Matemática
Contemporânea
Vol. 58, 264–315
http://doi.org/10.21711/231766362023/rmc588

Effective diffusions with intertwined structures

Xue-Mei Li [1,2]

[1]Imperial College London, South Kensington, London SW7 2AZ, U.K.
[2]École Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland

Abstract. The aim of the article is to explore multi-scale stochastic differential equations (SDEs) on manifolds. In these equations the variables of interest evolve at their natural speeds and interact with other variables that evolve at a faster pace. The primary objective is to identify the effective motion. This is an autonomous equation, its solutions approximate the evolution of the slowly moving variables as the scale of speed separation approaches zero. While multi-scale SDEs with linear state spaces have been a popular research topic for several decades, their study on manifolds is a more recent development. In this work, we illustrate the derivation of multi-scale equations by incorporating geometric information, such as symmetries and conserved quantities. Subsequently, we delve into these examples to investigate their effective dynamics.

Keywords: perturbation, stochastic differential equations, averaging, stochastic homogenisation, geodesic flows, Hopf fibration, frame bundle, holonomy bundle.
2020 Mathematics Subject Classification: 60H10, 58J65, 37Hxx, 53B05.

Contents

Research partially supported by EPSRC grant EP/E058124/1 and EP/V026100/1,
e-mail: xue-mei.li@epfl.ch

1 Introduction

A. The evolution of particles over time is often subject to various constraints and exhibits symmetries and invariants. Equations describing such particle motions calls for modifications to account for the influence of different types and sizes of factors, resulting in increased complexity and the loss of certain invariances. The presence of constraints confines the evolution of objects under investigation to a sub-manifold within Euclidean space. To fully leverage the inherent geometry of stochastic systems, we make the assumption that the state space of stochastic processes takes the form of smooth manifolds.

Consider these quantities, which remain constant along the trajectories of the solutions of the original equation. They no longer maintain this constancy within the context of the modified equations. Nevertheless, their evolution proceeds at a slower rate when compared to that of the particles.

If the magnitude of the perturbation is of the order ϵ, where ϵ is a positive number which we shall take to be close to zero, these slow variables evolve at most at a speed of order ϵ. Continuing from this point, we study the evolution of the slow variables over a time interval of scale $\frac{1}{\epsilon}$ during which it is no longer possible to track all aspects of the motion. Instead, we take advantage of the ergodic properties of the fast variables to determine an autonomous equation that approximates the behaviour of the slow variables within the interacting system. When stochasticity is involved, this approach is known as *stochastic averaging*. It is actually a dynamic functional law of large numbers. If the limiting motion is constant, we observe the motion on an even longer time scale and study the fluctuations. In other words, we extract information from an otherwise complex system by examining a relatively simpler one, referred to as the effective motion. The effective equation comprises autonomous equations whose solutions approximate the slow variables as the scale separation parameter ϵ approaches zero. We illustrate this concept and procedure below with two examples of Hamiltonian systems.

B. Examples. Before proceeding further, we explore the example of perturbation of completely integrable Hamiltonian systems.

Example 1. Let $H : M \to \mathbb{R}$ be a non-trivial Hamiltonian function on a two-dimensional symplectic manifold, and let X_H be the corresponding Hamiltonian vector field. We consider $H(x_t)$ as the energy of a particle evolving according to the Hamiltonian vector field: $\dot{x} = X_H(x)$. Since $dH(X_H) = 0$, one has $H(x_t) = H(x_0)$ for all t. The particle remains in the energy orbit in which it started. Introducing action-angle coordinates $x_t = (I_t, \theta_t)$, we observe that $\dot{I}_t = 0$ and $\dot{\theta}_t = g(\theta_t)$.

To account for additional influences on the particle's evolution, apart from the Hamiltonian, one modifies the equation driven by the Hamiltonian, the new equation takes the form

$$\dot{x}_t^\epsilon = X_H(x_t^\epsilon) + \epsilon V(x_t^\epsilon).$$

Consequently, $\dot{H}(x_t^\epsilon) = \epsilon \, dH(V(x_t^\epsilon))$. In the action-angle coordinates, this can be written as a system of a slow /fast ODE. A simple example is of the form

$$\begin{cases} \dot{I}_t^\epsilon = \epsilon \, f(\theta_t^\epsilon, I_t^\epsilon), \\ \dot{\theta}_t^\epsilon = g(I_t^\epsilon). \end{cases} \tag{1.1}$$

where the initial values are chosen to be independent of ϵ, the angle variable θ takes values in the circle S^1, and f, g are real valued functions. We want to consider this random dynamical system over the period of $[0, \frac{1}{\epsilon}]$. This is equivalent to, by a change of time variable,

$$\begin{cases} \dot{I}_t^\epsilon = & f(\theta_t^\epsilon, I_t^\epsilon), \\ \dot{\theta}_t^\epsilon = & \frac{1}{\epsilon} g(I_t^\epsilon). \end{cases} \tag{1.2}$$

As $\epsilon \to 0$, the solutions of the equations above gets closer to the solution of the ODE driven by a vector field $\bar{f}$. The vector field $\bar{f}$ is obtained by averaging the angle variable over one period. When $M = \mathbb{R}^2$, this can be extended to random perturbations of Hamiltonian dynamics, from which a beautiful theory emerges [4, 14].

Example 2. Let's consider extending the concept presented in Example 1 to include time-dependent random energies. We start with an $2n$-dimensional symplectic manifold, denoted as M. To each smooth function $H : M \to \mathbb{R}$, we associate a Hamiltonian / symplectic vector field X_H. If ω represents the symplectic form, then these vector fields satisfy $\omega(X_H, \cdot) = dH$. In particular, $dH(X_H) = 0$. In the case where $M = \mathbb{R}^2$, we can take $\omega = dx \wedge dy$. For instance if $H = \frac{1}{2}(x^2 + y^2)$, then $X_H = J \cdot \nabla H$ where ∇H is the gradient of H and J is the matrix: $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Suppose that $\{H_i(x)\}$ are independent at every point x, except for a set of measure zero. Additionally, assume that $dH_i(H_j) = 0$ whenever $i \neq j$. Such a system is known as a completely integrable Hamiltonian system. Let us define H as the vector $(H_1, \ldots, H_m)$. While there are only a few known examples of completely integrable Hamiltonian systems, many dynamical systems are approximately integrable. Therefore, we explore an approximately integrable system with random energy given by $\sum_{i=1}^n H_i \dot{W}_t^i$, where $\dot{W}_t^i$ is a set of independent white noise processes (the integrals of these processes represent independent Brownian motions). Let K be a vector field, and set $g = dH(K)$. For a small parameter $\epsilon \ll 1$, we consider the following stochastic differential equation (SDE):

$$dy_t^\epsilon = \sum_{i=1}^n X_{H_i}(y_t^\epsilon) \circ dW_t^i + \epsilon K(y_t^\epsilon)\, dt. \tag{1.3}$$

Here, the symbol $\circ$ denotes the Stratonovich integral. The solutions of these equations satisfy the following equations:

$$\frac{d}{dt} dH(y_t) = \epsilon g(y_t).$$

In the above equation, $dH = (dH^1, \ldots, dH^n)$, and dH_i represents the differential of $H_i : M \to \mathbb{R}$. Since stochastic integrals are non-commutative, this model constitutes a non-trivial higher-dimensional integrable system when $n > 1$. This model was introduced in [22].

We introduce a coordinate chart $\phi : \bar{U} \to \mathbb{T}^m \times D$, defining action-angle variables such that $\tilde{H}_k := H_k \circ \phi^{-1}$ depends solely on the I coordinates. Here, $\mathbb{T}^m$ represents the m-dimensional torus, and D is a compact subset of $\mathbb{R}^m$. In these coordinates, $\omega_k^i(I_t) = \frac{\partial \tilde{H}_i(I)}{\partial I_k}$. We set $\theta_t := (\theta_t^1, ..., \theta_t^m)$ and $I_t = (I_t^1, ..., I_t^m)$. One can express this system as follows:

$$\begin{cases} dI_t^i = \epsilon K_I^i(\theta_t, I_t)dt, \\ d\theta_t^i = \sum_{k=1}^m \omega_k^i(I_t)dB_t^k + \epsilon K_\theta^i(\theta_t, I_t)dt \end{cases} \tag{1.4}$$

In this system, the action variable $I = (I^1, \ldots, I^n)$ evolves slowly in comparison to the angle variables θ_t^i.

Let $\mathcal{M}_c$ denote the subset of $\mathcal{M}$ consisting of $p \in \mathcal{M}$ such that $H(p) = c$. For each value of c, $\mathcal{M}_c$ is diffeomorphic to the torus $\mathbb{T}^m$. The conditions on the vector fields imply that $d\theta_t^i = \sum_{k=1}^m \omega_k^i(I_t)dB_t^k$ has an invariant probability measure μ_a on each energy level $\phi(\mathcal{M}_a) \simeq \mathbb{T}^m$. The energy level sets are invariant under the flow of Equation (1.3), in the unperturbed case ($\epsilon = 0$). Furthermore, the invariant measures ν_f depend smoothly on f.

To investigate the slow evolution of I_t^i (or equivalently, the energies H_i), we rescale time by $t \to t/\epsilon$ to examine large time scales and consider motion averaged over the fast angle variables. Theorem 3.3 from [22] states that the energies H converge to, up to its first exit time from the coordinate chart, to $f = (f_1, ..., f_m)$, which are defined to be the solutions to the following system of ordinary differential equations:

$$\begin{cases} \dfrac{df}{dt} = Q(f(t)) \\ f(0) = H(y_0). \end{cases}.$$

Here $\tilde{g} := g \circ \phi^{-1}$, $\tilde{I}(f) = I$ is the inversion of $\tilde{H}(I) = f$ and $Q(a) = \int_{\phi(\mathcal{M}_a)} \tilde{g}(\theta, \tilde{I}(a))\mu_f(d\theta)$. The error bound given there was revised to $\epsilon^{\frac{1}{4}}$ in [35], where various interesting examples of Hamiltonian systems were investigated. The precise statement is: Theorem [22, Th. 3.3]. Let T^b be the first time at which the trajectory y_t leaves the domain U. Then for any $\beta > 1$ there exists a constant $c(t)$ such that

$$\mathbb{E}\left[\sup_{s\le t}\left|H_i(y_{\frac{s}{\epsilon}\wedge T^\epsilon}) - f_i(s\wedge(\epsilon T^\epsilon))\right|^\beta\right]^{1/\beta} \le c(t)\epsilon^{1/4}.$$

To conclude this example, we assume that the perturbation vector field K is furthermore a Hamiltonian vector field, the effective dynamics is trivial and over a longer time scale and we have a diffusion creation theory. Let $K = X_k$ where $k : M \to \mathbb{R}$ is a smooth function, then $Q(f)$ vanishes we study the following stochastic dynamics as the separation $\epsilon \to 0$ on the scale of $\frac{1}{\epsilon^2}$:

$$dI_t^i = \frac{1}{\epsilon} K_I^i(\theta_{\frac{t}{\epsilon^2}}, I_t)dt$$

$$d\theta_t^i = \sum_{k=1}^m \omega_k^i(I_t)dB_t^k + \epsilon K_\theta^i(\theta_t, I_t)dt.$$

Furthermore the effective dynamics for (1) is a Markov process, see [22, Th. 4.1].

C. Exploring Invariants and Symmetries. Beyond the scope of stochastic differential equations generated by Hamiltonian functions lies a rich landscape of reducible differential equations. These systems often reveal invariants, primarily stemming from underlying symmetries within their dynamics. For instance, consider rigid motions; they are invariant under rotational transformations. However, the existence of these conserved quantities isn't always immediately apparent and often becomes evident only through projection onto a quotient space. The action of a group brings in the structure of a principal bundle structure to the state space. The projection of perturbed equation helps us to separate the slowly varying and the rapidly varying variables in the perturbed equations. We delve into examples to show how these ideas operate and offer insights into potential avenues for future research.

D. Markov Processes with Diffusion-Type Generators. Our focus now shifts towards a specific class of stochastic processes: Markov processes with Markov generators of diffusion type. In this context, we consider a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ satisfying the usual assumptions. To define our system, we begin with smooth vector fields X_i and Y_i and construct Markov generators $\mathcal{L}_0$ and $\mathcal{L}_1$. We define $\mathcal{L}_0 = \frac{1}{2}\sum(X_i)^2 + X_0$ and $\mathcal{L}_1 = \frac{1}{2}\sum(Y_i)^2 + Y_0$.

Our objective is to examine the behaviour of $\mathcal{L}^\epsilon$-diffusions and consider its effective dynamics. Such $\mathcal{L}^\epsilon$-diffusions have a representation as the solution of the following system of stochastic differential equations (SDEs) :

$$dy_t^\epsilon = \frac{1}{\sqrt{\epsilon}} \sum_i X_i(y_t^\epsilon) \circ db_t^i + \frac{1}{\epsilon} X_0(y_t^\epsilon)dt + \sum_j Y_j(y_t^\epsilon) \circ dw_t^j + Y_0(y_t^\epsilon)dt, \qquad y_0^\epsilon = y_0$$

where (b_t^i, w_t^j) are independent one dimensional Brownian motions and $\circ$ denotes Stratonovich integration. One of the strategies is to seek out a quantity $F : M \to$

$\mathbb{R}$ such that it is invariant along the trajectories of (x_t) where

$$dx_t = \sum_i X_i(x_t) \circ db_t^i + Y_0(x_t)dt.$$

On a large time scale of order $1/\epsilon$, the deviation of the perturbed orbit becomes visible. Consequently, the variable $F(y_t^\epsilon)$ varies slowly in t when $\epsilon \to 0$. We can expect an effective dynamics when $F(y_{\frac{t}{\epsilon}}^\epsilon)$ converges as ϵ is taken to zero. The ansatz for the effective limit is a Markov process. It is therefore desirable to find all conserved quantities.

Let us write down the evolution equations :

$$\frac{\partial}{\partial t} f^\epsilon = (\frac{1}{\epsilon}\mathcal{L}^0 + \mathcal{L}_1) f^\epsilon.$$

We expand its solution f^ϵ in powers of ϵ:

$$f^\epsilon = \frac{1}{\epsilon} f_0 + f_1 + \epsilon f_2 + o(\epsilon).$$

Our goal is to identify the specific forms of f_0 and f_1 within this expansion. To accomplish this, we explore a novel aspect introduced by the intertwining property of two Markov operators. On the flat spaces, such problems are popular and remain so, see [18] and the following books [2, 4, 14, 36] and the references therein. This investigation promises insights on the behaviour of perturbed SDEs with intriguing geometric information of the system.

Some models presented in this article have large groups of symmetries rendering the dynamics over the length of $\frac{1}{\epsilon}$ trivial. As a consequence we work in a much longer time scale and employ an unconventional non-diffusive scaling which can easily mistaken as a wrong scale to use.

This article is based on the Arxiv article [24]. Since it was put on the Arxiv, I started to work on specific models to gauge interest, see [26, 27] and also [25]. In the summer of 2022 I attended the XXV Brazilian School of Probability and got in touch with the geometric stochastic dynamic group at Unicamp. This motivated me to polish up this article, as the models there remain relevant. Here we maintain the general framework of horizontal and vertical perturbation, but I removed some proofs. To conclude the introduction we refer to some recent work with limit theorems on manifolds [12, 17, 19, 31, 32, 37–39].

2 Main Results

We present the effective theory for several models of slow /fast stochastic differential equations. We shall explain the basic tools used to reduce a system of equations to slow/ fast systems and obtain an averaged system. As mentioned earlier, the symmetries of a dynamical system is fundamental in this theory. In those models, if the averaged dynamics vanishes (on the time interval $[0, \frac{1}{\epsilon}]$), which happen when the perturbed system has abundant symmetries, a diffusion creation theorem is expected.

2.1 Notation

Symmetries can be described by the action of a group. Let G be a Lie group. A right action R_g by $g \in G$ on a manifold P is a smooth map $R_g : P \to P$ such that $R_g(R_h u) = R_{hg}u$. It is convenient to denote $R_g u$ by ug, so $(ug)h = u(gh)$. For a dynamical system with a group of symmetries, it is convenient to use the principal bundle structure.

A principal bundle is a special fibre bundle, consisting of a total space P, a base space M, and a Lie group G acting on P, on the right. We denote by $\mathfrak{g}$ the Lie algebra of G. Both P and M are smooth manifolds. We denote $\pi : P \to M$ as the projection, and the fibre at x is the collection of element $\{\pi^{-1}(x)\}$.

A trivialization of a principal bundle is a collection of open sets $U \subset M$ such that the bundle over each set U, $\pi^{-1}(U)$, is expressed as $U \times G$. Let $u = (x, h)$ denote an element of P in a chart, then $ug = (x, hg)$. In particular $\pi(ug) = \pi(u)$.

An important principal bundle is the bundle of frames on a smooth manifold M. Given a Riemannian metric on M, the orthonormal frame bundle $\pi : OM \to M$ is a reduced bundle of the frame bundle. Let n denote the dimension of the manifold, we assume that $n > 1$. In this case, the structure group is the special orthogonal group $SO(n)$. Its Lie algebra $\mathfrak{g} = \mathfrak{so}(n)$ is the space of skew symmetric matrices. If M is oriented the orthonormal frame bundle consists of two components, in which case, we only need to consider the action by the component $SO(n)$ of the group that contains the identity. Let us assume that M is oriented.

The fibres of OM consist of isometries $u : \mathbb{R}^n \to T_{\pi(u)}M$. Let

$$O_x M = \{u : \mathbb{R}^n \to T_{\pi(u)}M\},$$

this is the fibre at x. The projection $\pi : OM \to M$ is the map taking a linear frame $u : \mathbb{R}^n \to T_xM$ to its base point x on the manifold.

Denote by $T\pi$ its differential $T\pi : TOM \to TM$ from the tangent space of OM to the tangent space of M. The vertical bundle is then defined as the fibre bundle whose fibre at x is the kernel of the linear map $T_u\pi : T_uOM \to T_{\pi(u)}M$.

To differentiate vector field on M we use a linear connection, allowing us to parallel transport a vector along a curve. A connection ∇ is Riemannian if it is compatible with the Riemannian metric, in this case we may neglect differentiating the inner products $x \mapsto \langle \cdot, \cdot \rangle_x$, when making computations. We are not restricted to Levi-Civita connections, allowing a non-vanishing torsion $\mathcal{T}$. We assume that the connection is complete which means that every geodesic extends to all finite time parameter. This is equivalent to every standard horizontal vector field being complete.

A linear connection on M is equivalent to a splitting of TOM, which means introducing a complementary horizontal bundle satisfying some constraints. We first define a connection 1-form $\varpi : T_uOM \to \mathfrak{so}(n)$ where $\mathfrak{so}(n)$ denotes the Lie algebra of $SO(n)$. The connection 1-form is adjoint invariance $(R_g)^*\varpi = \mathrm{ad}(g^{-1})\varpi$. Let A^* denote the vertical fundamental vector field on the orthonormal frame bundle generated by a vector $A \in \mathfrak{so}(n)$. It is defined by

$$A^*(u) = \lim_{t\to 0} u \exp(tA), \qquad u \in OM.$$

We define the value of ϖ on fundamental vertical vector fields by $\varpi(A^*) \equiv A$. The kernel of ϖ are called horizontal vectors.

Any vector in TOM can be written as the sum of a horizontal part and a vertical part. For any vector $v \in T_uOM$ and vector field U on OM, we define

$$\breve{\nabla}_v U = \varpi^{-1} d(\varpi(U))(v) + \theta^{-1} d(\theta(U))(v).$$

where θ denotes the canonical 1-form, we have $\theta_u : T_uOM \to \mathbb{R}^n$. This is a metric connection on OM.

Let us denote h_u the horizontal lift through u, it maps a vector in $T_{\pi(u)}M$ to a vector in T_uOM. It is in fact the right inverse to the projection T_uOM. To each $e \in \mathbb{R}^n$, we associate a standard (or basic) horizontal vector field H_e. This is defined by the formula:

$$H_e(u) := \mathfrak{h}_u(ue).$$

We shall take an orthonormal basis of $\mathbb{R}^n$, denoting it by $\{e_i\}$ and set

$$H_i = H_{e_i}.$$

The horizontal vector fields $\{H_i\}_{i=1}^n$ are known as either the basic horizontal vector fields or the fundamental horizontal vector fields. They span the horizontal bundle at every point. Finally we define the horizontal Laplacian $\Delta^H := \sum_{i=1}^n H_i H_i$ on the orthonormal frame bundle. It is a semi-elliptic operator with constant rank n.

2.2 Background

W define a bundle map $H : OM \times \mathbb{R}^n \to TOM$ as follows. For any $e \in \mathbb{R}^n$ and $u \in OM$, set:

$$H(u, e) = \sum_{i=1}^{n} H_i(u)\langle e, e_i\rangle. \tag{2.1}$$

Using this notation, $H_e(u) = \sum_{i=1}^n H_i(u)\langle e, e_i\rangle$. Note that for every $u \in OM$, $H(u, \cdot) : \mathbb{R}^n \to T_u OM$ is a linear map.

Similarly, if $b = (b^1, \ldots, b^n)$ is an n-dimensional Brownian motion, we introduce simpler notation for the stochastic differential:

$$H(u) \circ db_t := \sum_{i=1}^{n} H_i(u) \circ db_t^i.$$

The symbol $\circ$ denotes Stratonovich integral. The probability distribution of the solutions of $du_t = H(u_t) \circ db_t$ are independent of the choice of the orthonormal basis $\{e_i\}$.

2.2.1 Canonical Horizontal SDE and BMs.

In general we do not expect to find $n+1$ smooth vector fields $\{X_i\}$ such that the solutions of

$$dx_t = \sum_{i=1}^{n} X_i(x_t) db_t^i + X_0(x_t) dt$$

is a Brownian motion. A large dimensional driving noise is needed to construct a Brownian motion as the solution to an SDE on a generic manifold. The construction of such vector fields are in general not intrinsic, meaning they rely on other concepts such as isometric embeddings.

A construction of a Brownian motion, using only intrinsic properties of the Riemannian manifold, is given in [7] by the following canonical horizontal SDE

$$du_t = H(u_t) \circ db_t.$$

The solution u_t is the horizontal Brownian motion. By the rotational invariance of Gaussian measures, $\pi(u_t)$ is a Markov process. They are Brownian motions on M.

In this SDE, the dimension of the noise is equal to the dimension of the manifold, and $T\pi(H_e(u)) = ue$. The projection of H_e is not a vector field on M.

Let Z be vertical vector fields, so that $T\pi(Z) = 0$, and consider the SDE

$$du_t = H(u_t) \circ db_t + Z(u_t)\, dt.$$

We frequently rely on the following facts: The Lie bracket of two vertical vector fields is again vertical, indicating the vertical bundle is integrable. In contrast, the horizontal bundle is in general not integrable. For $e, \tilde{e} \in \mathbb{R}^n$ the vertical part of $[H(e), H(\tilde{e})]$ is given by the following formula: $\varpi([H(e), H(\tilde{e})]) = -2\Omega(H(e), H(\tilde{e}))$, where Ω is the curvature form. Additionally, if $A \in \mathfrak{so}(n)$, $[H(e), A^*]$ is again horizontal. This observation is a key element in our diffusion creation theorems, as illustrated in §4.1 for an example.

2.2.2 Perturbation to SDEs on OM.

In this section we study perturbations to SDEs on the orthonormal frame bundle over a manifold M. Let us consider a small perturbation of size ϵ to an horizontal SDE, denoting the solutions to the original SDE and to the perturbed SDE respectively by u_t and u_t^ϵ:

$$du_t^\epsilon = H(u_t^\epsilon) \circ db_t + \epsilon Z(u_t^\epsilon) dt.$$

Then for $t \in [0, 1]$, as $\epsilon \to 0$, the convergence $\pi(u_t^\epsilon) \to \pi(u_t)$ is expected. To be able to see the effect of the perturbation, we observe the dynamics for $t \in [0, \frac{1}{\epsilon}]$. This is equivalent to considering the equation

$$du_t^\epsilon = \frac{1}{\sqrt{\epsilon}} H(u_t^\epsilon) \circ db_t + Z(u_t^\epsilon) dt,$$

Here we take Z to be a vertical vector field. As an example, we take $Z(u) = A^*(u)$ where $A \in \mathfrak{so}(n)$. One may also consider

$$dx_t = Z(x_t) dt,$$

and

$$du_t^\epsilon = H(u_t^\epsilon) \circ db_t + \frac{1}{\sqrt{\epsilon}} Z(u_t^\epsilon) dt.$$

In both cases $\pi(u_t) = 0$. In the former, the the vertical part of u_t^ϵ corresponds to a slower movement, while in the latter, it represents the faster motion.

2.3 Main Results

Let $\epsilon > 0$ be a positive number, and let $\{b_t^l, w_t^j, 1 \le l \le m, 1 \le j \le p\}$ be pairwise independent one dimensional Brownian motions defined on a probability space. Consider a family of vertical vector fields $\{Z_j, 1 \le j \le p\}$ the orthonormal frame bundle OM over a manifold M. We assume that at each $u \in OM$, the set of vectors $\{Z_j(u)\}_{j=1}^p$ spans the vertical sub-space of T_uOM. Let $\{\mathbb{X}_l, 0 \le l \le m\}$ be a family of horizontal vector fields. We pay attention to two classes of horizontal vector fields: the basics horizontal vector fields and horizontal lifts of vector fields on M.

2.3.1 Perturbation to vertical motions

Let u^ϵ denotes the solution to the following equations with common initial values:

$$\begin{cases} du_t^\epsilon = \sqrt{\epsilon}\sum_{l=1}^m \mathbb{X}_l(u_t^\epsilon)\circ db_t^l + \epsilon\mathbb{X}_0(u_t^\epsilon)dt + \sum_{j=1}^p Z_j(u_t^\epsilon)\circ dw_t^j + Z_0(u_t^\epsilon)dt, \\ u_0^\epsilon = u_0. \end{cases} \tag{2.2}$$

We first use the connection 1-form ϖ and the canonical form θ to identify the slow component and the horizontal and the vertical components of u^ϵ.

For $g \in SO(n)$, L_g denotes its Left action on the Lie group $SO(n)$ on OM. Similarly R_g denotes the right action. We also use respectively ug for $R_g u$ and gu for $L_g u$. Note that $L_g : SO(n) \to SO(n)$ is a smooth map, we denote by TL_g its differential. Denote ϖ the connection 1-form. Then $\varpi[Z_j(ug_t)] \in \mathfrak{so}(n)$.

As we shall see, the slow exponent of the solution u_t^ϵ is its projection to the manifold, while the fast component is then represented by an SDE on the Lie group $SO(n)$ as follows:

$$dg_t = \sum_{j=1}^m TL_{g_t}\varpi[Z_j(ug_t)]\circ dw_t^j + TL_{g_t}\varpi[Z_0(ug_t)]dt.$$

See later sections for further detail. By assumption, for any u, the above equation has an invariant probability measure, which we denote by π_u.

Let $\pi_*\mathbb{X}$ denote the pushed forward vector field, meaning $\pi_*\mathbb{X}_l(u) = T\pi(\mathbb{X}_l(u))$. Set $\mathbf{X}_l(u,g) = TR_{g^{-1}}\mathbb{X}_l(u)$ and for any $g \in SO(n)$ we denote by $\mathcal{L}^g$ the Markov generator of the following SDE on OM: $d\tilde{u}_t = \sum_{l=1}^p \mathbf{X}_l(\tilde{u}_t, g)\circ db_t^l + \mathbf{X}_0(\tilde{u}_t, g)dt$. The following is taken from Theorem 4.2.

Theorem A. Assume that the vector fields $\{Z_i, \mathbb{X}_j\}$ and their covariant derivatives grow at most linearly at infinity. Suppose that M has positive injectivity radius. Let $x_t^\epsilon = \pi(u_t^\epsilon)$, we denote their horizontal lifts by $\tilde{x}_t^\epsilon$. Then $\tilde{x}_{\frac{t}{\epsilon}}^\epsilon$ converges in distribution to a Markov process, which we denote by $\bar{x}_t$. Consequently, $x_{\frac{t}{\epsilon}}^\epsilon$ converges. Furthermore, the generator of $\bar{x}_t$ is given by averaging the Markov generator of $\mathcal{L}^g$ with respect to π_u, meaning that for any test function $F : OM \to \mathbb{R}$, one has that

$$\bar{\mathcal{L}}F(u) = \int_{SO(n)} \mathcal{L}^g F(u)\pi_u(dg).$$

We now give two examples.

Example 2.1 (The Right Invariant Horizontal Vector Field Case)**.** Let $X_l, l = 0, 1, 2, \dots m$, be vector fields on M. Define $\mathbb{X}_l(u) = \mathfrak{h}_u(X_l(\pi(u)))$ and we have

$$du_t^\epsilon = \sqrt{\epsilon}\sum_{l=1}^p \mathbb{X}_l(u_t^\epsilon) \circ db_t^l + \epsilon\mathbb{X}_0(u_t^\epsilon)dt + \sum_{j=1}^m Z_j(u_t^\epsilon) \circ dw_t^j + Z_0(u_t^\epsilon)dt.$$

The projection $\pi(u_t^\epsilon)$ satisfies $dx_t^\epsilon = \sqrt{\epsilon}\sum_{l=1}^p X_l(x_t^\epsilon) \circ db_t^l + \epsilon X_0(x_t^\epsilon)dt$. For all ϵ, $x_{\frac{t}{\epsilon}}^\epsilon$ are $\frac{1}{2}\sum L_{X_i}L_{X_i} + L_{X_0}$-diffusions, independent of ϵ. The horizontal lifts $\tilde{x}_{\frac{t}{\epsilon}}$ of $x_{\frac{t}{\epsilon}}$ are $\frac{1}{2}\sum L_{\mathbb{X}_i}L_{\mathbb{X}_i} + L_{\mathbb{X}_0}$-diffusions.

Example 2.2 (The Canonical Horizontal Vector Field Case)**.**

$$du_t^\epsilon = \sqrt{\epsilon}\sum_{l=1}^n H_l(u_t^\epsilon) \circ db_t^l + \epsilon H_0(u_t^\epsilon)dt + \sum_{j=1}^m Z_j(u_t^\epsilon) \circ dw_t^j + Z_0(u_t^\epsilon)dt.$$

Since $\tilde{x}_t^\epsilon$ and u_t^ϵ belong to the same fibre, there exists a group element g_t^ϵ such that $\tilde{x}_t^\epsilon = u_t^\epsilon(g_t^\epsilon)^{-1}$. Then

$$d\tilde{x}_t^\epsilon = \sqrt{\epsilon}H(\tilde{x}_t^\epsilon)(g_t^\epsilon \circ db_t) + \epsilon H(\tilde{x}_t^\epsilon)(g_t^\epsilon e_0)dt. \tag{2.3}$$

Its 'formal' Stratonovich correction term vanishes. If $\tilde{x}_0^\epsilon = u_0$ then g_0^ϵ is the identity matrix. Let $Z_j = \sigma_k^j A_j^*$ where $\{A_j\}$ is an o.n.b of $\mathfrak{so}(n)$. Then

$$dg_t^\epsilon = \sum_{j,k} \sigma_k^j(u_t^\epsilon)g_t^\epsilon A_j \circ dw_t^k + \sum_j \sigma_0^j(u_t^\epsilon)g_t^\epsilon A_j \ dt.$$

If $e_0 = 0$, the law of $\tilde{x}_{\frac{t}{\epsilon}}^\epsilon$, and hence that of $x_{\frac{t}{\epsilon}}^\epsilon$, is independent of ϵ. This follows from the independence of g_t^ϵ and $\{b_t\}$. Finally $\tilde{x}_t^\epsilon$ is a horizontal Brownian motion with projection x_t a Markov process and a Brownian motion on M. This is the construction of Brownian motions of Eells-Elworthy [7]. The invariance is no longer true for $e_0 \neq 0$. In case where σ_k^j are constant functions, the process g_t^ϵ is independent of ϵ. Note that it is ergodic. As $\epsilon \to 0$, $\tilde{x}_{\frac{t}{\epsilon}}$ converges to a horizontal BM with a drift $\bar{H}$, meaning that it is a Markov process with generator $\frac{1}{2}\Delta^H + \bar{H}$. We expect that $\bar{H}_0(x) = \int H(x)ge_0 dg$ to be the effective drift where dg is the Haar measure on the group.

Remark 2.3. More generally if $\{\Phi_t(u)\}$ is a family of Markov processes on OM with the property that $\Phi_t(ug) \stackrel{law}{=} \Phi_t(u)\psi_t(g)$ for some $\psi_t(g) \in G$ and $\sigma\{\pi(\Phi_r(u))|r \leq s\} = \sigma\{\Phi_r(u) : r \leq s\}$, then $\pi(\Phi_t(u))$ is a Markov process. Denote by $Q_t(u_0, du)$ the law of $\Phi_t(u_0)$ and let $f : M \to \mathbb{R}$ be a Borel measurable function, $x_t = \pi(\Phi_t(u_0))$,

$$\mathbb{E}\{f(x_t)|\sigma\{x_r, r \leq s\}\} = \int (f \circ \pi)(u) Q_{t-s}(\tilde{x}_s, du).$$

It follows that $\int (f \circ \pi)(u) Q_{t-s}(\tilde{x}_s, du) = \int (f \circ \pi)(u\psi_s(g)) Q_{t-s}(\tilde{x}_s g, du) = \int (f \circ \pi)(u) Q_{t-s}(\tilde{x}_s g, du)$. So $\mathbb{E}\{f(x_t)|\sigma\{x_r, r \leq s\}\}$ depends only on $x_s = \pi(\tilde{x}_s)$. When $e_0 = 0$, the flow of (2.3) satisfies the rotational invariance condition and the horizontal lift of x_t is a function of the path $(x_r, r \leq t)$.

2.3.2 Perturbation to horizontal motions

We proceed to consider vertical perturbations to a left invariant SDE. Since a diffusion process is continuous we may assume that M is connected.

A connection on a manifold gives rise through the transport map to the notion of holonomy. Two points u and v of OM are equivalent if they are connected by a C^1 horizontal curve. If they are equivalent we write $u \sim v$. Fix $u_0 \in OM$ we denote by $\Phi(u_0)$ the restricted holonomy group at u_0. It consists of the sub-group of $SO(n)$ such that $u_0 g$ and u_0 are connected by a horizontal loop. It measures how much information is lost by parallel transport along a loop.

Fix u_0, we denote $H = \Phi(u_0)$. Since M is connected, the holonomy groups are conjugate to each other. Consequently the choice of u_0 is not significant.

We carry on defining the holonomy bundle. Let

$$P(u_0) = \{u \sim u_0 : u \in OM\}.$$

Then $P(u_0)$ is a principle bundle with the usual projection with structural group K. It is the holonomy bundle through u_0 of the connection. Holonomy bundles through two different points on the same fibre are related $P(u) = P(u_0)g$ for some g. We may consider OM as disjoint union of sets of the form $P(u)$. Let Π_1 denote the map taking a frame to a representative in the modulus space with respect to the equivalent classes determined by H.

Let H_0 be a fundamental horizontal vector field associated to a vector $e_0 \in \mathbb{R}^n$. Recall that the solution of $\dot{u}_t = H_0(u_t)$ project to a geodesic on M with initial speed $u_0(e_0)$. The following is taken from Theorem 4.6.

Theorem B. Let M be a connected compact Riemannian manifold with a Riemannian connection ∇. Consider

$$\begin{cases} du_t^\epsilon = H(u_t^\epsilon) \circ db_t + H_0(u_t^\epsilon)dt + \sqrt{\epsilon}\sum_{k=1}^m Z_k(u_t^\epsilon) \circ dw_t^k + \epsilon Z_0(u_t^\epsilon)dt, \\ u_0^\epsilon = u_0, \end{cases} \tag{2.4}$$

where Z_k are vertical vector fields. Then as $\epsilon \to 0$, $\Pi_1(u^\epsilon_{\frac{t}{\epsilon}})$ converges in law to a Markov process whose probability distribution is identified in (4.14) below.

In addition to the above two examples, in Theorem 3.10, we construct a Brownian motion on the two sphere by time homogenisation and a left action. In the language of Hopf fibration, the BM on S^2 is the effective motion in the slow /fast system we constructed by a perturbation in the cross fibre direction of a horizontal diffusion. We also include the example of the BM creation by spinning geodesics, its proof is omitted here as the result in [26] essentially cover the original proof in the original version of this article [24].

3 How to construct slow/fast random SDEs?

Our methodology for complex reduction is as follows: we first project the stochastic process u_t^ϵ to another manifold N. Then we construct an appropriate horizontally lift map and lifting the projection to M. The horizontal map is an ansatz for the slow variables. This procedure looses information, but we could in some cases construct also a complementary process keeping the lost information.

In light of this, we shall introduce a more general setting of two manifolds linked by a projection map. In a related context, a horizontal lift map was constructed from diffusions in [10]. This section is based on [8,11], where the focus is on defining horizontal lifts with diffusion operators. In particular these books are not concerned with perturbation of conservation laws nor with slow/fast systems. We also like to refer to [29].

We fix some notations on a principal bundle $p : P \to M$. If G is the structure group of transformations, recall that R_g denote the right translation, a diffeomorphism on N where $g \in G$. The identity transformation corresponds to the identity element of G. It has the associative property and the composition $R_g R_h = R_{gh}$. Let π denote the projection taking a fibre to its base point. Denote by VT_uP the naturally defined vertical sub-bundle, $VT_uP = \ker[T_u\pi]$. If $A \in \mathfrak{g}$, we define $A^*(u) = \frac{d}{dt}|_{t=0} u \exp(tA)$, for any $u \in P$. A connection ∇ on the tangent space of M induces a splitting of the tangent spaces of T_uP:

$$T_uP = HT_uP \oplus VT_uP.$$

The horizontal bundle HTP is a right invariant distribution and the splitting is an Ehresmann connection. This determines a connection 1-form

$$\varpi_u : T_uP \to \mathfrak{g},$$

and a horizontal lifting map $\mathfrak{h}_u : T_{\pi(u)}M \to T_uP$.

A connection 1-form is instrumental to our computations. It assigns any tangent vector on P a Lie algebra element. It is determined by adjoint invariance and its values on fundamental vertical vector fields.

(1) $\varpi(A^*) = A$, for all $A \in \mathfrak{g}$;

(2) $(R_g)^*\varpi = \mathrm{ad}(g^{-1})\varpi$ for all $g \in G$.

We recall that $R_{g_*}(A^*) = (Ad(g^{-1})A)^*$.

Note also that the connection 1-form ϖ on OM is equivalent to the set of Christoffel symbols defining a linear connection on M. Let $E = \{E_1, \dots, E_n\}$ be a local frame; we define the Christoffel symbols relative to E by $\nabla E_j = \sum_{ki} \Gamma_{ij}^k dx_i \otimes E_k$. Let θ^i be the set of dual differential 1-forms on M to $\{E_i\}$: $\theta^i(E_j) = \delta_{ij}$. We define $\omega_k^i = \Gamma_{lk}^i \theta^l$. Then $d\theta^i = -\sum_k \omega_k^i \wedge \theta^k$. Let $\{A_i^j\}$ be a basis of $\mathfrak{g}$. To each moving frame E we associate a 1-form, $\omega = \sum_{i,j} \omega_j^i A_i^j$, on M. If (O, x) is a chart of M and $s : O \to OM$ is a local section of OM, let us denote by ω_s the differential 1-form given above, then $\varpi(s_*v) = \omega_s(v)$. Conditions (1) and (2) are equivalent to the following: if $a : O \to G$ is a smooth function,

$$\varpi((s \cdot a)_*v) = a^{-1}(x)da(v) + a^{-1}(x)\varpi(s_*v)a(x).$$

This corresponds to the differentiation of $s \cdot a$. Tangent vectors or vector fields are called *horizontal* (res. *vertical*) if they take vales in HTP (respectively in VTP). Horizontal tangent vectors are in the kernel of ϖ. Any orthonormal basis $\{A_1, \dots, A_N\}$ of $\mathfrak{g}$ induces a family of basis for the vertical bundle VTP. Then the horizontal component of a vector w is given by $w^h = w - \sum_j \langle \varpi(w), A_j \rangle A_j^*$.

Further background and examples

A *diffusion operator* $\mathcal{A}$ on a smooth manifold is a second order differential operator with positive definite symbol and vanishing zero order term. In local coordinates $\mathcal{B}$ acts on a real valued function as follows:

$$\mathcal{A}g = \frac{1}{2} \sum_{i,j=1}^{n} a_{i,j} \frac{\partial^2 g}{\partial y_i \partial y_j} + \sum_{k=1}^{n} b_k \frac{\partial g}{\partial y_k}$$

where $a_{i,j}$ and b_k are smooth functions with $(a_{i,j})$ positive symmetric. For our analysis the local charts description would not be sufficient. For sufficiently smooth a_{ij} and b_k we can find X_i and X_0 such that

$$\mathcal{A} = \tfrac{1}{2} \sum_{i=1}^{m} X_i^2 + X_0$$

where write $Xf = L_{X_i} f$ for the derivative of f in the direction of X. Choose a sum square of representation determines the correlation of the family of diffusion process with generator $\mathcal{A}$. Such Markov processes are referred as $\mathcal{A}$-diffusion [16]. A sum square of representation gives a representation of the diffusion as solutions of the SDE:

$$dx_t = \sum X_i(x_t) dW_t^i + X_0(x_t) dt.$$

Such a representation is not unique, each representation defines the same probability law of the Markov process, but the correlation between different initial points may vary. If the manifold has a Riemannian metric, we denote by Δ the Laplace-Beltrami operator. The $\frac{1}{2}\Delta$-diffusions are Brownian motions.

Definition 3.1. Let $\mathcal{B}$ be a diffusion operator on N and $\mathcal{A}$ be diffusion operator on M, and $p : N \to M$ a smooth map.

- We say that $\mathcal{B}$ and $\mathcal{A}$ are *intertwined* or $\mathcal{B}$ is over $\mathcal{A}$, if for all real valued C^2 functions f on M,

$$\mathcal{B}(f \circ p) = (\mathcal{A}f) \circ p.$$

- A diffusion operator $\mathcal{B}$ on N is *vertical* if $\mathcal{B}(f \circ p) = 0$ for any C^2 function $f : N \to \mathbb{R}$.

Intertwined diffusion operator pairs are prevalent. For example take $\mathcal{B} = A + \mathcal{B}_2$ on $\mathbb{R}^2$ where $A = \frac{1}{2}a(x)\frac{\partial^2}{\partial x^2}$ and $\mathcal{B}_2 = \frac{1}{2}b(y)\frac{\partial^2}{\partial y^2}$.

For intertwined pairs of diffusions operators, $\mathcal{B}$ and $\mathcal{A}$, we can use a suitable choice of horizontal lift map to separate the intrinsic and extrinsic properties, and eventually eliminate the extrinsic noise. This programme started in [9] and was continued in [8, 11]. To explain the construction from the above references, we begin with the symbol of the operator, which is a linear map from the cotangent bundle to the tangent bundle: $\sigma^{\mathcal{A}} : T^*N \to TN$. For any C^2 functions f, g and any point x,

$$df(\sigma_x^{\mathcal{A}}(dg)) = \frac{1}{2}\left(\mathcal{A}(fg) - f\mathcal{A}g - g(\mathcal{A}f)\right)(x).$$

In local charts, $\sigma^{\mathcal{A}}(df, dg) = \sum_{i,j=1}^n a_{i,j}\partial_i f \partial_j g$.

Definition 3.2. Let Image$(\sigma_x^{\mathcal{A}})$ denote the image of the linear map $\sigma_x^{\mathcal{B}} : T_x^*M \to T_xM$. The operator $\mathcal{A}$ is said to be elliptic if its symbol is strictly positive at each point. It is said to have constant rank if Image$(\sigma_x^{\mathcal{A}})$ has constant dimension.

We are now ready to construct the horizontal lifting maps. To begin, consider a smooth surjective map $p : N \to M$ between smooth manifolds.

Theorem 3.3. *[11, Prop. 2.1.2] Suppose that $\{\sigma_x^{\mathcal{A}}, x \in M\}$ has constant rank and $\mathcal{B}$ an $\mathcal{A}$ are intertwined by $p : N \to M$. There is a unique linear map, called the horizontal lifting map, $\mathfrak{h}_u$: Image$(\sigma_{p(u)}^{\mathcal{A}}) \subset T_{p(u)}M \to$ Image$(\sigma_u^{\mathcal{B}}) \subset T_uN$, such that $T_up \circ \mathfrak{h}_u$ is the identity map. This is defined as below: Let $u \in N$, $v \in T_{p(u)}M$ and α a pre-image of v by $\sigma_{\pi(u)}^{\mathcal{A}}$:*

$$\mathfrak{h}_u(v) = \sigma_u^{\mathcal{B}}(T_up)^*(\alpha),$$

The image of $\mathfrak{h}_u$ induces a smooth distribution, called the horizontal distribution associated to $\mathcal{A}$. If $\mathcal{A}$ elliptic then the image HT_uN of $\mathfrak{h}_u$ is a horizontal complement to the vertical space:

$$T_uN = HT_uN \oplus VT_uN, \qquad u \in N.$$

If $p : N \to M$ admits furthermore a principal bundle structure with group G acting on N, this is an Ehresmann connection. A connection 1-form $\varpi : TN \to \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra of G, determines such a connection.

The theory in [8] allows to extract from the pair $(\mathcal{A}, \mathcal{B})$ a horizontal lift and the horizontal lift of $\mathcal{A}$:

$$\mathcal{A}^H = \tfrac{1}{2}\sum \tilde{X}_i^2 + \tilde{X}_0$$

where $\tilde{X}_i$ is the horizontal lift of the vector field X_i. The procedure of splitting $\mathcal{B}$ into a horizontal and vertical part is easier to describe. First recall that $\mathcal{A}$ is semi-elliptic, if it has a constant but not full rank.

Definition 3.4. If $X_0 \in \operatorname{Image} \sigma^{\mathcal{A}}$ we say $\mathcal{A}$ is cohesive.

We shall assume that the semi-elliptic operator is cohesive. The following construction is the underlying structure allowing us to study perturbation to motions that are horizontal and vertical.

Theorem 3.5. *[11, Theorem 2.2.5] There is a unique vertical diffusion $\mathcal{B}^v$ such that*

$$\mathcal{B} = \mathcal{A}^H + \mathcal{B}^v.$$

Indeed, if y is a regular point of the map p, the $\mathcal{B}^v$ diffusion starting from y stays in the sub-manifold $p^{-1}(p(y))$. These sub-manifolds are conserved quantities. To study perturbed systems, we make ergodic assumption. For example, for horizontal perturbation to vertical motions, we may want to assume that the vertical diffusion operator is 'elliptic', which means that its symbol $\sigma_y^{\mathcal{B}^v} : VT_yN \times VT_yN \to \mathbb{R}$ is strictly positive definite.

At this point we would like to note that decomposition of stochastic flows has been studied in the past, see e.g. [5, 30]. The following examples of diffusion pairs are given in [11].

Example 3.6. Consider the cylinder $\mathbb{R} \times S^1$. Let z denote the S^1 direction, $p(x, z) = z$, $\mathcal{A} = \frac{\partial^2}{\partial z^2}$, and $\mathcal{B} = \sin z \frac{\partial}{\partial x} + \frac{\partial^2}{\partial z^2}$. Set $\mathcal{A}^H = \frac{\partial^2}{\partial z^2}$. We may consider the following perturbed operators: $\mathcal{L}^\epsilon := \sin z \frac{\partial}{\partial x} + \frac{1}{\epsilon}\frac{\partial^2}{\partial z^2}$. The projection of the $\mathcal{L}^\epsilon$ is approximately a constant on the interval $[0, \frac{1}{\epsilon}]$.

Example 3.7. Let $N = \mathbb{R}^3$ with the Heisenberg group structure. This is isomorphic to the matrix group of 3×3 upper diagonal matrices :

$$(x, y, z) \mapsto \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \tag{3.1}$$

The Heisenberg group multiplication on $\mathbb{R}^3$ is defined by

$$(x,y,z)\cdot(x',y',z') = \big(x+x', y+y', z+z'+\frac{1}{2}(xy'-yx')\big).$$

For $(x,y,z)\in\mathbb{R}^3$ define $\pi(x,y,z)=(x,y)$. Let

$$Y_1 = \frac{\partial}{\partial x} - \frac{1}{2}y\frac{\partial}{\partial z}, \qquad Y_2 = \frac{\partial}{\partial y} + \frac{1}{2}x\frac{\partial}{\partial z}$$

be the left invariant vector fields on N. These are obtained respectively by left translations of the vectors $(1,0,0)$ and $(0,1,0)$ at the origin, hence the span of $\{Y_1,Y_2\}$ has constant rank 2 at any point. The vertical vector field is $\frac{\partial}{\partial z}$. Let

$$\mathcal{A}^H = \tfrac{1}{2}(Y_1^2+Y_2^2), \qquad \mathcal{B}^0 = \tfrac{1}{2}\frac{\partial^2}{\partial z^2} - \frac{\partial}{\partial z}.$$

Define horizontal lift :

$$(u,v)\mapsto (u, v, \frac{1}{2}xv - \frac{1}{2}yu).$$

Then the vector fields Y_1, Y_2 are respectively the horizontal lifts of the vector fields $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ on $\mathbb{R}^2$, and $\mathcal{A}^H$ is the horizontal lift of the linear operator $\mathcal{A} = \frac{1}{2}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})$ on $\mathbb{R}^2$. With the horizontal lift map , we may list a smooth curve $\sigma:[0,T]\to M$, with $\sigma(0)=0$, to a horizontal curve $\tilde{\sigma}$. This is given by

$$\tilde{\sigma}(t) = \left(\sigma^1(t), \sigma^2(t), \frac{1}{2}\int_0^t \left(\sigma^1(t)d\sigma^2(t) - \sigma^2(t)d\sigma^1(t)\right)\right). \tag{3.2}$$

Equation (3.2) remains valid for the horizontal lift of Brownian motion on $\mathbb{R}^2$, provided the integrals are interpreted as Stratonovich integral.

Let us define $\mathcal{L}^\epsilon = \frac{1}{2}(Y_1^2+Y_2^2) + \frac{1}{\epsilon}\mathcal{B}^0$, which is over $\mathcal{A}$ according to Definition 3.1. The slow part of the $\mathcal{L}^\epsilon$-diffusion is an $\mathcal{A}^H$ diffusion, it does not depend on ϵ. This example can extend to a more general case. Let us assume that $r_i:\mathbb{R}^3\to\mathbb{R}$ be real valued functions with $(\frac{\partial r_1}{\partial y} - \frac{\partial r_2}{\partial x})^2$ strictly positive. For each point $(x,y,z)\in\mathbb{R}^3$, we define the following map, from the tangent space of $\mathbb{R}^2$ to the tangent space of $\mathbb{R}^3$:

$$\mathfrak{h}_{(x,y,z)}(u,v) = (u, v, r_1(x,y,z)u + r_2(x,y,z)v).$$

See Elworthy-LeJan-Li [11, pp21]. Let $X_1 = \frac{\partial}{\partial x} + r_1\frac{\partial}{\partial z}$ and $X_2 = \frac{\partial}{\partial y} + r_2\frac{\partial}{\partial z}$. Set

$$\mathcal{A}^H := \frac{1}{2}(\frac{\partial}{\partial x} + r_1\frac{\partial}{\partial z})^2 + \tfrac{1}{2}(\frac{\partial}{\partial y} + r_2\frac{\partial}{\partial z})^2.$$

The first prolongation of span$\{X_1, X_2\}$, *i.e.* span$\{X_1, X_2, [X_1, X_2]\}$ has full rank at each point. Consider a function α such that the invariant measure μ of the Markov process with generator $(\gamma - r_1^2 - r_2^2)\frac{\partial^2}{\partial z^2} + \alpha \frac{\partial}{\partial z}$ is the standard Gaussian measure. Let u_t^ϵ denotes a $\mathcal{L}^\epsilon$ Markov process where

$$\mathcal{L}^\epsilon := \mathcal{A}^H + \frac{1}{\epsilon^2}(\gamma - r_1^2 - r_2^2)\frac{\partial^2}{\partial z} + \frac{1}{\epsilon}\alpha\frac{\partial}{\partial z}.$$

For each ϵ, $(\mathcal{A}, \mathcal{L}^\epsilon)$ determines a connection. Denote by $\tilde{u}_t^\epsilon$ the horizontal lift of $x_t^\epsilon := \pi(u_\cdot^\epsilon)$. Write $u_t^\epsilon = x_t^\epsilon g_t^\epsilon$. We say a few words about an auxiliary vertical process with parameter $u \in OM$. Fix $u \in OM$, the vertical motion with frozen u is a Markov process with infinitesimal generator $\frac{1}{2}(\frac{\partial}{\partial x} + r_1(ug)\frac{\partial}{\partial z})^2 + \frac{1}{2}(\frac{\partial}{\partial y} + r_2(ug)\frac{\partial}{\partial z})^2$. Let γ be such that $\gamma - r_1^2 - r_2^2 > c$ for some positive number c. If r_i are not constants a.s. with respect to μ, the slow part of diffusion converges to an elliptic diffusion on the Heisenberg group. Its invariant measures shall be denoted by μ^u. These measures are use for averaging out the diffusion generator of the slow variables.

Example 3.8. A non-relativistic quantum mechanical diffusion lives naturally in $\mathbb{R}^3 \times SO(3)$, the orthonormal frame bundle of $\mathbb{R}^3$. Its spatial projection lives in $\mathbb{R}^3$. Studies associated to quantum mechanical equations, mainly the continuity equation describing the probability density of the quantum equation, have intertwined structures on $p : \mathbb{R}^3 \times SO(3) \to \mathbb{R}^3$. I am grateful to D. Elworthy to bring my attention to the paper of Wallstrom [41] where limits of stochastic processes in $\mathbb{R}^3 \times SO(3)$ are discussed. The Bopp-Haag-equations have one free parameter I and its solutions converge to that of an equation with Pauli Hamiltonian as $I \to 0$. The Bopp-Haag-Dankel stochastic mechanical diffusions $\mathbb{R}^3 \times SO(3)$ were introduced by Dankel, describing a diffusion particle with definite position and orientation. The Bopp-Haag -Dankel diffusions on $\mathbb{R}^3 \times SO(3)$ are given by a simple SDE with drift given by a Pauli spinor (solution of quantum equation associated with Pauli Hamiltonian with parameter I). In [41] it was shown that for spin $\frac{1}{2}$ wave functions and regular potentials the process parametrized by I converge to a Markovian process onto $\mathbb{R}^3$, due to the averaging out of the orientational motion. The spatial projection describes the spatial motion of the particle without its orientation.

Example 3.9. Let $G = SO(n)$ and $\pi : \mathbb{R}^n \times G \to \mathbb{R}^n$ the projection to its first component. Let $\mathfrak{g} = \mathfrak{so}(n)$ be the Lie algebra of the Lie group G. For each $x \in \mathbb{R}^n$, let $h_x : T_x\mathbb{R}^n \sim \mathbb{R}^n \to \mathfrak{g}$ be a linear map varying smoothly in x. The map $(x,v)\mapsto(x,h_x(v))$ can be considered as the horizontal lifting map through (x,I) where I is the identity matrix. This induces on $\mathbb{R}^n$ a non-trivial covariant differentiation ∇. Let $e\in\mathbb{R}^n$, consider the SDE

$$\begin{aligned}
dx_t &= \epsilon_1 g_t\circ db_t+\epsilon g_t e\,dt\\
dg_t &= \epsilon_1 h_{x_t^\epsilon}(g_t\circ db_t)g_t+\epsilon h_{x_t}(g_t e)g_t dt+\sqrt{\delta}\sum_{k=1}^p Z_k(x_t,g_t)\circ dw_t^k\\
&\quad +\delta Z_0(x_t,g_t)dt.
\end{aligned}$$

where $b_t=(b_t^1,\dots,b_t^n)$ and $w_t=(w_t^1,\dots,w_t^p)$ for (b_t^i,w_t^k) independent 1-dimensional Brownian motions. Also, $Z_k:\mathbb{R}^n\times G\to TG$ with $Z_k(x,g)\in T_gG$. When $h=0$ this corresponds to the trivial lifting. We consider three types of scalings: 1) $\delta=1$, $\epsilon_1=\sqrt{\epsilon}$ and $\epsilon\to0$; 2) $\epsilon_1=\epsilon=1$ and $\delta\to0$. For the third type take $\epsilon_1=0$, $\epsilon=1$ and $\delta\to\infty$. In case 1) it turns out that the solution x_t is a slow variable, despite the involvement of g_t.

3.1 Hamiltonian Dynamics on the cotangent bundle.

We shall represent the equation on the orthonormal frame bundle and consider randomly perturbation to the second order geodesic equation which reduces to a slow /fast motion in which the fast motion satisfies strong Hörmander's conditions.

A geodesic $(x(t),t\in[0,1])$ is a solution to the second order ODE on M: $\frac{d}{dt}\dot{x}^k(t)=-\Gamma_{ij}^k(x(t))\dot{x}^i(t)\dot{x}^j(t)$, where the functions Γ_{ij}^k are the Christoffel symbols. A geodesic is constant speed length minimising curve. It is also the motion of a free particle minimizing the energy function $E(x_\cdot)=\frac{1}{2}\int_0^1|\dot{x}(t)|^2dt$. So a geodesic is intuitively the motion of a free particle that minimises the energy function. In fact they are critical points of E.

If we identify the tangent and the cotangent bundle, the geodesic flow is the Hamiltonian flow for $H(x,y)=\frac{1}{2}|y|_x^2$, $(x,y)\in T^*M$. To see this let (U,x) be a local coordinate for M, where U is an open set of M and by abuse of notation $x:U\to\mathbb{R}^n$ is a diffeomorphism to its image and $x=(x^1,\dots,x^n)$. Then (x,y) is the induced coordinate map for T^*M, and (x,y) represents the cotangent vector $\sum_i y_i dx_i$. Let (g^{ij}) denote the inverse matrix to the Riemannian metric (g_{ij}) then $H(x,y)=\frac{1}{2}\sum_{i,j}g^{ij}(x)y_iy_j$. Let $\omega=\sum_i dx_i\wedge dy_i$ be the non-degenerate symplectic 2-form on T^*M and we define the Hamiltonian vector field X_H on

T^*M by $\iota_{X_H}\omega = dH$ where ι denotes the interior product. The solution flow to

$$\dot{x}(t) = X_H(x(t))$$

are geodesics. To see this more clearly we will write the equation in local coordinates. Firstly

$$dH = \frac{1}{2}\sum_{i,j,k} \frac{\partial g^{i,j}}{\partial x_k} y_i y_j dx_k + \sum_{i,j} g^{ij} y_i dy_j.$$

If $X = \sum_k f_k \frac{\partial}{\partial x_k} + \sum_k h_k \frac{\partial}{\partial y_k}$, then $\iota_X\omega = \sum_k f_k dy_k - \sum_k h_k dx_k$. This means that X has the expression

$$X = \sum_k \sum_i g^{ik} y_i \frac{\partial}{\partial x_k} - \sum_k \frac{1}{2}\sum_{i,j} \frac{\partial g^{i,j}}{\partial x_k} y_i y_j \frac{\partial}{\partial y_k}.$$

Let (x_t, y_t) denote the integral curve of X, then

$$\dot{x}^k = \sum_j g^{kj} y_j, \qquad \dot{y}_k = -\frac{1}{2}\sum_{i,j} \frac{\partial g^{ij}}{\partial x_k} y_i y_j.$$

We differentiate $\dot{x}^k$ once more, transform y_k to $\dot{x}^k$'s by lowering the indexes, apply the formula for Christoffel symbols in terms of $(g_{i,j})$, and we see that this is indeed the geodesic equation given earlier.

To pass to the frame bundle, we recall the notion of associated bundles and the tangent bundle is a associated fibre bundle with fibre $\mathbb{R}^n$ to the frame bundle P. Fix $e_0 \in \mathbb{R}^n$, let $H = \{g \in G : ge_0 = e_0]$ denote the isotropy group at e_0. The total space is $P \times \mathbb{R}^n / \sim$ where the equivalent class is determined by $[u, e] \sim [ug^{-1}, ge]$, any $g \in G$. Elements of the form ug where $g \in H$ belong to the same equivalence class. Let P/H denote the quotient bundle which contains precisely the equivalence class of the form $\{ug, g \in H\}$. Then the tangent bundle is identified with the quotient bundle P/H. Denote by ξ_0 the coset H. Let α be the map:

$$\alpha_{e_0} : u \in P \to ue_0 \in TM.$$

Fix a unit vector $e_0 \in \mathbb{R}^n$ and consider $G = SO(n)$. Each element $v \in TM$ has a representation $[u, e_0]$ in $OM \times \mathbb{R}^n$ and it is unique up to right translation by elements of isotropy group at e_0. We may identify $OM \times \mathbb{R}^n / \sim$ with the quotient bundle OM/H, whose elements are the equivalence class of elements of the form $ug, g \in H$. Let α be the associated map:

$$\alpha_{e_0} : u \in OM \to ue_0 \equiv [u, e_0] \in TM.$$

The differential $D\alpha_{e_0}$ at u induces a map from T_uOM to $T_{ue_0}TM$. Any vector field W on OM, that is invariant by right translations of elements of G_{e_0}, induces a vector field on TM. If $v = ue_0' = ue_0$ there is $g \in G$ with $e_0' = g^{-1}e_0$. Set $u' = ug$. Since $\alpha_{e_0}(u) = \alpha_{e_0'}(R_g u)$,

$$D_u\alpha_{e_0}(W(u)) = D_{u'}\alpha_{e_0'}DR_g(W(u)) = D_{u'}\alpha_{e_0'}\big(W(u')\big),$$

the map $W \in \Gamma TP \mapsto D_u\alpha_{e_0}(W(u)) \in \Gamma TTM$, is independent of the choice of e_0. If $W(u) = H_u(e_0)$, the induced vector field X on the tangent bundle TM is a geodesic spray, i.e. in local co-ordinates $X(x,v) = (x, v, v, Z(x,v))$ and $Z(x, sv) = s^2 Z(s,v)$. This corresponds to the geodesic equation on TM: $dv_t^k = -\Gamma_{ij}^k(\sigma_t)v_t^i v_t^j$, $\dot\sigma_t = v_t, \sigma(0) = \pi(u_0), v(0) = u_0e_0$. A vector field on OM that is horizontal and invariant under translation by the action of G projects to a vector field on the base manifold.

Let $(u_t^{e_0})$ be the solution to

$$\dot u(t) = H_{u_t}(e_0), \quad u(0) = u_0,$$

then $\pi(u_t^{e_0})$ is the geodesic with initial velocity $u_0(e_0)$ and initial point $\pi(u_0)$. In other words a geodesic is the projection of a horizontal flow from the bundle of orthonormal frames of M to M. It is worth remarking that horizontal vector field $H(e_0)$ does not project to a vector field on M. This fact explains how random perturbation in the rotation group be transmitted to the geodesics.

We explain the horizontal lift in this setting. Let $c(t)$ be a curve and $\tilde c(t)$ a horizontal lift of $c(t)$. The principal part of $\tilde c(t)$ is a $n \times n$ matrix whose column vectors $\{\tilde c_1(t), \dots, \tilde c_n(t)\}$ form a frame. In components, write $\tilde c_l(t) = (\tilde c_l^1(t), \dots, \tilde c_l^n(t))^T$. Then

$$\frac{\partial \tilde c_l^k(t)}{\partial t} + \sum_{i=1,j=1}^n \frac{\partial c^i(t)}{\partial t}\Gamma_{ij}^k(c(t))\tilde c_l^j(t) = 0.$$

Take $c(t) = (0, \dots, t, \dots, 0)$, where the non-zero entry is in the ith-place. We obtain the principal part of the horizontal lift of $\frac{\partial}{\partial x_i}$ through $u = \tilde c(0) = (u_l^j)$:

$$\left(\mathfrak{h}_{\tilde c(0)}(\frac{\partial}{\partial x_i})\right)_l = \left(\frac{\partial \tilde c}{\partial t}(0)\right)_l = -\left(\sum_j \Gamma_{ij}^1\, u_l^j, \dots, \sum_{j=1}^n \Gamma_{ij}^n u_l^j\right)^T.$$

Denote by A_i the matrix whose element at the (b,l) position is $\sum_j \Gamma_{ij}^b u_l^j$. We conclude that A_i is the principal part of $H_u(\frac{\partial}{\partial x_i})$ and the horizontal space at u is spanned by the basis $\{(\frac{\partial}{\partial x_i}, A_i)\}$. This example will be continued in §4.1.

3.2 Horizontal diffusions on the hyperbolic plane

Below we compute the horizontal diffusions in case the base manifold is the hyperbolic plane and R^2.

On the hyperbolic plane

Let $H^2 = \{(x_1, x_2) : x_2 > 0\}$ be the hyperbolic plane and set $g_{ij} = \frac{1}{(x_2)^2}\delta_{ij}$. Its non-zero Christoffel symbols are:

$$\Gamma^1_{12} = \Gamma^1_{21} = -\frac{1}{x_2}, \quad \Gamma^2_{22} = -\frac{1}{x_2}, \quad \Gamma^2_{11} = \frac{1}{x_2}.$$

The total space of the orthonormal frames is a product space. We let u^0 denote the principal part of a frame u. Let $u^0 = (u_1, u_2) = \begin{pmatrix} u_1^1 & u_2^1 \\ u_1^2 & u_2^2 \end{pmatrix}$, it belongs to $SO(2)$. Let $x = \pi(u) = (x_1, x_2)$. Let $\{e_1, e_2\}$ be the standard basis of $\mathbb{R}^2$. Then

$$X_1(u^0) := \mathfrak{h}_u(e_1) = -\begin{pmatrix} \sum_j \Gamma^1_{1j} u_1^j & \sum_j \Gamma^1_{1j} u_2^j \\ \sum_j \Gamma^2_{1j} u_1^j & \sum_j \Gamma^2_{1j} u_2^j \end{pmatrix} = -\frac{1}{x_2}\begin{pmatrix} -u_1^2 & -u_2^2 \\ u_1^1 & u_2^1 \end{pmatrix}$$

$$X_2(u^0) := \mathfrak{h}_u(e_2) = -\begin{pmatrix} \sum_j \Gamma^1_{2j} u_1^j & \sum_j \Gamma^1_{2j} u_2^j \\ \sum_j \Gamma^2_{2j} u_1^j & \sum_j \Gamma^2_{2j} u_2^j \end{pmatrix} = \frac{1}{x_2}\begin{pmatrix} u_1^1 & u_2^1 \\ u_1^2 & u_2^2 \end{pmatrix}.$$

In equation (4.2) we take $e_0 = (1,0)^\perp$, $\bar{A} = 0$, $u_t = (x(t), u^0(t))$, and

$$A = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}.$$

Denote the components of $(u^0(t))$ by $(u_l^j(t))$, where the subscript denote the index for the column. Observe that ue_0 is the first column of u. Also

$$H_u(e_1) = \mathfrak{h}_u(ue_1) = u_1^1 \mathfrak{h}_u(e_1) + u_1^2 \mathfrak{h}_u(e_2) = u_1^1 X_1(u) + u_1^2 X_2(u).$$

Then equation (4.2) takes the following form:

$$\dot{x}(t) = T\pi(H_{u_t}(e_0)) = u^0(t)(e_0) = \begin{pmatrix} u_1^1(t) \\ u_1^2(t) \end{pmatrix},$$

$$du^0(t) = \big(u_1^1(t) X_1(u^0(t)) + u_1^2(t) X_2(u^0(t))\big)\, dt + \frac{1}{\sqrt{\epsilon}}\, u^0(t)\, A \circ dw_t.$$

In the equations we suppressed the superscript ϵ and the variable ω. In particular, $x_2(t) = \int_0^t u_1^2(s) ds$. Furthermore, the principal part of horizontal lift of $x(t)$ solves the following equation:

$$\dot{\tilde{x}}(t) = \mathfrak{h}_{\tilde{x}(t)}(\dot{x}(t)) = \mathfrak{h}_{\tilde{x}(t)}(u^0(t)(e_0)) = u_1^1(t)X_1(\tilde{x}(t)) + u_1^2(t)X_2(\tilde{x}(t)).$$

We see later that there is a rotation matrix $g_t(\omega)$ such that $u^0(t,\omega) = \tilde{x}_t(\omega)g_t(\omega)$. Again we have suppressed the superscript ϵ.

Horizontal diffusions on $\mathbb{R}^n$

Let us take $M = \mathbb{R}^d$ with the trivial Riemannian metric. Then $OM = \mathbb{R}^n \times O(n)$. The horizontal vectors in the tangent space of OM are those with vanishing Lie-algebra components. We write a frame u as (x, g). The horizontal lift at $u = (x, g)$ of a vector $v \in T_x\mathbb{R}^n$ is $((x, g), (v, 0)) \in (\mathbb{R}^n \times O(n)) \times (T_x\mathbb{R}^n \times \mathfrak{so}(n))$.

To ease the notation we omit the trivial component of the horizontal lift, we have $H_u(e) = (ge, 0)$ and the equation $\dot{u}_t = H_{u_t}(e_0)$ is equivalent to $\dot{x}_t = g_t e_0, \dot{g}_t = 0$, $g_0 e_0 = v_0$. Let $A_k \in \mathfrak{so}(n)$. We claim that $(x^\epsilon_{\frac{t}{\epsilon}})$ converges as $\epsilon \to 0$. At first glance this equation appears to have the wrong scaling. Let us set $y_t^\epsilon = \frac{1}{\epsilon} x^\epsilon_{\epsilon t}$ and $\tilde{g}_t^\epsilon = g^\epsilon_{\epsilon t}$. The above equation is equivalent to

$$\left\{ \begin{array}{lcl} \dot{y}_t^\epsilon & = & \tilde{g}_t^\epsilon \, e_0, \\ d\tilde{g}_t^\epsilon & = & \sum_k \tilde{g}_t^\epsilon \, A_k \circ d\tilde{w}_t^k + \epsilon \, \tilde{g}_t^\epsilon \, \bar{A} \, dt \end{array} \right.$$

with $y_0^\epsilon = \frac{1}{\epsilon} x_0$ and $\tilde{g}_0^\epsilon = I$. Here $\{\tilde{w}_t^k\}$ is a family of independent Brownian motions.

For simplicity, let us assume that $\bar{A} = 0$, then $\{\tilde{g}_t^\epsilon\}$ is a reversible ergodic Markov process on $SO(n)$ with the Haar measure as the invariant measure. We can apply central limit theorems for additive functionals of Markov processes. Let $e \in \mathbb{R}^n$ we set $V^e(g) = \langle g e_0, e\rangle$ and set $Y^e(t) = \int_0^t V^e(\tilde{g}_s^\epsilon) ds$. It is easy to check that for each e, V^e satisfies the conditions in Theorem 1.8 and Corollary 1.9 in Kipnis-Varadhan [20], in particular $\int_{O(n)} V^e(g) dg = 0$. Hence $\epsilon \, Y^e(\frac{t}{\epsilon^2})$ converges weakly, and so does $x^\epsilon_{\frac{t}{\epsilon}} = \epsilon \, y^\epsilon_{\frac{t}{\epsilon^2}}$. See also Helland [15]. For more general manifolds, the integrals $\int_0^t V^e(\tilde{g}_s^\epsilon) ds$ does not make sense as it stands, we 'transform' them into path integrals of differential 1-forms along semi-martingales.

3.3 An example of diffusion creation on S^2

Let S^n denote the n-sphere. We present a stochastic perturbation model on S^3 which we visualise as S^2 with a non-trivial fibration, on each point of S^1 we attach a circle, these circles are entangled. Hopf fibration is the principal bundle $\pi : S^3 \to S^2$ with S^1 acting on the right. It is convenient to consider the representation by unitary groups in which S^3 is identified with $SU(2)$, S^1 with $U(1)$, and S^2 with $SU(2)/U(1)$. We introduce a family of SDEs on S^3 with a small parameter ϵ and construct a Brownian Motion on S^2 by stochastic homogenisation.

Hopf fibration occurs in multiple situation in physics: in quantum systems and in mechanics, c.f. Urbantke [40]. Our interest in the Hopf fibration comes from its history as a rich background for constructing counter example. For example, the first non-trivial example of collapsing manifolds without blowing up the curvature was given on the Hopf fibration by M. Berger in 1962. This was achieved by shrinking the length of S^3 by a scale of ϵ along the Hopf fibration direction and leaving the orthogonal directions unchanged. Our model is related to this scaling. In a paper in preparation we study the dynamics associated to collapsing manifold [23].

An element of $SU(2)$ is of the form $\begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix}$, an element $e^{i\theta} \in U(1)$ shall be represented as the matrix $\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$. The Lie algebra $\mathfrak{su}(2)$ is the set of matrices such that $A + \bar{A}^T = 0$ and with zero trace:

$$\begin{pmatrix} ia & \beta \\ -\bar{\beta} & -ia \end{pmatrix}, \qquad a \in \mathbb{R}, \beta \in \mathbb{C}.$$

We shall denote by Y^* the left invariant vector fields associated to a vector $Y \in \mathfrak{su}(2)$, however the superscript will be often omitted for simplicity so Y denotes both an element of the tangent vector at the identity or the left invariant vector field it generated.

Consider the Pauli basis $\{X_i\}$ where

$$X_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad X_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Under the standard following inner product on $\mathfrak{su}(2)$, $\langle A, B\rangle := \frac{1}{2}\operatorname{trace} AB^*$, $\{X_1, X_2, X_3\}$ is an orthonormal basis. Note that X_1 is adjoint invariant under the circle action and so is the linear span of $\{X_2, X_3\}$. In fact $e^{i\theta}X_i = X_i e^{-i\theta}$. Let us define a distribution $D = \operatorname{span}\{X_2^*, X_3^*\}$, which is obviously left invariant with respect to the group action on S^3. The span of the left invariant vector fields is also right invariant under the circle action. This is due to the following fact: let $u \in SU(2)$, $\theta \in \mathbb{R}$, $ue^{i\theta}X_i \in D_{ue^{i\theta}}$ where $i = 2, 3$, we have $ue^{i\theta}X_i = u(e^{i\theta}X_i e^{-i\theta})e^{i\theta} \in TR_{e^{i\theta}}D_u$. Then $T_uS^3 = [\ker T_u\pi] \oplus D_u$ defines an Ehresmann connection on the principal bundle and a horizontal lifting map.

Let ∇^L be the left invariant linear connection and ∇ the Levi-Civita connection for the bi-invariant Riemannian metric on the Lie group $SU(2)$. Denote by Δ the Laplacian on S^3. Let $\Delta_H = \sum_{i=2}^3 \nabla^L df(X_i, X_i) = \sum_{i=2}^3 L_{X_i}L_{X_i}$ where L_{X_i} denotes Lie derivative in the direction of X_i, be the Horizontal Laplacian corresponding to the Horizontal distribution.

Theorem 3.10. *Let $Y_0 \in \mathrm{span}\{X_2, X_3\}$ and let (b_t) be a 1-dimensional Brownian motion. Let u_t^ϵ be the solution to the following SDE on $SU(2) \times U(1)$,*

$$du_t^\epsilon = u_t^\epsilon Y_0 dt + \frac{1}{\sqrt{\epsilon}} u_t^\epsilon X_1 \circ db_t, \qquad u_0^\epsilon = u_0.$$

where $u_0 \in SU(2)$. Let $x_t^\epsilon = \pi(u_t^\epsilon)$ and $\tilde{x}_t^\epsilon$ its horizontal lift. Then $\tilde{x}_{\frac{t}{\epsilon}}^\epsilon$ converges in probability to the hypo-elliptic diffusion with generator $\bar{\mathcal{L}}F = \frac{1}{2}|Y_0|^2\Delta_H$. If Y_0 is a unit vector, $x_{\frac{t}{\epsilon}}^\epsilon$ converges in law to the Brownian motion on S^2.

Note that $\{X_1, X_2, X_3\}$ is a Milnor frame [33] with structural constants $(-2, -2, -2)$,

$$[X_1, X_2] = -2X_3, \quad [X_2, X_3] = -2X_1, \quad [X_3, X_1] = -2X_2.$$

If $Y_0 = c_2X_2 + c_3X_3 \neq 0$, $\mathrm{span}\{X_1, Y_0, [Y_0, X_1]\} = \mathrm{span}\{X_1, X_2, X_3\}$. By the structural equations $[Y_0, X_1] = 2c_2X_3 - 2c_3X_2$ and $\{X_1, Y_0, [Y_0, X_1]\}$ is linearly independent. This is easily seen from the non-degeneracy of the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_2 & -2c_3 \\ 0 & c_3 & 2c_2 \end{pmatrix}.$$

It follows that the equation $du_t^\epsilon = u_t^\epsilon Y_0 dt + \frac{1}{\sqrt{\epsilon}} u_t^\epsilon X_1 \circ db_t$ satisfies Hörmander's condition and is hypo-elliptic.

For the theorem, the important observation is that if $Y_0 \in \mathrm{span}\{X_2, X_3\}$ so is the multiplication of Y_0 by an element of S^1. Hence the left invariant vector fields $X^*(gY_0)$ makes sense. This lead us to the following problem, which can be considered as perturbation of geodesic flows by rotating the direction of the geodesic flow very fast.

Let us examine the principal bundle in detail. A typical element of $SU(2)$ may be expressed as (z, w), where $z, w \in \mathbb{C}$ are such that $|z|^2 + |w|^2 = 1$, or as a matrix $\begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix}$. The right action by $e^{i\theta} \in U(1)$ is $(z, w) \mapsto (e^{i\theta} z, e^{i\theta} w)$, which can be considered as right multiplication in the group $SU(2)$ by elements of the form $e^{i\theta} \sim (e^{i\theta}, 0)$:

$$\begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix} \mapsto \begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

We introduce the Hopf map $\pi : SU(2) \to S^2$. It is a submersion given by the following formula:

$$\pi(z, w) = (Re(2z\bar{w}), Im(2z\bar{w}), |z|^2 - |w|^2).$$

The map $T_u\pi$ can be better visualised if S^3 is treated as a subset of $\mathbb{R}^4$, writing $z = y_1 + iy_2, w = y_3 + iy_4$, and $y = (y_1, y_2, y_3, y_4) \in R^4$, then

$$T_y\pi = 2 \begin{pmatrix} y_3 & y_4, & y_1 & y_2 \\ -y_4 & y_3 & y_2, & -y_1 \\ y_1, & y_2, & -y_3, & -y_4 \end{pmatrix}.$$

The vertical tangent spaces are the kernels of $T\pi$. It is easy to check that the vector field $V(y_1, y_2, y_3, y_4) = -y_2\partial_1 + y_1\partial_2 - y_4\partial_3 + y_3\partial_4$ is vertical, meaning $T\pi(V) = 0$. Back to the principal bundle picture, $V((z, w)) := (iz, iw)$ is the fundamental vertical vector field, associated to the element i in the Lie algebra of $U(1)$.

The Hopf map π projects a curve u_t in $SU(2)$ to one in S^2. A curve x_t in S^2 lifts to a horizontal curve $\tilde{x}_t$ in $SU(2)$ through the horizontal lifting map induced by the Ehresmann connection. Below we present the proof for Theorem 3.10.

Proof. Let x_t^ϵ denote the projection of u_t^ϵ. Then there exists a stochastic process $a_t^\epsilon \in S^1$ be such that $u_t^\epsilon = \tilde{x}_t^\epsilon a_t^\epsilon$ where $\tilde{x}_t^\epsilon$ is the horizontal lift of x_t^ϵ through u_0, using the connection determined by $\{X_2^*, X_3^*\}$. Then $a_0^\epsilon = 1$ and

$$d\tilde{x}_t^\epsilon = TR_{(a_t^\epsilon)^{-1}} \circ du_t^\epsilon + (a_t^\epsilon \circ d(a_t^\epsilon)^{-1})^*(\tilde{x}_t^\epsilon).$$

In the equation above, a_t^ϵ represents the action of the differential of the left multiplication $L_{a_t^\epsilon}$. The term $d(a_t^\epsilon)^{-1}$ denotes the stochastic differential, and the symbol $\circ$ signifies Stratonovich integration. Thus

$$d\tilde{x}_t^\epsilon = TR_{(a_t^\epsilon)^{-1}} \left(u_t^\epsilon Y_0 dt + \frac{1}{\sqrt{\epsilon}} u_t^\epsilon X_1 \circ db_t \right) + ((a_t^\epsilon) \circ d(a_t^\epsilon)^{-1})^*(\tilde{x}_t^\epsilon). \tag{3.3}$$

Since $\tilde{x}_t^\epsilon$ is horizontal, the connections form : $\varpi(d\tilde{x}_t^\epsilon) = 0$. Since the left invariant vector fields, $u_t^\epsilon Y_0$ or $v_t^\epsilon Y_0$, are horizontal, we obtain

$$(a_t^\epsilon) \circ d(a_t^\epsilon)^{-1} = -\varpi_{\tilde{x}_t^\epsilon}\left(\frac{1}{\sqrt{\epsilon}} u_t^\epsilon X_1 (a_t^\epsilon)^{-1} \circ db_t\right) = -\frac{1}{\sqrt{\epsilon}} a_t^\epsilon X_1 (a_t^\epsilon)^{-1} \circ db_t.$$

We have used the fact that on the fundamental vertical vector field A^*, $\varpi_u(A^*(u)) = A$, where A is any vertical vector in the Lie algebra $\mathfrak{su}(2)$. It follows that

$$d(a_t^\epsilon)^{-1} = -\frac{1}{\sqrt{\epsilon}} X_1 (a_t^\epsilon)^{-1} \circ db_t.$$

Consequently,

$$\begin{aligned} d\tilde{x}_t^\epsilon &= u_t^\epsilon Y_0 (a_t^\epsilon)^{-1} dt + \frac{1}{\sqrt{\epsilon}} u_t^\epsilon X_1 (a_t^\epsilon)^{-1} \circ db_t - \frac{1}{\sqrt{\epsilon}} \tilde{x}_t^\epsilon a_t^\epsilon X_1 (a_t^\epsilon)^{-1} \circ db_t \\ &= \tilde{x}_t^\epsilon a_t^\epsilon Y_0 (a_t^\epsilon)^{-1} dt. \end{aligned}$$

Since there is no Stratonovich correction term for $da_t^\epsilon = \frac{1}{\sqrt{\epsilon}} a_t^\epsilon X_1 \circ db_t$, the corresponding infinitesimal generator is $\frac{1}{2\epsilon}\Delta_{S^1}$ where Δ_{S^1} is the Laplacian on S^1. The Haar measure dg is the unique invariant measure for it. Then the stochastic processes $\tilde{x}_t^\epsilon$ converges to u_0. This is because the limiting diffusions would be the solution flow to the vector field $\int_{S^1} g Y_0 g dg$ which vanishes. We shall show that over the time scale $[0, \frac{1}{\epsilon}]$, x_t^ϵ converges. We begin with proving that $\{\tilde{x}_{\frac{t}{\epsilon}}^\epsilon, \epsilon \in (0,1]\}$ is tight. Its probability distribution is the same as that of the solution of

$$\frac{d}{dt} y_t^\epsilon = \frac{1}{\epsilon} y_t g_{\frac{t}{\epsilon}} Y_0 (g_{\frac{t}{\epsilon}})^{-1}.$$

where

$$dg_t = \frac{1}{\epsilon} g_t X_1 \circ db_t, \qquad g_0 = \mathrm{Id}\,.$$

Let $F : S^3 \to \mathbb{R}$ be any smooth function. Since $Y_0 \in \operatorname{span}\{X_2, X_3\}$,

$$\begin{aligned} F(y_t^\epsilon) &= F(u_0) + \frac{1}{\epsilon}\int_0^t dF(y_s^\epsilon g_s^\epsilon Y_0 (g_s^\epsilon)^{-1}) ds \\ &= F(u_0) + \frac{1}{\epsilon}\sum_{j=2}^3 \int_0^t dF(y_s^\epsilon X_j) \langle y_s^\epsilon X_j, y_s^\epsilon g_s^\epsilon Y_0 (g_s^\epsilon)^{-1}\rangle. \end{aligned}$$

Here $dF : TS^3 \to \mathbb{R}$ denotes the differential of F as a linear map. We omit the base variable: for $v \in T_x S^3$, we use $dF(v)$ for its devaluation at v. We have used the left invariant of the Riemannian metric on S^2. By left invariance of the Riemannian metric, we have

$$F(y_t^\epsilon) - F(u_0) = \frac{1}{\epsilon}\sum_{j=2}^{3}\int_0^t dF(y_s^\epsilon X_j)\langle X_j, g_s^\epsilon Y_0(g_s^\epsilon)^{-1}\rangle ds. \tag{3.4}$$

It is easy to see that

$$\begin{aligned} &dF(y_s^\epsilon X_j)\\ &\quad= dF(u_0X_j) + \frac{1}{\epsilon}\int_0^s \nabla^L dF(y_r^\epsilon g_r^\epsilon Y_0(g_s^\epsilon)^{-1},\ y_r^\epsilon X_j)dr\\ &\quad= dF(u_0X_j) + \frac{1}{\epsilon}\sum_{k=2}^{3}\int_0^s \nabla^L dF(y_r^\epsilon X_k,\ y_r^\epsilon X_j)\langle y_r^\epsilon g_r^\epsilon Y_0(g_s^\epsilon)^{-1}, y_r^\epsilon X_k\rangle\ dr\,. \end{aligned}$$

In the above formula, we use the left invariant connection ∇^L so that the covariant derivative of the left invariant vector fields vanish. Then for $j = 2, 3$,

$$\begin{aligned} &dF(y_t^\epsilon X_j)\langle X_j, g_t^\epsilon Y_0(g_t^\epsilon)^{-1}\rangle - dF(u_0X_j)\langle X_j, Y_0\rangle\\ =&\frac{1}{\epsilon}\sum_{k=2}^{3}\int_0^t \nabla^L dF(y_s^\epsilon X_k, y_s^\epsilon X_j)\langle X_k, g_s^\epsilon Y_0(g_s^\epsilon)^{-1}\rangle\ \langle X_j, g_s^\epsilon Y_0(g_s^\epsilon)^{-1}\rangle\ ds\\ &+\int_0^t dF(y_s^\epsilon X_j)d\langle X_j, g_s^\epsilon Y_0(g_s^\epsilon)^{-1}\rangle. \end{aligned} \tag{3.5}$$

By Itô's formula, it is easy to see that

$$d\langle X_j, g_t^\epsilon Y_0(g_t^\epsilon)^{-1}\rangle = \frac{1}{\epsilon}\langle X_j, g_t^\epsilon[X_1, Y_0](g_t^\epsilon)^{-1}\rangle \circ db_t - \frac{2}{\epsilon^2}\langle X_j, g_t^\epsilon[X_1, Y_0](g_t^\epsilon)^{-1}\rangle dt. \tag{3.6}$$

We have used the fact that $\frac{d}{dt}g_tY_0g_t = [X_1, Y_0]$ if $\dot g_t = g_tX_1$, the transfer principle, the fact that the inner product is linear. Observe also that $X_1^2 = -\mathrm{Id}$, $X_1Y_0X_1 = -Y_0$ and so $[X_1, [X_1, Y_0]] = -4Y_0$. It follows that

$$\begin{aligned} &dF(y_t^\epsilon X_j)\langle X_j, g_t^\epsilon Y_0(g_t^\epsilon)^{-1}\rangle - dF(u_0X_j)\langle X_j, Y_0\rangle\\ =&\frac{1}{\epsilon}\sum_{k=2}^{3}\int_0^t \nabla^L dF(y_s^\epsilon X_k, y_s^\epsilon X_j)\langle X_k, g_s^\epsilon Y_0(g_s^\epsilon)^{-1}\rangle\ \langle X_j, g_s^\epsilon Y_0(g_s^\epsilon)^{-1}\rangle\ ds\\ &+\frac{1}{\epsilon}\int_0^t dF(y_s^\epsilon X_j)\langle X_j, g_s^\epsilon[X_1, Y_0](g_s^\epsilon)^{-1}\rangle db_s\\ &-\frac{2}{\epsilon^2}\int_0^t dF(y_s^\epsilon X_j)\langle X_j, g_s^\epsilon[X_1, Y_0](g_s^\epsilon)^{-1}\rangle ds. \end{aligned}$$

Input the above information to the right hand side of (3.4), we obtain

$$\begin{aligned}F(y_t^\epsilon) &= F(u_0) - \frac{1}{2}\epsilon\left(dF(y_t^\epsilon g_t^\epsilon Y_0 (g_t^\epsilon)^{-1}) - dF(u_0 Y_0)\right)\\ &\quad + \frac{1}{2}\int_0^t \nabla^L dF(y_s^\epsilon g_s^\epsilon Y_0 (g_s^\epsilon)^{-1}, y_s^\epsilon g_s^\epsilon Y_0 (g_s^\epsilon)^{-1}) ds \\ &\quad + \frac{1}{2}\sum_{j=2}^3 \int_0^t dF(y_s^\epsilon X_j)\langle X_j, g_s^\epsilon [X_1, Y_0](g_s^\epsilon)^{-1}\rangle db_s.\end{aligned} \tag{3.7}$$

This computation is inspired by the fact that the two real valued functions on S^1, $\theta \mapsto \langle X_2, e^{i\theta} Y_0\rangle$ and $\theta \mapsto \langle X_3, e^{ß\theta} Y_0\rangle$, are eigenfunctions of the Laplace-Beltrami operator Δ_{S^1}. Also note that g_t^ϵ, as an element of S^1, takes the form $e^{t\theta_t^\epsilon}$.

Since F is a smooth function on compact manifolds, the probability distribution of $\{\tilde{x}_{\frac{t}{\epsilon}}^\epsilon, \epsilon > 0\}$ is tight, see Lemma 3.11 below. Thus there is a sequence ϵ_n decreasing to zero with the probability distribution of y^{ϵ_n} converges to a probability distribution on the space of paths over S^3 which we denote by μ.

By Theorem 4.2 [16], there is a probability space and a sub-sequence of numbers ϵ_{n_k} of ϵ_n, a family of stochastic processes $v_\cdot^{n_k}$ and $v_\cdot$, a one dimensional Brownian motion (w_t), such that $(v_\cdot^{n_k}, w_\cdot)$ equals to $(y_\cdot^{\epsilon_n}, b_\cdot)$ in law, u^k converges to v almost surely, whose probability distribution is μ, which we identify below. By Stroock-Varadhan's martingale method we first prove that $F(v_t) - F(v_0) - \int_0^t \bar{\mathcal{L}}F(v_s)ds$ is a martingale where

$$\bar{\mathcal{L}}F(u) = \sum_{j,k=2}^3 \nabla^L dF(uX_k, uX_j)\int_{S^1}\langle X_j, gY_0\rangle\langle X_k, gY_0\rangle dg.$$

We have to take care of two delicate points. The first point is that we have taken the liberty to enlarge the space of consideration from S^3 to $S^3 \times S^1$ and equation (3.7) involves the process $g_t^\epsilon \in S^1$. The second is to pass to the limit $k \to \infty$ in (3.7), where one of the terms is a stochastic integration. The first difficulty arises from the fact that the laws of the fast rotation $g_\cdot^\epsilon$ is not tight and we do not expect a convergent subsequence of the triple $(y_\cdot^\epsilon, g_\cdot^\epsilon, b_\cdot)$ and cannot legitimately assume that (3.7) holds with $(y_\cdot^{\epsilon_{n_k}}, b_\cdot)$ replaced by $(v_\cdot^k, w_\cdot)$.

Let us define a process h_t^k on S^1 by the following relation:

$$h_t^\epsilon(\omega) Y_0 (h_t^\epsilon(\omega))^{-1} = \epsilon (v_t^k)^{-1}(\omega)\dot{v}_t^\epsilon(\omega).$$

Then the triple (u_t^k, h_t^k, w_t) satisfies (3.7). We follow the above computation backwards, using (3.5) with g_t replaced by h_t. Replace the processes $(y_t^{\epsilon_{n_k}}, g_t^{\epsilon_{n_k}}, b_t)$

by (v_t^k, h_t^k, w_t) in (3.5) and (3.7) so we have:

$$\begin{aligned} F(v_t^k) =& F(u_0) - \frac{1}{2}\epsilon_{n_k}\left(dF(v_t^k h_t^k Y_0 (h_t^k)^{-1}) - dF(u_0 Y_0)\right) ds \\ &+ \frac{1}{2}\int_0^t \nabla^L dF(v_s^k h_s^k Y_0 (h_s^k)^{-1}, v_s^k h_s^k Y_0)(h_s^k)^{-1} ds \\ &+ \frac{1}{2}\sum_{j=2}^3 \int_0^t dF(v_s^k X_j)\langle X_j, h_s^k [X_1, Y_0](h_s^k)^{-1}\rangle db_s. \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} & dF(v_t^k h_t^k Y_0 (h_t^k)^{-1}) - dF(u_0 Y_0) \\ =& \frac{1}{\epsilon_{n_k}} \int_0^t \nabla^L dF(v_s^k h_s^k Y_0 (h_s^k)^{-1}, v_s^k h_s^k Y_0 (h_s^k)^{-1}) ds \\ &+ \int_0^t dF(v_s^{\epsilon_{n_k}} X_j) d\langle X_j, h_s^k Y_0 (h_s^k)^{-1}\rangle. \end{aligned}$$

Comparing with

$$F(v_t^k) = F(u_0) + \frac{1}{\epsilon_{n_k}} \int_0^t dF(v_s^k h_s^k Y_0 (h_s^k)^{-1}) ds$$

we obtain the following identity:

$$d\langle X_j, h_t^k Y_0 (h_t^k)^{-1}\rangle = \frac{1}{\epsilon}\langle X_j, g_t^\epsilon [X_1, Y_0](g_t^\epsilon)^{-1}\rangle \circ dw_t - 2\frac{1}{\epsilon^2}\langle X_j, g_t^\epsilon [X_1, Y_0](g_t^\epsilon)^{-1}\rangle dt.$$

Finally we recover that

$$dh_s^k = \frac{1}{\epsilon} h_s^k X_1 dw_t.$$

With this established we may wish to take conditional expectation of both sides of the identity (3.7) on the information up to time s. The appropriate conditional expectation would be with respect to

$$\begin{aligned} &\mathbb{E}\{F(v_t^k) - F(u_0) - \frac{1}{2}\epsilon_{n_k}\sum_{j=2}^3 \left(dF(v_t^k X_j)\langle X_j, v_t^k Y_0\rangle - dF(u_0 X_j)\langle X_j, Y_0\rangle\right) \\ &\quad + \sum_{j,l=2}^3 \int_0^t \nabla^L dF(v_s^k X_l, v_s^k X_j)\langle X_k, h_s^k Y_0 (h_s^k)^{-1}\rangle\langle X_j, h_s^k Y_0 (h_s^k)^{-1}\rangle\, ds - |\mathcal{F}_r\}. \end{aligned}$$

Conditioning on the filtration of the canonical process $\sigma\{v_r, r \le s\}$, the sequence

$$\frac{1}{2}\int_0^t \nabla^L dF(v_s^k h_s^k Y_0 (h_s^k)^{-1}, v_s^k h_s^k Y_0)(h_s^k)^{-1} ds$$

converges to

$$\sum_{j,k=2}^{3} \nabla^L dF(\bar v X_k, \bar v X_j) \int_{S^1} \langle X_j, gY_0g^{-1}\rangle \langle X_k, gY_0 g^{-1}\rangle dgds.$$

Here dg is the Haar measure on S^1. It is easy to check that

$$\int_{S^1} \langle X_2, gY_0\rangle\langle X_3, gY_0\rangle dg = 0,$$

either by direct computation or note that there is $g' \in U(1)$ such that $g'X_2 = -X_2$ and $g'X_3 = X_3$ and using the translation invariance of the Haar measure. Since there is an element of S^1 that maps X_2 to X_3,

$$\int_{S^1} \langle X_2, gY_0g^{-1}\rangle^2 dg = \int_{S^1} \langle X_3, gY_0g^{-1}\rangle^2 dg.$$

Note that

$$\sum_{j=1}^{2} \int_{S^1} \langle X_j, gY_0g^{-1}\rangle\langle X_j, gY_0g^{-1}\rangle dg = |gY_0g^{-1}|^2 = |Y_0|^2.$$

We conclude that $\tilde x^\epsilon_{\frac{t}{\epsilon}}$ converges in distribution and its law is determined by the generator $\bar{\mathcal{L}}F(u) = \frac{1}{2}|Y_0|^2\nabla^L dF(uX_2, uX_2) + \frac{1}{2}|Y_0|^2\nabla^L dF(uX_3, uX_3)$. Take $F = f\circ\pi$ with $f : S^2 \to \mathbb{R}$. Since $\nabla^L X_i^* = 0$ for $i = 2, 3$, $\sum_{i=2}^3 \nabla^L d(f\circ\pi)(T\pi(X_i^*), T\pi(X_i^*)) = \operatorname{trace}\nabla df$. Note also the Riemannian metric on S^2 is that induced from S^3, the limiting process has generator $\frac{1}{2}|Y_0|^2\Delta_{S^2}$ and is a Brownian motion when Y_0 is a unit vector. This together with the lemmas below concludes the proof of the theorem.

Lemma 3.11. *Let μ^ϵ be the probability distributions of the stochastic processes $(\tilde x^\epsilon_{\frac{t}{\epsilon}}, t \ge 0)$ in Theorem 3.10. Then $\{\mu^\epsilon, \epsilon > 0\}$ is relatively compact.*

Proof. Write $y_t^\epsilon = \tilde x^\epsilon_{\frac{t}{\epsilon}}$ for simplicity. Let μ_n be a subsequence from $\{\mu_\epsilon\}$ corresponding to a sequence of numbers ϵ_n. We wish to prove that it has a weakly convergent subsequence. It is sufficient to prove that the family of measures μ_n is tight, *i.e.* for every $\delta > 0$ there exists a compact set $K_\delta \subset M$ such that $\mu_n(K_\delta) > 1-\delta$ for all n. As probability measures on the space of continuous paths on M, $\mu_n(\sigma : \sigma(0) = y_0) = 1$ where $\sigma : [0,1] \to M$ is a continuous path on M. For any $y_1, y_2 \in M$, let $\phi : M\times M \to \mathbb{R}$ be a smooth function that agrees with the Riemannian distance function when $d(y_1,y_2) < a/2$ where a is the injectivity radius of M and $\phi(y_1, y_2) = 1$ when $d(y_1, y_2) > 2a$. This is possible by taking $\phi = \alpha \circ d$ where $\alpha : \mathbb{R}_+ \to \mathbb{R}$ is a suitable bump function with α the identity function on $[0, a/2]$. Then ϕ is a distance function on M that generates the same topology as d. The family of measures $\{\mu_n\}$ is tight if for any $a, \eta > 0$ there exists $0 < \delta < 1$ such that there is an $\epsilon_0 > 0$, with

$$\mathbb{P}\left(\omega : \sup_{|s-t|<\delta} \phi(y_s^{\epsilon_n}, y_t^{\epsilon_n}) > a\right) < \eta, \quad \text{when } \epsilon_n < \epsilon_0.$$

In the proof of the Theorem, take $F(y) = \phi(y, u)$. We omit the subscript n in ϵ_n. Then by formula (3.7),

$$\mathbb{E} \sup_{s \le \delta} \phi^2(y_s^\epsilon, u) \le \phi^2(y_0^\epsilon, u) + C + \delta$$

for come constant C. Let $\phi_t^\epsilon(y, \omega))$ denote $y_\cdot^\epsilon(\omega)$ with $y_0^\epsilon(\omega) = y$. Let θ_s denotes the shift operator in the Wiener space. By the Cocycle property, for $s < t$,

$$\mathbb{E} \sup_{|s-t| \le \delta} \phi^2(y_s^\epsilon, y_t^\epsilon) = \mathbb{E}\mathbb{E}\{ \sup_{|s-t| \le \delta} \phi^2(z, \phi_{t-s}^\epsilon(z, \theta_{t-s}(\omega))) | y_s^\epsilon = z\} \le C + \delta$$

and the required tightness holds. □

We return to explain in further detail the proof of the theorem.

Let $\{y_\cdot^\epsilon : \epsilon \in (0,1)\}$ be a family of Markov processes on a Riemannian manifold M that is relatively compact. We represent these processes as the coordinate process on path space with measure μ^ϵ, the distribution of $y_\cdot^\epsilon$. Let ϵ_n be sequence of numbers converging to zero. Suppose that μ^{ϵ_n} converges weakly to $\bar{\mu}$. Let $F : M \to \mathbb{R}$ be a smooth function with compact support. Let $\mathcal{A}$ be a diffusion operator. Suppose that

$$\int_\Omega f \left(F(X_t) - F(X_0) - \int_s^t \mathcal{A}F(X_r) dr \right) d\mu_n \to 0$$

for any bounded function f that is measurable with respect to $\mathcal{F}_s$ where $(\mathcal{F}_s, s \ge 0)$ is the canonical filtration. Then $\bar{\mu}$ is the probability distribution of a $\mathcal{A}$-diffusion. In fact letting $M_t^F = F(X_t) - F(X_0) - \int_s^t \mathcal{A}F(X_r) dr$, then M_t^F is a μ martingale and μ is the solution to the martingale problem associated to $\mathcal{A}$.

Finally we note that if

$$\mathcal{A}F(y, g) = 2 \sum_{j,k=2}^{3} \nabla^L dF(yX_k, yX_j) \langle X_j, gY_0 \rangle \langle X_k, gY_0 \rangle,$$

the averaged function with respect to dg is $\bar{\mathcal{L}}$. We also need the following dynamical law of large numbers (Birkhoff's ergodic theorem for Markov processes):

Lemma 3.12. *Let M, N be smooth compact manifolds. Suppose that $(y_t^n, g_{\frac{t}{n}})$ is a $M \times N$ valued stochastic processes on a probability space and y_t^n converges weakly to y. Suppose that g_t has a unique invariant measure dg. Let $f : M \to \mathbb{R}$ be a smooth function. Then the following convergence holds:*

$$\mathbb{E}\Big| \int_s^t f(y_r^n, g_{\frac{r}{n}})dr - \int_s^t \int f(y_r, g)dgdr \Big| \to 0$$

In addition, for any real valued bounded function ϕ on the path space,

$$\mathbb{E}\phi(v_r^n, r \leq s) \left(F(v_t^n) - F(v_s^n) - \int_s^t \bar{\mathcal{L}}F(v_r^n)dr \right) \to 0.$$

To prove this, we let $t_0 = s < t_1 < \cdots < t_N = t$ be a division of $[s, t]$, set $\Delta t_t = t_{i+1} - t_i$. Assume that $\Delta t_i = \frac{1}{\sqrt{N}}$ so $N\Delta t_i = \sqrt{N}$ is sufficiently large for the effect of the law of large numbers to take effect. On each interval $[Nt_i, Nt_{i+1}]$, we make an approximation to conclude.

4 Stochastic Perturbation On Frame Bundles

In this section we study the SDE on the orthonormal frame bundle

$$\left\{ \begin{array}{ll} du_t^\epsilon & = \sqrt{\epsilon}_1 \sum_{l=1}^m \mathbb{X}_l(u_t^\epsilon) \circ db_t^l + \epsilon \mathbb{X}_0(u_t^\epsilon)dt + \sqrt{\delta} \sum_{j=1}^p Z_j(u_t^\epsilon) \circ dw_t^j + \delta Z_0(u_t^\epsilon)dt, \\ u_0^\epsilon & = u_0. \end{array} \right. \tag{4.1}$$

where $\{\mathbb{X}_l, l = 0, 1, 2, \ldots, m\}$ is a family of horizontal vector fields and $\{Z_j, j = 0, 1, \ldots p\}$ is a family of vertical vector fields. The parameters ϵ_1, ϵ and δ take values in $\mathbb{R}_+$. The Markov generator is $\frac{1}{2}\epsilon_1 \sum_{l=1}^m \mathbb{X}_l^2 + \epsilon \mathbb{X}_0 + \delta \sum_{j=1}^m Z_j^2 + \delta Z_0$.

4.1 Perturbation of Ornstein-Uhlenbeck Type

Based on E. Nelson's Ornstein-Uhlenbeck theory of Brownian motions [34] we ask the following question. What happens if we replace the driving white noise $\dot{w}_t$ by an Ornstein-Uhlenbeck process? Consider the position process z_t in $\mathbb{R}^n$ with velocity process satisfy the Langevin equation:

$$\begin{aligned} dv_t^\epsilon &= -\frac{1}{\epsilon} v_t^\epsilon dt + \frac{1}{\epsilon} dw_t \\ \dot{z}_t^\epsilon &= v_t^\epsilon. \end{aligned}$$

Here w_t is a Brownian motion with values in $\mathbb{R}^n$ and $z_0 = 0$. The z_t^ϵ process converges to w_t as $\epsilon \to 0$. The convergence holds almost surely and in fact the result holds if w_t is replaced by any continuous function.

We now interpret the convergence in terms of homogenisation. First we rescale the variables in space and time and setting $\tilde{v}_t = \sqrt{\epsilon} v_t, \tilde{z}_t = \sqrt{\epsilon} z_t$. It is easy to see that $\tilde{z}_t^\epsilon$ is the slow variable and $\sqrt{\epsilon} z_{\frac{t}{\epsilon}}$ converges to a Brownian motion. In fact

$$d\tilde{v}_t^\epsilon = -\frac{1}{\epsilon}\tilde{v}_t^\epsilon dt + \frac{1}{\sqrt{\epsilon}} dw_t, \qquad \dot{\tilde{z}}_t^\epsilon = \tilde{v}_t^\epsilon.$$

We take this model to the orthonormal bundle. First we are not allowed to rescale variables in non-linear spaces. We should not rescale the frame variable in the orthonormal frame bundle, in space, either.

Let $e_0 \in \mathbb{R}^n$ be a unit vector and $\{A_0, A_k, k = 1, 2, \dots, N = n(n-1)/2\}$ be elements of $\mathfrak{g}$. Let A_k^* be the corresponding fundamental vertical vector field corresponding to A_k. Consider

$$du_t^\epsilon = H(u_t^\epsilon)(e_0)dt + \frac{1}{\sqrt{\epsilon}} \sum_{k=1}^N A_k^*(u_t^\epsilon) \circ dw_t^k + \frac{1}{\epsilon} A_0^*(u_t^\epsilon) dt.$$

For '$\epsilon = \infty$', the equation can be considered as the 'geodesic flow' equation, as explained earlier. If $x_t^\epsilon = \pi(u_t^\epsilon)$ then $\dot{x}_t^\epsilon = u_t^\epsilon e_0$. Note that the change of the velocity of the motion on M is always unitary. Due to the fast rotation, the geodesic has rapid changing directions and we expect to see a jittering motion and indeed we obtain a scaled Brownian motion in the limit if the rotational motion is elliptic.

A related theorem is given in Dowell [6] stating that an Ornstein-Uhlenbeck position process on 2-uniformly smooth Banach manifolds converges. Those are manifolds modelled on 2-uniformly smooth Banach spaces. By a 2-uniformly smooth Banach space B we mean one with the property that there is a constant $C > 0$ such that $||x+y||^2 + ||x-y||^2 \le 2||x||^2 + C||y||^2$ holds for all $x, y \in B$. The iterated Ornstein-Uhlenbeck processes in [6]. We expect that interesting results arise for processes with infinite-dimensional noise. For a related work, central limit theorem for geodesic flows, we refer to Enriquez-Franchi-LeJan [12].

As an example, we present a theorem from [28] generalising the main theorem in [26]. Let M be a complete Riemannian manifold of dimension $n > 1$ of positive injectivity radius. If $A \in \mathfrak{so}(n)$ we define

$$A^*(u) = \frac{d}{dt}\Big|_{t=0} u \exp(tA).$$

Then A^* is the left invariant vector field on G induced by $A \in \mathfrak{g}$, in other words $A^*(g) = gA$. Let $\{A_1, \dots, A_m\}$ be a set of normal vectors in $\mathfrak{so}(n)$ such that they and their brackets generate $\mathfrak{so}(n)$. Let $\mathcal{A} = \frac{1}{2}\sum_k (A_k^*)^2 + A_0^*$. In case $m = \frac{1}{2}n(n-1)$, $\mathcal{A}$ is the the left invariant Laplacian on G which we denote by Δ^L.

Let $x_0 \in M$ and $u_0 \in \pi^{-1}(x_0)$. Consider the SDE

$$\begin{aligned} du_t^\epsilon &= H(u_t^\epsilon)(e_0)dt + \frac{1}{\sqrt{\epsilon}}\sum_{k=1}^N A_k^*(u_t^\epsilon) \circ dw_t^k + \frac{1}{\epsilon}A_0^*(u_t^\epsilon)dt, \\ u_0^\epsilon &= u_0. \end{aligned} \tag{4.2}$$

where $u_0 \in OM$ is a common initial condition. Let (u_t^ϵ) be the solution to (4.2), $x_t^\epsilon = \pi(u_t^\epsilon)$ its projection, and let $(\tilde{x}_t^\epsilon)$ be the horizontal lift of (x_t^ϵ) to OM through u_0. As usual we define the bracket of $[A_j, A_k] = A_jA_k - A_kA_j$.

Theorem 4.1. *Let M be a compact smooth Riemannian manifold. Suppose that $\{A_k, k \geq 1\}$ and their iterated brackets spans $\mathfrak{g}$. Then the following results hold:*

(1) As $\epsilon \to 0$, the processes $(x_{\frac{t}{\epsilon}}^\epsilon)$ and $(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon)$ converge in law.

(2) The limiting law of $(x_{\frac{t}{\epsilon}}^\epsilon)$ is universal, independent of e_0. The generator can be explicitly given.

This result, for the elliptic fast motion case, was published in [26]. The generalisation to the case where the noise is not given in every direction is presented in [28]. In case $\{e_i\}$ is an orthonormal basis of $\mathbb{R}^n$ and we may assume that e_0 is a unit vector. Then $\pi(u_{\frac{t}{\epsilon}}^\epsilon)$ converges in law to a rescaled Brownian motion with generator $\frac{4}{n(n-1)}\Delta$. Its horizontal lift converges in law to the diffusion process on OM with generator $\frac{4}{n(n-1)}\Delta_H$. The main idea for the proof is to show that $d\tilde{x}_t^\epsilon = H(\tilde{x}_t^\epsilon)(g_t^\epsilon e_0)$ where $g_t^\epsilon = g_{\frac{t}{\epsilon}}$ and g_t is the solution to the SDE with diffusion coefficients the left invariant vector fields $\{A_k^* : 1 \leq k \leq N\}$ with drift A_0^*. Under the conditions of the theorem, g_t had a unique invariant probability measure, the Haar measure.

4.2 Perturbation to Vertical Flows

In (4.1) take $\epsilon_1 = \epsilon$ and $\delta = 1$. Let $\tilde{L}^\epsilon = \mathcal{L}_0 + \epsilon\mathcal{L}_1$ where

$$\mathcal{L}_1 = \frac{1}{2}\sum_{l=1}^m L_{\mathbb{X}_l}L_{\mathbb{X}_l} + L_{\mathbb{X}_0}, \mathcal{L}_0 = \frac{1}{2}\sum_{j=1}^p L_{Z_j}L_{Z_j} + \mathcal{L}_{Z_0}.$$

Define $\mathbf{X}_l(u,g) = TR_{g^{-1}}\mathbb{X}_l(u)$.

Theorem 4.2. *Assume that M has positive injectivity radius, $\{\varpi_u[Z_j(u)]\}_{j=1}^m$ spans $\mathfrak{g}$, and the vector fields $\{\mathbb{X}_l, l \ge 0\}$ and $\{|\breve{\nabla}_{\mathbb{X}_l}\mathbb{X}_l|, \ge 1\}$ have linear growth. Let u_t^ϵ be a solution with initial value $u_0 \in OM$, to the SDE*

$$du_t^\epsilon = \sqrt{\epsilon}\sum_{l=1}^m \mathbb{X}_l(u_t^\epsilon)\circ db_t^l + \epsilon\mathbb{X}_0(u_t^\epsilon)dt + \sum_{j=1}^p Z_j(u_t^\epsilon)\circ dw_t^j + Z_0(u_t^\epsilon)dt. \tag{4.3}$$

Let $x_t^\epsilon = \pi(u_t^\epsilon)$ and $\tilde{x}_t^\epsilon$ its horizontal lift. Then $\tilde{x}_{\frac{t}{\epsilon}}^\epsilon$ converges weakly to a Markov process. Furthermore, Let π_u be the invariant measure of the following SDE on G:

$$dg_t = \sum_{j=1}^m TL_{g_t}\varpi[Z_j(ug_t)]\circ dw_t^j + TL_{g_t}\varpi[Z_0(ug_t)]dt.$$

and let $\mathcal{L}^u$ denote the Markov generator of

$$d\tilde{u}_t = \sum_{l=1}^p \mathbf{X}_l(\tilde{u}_t, g)\circ db_t^l + \mathbf{X}_0(\tilde{u}_t, g)dt. \tag{4.4}$$

Then the generator $\bar{\mathcal{L}}$ of the limiting Markov process is obtained from averaging that of $\mathcal{L}^u$ with respect to π^u.

Remark 4.3. In other word, define

$$b(u) = \int_G \left(\frac{1}{2}\sum_{l=1}^p \breve{\nabla}_{\mathbf{X}_l}\mathbf{X}_l(ug) + \mathbf{X}_0(ug)\right) d\pi_u(g)$$

$$a_{i,j}(u) = \int_G \sum_{l=1}^p \langle TR_g^{-1}\mathbb{X}_l(ug), H_i(u)\rangle\langle TR_g^{-1}\mathbb{X}_l(ug), H_j(u)\rangle \; d\pi_u(g),$$

with limiting generator $\bar{\mathcal{L}}$. For $F: OM \to \mathbb{R}$ smooth with compact support,

$$\bar{\mathcal{L}}F(u) = dF(b(u)) + \frac{1}{2}\sum_{i,j=1}^p a_{i,j}(u)\breve{\nabla}dF(H_i(u), H_j(u)).$$

Proof. We first show that $\{\tilde{x}_t^\epsilon\}$ is tight, then identify their limiting theorem by the martingale problem by showing that every accumulation point is a Markov process with the same generator.

Since $\tilde{x}_t^\epsilon$ and u_t^ϵ belong to the same fibre we may define $g_t^\epsilon \in G$ by $u_t^\epsilon = \tilde{x}_t^\epsilon g_t^\epsilon$. If a_t is a C^1 curve in G

$$\frac{d}{dt}|_t u a_t = \frac{d}{dr}_{|_{r=0}} u a_t a_t^{-1} a_{r+t} = (a_t^{-1}\dot{a}_t)^*(u a_t).$$

It follows that

$$du_t^\epsilon = TR_{g_t^\epsilon} d\tilde{x}_t^\epsilon + (TL_{(g_t^\epsilon)^{-1}} dg_t^\epsilon)^*(u_t^\epsilon).$$

Since right translation of horizontal vectors are horizontal, $\varpi(du_t^\epsilon) = TL_{(g_t^\epsilon)^{-1}} dg_t^\epsilon$ and

$$dg_t^\epsilon = \sum_{j=1}^m TL_{g_t^\epsilon} \varpi[Z_j(\tilde{x}_t^\epsilon g_t^\epsilon)] \circ dw_t^j + TL_{g_t^\epsilon} \varpi[Z_0(\tilde{x}_t^\epsilon g_t^\epsilon)]dt. \tag{4.5}$$

For each $u \in OM$, we consider the auxiliary process

$$dg_t = \sum_{j=1}^m TL_g \varpi[Z_j(ug_t)] \circ dw_t^j + TL_{g_t} \varpi[Z_0(ug_t)]dt.$$

Since the SDE is elliptic and $SO(n)$ is compact, it has a unique invariant probability measure which we denote by π_u.

By Itô's formula, $dx_t^\epsilon = \sqrt{\epsilon} \sum_{l=1}^p T\pi(\mathbb{X}_l(u_t^\epsilon)) \circ db_t^l + \epsilon T\pi(\mathbb{X}_0(u_t^\epsilon))dt$ so

$$d\tilde{x}_t^\epsilon = \mathfrak{h}_{\tilde{x}_t}(\circ dx_t^\epsilon) = \sqrt{\epsilon} \sum_{l=1}^p \mathfrak{h}_{\tilde{x}_t^\epsilon}[T\pi(\mathbb{X}_l(u_t^\epsilon))] \circ db_t^l + \epsilon \mathfrak{h}_{\tilde{x}_t^\epsilon}[T\pi(\mathbb{X}_0(u_t^\epsilon))]dt.$$

By the assumptions on the vector fields $\mathbb{X}_l$, the above SDE is conservative and $\pi(u_t^\epsilon)$ exists for all time. We introduce the notation $\mathbf{X}_l(u, g) := \mathfrak{h}_u \pi_*(\mathbb{X}_l(ug))$. By the right invariance of the horizontal lift,

$$\mathbf{X}_l(\tilde{x}_t^\epsilon, g_t^\epsilon) := \mathfrak{h}_{\tilde{x}_t^\epsilon}[T_{\tilde{x}_t^\epsilon g_t^\epsilon} \pi(\mathbb{X}_l(\tilde{x}_t^\epsilon g_t^\epsilon)] = TR_{(g_t^\epsilon)^{-1}} \mathbb{X}_l(u_t^\epsilon).$$

We have

$$d\tilde{x}_t^\epsilon = \sqrt{\epsilon} \sum_{l=1}^p \mathbf{X}_l(\tilde{x}_t^\epsilon, g_t^\epsilon) \circ db_t^l + \epsilon \mathbf{X}_0(\tilde{x}_t^\epsilon, g_t^\epsilon)dt.$$

Changing to a different time scale $t \mapsto t/\epsilon$, we want to show that the probability distributions of the stochastic processes $\{\tilde{x}_{\frac{t}{\epsilon}}^\epsilon : 0 < \epsilon \le 1\}$ is relatively compact. By Prohorov's theorem it is sufficient to show that they is a tight family of probability measures. Since $\tilde{x}_0^\epsilon = u_0$ it suffices to estimate the modulus of continuity and show that for all positive numbers a, η, there exists $\delta > 0$ such that for all ϵ reasonably small, see Billingsley [3] and Ethier-Kurtz[13],

$$P(\omega : \sup_{|s-t|<\delta} d(\tilde{x}_{t/\epsilon}^\epsilon, \tilde{x}_{s/\epsilon}^\epsilon) > a) < \eta.$$

Here d denotes a distance function on OM. The Riemannian distance function is not smooth on the cut locus of OM. We construct a smooth distance function preserving the topology of OM.

Let $x \in M$ and $2a$ the minimum of 1 and the injectivity radius of M. Let $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ be a smooth concave function such that $\phi(r) = r$ when $r < a$ and $\phi(r) = 1$ when $r \geq 2a$, e.g. ϕ is the convolution of $\min(1,r)$ with a standard mollifier supported in the set $\{r : |r - \frac{3a}{2}| < a/2\}$. Let ρ and $\tilde{\rho}$ be respectively the Riemannian distance on M and OM. Then $\phi \circ \rho$ and $d := \phi \circ \tilde{\rho}$ are distance functions. For $u \in \pi^{-1}(x)$,

$$\begin{aligned}
\phi \circ \tilde{\rho}(u, \tilde{x}_t^\epsilon) =& (\phi \circ \tilde{\rho})(u, \tilde{x}_0^\epsilon) + \int_0^t d(\phi \circ \tilde{\rho}) \left(\sqrt{\epsilon} \sum_{l=1}^p \mathfrak{h}_{\tilde{x}_s^\epsilon}[T\pi(\mathbb{X}_l(u_s^\epsilon))] \circ dB_s^l \right) \\
&+ \int_0^t \epsilon \, d(\phi \circ \tilde{\rho}) \, \mathfrak{h}_{\tilde{x}_s^\epsilon}[T\pi(\mathbb{X}_0(u_s^\epsilon))] \, ds \\
=& (\phi \circ \tilde{\rho})(u, \tilde{x}_0^\epsilon) + \int_0^t d(\phi \circ \rho) \left(\sqrt{\epsilon} \sum_{l=1}^p [T\pi(\mathbb{X}_l(u_s^\epsilon))] dB_s^l \right) \\
&+ \epsilon \sum_{l=1}^p \int_0^t \nabla d(\phi \circ \rho) \left(T\pi(\mathbb{X}_l(u_s^\epsilon)), T\pi(\mathbb{X}_l(u_s^\epsilon)) \right) \, ds \\
&+ \epsilon \int_0^t d(\phi \circ \rho) \left(\frac{1}{2} \sum_{l=1}^p \nabla_{T\pi(\mathbb{X}_l)} (T\pi \circ \mathbb{X}_l)(u_s^\epsilon) + T\pi(\mathbb{X}_0(u_s^\epsilon)) \right) ds.
\end{aligned}$$

Since $\phi \circ \rho$ has compact support and the vector fields concerned have linear growth, $|T\pi(\mathbb{X}_l(u_s^\epsilon))| \leq C(1 + \rho(u_s^\epsilon, u)) \leq [C + C\tilde{\rho}(\tilde{x}_s^\epsilon, \tilde{u}_s^\epsilon)] + C\tilde{\rho}(u, \tilde{x}_s^\epsilon)$ for some fixed $u \in OM$. The quantity $C + C\tilde{\rho}(\tilde{x}_s^\epsilon, \tilde{u}_s^\epsilon)$ is bounded from the compactness of G and it follows that $\mathbb{E}[\phi \circ \tilde{\rho}(u, \tilde{x}_t^\epsilon)]^2) \leq C_1(t)((\phi \circ \tilde{\rho})^2(u, \tilde{x}_0^\epsilon) + \epsilon t)$ for some constant C. By the Markov property and the estimates below,

$$\mathbb{E}[\phi \circ \tilde{\rho}(\tilde{x}_{\frac{s}{\epsilon}}^\epsilon, \tilde{x}_{\frac{t}{\epsilon}}^\epsilon)]^2 \leq C_1 |t - s|.$$

Consequently $\{\tilde{x}_{\frac{t}{\epsilon}}^\epsilon) : \epsilon(0,1)\}$ is tight. By Prohorov's theorem, the probability distributions of $\{\tilde{x}_{\frac{t}{\epsilon}}^\epsilon, \epsilon \in (0,1]\}$ is relatively compact on any finite time interval. It is sufficient to identify their accumulation points.

For this purpose, we may take a sequence $\epsilon_n \to 0$ with the property that $\tilde{x}_{\frac{t}{\epsilon}}^{\epsilon_n}$ converges in law to a probability measure μ. Let $\mu_n = law(\tilde{x}_{\frac{t}{\epsilon}}^{\epsilon_n})$.

Let $F : OM \to \mathbb{R}$ be a smooth function. For $\breve{\nabla}$, the canonical direct sum connection on OM associated to ∇,

$$\begin{aligned}
F(\tilde{x}_t^\epsilon) \quad = \quad & F(u_0) + \sqrt{\epsilon} \textstyle\sum_{l=1}^p \int_0^t dF \left(TR_{(g_s^\epsilon)^{-1}} \mathbb{X}_l(u_s^\epsilon) \right) dB_s^l \\
& + \tfrac{1}{2} \epsilon \textstyle\sum_{l=1}^p \int_0^t \breve{\nabla} dF \left(TR_{(g_s^\epsilon)^{-1}} \mathbb{X}_l(u_s^\epsilon), TR_{(g_s^\epsilon)^{-1}} \mathbb{X}_l(u_s^\epsilon) \right) ds \\
& + \tfrac{1}{2} \epsilon \textstyle\sum_{l=1}^p \int_0^t dF \left(\breve{\nabla}_{\mathbf{X}_l} \mathbf{X}_l(u_s^\epsilon) + \mathbf{X}_0(u_s^\epsilon) \right) ds.
\end{aligned} \tag{4.6}$$

Note that $\tilde{x}_t^\epsilon = \tilde{u}_t^\epsilon g_t^\epsilon$. Suppose that it has a convergent subsequence which we denote by $\tilde{x}^{\epsilon_n}_{\frac{t}{\epsilon_n}}$. We define,

$$\begin{aligned}\int_0^{\frac{t}{\epsilon}} \mathcal{A}^\epsilon F(\tilde{u}_s^\epsilon, g_s^\epsilon)\, ds = \quad & \tfrac{1}{2}\epsilon \textstyle\sum_{l=1}^p \int_0^{\frac{t}{\epsilon}} \breve{\nabla} dF\left(\mathbf{X}_l(\tilde{x}_s^\epsilon, g_s^\epsilon), \mathbf{X}_l(\tilde{x}_s^\epsilon, g_s^\epsilon)\right) ds \\ & + \tfrac{1}{2}\epsilon \textstyle\sum_{l=1}^p \int_0^{\frac{t}{\epsilon}} dF\left(\breve{\nabla}_{\mathbf{X}_l}\mathbf{X}_l((\tilde{x}_s^\epsilon, g_s^\epsilon) + \mathbf{X}_0((\tilde{x}_s^\epsilon, g_s^\epsilon)\right) ds.\end{aligned} \tag{4.7}$$

Let $X.$ be the coordinate process on the path space and $\mathcal{G}_t = \sigma\{X_s : 0 \leq s \leq t\}$. To identify the limiting process it suffices to show that for all real-valued smooth functions F on OM with compact support and for any real-valued bounded $\mathcal{G}_s$-measurable function ϕ on the path space, the following holds:

$$\int \left(F(X_t) - F(X_s) - \int_s^t \bar{\mathcal{L}}F(X_r)dr\right) \phi\, d\mu^\epsilon \to 0.$$

Write $z_t^n = \tilde{x}^\epsilon_{\frac{t}{\epsilon}n}$. Let ϕ be a $\{z_s^n, s \leq t\}$-adapted bounded function. It is equivalent to show that for $t \geq s$,

$$\begin{aligned}&\mathbb{E}\phi\left(F(z_t^n) - F(z_s^n) - \int_s^t \bar{\mathcal{L}}F(z_r^n)dr\right) \\ &\qquad = \mathbb{E}\left[\phi \int_s^t (\mathcal{A}^{\epsilon_n} F(z_r^n, g_r^{\epsilon_n}) - \bar{\mathcal{L}}F(z_r^n))dr\right] \to 0,\end{aligned} \tag{4.8}$$

where $\mathcal{A}^{\epsilon_n} F$ is given by the bounded variation part in (4.7). This follows from the ergodic theorem. Since G is compact and also the invariant measure π_u for the elliptic SDE (4.5) is ergodic, Birkhoff's ergodic theorem shows that

$$\int_0^t \Phi(g_r^{\epsilon_n})(dr) \to t \int_{OM} \Phi(g)\pi_u(dg)$$

for any integrable function Φ. Now using the fact the z_n converges, the right hand side converges to zero. With this we conclude that $\tilde{x}^\epsilon_{\frac{\cdot}{\epsilon}}$ converge in distribution to a Markov process with generator $\bar{\mathcal{L}}$, completing the proof. □

Example 4.4. Let $\alpha_l : M \times \mathbb{R}^n \to \mathbb{R}^n$ be smooth maps so that $\alpha_l(x) \in \mathbb{L}(\mathbb{R}^n; \mathbb{R}^n)$. Let$\{e_i\}_{i=1}^n$ be an o.n.b. of $\mathbb{R}^n$, $e_0 \in \mathbb{R}^n$. Consider

$$\begin{aligned}du_t^\epsilon =& \sqrt{\epsilon}\sum_{l=1}^n \alpha_l(\pi(u_t^\epsilon))H_l(u_t^\epsilon) \circ db_t^l \\ &+ \epsilon\alpha_0(\pi(u_t^\epsilon))H_{e_0}(u_t^\epsilon)\, dt + \sum_{j=1}^m Z_j(u_t^\epsilon) \circ dw_t^j + Z_0(u_t^\epsilon)dt.\end{aligned} \tag{4.9}$$

The projection $x_t^\epsilon = \pi(u_t^\epsilon)$ satisfies:

$$dx_t^\epsilon = \sqrt{\epsilon}\sum_{l=1}^n \alpha_l(x_t^\epsilon)u_t^\epsilon \circ db_t^l + \epsilon\alpha_0(x_t^\epsilon)u_t^\epsilon(e_0)dt.$$

Let $\tilde{x}_t^\epsilon = \mathfrak{h}_{\tilde{x}_t^\epsilon}(x_{\frac{t}{\epsilon}}^\epsilon)$ and g_t^ϵ be an element of G determined by $u_{\frac{t}{\epsilon}}^\epsilon = \tilde{x}_t^\epsilon\, g_{\frac{t}{\epsilon}}^\epsilon$, where $g_{\frac{t}{\epsilon}}^\epsilon$ is the ergodic process on the group. Then $d\tilde{x}_t^\epsilon = \sqrt{\epsilon}\sum_l \alpha_l(x_t^\epsilon)H_l(\tilde{x}_t^\epsilon)g_t^\epsilon \circ db_t^l + \epsilon\alpha_0(x_t^\epsilon)H(\tilde{x}_t^\epsilon)g_t^\epsilon(e_0)dt$. Note that the limit of

$$\sum_i \nabla df(\alpha_i(\pi(u))ug_t^\epsilon e_i, \alpha_i(\pi(u))ug_t^\epsilon e_i) \tag{4.10}$$

where $f : M \to \mathbb{R}$ is a smooth function, is in general not a trace.

4.3 Another Intertwined Pair

At this point we discuss a question asked to me by J. Norris. Since the process on the orthonormal frame bundle encodes the Riemannian metric we expect to see the Riemannian metric manifesting itself in some form, e.g. in the form of the corresponding Laplacian operator. Does the system below have a non-degenerate limit which is not necessarily associated to the given Riemannian metric on M? In general the intertwined system would look like the following,

$$\begin{aligned}
du_t^\epsilon &= CH(u_t^\epsilon)\circ db_t^\epsilon + \frac{1}{\sqrt{\epsilon}}H(u_t^\epsilon)V(x_t^\epsilon, g_t^\epsilon)dt + \frac{1}{\sqrt{\epsilon}}A_k^*(u_t^\epsilon)\circ dw_t^k + \frac{1}{\epsilon}A_0^*(u_t^\epsilon)dt\\
dx_t^\epsilon &= Cu_t^\epsilon \circ db_t^\epsilon + \frac{1}{\sqrt{\epsilon}}u_t^\epsilon V(x_t^\epsilon, g_t^\epsilon)dt.
\end{aligned}$$

Below we compute a simple case. The argument, with suitable adjustments, remains valid for the general case.

Example 4.5. For simplicity consider $\mathbb{R}^n \times SO(n)$ with the standard connection, and the SDE

$$\begin{aligned}
dg_t^\epsilon &= \frac{1}{\sqrt{\epsilon}}g_t^\epsilon A_k \circ dw_t^k\\
dx_t^\epsilon &= \delta g_t^\epsilon \circ db_t + \frac{1}{\sqrt{\epsilon}}g_t^\epsilon V(x_t^\epsilon, g_t^\epsilon)dt.
\end{aligned} \tag{4.11}$$

Here V is a $\mathbb{R}^n$ valued function such that $\int_G gV(x,g)dg = 0$ where dg is the Haar measure. For example take $V(g)$ to be a function of even powers of g. We assume that V is suitably bounded with its partial derivatives in x suitably bounded. The parameter δ is to be chosen.

Letting $A_k^*(g) = gA_k$. $\mathcal{L}_0 = \frac{1}{2}\sum_k (A_k^*)^2$, assume that it is $\frac{1}{2}\Delta^L$. Taking $\delta = \sqrt{\epsilon}$, formal computation by multi scale analysis shows that :

Claim. The limiting law for x_t^ϵ is governed by the partial differential equation on $\mathbb{R}^n$:

$$\frac{\partial \rho}{\partial t} = -\int L_{gV(x,g)\partial_x} \mathcal{L}_0^{-1}(gV(x,g)\rho)\, dg, \tag{4.12}$$

where the integral is with respect to the Haar measure on $SO(n)$.

If $\delta = 1$ it ought to have, in addition, a Δ_M term on the right hand side:

$$\frac{\partial \rho}{\partial t} = \Delta_M f_t - \int L_{uV(x,u)} \mathcal{L}_0^{-1}(L_{uV(x,u)}\rho) d\nu(u),$$

which we do not discuss rigorously. A drift term in the g equation can also be added. Another interesting regime to consider is $\sum_i \delta_i g_t^\epsilon \circ db_t^i$ instead of $g_t^\epsilon \circ db_t$ with δ_i takes values from $\{1, \sqrt{\epsilon}\}$. In this case, a non-Laplacian like equation would follow. In the case that δ_i are all equal and $V(x,g)$ is independent of x, the system can be interpreted as an intertwined pair through time scaling.

Equation (4.12) can be deduced by the methodology below. Let $f : M \to \mathbb{R}$ be a smooth compactly supported function and Δ_M the Laplacian on M. Then

$$f(x_t^\epsilon) = f(x_0) + \delta \int_0^t df(g_s^\epsilon db_s) + \frac{1}{2}\delta^2 \int_0^t \Delta_M f(x_s^\epsilon) ds + \frac{1}{\sqrt{\epsilon}} \int_0^t df(g_s^\epsilon V(x_s^\epsilon, g_s^\epsilon)) ds.$$

If h is solution to $\mathcal{L}_0 h(x,g) = df_x(gV(x,g))$, then

$$\begin{aligned} &\frac{1}{\sqrt{\epsilon}} \int_0^t df(g_s^\epsilon V(x_s^\epsilon, g_s^\epsilon)) ds \\ &= \sqrt{\epsilon} h(x_t^\epsilon, u_t^\epsilon) - \sqrt{\epsilon} h(x_0, u_0) - \sqrt{\epsilon}\delta \int_0^t \partial_x h(x_s^\epsilon, g_s^\epsilon) g_s^\epsilon \, db_s \\ &- \int_0^t \partial_g h(g_s^\epsilon A_k dw_s^k) - \sqrt{\epsilon}\delta^2 \int_0^t \Delta_M h(x_s^\epsilon, g_s^\epsilon) ds \\ &- \int_0^t L_{g_s^\epsilon V_s^\epsilon \partial_x} h(x_s^\epsilon, g_s^\epsilon) ds. \end{aligned}$$

Since $\delta = \sqrt{\epsilon}$, it is now easy to observe that $\{x_t^\epsilon\}$ is a tight family. Since g_t^ϵ is a fast ergodic motion and x_t^ϵ does not move much as $t \to 0$, under suitable conditions,

$$\lim_{\epsilon \to 0} \lim_{t \to 0} \frac{\mathbb{E} f(x_t^\epsilon) - f(x_0)}{t} = \lim_{t \to 0} \lim_{\epsilon \to 0} \frac{\mathbb{E} f(x_t^\epsilon) - f(x_0)}{t} = \mathcal{L}_{gV\partial_x} h(x_0).$$

has the required limit.

4.4 Perturbation to Horizontal Diffusions and effective diffusion on the holonomy bundle

Let M be a compact connected n-dimensional smooth Riemannian manifold with a Riemannian connection ∇. The horizontal bundle is integrable when and only when the curvature tensor of ∇ vanishes. The Lie brackets of two fundamental horizontal vector fields will in general contribute to a vertical motion. However perturbation to horizontal flows can still be discussed and in this case we should consider not its projection to the manifold M unless the connection ∇ is flat, but only its motion transversal to the holonomy bundle.

Let ϖ be the corresponding connection 1-form on the orthonormal frame bundle OM with Lie group G taken to be $O(n)$ or $SO(n)$ depending whether M is oriented. This connection determines the horizontal maps $\mathfrak{h}_u : T_{\pi(u)}M \to T_uOM$. A curve σ on OM is horizontal if it is derivative is horizontal, meaning that $\varpi(\dot{\sigma}) = 0$. A curve $\tilde{\tau}$ on OM is a horizontal lift of a curve τ on M if $\tilde{\tau}$ is horizontal and such that $\pi(\tilde{\tau}(t)) = \tau(t)$. Given a smooth curve τ on M and any $u \in \pi^{-1}(\tau(0))$ there is a unique horizontal lift $\tilde{\tau}$ with the initial value u. These curves solve the equations

$$\frac{d}{dt}\tilde{\tau}(t) = \mathfrak{h}_{\tilde{\tau}(t)}(\dot{\tau}(t)) = H(\tilde{\tau}(t))(\tilde{\tau}(t)^{-1}\dot{\tau}(t)).$$

Let $u_0 \in OM$ and $\tau : [0,1] \to M$ be a closed C^1 curve with $\tau(0) = \tau(1) = \pi(u_0)$. Let $\tilde{\tau}$ be the horizontal lift of τ through u_0. The displacement, $\tilde{\tau}_1$, of u_0 can be written as $u_0 a$ some $a \in G$. The set of such a that represents parallel displacements of u_0 forms a subgroup of G and is called the holonomy group with reference point $u_0 \in OM$ which we denote by $\Phi(u_0)$. In other words $a \in \Phi(u_0)$ if u_0 and $u_0 a$ are connected by a horizontal curve. Furthermore we denote by $\Phi_0(u_0)$ the restricted holonomy group which contains group elements arising only from loops that are homotopic to the identity loop. By Theorem 4.2, in Kobayashi-Nomizu [21], $\Phi(u_0)$ is a Lie subgroup of $SO(n)$ with $\Phi_0(u_0)$ its identity component. The restricted holonomy group $\Phi_0(u)$ is a normal subgroup of $\phi(u)$ and a path connected Lie subgroup of G. Since M is connected all holonomy groups are isomorphic.

Two points u and v of OM are equivalent, which we denote by the symbol $u \sim v$, if they are connected by a C^1 horizontal curve. For each u in OM, let $P(u_0)$ be the holonomy bundle through u_0, it consists of all $u \in OM$ such that $u \sim u_0$. Then, $OM = \sqcup_u P(u)$, disjoint union of sets of the form $P(u)$.

Denote by OM/H the modulus space of OM with respect to the equivalent relation induced by $H = \Phi(u_0)$. It can be identified with the associated bundle with fibre G/H and the equivalent relation: $(uh^{-1}, h\xi) \sim (u, \xi)$. Let $\Pi_1 : OM \to OM/H$ denote the natural projection. Let $H : M \times \mathbb{R}^n \to TOM$ the bundle map introduced in (2.1), defined by the basic horizontal vector fields obtained from an orthonormal basis of $\mathbb{R}^n$. In particular, $H(u,e) \in HT_uOM$. Recall that the solutions of the SDE $du_t = H(u_t) \circ db_t$ are Brownian motions where b_t is a BM on $\mathbb{R}^n$. Denote by $[u]$ the holonomy bundle contains u. Then by the definition, $[u_t]$ is constant in time.

Theorem 4.6. *Let M be a connected and compact Riemannian manifold with a Riemannian connection ∇. Let H_0 be a horizontal vector field and $Z_k = A_k^*$ be fundamental vertical vector fields. Consider*

$$\begin{aligned} du_t^\epsilon &= H(u_t^\epsilon) \circ db_t + H_0(u_t^\epsilon)dt + \sqrt{\epsilon}\sum_{k=1}^m Z_k(u_t^\epsilon) \circ dw_t^k + \epsilon Z_0(u_t^\epsilon)dt, \\ u_0^\epsilon &= u_0, \end{aligned} \tag{4.13}$$

Then $[u^\epsilon_{\frac{t}{\epsilon}}]$ converges in law to a Markov process. Moreover, the Markov process is identified in (4.14) below.

Proof. By the holonomy theorem of Ambrose-Singer [1] the Lie algebra of $\Phi(u_0)$ is a subspace of $\mathfrak{g}$ and is spanned by matrices of the form $\Omega_v(w_1, w_2)$ where w_1, w_2 are horizontal vectors at $T_{u_0}OM$ and $v \in P(u_0)$. If $u \sim v$ then $\Phi(u) = \Phi(v)$. Let $H = \Phi(u_0)$, a group of dimension n_0. We define a distribution S on OM: $S = \{T(P(u)) : u \in OM\}$. It is of constant rank, $n + n_0$. This distribution is differentiable and involutive and $P(u)$ is the maximal integral manifold of S through u. Note that the holonomy bundles are translations of each other: $P(u_0a) = P(u_0)a$, $a \in G$. If u is equivalent to v, the maximal integral manifolds through them are identical.

Let u_t be the solution starting from u_0 of the equation

$$du_t = H(u_t) \circ db_t + H_0(u_t)dt.$$

Then $u_t \sim u_0$. To see this let f be a BC^∞ function on $OM/\Phi(u_0)$ and denote by $\Pi_1 : OM \to OM/\Phi(u_0)$ the projection from an element u to its equivalent class $[u]$. Then

$$f([u_t]) = f([u_0]) + \int_0^t df\,(T\Pi_1(H(u_s)) \circ db_s + \int_0^t df\,(T\Pi_1 H_0(u_s))\,ds.$$

By the Reduction Theorem, page 83 of Kobayashi-Nomizu [21], each holonomy bundle $P(u)$ is a reduced bundle with structure group $\Phi(u)$ and the connection in OM is reducible to a connection in $P(u)$. Hence

$$T_u(P(u)) = HT_uOM \oplus VT_u(p(u)).$$

In particular we have $T\Pi_1(HTOM) = 0$ and $f([u_t]) = f([u_0])$.

We have shown that the solution to the horizontal SDE stays in $P(u_0)$ for all times. The horizontal SDE, restricted to the maximal integrable manifold $P(u_0)$, satisfies the Hörmander conditions and is ergodic with a unique invariant measure $\mu_{P(u_0)}$. Fix a point u_0 with $x_0 = \pi(u_0)$. Let ν be the Haar measure on $H = \Phi(u_0)$. Denote by ν_a the Haar measure on $\Phi(u_0 a)$. Note that if $v = u_0 a$ some $a \in G$, let $u \in \Phi(u_0)$ and a horizontal curve α with $\alpha_0 = u_0, \alpha_1 = u_0 g, g \in \Phi(u_0)$. Then $\beta = \alpha_0 a$ is horizontal with $\beta_0 = v$ and $\beta_1 = u_0 g a = v a^{-1} g a$. Consequently $\Phi(u_0 a) = \mathrm{ad}(a^{-1})\Phi(u_0)$. If $a \in H$, $\Phi(u_0 a) = \Phi(u_0)$ and $\nu_a = \nu$.

Denote by N_u the following fibre of the holonomy bundle $P(u_0)$:

$$N_u = \pi^{-1}(x) \cap P(u_0), \qquad x = \pi(u).$$

Locally $N_{u_0 a} = M \times \{\mathrm{ad}(a^{-1}\Phi(u_0)\}$ and $\mu_{P(u_0 a)} = dx \times d\nu_a$ where dx is the volume measure of the manifold M.

Let $F : OM \to \mathbb{R}$ be a BC^1 function, the integral

$$\tilde{F}[P(u)] := \int_{P(u)} F d\mu_{P(u)}$$

is defined to be a number depending on a transversal of $P(u)$. On each fibre of the holonomy bundle $P(u_0)$ we choose a reference element $v(x)$, which determines reference elements on holonomy bundles $P(ua)$, due to that $v(x)a$ is an element of $P(ua)$ where $u \in \pi^{-1}(x)$. For any $u \in P(u_0)$ there is $g \in H$ such that $u = v(\pi(u))g$. We define

$$\int_{P(u_0)} F d\mu_{P(u_0)} := \int_M \int_{P(u_0)\cap\pi^{-1}(x)} F(v(x)g) d\nu(g) dx.$$

The resulting number is independent of the choice of v. To see this let v' be another choice then $v' = vh$ some $h \in H$ and $u = va = v'h^{-1}a$. Since G is a compact group, the Haar measure is bi-invariant,

$$\begin{aligned}\int_{P(u_0)\cap\pi^{-1}(x)} F(v(x)g) d\nu_x(g) &= \int_{P(u_0)\cap\pi^{-1}(x)} F(v'(x)h^{-1}a) d\nu_x(g)\\ &= \int_{P(u_0)\cap\pi^{-1}(x)} F(v'(x)a') d\nu_x(g').\end{aligned}$$

Similarly if $u = v(x)ag \in P(u_0 a)$ the following integral is well defined:

$$\int_{P(u)} F d\mu_{P(u)} := \int_M \int_{P(u)\cap\pi^{-1}(x)} F(v(x)ag)d\nu_a(g)dx.$$

Evaluate $f : P/\Phi(u_0) \to \mathbb{R}$ at u_t^ϵ, to see that

$$f([u_t^\epsilon]) = f([u_0^\epsilon]) + \sqrt{\epsilon}\int_0^t df\left(T\Pi_1\left(Z_k(u_s^\epsilon)\right)\right) \circ dw_s^k + \epsilon \int_0^t df\left(T\Pi_1\left(Z_0(u_s^\epsilon)\right)\right) ds.$$

Let $\mathfrak{m}$ be the Lie algebra of H and let $A_i, i = 1, \dots, n_0$ be an o.n.b. of $\mathfrak{m}$. Let $B_j, j = n_0+1, \dots, N$ be an o.n.b. of the vertical part of the distribution S at u_0. Define $A_j = \varpi_{u_0}(B_j) \in \mathfrak{g}$. Consider the family of fundamental vertical vector fields $\{A_j^*(u), j > n_0\}$, restricted to $P(u_0)$. Then $T\Pi_1(A_i^*) = 0$ for $i \le n_0$ and for $j > n_0$, $T_u\Pi_1(A_j^*) = A_j^*([u])$.

Writing Z_k in terms of the basis $\{A_k\}$, $Z_k = \sum_j \sigma_k^j A_j^*$, we have

$$\begin{aligned} f([u_t^\epsilon]) =& f([u_0^\epsilon]) + \sqrt{\epsilon}\sum_k \sum_{j=n_0+1}^N \int_0^t \sigma_k^j(u_s^\epsilon)\, df\left(A_j^*([u_s^\epsilon])\right) \circ dw_s^k \\ &+ \epsilon \sum_{j=n_0+1}^N \int_0^t \sigma_0^j(u_s^\epsilon)\, df\left(A_j^*([u_s^\epsilon])\right) ds. \end{aligned}$$

The process $[u_t^\epsilon]$ is in general not Markov. It is however clear, following the standard method as used earlier, that the probability distributions $\{[u^\epsilon_{\frac{\cdot}{\epsilon}}], \epsilon > 0\}$ is tight and any sequence of $[u^\epsilon_{\frac{t}{\epsilon}}]$ has a convergent sub-sequence with the same limit. The limit can be identified below. Define $a_{i,j}([u]) = \sum_{k\ge1}\int \sigma_k^i \sigma_k^j d\mu_{P(u)}$, and $\bar{Z} = \sum_{j=n_0+1}^N \bar{\sigma}_0^j A_j^*$. For $i, j \ge n_0$ define

$$\bar{\sigma}_0^j([u]) = \int_{P(u)} \left(\sigma_0^j + \frac{1}{2}\sum_{k\ge1} d\sigma_k^j(Z_k)\right) d\mu_{P(u)}.$$

Then

$$\mathcal{L}f([u]) = \sum_{i,j=n_0+1}^N a_{i,j}([u])\nabla df\left(A_j^*, A_i^*\right) + df(Z([u])). \tag{4.14}$$

This concludes the proof as it only remains to prove the relatively compactness which is similar to that of Theorem 4.2. □

In summary, we have introduced several random perturbation models to stochastic dynamics on manifolds which can be reduced to a system of slow/fast systems. Following these, we take the separation of scale constant to zero, and study whether the slow motions can be approximated by effective dynamics. In addition we identified tools in differential geometry useful for obtaining quantitative estimates and limit theorems on manifolds.

Acknowledgements

It is a pleasure to thank M. Hairer, Y. Maeda, J. Norris, and S. Rosenberg for helpful discussions. I would like to thank K. D. Elworthy and Y. LeJan for discussions and earlier collaborations from which I draw inspiration. I would like to thank the organisers of the XXV Brazilian School of Probability for the invitation.

References

[1] W. Ambrose and I. M. Singer. A theorem on holonomy. *Trans. Amer. Math. Soc.*, 75:428–443, 1953.

[2] V. I. Arnol′d. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989.

[3] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, 1968.

[4] A. N. Borodin and M. I. Freidlin. Fast oscillating random perturbations of dynamical systems with conservation laws. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(3):485–525, 1995.

[5] Pedro J. Catuogno, Fabiano B. da Silva, and Paulo R. Ruffino. Decomposition of stochastic flows in manifolds with complementary distributions. *Stoch. Dyn.*, 13(4):1350009, 12, 2013.

[6] R. M. Dowell. *Differentiable Approximations to Brownian Motion on Manifolds.* PhD thesis, University of Warwick, 1980.

[7] J. Eells and K. D. Elworthy. Stochastic dynamical systems. In *Control theory and topics in functional analysis (Internat. Sem., Internat. Centre Theoret. Phys., Trieste, 1974), Vol. III*, pages 179–185. Internat. Atomic Energy Agency, Vienna, 1976.

[8] K. D. Elworthy, Y. Le Jan, and Xue-Mei Li. *On the geometry of diffusion operators and stochastic flows*, volume 1720 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999.

[9] K. D. Elworthy, Yves Le Jan, and Xue-Mei Li. Concerning the geometry of stochastic differential equations and stochastic flows. In *New trends in stochastic analysis (Charingworth, 1994)*, pages 107–131. World Sci. Publ., River Edge, NJ, 1997.

[10] K. David Elworthy, Yves Le Jan, and Xue-Mei Li. Equivariant diffusions on principal bundles. In *Stochastic analysis and related topics in Kyoto*, volume 41 of *Adv. Stud. Pure Math.*, pages 31–47. Math. Soc. Japan, Tokyo, 2004.

[11] K. David Elworthy, Yves Le Jan, and Xue-Mei Li. *The geometry of filtering*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2010.

[12] N. Enriquez, J. Franchi, and Y. Le Jan. Central limit theorem for the geodesic flow associated with a Kleinian group, case $\delta > d/2$. *J. Math. Pures Appl. (9)*, 80(2):153–175, 2001.

[13] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[14] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260. Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.

[15] Inge S. Helland. Central limit theorems for martingales with discrete or continuous time. *Scand. J. Statist.*, 9(2):79–94, 1982.

[16] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1989.

[17] Paulo R. Ruffino Ivan I. Gonzales Gargate. An averaging principle for diffusions in foliated spaces. http://arxiv.org/abs/1212.1587, 2012.

[18] R. Z. Khas'minskii. The behavior of a self-oscillating system acted upon by slight noise. *J. Appl. Math. Mech.*, 27:1035–1044, 1963.

[19] Yuri Kifer. Averaging in dynamical systems and large deviations. *Invent. Math.*, 110(2):337–370, 1992.

[20] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.

[21] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol I.* Interscience Publishers, a division of John Wiley & Sons, New York-Lond on, 1963.

[22] Xue-Mei Li. An averaging principle for a completely integrable stochastic Hamiltonian system. *Nonlinearity*, 21(4):803–822, 2008.

[23] Xue-Mei Li. Collapsing Brownian motions on the Hopf fibration. In Preparation, 2012.

[24] Xue-Mei Li. Effective diffusions with intertwined structures. arXiv:1204.3250, 2012.

[25] Xue-Mei Li. Limits of random differential equations on manifolds. *Probab. Theory Related Fields*, 166(3-4):659–712, 2016.

[26] Xue-Mei Li. Random perturbation to the geodesic equation. *Ann. Probab.*, 44(1):544–566, 2016.

[27] Xue-Mei Li. Homogenisation on homogeneous spaces. *J. Math. Soc. Japan*, 70(2):519–572, 2018. With an appendix by Dmitriy Rumynin.

[28] Xue-Mei Li. Perturbation of conservation laws and averaging on manifolds. In *Computation and combinatorics in dynamics, stochastics and control*, volume 13 of *Abel Symp.*, pages 499–550. Springer, Cham, 2018.

[29] Ming Liao. The Brownian motion and the canonical stochastic flow on a symmetric space. *Trans. Amer. Math. Soc.*, 341(1):253–274, 1994.

[30] Ming Liao. Decomposition of stochastic flows and Lyapunov exponents. *Probab. Theory Related Fields*, 117(4):589–607, 2000.

[31] Ming Liao and Longmin Wang. Motion of a rigid body under random perturbation. *Electron. Comm. Probab.*, 10:235–243 (electronic), 2005.

[32] Shojiro Manabe and Yoko Ochi. The central limit theorem for current-valued processes induced by geodesic flows. *Osaka J. Math.*, 26(1):191–205, 1989.

[33] John Milnor. Curvatures of left invariant metrics on Lie groups. *Advances in Math.*, 21(3):293–329, 1976.

[34] Edward Nelson. *Dynamical theories of Brownian motion.* Princeton University Press, Princeton, N.J., 1967.

[35] Stuart Patching. Analysis of stochastic fast-slow systems. M.Sc. dissertation, Imperial College London, 2018.

[36] Grigorios A. Pavliotis and Andrew M. Stuart. *Multiscale methods*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.

[37] M. Ratner. The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. *Israel J. Math.*, 16:181–197, 1973.

[38] Ja. G. Sinai. The central limit theorem for geodesic flows on manifolds of constant negative curvature. *Soviet Math. Dokl.*, 1:983–987, 1960.

[39] Richard B. Sowers. Averaging of stochastic flows: twist maps and escape from resonance. *Stochastic Process. Appl.*, 119(10):3549–3582, 2009.

[40] H. K. Urbantke. The Hopf fibration—seven times in physics. *J. Geom. Phys.*, 46(2):125–150, 2003.

[41] Timothy C. Wallstrom. The stochastic mechanics of the Pauli equation. *Trans. Amer. Math. Soc.*, 318(2):749–762, 1990.