Noise and Dissipation on Coadjoint Orbits

Alexis Arnaudon¹ · Alex L. De Castro^{1,2} ·
Darryl D. Holm¹

Received: 11 June 2017 / Accepted: 12 June 2017 / Published online: 17 July 2017
© The Author(s) 2017. This article is an open access publication

Abstract We derive and study stochastic dissipative dynamics on coadjoint orbits by incorporating noise and dissipation into mechanical systems arising from the theory of reduction by symmetry, including a semidirect product extension. Random attractors are found for this general class of systems when the Lie algebra is semi-simple, provided the top Lyapunov exponent is positive. We study in details two canonical examples, the free rigid body and the heavy top, whose stochastic integrable reductions are found and numerical simulations of their random attractors are shown.

Keywords Stochastic geometric mechanics · Euler-Poincaré theory · Coadjoint orbits · Invariant measures · Random attractors · Lyapunov exponents

Mathematics Subject Classification 37H10 · 37J15 · 60H10

1 Introduction

Geometric mechanics, introduced in [Poincaré \(1901\)](#), is a powerful formalism for understanding dynamical systems whose Lagrangian and Hamiltonian are invariant

Communicated by Paul Newton.

✉ Darryl D. Holm
d.holm@imperial.ac.uk

Alexis Arnaudon
alexis.arnaudon@imperial.ac.uk

Alex L. De Castro
alex.lucio.castro@gmail.com

¹ Department of Mathematics, Imperial College, London SW7 2AZ, UK

² Departamento de Matemática, PUC-Rio, Rio de Janeiro 22451-900, Brazil

under the transformations of the configuration manifold M by a Lie group G . Examples of its applications range from the simple finite dimensional dynamics of the freely rotating rigid body to the infinite dimensional dynamics of the ideal fluid equations. For a historical review and basic references, see, e.g., [Holm et al. \(1998\)](#). See [Marsden and Ratiu \(1999\)](#), [Holm \(2008\)](#) and [Holm et al. \(2009\)](#) for textbook introductions to geometric mechanics and background references. One of the main approaches of geometric mechanics is the method of reduction of the motion equations of a mechanical system by a Lie group symmetry G in either its Lagrangian formulation on the tangent space TM of a configuration manifold M , or its Hamiltonian formulation on the cotangent space T^*M .

The method of reduction by symmetry yields reduced Lagrangian and Hamiltonian formulations of the Euler–Poincaré equations governing the dynamics of the momentum map $J : T^*M \rightarrow \mathfrak{g}^*$, where $\mathfrak{g}^*$ is the dual Lie algebra of the Lie symmetry group G . In general terms, Lie group reduction by symmetry simplifies the motion equations of a mechanical system with symmetry by transforming them into new dynamical variables in $\mathfrak{g}^*$ which are invariant under the same Lie group symmetries as the Lagrangian and Hamiltonian of the dynamics. More specifically, on the Lagrangian side, the new invariant variables under the Lie symmetries are obtained from Noether’s theorem, via the tangent lift of the infinitesimal action of the Lie symmetry group on the configuration manifold. The unreduced Euler–Lagrange equations are replaced by equivalent Euler–Poincaré equations expressed in the new invariant variables in $\mathfrak{g}^*$, plus an auxiliary reconstruction equation, which restores the information in the tangent space of the configuration space lost in transforming to group invariant dynamical variables. On the Hamiltonian side, after a Legendre transformation, equivalent new invariant variables in $\mathfrak{g}^*$ are defined by the momentum map $J : T^*M \rightarrow \mathfrak{g}^*$ from the phase-space T^*M of the original system on the configuration manifold M to the dual $\mathfrak{g}^*$ of the Lie symmetry algebra $\mathfrak{g} \simeq T_e G$, via the cotangent lift of the infinitesimal action of the Lie symmetry group on the configuration manifold. The cotangent lift momentum map is an equivariant Poisson map which reformulates the canonical Hamiltonian flow equations in phase space as noncanonical Lie–Poisson equations governing flow of the momentum map on an orbit of the coadjoint action of the Lie symmetry group on the dual of its Lie algebra $\mathfrak{g}^*$, plus an auxiliary reconstruction equation for lifting the Lie group reduced coadjoint motion back to phase space T^*M .

Thus, Lie symmetry reduction yields coadjoint motion of the corresponding momentum map. The dimension of the dynamical system reduces because its solutions are restricted to remain on certain subspaces of the original phase space, called coadjoint orbits. These are orbits of the action of the group G on $\mathfrak{g}^*$, the dual space of its Lie algebra $\mathfrak{g}$. Coadjoint orbits lie on level sets of the distinguished smooth functions $C \in \mathcal{F} : \mathfrak{g}^* \rightarrow \mathbb{R}$ of the symmetry-reduced dual Lie algebra variables $\mu \in \mathfrak{g}^*$ called Casimir functions. Thus, the Casimir functions are conserved quantities. Indeed, Casimir functions have null Lie–Poisson brackets $\{C, F\}(\mu) = 0$ with any other functions $F \in \mathcal{F}(\mathfrak{g}^*)$, including the reduced Hamiltonian $h(\mu)$. Furthermore, level sets of the Casimirs, on which the coadjoint orbits lie, are symplectic manifolds which provide the framework on which geometric mechanics is constructed. These symplectic manifolds have many applications in physics, as well as in symplectic geometry, whenever Lie symmetries are present. In particular, coadjoint motion of the momen-

tum map $J(t) = \text{Ad}_{g(t)}^* J(0)$ for a solution curve $g(t) \in C(G)$ takes place on the intersections of level sets of the Casimirs with level sets of the Hamiltonian.

Given this framework for Lie group reduction by symmetry in deterministic geometric mechanics, we seek strategies for adding stochasticity and dissipation in classical mechanical systems with symmetry which preserve the coadjoint motion structure of the unperturbed deterministic dynamics. Specifically, we seek stochastic coadjoint motion equations whose solutions do not preserve the energy while lying on the coadjoint orbits of the unperturbed equation. Consequently, our first goal in this paper will be to replace the deterministic equations for coadjoint motion by stochastic processes whose solutions lie on coadjoint orbits. However, simply inserting additive noise into the deterministic equations will not, in general, produce coadjoint motion on level sets of the Casimirs of a Lie–Poisson bracket. Instead, our approach in developing a systematic derivation of stochastic deformations that preserve coadjoint orbits will be to constrain the variations in Hamilton’s principle to preserve the transport relations for infinitesimal transformations defined by the action of a stochastic vector field on the configuration manifold, as in [Holm \(2015\)](#).

Having used the constrained Hamilton’s principle to derive the stochastic coadjoint motion equation, the study of the associated Fokker–Planck equation and its invariant measure will follow naturally, and be well defined, at least provided one restricts to finite dimensional mechanical systems. The resulting Fokker–Planck equation defines a probability density for coadjoint motion on Casimir surfaces since it takes the form of a Lie–Poisson equation for the transport part, and a double Lie–Poisson structure for the diffusion part, both of which generate motion along coadjoint orbits. As we will discover, this form of the Fokker–Planck equation in the absence of any additional energy dissipation will imply that the invariant measure (asymptotically in time) simply tends to a constant on Casimir surfaces.

Next, we shall include an energy dissipation mechanism, called double bracket dissipation, or selective decay, which preserves the coadjoint orbits, while it decays the energy towards its minimum value, usually associated with an equilibrium state of the deterministic system. We refer to [Bloch et al. \(1996\)](#), [Gay-Balmaz and Holm \(2013\)](#) and [Gay-Balmaz and Holm \(2014\)](#) and references therein for complete studies of double bracket dissipations. In a second step, we will include this double bracket dissipative term in our stochastic coadjoint motion equations and again study the associated Fokker–Planck equation and its invariant measure, which will no longer be constant. Instead, the invariant measure will be an exponential function of the energy (Gibbs measure).

The procedure we shall follow will produce stochastic dissipative dynamical systems on coadjoint orbits. The study of multiplicative noise and nonlinear dissipation in these systems is greatly facilitated by the geometric structure of the equations for coadjoint motion. Indeed, a large part of standard dynamical system theory will still apply in our setting. In particular, the proof of the existence of random attractors will follow a standard approach. We will mainly focus on this particular feature of random attractors of our systems, as it is an important diagnostic and has recently been an active field of research. The main idea behind the random attractor is the decomposition of the stationary distribution of the Fokker–Planck equation into random measures, called Sinai-Ruelle-Bowen, or SRB measures, whose expectation recovers

the stationary distribution of the Fokker–Planck equation. See, e.g., [Young \(2002\)](#) for a short insightful review. Random attractors can also help in understanding the notion of reliability in complex dynamical systems, see for example [\(Lin et al. 2009\)](#). These ideas have recently been developed and applied actively in the context of climate studies. For example, see [Chekroun et al. \(2011\)](#) for discussions and illustrations of how the notion of random SRB measures and random attractors enable the investigation of detailed geometric structures of the random attractors associated with nonlinear, stochastically perturbed systems. In particular, high-resolution numerical studies of two idealised models of interest for climate dynamics (the Jin97 ENSO model and the Lorenz63 model) are reported in [Chekroun et al. \(2011\)](#). The present work follows a similar line of investigation for a class of nonlinear, stochastically perturbed systems which exhibit coadjoint motion. The proof of existence of non-singular SRB measures requires some work, but it can be accomplished for our general class of mechanical systems written on semi-simple Lie algebras. Although geometric mechanics can also describe infinite dimensional systems such as fluid mechanics ([Holm et al. 1998](#)), we will only focus here on finite dimensional systems, and in particular on systems described by semi-simple Lie algebras. The natural non-degenerate and bi-invariant pairing admitted by semi-simple Lie algebras will facilitate the computations involved in proving our results, although some of the results may still apply more generally.

In the Euler–Poincaré theory, introducing a parameter into the Lagrangian or Hamiltonian which breaks the symmetry results in a semidirect product of groups acting on coset spaces representing the order parameters, or advected quantities, which are not invariant under the original symmetry group ([Holm et al. 1998](#)). We will apply the theory of stochastic deformations that preserve coadjoint orbits for a particular class of semidirect product systems whose advected quantities live in the underlying vector space of the Lie algebra $\mathfrak{g}$. With this particular structure, which can be viewed as a generalisation of the heavy top, we will be able to prove the existence of SRB measures, provided the top Lyapunov exponent is positive. Although much of the present theory may also apply to more general systems than we treat here, as a first investigation, we will show that these particular mechanical systems in geometric mechanics exhibit interesting random attractors on coadjoint orbits when both noise and a certain type of double bracket dissipation are included.

As illustrations, we will discuss in detail two canonical elementary examples in the science of stochastic dissipative geometric mechanics. These two examples are the rigid body and the heavy top, which are also the well-known canonical examples for understanding symmetry reduction for deterministic geometric mechanics ([Marsden and Ratiu 1999](#); [Holm 2008](#); [Holm et al. 2009](#)). As mentioned earlier, the extensions here to include stochasticity and dissipation which preserve coadjoint orbits may be regarded as natural counterparts for geometric mechanics of the standard nonlinear dissipative systems, such as the stochastic Lorenz systems, treated, e.g., in [Chekroun et al. \(2011\)](#), [Kondrashov et al. \(2015\)](#).

Main contributions of this work Section 2 uses the Clebsch approach of [Holm \(2015\)](#) to introduce noise into the Euler–Poincaré equation for the momentum map, including its extension for semidirect product Lie symmetry groups. By construction, the momentum map for the stochastic vector field is the same as that for the determinis-

tic vector field, so the stochastic and deterministic Euler–Poincaré equations for the momentum map may be compared directly. Section 3 introduces the selective decay mechanism for dissipation and studies the existence of random attractors. The first example of the Euler–Poincaré equation is treated in Sect. 5 with the free rigid body. Section 6 treats the Heavy Top as an example of the semidirect product extension. Finally, Sect. 7 briefly sketches the treatments of two other examples, the $SO(4)$ free rigid body and the spring pendulum.

2 Structure Preserving Stochastic Mechanics

Stochastic Hamilton equations were introduced along parallel lines with the deterministic canonical theory in Bismut (1982). These results were later extended to include reduction by symmetry in Lázaro-Camí and Ortega (2008). Reduction by symmetry of expected-value stochastic variational principles for Euler–Poincaré equations was developed in Arnaudon et al. (2014) and Chen et al. (2015). Stochastic variational principles were also used in constructing stochastic variational integrators in Bou-Rabee and Owhadi (2009).

The inclusion of noise in fluid equations has a long history in the scientific literature. For reviews and recent advances in stochastic turbulence models, see Kraichnan (1994) and Gawedzki and Kupiainen (1996); and in the analysis of stochastic Navier–Stokes equations, see Mikulevicius and Rozovskii (2001). These studies of the stochastic Navier–Stokes equation are fundamental in the analysis of fluid turbulence. Expected-value stochastic variational principles leading to the derivation of the Navier–Stokes motion equation for incompressible viscous fluids have been investigated in Arnaudon and Cruzeiro (2012). For further references, we refer to Holm (2015), on which the present work is based. This work has recently had a sequence of further developments, which we now briefly sketch.

- (1) In Holm (2015), the extension of geometric mechanics to include stochasticity in nonlinear fluid theories was accomplished by using Hamilton’s variational principle, constrained to enforce stochastic Lagrangian fluid trajectories arising from the stochastic Eulerian vector field

$$v(x, t, dW) := u(x, t) dt + \sum_{i=1}^N \xi_i(x) \circ dW^i(t), \quad (2.1)$$

regarded as a decomposition into the sum of a drift velocity $u(x, t)$ and a sum over stochastic terms. Imposing this decomposition as a constraint on the variations in Hamilton’s principle for fluid dynamics (Holm et al. 1998), led in Holm (2015) to applications in a variety of fluid theories, particularly for geophysical fluid dynamics (GFD).

- (2) The same stochastic fluid dynamics equations derived in Holm (2015) were also discovered in Cotter et al. (2017) to arise in a multi-scale decomposition of the deterministic Lagrange-to-Euler flow map, into a slow large-scale mean and a rapidly fluctuating small-scale map. Homogenisation theory was used to derive

effective slow stochastic particle dynamics for the resolved mean part, thereby justifying the stochastic fluid partial equations in the Eulerian formulation. The application of rigorous homogenisation theory required assumptions of mildly chaotic fast small-scale dynamics, as well as a centring condition, according to which the mean of the fluctuating deviations was small, when pulled back to the mean flow.

Paper (Cotter et al. 2017) justified regarding the Eulerian vector field in (2.1) as a genuine decomposition of the fluid velocity into a sum of drift and stochastic parts, rather than simply a perturbation of the dynamics meant to model unknown effects in uncertainty quantification. As a genuine decomposition of the solution, one should expect that the properties of the fluid equations with stochastic transport noise should closely track the properties of the unapproximated solutions of the fluid equations. For example, if the unapproximated model equations are Hamiltonian, then the model equations with stochastic transport noise should also be Hamiltonian, as shown in Holm (2015).

- (3) Paper (Crisan et al. 2017) showed that the same stochastic fluid dynamics derived in Holm (2015) naturally arises from an application of a stochastic Lagrange-to-Euler map to Newton's second law for a Lagrangian domain of fluid, acted on by an external force. In addition, local well-posedness in regular spaces and a Beale-Kato-Majda blow-up criterion are proved in Crisan et al. (2017) for the stochastic model of the 3D Euler fluid equation for incompressible flow derived in Holm (2015). Thus, the analytical properties of the 3D Euler fluid equations with stochastic transport noise derived in Holm (2015) closely mimic the corresponding analytical properties of the original deterministic 3D Euler fluid equations.
- (4) Inspired by spatiotemporal observations from satellites of the trajectories of objects drifting near the surface of the ocean in the National Oceanic and Atmospheric Administration's "Global Drifter Program", paper (Gay-Balmaz and Holm 2017) developed data-driven stochastic models of geophysical fluid dynamics (GFD) with non-stationary spatial correlations representing the dynamical behaviour of oceanic currents. These models were derived using reduction by symmetry of stochastic variational principles, leading to stochastic Hamiltonian systems, whose momentum maps, conservation laws and Lie–Poisson bracket structures were used in developing the new stochastic Hamiltonian models of GFD with nonlinearly evolving stochastic properties.

The present section incorporates stochasticity into finite dimensional mechanical systems admitting Lie group symmetry reduction, by using the standard Clebsch variational method for deriving cotangent lifted momentum maps. We review the standard approach to Lie group reduction by symmetry for finite dimensional systems in 2.1 and incorporate noise into this approach in Sect. 2.2. Next, we describe the semidirect extension in 2.3 and study the associated Fokker–Planck equations and stationary distribution in 2.4. The primary examples from classical mechanics with symmetry will be the free rigid body and the heavy top under gravity.

2.1 Euler–Poincaré Reduction

Classical mechanical systems with symmetry can often be understood geometrically in the context of Lagrangian or Hamiltonian reduction, by lifting the motion $m(t)$ on the configuration manifold M to a Lie symmetry group via the action of the symmetry group G on the configuration manifold, by setting $m(t) = g(t)m(0)$, where the multiplication is to be understood as the action of G to M . This procedure lifts the solution of the motion equation from a curve $m(t) \in M$ to a curve $g(t) \in G$, see Marsden and Ratiu (1999) and Holm (2008). The simplest case is when $M = G$. This case, called Euler–Poincaré reduction, will be described in the present section.

In the Lagrangian framework, reduction by symmetry may be implemented in Hamilton’s principle via restricted variations in the reduced variational principle arising from variations on the corresponding Lie group. In the standard approach, for an arbitrary variation δg of a curve $g(t) \in G$ in a Lie group G , the left-invariant reduced variables are $g^{-1}\dot{g} \in \mathfrak{g}$ in the Lie algebra $\mathfrak{g} = T_e G$. Their variations arise from variations on the Lie group and are given by

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

for $\eta := g^{-1}\delta g$. Here, the operation $\text{ad} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ represents the adjoint action of the Lie algebra on itself via the Lie bracket, denoted equivalently as $\text{ad}_{\xi} \eta = [\xi, \eta]$, and we will freely use either notation throughout the text. If the Lagrangian $L(g, \dot{g})$ is left-invariant under the action of G , the restricted variations $\delta \xi$ of the reduced Lagrangian $L(e, g^{-1}\dot{g}) =: l(\xi)$ inherited from admissible variations of the solution curves on the group yield the Euler–Poincaré equation

$$\frac{d}{dt} \frac{\partial l(\xi)}{\partial \xi} + \text{ad}_{\xi}^* \frac{\partial l(\xi)}{\partial \xi} = 0 . \quad (2.2)$$

In this equation, $\text{ad}^* : \mathfrak{g} \times \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ is the dual of the adjoint Lie algebra action, ad . That is, $\langle \text{ad}_{\xi}^* \mu, \eta \rangle = \langle \mu, \text{ad}_{\xi} \eta \rangle$ for $\mu \in \mathfrak{g}^*$ and $\xi, \eta \in \mathfrak{g}$, under the non-degenerate pairing $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g}^* \rightarrow \mathbb{R}$. Throughout this paper, we will restrict ourselves to semi-simple Lie algebras, so that the pairing is given by the Killing form, defined as

$$\kappa(\xi, \eta) := \text{Tr}(\text{ad}_{\xi} \text{ad}_{\eta}) . \quad (2.3)$$

In terms of the structure constants of the Lie algebra denoted as c_{jk}^i for a basis e_i , $i = 1, \dots, \dim(\mathfrak{g})$, so that $[e_i, e_j] = c_{jk}^i e_i$, in which $\xi = \xi^i e_i$ and $\eta = \eta^j e_j$, the Killing form takes the explicit form

$$\text{Tr}(\text{ad}_{\xi} \text{ad}_{\eta}) = c_{im}^n c_{jn}^m \xi^i \eta^j .$$

An important property of this pairing is its bi-invariance, written as

$$\kappa(\xi, \text{ad}_{\zeta} \eta) = \kappa(\text{ad}_{\xi} \zeta, \eta) , \quad (2.4)$$

for every $\xi, \eta, \zeta \in \mathfrak{g}$. This pairing allows us to identify the Lie algebra with its dual, as the Killing form of semi-simple Lie algebras is non-degenerate. We will also use the property for compact Lie algebras, that the Killing form is negative definite and thus induces a norm on the Lie algebra, $\|\xi\|^2 := -\kappa(\xi, \xi)$. This function turns out to always be an invariant function on the coadjoint orbit, for every semi-simple Lie algebra; that is, it is a Casimir function. Indeed, an invariant function is in the kernel of the Lie–Poisson bracket $\{F, C\}(\mu) := \kappa\left(\mu, \left[\frac{\partial F}{\partial \mu}, \frac{\partial C}{\partial \mu}\right]\right)$, $\forall F \in C(\mathfrak{g}^*, \mathbb{R})$. For $C(\mu) = \frac{1}{2}\kappa(\mu, \mu)$ it is straightforward to check using the bi-invariance (2.4) that this is true for any function F . In general, it is difficult to find other independent Casimir functions of semi-simple Lie algebra; see [Perelomov and Popov \(1968\)](#) and [Thiffeault and Morrison \(2000\)](#). Of course, the theory of semi-simple Lie algebras is standard and well developed, see for example [Varadarajan \(1984\)](#). However, for the sake of clarity, we will express the abstract notations of adjoint and coadjoint actions with respect to the Killing form. We may then identify $\mathfrak{g}^* \cong \mathfrak{g}$ for each semi-simple Lie algebra we treat here.

We now turn to the equivalent Clebsch formulation of the Euler–Poincaré equations via a constrained Hamilton’s principle, which we will use for implementing the noise in these systems. The Clebsch formulation of the Euler–Poincaré equation and its corresponding Lie–Poisson bracket on the Hamiltonian side has been explored extensively in ideal fluid dynamics ([Holm and Kupershmidt 1983](#); [Marsden and Weinstein 1983](#)) and more recently in optimal control problems [Gay-Balmaz and Ratiu \(2011\)](#) and stochastic fluid dynamics [Holm \(2015\)](#). This earlier work should be consulted for detailed derivations of Clebsch formulations of Euler–Poincaré equations in the contexts of ideal fluids and optimal control problems. We will briefly sketch the Clebsch approach, as specialised to the applications treated here; since we will rely on it for the introduction of noise in finite-dimensional mechanical systems by following the approach of [Holm \(2015\)](#) for stochastic fluid dynamics. We first introduce the Clebsch variables $q \in \mathfrak{g}$ and $p \in \mathfrak{g}^*$, where p will be a Lagrange multiplier which enforces the dynamical evolution of q given by the Lie algebra action of $\xi \in \mathfrak{g}$, as $\dot{q} + \text{ad}_\xi q = 0$. Note the similarity of this equation with the constrained variations of the Lagrangian reduction theory. The Clebsch method in fluid dynamics (resp. optimal control) introduces auxiliary equations for advected quantities (resp. Lie algebra actions on state variables) as constraints in the Hamilton (resp. Hamilton–Pontryagin) variational principle $\delta S = 0$ with constrained action

$$S(\xi, q, p) = \int l(\xi) dt + \int \langle p, \dot{q} + \text{ad}_\xi q \rangle dt. \quad (2.5)$$

Taking free variations of S with respect to ξ, q and p yields a set of equations for these three variables which can be shown to be equivalent to the Euler–Poincaré equation (2.2). The relation between the Lie algebra vector $\xi \in \mathfrak{g}$ and the phase-space variables $(q, p) \in T_e^*G$ is given by the variation of the action S in (2.5) with respect to the velocity ξ in (2.5). This variation yields the momentum map, $\mu : T_e^*G \rightarrow \mathfrak{g}^*$, given explicitly by

$$\mu := \frac{\partial l(\xi)}{\partial \xi} = \text{ad}_q^* p. \quad (2.6)$$

Unless specified otherwise, we will always use the notation μ for the conjugate variable to ξ . This version of the Clebsch theory is especially simple, as the Clebsch variables are also in the Lie algebra $\mathfrak{g}$. In general, it is enough for them to be in the cotangent bundle of a manifold T^*M on which the group G acts by cotangent lifts. In this more general case, the adjoint and coadjoint actions must be replaced by their corresponding actions on T^*M , but the method remains the same. Another generalisation, which will be useful for us later, allows the Lagrangian to depend on both ξ and q . In this case, the Euler–Poincaré equation will acquire additional terms depending on q and the Clebsch approach will be equivalent to semidirect product reduction (Holm et al. 1998). We will consider a simple case of this extension in Sect. 2.3 and in the treatment of the heavy top in Sect. 6.

2.2 Structure Preserving Stochastic Deformations

We are now ready to deform the Euler–Poincaré equation (2.2) by introducing noise in the constrained Clebsch variational principle (2.5). In order to do this stochastic deformation, we introduce n independent Wiener processes W_t^i indexed by $i = 1, 2, \dots, n$, and their associated stochastic potential fields $\Phi_i(q, p) \in \mathbb{R}$ which are prescribed functions of the Clebsch phase-space variables, (q, p) . The stochastic processes used here are standard Wiener processes, as discussed, e.g., in Chen et al. (2015) and Ikeda and Watanabe (2014). The number of stochastic processes can be arbitrary, but we will assume it is equal to the dimension of the Lie algebra, $n = \dim(\mathfrak{g})$. The constrained stochastic variational principle is then given by

$$S(\xi, q, p) = \int l(\xi) dt + \int \langle p, \circ dq + \text{ad}_\xi q \, dt \rangle + \int \sum_{i=1}^n \Phi_i(q, p) \circ dW_t^i. \quad (2.7)$$

In the stochastic action integral (2.7) and hereafter, the multiplication symbol $\circ$ denotes a stochastic integral in the Stratonovich sense (Kloeden and Platen 1991). The Stratonovich formulation is the only choice of stochastic integral that admits the classical rules of calculus (e.g., integration by parts, the change of variables formula, etc.). Therefore, the Stratonovich formulation is indispensable in variational calculus and optimal control. The free variations of the action functional (2.7) may now be taken, and they will yield stochastic processes for the three variables ξ , q and p .

For convenience in the next step of deriving a stochastic Euler–Poincaré equation, we will assume that the Lagrangian $l(\xi)$ in the action (2.7) is hyperregular, so that ξ may be obtained from the fibre derivative $\frac{\partial l(\xi)}{\partial \xi} = \text{ad}_q^* p$. We will also specify that the stochastic potentials $\Phi_i(q, p)$ should depend only on the momentum map $\mu = \text{ad}_q^* p$, so that the resulting stochastic equation will be independent of q and p . Following the detailed calculations in Holm (2015), we then find the stochastic Euler–Poincaré equation

$$d \frac{\partial l(\xi)}{\partial \xi} + \text{ad}_\xi^* \frac{\partial l(\xi)}{\partial \xi} dt - \sum_i \text{ad}_{\frac{\partial \Phi_i(\mu)}{\partial \mu}}^* \frac{\partial l(\xi)}{\partial \xi} \circ dW_t^i = 0. \quad (2.8)$$

In terms of the stochastic process

$$dX = \xi dt - \sum_i \frac{\partial \Phi_i(\mu)}{\partial \mu} \circ dW_t^i, \quad \text{with } \mu = \frac{\partial l(\xi)}{\partial \xi}, \quad (2.9)$$

the stochastic Euler–Poincaré equation (2.8) may be expressed in compact form, as

$$d\mu + \text{ad}_{dX}^* \mu = 0. \quad (2.10)$$

The introduction of noise in the Clebsch-constrained variational principle rather than using reduction theory provides a transparent approach for dealing with stochastic processes on Lie groups and constrained variations arising for such processes. In this approach, the momentum map plays the same central role in both the deterministic and stochastic formulations. See [Arnaudon et al. \(2014\)](#) for a different approach, resulting in the derivation and analysis of deterministic expectation-value Euler–Poincaré equations using reduction by symmetry with conditional expectation.

Remark 2.1 (Reduction with noise) The stochastic Euler–Poincaré equation (2.8) arises from a stochastic reduction by symmetry, as follows. First, the reconstruction relation $\dot{g} = g\xi$ in the deterministic case has its stochastic counterpart

$$dg = g\xi dt + \sum_i g\sigma_i \circ dW_t^i, \quad (2.11)$$

where $\sigma_i := -\frac{\partial \Phi_i(\mu)}{\partial \mu}$, and the expressions $g\xi$ and $g\sigma_i$ are understood as the tangent of the left action of the group on itself; or equivalently, the left action of the group on its Lie algebra. Equation (2.8) then results from taking the variation of $g^{-1}dg$ with (2.11) and comparing with the derivative of $g^{-1}\delta g$ while setting $d(\delta g) = \delta(dg)$. This gives the variation $\delta\xi$ as

$$\delta\xi = d\eta + \text{ad}_{(g^{-1}dg)}\eta = d\eta + \text{ad}_\xi \eta dt + \sum_i \text{ad}_{\sigma_i} \eta \circ dW_t^i, \quad (2.12)$$

where $\eta = g^{-1}\delta g \in \mathfrak{g}$ is arbitrary, except that $\delta g(0) = 0 = \delta g(T)$ at the endpoints in time $t \in [0, T]$. Then, using these constrained variations in the reduced variational principle $\delta \int l(\xi) dt = 0$ yields the stochastic Euler–Poincaré equation (2.8), by the following calculation,

$$\begin{aligned} 0 &= \delta \int l(\xi) dt = \int \left\langle \frac{\delta l}{\delta \xi}, \delta \xi \right\rangle dt = \int \left\langle \frac{\delta l}{\delta \xi}, d\eta + \text{ad}_{(g^{-1}dg)}\eta \right\rangle dt \\ &= \int \left\langle -d\frac{\delta l}{\delta \xi} + \text{ad}_{(g^{-1}dg)}^* \frac{\delta l}{\delta \xi}, \eta \right\rangle dt + \left\langle \frac{\delta l}{\delta \xi}, \eta \right\rangle \Big|_0^T \end{aligned} \quad (2.13)$$

upon imposing the condition that η vanishes at the endpoints in time.

As in the deterministic case, various generalisations of this theory are possible. For example, as mentioned earlier, the Clebsch phase-space variables can also be defined in T^*M , and the Lagrangian can depend on q for systems of semidirect product type (Gay-Balmaz and Holm 2017). Another generalisation is to let the stochastic potentials $\Phi_i(\mu)$ also depend separately on q in the semidirect product setting, as we will see later.

After having defined the Stratonovich stochastic process (2.8), one may compute its corresponding Itô form, which is readily given in terms of the ad^* operation by

$$d\frac{\partial l(\xi)}{\partial \xi} + \text{ad}_\xi^* \frac{\partial l(\xi)}{\partial \xi} dt + \sum_i \text{ad}_{\sigma_i}^* \frac{\partial l(\xi)}{\partial \xi} dW_t^i - \frac{1}{2} \sum_i \text{ad}_{\sigma_i}^* \left(\text{ad}_{\sigma_i}^* \frac{\partial l(\xi)}{\partial \xi} \right) dt = 0, \quad (2.14)$$

where $\sigma_i := -\frac{\partial \Phi_i(\mu)}{\partial \mu}$. Notice that the indices for σ_i in the Itô sum in (2.14) are the same, and the σ_i may be taken as a basis of the underlying vector space. In terms of $\mu := \frac{\partial l(\xi)}{\partial \xi}$ the Itô stochastic Euler–Poincaré equation (2.14) may be expressed equivalently as

$$d\mu + \text{ad}_\xi^* \mu dt + \sum_i \text{ad}_{\sigma_i}^* \mu dW_t^i - \frac{1}{2} \sum_i \text{ad}_{\sigma_i}^* (\text{ad}_{\sigma_i}^* \mu) dt = 0. \quad (2.15)$$

Another formulation of the stochastic Euler–Poincaré equation in (2.8) which will be used later in deriving the Fokker–Planck equation is the stochastic Lie–Poisson equation

$$df(\mu) = \left\langle \mu, \left[\frac{\partial f}{\partial \mu}, \frac{\partial h}{\partial \mu} \right] \right\rangle dt + \sum_i \left\langle \mu, \left[\frac{\partial f}{\partial \mu}, \frac{\partial \Phi_i}{\partial \mu} \right] \right\rangle \circ dW_t^i \quad (2.16)$$

$$=: \{f, h\}dt + \sum_i \{f, \Phi_i\} \circ dW_i, \quad (2.17)$$

where the Lie–Poisson bracket $\{\cdot, \cdot\}$ is defined just as in the deterministic case, from the adjoint action and the pairing on the Lie algebra $\mathfrak{g}$.

2.3 Extension to Semidirect Product Systems

As discussed in Holm et al. (1998), “It turns out that semidirect products occur under rather general circumstances when the symmetry in T^*G is broken”. The geometric mechanism for this remarkable fact is that under reduction by symmetry, a semidirect product of groups emerges whenever the symmetry in the phase space is broken. The symmetry breaking produces new dynamical variables, living in the coset space formed from taking the quotient G/G_a of the original unbroken symmetry G by the remaining symmetry G_a under the isotropy subgroup of the new variables. These new dynamical variables form a vector space $G/G_a \simeq V$ on which the unbroken symmetry acts as a semidirect product, $G \ltimes V$. In physics, elements of the vector space V corresponding to the new variables are called “order parameters”. Typically, in physics, the original

symmetry is broken by the introduction of potential energy depending on variables which reduce the symmetry to the isotropy subgroup of the new variables. Dynamics on the semidirect product $G \ltimes V$ results, and what had previously been flowed under the action of the unbroken symmetry now becomes flow plus waves, or oscillations, produced by the exchange of energy between its kinetic and potential forms. The heavy top is the basic example of this phenomenon, and it will be treated in Sect. 6. The semidirect product motion for the heavy top arises in the presence of gravity when the support point of a freely rotating rigid body is shifted away from its centre of mass.

With this connection between symmetry breaking and semidirect products in mind, we now extend the stochastic Euler–Poincaré equations to include semidirect product systems. We refer to [Holm et al. \(1998\)](#) for a complete study of these systems. Although the deterministic equations of motion in [Holm et al. \(1998\)](#) are derived from reduction by symmetry, we will instead incorporate noise by simply extending the Clebsch-constrained variational principle used in the previous section.

The generalisation proceeds, as follows. We will begin by assuming that the Clebsch phase-space variables comprise the elements of T^*V for a given vector space V on which the Lie group G acts freely and properly. In fact we will have $(q, p) \in V \times V^*$. However, in this work, we will restrict ourselves to the case where V is the underlying vector space of $\mathfrak{g}$. Following the notation of [Ratiu \(1981\)](#), we denote $\bar{\mathfrak{g}} = V$ in the sequel. Then, from the Killing form on $\mathfrak{g}$, denoted by $\kappa : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$, there is a bi-invariant extension of the Killing form on $\mathfrak{g} \ltimes \bar{\mathfrak{g}}$ defined as

$$\kappa_s((\xi_1, \xi_2), (\eta_1, \eta_2)) := \kappa(\xi_1, \eta_2) + \kappa(\xi_2, \eta_1). \quad (2.18)$$

Although this pairing is non-degenerate and bi-invariant, we will not use it for the definition of the dual of the semidirect algebra $\mathfrak{g} \ltimes V$. Instead, we will use the sum of both Killing forms, namely

$$\kappa_0((\xi_1, \xi_2), (\eta_1, \eta_2)) := \kappa(\xi_1, \eta_1) + \kappa(\xi_2, \eta_2). \quad (2.19)$$

The group action is defined via the adjoint representation of G on $V = \bar{\mathfrak{g}}$, given by $(g_1, \eta_1)(g_2, \eta_2) = (g_1 g_2, \eta_1 + \text{Ad}_{g_1} \eta_2)$. We then directly have the infinitesimal adjoint and coadjoint actions as

$$\begin{aligned} \text{ad}_{(\xi_1, q_1)}(\xi_2, q_2) &= (\text{ad}_{\xi_1} \xi_2, \text{ad}_{\xi_1} q_2 + \text{ad}_{q_1} \xi_2), \\ \text{ad}_{(\xi, q)}^*(\mu, p) &= (\text{ad}_{\xi}^* \mu + \text{ad}_q^* p, \text{ad}_{\xi}^* p), \end{aligned} \quad (2.20)$$

where the coadjoint action is taken with respect to κ_0 in (2.19). The extended Killing form κ_s defined in (2.18), gives, apart from $\kappa(\eta, \eta)$ with $\eta \in \bar{\mathfrak{g}}$, a second invariant function on the coadjoint orbit

$$\kappa_s((\xi, \eta), (\xi, \eta)) = 2\kappa(\xi, \eta),$$

found from the bi-invariance of the Killing form κ_s . One then replaces the corresponding Lie algebra actions in the Clebsch-constrained variational principle (2.7), to obtain the stochastic process with semidirect product

$$d(\mu, q) + \text{ad}_{(\xi, r)}^*(\mu, q) dt + \sum_i \text{ad}_{\left(\frac{\partial \Phi_i(\mu, q)}{\partial \mu}, \frac{\partial \Phi_i(\mu, q)}{\partial q}\right)}^*(\mu, q) \circ dW_t^i = 0, \quad (2.21)$$

where $l : \mathfrak{g} \otimes V \rightarrow \mathbb{R}$, $\Phi_i : \mathfrak{g}^* \rightarrow \mathbb{R}$,

$$\frac{\partial l(\xi, q)}{\partial \xi} =: \mu \quad \text{and} \quad \frac{\partial l(\xi, q)}{\partial q} =: r. \quad (2.22)$$

Remark 2.2 (q dependence on the stochastic potentials) Notice that the stochastic potentials do not depend on the advected variables. This is a consequence of the Clebsch construction. By using the Hamilton-Pontryagin derivation of this equation, the potentials may depend on the advected variables, see the remark 3.2 or Gay-Balmaz and Holm (2017) for more details. We will write here the dependence on the advected variables for completeness, but we will not use it in the examples.

After taking the Legendre transform of reduced Lagrangian l , we have the Hamiltonian derivatives

$$\frac{\partial h(\mu, q)}{\partial \mu} =: \xi \quad \text{and} \quad \frac{\partial h(\mu, q)}{\partial q} =: -r, \quad (2.23)$$

for $h : \mathfrak{g}^* \otimes V^* \rightarrow \mathbb{R}$. By substituting into (2.21) the expressions in (2.20) for the coadjoint action, we obtain the system

$$\begin{aligned} d\mu + \left(\text{ad}_{\xi}^* \mu + \text{ad}_r^* q \right) dt + \sum_i \left(\text{ad}_{\frac{\partial \Phi_i(\mu, q)}{\partial \mu}}^* \mu + \text{ad}_{\frac{\partial \Phi_i(\mu, q)}{\partial q}}^* q \right) \circ dW_t^i &= 0, \\ dq + \text{ad}_{\xi}^* q dt + \sum_i \text{ad}_{\frac{\partial \Phi_i(\mu, q)}{\partial \mu}}^* q \circ dW_t^i &= 0. \end{aligned} \quad (2.24)$$

Although the number of stochastic potentials Φ_i which one may consider is arbitrary, we shall find it convenient for our purposes to restrict to a maximum of $n = \dim(\mathfrak{g}) + \dim(V)$ such potentials. In fact, the potentials associated with V will not be fully treated here.

Remark 2.3 General setting The semidirect product theory we have described here is the simplest instance of it, as we are using a particular vector space V . In general, the advected quantities can also be in a Lie algebra, or an arbitrary manifold, provided the action of the group G on it is free and proper.

2.4 The Fokker–Planck Equation and Stationary Distributions

We derive here a geometric version of the classical Fokker–Planck equation (or forward Kolmogorov equation) using our SDE (2.8). Recall that the Fokker–Planck equation describes the time evolution of the probability distribution $\mathbb{P}$ for the process driven by (2.8). We refer to Ikeda and Watanabe (2014) for the standard textbook treatment of stochastic processes. Here, we will consider $\mathbb{P}$ as a function $C(\mathfrak{g}^*)$ with the additional

property that $\int_{\mathfrak{g}^*} \mathbb{P} d\mu = 1$. First, the generator of the process (2.8) can be readily found from the Lie–Poisson form (2.14) of the stochastic process (2.17) to be

$$Lf(\mu) = \left\langle \text{ad}_{\xi}^* \mu, \frac{\partial f}{\partial \mu} \right\rangle - \sum_i \left\langle \text{ad}_{\sigma_i}^* \mu, \frac{\partial}{\partial \mu} \left\langle \text{ad}_{\sigma_i}^* \mu, \frac{\partial f}{\partial \mu} \right\rangle \right\rangle, \quad (2.25)$$

where $\langle \cdot, \cdot \rangle$ still denotes the Killing form on the Lie algebra $\mathfrak{g}$ and $f \in C(\mathfrak{g}^*)$ is an arbitrary function of μ . The Fokker–Planck equation

$$\partial_t \mathbb{P} = L^* \mathbb{P}$$

obtained from the formula (2.25) is proved in the following proposition, cf. Holm et al. (2016).

Proposition 2.4 *The forward Kolmogorov operator is*

$$L^* \mathbb{P}(\mu) = -\{h, \mathbb{P}\} - \frac{1}{2} \sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}\}\}. \quad (2.26)$$

Proof We first consider the drift term of (2.25) written in Lie-Bracket form and use the Leibniz property of the Lie–Poisson bracket to get

$$\int \mathbb{P}\{h, f\} d\mu = - \int f\{h, \mathbb{P}\} d\mu + \int \{h, f\mathbb{P}\} d\mu. \quad (2.27)$$

Recall that a non-degenerate Poisson bracket can be expressed in term of a symplectic form as $\{F, G\} = \omega(X_F, X_G)$ where X_F and X_G are the Hamiltonian vector fields associated with the function F and G . Now, the symplectic form on the coadjoint orbit is an exact 2-form, that is $\omega = d\theta$ for some 1-form θ . These facts mean that provided that we have vanishing or periodic boundary conditions on the coadjoint orbits, the integration of the Lie–Poisson bracket corresponds to the integration of the exterior differential of a 1-form. This integral vanishes by using Stokes theorem and suitable boundary conditions. For the double bracket term, we similarly obtain

$$\begin{aligned} \int \mathbb{P}\{\Phi_i, \{\Phi_i, f\}\} d\mu &= - \int \{\Phi_i, f\}\{\Phi_i, \mathbb{P}\} d\mu + \int \{\Phi_i, \{\Phi_i, f\}\mathbb{P}\} d\mu \\ &= \int f\{\Phi_i, \{\Phi_i, \mathbb{P}\}\} d\mu - \int \{\Phi_i, f\{\Phi_i, \mathbb{P}\}\} d\mu \\ &\quad + \int \{\Phi_i, \{\Phi_i, f\}\mathbb{P}\} d\mu, \end{aligned}$$

where the last two terms vanishes from the same argument as before, that is using integration by parts. Collecting terms give the forward Kolmogorov operator (2.26). $\square$

In (2.26), we recover the Lie–Poisson formulation (2.17) of the Euler–Poincaré equation together with a dissipative term arising from the noise of the original SDE in a double Lie–Poisson bracket form. This formulation gives the following theorem for stationary distributions of (2.25):

Theorem 2.5 *The stationary distribution $\mathbb{P}_\infty$ of the Fokker–Planck equation (2.25), i.e., $L^*\mathbb{P}_\infty = 0$ is uniform on the coadjoint orbits on which the SDE (2.8) evolves.*

Proof By a standard result in functional analysis, see for example (Villani 2009), a linear differential operator of the form $L = B + \frac{1}{2} \sum_i A_i^2$ has the property that $\ker(L) = \ker(A_i) \cap \ker(B)$, where here $A_i = \{\Phi_i, \cdot\}$ and $B = \{h, \cdot\}$. Consequently, for every smooth function f , the only functions g which satisfy $\{f, g\} = 0$ are the Casimirs, or invariant functions, on the coadjoint orbits. When restricted to a coadjoint orbit, these functions become constants. Hence, the stationary distribution $\mathbb{P}_\infty$ is a constant on the coadjoint orbit identified by the initial conditions of the system. $\square$

Since the dynamics are restricted to the coadjoint orbits, for the probability distribution $\mathbb{P}$ to tend to a constant, yet remain normalisable, satisfying $\int_{\mathfrak{g}^*} \mathbb{P}(\mu) d\mu = 1$, the value of the density must tend to the inverse of the volume of the coadjoint orbit. Of course, the compactness of the coadjoint orbit is equivalent to $\mathbb{P}_\infty > 0$. For non-compact orbits, Theorem 2.5 is still valid, and it will imply an asymptotically vanishing stationary distribution, in the same sense as for the stationary solution of the heat equation on the real line. In this case, a more detailed analysis of the stationary distribution can be performed by studying marginals, or projections onto a compact subspace of the coadjoint orbit.

Examples of non-compact coadjoint orbits arise in the semidirect product setting. First, the Fokker–Planck equation for the semidirect product stochastic process (2.21) is given by

$$\begin{aligned} Lf(\mu, q) = & \left\langle \text{ad}_{(\xi, r)}^*(\mu, q), \left(\frac{\partial f(\mu, q)}{\partial \mu}, \frac{\partial f(\mu, q)}{\partial q} \right) \right\rangle - \\ & - \sum_i \left\langle \text{ad}_{(\sigma_i, \eta_i)}^*(\mu, q), \left\{ \frac{\partial}{\partial \mu} \left\langle \text{ad}_{(\sigma_i, \eta_i)}^*(\mu, q), \left(\frac{\partial f(\mu, q)}{\partial \mu}, \frac{\partial f(\mu, q)}{\partial q} \right) \right\rangle \right. \right. \\ & \times \left. \left. \frac{\partial}{\partial q} \left\langle \text{ad}_{(\sigma_i, \eta_i)}^*(\mu, q), \left(\frac{\partial f(\mu, q)}{\partial \mu}, \frac{\partial f(\mu, q)}{\partial q} \right) \right\rangle \right\} \right\rangle, \end{aligned} \quad (2.28)$$

where $\sigma_i := -\partial \Phi_i / \partial \mu$ and $\eta_i := -\partial \Phi_i / \partial q$. The pairing used here is the sum of the pairings on $\mathfrak{g}$ and on V , given by κ_0 in (2.19). Note that for some values of index i , the vector fields σ_i or η_i may be absent. One can check that $L^* = L$; so that L generates the Lie–Poisson Fokker–Planck equation for the probability density $\mathbb{P}(\mu, q)$. As before, upon using the explicit form of the coadjoint actions, one finds

$$\begin{aligned}
Lf(\mu, q) &= \left\langle \text{ad}_\xi^* \mu + \text{ad}_q^* r, \frac{\partial f}{\partial \mu} \right\rangle + \left\langle \text{ad}_\xi^* q, \frac{\partial f}{\partial q} \right\rangle - \\
&\quad - \sum_i \left\langle \text{ad}_{\sigma_i}^* \mu + \text{ad}_q^* \eta_i, \frac{\partial A_i}{\partial \mu} \right\rangle - \sum_i \left\langle \text{ad}_{\sigma_i}^* q, \frac{\partial A_i}{\partial q} \right\rangle, \quad (2.29)
\end{aligned}$$

where $A_i := \left\langle \text{ad}_{\sigma_i}^* \mu + \text{ad}_q^* \eta_i, \frac{\partial f}{\partial \mu} \right\rangle + \left\langle \text{ad}_{\sigma_i}^* q, \frac{\partial f}{\partial q} \right\rangle$.

The Fokker–Planck equation (2.28) provides a direct corollary of Theorem 2.5.

Corollary 2.5.1 *The stationary distribution $\mathbb{P}_\infty(\mu, q)$ of (2.28) is constant on the coadjoint orbit corresponding to the initial conditions of the stochastic process (2.21).*

As mentioned earlier, the coadjoint orbit of this system is not compact, even if it had been compact for the Lie algebra $\mathfrak{g}$. Nevertheless, we can study the marginal distributions

$$\mathbb{P}^1(\mu) := \int \mathbb{P}(\mu, q) dq \quad \text{and} \quad (2.30)$$

$$\mathbb{P}^2(q) := \int \mathbb{P}(\mu, q) d\mu, \quad (2.31)$$

which of course extend to stationary marginal distributions $\mathbb{P}_\infty^1$ and $\mathbb{P}_\infty^2$. With these marginal distributions, we can get more information about the stationary distribution of the semidirect product Lie–Poisson Fokker–Planck equation (2.28), as summarised in the next theorem.

Theorem 2.6 *For a semi-simple Lie algebra $\mathfrak{g}$ and $V = \bar{\mathfrak{g}}$, the marginal stationary distributions defined in (2.30) and (2.31) of the Fokker–Planck equation (2.28), with $\eta_i = 0$, for all i , have the following forms.*

- (1) *The stationary distribution $\mathbb{P}_\infty^2(q)$ is constant on the q -dependent subspace of the coadjoint orbit. If the Lie algebra $\mathfrak{g}$ is non-compact, the constant is zero.*
- (2) *The stationary distribution $\mathbb{P}_\infty^1(\mu)$ restricted to $\kappa(\mu, \mu)$ is constant.*
- (3) *If $\mathfrak{g}$ is compact, the stationary distribution $\mathbb{P}_\infty^1(\mu)$ is linearly bounded in time in the direction perpendicular to $\kappa(\mu, \mu)$.*

Proof We will compute the stationary marginal distributions separately, but first recall that the stationary distribution $\mathbb{P}(\mu, q)$ is constant on the Casimir level sets given by the initial conditions.

- (1) By integrating the Fokker–Planck equation (2.28) over μ , one obtains

$$L\mathbb{P}^2(q) = \int \left\langle \text{ad}_\xi^* q, \frac{\partial \mathbb{P}(\mu, q)}{\partial q} \right\rangle d\mu - \frac{1}{2} \left\langle \text{ad}_{\sigma_i}^* q, \frac{\partial}{\partial q} \left\langle \text{ad}_{\sigma_i}^* q, \frac{\partial \mathbb{P}^2(q)}{\partial q} \right\rangle \right\rangle, \quad (2.32)$$

where we have used the property that the coadjoint action is divergence-free (because of the anti-symmetry of the adjoint action, when identified with the coadjoint action via the Killing form) and have recalled that the Lie algebra is either compact, or $\mathbb{P}(\mu, q) = 0$ for the boundary conditions.

Only the advection term remains in (2.32), as $\xi = \frac{\partial h}{\partial \mu}$ depends on μ . Nevertheless, an argument similar to that for the proof of Theorem 2.5 may be applied here to give the result of constant marginal distribution on the q dependent part of the coadjoint orbits. Again, if the Lie algebra is non-compact, then the probability density $\mathbb{P}_\infty^2(q)$ must vanish because of the normalisation.

(2) We first integrate the Fokker–Planck equation (2.28) with respect to the q variable to find

$$L\mathbb{P}^1(\mu) = \left\langle \text{ad}_\xi^* \mu, \frac{\partial \mathbb{P}^1}{\partial \mu} \right\rangle - \frac{1}{2} \sum_i \left\langle \text{ad}_{\sigma_i}^* \mu, \frac{\partial}{\partial \mu} \left\langle \text{ad}_{\sigma_i}^* \mu, \frac{\partial \mathbb{P}^1}{\partial \mu} \right\rangle \right\rangle, \quad (2.33)$$

where we have again used that the coadjoint action is divergence-free, the same boundary conditions and the fact that $\langle \text{ad}_q \xi, \frac{\partial \mathbb{P}}{\partial \mu} \rangle = 0$, $\forall \xi$ since $\frac{\partial \mathbb{P}}{\partial \mu}$ is aligned with q . Indeed, $\mathbb{P}$ is a function of the Casimirs, and thus is a function of $\kappa_s((\mu, q), (\mu, q))$. This fact prevents us from directly invoking Theorem 2.5 as we would find that $\mathbb{P}^1$ is indeed constant on $\kappa(\mu, \mu)$, but μ does not have an invariant norm. Nevertheless, we can still use this theorem by restricting $\mathbb{P}^1$ to the sphere $\kappa(\mu, \mu)$, or equivalently simply considering polar coordinates for μ and discarding the radial coordinate. In this case, we can invoke Theorem 2.5 and obtain the result of a constant marginal distribution $\mathbb{P}_\infty^1$ projected on the coadjoint orbit of the Lie algebra $\mathfrak{g}$ alone.

(3) We compute the time derivative of the quantity $\|\mu\|^2 := -\kappa(\mu, \mu)$, which is positive definite and thus defines a norm, to get an upper estimate of the form

$$\frac{d}{dt} \frac{1}{2} \|\mu\|^2 = \langle \mu, \dot{\mu} \rangle = \langle \text{ad}_r q, \mu \rangle \leq \|r\| \|q\| \|\mu\|.$$

Then, because $\|q\| = \sqrt{-\kappa(q, q)}$ is constant, and provided that r is bounded, we can integrate to find

$$\|\mu(t)\| \leq \|\mu(0)\| + \alpha t, \quad (2.34)$$

where α is a constant depending on the Lie algebra and the Hamiltonian. $\square$

Remark 2.7 (On ergodicity) An important question about any given dynamical system is whether its solution is ergodic. This question needs some clarification for the systems considered here. First, notice that the deterministic systems are not ergodic, as they are Hamiltonian systems with extra conserved quantities given by the Casimirs. If the σ_i span the entire Lie algebra, there is a constant invariant measure on the level set of the Casimir given by the initial conditions. This means that we have the ergodicity property on the coadjoint orbits but not on the full Euclidian space in which the coadjoint orbits are embedded. The ergodicity must then only be defined with respect to the coadjoint orbit, or the system will not be seen to be ergodic. Finally, the cases where the σ_i do not span the Lie algebra must be treated individually, depending on the system in hand. For example with the rigid body in Sect. 5, having two independent non-trivial σ_i is sufficient for ergodicity, while having only one σ_i will make the system integrable, and thus non-ergodic on the coadjoint orbit, or momentum sphere.

Summary This section has reviewed the framework for the study of noise in dynamical systems defined on coadjoint orbits and has illustrated how noise may be included in these systems, so as to preserve the deterministic coadjoint orbits. This preservation property is seen clearly in the Clebsch formulation because the deterministic and stochastic systems share the same momentum map, whose level sets define the coadjoint orbits. The systems we have considered are the Euler–Poincaré equations on semi-simple finite-dimensional Lie groups and the semidirect product structures which appear when the advected quantities are introduced into the underlying vector space of the Lie algebra of the Lie group. These structures are not the most general. However, their study has allowed us to use the properties of the natural pairing given by the Killing form to prove a few illustrative results in a simple and transparent way. In particular, we showed that the stationary solution of the Fokker–Planck equation, written in Lie–Poisson form, is constant on the coadjoint orbits. In the semidirect product setting, a bit more care was needed to obtain similar results for the marginal distributions, as the coadjoint orbits are not compact in this case. We will illustrate our approach with the two basic examples of the rigid body and heavy top in Sects. 5 and 6, where more will be said about these systems, and in particular about their integrability.

3 Dissipation and Invariant Measures

In the previous section, we described a structure preserving stochastic deformation of mechanical systems with symmetries. The preserved structure is the coadjoint orbit of the deterministic system. Namely, the stochastic process still belongs to one of these orbits, characterised by the initial conditions of the system. This preservation is reflected in the strict conservation of particular integrals of motion, called Casimirs. In general, these are the only conserved quantities of our stochastic processes. Indeed, the energy is not conserved, apart from very particular choices of the energy and the stochastic potentials as we will see for some examples. The energy is not strictly decaying either but is subject to random fluctuations. The complexity of the energy evolution has hindered us from studying it in full generality in the previous sections. In the present section, however, we will investigate the energy behaviour for particular mechanical examples subject to dissipation and random fluctuations. The type of energy dissipation that we will introduce in Sect. 3.1 also preserves the coadjoint orbits. Consequently, the dissipation is compatible with our stochastic deformation. The main outcome after introducing this dissipation is the emergence of a balance between noise and dissipation which will make the stationary solution of the Fokker–Planck equation energy dependent, as we will see in Sect. 3.2 and in the proof of existence of random attractors in Sect. 4.

3.1 Double Bracket Dissipation

To augment the stochastic processes introduced in the previous section, we will add a type of dissipation for which the solutions of the stochastic process will still lie on the deterministic coadjoint orbit. For this purpose, we will use double bracket dissipation, which was studied in detail in Bloch et al. (1996) and was generalised recently in Gay-

Balmaz and Holm (2013) and Gay-Balmaz and Holm (2014). We will follow the latter works to include an energy dissipation which preserves the Casimir functions. We will not review this theory in detail here. Instead, we refer the reader to Gay-Balmaz and Holm (2013) for a detailed discussion of Euler–Poincaré selective decay dissipation and Gay-Balmaz and Holm (2014) for the semidirect product extension.

For the stochastic process (2.10), the dissipative stochastic Euler–Poincaré equation written in Hamiltonian form is

$$d\mu + \text{ad}_{\frac{\partial h}{\partial \mu}}^* \mu \, dt + \theta \, \text{ad}_{\frac{\partial C}{\partial \mu}}^* \left[\frac{\partial C}{\partial \mu}, \frac{\partial h}{\partial \mu} \right]^b \, dt + \sum_i \text{ad}_{\sigma_i}^* \mu \circ dW_t^i = 0, \quad (3.1)$$

where $\theta > 0$ parametrises the rate of energy dissipation and C is a chosen Casimir of the coadjoint orbit. For convenience, we are using the isomorphism $\flat : \mathfrak{g} \rightarrow \mathfrak{g}^*$ defined via the Killing form of $\mathfrak{g}$. The converse isomorphism will be denoted $\sharp : \mathfrak{g}^* \rightarrow \mathfrak{g}$. The corresponding generalisation of selective decay for the semidirect product stochastic process (2.21), following (Gay-Balmaz and Holm 2014), is given by

$$\begin{aligned} d(\mu, q) + \text{ad}_{(\xi, r)}^*(\mu, q) \, dt + \theta \, \text{ad}_{\left(\frac{\partial C}{\partial \mu}, \frac{\partial C}{\partial q}\right)}^* \left[\left(\frac{\partial C}{\partial \mu}, \frac{\partial C}{\partial q} \right), (\xi, r) \right]^b \, dt \\ + \sum_i \text{ad}_{(\sigma_i, \eta_i)}^*(\mu, q) \circ dW_t^i = 0, \end{aligned} \quad (3.2)$$

where $\xi = \frac{\partial h}{\partial \mu}$, and the quantities h and r are defined in equation (2.23). Equation (3.2) may be written equivalently as a system of equations, by using the actions given in (2.20). Namely,

$$\begin{aligned} d\mu + (\text{ad}_{\xi}^* \mu + \text{ad}_r^* q) \, dt + \theta \, \text{ad}_{\frac{\partial C}{\partial \mu}}^* \left[\frac{\partial C}{\partial \mu}, \xi \right]^b \, dt + \theta \, \text{ad}_{\frac{\partial C}{\partial q}}^* \left(\text{ad}_{\frac{\partial C}{\partial \mu}} r + \text{ad}_{\frac{\partial C}{\partial q}} \xi \right)^b \, dt \\ + \sum_i \left(\text{ad}_{\sigma_i}^* \mu + \text{ad}_{\eta_i}^* q \right) \circ dW_t^i = 0, \\ dq + \text{ad}_{\xi}^* q \, dt + \theta \, \text{ad}_{\frac{\partial C}{\partial \mu}}^* \left(\text{ad}_{\frac{\partial C}{\partial \mu}} r - \text{ad}_{\frac{\partial C}{\partial q}} \xi \right)^b \, dt + \sum_i \text{ad}_{\sigma_i}^* q \circ dW_t^i = 0. \end{aligned} \quad (3.3)$$

Recall for the deterministic equations that the energy decays for $\theta > 0$ as

$$\frac{d}{dt} h(\mu, q) = -\theta \left\| \text{ad}_{\frac{\partial C}{\partial \mu}} \xi \right\|^2 - \theta \left\| \text{ad}_{\frac{\partial C}{\partial \mu}} r + \text{ad}_{\frac{\partial C}{\partial q}} \xi \right\|^2, \quad (3.4)$$

where the second term is present only in the semidirect product setting (Gay-Balmaz and Holm 2014).

Remark 3.1 (Choice of Casimir) The selective decay approach preserves the entire coadjoint orbit, and the speed of decay depends upon which invariant function C one

uses in implementing it. Indeed, either the first or second term of (3.4) can vanish depending on the choice of Casimir. We refer to the heavy top example in Sect. 6 for more details.

Remark 3.2 (Variational principle and reduction) The reader may have noticed already that although we introduced noise via variational principles, the dissipation is added as an extra term in the equations of motions. Indeed, the introduction of noise can also be seen as replacing the Hamiltonian h with the stochastic quantity $h dt + \Phi_i \circ dW_t^i$, and we expect to replace it in both the advection and dissipation. In fact, such an equation also fits a variational principle, the so-called Hamilton-Pontryagin principle, combined with Lagrange-d'Alembert for the dissipative force.

Proposition 3.3 *The variational principle*

$$\begin{aligned} & \delta \left(\int l(\xi) dt \right) + \int \left\langle \theta \left[\frac{\partial C}{\partial \mu}, g^{-1} dg \right], \left[\frac{\partial C}{\partial \mu}, g^{-1} \delta g \right] \right\rangle \\ & + \delta \left(\sum_i \int \langle \mu, g^{-1} dg - \xi dt - \sigma_i \circ dW_t^i \rangle \right) = 0, \end{aligned} \quad (3.5)$$

yields the stochastic process

$$\begin{aligned} & d\mu + \text{ad}_{\frac{\delta h}{\delta \mu}}^* \mu dt + \theta \text{ad}_{\frac{\delta C}{\delta \mu}}^* \left[\frac{\delta C}{\delta \mu}, \frac{\delta h}{\delta \xi} \right]^b dt + \theta \sum_i \text{ad}_{\frac{\delta C}{\delta \mu}}^* \left[\frac{\delta C}{\delta \mu}, \sigma_i \right]^b \circ dW_t^i \\ & + \sum_i \text{ad}_{\sigma_i}^* \mu \circ dW_t^i = 0, \end{aligned} \quad (3.6)$$

for free variations δg , $\delta \xi$ and $\delta \mu$.

Proof By a direct computation, using the theory of Hamilton-Pontryagin for the noise and Lagrange-d'Alembert for the double bracket dissipation. $\square$

This term would be admissible since it preserves the coadjoint orbit. However, we do not include it here. Instead, we leave it for treatment elsewhere, as it complicates the calculations to follow without significantly affecting the solution behaviour; since it is proportional to $\theta \sigma^2$, and θ and σ^2 are taken as being small compared to the original dynamics so that the noise and dissipation are viewed as perturbations.

Interestingly, the Hamilton-Pontryagin principle produces a more physically consistent dissipative equation. This suggests that this principle is the more general than the Clebsch derivation. On top of this remark, we already mentioned that for semidirect product structures, the Clebsch derivation does not allow the noise fields to depend on the advected quantities, whereas the Hamilton-Pontryagin does not have this restriction.

Asymptotically in time, $t \rightarrow \infty$, the deterministic equations with selective decay will tend towards a state which is compatible with the state of minimal energy, as

shown in Gay-Balmaz and Holm (2014). However, the presence of noise will balance the dissipation due to selective decay and prevent the system from reaching this deterministic equilibrium position. This feature will imply a non-constant stationary distribution of the corresponding Fokker–Planck solution to be studied in the next section, as well as the existence of random attractors, for which we refer to Kondrashov et al. (2015) and Schenk-Hoppé (1998) for background information.

3.2 The Fokker–Planck Equation and Stationary Distributions

To study the balance between multiplicative noise and nonlinear dissipation, we compute the Fokker–Planck equation associated with the process (3.1) or, equivalently, (3.3), and its stationary solutions. We obtain the Fokker–Planck equation

$$\frac{d}{dt} \mathbb{P}(\mu) + \{h, \mathbb{P}\} + \theta \left\langle \left[\frac{\partial \mathbb{P}}{\partial \mu}, \frac{\partial C}{\partial \mu} \right], \left[\frac{\partial h}{\partial \mu}, \frac{\partial C}{\partial \mu} \right] \right\rangle - \frac{1}{2} \sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}\}\} = 0. \quad (3.7)$$

The stationary distribution of this Fokker–Planck equation is no longer a constant on the coadjoint orbits. Instead, it now depends on the energy, as summarised in the following theorem.

Theorem 3.4 *For a compact semi-simple Lie algebra $\mathfrak{g}$, let the noise be isotropic, that is $\sigma_i = \sigma e_i$ for an arbitrary $\sigma \in \mathbb{R}$, where the e_i 's span the underlying vector space of the dual Lie algebra $\mathfrak{g}^* \cong \mathfrak{g}$. The stationary distribution of the Fokker–Planck equation (3.7) associated to (3.1) with Casimir $C = \kappa(\mu, \mu)$ is given by*

$$\mathbb{P}_\infty(\mu) = Z^{-1} e^{-\frac{2\theta}{\sigma^2} h(\mu)}, \quad (3.8)$$

where Z is the normalisation constant, or partition function, that enforces $\int \mathbb{P}_\infty(\mu) d\mu = 1$.

Proof The stationary distribution is given by solving $\frac{d}{dt} \mathbb{P}_\infty(\mu) = 0$. The isotropic noise enforces the advection term to vanish independently of the other terms. We therefore use the Ansatz $\mathbb{P}_\infty(\mu) = f(h(\mu))$, where the function f is to be determined. Consequently, only the selective decay and the double bracket term remain. The selective decay is first rewritten, using the bi-invariance property of the Killing form (2.4), as

$$\begin{aligned} \theta \left\langle \left[\frac{\partial \mathbb{P}}{\partial \mu}, \frac{\partial C}{\partial \mu} \right], \left[\frac{\partial h}{\partial \mu}, \frac{\partial C}{\partial \mu} \right] \right\rangle &= \theta \left\langle \frac{\partial \mathbb{P}}{\partial \mu}, \text{ad}_{\frac{\partial C}{\partial \mu}} \left[\frac{\partial C}{\partial \mu}, \frac{\partial h}{\partial \mu} \right] \right\rangle \\ &= \theta \mathbf{d} \left(f(h) \text{ad}_{\frac{\partial C}{\partial \mu}} \left[\frac{\partial C}{\partial \mu}, \frac{\partial h}{\partial \mu} \right] \right), \end{aligned}$$

where we have used the property that the coadjoint action for semi-simple Lie algebras is divergence-free. (Notice that the exterior derivative $\mathbf{d}$ is a divergence operation here.)

Since $\kappa(\mu, \mu)$ is a Casimir and $\mu^\sharp = \frac{\partial C}{\partial \mu}$, we can rewrite the double bracket due to the noise as

$$\begin{aligned} -\frac{1}{2} \sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}\}\} &= -\sigma^2 \frac{1}{2} \sum_i \left\langle \text{ad}_{e_i}^* \frac{\partial C}{\partial \mu}, \frac{\partial}{\partial \mu} \left\langle \text{ad}_{e_i}^* \frac{\partial C}{\partial \mu}, \frac{\partial \mathbb{P}}{\partial \mu} \right\rangle \right\rangle \\ (\text{From bi-invariance of } \kappa) &= \sigma^2 \frac{1}{2} \sum_i \mathbf{d} \left(f'(h) \text{ad}_{e_i} \frac{\partial C}{\partial \mu} \left\langle \text{ad}_{\frac{\partial h}{\partial \mu}} \frac{\partial C}{\partial \mu}, e_i \right\rangle \right) \\ &= \sigma^2 \frac{1}{2} \mathbf{d} \left(f'(h) \text{ad}_{\text{ad}_{\frac{\partial h}{\partial \mu}} \frac{\partial C}{\partial \mu}} \frac{\partial C}{\partial \mu} \right). \end{aligned}$$

We have used the bi-invariance of the pairing to enforce the relation $\text{ad}_\xi^\dagger \eta := \text{ad}_\xi^* \eta^\flat = -\text{ad}_\xi \eta$. See for example (Varadarajan 1984) for more details. The result (3.8) for the equilibrium distribution then follows by comparing the selective decay term with the double bracket term and noticing that the two terms will cancel, provided $f(x) = e^{-2\theta x/\sigma^2}$. $\square$

The Fokker–Planck equation with dissipation in the semidirect product setting directly gives

$$\begin{aligned} \frac{d}{dt} \mathbb{P}(\mu, q) + \{h, \mathbb{P}\} - \frac{1}{2} \sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}\}\} \\ + \theta \left\langle \left[\left(\frac{\partial \mathbb{P}}{\partial \mu}, \frac{\partial \mathbb{P}}{\partial q} \right), \left(\frac{\partial C}{\partial \mu}, \frac{\partial C}{\partial q} \right) \right], \left[\left(\frac{\partial h}{\partial \mu}, \frac{\partial h}{\partial q} \right), \left(\frac{\partial C}{\partial \mu}, \frac{\partial C}{\partial q} \right) \right] \right\rangle = 0. \end{aligned} \quad (3.9)$$

Consequently, for semidirect products, we have the analogue of the previous theorem, but for the marginal stationary distribution on the advected quantities.

Theorem 3.5 *Provided that the Hamiltonian is of the form $h(\mu, q) = K(\mu) + V(q)$ for two functions K and V and that the noise is isotropic, the stationary marginal distribution $\mathbb{P}_\infty^2(q)$ with the selective decay from the Casimir $\kappa(\mu, q)$ is given by*

$$\mathbb{P}_\infty^2(q) = Z^{-1} e^{-\frac{2\theta}{\sigma^2} V(q)}, \quad (3.10)$$

where Z is the normalisation constant, or partition function.

Proof The proof here is similar to the proof for Theorem 3.4. Thus, we only show the main calculations. First, the selective decay term is given explicitly, using the Casimir $\kappa(\mu, q)$, by

$$\theta \left\langle \text{ad}_{\frac{\partial \mathbb{P}}{\partial \mu}} q, \text{ad}_\xi q \right\rangle + \theta \left\langle \text{ad}_{\frac{\partial \mathbb{P}}{\partial \mu}} \mu + \text{ad}_{\frac{\partial \mathbb{P}}{\partial q}} q, \text{ad}_\xi \mu + \text{ad}_r q \right\rangle.$$

Integrating the selective decay term of (3.9) in μ and assuming $\mathbb{P}^2(q) = f(V(q))$, gives

$$\theta \left\langle \operatorname{ad}_{\frac{\partial \mathbb{P}^2}{\partial q}} q, \operatorname{ad}_r q \right\rangle = -\theta \left\langle \operatorname{ad}_{\operatorname{ad}_r q}^* q, \frac{\partial \mathbb{P}^2}{\partial q} \right\rangle = -\theta \mathbf{d} \left(\operatorname{ad}_{\operatorname{ad}_r q}^* q f \right),$$

where we have used the bi-invariance property of κ (2.4), as well as the divergence-free property of the coadjoint action. Then, after integration over μ , the double bracket term becomes

$$\begin{aligned} -\frac{1}{2} \sigma^2 \mathbf{d} \left(\operatorname{ad}_{e_i}^* q \left\langle \operatorname{ad}_{e_i}^* q, \frac{\partial \mathbb{P}^2}{\partial q} \right\rangle \right) &= \frac{1}{2} \sigma^2 \mathbf{d} \left(f' \operatorname{ad}_{e_i}^* q \langle \operatorname{ad}_r q, e_i \rangle \right) \\ &= -\frac{1}{2} \sigma^2 \mathbf{d} \left(f' \operatorname{ad}_{\operatorname{ad}_r q}^* q \right), \end{aligned}$$

upon again using bi-invariance. Thus, the result follows, as f must satisfy $\theta f = \frac{1}{2} \sigma^2 f'$. $\square$

In the Euler–Poincaré setting, the stationary distribution was concentrated around the positions of minimum energy, and here the advected quantity q is concentrated around the position of minimal potential energy. We conjecture that the complete stationary distribution is concentrated around the minimal energy region, as in the Euler–Poincaré setting. However, we will not investigate this conjecture here, as we will be mainly interested in the dynamics of the advected quantities.

Remark 3.6 (Gibbs measure) This calculation only uses the bi-invariance of the Killing form, which holds in general for semi-simple Lie algebras. Therefore, the same conclusion applies for other Lie algebras which admit a bi-invariant pairing. In statistical physics, the invariant measure (3.8) is often called a Gibbs measure. This association provides a natural identification of the quantity $\sigma^2/2k_B\theta$ with a Kelvin temperature T , where k_B is the Boltzmann constant. This notion of temperature arises via coupling the system with a heat bath, at temperature T . Such an open system in statistical physics is referred to as a canonical ensemble, whereas the system without dissipation is closed, and hence fits in the traditional category of micro-canonical ensembles.

Remark 3.7 (Non-isotropic noise) To obtain the Gibbs measure, we had to assume an isotropic noise, but stationary distribution also exists with non-isotropic noise. The exact form of these distribution is more involved but will remain close to be a Gibbs measure. We will not compute it here, and only show an example of such distribution in the case of the rigid body, in Sect. 5.

Remark 3.8 (Time reversal) The results of this section also hold when evolving backwards in time, using the change of variable $t \rightarrow -t$. Indeed, since the noise dW is centred, only the dissipation will be affected by time reversal, and it will have the opposite effect, namely the system will tend towards the highest energy equilibrium position.

4 Random Attractors

We now turn to the study of the existence of random attractors (RAs) in our stochastic dissipative systems, in connection with the theory of random dynamical systems (RDS). The classic approach in studying the effect of stochastic forcing of nonlinear dynamical systems proceeds by integrating the system forward in time and performing averages, then studying the Fokker–Planck equation, as we have done up to now. Another approach studies random dynamical systems via the so-called pull-back method. We will not fully explore the theory of random dynamical systems and pullback attractors here. Instead, we will only invoke the main results from the theory and refer the interested reader to [Crauel and Flandoli \(1994\)](#), [Crauel et al. \(1997\)](#), [Arnold \(1995\)](#), [Bonatti et al. \(2006\)](#) and [Kloeden and Rasmussen \(2011\)](#) for in-depth accounts of these subjects. In a nutshell, for a given fixed realisation of the noise, the average is taken over the initial conditions. The noise makes the system time-dependent; so the notion of an attractor should be defined in the pull-back sense, such that for large times the attractive set does not depend on time. That is, the pullback attractor is defined by pulling back a given set of initial conditions from $t = 0$ to $t \rightarrow -\infty$ and letting the system evolve to $t = 0$. In this limit, the set obtained at $t = 0$ is the pullback attractor. In random dynamical systems theory, the pullback attractor is usually called a random attractor, and if it is not singular, it may admit a particular type of measure, the Sinai-Ruelle-Bowen measure (SRB), which is also called a physical measure, see [Young \(2002\)](#). We will denote the physical measure by $\mathbb{P}_\omega(\mu)$ for a given realisation of the noise ω . There is a fundamental relation between this SRB measure and the stationary distribution of the Fokker–Planck equation which was first discovered in [Crauel \(1991\)](#) and later in a theorem of [Crauel and Flandoli \(1998\)](#). This relation is informally given by

$$\int_{\Omega} \mathbb{P}_\omega(\mu) d\omega = \mathbb{P}_\infty(\mu), \quad (4.1)$$

for the probability space Ω . Here we are referring to probability densities, and the SRB measure can be seen as the invariant measure most compatible with volume, although volume in phase space is not preserved, because of dissipation. For more explanation, see [Young \(2002\)](#).

Remark 4.1 (Periodic kicking) We are only considering here the interaction of noise and dissipation. However, if the noise were replaced by a simpler deterministic forcing, similar results would emerge. In particular, periodic forcing or kicking of dissipative dynamical systems has been studied in great detail in numerous works, e.g., in [Lin and Young \(2010\)](#) and [Lu et al. \(2013\)](#). In Sect. 5.5, we will implement periodic kicking and damping in the rigid body, and will numerically demonstrate the existence of non-singular attractors and chaos. We have left deeper theoretical studies of these systems for future investigations.

4.1 Existence of Attractors

We first determine that the stochastic processes (3.1) and (3.3) do indeed admit random attractors, provided the top Lyapunov exponent is positive. See Kondrashov et al. (2015) and Schenk-Hoppé (1998) and references therein for more details about this type of approach. Then, we will estimate the value of the top Lyapunov exponent using numerical simulations for the rigid body in Sect. 5.

Theorem 4.2 *The stochastic process (3.1) admits a random attractor, for every Lie group G .*

Proof The SDE (3.1) may be recast as a random dynamical equation (RDE) by using the following vector Wiener processes z_i ,

$$dz_i = \sigma_i dW_t^i, \quad (4.2)$$

where z_i is understood as a vector process in the direction along σ_i . In the sequel, we will denote $z(t, \omega) = \sum_i z_i(t, \omega) \in \mathfrak{g}$. The process $z(t)$ thus defines a random path in the Lie algebra $\mathfrak{g}$ and, via the exponential map, a random path in the group G as $g(t, \omega) = e^{z(t, \omega)}$.

We then define a new variable $\tilde{\mu}(t) = g(t)\mu(t) := \text{Ad}_{g(t)}^* \mu(t)$ and we have, from (3.1) (see for example Marsden and Ratiu 1999),

$$d\tilde{\mu}(t) = \text{Ad}_{g(t)}^* \left(- \sum_i \text{ad}_{\sigma_i}^* \mu \circ dW + d\mu \right) = \text{Ad}_{g(t)}^* \left(F(\text{Ad}_{g(t)^{-1}}^* \tilde{\mu}(t)) \right) dt,$$

where our stochastic process is generically written $d\mu = F(\mu)dt + G_i(\mu) \circ dW_i$ for convenience. From here, we have the RDE associated to (3.1) of the form

$$\frac{d}{dt} \tilde{\mu}(t) = \tilde{F}(\tilde{\mu}(t), g(t)), \quad (4.3)$$

where $\tilde{F}$ is defined in the previous calculation as the drift part of the process. Recall that from the theory of selective decay we have (Gay-Balmaz and Holm 2013)

$$\frac{d}{dt} h(\mu) = -\theta \left\| \left[\frac{\partial C}{\partial \mu}, \frac{\partial h}{\partial \mu} \right] \right\|^2,$$

and $h(\text{Ad}_g^* \mu) = h(\mu)$, because h is G -invariant, so that this equality becomes for (4.3),

$$\frac{d}{dt} h(\tilde{\mu}) = -\theta \left\| \left[\text{Ad}_{g^{-1}} \frac{\partial C(\tilde{\mu})}{\partial \tilde{\mu}}, \text{Ad}_{g^{-1}} \frac{\partial h(\tilde{\mu})}{\partial \tilde{\mu}} \right] \right\|^2 \leq 0. \quad (4.4)$$

This inequality assures that the energy decays at a random strictly negative rate. The existence of the random attractor then follows from standard arguments, demonstrated,

for example, in the linear case by [Schenk-Hoppé \(1998\)](#) and in a more general nonlinear setting by [Chekroun et al. \(2011\)](#). $\square$

The idea of this proof is to generalise the linear change of variables used to recast the original stochastic process as a random dynamical equation, by using a nonlinear group theoretical change of variable. The dissipative property follows from the selective decay theory and the invariance of the Hamiltonian under the group action. This theorem is general, in that no specific assumptions on the Lie group need to be imposed. In particular, modulo difficulties in analysis, the theorem should also apply for the diffeomorphism group used in the description of compressible fluid equations. However, we have no intention of investigating the infinite dimensional theory here.

The same result persists in the semidirect product theory, as developed earlier.

Corollary 4.2.1 *Theorem 4.2 applies to semidirect product stochastic processes (3.3).*

Proof The proof follows the same argument, upon using the action of the group G and the Lie algebra $\mathfrak{g}$ and the advected quantities in V to define the change of variables. The decay rate of the energy is given by using the deterministic selective decay formulae (3.4). $\square$

4.2 Existence of the SRB Measure

We now turn to the existence of the SRB measure. Theorem 4.3 below for the existence of SRB measures will invoke Hörmander's theorem about the smoothness of transition probabilities for a diffusion satisfying the so-called Hörmander (Lie) bracket conditions. The Lie bracket $[v, w](x)$ of two vector fields $v(x)$, $w(x)$ in $\mathbb{R}^n$ is defined as

$$[v, w](x) = Dv(x)w(x) - Dw(x)v(x), \quad (4.5)$$

where we denote by Dv the derivative matrix given by $(Dv)_{ij} = \partial_j v_i = v_{i,j}$. Given an SDE of the form

$$dx = A_0(x)dt + \sum A_i(x) \circ dW_t^i, \quad (4.6)$$

the Hörmander condition we use states that if the following condition is satisfied

$$\cup_{k \geq 1} V_k(x) = \mathbb{R}^n, \quad \text{for all } x, \quad (4.7)$$

where

$$\begin{aligned} V_k(x) &= V_{k-1}(x) \cup \text{span}\{[v(x), A_j(x)] : v \in V_{k-1}, j \geq 0\} \quad \text{and} \\ V_0(x) &= \text{span}\{A_j, j \geq 1\}, \end{aligned} \quad (4.8)$$

then the invariant measure of (4.6) is smooth with respect to the Lebesgue measure. The Hörmander condition implies the following standard theorem for stochastic dissipative systems.

Theorem 4.3 *If the largest Lyapunov exponent of (3.1) is positive, the random attractor is the support of a Sinai-Ruelle-Bowen (SRB) measure.*

Proof The proof uses the corollary of Theorem B in (Ledrappier and Young 1988), which assumes the existence of a random attractor. The only point left to show here is that the parabolic Hörmander condition (4.7) is fulfilled. Given the Stratonovich process (4.6) in $\mathfrak{g}^*$, we only need to check that the vector fields $A_1, \dots, A_N$ will span the tangent space to the coadjoint orbits as long as N is sufficiently large. Since $A_i(\mu) := \text{ad}_{\sigma_i}^* \mu$, they are tangent to the coadjoint orbits. The minimal number of A_i needed cannot be found, in general, as it will depend on the Lie symmetry algebra and the form of the Hamiltonian. Nevertheless, the σ_i span the vector space $\mathfrak{g}$, and the Hörmander condition is fulfilled. $\square$

Corollary 4.3.1 *Theorem 4.3 also applies for the semidirect product case, even with $\eta_i = 0$.*

Proof The same argument applies here, even if $\eta_i = 0$, as the semidirect product structure will automatically span the whole space, provided $\mathfrak{g}$ is already spanned, and h is not too degenerate on V . $\square$

4.3 Estimation of Lyapunov Exponents

In principle, it is possible to compute the value of the top Lyapunov exponent as a function of the parameters of the system. However, this turns out to be a very challenging computation. Nonetheless, positivity of the top Lyapunov exponent is important to determine, as it allows us to use the previous Theorem 4.3 to prove the existence of a non-singular random attractor with an SRB measure and positive entropy. We will restrict ourselves to the first step of the calculation and explain why it is a difficult problem. We will then estimate the top Lyapunov exponent numerically in the example section, and leave the rigorous proof as an open problem.

The very first step is to estimate the sum of the Lyapunov exponents using the multiplicative ergodic theorem (MET), that we state here in its simplest form. We will denote by Ω the probability space of the stochastic dynamical system.

Theorem 4.4 (MET theorem) *For an ergodic stochastic process μ with invariant measure $\mathbb{P}_\infty$ there exists an flow invariant subset $\Delta \subset \mathfrak{g}^* \times \Omega$ of the n -dimensional phase space and a sequence of ordered Lyapunov exponents λ_i for $i = 1, \dots, n$ such that the following properties hold for all $(\mu, \omega) \in \Delta$:*

(1) *The dual of the Lie algebra can be decomposed into a direct sum*

$$T_\mu \mathfrak{g}^* = E_1(\omega, \mu) \oplus \dots \oplus E_n(\omega, \mu),$$

where

$$\delta\mu \in E_i \Leftrightarrow \lim_{t \rightarrow \infty} \frac{1}{t} \log \|DF(t, \omega, \mu)\delta\mu\| = \lambda_i. \quad (4.9)$$

- $d(\delta\mu) = DF(t, \omega, \mu)\delta\mu$ is the linearisation of the flow equation $d\mu = F(t, \omega, \mu)$ along a particular solution $\mu(t)$.
- (2) For a generic $\delta\mu$, the associated Lyapunov exponent is $\lambda(\omega, x, v) = \lambda_+$, the largest one almost surely.
- (3) The sum of the Lyapunov exponents is

$$\sum_i \lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log \det DF(t, \omega, \mu). \quad (4.10)$$

For simplicity here,, we have assumed that the multiplicity of the Lyapunov exponents is always 1; that is, they are all distinct. We refer to [Arnold \(1995\)](#) for more details on this theorem and its generalisations. The stochastic systems which we consider here are written on compact semi-simple Lie algebras, such that $c^2 = \|\mu\|^2$ is constant and defines a bounded set. The energy functional $h(\mu)$ is also a generic quadratic kinetic energy term, with a given inertia tensor $\mathbb{I}^{-1}$, corresponding to the Hessian matrix of $h(\mu)$. We can then prove the following formula for the sum of the Lyapunov exponents.

Proposition 4.5 *Provided the Lie algebra is compact semisimple, the sum of the Lyapunov exponents is estimated from below by*

$$\sum_i d_i \lambda_i \geq -\frac{1}{2} |\epsilon| n \sigma^2 - \theta |\epsilon| \left(c^2 \mathbb{I}_{\min}^{-1} - \mathbb{E}_{\infty} h(\mu) \right), \quad (4.11)$$

where $c = \|\mu\|^2$, ϵ is the Killing form constant, n is the number of $\sigma_i = \sigma e_i$ spanning the Lie algebra and d_i are the multiplicity of each Lyapunov exponent. Thus, the dimension of the Lie algebra, $n = \dim(\mathfrak{g})$. The quantity $\mathbb{I}_{\min}^{-1}$ is the largest eigenvalue of the Hessian of the Hamiltonian. The expectation $\mathbb{E}_{\infty}$ is taken with respect to the invariant measure $\mathbb{P}_{\infty}$. An estimation from above is also available, upon using $\mathbb{I}_{\max}^{-1}$, the minimal eigenvalue, instead of $\mathbb{I}_{\min}^{-1}$.

Proof Let us denote the stochastic process (3.1) in Itô form by

$$d\mu = F(\mu)dt + \sum_i G_i(\mu) dW_t^i.$$

We can now directly apply the MET Theorem 4.4 to compute the sum of the Lyapunov exponents

$$\sum_i \lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log \det \delta\mu(t, \omega, x). \quad (4.12)$$

We can then use Jacobi formula to rewrite (4.12) as

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \det \delta\mu(t, \omega, x) = \lim_{t \rightarrow \infty} \frac{1}{t} \text{Tr} \int_t^t DF(\varphi(t, x, \omega)) ds. \quad (4.13)$$

Here, we have intentionally dropped the linearisation of the noise amplitude, since this term will vanish (the trace of the adjoint action always vanishes). Finally, ergodicity of this process gives

$$\sum_i \lambda_i = \int \text{Tr}(DF(\mu)) \mathbb{P}_\infty(\mu) d\mu, \quad (4.14)$$

where $\mathbb{P}_\infty$ is the invariant measure of the underlying stochastic process. The calculation of the trace simplifies in the case of a compact semi-simple Lie algebra with the Killing form $\text{Tr}(\text{ad}_A \text{ad}_B) = \epsilon A \cdot B$, where $\epsilon < 0$ depends on the Lie algebra. Then, using the explicit form of F along with semi-simplicity for $\mathfrak{g}$, yields

$$F(\mu) = \text{ad}_{\frac{\partial h}{\partial \mu}} \mu + \theta \text{ad}_\mu \text{ad}_\mu \frac{\partial h}{\partial \mu} + \frac{1}{2} \sum_i \text{ad}_{\sigma_i} \text{ad}_{\sigma_i} \mu.$$

Consequently, we arrive at

$$\text{Tr}(DF(\mu)) = \text{Tr} \left(-\theta \text{ad}_\mu \text{ad}_{\frac{\partial h}{\partial \mu}} + \theta \text{ad}_\mu \text{ad}_\mu \frac{\partial^2 h}{\partial \mu^2} + \frac{1}{2} \sum_i \sigma_i^2 \text{ad}_{e_i} \text{ad}_{e_i} \right) \quad (4.15)$$

$$= |\epsilon| \theta h(\mu) + \theta A(\mu, \mu) - \frac{1}{2} |\epsilon| n \sigma^2, \quad (4.16)$$

where n is the number of σ_i fields. The $A(\mu, \mu)$ term depends on the Lie algebra structure constants and is difficult to obtain explicitly for every compact semi-simple Lie algebra. However, we can estimate it here using

$$|\epsilon| \theta \kappa(\mu, \mu) \mathbb{I}_{\max}^{-1} \geq \theta \text{Tr} \left(\theta \text{ad}_\mu \text{ad}_\mu \frac{\partial^2 h}{\partial \mu^2} \right) \geq |\epsilon| \theta \kappa(\mu, \mu) \mathbb{I}_{\min}^{-1}. \quad (4.17)$$

Collecting terms then gives a lower and upper bound for the sum of the Lyapunov exponents (4.11). $\square$

Unfortunately, Proposition 4.5 only gives a negative lower bound for the sum of the Lyapunov exponents of an arbitrary compact semi-simple Lie algebra. Indeed, the last term can be bounded from above by

$$\mathbb{E}_\infty h(\mu) = \int h(\mu) e^{-\frac{2\theta}{\sigma^2} h(\mu)} d\mu \leq c^2 \mathbb{I}_{\min}^{-1}.$$

Thus, the sum of the Lyapunov exponents is found to be negative.

A precise value can be computed explicitly for each Lie algebra, by using the structure constants to calculate the term $A(\mu, \mu)$ in (4.16). We will show this calculation in the case of $SO(3)$ in the free rigid body example in Sect. 5.

Remark 4.6 (Non-compact Lie algebras) This argument does not apply for non-compact semi-simple Lie algebras, as the Killing form is not sign-definite, so it does not provide us with a norm.

Having only a negative lower bound for the sum of the Lyapunov exponent means we must proceed further by estimating the top Lyapunov exponent, to obtain an SRB measure. Obtaining a precise estimate for our general class of system is still an open problem, but we will show here the main steps towards this result. We will then numerically estimate it for the simple case of $SO(3)$ in the Sect. 5.

First, recall the linearisation of the flow, in Itô form

$$d\delta\mu = DF(\mu)\delta\mu dt + \sum_i DG_i(\mu)\delta\mu dW_t^i, \quad (4.18)$$

where $DG_i(\mu) = \text{ad}_{\sigma_i}$. We then want to apply the Furstenberg-Kasminskii formula. (See Arnold 1995 for details.) For this purpose, we introduce the following change of variables

$$R = \|\delta\mu\| \in \mathbb{R} \quad \text{and} \quad \Phi = \frac{\delta\mu}{\|\delta\mu\|} \in \mathbb{S}^{n-1}, \quad (4.19)$$

where $\mathbb{S}^{n-1}$ is the $(n-1)$ -sphere. The corresponding stochastic processes are

$$dR = \langle \Phi, (DF(\mu)dt + DG_i dW_t^i) \Phi \rangle R := Q(\mu, \Phi)R, \quad (4.20)$$

$$\begin{aligned} d\Phi &= (DF(\mu)dt + DG_i dW_t^i)\Phi - \langle \Phi, (DF(\mu)dt + DG_i dW_t^i) \Phi \rangle \Phi \\ &:= P(\mu, \Phi)dt + \sum_i P_i(\mu, \Phi)dW_t^i. \end{aligned} \quad (4.21)$$

As a direct consequence of ergodicity and the multiplicative ergodic Theorem 4.4 [Item 1 and 2], the Furstenberg-Kasminskii formula then gives the top Lyapunov exponent by evaluating the following integral

$$\lambda_+ = \int Q(\mu, \Phi) \mathbb{Q}(\mu, \Phi) d\mu d\Phi, \quad (4.22)$$

where $\mathbb{Q}(\mu, \Phi)$ is the joint stationary distribution of the processes for μ and Φ , where the processes for Φ depend on μ , but not the reverse.

One can see that a rough estimate of the integrand from below would always be negative. Consequently, we need more work to obtain a positive bound. Let us evaluate the quantity Q at the positions (μ_i, Φ_j) , where we use the basis corresponding to the eigenvalues of the Hessian of h , i.e., the inverse of the moment of inertia $\mathbb{I}^{-1}$. We also order the eigenvalues as $\mathbb{I}_1 > \dots > \mathbb{I}_n$. Note that because the dynamics of μ takes place on the coadjoint orbit, the linearisation is tangent to it and thus $i \neq j$ for the choice of positions (μ_i, Φ_j) . A direct calculation gives the following simplifications of the quantity $Q(\mu_i, \Phi_j)$

$$\begin{aligned}
Q(\mu_i, \Phi_j) &= \langle \Phi_j, DF(\mu_i)\Phi_j \rangle \\
&= \langle \text{ad}_{\mu_i} \Phi_j, \mathbb{I}^{-1} \Phi_j \rangle - \frac{1}{2} \sigma^2 \sum_k \langle \text{ad}_{e_k} \Phi_j, \text{ad}_{e_k} \Phi_j \rangle \\
&\quad + \theta \langle \text{ad}_{\mu_i} \Phi_j, \text{ad}_{\mathbb{I}^{-1} \mu_i} \Phi_j \rangle - \theta \langle \text{ad}_{\mu_i} \Phi_j, \text{ad}_{\mu_i} \mathbb{I}^{-1} \Phi_j \rangle \\
&= \langle \mu_i, \text{ad}_{\Phi_j} \mathbb{I}^{-1} \Phi_j \rangle - \frac{1}{2} \sigma^2 \sum_k \|\text{ad}_{e_k} \Phi_j\|^2 \\
&\quad + \theta \langle \text{ad}_{\mu_i} \Phi_j, \text{ad}_{\mu_i} (\mathbb{I}_i^{-1} \text{Id} - \mathbb{I}^{-1}) \Phi_j \rangle \\
&= -\frac{1}{2} \sigma^2 \sum_k \|\text{ad}_{e_k} \Phi_j\|^2 + \theta (\mathbb{I}_i^{-1} - \mathbb{I}_j^{-1}) \|\text{ad}_{\mu_i} \Phi_j\|^2 \\
&= -\frac{n-1}{2} \sigma^2 + \theta c^2 (\mathbb{I}_i^{-1} - \mathbb{I}_j^{-1}),
\end{aligned}$$

where we have used the Casimir sphere radius c (thus $\mu_i = c e_i$) and the fact that $\|\Phi\|^2 = 1$. The last formula conveys a lot of information about the possible sign for λ_+ . Indeed, depending on the choice of i and j , it is possible that $Q(\mu_i, \Phi_j)$ is positive, provided the difference between the moments of inertia is large enough with respect to the σ , θ and c . That difference between the moments of inertia is important and is tied to the nature of the random attractor. Indeed, differences in the moments of inertia imply different speeds for nearby orbits, and thus a shear effect. This shear effect is a common source of random attractors, which has appeared in a number of recent papers such as (Wang and Young 2003; Lin and Young 2008). The necessity of sufficiently large shear for the existence of the random attractor is thus implied by the last formula in the computation above. Of course, this is not the whole story, as the original dynamics of the system is also an important factor. Indeed, without noise and dissipation, the system is Hamiltonian. Thus the sum of Lyapunov exponents vanishes in this case, and the top Lyapunov exponent must be positive for a large set of initial conditions. In the present case, integrating the deterministic part of Q against the Gibbs measure is already analytically difficult, and no particular sign can be easily expected from examining this term. The only thing we can expect at this stage is that for large dissipation when the Gibbs measure is localised around the equilibrium positions of minimum energy, the dynamics of the deterministic system is negligible, whereas for a small dissipation the deterministic dynamics will be important. We refer to Sect. 5, where we will evaluate the top Lyapunov exponent in the rigid body example by numerical simulation.

We now turn to the semidirect product structure and also estimate the sum of the Lyapunov exponents in the following proposition, where we choose to use the Casimir $C(\mu, q) = \frac{1}{\epsilon} \kappa(q, q) = c^2$ for the dissipative term for simplicity only.

Proposition 4.7 *The sum of the Lyapunov exponents for the semidirect stochastic process (3.3) with Casimir $C(\mu, q) = \frac{1}{\epsilon} \kappa(q, q) = c^2$ is given by*

$$\sum_i \lambda_i \geq -|\epsilon| n \sigma^2 - \theta c^2 \mathbb{I}_{\min}^{-1}. \quad (4.23)$$

Proof We follow the proof for the Euler–Poincaré case. Let us denote the stochastic process (3.3) in Itô form by

$$d(\mu, q) = [F^\mu(\mu, q) + F^q(\mu, q)] dt + \sum_i [G_i^\mu(\mu, q) + G_i^q(\mu, q)] dW_t^i,$$

where we denoted F^μ (resp. F^q) the μ (resp. q) component of F . The MET theorem, Jacobi's formula and ergodicity of this process gives

$$\sum_i \lambda_i = \int \text{Tr} (D_\mu F^\mu(\mu, q) + D_q F^q(\mu, q)) \mathbb{P}_\infty(\mu, q) d(\mu, q), \quad (4.24)$$

where $\mathbb{P}_\infty(\mu, q)$ is the invariant measure of the underlying stochastic process, and D_μ and D_q denote the Jacobian matrices taken with respect to μ or q , respectively. Consequently, after substituting the Casimir $C(q)$ in the general formula (3.3), we obtain

$$\text{Tr}(D_\mu F^\mu + D_q F^q) = \text{Tr} \left(-\theta \text{ad}_q \text{ad}_q \mathbb{I}^{-1} - \sigma^2 \sum_i \text{ad}_{e_i} \text{ad}_{e_i} \right),$$

and, using again (4.17), we have the result (4.23). $\square$

As before, the sum of the Lyapunov exponents is negative. Thus, we must estimate the top Lyapunov exponent to prove the existence of a non-singular SRB measure for this system. The same difficulty as before remains in this case for estimating the top Lyapunov exponent using the Furstenberg–Kasminskii formula.

Summary In this section, we have studied the interaction of multiplicative noise and nonlinear dissipation on coadjoint orbits. For this purpose, we added a double bracket dissipation mechanism to the previously derived stochastic process to preserve the coadjoint orbit structure on which the solutions of the stochastic process are supported. In the case of semi-simple Lie algebras, we obtained the invariant measure of the Fokker–Planck equation and found the associated Gibbs measure on the coadjoint orbits. In the semidirect product case, this result was shown to hold for the marginal distribution of the advected quantity only, where the Gibbs distribution depends only on the potential energy. We then proved the existence of random attractors for a wide class of systems by using the dissipative property of the double bracket and the Hörmander condition on the generating vector fields, provided the top Lyapunov exponent is positive. Unfortunately, we were not able to derive an exact lower positive bound for the top Lyapunov exponent. However, numerical investigations in the example Sects. 5 and 6 will provide us with strong pieces of evidence of the positivity of the top Lyapunov exponent for some region of the parameter space (θ, σ) . In the next two sections, we will study two specific examples of stochastic deformations of the Euler–Poincaré dynamical equation, for the free rigid body and the heavy top, using both analytical and numerical tools.

5 Euler–Poincaré Example: The Stochastic Free Rigid Body

This section introduces stochastic dynamics for the classic example of the Euler–Poincaré dynamical equation; namely, the equation for free rigid body motion. Stochastic rigid body models have arisen in various fields of application, such as nanoparticles (Shrestha et al. 2015; Blum et al. 2006), molecular biology (Gordon et al. 2009), polymer dynamics (Chirikjian 2009[Section 13.7]), filtering in aeronautics: guidance and tracking (Willsky 1974). We refer to Chirikjian (2012) and Chirikjian (2009) for more applications. One source of models for stochastic dynamics stems from the so-called rotational Brownian motion of molecules. Rotational Brownian motion comprises the random change in the orientation of a polar molecule due to collisions with other molecules and is an important element in the theory of dielectric materials. Perrin and Debye’s non-inertial theories are the most well-known models, see for example Chirikjian (2009)[section 16.3]. Rotational Brownian motions have also been observed in a laboratory setting and have been properly documented in Han et al. (2006). Much of the current research in rotational Brownian motions has been devoted to inertial models, non-spherical molecules and the possibility of dipole-dipole interactions. Walter et al. (2010) took a step further in proposing an inertial, Langevin type of generalisation to the rigid body equations aiming at studying systems of rigid bodies as models for polymer dynamics. The coupling between linear and rotational dynamics was important in this case, to capture the motion features of long polymeric chains. Their models assume linearity in the noise for both linear and angular momentum variables, whereas the model used here is fully nonlinear with multiplicative noise and preserves strong geometrical features such as the coadjoint orbits.

Remark 5.1 (The LLG equation) We mention that the stochastic Landau-Lipschitz-Gilbert (LLG) equation studied for example in Garanin (1997), Brzeźniak et al. (2012) and Kohn et al. (2005) has the same structure as our stochastic dissipative rigid body equation. Indeed, we preserve the coadjoint orbit, thus the amplitude of the momentum variable, which corresponds to the strength of magnetic moments in the LLG model. We will not study this link further here as the LLG equation is a PDE in two or three dimensions and requires different analytical methods than the rigid body equation.

5.1 The Stochastic Rigid Body

The canonical example for illustrating the Euler–Poincaré reduction by symmetry is the free rigid body, whose configuration space is the group of rotations $SO(3)$. For a complete treatment from the viewpoint of reduction, we refer to Marsden and Ratiu (1999). For simplicity here, we rely on the isomorphism $\mathfrak{so}(3) \cong \mathbb{R}^3$ which translates the commutator in the Lie algebra to the cross product of three-dimensional vectors, via $[A, B] \rightarrow \mathbf{A} \times \mathbf{B}$, where $\mathbb{R}^3$ vectors are denoted with bold font. This map allows us to use a slightly different Killing form than the canonical one. Namely, we shall use the scalar product as our pairing, via the formula $\mathbf{A} \cdot \mathbf{B} = -\frac{1}{2}\kappa(\mathbf{A}, \mathbf{B})$.

The reduced Lagrangian of the free rigid body is written in terms of the angular velocity $\Omega \in \mathfrak{so}(3)$ and a prescribed moment of inertia $\mathbb{I} \in \text{Sym}(3)$ as

$$l(\Omega) = \frac{1}{2} \Omega \cdot \mathbb{I} \Omega := \frac{1}{2} \Omega \cdot \Pi, \quad (5.1)$$

where the angular momentum Π is defined accordingly and the Legendre transform gives the reduced Hamiltonian $h(\Pi) = \frac{1}{2} \Pi \mathbb{I}^{-1} \Pi$. We take the stochastic potential to be linear in the momentum variable Π

$$\Phi_i(\Pi) = \sum_{i=0}^3 \sigma_i \cdot \Pi, \quad (5.2)$$

where the constants σ_i generically span $\mathbb{R}^3$ but can be chosen in various ways. The stochastic process for Π is then computed from (2.8) to be

$$d\Pi + \Pi \times \Omega dt + \sum_i \Pi \times \sigma_i \circ dW_t^i = 0, \quad (5.3)$$

and the corresponding Itô process is

$$d\Pi + \Pi \times \Omega dt + \frac{1}{2} \sum_i (\Pi \times \sigma_i) \times \sigma_i dt + \sum_i \Pi \times \sigma_i dW_t^i = 0. \quad (5.4)$$

The coadjoint orbit defined by a level set of the quadratic Casimir $\|\Pi\|^2 = c^2$ is preserved in our geometrical construction, as may be checked by a direct computation in both the Stratonovich and the Itô stochastic representations. Although the Casimir is conserved, the energy $h(\Pi) = l(\Omega)$ is not a conserved quantity in general. Indeed, since the moment of inertia, $\mathbb{I}$ is a symmetric matrix, the stochastic process associated to h can be found to be

$$dh = \sum_i (\Pi \times \sigma_i) \cdot [\mathbb{I}^{-1}(\Pi \times \sigma_i) - (\Omega \times \sigma_i)] dt + 2 \sum_i (\Pi \times \sigma_i) \cdot \Omega dW_t^i. \quad (5.5)$$

In the general case, one only has bounds for the energy given by the two stable equilibrium positions of the rigid body, namely $E_{\min} = \frac{1}{2I_3} |\Pi_3(0)|^2$ and $E_{\max} = \frac{1}{2I_1} |\Pi_1(0)|^2$ if $I_1 \leq I_2 \leq I_3$. Thus, the energy may randomly fluctuate within these bounds.

Apart from the obvious case of $\mathbb{I} = Id$, one can check that the system with $\mathbb{I} = (I_1, I_1, I_3)$ and $\sigma = (0, 0, \sigma_3)$ conserves the energy for every values of I_1, I_3 and σ_3 . In this case, the stochastic rigid body reduces to the Kubo oscillator of Kubo et al. (1991)

$$d\Pi_1 = \Pi_2(a\Pi_3 dt + \chi_3 \circ dW), \quad d\Pi_2 = -\Pi_1(a\Pi_3 dt + \chi_3 \circ dW) \quad \text{and} \quad d\Pi_3 = 0,$$

where $a := \frac{I-I_3}{I I_3}$. This system is integrable by quadratures and a solution is

$$\begin{aligned}\Pi_1(t) &= \Pi_1(0) \cos(\gamma t + \chi_3 W_t) - \Pi_2(0) \sin(\gamma t + \chi W_t) \\ \Pi_2(t) &= \Pi_2(0) \cos(\gamma t + \chi_3 W_t) + \Pi_1(0) \sin(\gamma t + \chi W_t),\end{aligned}$$

where $\gamma := a\Pi_3$. Although the deterministic free rigid body is integrable, the only known integrable stochastic rigid body is this particular case.

5.2 The Fokker–Planck Equation

The Fokker–Planck equation of the process (5.3) is simply given for a probability density $\mathbb{P}$ by

$$\frac{d}{dt}\mathbb{P} + (\mathbf{\Pi} \times \mathbf{\Omega}) \cdot \nabla \mathbb{P} + \frac{1}{2} \sum_i (\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla [(\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla \mathbb{P}] = 0, \quad (5.6)$$

where ∇ is the gradient with respect to the independent variable $\mathbf{\Pi} \in \mathbb{R}^3$. According to Theorem 2.5, the stationary distribution $\mathbb{P}_\infty$ is constant on coadjoint orbits, which are spheres.

Based on this result, more can be said about the energy evolution of the stochastic rigid body, without embarking on any deeper studies into the coupled stochastic processes (5.5) and (5.3). For example, by ergodicity of (5.3), the long time average of the stochastic rigid body motion follows the limiting distribution $\mathbb{P}_\infty$. In terms of energy, the distribution is not uniform but will be proportional, at a given energy, to the length of the deterministic trajectory of the rigid body with this energy. The energy will thus randomly oscillate between two bounds, with maximum probability to be near the energy of the unstable equilibrium.

5.3 Double Bracket Dissipation

The double bracket dissipation for the rigid body involves the only Casimir $\|\mathbf{\Pi}\|^2$ and yields, with noise, the dissipative stochastic process

$$d\mathbf{\Pi} + \mathbf{\Pi} \times \mathbf{\Omega} dt + \theta \mathbf{\Pi} \times (\mathbf{\Pi} \times \mathbf{\Omega}) dt + \sum_i \mathbf{\Pi} \times \boldsymbol{\sigma}_i \circ dW_t^i = 0. \quad (5.7)$$

The Itô formulation is similar to (5.4) and will not be written here. The corresponding the Fokker–Planck equation is

$$\frac{d}{dt}\mathbb{P} + (\mathbf{\Pi} \times \mathbf{\Omega}) \cdot (\nabla \mathbb{P} - \theta \mathbf{\Pi} \times \nabla \mathbb{P}) + \frac{1}{2} \sum_i (\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla [(\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla \mathbb{P}] = 0. \quad (5.8)$$

The Fokker–Planck equation for stochastic rigid body dynamics with selective decay may be found by specialising the general proof given for Theorem 3.4. Indeed, we can rewrite the Fokker–Planck equation as

$$\frac{d}{dt}\mathbb{P} + (\mathbf{\Pi} \times \mathbf{\Omega}) \cdot \nabla \mathbb{P} + \nabla \cdot \left(\theta \mathbf{\Pi} \times (\mathbf{\Pi} \times \mathbf{\Omega}) \mathbb{P} - \frac{1}{2} \sigma^2 \mathbf{\Pi} \times (\mathbf{\Pi} \times \nabla \mathbb{P}) \right) = 0, \quad (5.9)$$

where we have used $\nabla \cdot (\mathbf{\Pi} \times (\mathbf{\Pi} \times \mathbf{\Omega})) = 0$. The last term in (5.8) simplifies as

$$\sum_i (\mathbf{\Pi} \times \mathbf{e}_i) [(\mathbf{\Pi} \times \mathbf{e}_i) \cdot \nabla \mathbb{P}] = \sum_i (\mathbf{\Pi} \times \mathbf{e}_i) [(\nabla \mathbb{P} \times \mathbf{\Pi}) \cdot \mathbf{e}_i] = \mathbf{\Pi} \times (\nabla \mathbb{P} \times \mathbf{\Pi}),$$

since the sum over i is simply the decomposition of the vector $(\nabla \mathbb{P} \times \mathbf{\Pi})$ into its $\mathbf{e}_i$ components. Consequently, the asymptotic equilibrium solution tends to

$$\mathbb{P}_\infty = Z^{-1} e^{-\frac{\theta}{\sigma^2} h(\mathbf{\Pi})}, \quad (5.10)$$

in which the overall sign of the exponential argument is negative, since $\theta > 0$.

We illustrate the Gibbs measure in Fig. 1a obtained by Monte–Carlo sampling. To illustrate the assumption of isotropic noise necessary to obtain the Gibbs measure, we performed the same simulation but with a non-isotropic noise, displayed on Fig. 1b. The distribution is no longer symmetric, and its analytical form is difficult to obtain. Nevertheless, it remains close to a Gibbs measure.

5.4 Random Attractor

For $\mathfrak{so}(3)$, we can go beyond Theorem 4.5 to obtain the exact value of the sum of the Lyapunov exponents.

Fig. 1 We display two invariant measures of the Fokker–Planck equation of the stochastic free rigid body. We plot in *dark blue* the low density regions and yellow the high density regions. The *left panel* shows a simulation with an isotropic noise which illustrates the Gibbs measure and the *right panel* shows a non-isotropic noise with a near Gibbs measure. The *white lines* show a few deterministic trajectories of the rigid body. **a** Isotropic noise with $\sigma \propto (1, 0, 0)$, $\sigma \propto (0, 1, 0)$ and $\sigma \propto (0, 0, 1)$. **b** Non-isotropic noise with $\sigma \propto (1, 0, 0)$, and $\sigma \propto (0, 0, 1)$ (Color figure online)

Proposition 5.2 *The sum of the Lyapunov exponents can be given exactly as*

$$\sum_i \lambda_i = -3\sigma^2 - \theta \left(c^2 \text{Tr } \mathbb{I}^{-1} - 6\mathbb{E}_\infty h \right), \quad (5.11)$$

where c is the value of the Casimir function, and $\theta > 0$.

Proof We can compute the term A of Theorem 4.5 exactly with the structure constants $c_k^{ij} = \epsilon_{ijk}$

$$\begin{aligned} -A(\Pi, \Pi) &= -\text{Tr}(\text{ad}_\Pi \text{ad}_\Pi \mathbb{I}^{-1}) = -c_n^{im} c_m^{jn} \mathbb{I}^{-1} \Pi_i \Pi_j \\ &= \Pi_1^2 (\mathbb{I}_2^{-1} + \mathbb{I}_3^{-1}) + \Pi_2^2 (\mathbb{I}_1^{-1} + \mathbb{I}_3^{-1}) + \Pi_3^2 (\mathbb{I}_1^{-1} + \mathbb{I}_2^{-1}), \end{aligned}$$

which, when combined with the Hamiltonian, yields the result in equation (5.11). $\square$

We now turn to the numerical estimation of the top Lyapunov exponent for the stochastic damped rigid body. More explanation of the numerical scheme used can be found in the appendix A. The result is displayed in Fig. 2 where we sampled θ and σ with 0.1 steps and used a spline interpolation for smoothing data. This result must not be taken to be exact, since, for example, the regions of large or small noise are the least accurate, as larger noise requires smaller time steps and a sufficiently small noise loses the ergodicity property sooner, as the simulations are run for a finite time. Nevertheless, these results numerically demonstrate that the top Lyapunov exponent is positive over a large region of the parameter space (θ, σ^2) . Based on these numerical results, it is of course not possible to show that the observed chaos is not transient, but the longer runs suggest that the positive top Lyapunov exponent reaches a stable constant value.

The green iso-temperature lines (higher temperatures on the left) in Fig. 2 are tangent to the iso-top-lyapunov lines in the upper left quadrant, where the top Lyapunov exponent is positive. Perpendicular isolines indicate that these two variables can be used as a coordinate system in this figure. Thus we can change from (θ, σ) to (λ_+, T) . In fact, we already know that we can use (λ_+, λ_-) as a coordinate system, which means that in the upper left region of the plot we have proportionality, $\lambda_- \propto T$.

Fig. 2 This figure displays the value of the top Lyapunov exponent of the stochastic damped rigid body with $\mathbb{I} = (1, 2, 3)$ and $c = 1$ in the parameter space (θ, σ) . This clearly shows a large region of positive Lyapunov exponents, implying chaotic behaviour in the system. See the text for a more detailed discussion

Apart from demonstrating that the top Lyapunov exponent can be positive, this figure also provides us with a better understanding of this system, which we now discuss briefly.

- For $\theta = 0$, the attractor is the entire space, that is the momentum sphere, and the stationary distribution is uniform on it. No random attractor or SRB measure exists in this case. Nevertheless, we learn that by increasing the noise, we first observe an increase of the amount of chaos, and then a decrease. The decrease for large noise is rather slow, and it is not expected that the top Lyapunov exponent will ever become negative under further increase of the noise. This is because the shear of the rigid body is bounded by the difference between the two opposite moment of inertia. In turn, the magnitude of the top Lyapunov exponent is bounded to the extent that it is linked to this shear. This is in contrast with random attractors in the plane, or more generally in non-compact spaces, which can possess an arbitrarily large shear. For example, in the case of planar systems with limit cycles studied numerically by Lin and Young (2008) and analytically by Engel et al. (2016), the top Lyapunov is not bounded for large noise amplitudes.
- The limit $\sigma \rightarrow 0$ cannot be numerically computed, but it is easy to extrapolate it from this graphic. First, notice that if $\theta = \sigma = 0$, we are in the Hamiltonian case of rigid body dynamics; hence, the sum of the Lyapunov exponents must be zero and the top Lyapunov exponent is therefore positive. Furthermore, the Lyapunov exponents will depend on the initial conditions, as the system is not ergodic anymore, so the limit $\sigma \rightarrow 0$ is not very well defined. Nevertheless, if $\theta > 0$, the top Lyapunov exponent converges to a single negative value.
- Upon examining the plots for various values of the Lyapunov exponents, one notices that the dark region of negative Lyapunov exponent varies rapidly with the parameters θ and σ , whereas the rest of the plot shows slower variations. It is interesting to remark that the slope of the line $\lambda^+ = -1$, for example, is close to 0.4, which is smaller than 1. That means that to balance an increase in the noise for some value, the damping must be increased by a larger value.

To conclude, in light of this numerical result, and along with the theorem for the existence of SRB measures, provided σ_i spans $\mathbb{R}^3$, the proper dissipation of energy will imply the existence of a non-singular random attractor.

The final analysis at this stage of the investigation concerns the nature of the random attractors of this system. From numerical simulations, we display in Fig. 3 a realisation of a random attractor of the rigid body.¹

The plots in Fig. 3 show the SRB measure, in log scale and exhibit the phenomena of stretching and folding, typical of strange attractors with positive and negative Lyapunov exponents. The positive exponent produces the stretching mechanism, and the negative one produces the folding process. Asymptotically in time, these mechanisms may create a fractal structure, similar to the Smale horseshoe structure for the Duffing attractor. The mechanisms for the creation of the rigid body random attractor can be understood from the underlying rigid body dynamics. The heteroclinic orbits, linking the two saddle equilibrium points which correspond to the direction of the

¹ See <http://wwwf.imperial.ac.uk/~aa10213/> for a video of this random attractor.

Fig. 3 The four panels display snapshots of the same rigid body random attractor with $\mathbb{I} = \text{diag}(1, 2, 3)$, $\theta = 0.5$ and $\sigma = 0.5$. The simulation started with a uniform distribution of rigid bodies on the momentum sphere and created finer and finer structures. The *colour* is in log scale, and we simulated 400,000 rigid body initial conditions with a split-step scheme, see “Appendix A”

second moment of inertia are the longest orbits, with the fastest dynamics along them. The speed of the motion for each orbit then decreases to reach the stable equilibrium points, associated with the largest or smallest moment of inertia. This change in speed creates a shear on a given non-singular set evolving with the rigid body dynamics. Together with the compactness of the sphere, the combination of noise and dissipation produces the complicated structure of the attractor. One can remark that the attractor of the Duffing oscillator is similar and provides a good example in the deterministic context of the creation of these structures.

A more detailed study of this random attractor will certainly be interesting but is out of the present scope of this work as it would require deeper dynamical systems analysis. Nevertheless, we will briefly study in the next section a simplification of this model which considers periodic kicking instead of Brownian motion.

5.5 Periodic Kicking

We finish this section devoted to the stochastic three-dimensional rigid body by making the simple replacement of the noise by a periodic kicking. It turns out that this type of forced dynamical systems also shows chaotic behaviours, and furthermore the theoretical understanding of this phenomenon is more advanced than for the pure noise case. We refer to Wang and Young (2003) and Lin and Young (2008) and references therein for such studies.

The periodic kicking is achieved by simply replacing the noise dW_t^i by the sum of Dirac delta functions

$$dW_t^i \Rightarrow \sigma_i \sum_{n=1}^{\infty} \delta(t - nT), \quad (5.12)$$

where T is the period for the kicking and $\sigma \in \mathbb{R}^3$ represents the amplitude and direction of the kick, as in the noisy system. It is interesting to remark that the kicking corresponds to a rotation around the fixed axis σ and with an angle proportional to $\|\sigma\|$. Notice that in the noisy case, the axis of rotation was also random, so it was not relevant to consider it for our studies. The double bracket dissipation term is still there; thus, some attractors are expected to emerge, but they will surely not be random. In fact, due to the periodic kicking, the notion of attractor must be modified slightly. Recall that in the noisy case, to have a fixed attractor in time, we needed the notion of pullback attractor. Here, we will fix the attractor by just observing it at discrete times nT . Indeed, between each kick, the dynamics relaxes following the damped deterministic rigid body equation.

We will not attempt a theoretical study of this system but rather illustrate its complexity with the help of numerical investigations. Since the parameter space is rather large, we will just highlight the most typical behaviour of the system while only varying the amplitude of the kick, $\|\sigma\|$ for fixed $T = 1$, $\theta = 0.2$, $\sigma = \|\sigma\|(1, 1, 1)$ and $\mathbb{I} = \text{diag}(1, 2, 3)$. The fact that σ is not aligned with any eigenvalue of $\mathbb{I}$ makes the configuration of the rigid body generic enough for the present study. Remarkably, this simple system can undergo many different types of dynamical behaviour by simply varying the kicking amplitude $\|\sigma\|$.

We have numerically scanned the attractors for various values of $\|\sigma\|$ and have displayed the results in Fig. 4, which we will analyse qualitatively below.

The first thing to remark to understand these different types of behaviour is that the kick will rotate the momentum of the rigid body in the same direction as the original rigid body flow for the lower right region, and in the opposite direction for the upper left region in Fig. 4. For this reason, the effect of the periodic kicking will be different on the opposite sides of the sphere, and this will lead to two different types of attractors near these regions. We now describe the different types of attractors which appear in this system, upon varying the kicking amplitude. Let $\mathcal{A}^+$ denote the attractor where the dynamical flow and the kicking are in the same direction and let $\mathcal{A}^-$ denote the other attractor. They will be understood as being the same if only one attractor emerges. Recall that all the parameters are fixed, except the kicking amplitude $\|\sigma\|$. We thus classify the different regimes according to an ordered sequence $\epsilon_0 < \dots < \epsilon_5$ of values of $\|\sigma\|$, that we will find later.

The analysis was done by looking at individual trajectories of rigid bodies and the values of the Lyapunov exponents, which we display in Fig. 5. From all this data, we can estimate the values of the ϵ in the ordered sequence as approximatively (0.02, 0.08, 0.11, 0.13, 0.16), as illustrated in Fig. 5.

The different regimes are as follow.

- (1) $\|\sigma\| < 0.02$: Both attractors $\mathcal{A}^+$ and $\mathcal{A}^-$ are singular and near the original rigid body equilibrium positions.
- (2) $0.02 < \|\sigma\| < 0.08$: The points near the region of the previous $\mathcal{A}^-$ eventually reach the $\mathcal{A}^+$ region to form one singular attractor near the previous $\mathcal{A}^+$. Before

Fig. 4 This figure displays several attractors of the periodically kicked rigid body for the different regimes parametrised by the kicking amplitude $\|\sigma\|$ as described in the text. Each snapshot is taken after the system has reached an equilibrium position, except for $\|\sigma\| = 0.8$. In this case, we have illustrated the transient limit cycle structure which quickly disappears as the solution converges to a singular attractor. For $\sigma = 0.12$, the three singular attractors (except the *lower right one*) which are slightly thicker are in fact a single attractor, and the kicking makes these points periodically switch amongst each other. The *last panel* with $\sigma = 0.18$ shows a similar behaviour for a single attractor composed of two points that switch between each other

Fig. 5 We display the result of the numerical computation of the top Lyapunov exponent for the kicked rigid body as a function of the kicking amplitude $\|\sigma\|$ for the possibly two attractors. The $\mathcal{A}^+$ attractor is always singular and thus always has a negative top Lyapunov exponent, whereas the other attractor sometimes shows chaos on the limit cycle

the collapse of the point near $\mathcal{A}^-$ to the region near $\mathcal{A}^+$, we observe a short transient chaos, close to being a limit cycle. The chaos is revealed by a positive Lyapunov exponent during this period.

- (3) $0.08 < \|\sigma\| < 0.11$: The attractor $\mathcal{A}^-$ is a limit cycle which passes near both original equilibrium of minimum energy and thus is driven by the kicks on both sides of the sphere. The chaos on this limit cycle is not clear from the Lyapunov exponent computation, as the top Lyapunov exponent is very close to 0, see Fig. 5. The other attractor $\mathcal{A}^+$ is singular, near its previous position.
- (4) $0.11 < \|\sigma\| < 0.13$: The attractor $\mathcal{A}^-$ consists of 3 singular points at roughly equal distance from the kicking axis $(-1, -1, -1)$. These three points form a single attractor as the kick makes then periodically switch between themselves. The other attractor $\mathcal{A}^+$ is still singular.
- (5) $0.13 < \|\sigma\| < 0.17$: The attractor $\mathcal{A}^-$ consists of a chaotic limit cycle centred around the kicking axis $(-1, -1, -1)$ and $\mathcal{A}^+$ is still singular. As compared to the previous limit cycle, this one remains near the region where the kicking is opposite to the flow direction and has a stronger chaos, as seen from the Lyapunov exponent computation, see Fig. 5.
- (6) $0.17 < \|\sigma\| < 0.18$: The last region explored here shows that both attractors merge to a single attractor that consists of two fix points. The periodic kicking switches one to the other, as in case (4).

From these findings, the most remarkable result is not the existence of chaos on the limit cycles, but rather the existence of the limit cycles themselves. The chaos can be understood in the same way as for the stochastic case, namely by the shear of the system. The existence of a stable limit cycle is in fact rather subtle as it requires a fine balance between the kicking, the shear and the damping of the rigid body. A precise analytical estimation for the emergence of such limit cycles is of course out of the scope of this work, and we leave it for further studies.

6 Semidirect Product Example: The Stochastic Heavy Top

The classic example of semidirect product motion is the heavy top, which arises in the presence of gravity when the support point of a freely rotating rigid body is no longer at its centre of mass. The starting phase space for the heavy top is $T^*SO(3)$, just as for the free rigid body. When the support point is shifted away from the centre of mass, gravity breaks the symmetry, and the system is no longer $SO(3)$ invariant. Consequently, the motion can no longer be written entirely in terms of the body angular momentum $\mathbf{\Pi} \in \mathfrak{so}(3)^*$. One also needs to keep track of the unit vector $\mathbf{\Gamma}$, the “direction of gravity” as seen from the body ($\mathbf{\Gamma} = \mathbf{R}^{-1}\mathbf{k}$ where the unit vector $\mathbf{k}$ points upward in space and $\mathbf{R}$ is the element of $SO(3)$ describing the current configuration of the body). The variable $\mathbf{\Gamma}$ may be identified with elements in the coset space $SO(3)/SO(2)$, where $SO(3)$ is the symmetry broken by introducing a special vertical direction for gravity, and $SO(2)$ is the remaining symmetry. This $SO(2)$ is the isotropy subgroup of $SO(3)$ corresponding to rotations around the unit vector $\mathbf{k}$ which leave the direction of gravity invariant.

6.1 The Stochastic Heavy Top

The Lagrangian for the heavy top is the difference between the kinetic energy and the work against gravity, where the fixed vector χ represents the position of the centre of mass of the body with respect to the fixed point. In body coordinates, the reduced Lagrangian is

$$l(\Omega, \Gamma) = \frac{1}{2} \Omega \cdot \mathbb{I} \Omega - mgl \Gamma \cdot \chi. \quad (6.1)$$

We refer to see [Holm et al. \(1998\)](#) and [Marsden and Ratiu \(1999\)](#) for a complete description of the semidirect product reduction for the heavy top, which we will not explain here. The stochastic potential will be taken to be linear in both the Γ and Π :

$$\Phi_i(\Gamma, \Pi) = \sigma_i \cdot \Pi + \eta_i \cdot \Gamma, \quad (6.2)$$

where σ_i and η_i need not span $\mathbb{R}^3$. The stochastic process describing the stochastic heavy top is then

$$\begin{aligned} d\Pi + (\Omega dt + \sum_i \sigma_i \circ dW_t^i) \times \Pi + mgl(\Gamma \times \chi) dt + \sum_i mgl(\Gamma \times \eta_i) \circ dW_t^i &= 0, \\ d\Gamma + (\Omega dt + \sum_i \sigma_i \circ dW_t^i) \times \Gamma &= 0, \end{aligned} \quad (6.3)$$

and the corresponding Itô process is

$$\begin{aligned} d\Pi + (\Omega dt + \sum_i \sigma_i dW_t^i) \times \Pi + (\Gamma \times mgl\chi) dt \\ + \sum_i mgl(\Gamma \times \eta_i) \circ dW_t^i - \frac{1}{2} \sum_i \sigma_i \times (\sigma_i \times \Pi) dt &= 0, \\ d\Gamma + (\Omega dt + \sum_i \sigma_i dW_t^i) \times \Gamma - \frac{1}{2} \sum_i \sigma_i \times (\sigma_i \times \Gamma) dt &= 0. \end{aligned} \quad (6.4)$$

The two Casimirs of the heavy top are conserved, $\|\Gamma\|^2 = k$ and $\Pi \cdot \Gamma = c$. However, the energy is not conserved, as it satisfies the following stochastic process

$$\begin{aligned} \frac{d}{dt} E &= \frac{1}{4} \sum_i \left[\Omega \cdot (\sigma_i \times (\sigma_i \times \Pi)) + \Pi \cdot \mathbb{I}^{-1}(\sigma_i \times (\sigma_i \times \Pi)) \right] dt \\ &+ \frac{1}{2} \sum_i \left[(\Pi \times \sigma_i) \cdot \mathbb{I}^{-1}(\Pi \times \sigma_i) - mgl(\sigma \times \Gamma) \cdot (\chi \times \sigma_i) \right] dt \\ &+ \frac{1}{2} \sum_i \left[\Omega \cdot (\Pi \times \sigma_i) + \Pi \cdot \mathbb{I}^{-1}(\Pi \times \sigma_i) + 2\chi \cdot (\Gamma \times \sigma_i) \right] dW_t^i. \end{aligned} \quad (6.5)$$

The energy being only bounded from below, this stochastic process can lead to arbitrary large value for the energy, over a long enough time.

6.2 The Integrable Stochastic Lagrange Top

When $\mathbb{I}$ is of the form $\mathbb{I} = \text{diag}(I_1, I_1, I_3)$ and $\chi = (0, 0, \chi_3)$, the deterministic heavy top is called the Lagrange top and is integrable. The integrability comes from the extra conserved quantity $\Pi \cdot \chi$, in this case. For noise, the stochastic process for this quantity is

$$\frac{d}{dt}(\Pi \cdot \chi) = -\frac{1}{2} \sum_i (\chi \times \sigma_i) \cdot (\sigma_i \times \Pi) dt - \sum_i \chi \cdot (\Pi \times \sigma_i) dW_t^i, \quad (6.6)$$

which is not a conserved quantity in general. However, the form of this equation implies that if one selects $\sigma_i = \chi$ then $\Pi \cdot \chi$ is a conserved quantity. It is remarkable that with this choice of noise, the energy is also a conserved quantity, as one can check from equation (6.5). We thus have a stochastic integrable Lagrange top, with a stochastic Lax pair given by

$$d(\lambda^2 \chi + \lambda \Pi + \Gamma) = ((\lambda \chi + \Omega)dt + \chi \circ dW) \times (\lambda^2 \chi + \lambda \Pi + \Gamma), \quad (6.7)$$

where λ is arbitrary and called a spectral parameter. We refer to [Ratiu \(1981\)](#) for more details about the integrability of the Lagrange top. Following the framework of integrable hierarchies, further developed for infinite dimensional integrable hierarchies in [Arnaudon \(2015\)](#), there exists another integrable stochastic Lagrange top where the stochastic potential is the same as the Hamiltonian. The explanation for the integrability is straightforward, as the change of variable $t \rightarrow t + W_t$ maps the stochastic Lagrange top to the deterministic one; so we will not discuss it in more detail here.

We want to study this stochastic system further, as integrability means that an explicit solution can be found. Indeed, from the standard theory of the heavy top, see for example [Arnol'd \(1989\)](#) and [Audin \(1996\)](#), the equation for Γ_3 can be found to be of the form $\dot{\Gamma}_3^2 = f(\Gamma_3)$, where f depends only on the constants of motion k and c . Then, a straightforward calculation with Euler angles gives

$$\begin{aligned} \dot{\psi} &= \frac{c - k\Gamma_3}{(1 - \Gamma_3^2)I} \\ d\phi &= \left[\frac{c}{I_3\Gamma_3} - \frac{c - k\Gamma_3}{I_3\Gamma_3(1 - \Gamma_3^2)I} ((1 - \Gamma_3^2)I - I_3\Gamma_3^2) \right] dt - \chi_3 \circ dW, \end{aligned} \quad (6.8)$$

where $\cos(\theta) = \Gamma_3$ gives the third Euler angle. Surprisingly, only ϕ has a stochastic motion, while ψ and θ follow the deterministic Lagrange top motion. This is illustrated in Fig. 6 via a numerical integration of the stochastic Lagrange top equations. The conservation of all the Lagrange top quantities is reproduced, as well as the fact that the noise only influences the ϕ component of the Euler angles.

Fig. 6 This figure displays a realisation of the motion of the integrable stochastic Lagrange top. The conserved quantities are displayed in the *right panel*. **a** Γ motion, **b** conserved quantities

6.3 The Fokker–Planck Equation and Stationary Distributions

We now analyse the associated Fokker–Planck equation for the stochastic heavy top, which is given by

$$\begin{aligned}
 \frac{d}{dt} \mathbb{P}(\mathbf{\Pi}, \mathbf{\Gamma}) &= (\mathbf{\Pi} \times \mathbf{\Omega}) \cdot \nabla_{\mathbf{\Pi}} \mathbb{P} + (\mathbf{\Gamma} \times \mathbf{\Omega}) \cdot \nabla_{\mathbf{\Gamma}} \mathbb{P} \\
 &+ \sum_i \frac{1}{2} (\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla_{\mathbf{\Pi}} [(\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla_{\mathbf{\Pi}} \mathbb{P}] \\
 &+ \sum_i \frac{1}{2} (\mathbf{\Gamma} \times \boldsymbol{\sigma}_i) \cdot \nabla_{\mathbf{\Gamma}} [(\mathbf{\Gamma} \times \boldsymbol{\sigma}_i) \cdot \nabla_{\mathbf{\Gamma}} \mathbb{P}] \\
 &+ \sum_i \frac{1}{2} (\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla_{\mathbf{\Pi}} [(\mathbf{\Gamma} \times \boldsymbol{\sigma}_i) \cdot \nabla_{\mathbf{\Gamma}} \mathbb{P}] \\
 &+ \sum_i \frac{1}{2} (\mathbf{\Gamma} \times \boldsymbol{\sigma}_i) \cdot \nabla_{\mathbf{\Gamma}} [(\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla_{\mathbf{\Pi}} \mathbb{P}] ,
 \end{aligned} \tag{6.9}$$

where in our notation $\nabla_{\mathbf{\Pi}}$ denotes the gradient with respect to the $\mathbf{\Pi}$ variable only and similarly for $\nabla_{\mathbf{\Gamma}}$. By using the semidirect product Lie–Poisson structure of the heavy top

$$\{H, G\}_{HT} := \begin{bmatrix} \nabla_{\mathbf{\Pi}} H & \nabla_{\mathbf{\Gamma}} H \end{bmatrix} \begin{bmatrix} \mathbf{\Pi} \times \mathbf{\Gamma} \times \\ \mathbf{\Gamma} \times & 0 \end{bmatrix} \begin{bmatrix} \nabla_{\mathbf{\Pi}} G \\ \nabla_{\mathbf{\Gamma}} G \end{bmatrix} , \tag{6.10}$$

the Fokker–Planck equation (6.9) can be written in the double bracket form

$$\frac{d}{dt} \mathbb{P} = \{h, \mathbb{P}\}_{HT} + \frac{1}{2} \{\Phi, \{\Phi, \mathbb{P}\}_{HT}\}_{HT} , \tag{6.11}$$

where $h(\mathbf{\Pi}, \mathbf{\Gamma})$ is the Legendre transform of (6.1).

Fig. 7 We display the probability distribution of the norm of the momentum of the heavy top, as a function of time. The distribution tends to 0 as time goes to ∞ , but only linearly as shown by Eq. (6.12) and the white line in this Figure. The expansion is larger for small times, as the distribution is not yet uniform on the angles of the momentum but linearly bounded in time. After this rapid early expansion, the diffusion slows considerably

Recall that the stationary marginal distribution on the Γ sphere is constant. We study here the distribution in the Π coordinate, following the general argument of Theorem 2.6, which gives the bound

$$0 \leq \|\Pi\|(t) \leq \|\Pi_0\| + (mglc)t. \quad (6.12)$$

This bound increases linearly with time and is unbounded only when $t \rightarrow \infty$. This effect is clearly illustrated in the Fig. 7 where the probability distribution of $\|\Pi\|^2$ is plotted. The initial conditions are uniform distributions on the Γ sphere and a single position for all the momentum, with unit norm. Our system parameters are $m = g = c = 1$. Consequently, the linear bound is directly proportional to the time. According to Fig. 7, the bound is reached almost immediately in the first stage of the diffusion, where the Γ and Π sphere are not yet uniformly covered. After this first short temporal regime, however, the diffusion rate slows considerably below this linear bound.

6.4 Random Attractor

The dissipative heavy top equations can be computed directly from the semidirect theory (see also Bloch et al. 1996) and in Stratonovich form they read, when the Casimir $\Pi \cdot \Gamma$ is used,

$$\begin{aligned} d\Pi + (\Omega dt + \sum_i \sigma_i \circ dW_t^i) \times \Pi + mgl(\Gamma \times \chi)dt \\ + \theta \Gamma \times (\Omega \times \Gamma)dt + \theta [mgl\Pi \times (\chi \times \Gamma) - \Pi \times (\Pi \times \Omega)]dt = 0, \\ d\Gamma + (\Omega dt + \sum_i \sigma_i \circ dW_t^i) \times \Gamma + \theta [mgl\Gamma \times (\chi \times \Gamma) - \Gamma \times (\Pi \times \Omega)]dt = 0. \end{aligned} \quad (6.13)$$

Notice that the two Casimirs which define the coadjoint orbits are preserved by both the noise and the dissipation, as expected. Also, recall the form of the deterministic energy decay

$$\frac{dh}{dt} = -\theta \|\boldsymbol{\Omega} \times \boldsymbol{\Gamma}\|^2 - \theta \|\boldsymbol{\Omega} \times \boldsymbol{\Pi} + mgl\boldsymbol{\chi} \times \boldsymbol{\Gamma}\|^2, \quad (6.14)$$

which was used earlier to prove the existence of the random attractor after a nonlinear change of variables. The other Casimir $\|\boldsymbol{\Gamma}\|^2$ can also be used to derive dissipative equations, but energy dissipation will be slower, as only the first term in (6.14) and the first decay term of the $d\boldsymbol{\Pi}$ are left. The equilibrium solution of the purely dissipative system are found by setting the right-hand side of (6.14) to 0 and are always of the form $\boldsymbol{\Omega} = \boldsymbol{\Gamma} = \boldsymbol{\chi}$ if $\boldsymbol{\chi}$ is an eigenvalue of $\mathbb{I}$. If not, the equilibrium is aligned to another direction that we will not compute here, as we will stick to the simple generic case.

We can compute the lower bound for the value of the sum of the Lyapunov exponents using Theorem 4.7 to find

$$\sum_i \lambda_i \geq -6\sigma^2 - \theta c^2 \text{Tr}(\mathbb{I}^{-1}), \quad (6.15)$$

which is always negative. We will not study here the parameter space of the top Lyapunov exponent as done for the rigid body, but only display two instances of an attractor of the heavy top in Figure 8.² The formation of the attractors seems not to be of horseshoe type, as occurs for the rigid body. This may be explained by the higher dimensionality of the coadjoint orbit (dimension 4) on which the attractor is supported.

7 Two Other Examples

This section briefly sketches two other stochastic symmetry-reduced examples of the present theory which follow immediately from the examples of the $SO(3)$ rigid body and the heavy top, treated in the previous sections. These are the $SO(4)$ rigid body and the spring pendulum.

7.1 The $SO(4)$ Rigid Body

For a complete study of the rigid body motion on $SO(4)$ we refer to Birtea et al. (2012) and references therein. We use the generic elements

$$X = \begin{pmatrix} 0 & x_1 & x_2 & x_3 \\ -x_1 & 0 & x_4 & -x_5 \\ -x_2 & -x_4 & 0 & x_6 \\ -x_3 & x_5 & -x_6 & 0 \end{pmatrix},$$

² See <http://www.imperial.ac.uk/~aa10213/> for a video of this random attractor.

Fig. 8 We display here three projections of the attractor on the coadjoint orbit of the heavy top at two different times. The *left panel* is a projection on the sphere $\|\Gamma\|^2 = 1$, the *second panel* on the sphere $\|\Pi\|^2 = 1$ and the *third panel* is the amplitude of the momentum. We used $\theta = 0.2$, $\sigma = 0.1$, $\mathbb{I} = \text{diag}(1, 2, 3)$, $g = 1$ and 20,000 realisations of the stochastic heavy top, with initial conditions uniformly distributed on a subset of the coadjoint orbit defined by $\|\Pi\|^2 = 1$

or $X = (X_1, X_2) \in \mathbb{R}^6$. In terms of vectors $(X_1, X_2) \in \mathbb{R}^6$ and $(X'_1, X'_2) \in \mathbb{R}^6$ we have

$$[(X_1, X_2), (X'_1, X'_2)] = (X_1 \times X'_1 + X_2 \times X'_2, X_1 \times X'_2 + X_2 \times X'_1).$$

The coadjoint action is the same, under the trace-pairing. The Casimir for $SO(4)$ are given by

$$C_1 = \text{Tr}(X^2) = \sum_i x_i^2 = \|X_1\|^2 + \|X_2\|^2,$$

$$C_2 = \sqrt{\det(X)} = x_1 x_6 + x_2 x_5 + x_3 x_4 = X_1 \cdot X_2.$$

The first Casimir is a 4-dimensional sphere and the second is the Pfaffian, or scalar product between two vectors. The momentum-velocity relation is $\Pi = J\Omega + \Omega J$ where $J = \text{diag}(\lambda_1, \dots, \lambda_6)$ and the Hamiltonian $H(\Pi) = \frac{1}{2}(\Pi_1 \cdot \Omega_1 + \Pi_2 \cdot \Omega_2)$. We thus have the following stochastic 4-dimensional rigid body equations

$$d(\Pi_1, \Pi_2) = (\Pi_1 \times \Omega_1 + \Pi_2 \times \Omega_2, \Pi_1 \times \Omega_2 + \Pi_2 \times \Omega_1) dt$$

$$+ \sum_i \left(\Pi_1 \times \sigma_1^i + \Pi_2 \times \sigma_2^i, \Pi_1 \times \sigma_2^i + \Pi_2 \times \sigma_1^i \right) \circ dW_i, \quad (7.1)$$

which preserve the coadjoint orbit.

We now look at the selective decay term for the Casimir $C_2(\Pi) = \Pi_1 \cdot \Pi_2$. This term reads, upon using semi-simplicity,

$$SD = \text{ad}_{(\Pi_2, \Pi_1)} \text{ad}_{(\Pi_2, \Pi_1)}(\Omega_1, \Omega_2)$$

$$= (\Pi_2 \times (\Pi_2 \times \Omega_1 + \Pi_1 \times \Omega_2) + \Pi_1 \times (\Pi_2 \times \Omega_2 + \Pi_1 \times \Omega_1),$$

$$= (\Pi_2 \times \Pi_2 \times \Omega_1 + \Pi_2 \times \Pi_1 \times \Omega_2 + \Pi_1 \times \Pi_2 \times \Omega_2 + \Pi_1 \times \Pi_1 \times \Omega_1,$$

$$\Pi_2 \times \Pi_2 \times \Omega_2 + \Pi_2 \times \Pi_1 \times \Omega_1 + \Pi_1 \times \Pi_2 \times \Omega_1 + \Pi_1 \times \Pi_1 \times \Omega_2).$$

One can directly check that the first Casimir C_1 is also preserved by this flow.

Proposition 7.1 *This stochastic dissipative $SO(4)$ free rigid body admits a random attractor.*

Proof This is a direct application of the theory developed in Sect. 3. □

The stationary distribution will be centred around the minimal energy position, associated with the direction of the maximal moment of inertia. We will not numerically investigate the random attractors for this system here. However, further theoretical studies are indeed possible, and these would be interesting to discuss elsewhere, especially the integrable case, with a particular choice of the noise.

7.2 Spring Pendulum

From the heavy top equation, one can derive the spherical pendulum by letting one of the components of the diagonal inertia tensor in body coordinates tend to zero, e.g., $I_3 \rightarrow 0$. This follows, because the spherical pendulum is infinitely thin and, hence, does not have any inertia for rotations around its axis. We shall choose $\mathbb{I} = \text{diag}(I, I, \epsilon)$ in the heavy top equations and then take the limit $\epsilon \rightarrow 0$ so that the dynamics on Π_3 vanishes. The similarity of this system with the rigid body allows us to consider an extension of the spherical pendulum which is called the spring pendulum (Lynch 2002). To include the dynamics of the length of the spring pendulum, we introduce a new variable $R(t) \in \mathbb{R} \setminus \{0\}$ and enforce its dynamical evolution in the variational principle by adding $P(\dot{R} - v)dt$ where v denotes the velocity of the mass along the pendulum, and P denotes its associated momentum. The Lagrangian is then found to be

$$l(\Omega, \Gamma, R, v) = \frac{m}{2} R^2 \Omega \cdot \mathbb{I} \Omega - mg R \Gamma \cdot \chi + \frac{1}{2} m v^2 - \frac{k}{2} (R - 1)^2, \quad (7.2)$$

where χ represents the initial position of the pendulum which is taken to be $(0, 0, 1)$ in accordance with our choice of inertia tensor. In (7.2), we denote the spring constant by k and the mass of the pendulum by m .

We shall assume a general linear stochastic potential of the form,

$$\Phi(\Pi, \Gamma, R, P) := \sigma \cdot \Pi + \eta \cdot \Gamma + \alpha R + \beta P, \quad (7.3)$$

for constant vectors σ, η , and constant scalars α, β . Consequently, the stochastic spring pendulum equations are given by

$$\begin{aligned} d\Pi &= \Pi \times \Omega dt + mg R \Gamma \times \chi dt + \Pi \times \sigma_i \circ dW_t^i + \Gamma \times \eta_i \circ dW_t^i \\ d\Gamma &= \Gamma \times \Omega dt + \Gamma \times \sigma_i \circ dW_t^i \\ dR &= \frac{P}{m|\chi|^2} dt + \beta dW_t \\ dP &= -mg \Gamma \cdot \chi dt - k(R - 1)|\chi|^2 dt + \frac{1}{mR^3} \Pi \cdot \mathbb{I}^{-1} \Pi - \alpha dW_t. \end{aligned} \quad (7.4)$$

The analysis above is valid, provided $\epsilon > 0$ in the inertia tensor. In the limit $\epsilon \rightarrow 0$, we may set $\Omega_3 = 0$ and thereby recover the stochastic elastic spherical pendulum equations.

The equation set in (7.4) consists of two parts: the stochastic heavy top equations, coupled to a pair of stochastic canonical Hamilton equations for the (R, P) variables. The coupling between the two subsets of equations occurs through the dependence on R together with Ω and Γ in the Lagrangian (7.2).

The Fokker–Planck equation is now easily derived, and it reads

$$\begin{aligned} \frac{d}{dt} \mathbb{P} = & \{H, \mathbb{P}\}_{HT} + \{H, \mathbb{P}\}_{can} + \frac{1}{2} \{\Phi, \{\Phi, \mathbb{P}\}_{HT}\}_{HT} \\ & + \{\Phi, \{\Phi, \mathbb{P}\}_{HT}\}_{can} + \frac{1}{2} \{\Phi, \{\Phi, \mathbb{P}\}_{can}\}_{can}, \end{aligned} \quad (7.5)$$

where $\{\cdot, \cdot\}_{can}$ is the canonical Poisson bracket with respect to the (R, P) variables. The coupling between the elastic and pendulum motions is too complicated to extract any information from the Fokker–Planck equation. Indeed, inspection of the motion on (R, P) shows that the advection equation for (R, P) depends on the other variables. This inextricable complex dependence precludes finding the limiting distribution explicitly, despite the simple Laplacian form of the diffusion operator.

As pointed out by Lynch (2002), the deterministic elastic spherical pendulum system is a toy model for the lowest modes of atmosphere dynamics. For this application, the motion of the spring oscillations encoded in R is considerably faster than the pendulum motion and smaller in amplitude. Averaging the deterministic Lagrangian over the relatively rapid oscillations of the spring yields a nonlinear resonance between the modes of a type which also appears in the atmosphere. The noise can be included in either of the two types of dynamics, and each will influence the other through the nonlinear coupling. Also, for small oscillations around the equilibrium, the deterministic nonlinear coupling produces star-shaped orbits (Holm and Lynch 2002; Lynch 2002), which can be perturbed, or even entirely destroyed, by the introduction of the noise, depending on its amplitude.

8 Conclusion and Open Problems

Before stating some open problems arising from this work, we will briefly summarise it. In the first section, we reviewed and developed the new machinery of stochastic geometric mechanics, in the context of finite dimensional systems which admit a group of symmetry of semi-simple type. The main results emerged from the introduction of a particular type of noise that preserves the coadjoint motion of the deterministic equations. The associated Fokker–Planck equation was found to possess interesting geometrical properties, related to the Lie–Poisson formulation of the equation of motion. The Lie–Poisson formulation was used to derive its stationary solution which is constant on the level set of the Casimirs prescribed by the initial conditions. The second section was devoted to the introduction of dissipation with the double bracket term, for which the coadjoint orbits are still preserved by the flow of the equa-

tion. This particular combination of multiplicative noise and nonlinear dissipation on coadjoint orbits yields non-constant invariant measures often referred to as Gibbs measures. This type of invariant measure makes an interesting connection to statistical physics and naturally, provides us with a notion of temperature for these systems. The second outcome of the noise-dissipation interaction is the existence of so-called random attractors, which are mathematical objects deeply connected to the theory of random dynamical systems. We demonstrated, by adapting the standard tools from the random dynamical system theory, that such objects do exist in the mechanical systems studied here. Furthermore, we gave conditions on the dissipation and noise amplitudes for the existence of non-singular such attractors which will in turn support an SRB measure. The next two sections were devoted to the application of this theory to the standard examples in geometric mechanics, which are the free rigid body and the heavy top for the semidirect product extension, also developed here. We studied these explicit stochastic processes in detail, and in particular with illustrative numerical simulations. The final section touched upon other related examples such as a spring pendulum which can be viewed as an extension of the heavy top with a direct product structure, and a higher dimensional rigid body, written on $SO(4)$.

We now end by listing several open problems which have been formulated during the course of this work.

- Some of the results presented here relied on the assumption of compactness of the coadjoint orbits; for example, in estimating the sum of the Lyapunov exponents, and for the study of the stationary distributions. This assumption is probably unnecessary, but properly addressing its removal would require more advanced mathematical tools than we have used here.
- We were only able to obtain a numerical demonstration that the top Lyapunov exponent is positive. This numerical demonstration could be made considerably more refined, and possibly analytical results could also be derived in future studies.
- We only touched upon the analysis of random attractors via numerical simulations, but much more may be said about these objects by, for example, studying their Lyapunov exponents in more detail, and studying the underlying dynamical process of their formation. This is motivated by the fact that even in the two simple illustrative examples of random attractors treated here, we observed two rather different solution behaviours.
- Although we restricted ourselves to semisimple Lie algebras, we were able to write most of the equations for more general Lie algebras. Hence, the present line of reasoning should be valid for other similar systems by modifying the proofs accordingly. Examples of such systems would include semidirect products with arbitrary advected quantities, solvable or nilpotent Lie algebras, the Toda lattice and possibly infinite dimensional Lie groups such as the diffeomorphism group.

Acknowledgements We are grateful to many people for fruitful and encouraging discussions, including S. Albeverio, J.-M. Bismut, N. Bou-Rabee, M. D. Chekroun, G. Chirikjian, D. O. Crisan, A. B. Cruzeiro, J. Eldering, M. Engel, N. Grandchamps, P. Lynch, J. Newman, J.-P. Ortega, G. Pavliotis, V. Putkaradze, T. Ratiu and C. Tronci. The simulations were run with the Imperial College High Performance Computing Service. We also acknowledge the Bernoulli Centre for Advanced Studies at EPFL where parts of this work were elaborated. AA acknowledges partial support from an Imperial College London Roth Award and AC

from a CAPES Research Award BEX 11784-13-0. All the authors are also supported by the European Research Council Advanced Grant 267382 FCCA held by DH.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A: Numerical Scheme

Here we briefly discuss our numerical scheme for the integration of the stochastic rigid body equations. The heavy top being a direct extension of this scheme, we will not treat it here.

We use a split-step scheme where the deterministic part of the equation is integrated with the Python solver *odeint* of Scipy that used the classic scheme *lsoda*. The stochastic term is integrated via the exact solution of $d\Pi = \Pi \times \circ d\mathbf{W}$ where $d\mathbf{W} = \sum_i \sigma_i dW_t^i$, as

$$\Pi(t + dt) = e^{\widehat{d\mathbf{W}}} \Pi(t), \quad (\text{A.1})$$

where $\widehat{d\mathbf{W}}$ is corresponding anti-symmetric matrix, or $\mathfrak{so}(3)$ Lie algebra element, and the exponential is the matrix exponential.

The computation of the top Lyapunov exponent is done with this scheme extended to also compute the evolution of the linearisation, $\delta\Pi$. The form of the equation for $\delta\Pi$ is similar to the rigid body equation, and the previous split-step algorithm is implemented in the same way. Despite having a good preservation of the coadjoint orbit for the original stochastic process, this scheme does not accurately preserve the restriction of the linearisation to be tangent to the coadjoint orbit. The scheme *lsoda* being implicit, we need an interpolation between $\Pi(t)$ and $\Pi(t + dt)$, which is achieved by a spherical linear interpolation, or *slerp*. Even with this interpolation, the condition $\delta\Pi \cdot \Pi = 0$ is not precise enough and leads to an incorrect estimate of the norm $\|\delta\Pi\|$ in the computation of the Lyapunov exponent. We thus project out the part of $\delta\Pi$ aligned with Π at each time step such that $\delta\Pi$ remains in the tangent space to the coadjoint orbit. This projection gives coherent results, but it affects the convergence of the different estimations of the top Lyapunov exponent. Namely, for each different initial condition, the corresponding top Lyapunov exponent will take more time to converge to its true value. We overcome this effect by averaging over several realisations of the top exponent computation, which is 50 in our case.

References

- Ana Bela D.D.H., Ratiu, T.S.: Momentum maps and stochastic Clebsch action principles. arXiv preprint [arXiv:1604.04554](https://arxiv.org/abs/1604.04554) (2016)
- Arnaudon, M., Cruzeiro, A.B.: Lagrangian Navier–Stokes diffusions on manifolds: variational principle and stability. *Bull. Sci. Math.* **136**(8), 857–881 (2012)
- Arnaudon, M., Chen, X., Cruzeiro, A.B.: Stochastic Euler–Poincaré reduction. *J. Math. Phys.* **55**(8), 081507 (2014)

- Arnol'd, V.I.: *Mathematical Methods of Classical Mechanics*, volume 60 of Graduate Texts in Mathematics. Springer-Verlag, New York, 1989. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second edition (1989)
- Arnold, L.: *Random Dynamical Systems*. Springer, Berlin (1995)
- Arnaudon, A.: The integrable stochastic AKNS hierarchy. arXiv preprint [arXiv:1511.07080](https://arxiv.org/abs/1511.07080) (2015)
- Audin, M.: *A course on integrable systems*. Spinning Tops, Volume 51 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (1996)
- Blum, J., Bruns, S., Rademacher, D., Voss, A., Willenberg, B., Krause, M.: Measurement of the translational and rotational brownian motion of individual particles in a rarefied gas. *Phys. Rev. Lett.* **97**, 230601 (2006)
- Birtea, P., Caşu, I., Ratiu, T.S., Turhan, M.: Stability of equilibria for the $so(4)$ free rigid body. *J. Nonlinear Sci.* **22**(2), 187–212 (2012)
- Bonatti, C., Díaz, L.J., Viana, M.: *Dynamics Beyond Uniform Hyperbolicity: A Global Geometric and Probabilistic Perspective*, vol. 102. Springer, Berlin (2006)
- Brzeźniak, Z., Goldys, B., Jegaraj, T.: Large deviations for a stochastic Landau-Lifshitz equation, extended version. arXiv preprint [arXiv:1202.0370](https://arxiv.org/abs/1202.0370) (2012)
- Bismut, J.-M.: *Mécanique aléatoire*. In: 10 Saint Flour Probability Summer School—1980 (Saint Flour, 1980), volume 929 of Lecture Notes in Math., pp. 1–100. Springer, Berlin (1982)
- Bloch, A., Krishnaprasad, P.S., Marsden, J.E., Ratiu, T.S.: The Euler-Poincaré equations and double bracket dissipation. *Commun. Math. Phys.* **175**(1), 1–42 (1996)
- Bou-Rabee, N., Owadi, H.: Stochastic variational integrators. *IMA J. Numer. Anal.* **29**(2), 421–443 (2009)
- Chen, X., Cruzeiro, A.B., Ratiu, T.S.: Constrained and stochastic variational principles for dissipative equations with advected quantities. arXiv preprint [arXiv:1506.05024](https://arxiv.org/abs/1506.05024) (2015)
- Crauel, H., Debussche, A., Flandoli, F.: Random attractors. *J. Dyn. Differ. Equ.* **9**(2), 307–341 (1997)
- Crauel, H., Flandoli, F.: Attractors for random dynamical systems. *Probab. Theory Relat. Fields* **100**(3), 365–393 (1994)
- Crauel, H., Flandoli, F.: Additive noise destroys a pitchfork bifurcation. *J. Dyn. Differ. Equ.* **10**(2), 259–274 (1998)
- Crisan, D., Flandoli, F., Holm, D.D.: Solution properties of a 3d stochastic Euler fluid equation. arXiv preprint [arXiv:1704.06989](https://arxiv.org/abs/1704.06989). (2017)
- Cotter, C.J., Gottwald, G.A., Holm, D.D.: Stochastic partial differential fluid equations as a diffusive limit of deterministic Lagrangian multi-time dynamics. arXiv preprint [arXiv:1706.00287](https://arxiv.org/abs/1706.00287) (2017)
- Chirikjian, G.S.: *Classical results and geometric methods. Stochastic models, information theory, and Lie groups*. Vol. 1. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc, Boston (2009)
- Chirikjian, G.S.: *Analytic methods and modern applications. Stochastic Models, Information Theory, and Lie Groups*. Volume 2. Applied and Numerical Harmonic Analysis. Springer, New York (2012)
- Crauel, H.: Markov measures for random dynamical systems. *Stoch.: Int. J. Probab. Stoch. Process.* **37**(3), 153–173 (1991)
- Chekroun, M.D., Simonnet, E., Ghil, M.: Stochastic climate dynamics: random attractors and time-dependent invariant measures. *Phys. D: Nonlinear Phenom.* **240**(21), 1685–1700 (2011)
- Engel, M., Lamb, J.S.W., Rasmussen, M.: Bifurcation analysis of a stochastically driven limit cycle. arXiv preprint [arXiv:1606.01137](https://arxiv.org/abs/1606.01137) (2016)
- Garanin, D.A.: Fokker-Planck and Landau-Lifshitz-Bloch equations for classical ferromagnets. *Phys. Rev. B* **55**, 3050–3057 (1997)
- Gay-Balmaz, F., Holm, D.D.: Selective decay by Casimir dissipation in inviscid fluids. *Nonlinearity* **26**(2), 495 (2013)
- Gay-Balmaz, F., Holm, D.D.: A geometric theory of selective decay with applications in MHD. *Nonlinearity* **27**(8), 1747 (2014)
- Gay-Balmaz, F., Holm, D.D.: Stochastic geometric models with non-stationary spatial correlations in lagrangian fluid flows. arXiv preprint [arXiv:1703.06774](https://arxiv.org/abs/1703.06774) (2017)
- Gay-Balmaz, F., Ratiu, T.S.: Clebsch optimal control formulation in mechanics. *J. Geom. Mech.* **3**(1), 41–79 (2011)
- Gordon, D., Hoyles, M., Chung, S.-H.: Algorithm for rigid-body brownian dynamics. *Phys. Rev. E* **80**, 066703 (2009)
- Gawedzki, K., Kupiainen, A.: Universality in turbulence: an exactly solvable model. In: *Low-dimensional models in statistical physics and quantum field theory*, pp. 71–105. Springer (1996)

- Han, Y., Alsayed, A.M., Nobili, M., Zhang, J., Lubensky, T.C., Yodh, A.G.: Brownian motion of an ellipsoid. *Science* **314**(5799), 626–630 (2006)
- Holm, D.D., Kupersmidt, B.A.: Poisson brackets and Clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity. *Phys. D* **6**(3), 347–363 (1983)
- Holm, D.D., Lynch, P.: Stepwise precession of the resonant swinging spring. *SIAM J. Appl. Dyn. Syst.* **1**(1), 44–64 (2002). (**electronic**)
- Holm, D.D., Marsden, J.E., Ratiu, T.S.: The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.* **137**(1), 1–81 (1998)
- Holm, D.D.: Geometric mechanics. Part II. Rotating, Translating and Rolling. Imperial College Press, London (2008). (distributed by World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ)
- Holm, D.D.: Variational principles for stochastic fluid dynamics. *Proc. R. Soc. Lond. Math. Phys. Eng. Sci.* **471**(2176), 20140963 (2015)
- Holm, D., Schmah, T., Stoica, C.: Geometry, Symmetry and Mechanics. Oxford University Press, Oxford (2009)
- Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. Elsevier, Amsterdam (2014)
- Kondrashov, D., Chekroun, M.D., Ghil, M.: Data-driven non-Markovian closure models. *Phys. D* **297**, 33–55 (2015)
- Kloeden, P.E., Platen, E.: Stratonovich and Itô stochastic Taylor expansions. *Math. Nachr.* **151**(1), 33–50 (1991)
- Kloeden, P.E., Rasmussen, M.: Nonautonomous dynamical systems. Number 176. American Mathematical Soc. (2011)
- Kraichnan, R.H.: Anomalous scaling of a randomly advected passive scalar. *Phys. Rev. Lett.* **72**, 1016–1019 (1994)
- Kohn, R.V., Reznikoff, M.G., Vanden-Eijnden, E.: Magnetic elements at finite temperature and large deviation theory. *J. Nonlinear Sci.* **15**(4), 223–253 (2005)
- Kubo, R., Toda, M., Hashitsume, N.: Nonequilibrium statistical mechanics. Statistical physics. II, Volume 31 of Springer Series in Solid-State Sciences, 2nd edn. Springer, Berlin (1991)
- Lázaro-Camí, J.-A., Ortega, J.-P.: Stochastic Hamiltonian dynamical systems. *Rep. Math. Phys.* **61**(1), 65–122 (2008)
- Lin, K.K., Shea-Brown, E., Young, L.-S.: Reliability of coupled oscillators. *J. Nonlinear Sci.* **19**(5), 497–545 (2009)
- Lu, K., Wang, Q., Young, L.-S.: Strange Attractors for Periodically Forced Parabolic Equations, vol. 224. American Mathematical Soc., Providence (2013)
- Ledrappier, F., Young, L.-S.: Entropy formula for random transformations. *Probab. Theory Relat. Fields* **80**(2), 217–240 (1988)
- Lin, K.K., Young, L.-S.: Shear-induced chaos. *Nonlinearity* **21**(5), 899 (2008)
- Lin, K.K., Young, L.-S.: Dynamics of periodically kicked oscillators. *J. Fixed Point Theory Appl.* **7**(2), 291–312 (2010)
- Lynch, P.: The swinging spring: a simple model for atmospheric balance. *Large-Scale Atmos-Ocean Dyn.* **2**, 64–108 (2002)
- Marsden, J.E., Ratiu, T.S.: A basic exposition of classical mechanical systems. Introduction to Mechanics and Symmetry, Volume 17 of Texts in Applied Mathematics, 2nd edn. Springer, New York (1999)
- Mikulevicius, R., Rozovskii, B.: On Equations of Stochastic Fluid Mechanics. Springer, Berlin (2001)
- Marsden, J., Weinstein, A.: Coadjoint orbits, vortices, and Clebsch variables for incompressible fluids. *Phys. D* **7**(1), 305–323 (1983)
- Poincaré, H.: Sur une forme nouvelle des équations de la mécanique. *C.R. Acad. Sci.* **132**, 369–371 (1901)
- Perelomov, A.M., Popov, V.S.: Casimir operators for semisimple Lie groups. *Math. USSR-Izvestiya* **2**(6), 1313 (1968)
- Ratiu, T.: Euler–Poisson equations on Lie algebras and the N -dimensional heavy rigid body. *Proc. Nat. Acad. Sci. USA* **78**(3, part 1), 1327–1328 (1981)
- Schenk-Hoppé, K.R.: Random attractors-general properties, existence and applications to stochastic bifurcation theory. *Discret. Cont. Dyn. Syst.* **4**, 99–130 (1998)
- Shrestha, S., Tiwari, S., Klar, A., Hardt, S.: Numerical simulation of a moving rigid body in a rarefied gas. *J. Comput. Phys.* **292**, 239–252 (2015)
- Thiffeault, J.-L., Morrison, P.J.: Classification and Casimir invariants of Lie–Poisson brackets. *Phys. D* **136**(3–4), 205–244 (2000)

- Varadarajan, V.S.: Lie Groups, Lie Algebras, and their Representations, Volume 102 of Graduate Texts in Mathematics. Springer, Berlin (1984)
- Villani, C.: Hypocoercivity. *Mem. Amer. Math. Soc.* **202**(950), iv+141 (2009)
- Walter, J., Gonzalez, O., Maddocks, J.H.: On the stochastic modeling of rigid body systems with application to polymer dynamics. *Multiscale Model. Simul.* **8**(3), 1018–1053 (2010)
- Willsky, A.S.: Estimation and detection of signals in multiplicative noise. In: IEEE Conference on Decision and Control including the 13th Symposium on Adaptive Processes, 1974, pp. 728–734 (1974)
- Wang, Q., Young, L.-S.: Strange attractors in periodically-kicked limit cycles and hopf bifurcations. *Commun. Math. Phys.* **240**(3), 509–529 (2003)
- Young, L.-S.: What are SRB measures, and which dynamical systems have them? *J. Stat. Phys.* **108**(5), 733–754 (2002)