\right) \bullet dX^i_t \\
&\quad + \bar{\mu}_a B^a_{\alpha\beta} \bullet dx^\beta_t - \frac{\partial \ell}{\partial \bar{\zeta}^a} C^a_{db} A^b_\alpha \bullet d\bar{\xi}^d_t \\
\bullet dx^\alpha_t &= u^\alpha_t \bullet dX^0_t + \sum_{i=1}^k V_i^\alpha(x_t) \bullet dX^i_t \\
\left(y_{\alpha_t} - \frac{\partial \ell}{\partial u^\alpha_t}\right) &\bullet dX^0_t = 0
\end{aligned}$$

and the stochastic vertical implicit Lagrange-Poincaré equations are given by

$$\begin{aligned}
\bullet d\bar{\mu}_{b_t} &= \bar{\mu}_{a_t} C^a_{db}(\bullet d\bar{\xi}^d_t - A^d_\alpha \bullet dx^\alpha_t) \\
\bullet d\bar{\xi}^d_t &= \bar{\zeta}^d_t \bullet dX^0_t + \sum_{i=1}^k \bar{\beta}^d_i(x_t) \bullet dX^i_t \\
\left(\bar{\mu}_{b_t} - \frac{\partial \ell}{\partial \bar{\zeta}^b_t}\right) &\bullet dX^0_t = 0.
\end{aligned}$$

Remark 10.2.9. If $X^0_t = t$ and $X^i_t = 0$ then $\bar{\mu} = \frac{\partial \ell}{\partial \bar{\zeta}}$ and $\bullet d\bar{\xi}$ is replaced by the deterministic $\mathfrak{g}$-valued curve $\bar{\xi} = [q, A(q,\dot{q})]_G$. In this case, the local form of the reduced equations agree with the local form of the implicit Lagrange-Poincaré equations given in Yoshimura and Marsden [95].

If the bundle $Q \to Q/G$ is equipped with a trivial connection in local coordinates, that is, $A = 0$, then the horizontal equations in local coordinates are given by

$$\bullet dy_{\alpha_t} = \frac{\partial}{\partial x^\alpha_t}\left(\ell \bullet dX^0_t + \sum_{i=1}^k l_i \bullet dX^i_t\right) - \sum_{i=1}^k \left(y_{\beta_t}\frac{\partial V_i^\beta(x_t)}{\partial x^\alpha_t} + \right.$$

$$
\begin{aligned}
&\qquad + \bar{\mu}_a \frac{\partial \beta_i^a(x_t)}{\partial x_t^\alpha}\Bigg) \bullet dX_t^i \\
&\bullet dx_t^\alpha = u_t^\alpha \bullet dX_t^0 + \sum_{i=1}^k V_i^\alpha(x_t) \bullet dX_t^i \\
&\left(y_{\alpha_t} - \frac{\partial \ell}{\partial x_t^\alpha}\right) \bullet dX_t^0 = 0
\end{aligned}
$$

and the vertical equations are given by

$$
\begin{aligned}
\bullet d\bar{\mu}_{b_t} &= \bar{\mu}_{a_t} C_{db}^a (\bullet d\bar{\xi}_t^d) \\
\bullet d\bar{\xi}_t^d &= \bar{\zeta}_t^d \bullet dX_t^0 + \sum_{i=1}^k \beta_i^d(x_t) \bullet dX_t^i \\
\left(\bar{\mu}_{b_t} - \frac{\partial \ell}{\partial \bar{\zeta}_t^b}\right) &\bullet dX_t^0 = 0.
\end{aligned}
$$

We will call these equations the **stochastic Hamel's equations**. When $X_t^0 = t$ and $X_t^i = 0$ then these equations correspond to Hamel's equations described in Marsden and Scheurle [64].

Special Cases

We now discuss four special cases.

1. Suppose $Q = G$. In this case, Q/G is a point and hence the horizontal implicit stochastic Lagrange-Poincaré equations vanish. The G-invariant vector fields V_i are of the form $V_i(g) = T_e L_g \beta_i$, where $\beta_i \in \mathfrak{g}$ is a fixed element, and $g \in G$. We also have $\tilde{\mathfrak{g}} \cong \mathfrak{g}$ and $\tilde{\mathfrak{g}}^* \cong \mathfrak{g}^*$. Then, the vertical implicit stochastic Lagrange-Poincaré equations are given by

$$
\begin{aligned}
\bullet d\mu_t &= \mathrm{ad}^*_{\bullet d\xi_t} \mu_t \\
\bullet d\xi_t &= \zeta_t \bullet dX_t^0 + \sum_{i=1}^k \beta_i \bullet dX_t^i \\
\left(\mu_t - \frac{\partial \ell}{\partial \zeta_t}\right) &\bullet dX_t^0 = 0.
\end{aligned}
$$

When $X_t^0 = t$ and X^i is a Brownian motion, these equations agree with the stochastic Euler-Poincaré equations given in Street and Takao [89].

Remark 10.2.10. In Arnaudon, Chen, and Cruzeiro [7], the authors carry out stochastic Euler-Poincaré reduction with the generalized derivative of a semimartingale playing the role of the velocity. The action is given by $\mathbb{E}\left[\int_0^T \ell\left(T_{g_t} L_{g_t^{-1}} \frac{D^\nabla g_t}{Dt}\right) dt\right]$ and computed on semimartingales of the form

$$
\bullet dg_t = T_e L_{g_t}\left[u(t)dt - \sum_{i=1}^k \left(\nabla_{H_i} H_i \, dt - H_i \bullet dB_t^i\right)\right],
$$

where $u(t)$ is a curve in $\mathfrak{g}$ and $H_1, \cdots, H_k \in \mathfrak{g}$. The reduced equations of motion in that case are deterministic, and given by

$$\frac{d}{dt}\left(\frac{\delta \ell}{\delta u}\right) - \mathrm{ad}^*_{\tilde{u}}\frac{\delta \ell}{\delta u} = K\left(\frac{\delta \ell}{\delta u}\right),$$

where $\tilde{u} = u - \frac{1}{2}\sum_{i=1}^k \nabla_{H_i} H_i$ and $K : \mathfrak{g}^* \to \mathfrak{g}^*$ is given by

$$\langle K(\mu), \eta \rangle = -\frac{1}{2}\sum_{i=1}^k \left\langle \mu, \nabla_{[\eta, H_i]} H_i \right\rangle + \nabla_{H_i}[\eta, H_i]$$

for all $\eta \in \mathfrak{g}$. This determines the curve $u(t)$, which corresponds to the generalized derivative of g_t. In contrast, in our case the reduced velocity is stochastic and is determined via the implicit form of the stochastic Euler-Poincaré equations given above.

Additionally, if the connection on G is the canonical bi-invariant connection, then the Euler-Poincaré equations obtained in [7] agree with the classical Euler-Poincaré equation

$$\frac{d}{dt}\left(\frac{\delta \ell}{\delta u}\right) - \mathrm{ad}^*_{u}\frac{\delta \ell}{\delta u} = 0.$$

Then, for every $H_1, \cdots, H_k \in \mathfrak{g}$,

$$\bullet dg_t = T_e L_{g_t}\left[u(t)dt + \sum_{i=1}^k H_i \bullet dB_t^i\right]$$

is a critical point of the unreduced action. The critical point of unreduced action is unique for a given choice of $H_1, \cdots, H_k$ is made. The necessity of this choice arises from the fact that the generalized derivative of g_t determines its drift, but not the complete semimartingale. On the other hand, in our case the action is given by stochastic integrals, and the uniqueness of critical points is given by the fundamental lemma of the stochastic calculus of variations.

2. Assume that G is an abelian group. From Equation (2.5.2), the covariant derivative in the associated bundle agrees with the fibre derivative since the Lie bracket vanishes. Then, by Equation (10.1.8), it follows that the covariant derivative on the dual bundle agrees with the fibre derivative as well. Also, from the stochastic vertical implicit Lagrange-Poincaré equations, we see that $\bullet D\bar{\mu}_t = 0$. Writing $\bar{\mu}_t = [q_t, \mu_t]_G$, we see that μ_t is conserved along solutions.

   As a special case, note that if $G = \{e\}$ then the vertical stochastic implicit Lagrange-Poincaré equations vanish and the horizontal stochastic implicit Lagrange-Poincaré equations agree with the stochastic implicit Euler-Lagrange equations.

3. We now assume that the noise vector fields V_i are horizontal. In this case, $\bar{\beta}_i = 0$, so the stochastic horizontal and vertical implicit Lagrange-Poincaré equations are given by

$$\begin{aligned}
\bullet D y_t &= \frac{\partial}{\partial x_t}\left(\ell \bullet dX_t^0 + \sum_{i=1}^k \left(l_i - \left\langle y_t^{|T}, V_i^{\mathrm{red}}(x_t)\right\rangle\right) \bullet dX_t^i\right) \\
&\quad - \left\langle \bar{\mu}_t, i_{\bullet dx_t}\tilde{B}(x_t)\right\rangle \\
\bullet dx_t &= u_t \bullet dX_t^0 + \sum_{i=1}^k V_i^{\mathrm{red}}(x_t) \bullet dX_t^i \\
\left(y_t - \frac{\partial \ell}{\partial u_t}\right) \bullet dX_t^0 &= 0
\end{aligned}$$

and

$$\begin{aligned}
\bullet D\bar{\mu}_t &= \mathrm{ad}^*_{\bullet d\bar{\xi}_t}\bar{\mu}_t \\
\bullet d\bar{\xi}_t &= \bar{\zeta}_t \bullet dX_t^0 \\
\left(\bar{\mu}_t - \frac{\partial \ell}{\partial \bar{\zeta}_t}\right) \bullet dX_t^0 &= 0
\end{aligned}$$

respectively. As a result, if $X_t^0 = t$ then the vertical equations are noise-free.

4. Finally, suppose that $\beta_1, \cdots, \beta_k$ are Ad_G-invariant elements of $\mathfrak{g}$, and let $V_i = (\beta_i)_Q$. Then V_i is G-invariant, $V_i^{\mathrm{red}} = 0$ and $A(V_i) = \beta_i$. We also have $\bar{\beta}_i(x) = [q, \beta_i]_G$, where $x \in Q/G$ and $q \in \pi^{-1}(x)$. Note that this is independent of the choice of the representative q since β_i is Ad_G-invariant. As a result $\frac{\partial}{\partial x_t}\left\langle \bar{\mu}_t, \bar{\beta}_i(x_t)\right\rangle$ vanishes. We also assume that $l_1, \cdots, l_k = 0$ and $X_t^0 = t$. Then the horizontal equations are given by

$$\begin{aligned}
\frac{Dy}{Dt} &= \frac{\partial \ell}{\partial x} - \left\langle \bar{\mu}_t, i_{\dot{x}}\tilde{B}(x)\right\rangle \\
\dot{x} &= u \\
y &= \frac{\partial \ell}{\partial u}
\end{aligned}$$

and the vertical equations are given by

$$\begin{aligned}
\bullet D\bar{\mu}_t &= \mathrm{ad}^*_{\bullet d\bar{\xi}_t}\bar{\mu}_t \\
\bullet d\bar{\xi}_t &= \bar{\zeta}_t dt + \sum_{i=1}^k \bar{\beta}_i \bullet dX_t^i \\
\bar{\mu}_t &= \frac{\partial \ell}{\partial \bar{\zeta}_t}.
\end{aligned}$$

This shows that we have a stochastic curvature-induced force in the horizontal equations.

10.3 Examples

10.3.1 Rigid Body with a Rotor

Let us look at the example of a rigid body with a rotor aligned with its third principal axis. We will consider external stochastic forces on the rotor and the body, but for simplicity, we will assume that these forces do not affect the moments of inertia of the rotor and the body.

The Deterministic Free Rigid Body with a Rotor

Following Marsden [60] and Yoshimura and Marsden [95], we briefly recollect the setup of the deterministic problem. Let $I_1 > I_2 > I_3$ denote the principal rigid body moments of inertia, $K_1 = K_2$ denote the rotor moments of inertia about the first two principal axes of the rigid body and K_3 denote the moment of inertia of the rigid body about the third principal axis. Let K be the diagonal matrix $K = \operatorname{diag}(K_1, K_2, K_3)$. The configuration space is $Q = S^1 \times SO(3)$, where $G = SO(3)$ acts on the second factor by matrix multiplication. Then $Q/G = S^1$. The Lagrangian of this system is given by

$$\mathcal{L}(\theta, R, v_\theta, v_R) = \frac{1}{2}\left[\langle \Sigma, I\Sigma\rangle + \langle \Sigma + \mathbf{v}_\theta, K(\Sigma + \mathbf{v}_\theta)\rangle\right], \tag{10.3.1}$$

where $\mathbf{v}_\theta = (0, 0, v_\theta)^T \in \mathbb{R}^3$, $\Sigma \in \mathbb{R}^3$ is defined by $\hat{\Sigma} = R^{-1}v_R \in \mathfrak{so}(3)$ with the usual 'hat' map identification of $\mathbb{R}^3$ with $\mathfrak{so}(3)$ given by $\hat{\Sigma}w = \Sigma \times w$, for any $w \in \mathbb{R}^3$.

We let A be a trivializing connection on the bundle $Q \to Q/G$, so that $TQ/G \cong TS^1 \times \mathfrak{so}(3)$. Concretely, $A(\theta, R, v_\theta, v_R) = R^{-1}v_R \in \mathfrak{so}(3)$. Then, in terms the coordinates (θ, u, Σ) on $TS^1 \times \mathfrak{so}(3)$, the reduced Lagrangian is given by

$$\ell(\theta, u, \Sigma) = \frac{1}{2}\left[\lambda_1\Sigma_1^2 + \lambda_2\Sigma_2^2 + I_3\Sigma_3^2 + K_3(\Sigma_3 + u)^2\right], \tag{10.3.2}$$

with $\lambda_i = I_i + K_i$. The (deterministic) implicit horizontal Lagrange-Poincaré equations are given by

$$\dot{y} = 0,\ \ \dot{\theta} = u,\ \ y = \frac{\partial \ell}{\partial u} = K_3(\Sigma_3 + u)$$

and the vertical implicit Lagrange-Poincaré equations are given by

$$\dot{\Pi} = \Pi \times \Omega,\ \ \Omega = \Sigma,\ \ \Pi = \frac{\partial \ell}{\partial \Sigma} = (\lambda_1\Sigma_1, \lambda_2\Sigma_2, I_3\Sigma_3 + K_3(\Sigma_3 + u)).$$

Note that the conjugate momentum $y = K_3(\Sigma_3 + u)$ is conserved.

Stochastic Perturbations of a Free Rigid Body

Before considering a stochastic pertubation of the deterministic problem, we recall from Lázaro-Camí and Ortega [54] and Arnaudon, De Castro and Holm [4], the Stratonovich equations of motion describing a free rigid body under small random impacts. Let $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ be elements in the Lie algebra $\mathfrak{so}(3)$ and X_t^1, X_t^2 and X_t^3 be semimartingales. The body angular momentum of the stochastic free rigid body is given by

$$\bullet d\Pi_t = (\Pi_t \times I^{-1}\Pi_t)dt + \sum_{i=1}^{3}(\Pi_t \times \beta_i) \bullet dX_t^i, \tag{10.3.3}$$

and the attitude is given by

$$\bullet dR_t = R_t\widehat{I^{-1}\Pi_t}dt + \sum_{i=1}^{3} R_t\hat{\beta}_i \bullet dX_t^i.$$

Note that Equation (10.3.3) is equivalent to

$$\begin{aligned} \bullet d\Pi_t &= \Pi_t \times \bullet d\Omega_t \\ \bullet d\Omega_t &= \Sigma_t dt + \sum_{i=1}^{3} \beta_i \bullet dX_t^i \\ \Pi_t &= I\Sigma_t. \end{aligned}$$

These equations can be obtained by stochastic Euler-Poincaré reduction for the action

$$\mathcal{S}(R_t, v_{R_t}, p_{R_t}) = \int_0^T \left[\frac{1}{2}\left\langle \Sigma_t, I\Sigma_t\right\rangle dt + \left\langle p_t, \bullet dR_t - v_{R_t}dt - \sum_{i=1}^{3} R_t\hat{\beta}_i \bullet dX_t^i\right\rangle\right],$$

where $\hat{\Sigma}_t = R_t^{-1}v_{R_t}$. Comparing this to the general structure of the stochastic Hamilton-Pontryagin action, we note that $\mathcal{L} = \frac{1}{2}\left\langle \Sigma, I\Sigma\right\rangle$, $X_t^0 = t$, the noise Lagrangian l_i are zero, and the noise vector fields are the left invariant vector fields corresponding to $\hat{\beta}_i$.

Stochastic Perturbations of a Rigid Body with a Rotor

Let $\mathcal{L}$ be the Lagrangian given in Equation (10.3.1) and X_t^1, X_t^2 and X_t^3 be semimartingales. Suppose $\hat{\beta}_1, \hat{\beta}_2$ and $\hat{\beta}_3$ are fixed elements of $\mathfrak{so}(3)$. On $Q = S^1 \times SO(3)$ we consider the action

$$\begin{aligned} &\mathcal{S}(\theta_t, R_t, v_{\theta_t}, v_{R_t}, p_{\theta_t}, p_{R_t}) \\ &= \int_0^T \left[\mathcal{L}\, dt + \sum_{i=1}^{3} L_i(\theta_t) \bullet dX_t^i + \left\langle p_{\theta_t}, \bullet d\theta_t - v_{\theta_t}dt - \sum_{i=1}^{3} V_i(\theta_t) \bullet dX_t^i\right\rangle\right. \\ &\left. + \left\langle p_{R_t}, \bullet dR_t - v_{R_t}dt - \sum_{i=1}^{3} R_t\hat{\beta}_i \bullet dX_t^i\right\rangle, \right] \end{aligned}$$

where L_1, L_2, L_3 are smooth functions and V_i is a vector field on S^1.

Remark 10.3.1. The Lagrangians L_1, L_2, L_3 are interpreted as the contribution of the external force to the potential energy of the rotor, and the vector fields V_1, V_2, V_3 stochasticize the relation $\dot{\theta} = v_\theta$ to

$$\bullet d\theta_t = v_{\theta_t} dt + \sum_{i=1}^{3} V_i(\theta_t) \bullet dX_t^i.$$

The term

$$\int_0^T \left\langle p_{R_t}, \bullet dR_t - v_{R_t} dt - \sum_{i=1}^{3} R_t \hat{\beta}_i \bullet dX_t^i \right\rangle$$

accounts for random perturbation of the free rigid body.

We consider a trivializing connection A on $S^1 \times SO(3) \to S^1$. Then, the reduced action is given by

$$\begin{aligned} \mathcal{S}^{\mathrm{red}}(\theta_t, u_t, \Sigma_t, y_t, \Pi_t) = \int_0^T \Bigg[\ell\, dt &+ \sum_{i=1}^{3} L_i(\theta_t) \bullet dX_t^i + \left\langle y_t, \bullet d\theta_t - u_t dt - \sum_{i=1}^{3} V_i(\theta_t) \bullet dX_t^i \right\rangle \\ &+ \left\langle \Pi_t, \bullet d\Omega - \Sigma_t dt - \sum_{i=1}^{3} \hat{\beta}_i \bullet dX_t^i \right\rangle \Bigg], \end{aligned}$$

where ℓ is the reduced Lagrangian in Equation (10.3.2) and $\bullet d\Omega_t = R_t^{-1} \bullet dR_t$. It follows that the stochastic implicit horizontal Lagrange-Poincaré equations are given by

$$\begin{aligned} \bullet dy_t &= \frac{\partial}{\partial \theta_t} \sum_{i=1}^{3} (L_i(\theta_t) - y_t V_i(\theta_t)) \bullet dX_t^i \\ \bullet d\theta_t &= u_t dt + \sum_{i=1}^{3} V_i(\theta_t) \bullet dX_t^i \\ y_t &= \frac{\partial \ell}{\partial u_t} = K_3(\Sigma_{3_t} + u_t), \end{aligned}$$

and the stochastic implicit vertical Lagrange-Poincaré equations are given by

$$\begin{aligned} \bullet d\Pi_t &= \Pi_t \times \bullet d\Omega_t \\ \bullet d\Omega_t &= \Sigma_t dt + \sum_{i=1}^{3} \beta_i \bullet dX_t^i \\ \Pi_t &= \frac{\partial \ell}{\partial \Sigma_t} = (\lambda_1 \Sigma_{1_t}, \lambda_2 \Sigma_{2_t}, I_3 \Sigma_{3_t} + K_3(\Sigma_{3_t} + u_t)). \end{aligned}$$

We observe that the conjugate momentum $y = K_3(\Sigma_3 + u)$ is no longer conserved in the stochastic case.

Let us mention three special cases.

1. $L_i = 0$ and $V_i = 0$: We consider $L_i = 0$ and $V_i = 0$. In this case, the horizontal equations are given by

$$\begin{aligned} \dot{y}_t &= 0 \\ \dot{\theta}_t &= u_t \\ y_t &= \frac{\partial \ell}{\partial u_t} = K_3(\Sigma_{3_t} + u_t). \end{aligned}$$

   This implies the conjugate momentum $y = K_3(\Sigma_3 + u)$ is conserved, and the rotor angle of rotation evolves along differentiable trajectories. Physically, we interpret this as the case where the external noise only impacts the rigid body and not the rotor.

2. $\beta_i = 0$: In this case, the vertical equations are same as the vertical implicit Lagrange-Poincaré equations in the deterministic case. Physically, this models a system where the external noise only impacts the rotor and not the rigid body.

3. $V_i = 0$, $L_i(\theta) = \theta$ and X_t^i is a Brownian motion: In this case, the horizontal equations are given by

$$\begin{aligned} \bullet dy_t &= \sum_{i=1}^{3} \bullet dB_t^i \\ \dot{\theta}_t &= u_t \\ y_t &= \frac{\partial \ell}{\partial u_t} = K_3(\Sigma_{3_t} + u_t), \end{aligned}$$

   where B^1, B^2 and B^3 are independent Brownian motions. This implies that $y_t = B_t^1 + B_t^2 + B_t^3$. Hence $\mathbb{E}[y_t] = 0$ for all t. This shows that y_t is a weakly conserved quantity in the sense of Lázaro-Camí and Ortega [55, Definition 2.2], but not a conserved quantity.

10.3.2 Charged Particle in a Magnetic Field with Stochastic Perturbations

In this section we provide a Kaluza-Klein description of a stochastically perturbed charged particle in a magnetic field.

The Deterministic Kaluza-Klein Description for a Charged Particle in a Magnetic Field

The equation of motion for a charged particle (of unit mass) in a magnetic field $\mathbf{B}$ is given by $\dot{\mathbf{u}} = \frac{e}{c}\mathbf{u} \times \mathbf{B}$. Following Marsden and Ratiu [62], we show that this can be viewed as a reduction of the geodesic flow on $Q_K = \mathbb{R}^3 \times S^1$ under a certain metric. Here S^1 acts on the second factor by rotations.

Let $G = S^1$ with its standard bi-invariant metric κ and consider $\mathbb{R}^3$ with its standard metric given by the inner product $\langle \cdot, \cdot \rangle$. Let $\mathbf{A}$ be a vector in $\mathbb{R}^3$ and identify $\mathbf{A}$ with a 1-form A on $\mathbb{R}^3$. Let

$$\alpha = A + d\theta$$

be a connection 1-form on the bundle $\pi : \mathbb{R}^3 \times S^1 \to \mathbb{R}^3$. Consider the metric on Q_K given by

$$g((\mathbf{u}_q, u_\theta), (\mathbf{v}_q, v_\theta)) = \langle \mathbf{u}_q, \mathbf{v}_q \rangle + \kappa(\alpha(\mathbf{u}_q, v_\theta), \alpha(\mathbf{v}_q, v_\theta)).$$

The Lagrangian for the geodesic flow on (Q_K, g) is given by

$$\mathcal{L}_K(\mathbf{q}, \theta, \mathbf{v}_q, v_\theta) = \frac{1}{2}\left(||\mathbf{v}_q||^2 + (\mathbf{A} \cdot \mathbf{v}_q + v_\theta)^2\right).$$

We will call it the **Kaluza-Klein Lagrangian**. Let $B = d\alpha = dA$ and identify B with the vector $\mathbf{B} = \nabla \times \mathbf{A}$. The reduced curvature 2-form on $Q_K/S^1 \cong \mathbb{R}^3$ is identified with B or the vector $\mathbf{B}$. Let $(\mathbf{x}, \mathbf{u}, \lambda) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$ denote local coordinates on the bundle $T\mathbb{R}^3 \oplus \mathbb{R}$. The reduced Lagrangian is

$$\ell_K(\mathbf{x}, \mathbf{u}, \lambda) = \frac{1}{2}\left(||\mathbf{u}||^2 + \lambda^2\right).$$

The vertical implicit Lagrange-Poincaré equations are given by

$$\dot{p}_\theta = 0, \;\; \chi = \lambda, \;\; p_\theta = \frac{\partial \ell_K}{\partial \lambda} = \lambda,$$

where $\chi = A \cdot \dot{q} + \dot{\theta}$. Since $\dot{p}_\theta = 0$, it follows that p_θ is a constant and we set $p_\theta = \frac{e}{c}$. The horizontal Lagrange-Poincaré equations are then given by

$$\dot{\mathbf{y}} = \frac{e}{c}(\dot{\mathbf{x}} \times \mathbf{B}), \;\; \dot{\mathbf{x}} = \mathbf{u}, \;\; \mathbf{y} = \frac{\partial \ell_K}{\partial \mathbf{u}} = \mathbf{u},$$

which yields $\dot{\mathbf{u}} = \frac{e}{c}\mathbf{u} \times \mathbf{B}$.

The Stochastic Case

Suppose that $X^1, \cdots, X^k$ are arbitrary semimartingales, $L_1, \cdots, L_k \in C^\infty(R^3)$ and $V_1, \cdots, V_k$ are vectors fields on $\mathbb{R}^3$. We consider the stochastic Hamilton-Pontryagin action functional

$$\begin{aligned}
&\mathcal{S}(\mathbf{q}_t, \theta_t, \mathbf{u}_{q_t}, u_{\theta_t}, \mathbf{p}_{q_t}, p_{\theta_t}) \\
&= \int_0^T \left[\mathcal{L}_K \, dt + \sum_{i=1}^k L_i(\mathbf{q}_t) \bullet dX_t^i + \left\langle \mathbf{p}_{q_t}, \bullet d\mathbf{q}_t - \mathbf{u}_{q_t} dt - \sum_{i=1}^k V_i(\mathbf{q}_t) \bullet dX_t^i \right\rangle \right. \\
&\left. + \langle p_{\theta_t}, \bullet d\theta_t - u_{\theta_t} dt \rangle \right].
\end{aligned}$$

Then, in terms of the Kaluza-Klein description, the reduced action corresponding to $\mathcal{S}$ is given by

$$\mathcal{S}^{\text{red}}(\mathbf{x}_t, \mathbf{u}_t, \lambda_t, \mathbf{y}_t, p_{\theta_t}) = \int_0^T \left[\ell_K \, dt + \sum_{i=1}^k L_i(\mathbf{x}_t) \bullet dX_t^i + \left\langle \mathbf{y}_t, \bullet d\mathbf{x}_t - \mathbf{u}_t dt - \sum_{i=1}^k V_i(\mathbf{x}_t) \bullet dX_t^i \right\rangle \right.$$

$$+ \langle p_{\theta_t}, \bullet d\chi_t - \lambda_t dt \rangle] ,$$

where $\bullet d\chi_t = \mathbf{A} \bullet d\mathbf{q}_t + \bullet d\theta_t$. The stochastic implicit vertical Lagrange-Poincaré equations are given by

$$\begin{aligned} \bullet dp_{\theta_t} &= 0 \\ \dot{\chi}_t &= \lambda_t \\ p_{\theta_t} &= \lambda_t \end{aligned}$$

which shows that p_{θ_t} is conserved. As before, we set $p_{\theta_t} = \frac{e}{c}$. Then the stochastic implicit horizontal Lagrange-Poincaré equations are given by

$$\begin{aligned} \bullet d\mathbf{y}_t &= \frac{e}{c}(\bullet d\mathbf{x}_t \times \mathbf{B})dt + \sum_{i=1}^{k} \frac{\partial}{\partial \mathbf{x}_t} \left(L_i(\mathbf{x}_t) - \mathbf{y}_t \cdot V_i(\mathbf{x}_t) \right) \bullet dX_t^i \\ \bullet d\mathbf{x}_t &= \mathbf{u}_t dt + \sum_{i=1}^{k} V_i(\mathbf{x}_t) \bullet dX_t^i \\ \mathbf{y}_t &= \mathbf{u}_t . \end{aligned}$$

Equivalently, we can solve for

$$\begin{aligned} \bullet d\mathbf{u}_t &= \frac{e}{c}(\mathbf{u}_t \times \mathbf{B})dt + \sum_{i=1}^{k} \left(\frac{e}{c}(V_i(\mathbf{x}_t) \times \mathbf{B}) + \frac{\partial}{\partial \mathbf{x}_t} \left(L_i(\mathbf{x}_t) - \mathbf{u}_t \cdot V_i(\mathbf{x}_t) \right) \right) \bullet dX_t^i \\ \bullet d\mathbf{x}_t &= \mathbf{u}_t dt + \sum_{i=1}^{k} V_i(\mathbf{x}_t) \bullet dX_t^i , \end{aligned}$$

where the first equation represents the Lorentz force law with a stochastic perturbation. Note that if $k = 3$, $(X_t^1, X_t^2, X_t^3) = \mathbf{W}_t$, where $\mathbf{W}_t$ is a Brownian motion in $\mathbb{R}^3$, $L_i(x) = x^i$ and $V_i = 0$ then these equations become

$$\begin{aligned} \bullet d\mathbf{u}_t &= \frac{e}{c}(\mathbf{u}_t \times \mathbf{B})dt + \bullet d\mathbf{W}_t \\ \dot{\mathbf{x}}_t &= \mathbf{u}_t . \end{aligned}$$

Then, $(\mathbb{E}[\mathbf{x}_t], \mathbb{E}[\mathbf{u}_t])$ satisfies the equations for the charged particle in a magnetic field.

Chapter 11

Variational Principles for Stochastic Collective Motion

In this chapter we look at the variational principle counterpart of stochastic collective Hamiltonian systems. In Section 11.1, we show that the variational viewpoint is similar to the Hamiltonian one, namely that if one starts with a deterministic invariant Lagrangian then the stochastic system is given by an action of a Lie group valued semimartingale (i.e. the stochastic phase) on critical points of the deterministic action integral. This forms the Lagrangian counterpart of the corresponding theorem on stochastic collective Hamiltonian systems. For simplicity, we will restrict our attention to linear collectives, that is, collective Hamiltonians of the form $f_i(z) = \langle J(z), \xi_i \rangle$ for a fixed element $\xi_i \in \mathfrak{g}$. In this case, the Hamiltonian vector field corresponds to the infinitesimally generated vector field associated to ξ_i. This allows us to use the stochastic Hamilton-Pontryagin principle.

The more general case, in which one uses arbitrary collectives, is handled by a coupling mechanism similar to the Hamiltonian case. The variational principle we consider is a combination of Hamilton's principle in phase space on the Lie group side, the Hamilton-Pontryagin principle on the configuration manifold side, and a coupling term which corresponds to the coupling between the momentum map for the cotangent lifted left action (i.e. the spatial angular momentum) on the Lie group side and the momentum map on the phase space of the configuration manifold. The critical points of the action integral yield the stochastic phase. Even when the Lagrangian is not invariant, this variational principle is invariant under the partial action of the Lie group on the cotangent bundle factor (*not on the entire product space*) by cotangent lifts of right translations. Thus it corresponds to a reduced variational principle. In the reduced picture, the two momentum maps can be identified.

In Section 11.2 we describe the coupling mechanism in the deterministic case, while in Section 11.3 we extend it to the stochastic case. We wish to point out that unlike the Hamiltonian case, here the evolution of the momentum map on the phase space coincides with the evolution of J_L instead of J_R. This difference is due to the fact that we are using the cotangent lifted right action on the Lie group side.

11.1 Lagrangian Systems with Collective Perturbations

Let Q be a configuration manifold, and denote by $\mathcal{P}Q := TQ \oplus T^*Q$ its Pontryagin bundle. Let $\mathcal{L} \in C^\infty(TQ)$. Suppose G acts on Q, on TQ by tangent lifts and on T^*Q by cotangent lifts. We let gq, gv and gp denote the images of $q \in Q$, $v \in T_qQ$ and $p \in T^*_qQ$ under the action of $g \in G$. Let $J : T^*Q \to \mathfrak{g}^*$ denote the momentum map. Then, by Noether's formula (see, for instance, [42]), given any $\xi \in \mathfrak{g}$, we have $\langle J(q,p), \xi\rangle = \langle p, \xi_Q(q)\rangle$, where ξ_Q denotes the infinitesimal vector field generated on Q. Let $\xi_1, \cdots, \xi_k$ be fixed elements of $\mathfrak{g}$, $Y^1, \cdots, Y^k$ be real-valued semimartingales. Consider the stochastic action integral:

$$\mathcal{S}(q_t, v_t, p_t) = \int_0^T \left[\mathcal{L}(q_t, v_t)dt + \left\langle p_t, \bullet dq_t - v_t dt - \sum_{i=1}^k (\xi_i)_Q(q_t) \bullet dY_t^i \right\rangle\right], \qquad (11.1.1)$$

where (q_t, v_t, p_t) is a semimartingale in $\mathcal{P}Q$. We will always assume the admissibility hypothesis on semimartingales. The following theorem describes the critical points of $\mathcal{S}$ under fixed endpoint variation of q_t.

Theorem 11.1.1. The following are equivalent:

1. The $\mathcal{P}Q$-valued semimartingale $\Gamma_t = (q_t, v_t, p_t) \in \mathscr{S}(\mathcal{P}Q)$ is a critical point of $\mathcal{S}$ for all deformations $\epsilon \mapsto (q_{\epsilon,t}, v_{\epsilon,t}, p_{\epsilon,t})$ such that $\delta q_t = 0$ at $t = 0$ and $t = T$.
2. The $\mathcal{P}Q$-valued semimartingale $\Gamma_t = (q_t, v_t, p_t)$ satisfies the **stochastic implicit Euler-Lagrange equations** given by

$$\begin{aligned}
\bullet dq_t &= v_t dt + \sum_{i=1}^k (\xi_i)_Q(q_t) \bullet dY_t^i \\
\bullet dp_t &= \frac{\partial}{\partial q_t}\left(\mathcal{L}\, dt - \sum_{i=1}^k \langle p_t, (\xi_i)_Q(q_t)\rangle \bullet dY_t^i\right) \\
p_t &= \frac{\partial \mathcal{L}}{\partial v_t}.
\end{aligned} \qquad (11.1.2)$$

up to time T.

Moreover, suppose $\mathcal{L}$ is G-invariant and g^S solves the following Stratonovich differential equation in G:

$$\bullet dg = \sum_{i=1}^k T_e R_g(\xi_i) \bullet dY^i, \quad g_0 = e.$$

Then (1) and (2) are equivalent to the following statement:

1. The $\mathcal{P}Q$-valued curve $\gamma^{\det}(t) = (q(t), v(t), p(t))$ is a critical point of the deterministic action

$$\mathcal{S}^{\det}(q(t), v(t), p(t)) = \int_0^T [\mathcal{L}(q(t), v(t)) + \langle p(t), \dot{q}(t) - v(t)\rangle]dt$$

for all deformations $\epsilon \mapsto (q_\epsilon(t), v_\epsilon(t), p_\epsilon(t))$ such that $\delta q(t) = 0$ at $t = 0$ and $t = T$, and $\Gamma = g^S \cdot \gamma^{\det}$, that is, $(q_t, v_t, p_t) = (g_t^S q(t), g_t^S v(t), g_t^S p(t))$.

2. The $\mathcal{P}Q$-valued curve $\gamma^{\mathrm{det}}(t) = (q(t), v(t), p(t))$ solves the deterministic implicit Euler-Lagrange equations
$$\dot{q} = v, \quad \dot{p} = \frac{\partial \mathcal{L}}{\partial q}, \quad p = \frac{\partial \mathcal{L}}{\partial v}$$
and $\Gamma = g^S \cdot \gamma^{\mathrm{det}}$.

Proof. The equivalence between (1) and (2) (respectively (3) and (4)) follows from the stochastic (respectively, deterministic) Hamilton-Pontryagin principle.

To prove that (4) implies (2), we let $(q(t), v(t), p(t))$ solve the deterministic implicit Euler-Lagrange equations up to time T, that is,

$$\dot{p} = \frac{\partial \mathcal{L}}{\partial q}, \quad \dot{q} = v, \quad p = \frac{\partial \mathcal{L}}{\partial v},$$

and g^S solve the stochastic differential equation

$$\bullet dg = \sum_{i=1}^{k} T_e R_g(\xi_i) \bullet dY^i, \quad g_0 = e.$$

We will show that $(q_t, v_t, p_t) := g^S \cdot (q(t), v(t), p(t))$ solves the stochastic implicit Euler-Lagrange equations. First, we prove that $p_t = \frac{\partial \mathcal{L}}{\partial v_t}$. By G-invariance of $\mathcal{L}$, given any $g \in G$ and $v, w \in T_qQ$, we have

$$\begin{aligned}\langle F\mathcal{L}(gq, gv), gw\rangle &= \left.\frac{d}{ds}\right|_{s=0} \mathcal{L}(gq, gv + sgw) \\ &= \left.\frac{d}{ds}\right|_{s=0} \mathcal{L}(q, v + sw) \\ &= \langle F\mathcal{L}(q, v), w\rangle,\end{aligned}$$

that is $F\mathcal{L}(gq, gv) = g\, F\mathcal{L}(q, v)$. Hence, we obtain

$$p_t = g_t^S p(t) = g_t^S \frac{\partial \mathcal{L}}{\partial v} = \frac{\partial \mathcal{L}}{\partial v_t}.$$

Next, by the Stratonovich product rule for group actions, we have

$$\begin{aligned}\bullet dq_t &= \bullet d(g_t^S q(t)) \\ &= (\bullet dg_t^S)\, q(t) + g_t^S \dot{q}(t) dt \\ &= ((\bullet dg_t^S)(g_t^S)^{-1})_Q (g_t^S q(t)) + g_t^S v(t) dt \\ &= \sum_{i=1}^{k} (\xi_i)_Q(q_t) \bullet dY_t^i + v_t dt.\end{aligned}$$

Similarly, we have

$$\bullet dp_t = \bullet d(g_t^S p(t))$$

$$
\begin{aligned}
&= (\bullet dg_t^S)\, p(t) + g_t^S\, \dot{p}(t)dt \\
&= ((\bullet dg_t^S)(g_t^S)^{-1})_{T^*Q}(g_t^S p(t)) + g_t^S \frac{\partial \mathcal{L}}{\partial q}(t)dt \\
&= \sum_{i=1}^{k} (\xi_i)_{T^*Q}(p_t) \bullet dY_t^i + g_t^S \frac{\partial \mathcal{L}}{\partial q}(t)dt.
\end{aligned}
$$

First we recall that given any point $(q,p) \in T^*Q$, we have

$$(\xi_i)_{T^*Q}(q,p) = \left((\xi_i)_Q(q), -\frac{\partial}{\partial q} \langle p, (\xi_i)_Q(q) \rangle \right).$$

Next, we prove that

$$g_t^S\, \frac{\partial \mathcal{L}}{\partial q}(t) = \frac{\partial \mathcal{L}}{\partial q_t}.$$

For this, we will show that given any $g \in G$, $q \in Q$ and $v \in T_qQ$,

$$g\, \frac{\partial \mathcal{L}}{\partial q} = \frac{\partial \mathcal{L}}{\partial (gq)}.$$

Let $w \in T_qQ$ and $(q(s), v(s))$ be a curve in TQ with $v(0) = v$, $q(0) = q$ and $q'(0) = w$, where $(\cdot)'$ denotes the derivative with respect to s. Then, by definition

$$\left\langle \frac{\partial \mathcal{L}}{\partial q}, w \right\rangle = \frac{d}{ds}\Big|_{s=0} \mathcal{L}(q(s), v(s)).$$

Note that the curve $(gq(s), gv(s)) =: (\bar{q}(s), \bar{v}(s))$ satisfies $\bar{q}(0) = gq$, $\bar{v}(0) = gv$ and $\bar{q}'(0) = gw$. Hence

$$
\begin{aligned}
\left\langle \frac{\partial \mathcal{L}}{\partial \bar{q}}, gw \right\rangle &= \frac{d}{ds}\Big|_{s=0} \mathcal{L}(\bar{q}(s), \bar{v}(s)) \\
&= \frac{d}{ds}\Big|_{s=0} \mathcal{L}(q(s), v(s)) \\
&= \left\langle \frac{\partial \mathcal{L}}{\partial q}, w \right\rangle
\end{aligned}
$$

and hence, $g\frac{\partial \mathcal{L}}{\partial q} = \frac{\partial \mathcal{L}}{\partial (gq)}$. As a result, we obtain

$$
\begin{aligned}
\bullet dp_t &= \sum_{i=1}^{k} (\xi_i)_{T^*Q}(p_t) \bullet dY_t^i + g_t^S\, \frac{\partial \mathcal{L}}{\partial q}(t)dt \\
&= \frac{\partial \mathcal{L}}{\partial q_t} dt - \sum_{i=1}^{k} \frac{\partial}{\partial q_t} \langle p_t, (\xi_i)_Q(q_t) \rangle \bullet dY_t^i \\
&= \frac{\partial}{\partial q_t} \left(\mathcal{L}\, dt - \sum_{i=1}^{k} \langle p_t, (\xi_i)_Q(q_t) \rangle \bullet dY^i \right).
\end{aligned}
$$

This shows that (4) implies (2). Conversely, if (q_t, v_t, p_t) solves the stochastic implicit Euler-Lagrange equations then a similar reasoning shows that $(q(t), v(t), p(t)) := (g^S)^{-1}(q_t, v_t, p_t)$ solves the deterministic implicit Euler-Lagrange equations. Hence (2) and (4) are equivalent. This completes the proof. ∎

11.2 Coupling to a Lie Group: The Deterministic Case

Let $f \in C^\infty(\mathfrak{g}^*)$ and $J_L : T^*G \to \mathfrak{g}^*$ be the momentum map for the cotangent lifted left action of G on T^*G. Denote by $\omega^R : TG \to \mathfrak{g}^*$ the right invariant Maurer-Cartan 1-form, given by $\omega^R(v_g) = T_gR_{g^{-1}}(v_g)$ for all $g \in G$ and $v_g \in T_gG$. On the product of the Pontryagin bundle $\mathcal{P}Q$ with T^*G, consider the deterministic variational principle

$$\begin{aligned}\mathcal{S}^{\mathrm{det}}(q(t), v(t), p(t), g(t), p_g(t)) = & \underbrace{\int_0^T [\mathcal{L}(q(t), v(t)) + \langle p(t), \dot{q}(t) - v(t)\rangle]dt}_{\text{Hamilton-Pontryagin Principle on } \mathcal{P}Q} \\ & - \underbrace{\int_0^T J(q(t), p(t)) \circ \omega^R(\dot{g}(t))dt}_{\text{coupling term}} \\ & + \underbrace{\int_0^T [\langle p_g(t), \dot{g}(t)\rangle - f \circ J_L(g(t), p_g(t))]dt}_{\text{Hamilton's principle in phase space for } T^*G}\end{aligned}$$

We do not assume that $\mathcal{L}$ is G-invariant. Here we will consider fixed endpoint variations in $q(t)$ and $g(t)$. We suppose that G acts on $\mathcal{P}Q \times T^*G$ by acting only on the T^*G-factor by cotangent lifts of right translations. Then, since J_L is right invariant, the entire action integral is right invariant. Hence, it drops to a reduced action on $\mathcal{P}Q \times \mathfrak{g} \times \mathfrak{g}^*$ given by

$$\begin{aligned}\mathcal{S}^{\mathrm{red}}(q(t), v(t), p(t), \xi(t), \mu(t)) = & \int_0^T [\mathcal{L}(q(t), v(t)) + \langle p(t), \dot{q}(t) - v(t)\rangle]dt \qquad (11.2.1) \\ & - \int_0^T \langle J(q(t), p(t)), \xi(t)\rangle \, dt \\ & + \int_0^T [\langle \mu(t), \xi(t)\rangle - f(\mu(t))]dt, \qquad (11.2.2)\end{aligned}$$

where $\mu(t) = J_L(g(t), p_g(t))$ and $\xi(t) = \omega^R(\dot{g}(t))$. To find variations in $\xi(t)$, we can simply use the variations in the case of (right invariant) Euler-Poincaré reduction to obtain

$$\delta\xi(t) = \dot{\eta}(t) + [\eta(t), \xi(t)],$$

where $\eta(t) = \omega^R(\delta g(t))$, and if we assume fixed endpoint variations in $g(t)$, then $\eta(0) = \eta(T) = 0$.

Under such constrained variations, we obtain the following theorem:

Theorem 11.2.1. Given a curve $(q(t), v(t), p(t), \xi(t), \mu(t))$ in $\mathcal{P}Q \times \mathfrak{g} \times \mathfrak{g}^*$, the following are equivalent:

1. $(q(t), v(t), p(t), \xi(t), \mu(t))$ is a critical point of $\mathcal{S}^{\mathrm{red}}$ for all variations

$$(\delta q(t), \delta v(t), \delta p(t), \delta\xi(t), \delta\mu(t))$$

such that $\delta q(0) = \delta q(T) = 0$ and

$$\delta\xi(t) = \dot{\eta}(t) + [\eta(t), \xi(t)],$$

where $\eta(t)$ is an arbitrary $\mathfrak{g}$-valued curve such that $\eta(0) = \eta(T) = 0$.

2. $(q(t), v(t), p(t), \xi(t), \mu(t))$ satisfies the following equations:

$$\begin{aligned} \dot{q} &= v + \xi_Q(q) \\ \dot{p} &= \frac{\partial}{\partial q}[\mathcal{L}(q, v) - \langle p, \xi_Q(q)\rangle] \\ p &= \frac{\partial \mathcal{L}}{\partial v} \\ \xi &= \frac{\delta f}{\delta \mu} \\ \dot{\mu} - \dot{J}(q, p) &= \operatorname{ad}^*_\xi(\mu - J(q, p)). \end{aligned}$$

Proof. First, we derive the equations for $q(t), v(t)$, and $p(t)$. Note that the term

$$\int_0^T [\langle \mu(t), \xi(t)\rangle - f(\mu(t))]dt$$

does not involve $q(t), v(t)$, and $p(t)$. Taking arbitrary variations of $v(t)$ and $p(t)$ and setting $\delta\mathcal{S}^{\text{red}} = 0$ gives us

$$\begin{aligned} \delta v &: p - \frac{\partial \mathcal{L}}{\partial v} = 0 \\ \delta p &: \dot{q} - v - \xi_Q(q) = 0. \end{aligned}$$

where we used the fact that $\langle J(q(t), p(t)), \xi(t)\rangle = \langle p(t), (\xi(t))_Q(q(t))\rangle$. For the q-equation, we use integration by parts for the term $\int_0^T \langle p(t), \dot{q}(t)\rangle\, dt$ and the fixed endpoint conditions $\delta q(0) = \delta q(T) = 0$. Then, corresponding to the variations δq, we obtain

$$\delta q :\ \dot{p} - \frac{\partial}{\partial q}[\mathcal{L}(q, v) - \langle p, \xi_Q(q)\rangle] = 0$$

Let $\zeta = \frac{\delta f}{\delta \mu}$. Now we calculate variations of

$$-\int_0^T \langle J(q, p), \xi\rangle\, dt + \int_0^T [\langle \mu, \xi\rangle - f(\mu(t))]dt$$

for variations of the form $\delta\xi = \dot{\eta} + [\eta, \xi]$, where $\eta(0) = \eta(T) = 0$ and arbitrary variations in μ. We have,

$$\delta\left(-\int_0^T \langle J(q, p), \xi\rangle\, dt + \int_0^T \langle \mu, \xi - \zeta\rangle\, dt\right)$$

$$\begin{aligned}
&= \int_0^T \langle \mu - J(q,p), \dot\eta + [\eta, \xi]\rangle \, dt + \int_0^T \langle \delta\mu, \xi - \zeta\rangle \, dt \\
&= -\int_0^T \left\langle \dot\mu - \dot J(q,p), \eta \right\rangle dt + \int_0^T \langle \mu - J(q,p), \mathrm{ad}_\xi \eta\rangle \, dt + \int_0^T \langle \delta\mu, \xi - \zeta\rangle \, dt \\
&= \int_0^T \langle \delta\mu, \xi - \zeta\rangle \, dt - \int_0^T \left\langle \dot\mu - \dot J(q,p) - \mathrm{ad}^*_\xi(\mu - J(q,p)), \eta \right\rangle dt,
\end{aligned}$$

where we have used integration by parts and $\eta(0) = \eta(T) = 0$ to write $\langle \mu - J(q,p), \dot\eta\rangle = -\left\langle \dot\mu - \dot J(q,p), \eta \right\rangle$. Consequently, from $\delta\mathcal{S}^{\mathrm{red}} = 0$, the remaining equations emerge:

$$\begin{aligned}
\xi &= \frac{\delta f}{\delta \mu} \\
\dot\mu - \dot J(q,p) &= \mathrm{ad}^*_\xi(\mu - J(q,p)).
\end{aligned}$$

This completes the proof. ∎

Now we are ready to couple the μ with $J(q,p)$. We solve the equation

$$\dot\mu - \dot J(q,p) = \mathrm{ad}^*_\xi(\mu - J(q,p))$$

with the initial condition $\mu(0) = J(q(0), p(0))$. Then, the solution is $\mu(t) = J(q(t), p(t))$. Thus, the evolution of J can be identified with that of J_L on T^*G.

Remark 11.2.2. Suppose $\mathcal{L}$ is hyperregular and let h denote the Hamiltonian corresponding to $\mathcal{L}$. Then, since $\frac{\partial h}{\partial p} = \frac{\partial \mathcal{L}}{\partial v}$ and $\frac{\partial h}{\partial q} = -\frac{\partial \mathcal{L}}{\partial q}$, we obtain the equations:

$$\begin{aligned}
\dot q &= \frac{\partial h}{\partial p} + \xi_Q(q) \\
\dot p &= -\frac{\partial h}{\partial q} - \frac{\partial}{\partial q}\left\langle p, \xi_Q(q)\right\rangle,
\end{aligned}$$

where $\xi = \frac{\delta f}{\delta \mu}$. Moreover, using the Collective Hamiltonian Theorem,

$$\xi_{T^*Q}(q,p) = \left(\xi_Q(q), -\frac{\partial}{\partial q}\left\langle p, \xi_Q(q)\right\rangle\right) = X_{f\circ J}(q,p).$$

Therefore, letting $z(t) = (q(t), p(t))$, we obtain

$$\dot z = X_h(z) + X_{f\circ J}(z) = X_{h + f\circ J}(z). \tag{11.2.3}$$

Moreover, if $\mathcal{L}$ is G-invariant (and hence, h is G-invariant), we have the reconstruction equation

$$\dot g = T_e R_g(\xi(t)) = T_e R_g\left(\frac{\delta f}{\delta \mu}\right), \quad g(0) = e, \tag{11.2.4}$$

corresponds to the phase added to the integral curve $z^0(t)$ of X_h while solving for Equation (11.2.3), that is, $z(t)$ is a solution of Equation (11.2.3) if $z(t) = g(t)\cdot z^0(t)$, where $g(t)$ solves Equation (11.2.4).

11.3 Coupling to a Lie Group: The Stochastic Case

Now we extend the discussion in the previous section to the stochastic case. Let $f_1, \cdots, f_k \in C^\infty(\mathfrak{g}^*)$. Let $Y^1, \cdots, Y^k$ be semimartingales. On $\mathcal{P}Q \times T^*G$ consider the stochastic action integral:

$$\begin{aligned}
\mathcal{S}(q_t, v_t, p_t, g_t, p_{g_t}) = &\underbrace{\int_0^T [\mathcal{L}(q_t, v_t)dt + \langle p_t, \bullet dq_t - v_t dt\rangle]}_{\text{Hamilton-Pontryagin Principle on } \mathcal{P}Q} \\
&- \underbrace{\int_0^T J(q_t, p_t) \circ \omega^R \bullet dg_t}_{\text{coupling term}} \\
&+ \underbrace{\int_0^T \left[\langle p_{g_t}, \bullet dg_t\rangle - \sum_{i=1}^k f_i \circ J_L(g_t, p_{g_t}) \bullet dY_t^i\right]}_{\text{stochastic Hamilton's principle in phase space for } T^*G}
\end{aligned}$$

As in the deterministic case, the stochastic action integral is invariant under the action of G on $\mathcal{P}Q \times T^*G$, where G only acts on the T^*G factor via cotangent lifts of right translations. As before, using right invariance of J_L, we obtain a reduced stochastic action integral on $\mathcal{P}Q \times \mathfrak{g} \times \mathfrak{g}^*$, given by

$$\begin{aligned}
\mathcal{S}^{\text{red}}(q_t, v_t, p_t, \xi_t, \mu_t) = &\int_0^T [\mathcal{L}(q_t, v_t)dt + \langle p_t, \bullet dq_t - v_t dt\rangle] \\
&- \int_0^T J(q_t, p_t) \bullet d\xi_t \\
&+ \int_0^T \left[\langle \mu_t, \bullet d\xi_t\rangle - \sum_{i=1}^k f_i(\mu_t) \bullet dY_t^i\right],
\end{aligned}$$

where $\xi_t = \int \omega^R \bullet dg_t$. The variations of the term $\int_0^T \langle \mu_t, \bullet d\xi_t\rangle$ can be calculated directly by using the Maurer-Cartan equation

$$\mathbf{d}\omega^R(X, Y) = [\omega^R(X), \omega^R(Y)],$$

for all vector fields X and Y on G, where $\mathbf{d}\omega^R$ is the exterior derivative of ω^R viewed as a Lie algebra valued 2-form. Indeed, let $\eta_t = \langle \omega^R(g_t), \delta g_t\rangle$. Then, we have

$$\begin{aligned}
\delta\xi_t &= \delta \int \omega^R \bullet dg_t \\
&= \int i_{\delta g_t} \mathbf{d}\omega^R \bullet dg_t + \eta_t - \eta_0 \\
&= \int [\eta_t, \bullet d\xi_t] + \eta_t - \eta_0.
\end{aligned}$$

Thus, we get

$$\bullet d(\delta\xi_t) = [\eta_t, \bullet d\xi_t] + \bullet d\eta_t.$$

Note that η_t vanishes at $t = 0$ and $t = T$ if δg_t vanishes at $t = 0$ and $t = T$. Then we have the following theorem:

Theorem 11.3.1. Given a semimartingale $(q_t, v_t, p_t, \xi_t, \mu_t)$ in $\mathcal{P}Q \times \mathfrak{g} \times \mathfrak{g}^*$, the following are equivalent:

1. $(q_t, v_t, p_t, \xi_t, \mu_t)$ is a critical point of $\mathcal{S}^{\mathrm{red}}$ for all variations

   $$(\delta q(t), \delta v(t), \delta p(t), \delta\xi(t), \delta\mu(t))$$

   such that $\delta q_0 = \delta q_T = 0$ and

   $$\bullet d(\delta\xi_t) = \bullet d\eta_t + [\eta_t, \bullet d\xi_t],$$

   where $\eta(t)$ is an arbitrary $\mathfrak{g}$-valued semimartingale such that $\eta(0) = \eta(T) = 0$.

2. $(q_t, v_t, p_t, \xi_t, \mu_t)$ satisfies the following stochastic differential equations:

   $$\begin{aligned} \bullet dq_t &= v_t dt + (\bullet d\xi_t)_Q(q_t) \\ \bullet dp_t &= \frac{\partial}{\partial q}[\mathcal{L}(q_t, v_t)dt - \langle p, (\bullet d\xi_t)_Q(q_t)\rangle] \\ p_t &= \frac{\partial \mathcal{L}}{\partial v_t} \\ \bullet d\xi_t &= \sum_{i=1}^{k} \frac{\delta f_i}{\delta \mu_t} \bullet dY_t^i \\ \bullet d\mu_t - \bullet dJ(q_t, p_t) &= \mathrm{ad}^*_{\bullet d\xi}(\mu_t - J(q_t, p_t)). \end{aligned}$$

Remark 11.3.2. Here $(\bullet d\xi_t)_Q$ is interpreted in the sense of Equation (8.1.3). Namely, using the equation

$$\bullet d\xi_t = \sum_{i=1}^{k} \frac{\delta f_i}{\delta \mu_t} \bullet dY_t^i,$$

we have

$$\bullet d\xi_Q(q_t) = \sum_{i=1}^{k} \left(\frac{\delta f_i}{\delta \mu_t}\right)_Q (q_t) \bullet dY_t^i.$$

Proof. For simplicity, we proceed with formal analogy to the deterministic case. By noting that

$$\int_0^T J(q_t, p_t) \bullet d\xi_t = \int_0^T \langle p_t, (\bullet d\xi_t)_Q(q_t)\rangle\,,$$

the equations for q_t, v_t and p_t can be derived exactly as in the case of the stochastic Hamilton-Pontryagin principle. The equations for q_t and v_t follow directly, while the equation for p_t involves using integration by parts on the term $\delta\langle p_t, \bullet dq_t\rangle$ and the fact that $\delta q_0 = \delta q_T = 0$.

For the remaining two equations, we need to consider the variation of the terms

$$-\int_0^T J(q_t,p_t)\bullet d\xi_t + \int_0^T \left[\langle \mu_t, \bullet d\xi_t\rangle - \sum_{i=1}^k f_i(\mu_t)\bullet dY_t^i\right]$$

under arbitrary variations of μ_t and variations of ξ_t given by

$$\bullet d\delta\xi_t = \bullet d\eta_t + \mathrm{ad}_{\eta_t}\bullet d\xi_t,$$

where η_t is a $\mathfrak{g}$-valued semimartingale that vanishes at $t=0$ and $t=T$. Let $\xi_i = \frac{\delta f_i}{\delta \mu_t}$. From arbitrary variations in μ_t, we get

$$\delta\mu_t : \ \bullet d\xi_t - \sum_{i=1}^k \xi_{i_t}\bullet dY_t^i = 0,$$

that is $\bullet d\xi_t = \sum_{i=1}^k \frac{\delta f_i}{\delta\mu_t}\bullet dY^i$. Finally, for the remaining equation, for simplicity, we proceed by formally putting $\bullet d(\delta\xi_t) = \delta(\bullet d\xi_t)$. Using integration by parts and fixed endpoint conditions in η, this gives us

$$\begin{aligned}
&\mathcal{D}\left(-\int_0^T J(q_t,p_t)\bullet d\xi_t + \int_0^T \langle\mu_t, \bullet d\xi_t\rangle\right)\\
&= -\int_0^T J(q_t,p_t)\bullet d\eta_t + \int_0^T \left\langle \mathrm{ad}^*_{\bullet d\xi_t}(J(q_t,p_t)-\mu_t), \eta_t\right\rangle\\
&+ \int_0^T \langle \delta\mu_t, \eta_t\rangle\\
&= \int_0^T \left\langle \bullet d(J(q_t,p_t)-\mu_t) - \mathrm{ad}^*_{\bullet d\xi_t}(J(q_t,p_t)-\mu_t), \eta_t\right\rangle,
\end{aligned}$$

which yields $\bullet d(\mu_t - J(q_t,p_t)) = \mathrm{ad}^*_{\bullet d\xi_t}(\mu_t - J(q_t,p_t))$ after setting $\mathcal{D}\mathcal{S}^{\mathrm{red}} = 0$. ∎

Remark 11.3.3. The variation

$$\mathcal{D}\left(-\int_0^T J(q_t,p_t)\bullet d\xi_t + \int_0^T \langle\mu_t, \bullet d\xi_t\rangle\right)$$

can be computed more systematically using the methods of the previous chapter.

Similar to the deterministic setup, we can show that the evolution of the momentum map J coincides with that of J_L by setting $\mu_0 = J(q_0,p_0)$.

Remark 11.3.4 (Relation with stochastic collective Hamiltonian Systems)**.** Suppose $\mathcal{L}$ is hyperregular and let h denote the Hamiltonian corresponding to $\mathcal{L}$. Then, after setting $\mu_0 = J(q_0,p_0)$, we have,

$$\bullet dq_t = \frac{\partial h}{\partial p_t}dt + (\bullet d\xi)_Q(q_t)$$

$$
\begin{aligned}
&= \frac{\partial h}{\partial p_t} dt + \sum_{i=1}^{k} \left(\frac{\delta f_i}{\delta \mu_t}\Big|_{\mu = J(q_t, p_t)} \right)_Q (q_t) \bullet dY_t^i \\
\bullet dp &= -\frac{\partial h}{\partial q_t} dt - \frac{\partial}{\partial q_t} \langle p_t, (\bullet d\xi)_Q (q_t) \rangle \\
&= -\frac{\partial h}{\partial q_t} dt - \sum_{i=1}^{k} \frac{\partial}{\partial q_t} \left\langle p_t, \left(\frac{\delta f_i}{\delta \mu_t}\Big|_{\mu = J(q_t, p_t)} \right)_Q (q_t) \right\rangle \bullet dY_t^i.
\end{aligned}
$$

The Collective Hamiltonian Theorem gives us

$$
\left(\frac{\delta f}{\delta \mu} \right)_{T^*Q} (q, p) = \left(\left(\frac{\delta f}{\delta \mu} \right)_Q (q), -\frac{\partial}{\partial q} \left\langle p, \left(\frac{\delta f}{\delta \mu} \right)_Q (q) \right\rangle \right) = X_{f \circ J}(q, p).
$$

Therefore, letting $\Gamma_t = (q_t, p_t)$, we obtain the stochastic Hamiltonian system

$$
\bullet d\Gamma = X_h(\Gamma) dt + \sum_{i=1}^{k} X_{f_i \circ J}(\Gamma) \bullet dY^i. \tag{11.3.1}
$$

Moreover, if $\mathcal{L}$ is G-invariant then so is h, and hence the stochastic reconstruction equation

$$
\bullet dg = \sum_{i=1}^{k} T_e R_g \left(\frac{\delta f}{\delta \mu}\Big|_{\mu = J(\Gamma)} \right) \bullet dY^i, \quad g(0) = e, \tag{11.3.2}
$$

corresponds to the stochastic phase for Equation (11.3.1).

Future Areas of Research

In this thesis we have explored the role of symmetries in both the Hamiltonian side and the variational principle side of stochastic geometric mechanics by using the geometric nature of the Stratonovich integral to develop analogues of deterministic conservation laws as well as reduction techniques. Let us describe some ongoing and future areas of research connected with this thesis.

1. **Reduction and reconstruction of Schwartz operators and their use in stochastic geometric mechanics**: In Chapter 3, Section 3.5, we have carried out the reduction of stochastic differential equations given by Schwartz operators. This generalizes the reduction for Stratonovich differential equations carried out Lázaro-Camí and Ortega [54]. Schwartz operators provide a unified framework that simultaneously encompasses both Itô and Stratonovich formulations within a single geometric setting, and hence, we expect that the reduction techniques also carry over to symmetric Itô equations. The problems of reconstruction and skew-product decomposition for Schwartz operators remain open. It is also interesting to explore whether Schwartz operators can serve as a geometric framework that reconciles the two prevailing approaches to stochastic geometric mechanics: one based on Stratonovich integration, which is based on the change of coordinates formula but does not take martingales into account, and the other based on interpreting semimartingale drifts as velocities, which naturally accounts for martingales but is geometrically less straightforward.

2. **Block regularization in stochastic celestial mechanics**: In Chapter 5, we described the Moser regularization of the angularly perturbed stochastic Kepler problem. The key observation here was that the radial distance and the speeds agreed with the deterministic Kepler problem, and hence, on the 'dt' part of the equations, we could use Moser regularization methods.

   Another important regularization method in celestial mechanics, which also depends on radial distance and speed, is called block regularization. Block regularization was developed by Conley and Easton [22] and used by McGehee [65] to show that among potentials of the form $V(r) = -\frac{1}{r^\alpha}$, $\alpha > 0$, the ones for which the flow can be extended at least continuously over the collisional singularity, are those with $\alpha = 2 - 2/n$. More recently, in Saha and Stoica [82], we show that the logarithm central force problem can be block regularized.

We believe that block regularization methods can also be applied to central force problems with central force problems stochastically perturbed in the angular direction, and moreover, the results should be similar to the corresponding deterministic case.

3. **Non-holonomic Mechanics and Constrained Systems**: While the mechanical systems considered here are holonomic, it remains to study potential applications to non-holonomic mechanics. In this regard, there has been some recent developments in non-holonomic stochastic mechanics by Li, Gay-Balmaz, Shi, and Wang [50], [51], with applications to stochastic thermodynamics by Vaquero del Pino, Gay-Balmaz, Yoshimura, and Chew [31]. There are several research areas within this field that could be explored, for instance, the 'average' effects of stochastic forces on a deterministic constrained system, applications to control theory, reduction and reconstruction methods for stochastic constrained systems, and relations with thermodynamics.

4. **The drift of stochastic Hamiltonian systems**: The stochastic Euler-Poincaré reduction developed by Arnaudon, Chen and Cruzeiro [7] is useful for modelling systems with viscous forces. By using the generalized derivative of semimartingales that accounts for the drift of semimartingales, they are able to obtain deterministic equations of motion on the Lie algebra of a Lie group from stochastic variational principles. These equations resemble Euler-Poincaré equations with an additional forcing term. The Lie-Poisson counterpart of this drift-based Euler-Poincaré reduction remains open.

5. **Applications to Infinite Dimensional Systems**: In this thesis, the focus has been in the context of finite-dimensional systems. We believe that the ideas presented in this thesis can be applied to infinite-dimensional systems as well, for instance, in fluids, and image registration. Applications of the stochastic collective Hamiltonian framework to image registration is already being considered in a joint work with Tanya Schmah and Cristina Stoica. We also wish to suggest an infinite-dimensional mechanical setting for stochastic collective Hamiltonian systems. In the deterministic case, a problem that involves coupling to a Lie group is the attachment of a flexible linear elastic shear beam to a rigid body such as a satellite. Here the phase space is $\mathfrak{so}(3)^* \times T^*Q$, where

$$Q = \{\sigma : [0, L] \to \mathbb{R}^3 \mid \sigma \text{ is smooth}\},$$

and L is the maximum length of the attachment. If the rigid body evolves stochastically, then the results in Chapter 7 imply that one obtains a stochastic collective Hamiltonian system on T^*Q via the coupling mechanism. Moreover, it would be interesting to extend this to the case where the flexible attachment is also subjected to $SO(3)$-invariant noise.

Bibliography

[1] Sergio Albeverio, Ph Blanchard, and R Høegh-Krohn. Newtonian diffusions and planets, with a remark on non-standard dirichlet forms and polymers. In *Stochastic Analysis and Applications: Proceedings of the International Conference held in Swansea, April 11–15, 1983*, pages 1–24. Springer, 1984.

[2] Sergio A Albeverio, P Blanchard, and R Høegh-Krohn. A stochastic model for the orbits of planets and satellites: an interpretation of Titius-Bode law. Technical report, Bielefeld TU. Fak. Phys., Bielefeld, 1982.

[3] Alexis Arnaudon, Alex L. De Castro, and Darryl D. Holm. Noise and dissipation on coadjoint orbits. *Journal of Nonlinear Science*, 28:91 – 145, 2016.

[4] Alexis Arnaudon, Alex L. De Castro, and Darryl D. Holm. Noise and dissipation in rigid body motion. In Sergio Albeverio, Ana Bela Cruzeiro, and Darryl Holm, editors, *Stochastic Geometric Mechanics*, pages 1–12, Cham, 2017. Springer International Publishing.

[5] M. Arnaudon and A. Thalmaier. Stability of stochastic differential equations in manifolds. In *Séminaire de Probabilités XXXII*, volume 188 of *Lecture Notes in Mathematics*, pages 188–214. Springer, Berlin, Heidelberg, 1998.

[6] Marc Arnaudon. Connexions et martingales dans les groupes de lie. *Séminaire de probabilités de Strasbourg*, 26:146–154, 1992.

[7] Marc Arnaudon, Xin Chen, and Ana Bela Cruzeiro. Stochastic Euler-Poincaré reduction. *Journal of Mathematical Physics*, 55(8):081507, 08 2014.

[8] Marc Arnaudon and Ana Bela Cruzeiro. Generalized navier-stokes flows. 2012.

[9] Marc Arnaudon and Ana Bela Cruzeiro. Lagrangian Navier–Stokes diffusions on manifolds: Variational principle and stability. *Bulletin des Sciences Mathématiques*, 136(8):857–881, 2012.

[10] Richard F. Bass. The measurability of hitting times, 2023.

[11] J.-M Bismut. *Mecanique Aleatoire*, volume 866 of *Lecture Notes in Mathematics*. Springer Berlin / Heidelberg, Berlin, Heidelberg, 1981.

[12] Jean-Michel Bismut. An introduction to duality in random mechanics. In *Stochastic Control Theory and Stochastic Differential Systems: Proceedings of a Workshop of the „Sonderforschungsbereich 72 der Deutschen Forschungsgemeinschaft an der Universität Bonn "which took place in January 1979 at Bad Honnef*, pages 42–60. Springer, 2005.

[13] N. Bou-Rabee and H. Owhadi. Stochastic variational integrators. *IMA Journal of Numerical Analysis*, 29(2):421–443, April 2008.

[14] N. Bou-Rabee and H. Owhadi. Stochastic variational integrators. *IMA Journal of Numerical Analysis*, 29(2):421–443, April 2008.

[15] Pedro Catuogno. Second order connections and stochastic horizontal lifts. *Journal of Geometry and Physics*, 56(9):1637–1653, 2006.

[16] Pedro J Catuogno, Diego S Ledesma, and Paulo R Ruffino. A note on stochastic calculus in vector bundles. In *Séminaire de Probabilités XLV*, pages 353–364. Springer, 2013.

[17] Pedro J Catuogno and Paulo RC Ruffino. *Stochastic Exponential in Lie Groups and its Applications*. UNICAMP, 2003.

[18] Hernan Cendra, Jerrold Marsden, Sergey Pekarsky, and Tudor Ratiu. Variational principles for lie—poisson and hamilton—poincaré equations. *Moscow Mathematical Journal*, 3, 07 2003.

[19] Hernán Cendra, Jerrold E Marsden, and Tudor S Ratiu. Lagrangian reduction by stages. *Memoirs of the American Mathematical Society*, 126(601):1–130, 1997.

[20] Xin Chen, Ana Bela Cruzeiro, and Tudor S Ratiu. Stochastic variational principles for dissipative equations with advected quantities. *Journal of Nonlinear Science*, 33(1):5, 2023.

[21] F. Cipriano and A. B. Cruzeiro. Navier–stokes equation and diffusions on the group of homeomorphisms of the torus. *Communications in Mathematical Physics*, 275:255–269, 2007.

[22] C. Conley and R. Easton. Isolated invariant sets and isolating blocks. *Trans. Amer. Math. Soc*, 158:35–61, 1971.

[23] Colin Cotter, Dan Crisan, Darryl D Holm, Wei Pan, and Igor Shevchenko. Numerically modeling stochastic lie transport in fluid dynamics. *Multiscale Modeling & Simulation*, 17(1):192–232, 2019.

[24] Colin J Cotter, Georg A Gottwald, and Darryl D Holm. Stochastic partial differential fluid equations as a diffusive limit of deterministic lagrangian multi-time dynamics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2205):20170388, 2017.

[25] Jacky Cresson. The stochastisation hypothesis and the spacing of planetary systems. *Journal of Mathematical Physics*, 52(11), 2011.

[26] Jacky Cresson and Sébastien Darses. Stochastic embedding of dynamical systems. *Journal of Mathematical Physics*, 48(7), 2007.

[27] Jacky Cresson, Laurent Nottale, and Thierry Lehner. Stochastic modification of Newtonian dynamics and induced potential—application to spiral galaxies and the dark potential. *Journal of Mathematical Physics*, 62(7), 2021.

[28] Jacky Cresson, Frédéric Pierret, and Bénédicte Puig. The Sharma-Parthasarathy stochastic two-body problem. *Journal of Mathematical Physics*, 56(3), 2015.

[29] Ana Bela Cruzeiro, Darryl D. Holm, and Tudor S. Ratiu. Momentum maps and stochastic clebsch action principles. *Communications in Mathematical Physics*, 357:873–912, 2016.

[30] Richard H. Cushman and Larry M. Bates. *Global Aspects of Classical Integrable Systems*. Birkhäuser Basel, 2015.

[31] Héctor Vaquero del Pino, François Gay-Balmaz, Hiroaki Yoshimura, and Lock Yue Chew. A variational formulation of stochastic thermodynamics. part i: Finite-dimensional systems, 2025.

[32] Claude. Dellacherie and Paul Andre. Meyer. *Probabilities and potential*. North-Holland mathematics studies ; 29. Hermann, Paris, 1978.

[33] Tobias Diez and Lukas Miaskiwskyi. Expectation values of polynomials and moments on general compact lie groups, 2022.

[34] Theodore D Drivas, Darryl D Holm, and James-Michael Leahy. Lagrangian averaged stochastic advection by lie transport for fluids. *Journal of Statistical Physics*, 179(5):1304–1342, 2020.

[35] Johannes Jisse Duistermaat and Johan AC Kolk. *Lie groups*. Springer Science & Business Media, 2012.

[36] M. Emery. *Stochastic Calculus in Manifolds*. Universitext. Springer Berlin Heidelberg, 2012.

[37] Michel Emery. An invitation to second-order stochastic differential geometry. 2007.

[38] E. Fernández-Saiz, J. de Lucas, X. Rivas, and M. Zajac. Hamiltonian stochastic lie systems and applications, 2025.

[39] François Gay-Balmaz and Darryl D Holm. Stochastic geometric models with non-stationary spatial correlations in lagrangian fluid flows. *Journal of nonlinear science*, 28(3):873–904, 2018.

[40] M Hakim-Dowek and Dominique Lépingle. L'exponentielle stochastique des groupes de lie. In *Séminaire de Probabilités XX 1984/85: Proceedings*, pages 352–374. Springer, 2006.

[41] Gert Heckman and Tim de Laat. On the regularization of the Kepler problem. *Journal of Symplectic Geometry*, 10(3):463 – 473, 2012.

[42] Darryl Holm, Tanya Schmah, and Cristina Stoica. *Geometric Mechanics and Symmetry.* Oxford University Press, Oxford, 2009.

[43] Darryl D. Holm. Variational principles for stochastic fluid dynamics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2176):20140963, 2015.

[44] Qiao Huang and Jean-Claude Zambrini. From second-order differential geometry to stochastic geometric mechanics. *Journal of nonlinear science*, 33(4), 2023.

[45] D. Kazhdan, B. Kostant, and S. Sternberg. Hamiltonian group actions and dynamical systems of calogero type. *Communications on Pure and Applied Mathematics*, 31(4):481–507, 1978.

[46] Perinkulam S Krishnaprasad and Jerrold E Marsden. Hamiltonian structures and stability for rigid bodies with flexible attachments. *Archive for Rational Mechanics and Analysis*, 98(1):71–93, 1987.

[47] P. Kustaanheimo and E. Stiefel. Perturbation theory of kepler motion based on spinor regularization. *J. Reine Angew. Math.*, 218:204–219, 1965.

[48] Joan-Andreu Lázaro-Camí and Juan-Pablo Ortega. The stochastic hamilton-jacobi equation. *Journal of Geometric Mechanics*, 1(3):295–315, 2009.

[49] Melvin Leok, Tomoki Ohsawa, and Diana Sosa. Hamilton–jacobi theory for degenerate lagrangian systems with holonomic and nonholonomic constraints. *Journal of Mathematical Physics*, 53(7), 2012.

[50] Tianzhi Li, François Gay-Balmaz, Donghua Shi, and Jinzhi Wang. Variational principle for stochastic nonholonomic systems part i: Continuous-time formulation. In *International Conference on Geometric Science of Information*, pages 204–213. Springer, 2025.

[51] Tianzhi Li, François Gay-Balmaz, Donghua Shi, and Jinzhi Wang. Variational principle for stochastic nonholonomic systems part ii: Stochastic nonholonomic integrator. In *International Conference on Geometric Science of Information*, pages 225–233, 2025.

[52] Ming Liao et al. *Invariant Markov processes under Lie group actions.* Springer, 2018.

[53] Ottmar Loos. *Symmetric spaces: General theory*, volume 1. WA benjamin, 1969.

[54] Joan-Andreu Lázaro-Camí and Juan-Pablo Ortega. Reduction, reconstruction, and skew-product decomposition of symmetric stochastic differential equations. *Stochastics and Dynamics*, 9, 05 2007.

[55] Joan-Andreu Lázaro-Camí and Juan-Pablo Ortega. Stochastic hamiltonian dynamical systems. *Reports on Mathematical Physics*, 61(1):65–122, 2008.

[56] Kirill C. H. Mackenzie. *General Theory of Lie Groupoids and Lie Algebroids.* London Mathematical Society Lecture Note Series. Cambridge University Press, 2005.

[57] G.-M. Marle. Symplectic manifolds, dynamical groups, and hamiltonian mechanics. In *Differential Geometry and Relativity*, pages 249–269. Springer, 1976.

[58] Jerrold Marsden and Alan Weinstein. Reduction of symplectic manifolds with symmetry. *Reports on Mathematical Physics*, 5(1):121–130, 1974.

[59] Jerrold Marsden and Alan Weinstein. Coadjoint orbits, vortices, and clebsch variables for incompressible fluids. *Physica D: Nonlinear Phenomena*, 7(1-3):305–323, 1983.

[60] Jerrold E. Marsden. *Lectures on Mechanics.* London Mathematical Society Lecture Note Series. Cambridge University Press, 1992.

[61] Jerrold E Marsden, Tudor Raţiu, and Alan Weinstein. Semidirect products and reduction in mechanics. *Transactions of the american mathematical society*, 281(1):147–177, 1984.

[62] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to mechanics and symmetry : a basic exposition of classical mechanical systems.* Texts in applied mathematics ; 17. Springer, New York, 2nd ed. edition, 1999.

[63] Jerrold E Marsden and Jürgen Scheurle. Lagrangian reduction and the double spherical pendulum. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 44(1):17–43, 1993.

[64] Jerrold E. Marsden and Jürgen Scheurle. The reduced euler-lagrange equations. *Fields Institute Communications*, 1:139–164, 1993.

[65] R. McGehee. Double collisions for a classical particle system with non-gravitational interaction. *Comment. Math. Helvetici*, 56:524–557, 1981.

[66] P. A. Meyer. A differential geometric formalism for the ito calculus. In David Williams, editor, *Stochastic Integrals*, pages 256–270, Berlin, Heidelberg, 1981. Springer Berlin Heidelberg.

[67] Paul-André Meyer. Géométrie stochastique sans larmes. In *Séminaire de Probabilités XV 1979/80: Avec table générale des exposés de 1966/67 à 1978/79*, pages 44–102. Springer, 2006.

[68] Peter W Michor. *Topics in differential geometry*, volume 93. American Mathematical Soc., 2008.

[69] John Milnor. On the geometry of the Kepler problem. *The American Mathematical Monthly*, 90(6):353–365, 1983.

[70] J. Moser. Regularization of kepler's problem and the averaging method on a manifold. *Communications in Pure and Applied Mathematics*, 23:609–636, 1970.

[71] T Nakagomi, K Yasue, and J-C Zambrini. Stochastic variational derivations of the navier-stokes equation. *Letters in Mathematical Physics*, 5(6):545–552, 1981.

[72] Edward Nelson. Derivation of the Schrödinger equation from Newtonian mechanics. *Phys. Rev.*, 150:1079–1085, Oct 1966.

[73] Edward Nelson. *Dynamical theories of Brownian motion.* Mathematical notes. Princeton University Press, Princeton, N. J, 1967.

[74] Katsumi Nomizu. Invariant affine connections on homogeneous spaces. *American Journal of Mathematics*, 76(1):33–65, 1954.

[75] James R Norris. A complete differential formalism for stochastic calculus in manifolds. *Séminaire de probabilités de Strasbourg*, 26:189–209, 1992.

[76] Laurent Nottale. New formulation of stochastic mechanics. Application to chaos. *Chaos and diffusion in Hamiltonian systems*, 1:173, 1995.

[77] Laurent Nottale, Gérard Schumacher, and Jean Gay. Scale relativity and quantization of the solar system. *Astronomy and Astrophysics, v. 322, p. 1018-1025*, 322:1018–1025, 1997.

[78] Juan-Pablo Ortega and Tudor S Ratiu. *Momentum maps and Hamiltonian reduction*, volume 222. Springer Science & Business Media, 2013.

[79] Philip. Protter. *Stochastic Integration and Differential Equations.* Stochastic Modelling and Applied Probability, 21. Springer Berlin Heidelberg, Berlin, Heidelberg, 2nd ed. 2005. edition, 2005.

[80] Archishman Saha. Moser regularization of a stochastically perturbed kepler problem. *Celestial Mechanics and Dynamical Astronomy*, 137(6):36, 2025.

[81] Archishman Saha. On stochastic variational principles. *Geometric Mechanics*, 2(02):195–220, 2025.

[82] Archishman Saha and Cristina Stoica. Block regularisation of the logarithm central force problem. *Journal of Dynamics and Differential Equations*, 37(2):1625–1636, 2025.

[83] Laurent Schwartz. Géométrie différentielle du 2eme ordre, semi-martingales et équations différentielles stochastiques sur une variété différentielle. In *Séminaire de Probabilités XVI, 1980/81 Supplément: Géométrie Différentielle Stochastique*, pages 1–148. Springer, 2006.

[84] Ichiro Shigekawa. On stochastic horizontal lifts. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 59(2):211–221, 1982.

[85] Juan C Simo, TA Posbergh, and Jerrold E Marsden. Stability of coupled rigid body and geometrically exact rods: block diagonalization and the energy-momentum method. *Physics Reports*, 193(6):279–360, 1990.

[86] Simao N Stelmastchuk. Martingales in reductive homogeneous spaces. *arXiv preprint arXiv:1106.6246*, 2011.

[87] Simao N Stelmastchuk. The itô exponential on lie groups. *Int. J. Contemp. Math. Sci*, 8:307–326, 2013.

[88] O. D. Street and D. Crisan. Semi-martingale driven variational principles. *Proceedings of the Royal Society A*, 477:20200957, 2021.

[89] Oliver D. Street and So Takao. Semimartingale driven mechanics and reduction by symmetry for stochastic and dissipative dynamical systems, 2023.

[90] Micheele Thieullen and Jean-Claude Zambrini. Symmetries in the stochastic calculus of variations. *Probability theory and related fields*, 107(3):401–427, 1997.

[91] Kentaro Yano and EM Patterson. Vertical and complete lifts from a manifold to its cotangent bundle. *Journal of the Mathematical Society of Japan*, 19(1):91–113, 1967.

[92] Kunio Yasue. Stochastic calculus of variations. *Journal of Functional Analysis*, 41(3):327–340, 1981.

[93] Hiroaki Yoshimura and Jerrold E. Marsden. Dirac structures in lagrangian mechanics part i: Implicit lagrangian systems. *Journal of Geometry and Physics*, 57(1):133–156, 2006.

[94] Hiroaki Yoshimura and Jerrold E. Marsden. Dirac structures in lagrangian mechanics part ii: Variational structures. *Journal of Geometry and Physics*, 57(1):209–250, 2006.

[95] Hiroaki Yoshimura and Jerrold E. Marsden. Dirac cotangent bundle reduction. *Journal of Geometric Mechanics*, 1(1):87–158, 2009.