UNIVERSITY OF CALIFORNIA, SAN DIEGO

Hypoelliptic heat kernel inequalities on Lie groups

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Tai Alexis Melcher

Committee in charge:

Professor Bruce Driver, Chair
Professor Peter Ebenfelt
Professor Patrick Fitzsimmons
Professor George Fuller
Professor Costas Pozrikidis

2004

The dissertation of Tai Alexis Melcher is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2004

To my parents,
Michael and Sue May.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I thank my advisor Bruce Driver for patience, kindness, and the incredible gift of his knowledge. He has taught me how to be a mathematician. I thank Mark Lynch for suggesting that I become one in the first place. And I thank the members of my thesis committee, Peter Ebenfelt, Patrick Fitzsimmons, George Fuller, and Costas Pozrikidis.

The love and support of my parents, Michael and Sue May; of my sister and brothers, Li, Glenn, Robert, Danny, and Philip; and of their families, have carried me this far and will surely take me over many other obstacles. I hope I have made them proud. I thank Christian Gromoll for his strength and commitment and love, even when so far away.

I give my thanks to Imre Tuba, Ben Raphael, Alison Raphael, Arif Dowla, Rino Sanchez, Graham Hazel, Tim McMurry, Josh Griffin, Jason Bell, Brian Kurkoski, Joseph Soriaga, Jon Callihan, and Karen Warner – the friends who have been there for my graduate school years. And to Katy Stowe and Will and Angela Duke – the friends who have been there for everything.

To the people I have named here and to the people I have not – the usual suspects for Sunday night dinner, the weekend ultimate players, Doug Trenton and the staff at RT's Longboard, and many others – I could not have done this without you.

VITA

November 3, 1977	Born, Portsmouth, Virginia.
1999	B. S., Tufts University.
2001	M. A., University of California, San Diego.
2004	Ph. D., University of California, San Diego.

PUBLICATIONS

Hypoelliptic heat kernel inequalities on the Heisenberg group (with B. K. Driver), 2003. Submitted.

ABSTRACT OF THE DISSERTATION

Hypoelliptic heat kernel inequalities on Lie groups

by

Tai Alexis Melcher
Doctor of Philosophy in Mathematics

University of California San Diego, 2004

Professor Bruce Driver, Chair

We study the existence of gradient estimates for second order hypoelliptic heat kernels on manifolds. It is now standard that such inequalities, in the elliptic case, are equivalent to a lower bound on the Ricci tensor of the Riemannian metric. For hypoelliptic operators, the associated "Ricci curvature" takes on the value $-\infty$ at points of degeneration of the semi-Riemannian metric associated to the operator. For this reason, many of the standard proofs for the elliptic theory no longer work in the hypoelliptic setting.

This thesis gives recent results for hypoelliptic operators. Using Malliavin calculus methods, we transfer the problem to one of determining certain infinite dimensional estimates. We study the case where the underlying manifold is a Lie group and the hypoelliptic operators are invariant under left translations. In particular, we are able to show that "L^p-type" gradient estimates hold for $p \in (1, \infty)$, and the $p = 2$ gradient estimate implies a Poincaré estimate in this context. The case $p = 1$ (which would imply a logarithmic Sobolev inequality) is still under investigation; however, in the special case of the Heisenberg Lie group, we are able to determine several large classes of functions for which the inequality holds.

Chapter 1

Introduction

1.1 Background

Let M^d be a manifold of dimension d. A second order differential operator L is *subelliptic* if, given a chart x on M, one may write

$$L = a^{ij}(x)\partial_{x^i}\partial_{x^j} + b^i(x)\partial_{x^i},$$

where $\partial_{x^i} = \frac{\partial}{\partial x^i}$ and (a^{ij}) is a real symmetric non-negative matrix. We observe here the summation convention of summing over repeated upper and lower indices. In the nondegenerate case, that is, $(a^{ij}) > 0$, L is said to be *elliptic*.

On a Riemannian manifold M with metric g, a standard example of a subelliptic operator is the Laplace-Beltrami operator Δ_g. Given a chart x on M, Δ_g may be written as

$$L = \Delta_g = \frac{1}{\sqrt{g}}\partial_{x^i}\left(\sqrt{g}g^{ij}\partial_{x^j}\right),$$

where $g = g_{ij}dx^i dx^j = g(\partial_{x^i}, \partial_{x^j})dx^i dx^j$, $\sqrt{g} = \sqrt{\det(g_{ij})}$, and (g^{ij}) is the inverse matrix to the matrix (g_{ij}). It is well known that Δ_g is in fact an elliptic operator. On a general manifold M, given a set of smooth vector fields $\{X_1, \ldots, X_k\}$ on M, the operator

$$L = \sum_{i=1}^{k} X_i^2.$$

is also subelliptic, which may be seen by writing the vector fields in local coordinates, $X_i = X_i^j \partial_{x^j}$.

Notation 1.1. Let $C_c^\infty(M)$ denote the set of smooth functions on M with compact support. When $M = \mathbb{R}^n$, let $C_p^\infty(\mathbb{R}^n)$ denote those functions $f \in C^\infty(\mathbb{R}^n)$ such that f and all of its partial derivatives have at most polynomial growth.

For any subelliptic operator L, we may define

$$\Gamma_1(f,h) := \frac{1}{2}[L(fh) - fLh - hLf],$$

for $f, h \in C^\infty(M)$. By iterating this construction, we also have

$$\Gamma_2(f,h) := \frac{1}{2}[L\Gamma_1(f,h) - \Gamma_1(Lf,h) - \Gamma_1(f,Lh)].$$

If M is a Riemannian manifold and $\nabla := (X_1, \ldots, X_k)$, for $\{X_i\}_{i=1}^k$ a collection of smooth vector fields on M, then

$$\Gamma_1(f,h) = (\nabla f, \nabla h).$$

In particular, along the diagonal $f = h$, we write

$$\Gamma_1(f) := \Gamma_1(f,f) = \frac{1}{2}Lf^2 - fLf = |\nabla f|^2. \tag{1.1}$$

Similarly,

$$\Gamma_2(f) := \Gamma_2(f,f) = \frac{1}{2}L\Gamma_1(f) - \Gamma_1(f,Lf) = \frac{1}{2}L|\nabla f|^2 - (\nabla f, \nabla Lf).$$

When L is the Laplace-Beltrami operator Δ_g,

$$\Gamma_2(f) = \|\nabla^2 f\|^2 + (\mathrm{Ric}\nabla f, \nabla f), \tag{1.2}$$

where ∇^2 is the Hessian tensor, Ric is the Ricci tensor, and $\|\cdot\|$ is the Hilbert-Schmidt norm in the Riemannian metric.

Over the last twenty years or more, a fairly complete and very beautiful theory has been developed applying to elliptic operators on Riemannian manifolds. This theory relates properties of the solutions of elliptic and parabolic equations to properties of the Riemannian geometry. These geometric properties are determined by the principal symbol of the underlying elliptic operator. The following theorem is a typical example of the type of result we have in mind here; see for example [4, 5, 6]. In particular, [35] gives a nice survey of the progress made on the geometry of diffusion generators. We include the proof of this theorem in the appendix to the thesis.

Theorem 1.2. *Suppose (M,g) is a complete Riemannian manifold, and ∇ and Δ are the gradient and Laplace-Beltrami operators acting on $C^\infty(M)$. We write $e^{t\Delta/2}$ to denote the heat semigroup $e^{t\bar{\Delta}/2}$, where $\bar{\Delta}$ is the self-adjoint extension of $\Delta|_{C_c^\infty(M)}$ to $L^2(M,dV)$ with dV volume measure on M. Let $|v| := \sqrt{g(v,v)}$ for all $v \in TM$, Ric denote the Ricci curvature tensor, and k be a constant. Then the following are equivalent:*

1. $\mathrm{Ric}(\nabla f, \nabla f) \geq -2k|\nabla f|^2$, *for all* $f \in C_c^\infty(M)$,
2. $\Gamma_2(f) \geq -2k\Gamma_1(f)$, *for all* $f \in C_c^\infty(M)$,
3. $|\nabla e^{t\Delta/2} f| \leq e^{kt} e^{t\Delta/2} |\nabla f|$, *for all* $f \in C_c^\infty(M)$ *and* $t > 0$,
4. $|\nabla e^{t\Delta/2} f|^2 \leq e^{2kt} e^{t\Delta/2} |\nabla f|^2$, *for all* $f \in C_c^\infty(M)$ *and* $t > 0$,
5. *there is a function* $K(t) > 0$ *such that* $K(0) = 1$, $\dot{K}(0)$ *exists, and*

$$|\nabla e^{t\Delta/2} f|^2 \leq K(t) e^{t\Delta/2} |\nabla f|^2, \tag{1.3}$$

for all $f \in C_c^\infty(M)$ *and* $t > 0$.

Remark 1.3. This theorem leaves open the possibility that a function $K(t) > 0$ exists such that the inequality (1.3) is satisfied for all $f \in C_c^\infty(M)$ and $t > 0$, but $K(0) > 1$, or $\dot{K}(0) = \infty$ or does not exist, while at the same time there is no constant k such that $\mathrm{Ric}(\nabla f, \nabla f) \geq -2k|\nabla f|^2$, for all $f \in C_c^\infty(M)$.

Note that the estimate in (3) implies that

$$|\nabla e^{t\Delta/2} f|^p \leq e^{pkt} e^{t\Delta/2} |\nabla f|^p, \quad p \in [1,\infty), \tag{1.4}$$

for all $f \in C_c^\infty(M)$ and $t > 0$, by Hölder's inequality. Estimates like (1)–(5) are also equivalent to one parameter families of Poincaré and log Sobolev estimates for the heat kernel. The latter has implications for hypercontractivity of an associated semigroup; see Gross [23, 24].

As a simple illustration of this theorem, consider the manifold $M = \mathbb{R}^d$ with the usual vector fields $\partial_{x^1}, \ldots, \partial_{x^d}$. Let ∇ and Δ be the standard gradient and Laplacian,

$$\nabla = (\partial_{x^1}, \ldots, \partial_{x^d}) \text{ and } \Delta = \partial_{x^1}^2 + \ldots + \partial_{x^d}^2.$$

On $\mathbb{R}^d$, the Ricci curvature is 0, and the standard Laplacian is the Laplace-Beltrami operator. Also, $\Gamma_2(f) = \|\nabla^2 f\|^2 \geq 0$. In this case $e^{t\Delta/2}$ is convolution by the well known probability density

$$p_t\left(x\right) := \frac{1}{\left(2\pi t\right)^{d/2}} e^{-\frac{1}{2t}|x|^2_{\mathbb{R}^d}},$$

and

$$\nabla e^{t\Delta/2} f = e^{t\Delta/2} \nabla f, \tag{1.5}$$

for all $f \in C_c^1(\mathbb{R}^d)$, as follows from basic properties of convolutions; more abstractly, this follows from the commutativity of the Euclidean gradient and Laplacian. Equation (1.5) and an application of Hölder's inequality then imply that

$$\left|\nabla e^{t\Delta/2} f\right|^p = \left[e^{t\Delta/2}\left|\nabla f\right|\right]^p \leq e^{t\Delta/2} |\nabla f|^p,$$

for all $f \in C_c^1(\mathbb{R}^d)$, where $|\nabla f| := \left\{\sum_{i=1}^d (\partial_{x^i} f)^2\right\}^{1/2}$.

This research program moves toward extending Theorem 1.2 to operators of the form

$$L = \sum_{i=1}^k X_i^2, \tag{1.6}$$

where $\{X_i\}_{i=1}^k$ is a set of smooth vector fields on M satisfying

$$T_m M = \operatorname{span}\left(\{X(m) : X \in \mathcal{L}\}\right), \quad \forall\, m \in M, \tag{HC}$$

where $\mathcal{L}$ is the Lie algebra of vector fields generated by the collection $\{X_i\}_{i=1}^k$. This assumption is the *Hörmander bracket condition*, and we say that $\{X_i\}_{i=1}^k$ is a *Hörmander set*. Under this assumption, by a celebrated theorem of Hörmander, the operator L given in Equation (1.6) is hypoelliptic. Recall that a subelliptic operator L is said to be *hypoelliptic* if $Lu \in C^\infty(\Omega)$ implies that $u \in C^\infty(\Omega)$, for all distributions $u \in C^\infty(\Omega)'$ on any open set $\Omega \subset_o M$.

However, L need not be elliptic. The principle symbol of L at $\xi \in T_m^*M$ is given by $\sigma_L(\xi) = \sum_{i=1}^k [\xi(X_i)]^2$. By definition, the operator L is degenerate at points $m \in M$ where there exists $0 \neq \xi \in T_m^*M$ such that $\sigma_L(\xi) = 0$. At points of degeneracy of L, the Ricci tensor is not well defined and should be interpreted to take the value $-\infty$ in some directions. Thus, there exist no lower bounds in the Ricci curvature in this case. Γ_1 and Γ_2 are well defined for any subelliptic operator L; however, it can be shown that inequalities analogous to that in item (2) of Theorem 1.2 also fail in the non-elliptic case. Hence it is not possible to directly generalize Theorem 1.2 in this setting. Nevertheless it is reasonable to ask if inequalities of the form (1.4) might still hold. More precisely, we let $\nabla = (X_1, \ldots, X_k)$ and address the following question: do functions $K_p(t) < \infty$ exist such that

$$|\nabla e^{tL/2} f|^p \leq K_p(t) e^{tL/2} |\nabla f|^p, \quad p \in [1, \infty),$$

for all $f \in C_c^\infty(M)$ and $t > 0$?

Related results appear in Kusuoka and Stroock [34]. In particular, a special case of Theorem 2.18 of that paper implies that, for all $p \in (1, \infty)$, there exist finite constants C_p such that

$$|\nabla e^{tL/2} f|^p \leq C_p t^{-p/2} e^{tL/2} |f|^p,$$

for all smooth, bounded functions f with bounded derivatives of all orders and all $t > 0$. A similar result may be found in Picard [42].

Additionally, Auscher, Coulhon, Duong, and Hofmann [3] show that on a non-compact connected Riemannian manifold satisfying certain spectral gap conditions, the

Laplace-Beltrami operator satisfies

$$|\nabla e^{t\Delta} f|^2 \leq C e^{-C' t\Delta} |\nabla f|^2$$

for some $C, C' > 0$, for all $f \in C_c^\infty(M)$ such that $f, \nabla f \in L^2(M, dV)$, and $t > 0$. Their methods may be adapted for the sum of squares of vectors fields satisfying the Hörmander bracket condition (HC). In the case when the manifold is a Lie group, this is carried out in Alexopoulos [2]. The papers [3, 11] also include some potential nice applications of our own result.

1.2 Statement of results

Let G be a d-dimensional Lie group with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$ and identity element e. Suppose $\{X_i\}_{i=1}^k \subset \mathfrak{g}$ is a Lie generating set; that is, there exists some $m \in \mathbb{N}$ such that

$$\operatorname{span}\big\{X_i, [X_{i_1}, X_{i_2}], [X_{i_1}, [X_{i_2}, X_{i_3}]], \ldots, [X_{i_1}, [\cdots, [X_{i_{m-1}}, X_{i_m}] \cdots]] : \\ i, i_r \in \{1, \ldots, k\}, r \in \{1, \ldots, m\}\big\} = \mathfrak{g}. \quad (1.7)$$

Notation 1.4. Let $\Sigma = \Sigma_0 := \{X_1, \ldots, X_k\}$ and Σ_r be defined inductively by

$$\Sigma_r := \{[X_i, V] : V \in \Sigma_{r-1}, i = 1, \ldots, k\},$$

for all $r \in \mathbb{N}$. Since $\{X_i\}_{i=1}^k$ is a Lie generating set, there exists a finite m such that

$$\operatorname{span}\left(\cup_{r=0}^m \Sigma_r\right) = \mathfrak{g}.$$

Let $\mathfrak{g}_0 := \operatorname{span}(\Sigma_0)$, and let $\{Y_j\}_{j=1}^{d-k} \subset \cup_{r=1}^m \Sigma_r$ be a basis of $\mathfrak{g}/\mathfrak{g}_0$. Define an inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$ by making $\{X_i\}_{i=1}^k \cup \{Y_j\}_{j=1}^{d-k}$ an orthonormal set. Note then that $\{X_i\}_{i=1}^k$ is an orthonormal basis of $\mathfrak{g}_0$. We may extend $\langle \cdot, \cdot \rangle$ to a right invariant metric on G by defining $\langle \cdot, \cdot \rangle_g : T_g G \times T_g G \to \mathbb{R}$ as

$$\langle v, w \rangle_g := \left\langle R_{g^{-1}*} v, R_{g^{-1}*} w \right\rangle, \qquad \text{for all } v, w \in T_g G.$$

The g subscript will be suppressed when there is no chance of confusion.

Notation 1.5. Let L_g denote left translation by an element $g \in G$, and let R_g denote right translation. Given an element $X \in \mathfrak{g}$, let $\tilde{X}$ denote the left invariant vector field on G such that $\tilde{X}(e) = X$, where e is the identity of G. Recall that $\tilde{X}$ being left invariant means that the vector field commutes with left translation in the following way:

$$\tilde{X}(f \circ L_g) = (\tilde{X} f) \circ L_g,$$

for all $f \in C^1(G)$. Similarly, let $\hat{X}$ denote the right invariant vector field associated to X.

Definition 1.6. The left invariant *gradient* on G is the operator on $C^1(G)$ defined by

$$\nabla := (\tilde{X}_1, \ldots, \tilde{X}_k).$$

The *subLaplacian* on G is the second-order operator acting on $C^2(G)$

$$L := \sum_{i=1}^{k} \tilde{X}_i^2.$$

Remark 1.7. Since $\{X_i\}_{i=1}^k$ is a Lie generating set, $\{\tilde{X}_i\}_{i=1}^k$ satisfies the Hörmander condition (HC) and Hörmander's theorem [26] implies that L is a hypoelliptic operator.

Let $L^2(G)$ denote the space of twice integrable functions on G with respect to right invariant Haar measure. Note that for any $X \in \mathfrak{g}$, $\tilde{X}$ is formally skew-adjoint,

$$\begin{aligned}
(\tilde{X} f_1, f_2)_{L^2(G)} &= \int_0^t \tilde{X} f_1(g) f_2(g)\, dg = \int_0^t \frac{d}{d\epsilon}\bigg|_0 f_1(g e^{\epsilon X}) f_2(g)\, dg \\
&= \frac{d}{d\epsilon}\bigg|_0 \int_0^t f_1(g e^{\epsilon X}) f_2(g)\, dg \\
&= \frac{d}{d\epsilon}\bigg|_0 \int_0^t f_1(g) f_2(g e^{-\epsilon X})\, dg \\
&= \int_0^t \frac{d}{d\epsilon}\bigg|_0 f_1(g) f_2(g e^{-\epsilon X})\, dg = -(f_1, \tilde{X} f_2)_{L^2(G)},
\end{aligned}$$

for all $f_1, f_2 \in C_c^\infty(G)$. Thus,

$$(f_1, L f_2)_{L^2(G)} = (L f_1, f_2)_{L^2(G)},$$

for all $f_1, f_2 \in C_c^\infty(G)$, and L is a densely defined, symmetric operator on $L^2(G, dg)$. So we may associate to L the symmetric bilinear form $\mathcal{E}^0(f_1, f_2) := (-Lf_1, f_2)_{L^2(G)}$. Note that $\mathcal{E}^0$ is positive, since

$$(-Lf, f)_{L^2(G)} = \sum_{i=1}^k (\tilde{X}_i f, \tilde{X}_i f)_{L^2(G)} = \|\nabla f\|^2_{L^2(G)} \geq 0,$$

for all $f \in C_c^\infty(G)$. Thus, $\mathcal{E}^0$ is closable, and its minimal closure $\mathcal{E}$ is associated to a self-adjoint operator $\bar{L}$ which is an extension of L, called the Friedrichs extension of L. Now we may define the following.

Definition 1.8. Let P_t denote the *heat semigroup* $e^{t\bar{L}/2}$, where $\bar{L}$ is the Friedrichs extension of $L|_{C_c^\infty(G)}$ to $L^2(G, dg)$ with dg right Haar measure on G. By the left invariance of L and the satisfaction of the Hörmander condition, P_t admits a left convolution kernel p_t such that

$$P_t f(g) = f * p_t(g) = \int_G f(gh) p_t(h)\, dh, \tag{1.8}$$

for all $f \in C_c^\infty(G)$, where gh is defined by the group operation of G and dh again denotes right Haar measure on G. We will call p_t the *heat kernel* of G.

P_t is a symmetric Markov semigroup. By Remark 1.7, L is a hypoelliptic operator, and so p_t is a smooth density on G. In the sequel, we will let L denote its own Friedrichs extension. For the semigroup theory used here, see [13, 20, 41, 55].

Notation 1.9. Let $K_p(t)$ be the best function such that

$$|\nabla P_t f|^p \leq K_p(t) P_t |\nabla f|^p, \quad p \in [1, \infty), \tag{I_p}$$

for all $f \in C_c^\infty(G)$ and $t > 0$.

Note that in this context the inequality (I_p) could be equivalently written as

$$\Gamma_1(P_t f)^{p/2} \leq K_p(t) P_t \Gamma_1(f)^{p/2}, \quad p \in [1, \infty),$$

for all $f \in C_c^\infty(G)$ and $t > 0$.

Theorem 1.10. *Let G be a Lie group. Then for all $p \in (1, \infty)$, $K_p(t) < \infty$ for all $t > 0$. In particular, if G is a nilpotent Lie group, then there exists a constant $K_p < \infty$ such that $K_p(t) < K_p$ for all $t > 0$.*

This statement is verified in Chapter 2 in the context of the real 3-dimensional Heisenberg Lie group, before being treated in full generality in Chapter 4. We approach this case first because the basic idea of the proof can be seen without the added geometric complications appearing in the more general formulation. So for now taking G to be the Heisenberg group, one could attempt to determine finite constants satisfying (I_p) on G by proving bounds for the integral representation of P_t, Equation (1.8). In particular, there is an explicit integral formula for $p_t(g)$ on the Heisenberg group, given in Equation (2.5). However, our attempts at proving such bounds using this formula have been unsuccessful.

The Heisenberg heat kernel $p_t(g)\,dg$ may also be realized as the distribution in t of the Cartan rolling map on G, the process ξ satisfying Equation (2.20), the Stratonovich differential equation

$$\sum_{i=1}^{k} \tilde{X}_i(\xi_t) \circ db_t^i, \text{ with } \xi_0 = e,$$

where $b^1, \ldots, b^k$ are k independent real-value Brownian motions. Thus, applying the heat semigroup P_t to a function $f \in C_c^\infty(G)$ at the identity is equivalent to taking the expectation of $f(\xi_t)$; that is, $P_t f(e) = \mathbb{E}\{f(\xi_t)\}$. Using this representation of P_t, we may transform our finite dimensional problem to a problem on Wiener space. In a "lifting" procedure described in Section 2.7, we are able to construct vector fields $\mathbf{X}_i$ on Wiener space from the Heisenberg vector fields $\tilde{X}_i$ and the map ξ. Then we can apply Malliavin's probabilistic techniques on proving hypoellipticity to determine bounds of expressions like $\mathbb{E}\{\mathbf{X}_i[f(\xi_t)]\}$. Section 2.5 reviews some infinite dimensional calculus on Wiener space necessary for this argument.

Before proceeding with this proof, in Chapter 2 we show that the left invariance of the vector fields leaves the inequality (I_p) invariant under group translation, a result which also holds in the general Lie group context. The Heisenberg group admits a family of dilations adapted to its group structure. These dilations allow a scaling argument which additionally proves that in this case the functions $K_p(t) = K_p$ for all $t > 0$, for some constant K_p. We prove that $K_p \geq \sqrt{2}$ when $1 \leq p \leq 2$ and, in general, that $K_p > 1$. Note that at $t = 0$ the inequality is an empty statement and certainly holds for constant 1. Unlike the elliptic case where the inequality holds with a continuous coefficient (see Equation (1.4)), there is now a jump discontinuity in $K_p(t)$ at $t = 0$. This discontinuity at $t = 0$ should be a feature which persists in the general hypoelliptic setting, although that is not verified in this thesis. Sections 2.6 to 2.8 give the proof of Theorem 1.10 in the Heisenberg group, following the outline described above. Finally, Section 2.9.1 shows that our method can not, without modification, be used to prove $K_1 < \infty$. However, Section 2.9.2 gives hope, giving several large classes of functions on the Heisenberg group for which the estimate holds with $p = 1$.

Using the notation for calculus on Wiener space introduced in Section 2.5, Chapter 3 begins the generalization of the argument made in the Heisenberg case to a general Lie group G. Section 3.1.1 introduces the rolling map on G, the solution ξ to the Stratonovich equation (3.3) analogous to the Heisenberg equation (2.20). Section 3.2 provides the additional proofs of existence and differentiability on the path space over a group, not covered in the standard literature (which primarily addresses the case of Euclidean space or compact manifolds). In particular, a large part of this chapter is devoted to proving Theorem 3.2, which states that, for the solution ξ_t to Equation (3.3) at fixed time $t > 0$, $f(\xi_t)$ is smooth in some sense for reasonable functions f and has derivatives in all L^p with respect to Wiener measure. Section 3.1.3 generalizes the procedure of "lifting" vector fields given in Section 2.7 and, together with Section 3.1.2, gives the results necessary for bounding expectations of the lifted vector fields.

Chapter 4 contains the proof of Theorem 1.10 in the most generality we have been able to prove it in. In Section 4.1, using the results from Chapter 3, we show that for a Lie group G, $K_p(t) < \infty$ for all $t > 0$, although we are not able to estimate the behavior of K_p with respect to t. In a generalization of the Heisenberg scaling argument in Section 2.3, Section 4.2 addresses the special case of nilpotent and stratified groups. When G is stratified, dilation arguments imply that the K_p are independent of the t parameter. When G is nilpotent, we are able to prove that there is a constant K_p such that $K_p(t) < K_p$ for all $t > 0$, completing the proof of Theorem 1.10. This also allows us to prove the following Poincaré estimate for the heat kernel measure in this context.

Theorem 1.11. *Suppose G is a nilpotent Lie group with identity element e. Then*

$$P_t f^2(e) - (P_t f)^2(e) \leq K_2 t P_t |\nabla f|^2(e),$$

for all $f \in C_c^\infty(G)$ and $t > 0$, where K_2 is the constant in Theorem 1.10 for $p = 2$.

Note that this theorem gives a slight improvement in the elliptic case, where the estimate is known only to hold with coefficients of exponential growth.

Chapter 2

A special case: the Heisenberg group

Let $M = G$ be $\mathbb{R}^3$ equipped with the Heisenberg group multiplication given in Equation (2.2). In this setting, we will take $L = \tilde{X}^2 + \tilde{Y}^2$, where $\tilde{X}$ and $\tilde{Y}$ are the vector fields

$$\tilde{X} := \partial_x - \frac{1}{2} y \partial_z \text{ and } \tilde{Y} := \partial_y + \frac{1}{2} x \partial_z. \tag{2.1}$$

2.1 Realization of the Heisenberg Lie group

Recall that the real Heisenberg Lie algebra is $\mathfrak{g} = \operatorname{span}\{X, Y, Z\}$ where $Z = [X, Y]$ and Z is in the center of $\mathfrak{g}$. Thus, $\mathfrak{g}_0 := \operatorname{span}\{X, Y\}$ is a hypoelliptic subspace of $\mathfrak{g}$; that is, the Lie algebra generated by $\mathfrak{g}_0$ is $\mathfrak{g}$. The Heisenberg group G is the simply connected real Lie group such that $\operatorname{Lie}(G) = \mathfrak{g}$. Letting $A = aX + bY + cZ$ and $A' = a'X + b'Y + c'Z$, we have by the Baker-Campbell-Hausdorff formula that

$$e^A e^{A'} = e^{A + A' + \frac{1}{2}[A, A']}.$$

Thus we may realize G as $\mathbb{R}^3$ with the following group multiplication

$$(a, b, c)(a', b', c') = (a + a', b + b', c + c' + \frac{1}{2}(ab' - a'b)). \tag{2.2}$$

2.2 Differential operators on G

Let $X = (1,0,0)$, $Y = (0,1,0)$, and $Z = (0,0,1)$ at the identity $0 \in G$. We extend these to left invariant vector fields on G in the standard way. For $g = (a,b,c) \in G$, let L_g denote left translation by g, and compute as follows:

$$\begin{aligned}\tilde{X}(a,b,c) = L_{(a,b,c)*}X &= \frac{d}{dt}\bigg|_0 (a,b,c)(t,0,0)\\ &= \frac{d}{dt}\bigg|_0 (a+t, b, c - \frac{1}{2}bt) = (1, 0, -\frac{1}{2}b).\end{aligned}$$

So if (x,y,z) are the standard coordinates on $G = \mathbb{R}^3$, for $f \in C^1(G)$,

$$(\tilde{X}f)(g) = \frac{d}{dt}\bigg|_0 f(g \cdot tX) = \frac{\partial f}{\partial x}(g) - \frac{1}{2}y\frac{\partial f}{\partial z}(g).$$

Performing similar computations for Y and Z, we then have

$$\tilde{X} = \partial_x - \frac{1}{2}y\partial_z,\ \tilde{Y} = \partial_y + \frac{1}{2}x\partial_z,\ \text{and } [\tilde{X},\tilde{Y}] = \tilde{Z} = \partial_z, \tag{2.3}$$

where $\partial_x = \frac{\partial}{\partial x}$, $\partial_y = \frac{\partial}{\partial y}$, and $\partial_z = \frac{\partial}{\partial z}$; compare with Equation (2.1). Note then that $\{\tilde{X},\tilde{Y},\tilde{Z}\}$ forms a basis for the tangent space at every point of G. This combined with $[\tilde{X},\tilde{Y}] = \tilde{Z}$ implies that $\{\tilde{X},\tilde{Y}\}$ satisfies the Hörmander bracket condition (HC). In the same manner one shows the right invariant vector fields associated to X, Y, and Z are given by

$$\hat{X} = \partial_x + \frac{1}{2}y\partial_z, \hat{Y} = \partial_y - \frac{1}{2}x\partial_z,\ \text{and } [\hat{X},\hat{Y}] = \hat{Z} = -\partial_z. \tag{2.4}$$

Remark 2.1. The right invariant vector fields associated to X and Y may be expressed as the following linear combinations,

$$\hat{X} = \tilde{X} + y\tilde{Z} \text{ and } \hat{Y} = \tilde{Y} - x\tilde{Z}.$$

Definition 2.2. The left invariant *gradient* on $G = \mathbb{R}^3$ is the operator

$$\nabla = (\tilde{X},\tilde{Y}).$$

The *subLaplacian* is

$$L = \tilde{X}^2 + \tilde{Y}^2,$$

and we let $P_t = e^{tL/2}$ be the heat semigroup associated to L, as given in Definition 1.8. Note that here

$$L = \partial_x^2 + \partial_y^2 + (x\partial_y - y\partial_x)\partial_z + \frac{1}{4}(x^2+y^2)\partial_z^2,$$

and it is easy to see that L is a positive, symmetric operator on $C_c^\infty(G)$. Finally, $p_t(g) = P_t\delta_0(g) = e^{tL/2}\delta_0(g)$ denotes the fundamental solution associated to L, so that for $f \in C_p^\infty(G)$,

$$P_t f(g) = p_t * f\,(g) := \int_G f(gh)p_t(h)\,dh$$

where dh denotes right Haar measure and gh is computed relative to the Heisenberg group multiplication in Equation (2.2).

Remark 2.3. Since $\{X, Y\}$ is a Lie generating set, Hörmander's theorem [26] implies that L is a hypoelliptic operator. Also Malliavin's techniques show p_t is a smooth positive function on $\mathbb{R}^3$; see Section 2.6. In this simple setting, an explicit formula for $p_t(g)$ is

$$p_t(g) = \frac{1}{8\pi^2}\int_{\mathbb{R}} \frac{w}{\sinh\left(\frac{wt}{2}\right)} \exp\left(-\frac{1}{4}|\vec{x}|^2 w \coth\left(\frac{wt}{2}\right)\right) e^{iwz}\,dw, \tag{2.5}$$

where $g = (x, y, z) \in G$ and $\vec{x} = (x, y)$; see for example [53]. This formula for the heat kernel demonstrates the potential advantage of considering the infinite dimensional Wiener space representation of $p_t(g)\,dg$, since the Wiener space representation no longer involves the oscillatory integral that appears in (2.5).

We will need the following straightforward results.

Lemma 2.4. *By the left invariance of ∇ and P_t, the inequality (I_p) holds for all $g \in G$, $f \in C_p^\infty(G)$, and $t > 0$, if and only if,*

$$|\nabla P_t f|^p(0) \le K_p(t) P_t |\nabla f|^p(0), \tag{2.6}$$

for all $f \in C_p^\infty(G)$ and $t > 0$.

Proof. If (2.6) holds, then

$$\begin{aligned} |\nabla P_t f|^p(g) &= |(\nabla P_t f)\circ L_g|^p(0) = |\nabla(P_t f\circ L_g)|^p(0) \\ &= |\nabla(P_t(f\circ L_g))|^p(0) \le K_p P_t |\nabla(f\circ L_g)|^p(0) \\ &= K_p P_t |(\nabla f)\circ L_g|^p(0) = K_p P_t |\nabla f|^p \circ L_g(0) \\ &= K_p P_t |\nabla f|^p(g). \end{aligned}$$

The converse is trivial. □

Note that this result only depended on the left invariance of the ∇ and P_t, and not the specific Heisenberg group structure. Thus, this result holds in general and will be used later in the general Lie group context.

Lemma 2.5. *For $A \in \mathfrak{g}$,*

$$\tilde{A} P_t f(0) = P_t \hat{A} f(0),$$

for all $f \in C_p^\infty(G)$ and $t > 0$. More generally,

$$\hat{A} P_t f = P_t \hat{A} f,$$

from which the previous equation follows, since $\hat{A} = \tilde{A}$ at 0.

Proof. If we ignore domain issues, then $[\hat{A}, \tilde{B}] = 0$ for all $B \in \mathfrak{g}$, so that $[\hat{A}, L] = 0$, and thus we expect $[\hat{A}, e^{tL/2}] = 0$. Formally, we have

$$\begin{aligned} \tilde{A} P_t f(0) &= \frac{d}{d\epsilon}\bigg|_0 P_t f(e^{\epsilon A}) = \frac{d}{d\epsilon}\bigg|_0 \int_G f(e^{\epsilon A} g) p_t(g)\, dg \\ &= \int_G \frac{d}{d\epsilon}\bigg|_0 f(e^{\epsilon A} g) p_t(g)\, dg \\ &= \int_G \hat{A} f(g) p_t(g)\, dg = P_t \hat{A} f(0). \end{aligned}$$

To differentiate under the integral, we have used the translation invariance of Haar measure (which is Lebesgue measure on $\mathbb{R}^3$) and the heat kernel bound

$$p_t(g) \le C t^{-2} e^{-\rho^2(g)/Ct},$$

where $\rho(g) \ge C'(|x| + |y| + |z|^{1/2})$ is the Carnot-Carathéodory distance on G, and C and C' are some positive constants; see Theorem 5.4.3 in [51] and page 27 of [9]. □

2.3 Scaling on G

So let $r > 0$ and $g = (x, y, z)$, and define $\phi_r : G \to G$ by $\phi_r(x, y, z) = (rx, ry, r^2 z)$. Notice that

$$\begin{aligned}\phi_r((a, b, c) \cdot (x, y, z)) &= \phi_r((a + x, b + y, c + z + \frac{1}{2}(ay - xb)) \\ &= \phi_r((ra + rx, rb + ry, r^2 c + r^2 z + \frac{r^2}{2}(ay - xb)) \\ &= \phi_r(a, b, c)\phi_r(x, y, z).\end{aligned}$$

Thus ϕ_r is a group isomorphism on G. (In fact, ϕ_r is a dilation on G with generator W given by,

$$\begin{aligned}W(x, y, z) &= \frac{d}{dr}\bigg|_{r=1} \phi_r(x, y, z) = (x, y, 2z)_{(x,y,z)} \\ &= x\partial_x + y\partial_y + 2z\partial_z \\ &= x\left(\tilde{X} + \frac{1}{2}y\partial_z\right) + y\left(\tilde{Y} - \frac{1}{2}x\partial_z\right) + 2z\partial_z = x\tilde{X} + y\tilde{Y} + 2z\tilde{Z};\end{aligned}$$

see Definition 4.8.) Using $e^{t\tilde{X}}(g) = g \cdot (t, 0, 0)$ and

$$\phi_{r*}\tilde{X} \circ \phi_r^{-1}(g) = \frac{d}{dt}\bigg|_0 \phi_r(e^{t\tilde{X}}(\phi_r^{-1}(g))),$$

along with similar formulas involving $\tilde{Y}$, one shows

$$\phi_{r*}\tilde{X} \circ \phi_r^{-1} = r\tilde{X} \text{ and } \phi_{r*}\tilde{Y} \circ \phi_r^{-1} = r\tilde{Y}. \tag{2.7}$$

The equations in (2.7) are equivalent to

$$\tilde{X}(f \circ \phi_r) = r(\tilde{X}f) \circ \phi_r \text{ and } \tilde{Y}(f \circ \phi_r) = r(\tilde{Y}f) \circ \phi_r.$$

Therefore,

$$\begin{aligned}\nabla(f \circ \phi_r) &= r(\nabla f) \circ \phi_r \text{ and} \\ L(f \circ \phi_r) &= r^2(Lf) \circ \phi_r.\end{aligned} \tag{2.8}$$

Also, from Equation (2.5), for $g = (x, y, z)$,

$$\begin{aligned} p_{r^2 t}(g) &= \frac{1}{8\pi^2} \int_{\mathbb{R}} \frac{w}{\sinh\left(\frac{wr^2 t}{2}\right)} \exp\left(-\frac{1}{4}|\vec{x}|^2 w \coth\left(\frac{wr^2 t}{2}\right)\right) e^{iwz}\, dw \\ &= \frac{1}{8\pi^2} \int_{\mathbb{R}} \frac{w}{r^2 \sinh\left(\frac{wt}{2}\right)} \exp\left(-\frac{1}{4r^2}|\vec{x}|^2 w \coth\left(\frac{wt}{2}\right)\right) e^{iwz/r^2}\, r^{-2} dw \\ &= r^{-4}(p_t \circ \phi_{r^{-1}})(g) \end{aligned} \tag{2.9}$$

through the change of variables $w \mapsto r^{-2} w$. Thus,

$$\begin{aligned} P_t(f \circ \phi_r)(g) &= \int_G (f \circ \phi_r)(gh) p_t(h)\, dh = \int_G f(\phi_r(g)\phi_r(h)) p_t(h)\, dh \\ &= \int_G f(\phi_r(g)h)(p_t \circ \phi_{r^{-1}})(h) r^{-4}\, dh = \int_G f(\phi_r(g)h) p_{r^2 t}(h)\, dh = (P_{r^2 t} f \circ \phi_r)(g); \end{aligned}$$

that is,

$$P_t(f \circ \phi_r) = e^{tL/2}(f \circ \phi_r) = (e^{r^2 tL/2} f) \circ \phi_r = (P_{r^2 t} f) \circ \phi_r. \tag{2.10}$$

The above remarks lead to the following proposition.

Proposition 2.6. *If K_p is the best constant such that*

$$|\nabla P_1 f|^p \leq K_p P_1 |\nabla f|^p, \tag{2.11}$$

for all $f \in C_p^\infty(G)$, then $K_p(t) = K_p$ for all $t > 0$, where $K_p(t)$ is the function introduced in Notation 1.9.

Proof. Fix $t > 0$. Then by Equations (2.8) and (2.10),

$$\begin{aligned} |\nabla P_t(f \circ \phi_{t^{-1/2}})|^p &= |\nabla[(P_1 f) \circ \phi_{t^{-1/2}}]|^p = |t^{-1/2}(\nabla P_1 f) \circ \phi_{t^{-1/2}}|^p \\ &\leq K_p t^{-p/2}(P_1 |\nabla f|^p) \circ \phi_{t^{-1/2}} = K_p t^{-p/2} P_t(|\nabla f|^p \circ \phi_{t^{-1/2}}) \\ &= K_p P_t |\nabla(f \circ \phi_{t^{-1/2}})|^p, \end{aligned}$$

Replacing f by $f \circ \phi_{t^{1/2}}$ in the above computations completes the proof. Moreover, reversing the above argument shows that $|\nabla P_t f|^p \leq K_p P_t |\nabla f|^p$ implies that $|\nabla P_1 f|^p \leq K_p P_1 |\nabla f|^p$. □

2.4 The constant $K_p > 1$

Proposition 2.7. *For $p \in [1, \infty)$, let K_p be the best constant such that*

$$|\nabla P_t f|^p \le K_p P_t |\nabla f|^p \tag{2.12}$$

for all $f \in C_p^\infty(G)$ and $t > 0$. Then $K_p > 1$. In particular, $K_2 \ge 2$.

Proof. First consider the case $p = 2k$ for some positive integer k, and suppose that the constant $K_{2k} = 1$. Then

$$|\nabla P_t f|^{2k} \le P_t |\nabla f|^{2k},$$

for all $t \ge 0$, and $|\nabla f|^{2k} = |\nabla P_0 f|^{2k} = P_0 |\nabla f|^{2k} = |\nabla f|^{2k}$, together would imply that

$$k|\nabla f|^{2(k-1)} \nabla f \cdot \nabla L f = \frac{d}{dt}\bigg|_0 |\nabla P_t f|^{2k} \le \frac{d}{dt}\bigg|_0 P_t |\nabla f|^{2k} = \frac{1}{2} L |\nabla f|^{2k}. \tag{2.13}$$

We now show that the function $f(x, y, z) = x + yz$ violates this inequality. Note that

$$Lf = \nabla \cdot \nabla f = \begin{pmatrix} \tilde{X} \\ \tilde{Y} \end{pmatrix} \cdot \nabla f = \begin{pmatrix} \tilde{X} \\ \tilde{Y} \end{pmatrix} \cdot \begin{pmatrix} 1 - \frac{1}{2} y \cdot y \\ z + \frac{1}{2} x \cdot y \end{pmatrix} = \frac{1}{2} x,$$

$$\nabla L f = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \nabla f \cdot \nabla L f = \frac{1}{2}\left(1 - \frac{1}{2} y \cdot y\right), \text{ and } |\nabla f|^2 (0) = 1.$$

Hence,

$$\left(k|\nabla f|^{2(k-1)} \nabla f \cdot \nabla L f\right)(0) = \frac{k}{2}. \tag{2.14}$$

On the other hand,

$$L\phi(g) = \phi'(g) Lg + \phi''(g) |\nabla g|^2,$$

and so setting $\phi(t) = t^k$ and $g = |\nabla f|^2$ gives

$$L|\nabla f|^{2k} = k|\nabla f|^{2(k-1)} L |\nabla f|^2 + k(k-1) |\nabla f|^{2(k-2)} \left|\nabla |\nabla f|^2\right|^2.$$

From the above,

$$\left|\nabla |\nabla f|^2\right|^2 = \left|\begin{pmatrix} yz + \frac{1}{2} x y^2 - \frac{1}{2} y (2z + xy) \\ -2y + y^3 + xz + \frac{1}{2} 2x^2 y + \frac{1}{2} x (2z + xy) \end{pmatrix}\right|^2,$$

and hence

$$\left| \nabla \left| \nabla f \right|^2 \right|^2 (0) = 0,$$

while $\left(L \left| \nabla f \right|^2 \right) (0) = -2$. Therefore

$$\frac{1}{2} \left(L |\nabla f|^{2k} \right) (0) = -k. \tag{2.15}$$

Inserting the results of Equations (2.14) and (2.15) into Equation (2.13) would imply that $\frac{k}{2} \leq -k$, which is absurd. Thus, $K_{2k} > 1$ for any positive integer k.

For any $p \in [1, \infty)$, there is some integer k such that $p \leq 2k$. Thus,

$$\begin{aligned} |\nabla P_t f|^{2k} &= (|\nabla P_t f|^p)^{2k/p} \\ &\leq K_p^{2k/p} \left(P_t |\nabla f|^p \right)^{2k/p} \leq K_p^{2k/p} P_t |\nabla f|^{2k}. \end{aligned} \tag{2.16}$$

Since K_{2k} is the optimal constant for which (2.16) holds and $K_{2k} > 1$,

$$1 < K_{2k} \leq K_p^{2k/p}$$

implies that $K_p > 1$.

We now quantify this estimate this estimate for $p = 2$. Since

$$K_2 = \sup_{F \in C_p^\infty(G)} \frac{|\nabla P_t F|^2}{P_t |\nabla F|^2}(0) \geq \frac{|\nabla P_t f|^2}{P_t |\nabla f|^2}(0) =: C(t),$$

where $f(x, y, z) = x + yz$, it follows that $K_2 \geq \sup_{t>0} C(t)$. To finish the proof we compute $C(t)$ explicitly. Observe that P_t, when acting on polynomials, may be computed using

$$P_t = e^{tL/2} = \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{tL}{2} \right)^n = I + \frac{t}{2} L + \frac{1}{2!} \cdot \frac{t^2}{4} L^2 + \cdots .$$

We then have

$$P_t f = f + \frac{t}{2} L f = (x + yz) + \frac{t}{2} x, \quad \nabla P_t f = \begin{pmatrix} \left(1 + \frac{t}{2}\right) - \frac{1}{2} y \cdot y \\ z + \frac{1}{2} x \cdot y \end{pmatrix}, \text{ and}$$

$$|\nabla P_t f|^2 = \left(\left(1+\frac{t}{2}\right) - \frac{1}{2}y^2\right)^2 + \left(z + \frac{1}{2}xy\right)^2$$
$$= \left(1 - y^2 + \frac{1}{4}y^4 + z^2 + xyz + \frac{1}{4}x^2y^2\right) + \frac{t}{2}\left(2 - y^2\right) + \frac{t^2}{8}\cdot 2.$$

Also, from before we have,

$$\nabla f = \begin{pmatrix} 1 - \frac{1}{2}y\cdot y \\ z + \frac{1}{2}x\cdot y \end{pmatrix},$$

and so

$$|\nabla f|^2 = \left(1 - \frac{1}{2}y^2\right)^2 + \left(z + \frac{1}{2}xy\right)^2 = 1 - y^2 + \frac{1}{4}y^4 + z^2 + xyz + \frac{1}{4}x^2y^2,$$

$$L|\nabla f|^2 = \left(\frac{1}{2}y^2 + \left(-2 + 3y^2 + \frac{1}{2}x^2\right)\right) + (x\cdot x - y\cdot y) + \frac{1}{4}r^2\cdot 2 = -2 + 3y^2 + 2x^2,$$
$$L^2|\nabla f|^2 = 4 + 6 = 10.$$

Thus,

$$P_t|\nabla f|^2 = \left(1 - y^2 + \frac{1}{4}y^4 + z^2 + xyz + \frac{1}{4}x^2y^2\right) + \frac{t}{2}\left(-2 + 3y^2 + 2x^2\right) + \frac{t^2}{8}\cdot 10.$$

So evaluating at the identity, we have

$$P_t|\nabla f|^2(0) = |\nabla f|^2(0) + \frac{t}{2}L|\nabla f|^2(0) + \frac{t^2}{8}L^2|\nabla f|^2(0) = 1 - t + \frac{5}{4}t^2,$$

and also

$$|\nabla P_t f|^2(0) = 1 + t + \frac{1}{4}t^2.$$

We can find the maximum value of

$$C(t) = \frac{1 + t + \frac{1}{4}t^2}{1 - t + \frac{5}{4}t^2}$$

for $t \geq 0$ by taking derivatives in t to show that $C(t)$ takes on maximum value 2 at $t = \frac{2}{3}$. □

2.5 Review of calculus on Wiener space

This section contains a brief introduction to basic Wiener space definitions and notions of differentiability. For a more complete exposition, we refer the reader to [28, 29, 30, 32, 33, 38, 39, 40, 52, 53, 54] and references contained therein. In particular, the first two chapters of Nualart [40] and Chapter V of Ikeda and Watanabe [30] are cited often here.

Let $(\mathscr{W}(\mathbb{R}^k), \mathcal{F}, \mu)$ denote classical k-dimensional Wiener space. That is, $\mathscr{W} = \mathscr{W}(\mathbb{R}^k)$ is the Banach space of continuous paths $\omega : [0,1] \to \mathbb{R}^k$ such that $\omega_0 = 0$, equipped with the supremum norm

$$\|\omega\| = \max_{t\in[0,1]} |\omega_t|,$$

μ is standard Wiener measure, and $\mathcal{F}$ is the completion of the Borel σ-field on $\mathscr{W}$ with respect to μ. By definition of μ, the process

$$b_t(\omega) = (b_t^1(\omega), \ldots, b_t^k(\omega)) = \omega_t$$

is an $\mathbb{R}^k$ Brownian motion. For those $\omega \in \mathscr{W}$ which are absolutely continuous, let

$$E(\omega) := \int_0^1 |\dot{\omega}_s|^2\, ds$$

denote the energy of ω. The Cameron-Martin Hilbert space is the space of finite energy paths,

$$\mathscr{H} = \mathscr{H}(\mathbb{R}^k) := \{\omega \in \mathscr{W}(\mathbb{R}^k) : \omega \text{ is absolutely continuous and } E(\omega) < \infty\},$$

equipped with the inner product

$$(h,k)_{\mathscr{H}} := \int_0^1 \dot{h}_s \cdot \dot{k}_s\, ds, \quad \text{for all } h,k \in \mathscr{H}.$$

Definition 2.8. Denote by $\mathcal{S}$ the class of *smooth cylinder functionals*; that is, random variables $F : \mathscr{W} \to \mathbb{R}$ such that

$$F(\omega) = f(\omega_{t_1}, \ldots, \omega_{t_n}), \tag{2.17}$$

for some $n \geq 1$, $0 < t_1 < \cdots < t_n \leq 1$, and function $f \in C_p^\infty((\mathbb{R}^k)^n)$ (see Notation 1.1). For E be a real separable Hilbert space, let $\mathcal{S}_E$ be the set of E-valued smooth cylinder functions $F : W \to E$ of the form

$$F = \sum_{j=1}^{m} F_j e_j, \tag{2.18}$$

for some $m \geq 1$, $e_j \in E$, and $F_j \in \mathcal{S}$.

Definition 2.9. Fix $h \in \mathscr{H}$. The directional derivative of a smooth cylinder functional $F \in \mathcal{S}$ of the form (2.17) along h is given by

$$\partial_h F(\omega) := \frac{d}{d\epsilon}\bigg|_0 F(\omega + \epsilon h) = \sum_{i=1}^{n} \nabla^i f(\omega_{t_1}, \ldots, \omega_{t_n}) \cdot h_{t_i}, \tag{2.19}$$

where $\nabla^i f$ is the gradient of f with respect to the i^{th} variable.

The following integration by parts result is standard; see for example Theorem 8.2.2 of Hsu [27].

Proposition 2.10. *Let $F, G \in \mathcal{S}$ and $h \in \mathscr{H}$. Then*

$$(\partial_h F, G)_{\mathscr{H}} = (F, \partial_h^* G)_{\mathscr{H}},$$

where $\partial_h^* = -\partial_h + \int_0^1 \dot{h}_s \cdot db_s$.

Definition 2.11. The gradient of a smooth cylinder functional $F \in \mathcal{S}$ is the random process $D_t F$ taking values in $\mathscr{H}$ such that $(DF, h)_{\mathscr{H}} = \partial_h F$. It may be determined that, for F of the form (2.17),

$$D_t F = \sum_{i=1}^{n} \nabla^i f(\omega_{t_1}, \ldots, \omega_{t_n})(t_i \wedge t),$$

where $s \wedge t = \min\{s, t\}$. For $F \in \mathcal{S}_E$ of the form (2.18), we define the derivative $D_t F$ to be the random process taking values in $\mathscr{H} \otimes E$ given by

$$D_t F := \sum_{j=1}^{m} D_t F \otimes e_j.$$

Iterations of the derivative for smooth functionals $F \in \mathcal{S}$ are given by

$$D^k_{t_1,\dots,t_k} F = D_{t_1} \cdots D_{t_k} F \in \mathscr{H}^{\otimes k},$$

for $k \in \mathbb{N}$. For $F \in \mathcal{S}_E$,

$$D^k F = \sum_{j=1}^{m} D^k F_j \otimes e_j,$$

and these are measurable functions defined almost everywhere on $[0,1]^k \times \mathscr{W}$. The operator D on $\mathcal{S}_E$ is closable, and there exist closed extensions D^k to $L^p(\mathscr{W}, \mathscr{H}^{\otimes k} \otimes E)$; see, for example [40], Theorem 8.28 of [27], or Theorem 8.5 of [30]. We will denote the closure of the derivative operator also by D and the domain of D^k in $L^p([0,1]^k \times \mathscr{W})$ by $\mathcal{D}^{k,p}$, which is the completion of the family of smooth Wiener functionals $\mathcal{S}$ with respect to the seminorm $\|\cdot\|_{k,p,E}$ on $\mathcal{S}_E$ given by

$$\|F\|_{k,p,E} := \left(\sum_{j=0}^{k} \mathbb{E}(\|D^j F\|^p_{\mathscr{H}^{\otimes j} \otimes E}) \right)^{1/p},$$

for any $p \geq 1$. Let

$$\mathcal{D}^{k,\infty}(E) := \bigcap_{p>1} \mathcal{D}^{k,p}(E) \text{ and } \mathcal{D}^{\infty}(E) := \bigcap_{p>1} \bigcap_{k \geq 1} \mathcal{D}^{k,p}(E).$$

When $E = \mathbb{R}$, we write $\mathcal{D}^{k,p}(\mathbb{R}) = \mathcal{D}^{k,p}$, $\mathcal{D}^{k,\infty}(\mathbb{R}) = \mathcal{D}^{k,\infty}$, and $\mathcal{D}^{\infty}(\mathbb{R}) = \mathcal{D}^{\infty}$.

The operator ∂_h on $\mathcal{S}$ is also closable, and there exists a closed extension of $\partial_{h^1} \cdots \partial_{h^k}$ to $L^p(\mu)$. We will denote the closure of ∂_h also by ∂_h, with domain $\mathrm{Dom}(\partial_h)$. We will denote by $G^{k,p}$ the class of functions $F \in L^p(\mu)$ such that $\partial_{h^1} \cdots \partial_{h^j} F \in L^p(\mu)$, for all $h^1, \dots, h^j \in \mathscr{H}$, $j = 1, \dots, k$. The norm on $G^{k,p}$ is given by

$$\|F\|_{G^{k,p}} := \sum_{j=0}^{k} \left(\mathbb{E} \|D^j F\|^p_{(\mathscr{H}^{\otimes j})^*} \right)^{1/p},$$

where

$$\|D^j F\|_{(\mathscr{H}^{\otimes j})^*} = \sup\{ |\partial_{h^1} \cdots \partial_{h^j} F| : h^i \in \mathscr{H}, |h^i|_{\mathscr{H}} \leq 1 \}$$

is the operator norm on the space of continuous linear functionals $(\mathscr{H}^{\otimes j})^*$.

The following result follows from Proposition 5.4.6 and Corollary 5.4.7 of Bogachev [10]. (Note well that here the space $\mathcal{D}^{k,p}$ corresponds to the space $W^{k,p}$ in that text.)

Theorem 2.12. *For all $k \in \mathbb{N}$ and $p \in (1, \infty)$, $\mathcal{D}^{k,p} \subset G^{k,p}$. In particular, for $k = 1$, $\mathcal{D}^{1,p} = G^{1,p}$, for all $p \in (1, \infty)$.*

Definition 2.13. Let D^* denote the $L^2(\mu)$-adjoint of the derivative operator D, which has domain in $L^2(\mathscr{W} \times [0,1], \mathscr{H})$ consisting of functions G such that

$$|\mathbb{E}[(DF, G)_{\mathscr{H}}]| \leq C\|F\|_{L^2(\mu)},$$

for all $F \in \mathcal{D}^{1,2}$, where C is a constant depending on G. For those functions G in the domain of D^*, D^*G is the element of $L^2(\mu)$ such that

$$\mathbb{E}[FD^*G] = \mathbb{E}[(DF, G)_{\mathscr{H}}].$$

It is known that D is a continuous operator from $\mathcal{D}^\infty$ to $\mathcal{D}^\infty(\mathscr{H})$, and similarly, D^* is continuous from $\mathcal{D}^\infty(\mathscr{H})$ to $\mathcal{D}^\infty$; see for example Theorem V-8.1 and its corollary in [30] .

Malliavin [31, 36, 37] introduced the notion of derivatives of Wiener functionals and applied it to the regularity of probability laws induced by the solutions to stochastic differential equations at fixed times. The notion of Sobolev spaces of Wiener functionals was first introduced by Shigekawa [44] and Stroock [45, 46].

2.6 Heisenberg rolling map

For the rest of this chapter, let $\mathscr{W} = \mathscr{W}(\mathbb{R}^2)$ and $\mathscr{H} = \mathscr{H}(\mathbb{R}^2)$. Let $\xi : [0,1] \times \mathscr{W}(\mathbb{R}^2) \to G$ denote the solution to the Stratonovich stochastic differential equation

$$\begin{aligned} d\xi_t &= L_{\xi_t *}X \circ db_t^1 + L_{\xi_t *}Y \circ db_t^2 \\ &= \tilde{X}(\xi_t) \circ db_t^1 + \tilde{Y}(\xi_t) \circ db_t^2 \qquad (2.20) \\ \xi_0 &= 0. \end{aligned}$$

Remark 2.14. Since $\tilde{X}$ and $\tilde{Y}$ have smooth coefficients with bounded partial derivatives, Theorem 2.2.2 in Nualart [40] implies that $\xi_t^i \in \mathcal{D}^\infty$, for $i = 1, 2, 3$ and all $t \in [0, 1]$.

Because G is a nilpotent Lie group, we may determine an "explicit" solution of (2.20),

$$\begin{aligned} d\xi_t &= \tilde{X}(\xi_t^1, \xi_t^2, \xi_t^3) \circ db_t^1 + \tilde{Y}(\xi_t^1, \xi_t^2, \xi_t^3) \circ db_t^2 \\ &= \begin{pmatrix} 1 \\ 0 \\ -\frac{1}{2}\xi_t^2 \end{pmatrix} \circ db_t^1 + \begin{pmatrix} 0 \\ 1 \\ \frac{1}{2}\xi_t^1 \end{pmatrix} \circ db_t^2. \end{aligned}$$

Thus,

$$d\xi_t^1 = db_t^1, d\xi_t^2 = db_t^2, \text{ and } d\xi_t^3 = -\frac{1}{2}\xi_t^2 \circ db_t^1 + \frac{1}{2}\xi_t^1 \circ db_t^2,$$

and one may verify directly that

$$\xi_t = \left(b_t^1, b_t^2, \frac{1}{2} \int_0^t \left[b_s^1 \, db_s^2 - b_s^2 \, db_s^1 \right] \right) \tag{2.21}$$

satisfies the required equation. Note that the third component of ξ may be recognized as Lévy's stochastic area integral.

From Section 3.9 in Gīhman and Skorohod [21] and Theorem 1.22 in Bell [7], the solution ξ is a time homogenous Markov process, and $P_t = e^{tL/2}$ with $L = \tilde{X}^2 + \tilde{Y}^2$ is the associated Markov diffusion semigroup to ξ; that is, $\nu_t := (\xi_t)_*\mu = p_t(g)\, dg$ is the density of the transition probability of the diffusion process ξ_t, and

$$(P_t f)(0) = \mathbb{E}[f(\xi_t)], \tag{2.22}$$

for any $f \in C^\infty(G)$, where the right hand side is expectation for ξ_t started at 0.

Theorem 2.15. *The Malliavin covariance matrix of* ξ_t*,*

$$\sigma_t = \left((D\xi_t^i, D\xi_t^j)_{\mathscr{H}} \right)_{1 \le i,j \le 3},$$

is invertible a.s. for $t > 0$*, and*

$$(\det \sigma)^{-1} \in \cap_{p \ge 1} L^p(\mu) =: L^{\infty-}(\mu).$$

This statement follows from the proof of Theorem 2.3.3 in Nualart [40] and is a special case of Theorem 3.5 in the Heisenberg setting, which will be proved in Section 3.1.2; however, we include the following self-contained proof.

Proof. From Equation (5.19) of the appendix, we have

$$\det \sigma = t^2 \int_0^t |b_s|^2 \, ds - t \left(\int_0^t b_s^1 \, ds \right)^2 - t \left(\int_0^t b_s^2 \, ds \right)^2 .$$

Thus $\det \sigma(\omega) = 0$ if and only if $\omega = 0$, since the Cauchy-Schwarz inequality implies that

$$\det \sigma \geq t^2 \int_0^t |b_s|^2 \, ds - t^2 \int_0^t (b_s^1)^2 \, ds - t^2 \int_0^t (b_s^2)^2 \, ds = 0,$$

with equality if and only if $b_1 = c_1$ and $b_2 = c_2$ for some constants c_1 and c_2. Since $b(0) = 0$, then $b_1 = 0$ and $b_2 = 0$.

To simplify notation, we consider the quadratic form Q on one-dimensional path space $\mathscr{W}([0,T],\mathbb{R})$ defined by

$$Q(\omega) := T^2 \int_0^T \omega_s^2 \, ds - T \left(\int_0^T \omega_s \, ds \right)^2 . \tag{2.23}$$

The argument is similar on $\mathscr{W}([0,T],\mathbb{R}^2)$. We need to show that

$$\int_{\mathscr{W}} \frac{1}{[Q(\omega)]^p} d\mu(\omega) < \infty$$

for any $p \geq 1$. First note that

$$\int_0^\infty e^{-\lambda t} t^p \, \frac{dt}{t} = \frac{1}{\lambda^p} \int_0^\infty e^{-t} t^p \, \frac{dt}{t} =: \frac{1}{\lambda^p} \Gamma(p),$$

under the change of variables $t \mapsto \frac{t}{\lambda}$. Letting $\lambda = Q(\omega)$ then yields

$$\frac{1}{[Q(\omega)]^p} = \frac{1}{\Gamma(p)} \int_0^\infty e^{-tQ(\omega)} t^{p-1} \, dt$$

Thus,

$$\int_{\mathscr{W}} \frac{1}{[Q(\omega)]^p} d\mu(\omega) = \frac{1}{\Gamma(p)} \int_0^\infty dt \, t^{p-1} \int_W d\mu(\omega) e^{-tQ(\omega)} .$$

Define $F(t) := \int_{\mathscr{W}} e^{-tQ(\omega)}\, d\mu(\omega)$. To show that the above integral is finite for any $p \geq 1$, we must show that $F(t)$ decays faster than any polynomial as $t \to \infty$.

Consider first the finite dimensional analogue. That is, for any $a \in \mathbb{R}^N$, consider a quadratic form Q such that

$$Q(a) := \tilde{Q}a \cdot a$$

for some $\tilde{Q} > 0$ and $d\mu(a) = e^{-\frac{1}{2}a\cdot a} \frac{1}{(2\pi)^{N/2}}\, da$. Since $\tilde{Q}$ is compact ($\tilde{Q}$ is an operator on a finite dimensional vector space), there exists an orthonormal basis of eigenvectors $u_1, \ldots, u_N$ such that

$$\tilde{Q}u_i = \lambda_i u_i, \qquad \lambda_i > 0.$$

In this basis, for any $a \in \mathbb{R}^N$, we may write $a = \sum_{i=1}^N \alpha_i u_i$, where $\alpha_i = a \cdot u_i$. Then

$$\begin{aligned} F(t) &= \int_{\mathbb{R}^N} e^{-t\tilde{Q}a\cdot a} e^{-\frac{1}{2}a\cdot a} \frac{1}{(2\pi)^{N/2}}\, da \\ &= \int_{\mathbb{R}^N} e^{-t\sum_{i=1}^N \lambda\alpha_i^2} e^{-\frac{1}{2}\sum_{i=1}^N \alpha_i^2} \prod_{i=1}^N \frac{1}{\sqrt{2\pi}}\, d\alpha_i \\ &= \prod_{i=1}^N \int_{\mathbb{R}} e^{-\frac{1}{2}(1+2t\lambda_i)\alpha_i^2} \frac{1}{\sqrt{2\pi}}\, d\alpha_i \\ &= \prod_{i=1}^N \sqrt{\frac{2\pi}{1+2t\lambda_i}} \cdot \frac{1}{\sqrt{2\pi}} = \prod_{i=1}^N \frac{1}{\sqrt{1+2t\lambda_i}} \leq C_N t^{-N/2}. \end{aligned}$$

Returning then to the infinite dimensional case, Lemma 2.16 below implies that, for $Q : \mathscr{H} \to \mathbb{R}$ defined as in Equation (2.23),

$$Q(h) = (\tilde{Q}h, h)_{\mathscr{H}(\mathbb{R})},$$

where $\tilde{Q} : \mathscr{H}(\mathbb{R}) \to \mathscr{H}(\mathbb{R})$ is a positive compact operator. Thus, there exists an orthonormal basis of eigenvectors $\{h_i\}_{i=1}^\infty$ such that

$$\tilde{Q}h_i = \lambda_i h_i, \quad \lambda_i > 0 \text{ and } \sum_{i=1}^\infty \lambda_i < \infty,$$

and so, for any $h \in \mathscr{H}(\mathbb{R})$, we may write $h = \sum_{i=1}^{\infty} (h, h_i)_{\mathscr{H}} h_i$ and

$$Q(h) = (\tilde{Q}h, h)_{\mathscr{H}} = \left(\sum_{i=1}^{\infty} (h, h_i)_{\mathscr{H}} \tilde{Q} h_i, \sum_{i=1}^{\infty} (h, h_i)_{\mathscr{H}} h_i \right)_{\mathscr{H}} = \sum_{i=1}^{\infty} \lambda_i (h, h_i)_{\mathscr{H}}^2.$$

By Theorem 2.12 from Da Prato and Zabczyk [12], if $\{N_i\}_{i=1}^{\infty}$ is a sequence of independent, identically distributed normal random variables, and we define for $N \in \mathbb{N}$

$$B_N := \sum_{i=1}^{N} N_i h_i,$$

then

$$\lim_{N \to \infty} \left\| B_N - \sum_{i=1}^{\infty} N_i h_i \right\|_{\infty} = 0, \quad \text{a.s.},$$

and $B := \sum_{i=1}^{\infty} N_i h_i$ is Brownian motion. Then

$$Q(B) = \lim_{N \to \infty} Q(B_N) = \lim_{N \to \infty} \sum_{i=1}^{N} \lambda_i N_i^2,$$

and

$$\begin{aligned} \int_W e^{-tQ(\omega)} d\mu(\omega) &= \mathbb{E}\left[e^{-tQ(B)} \right] \\ &= \mathbb{E}\left[\exp\left(-t \lim_{N \to \infty} \sum_{i=1}^{N} \lambda_i N_i^2 \right) \right] \le \mathbb{E}\left[\exp\left(-t \sum_{i=1}^{N} \lambda_i N_i^2 \right) \right] \le C_N t^{-N/2}, \end{aligned}$$

for all $N \in \mathbb{N}$. □

Lemma 2.16. *For $\tilde{Q} : \mathscr{H}(\mathbb{R}) \to \mathscr{H}(\mathbb{R})$ such that*

$$(\tilde{Q}h, h)_{\mathscr{H}(\mathbb{R})} = Q(h) = T^2 \int_0^T \omega_s^2 \, ds - T \left(\int_0^T \omega_s \, ds \right)^2,$$

$\tilde{Q}$ is a trace class operator, and so is compact.

Proof. Let S be an orthonormal basis of $\mathscr{H}(\mathbb{R})$. Then

$$\begin{aligned}\operatorname{tr}(\tilde{Q}) = \sum_{h\in S} Q(h) &= \sum_{h\in S}\left[T^2\int_0^T h^2(s)\,ds - T\left(\int_0^T h(s)\,ds\right)^2\right]\\ &= \sum_{h\in S}\left[T^2\int_0^T h^2(s)\,ds - T\int_0^T\int_0^T h(r)h(s)\,dr\,ds\right]\\ &= T^2\int_0^T s\,ds - T\int_0^T\int_0^T r\wedge s\,dr\,ds\\ &= T^2\cdot\frac{1}{2}T^2 - T\cdot\left(\frac{1}{2}Ts^2 - \frac{1}{6}s^3\right)\Big|_0^T = \frac{1}{6}T^4 < \infty,\end{aligned}$$

where in the third equality, we have used that for any orthonormal basis S of $\mathscr{H}(\mathbb{R})$,

$$\sum_{h\in S} h(r)h(s) = r\wedge s.$$

□

Remark 2.17. By the general theory, Theorem 2.15 implies $\nu_t = \operatorname{Law}(\xi_t)$ is a smooth measure; see for example Theorem 2.12 and Remark 2.13 in Bell [7].

2.7 Lifted vector fields and their L^2-adjoints

Given $A \in \mathfrak{g}$, let $\tilde{A} = (\tilde{A}^1, \tilde{A}^2, \tilde{A}^3)$ be the associated left invariant vector field on G. In particular, we are interested in the vector fields $\tilde{X} = (1, 0, -\frac{1}{2}y)$ and $\tilde{Y} = (0, 1, \frac{1}{2}x)$. We define the "lifted vector field" $\mathbf{A}$ of $\tilde{A}$ as

$$\mathbf{A} = \mathbf{A}^t := \sum_{i,j=1}^{3} \sigma_{ij}^{-1}\tilde{A}^j(\xi_t)D\xi_t^i \in \mathscr{H}, \tag{2.24}$$

acting on functions $F \in \mathcal{D}^{1,2}$ by

$$\mathbf{A}F = (DF, \mathbf{A})_{\mathscr{H}}.$$

Remark 2.18. Recall that D is a continuous operator from $\mathcal{D}^\infty$ to $\mathcal{D}^\infty(\mathscr{H})$. Thus, Remark 2.14 implies that $\tilde{A}^j(\xi_t) \in \mathcal{D}^\infty$ and $D\xi_t^i \in \mathcal{D}^\infty(\mathscr{H})$, for all $t \in [0,1]$. So $\sigma_{ij} \in \mathcal{D}^\infty$ for $i,j = 1,2,3$, and this along with Theorem 2.15 implies that $\sigma_{ij}^{-1} \in \mathcal{D}^\infty$. Hence, $\mathbf{A} \in \mathcal{D}^\infty(\mathscr{H})$.

Proposition 2.19. *For all $f \in C_p^\infty(G)$,*

$$\mathbf{A}[f(\xi_t)] = (\tilde{A}f)(\xi_t).$$

Proof. For any function $f \in C_p^\infty(G)$, $f(\xi_t) \in \mathcal{D}^\infty$ and

$$D[f(\xi_t)] = \sum_{k=1}^{3} \frac{\partial f}{\partial x_k}(\xi_t) D\xi_t^k;$$

see Proposition 1.2.3 from Nualart [40]. Then using Equation (2.24) and for the lifted vector field $\mathbf{A}$ and the Malliavin matrix σ, we have

$$\begin{aligned}
\mathbf{A}[f(\xi_t)] &= (Df(\xi_t), \mathbf{A})_{\mathscr{H}} \\
&= \sum_{i,j,k=1}^{3} \left(\frac{\partial f}{\partial x_k}(\xi_t) D\xi_t^k, \sigma_{ij}^{-1} \tilde{A}^j(\xi_t) D\xi_t^i \right)_{\mathscr{H}} \\
&= \sum_{i,j,k=1}^{3} \tilde{A}^j(\xi_t) \frac{\partial f}{\partial x_k}(\xi_t) \left(D\xi_t^k, D\xi_t^i \right)_{\mathscr{H}} \sigma_{ij}^{-1} \\
&= \sum_{j,k=1}^{3} \tilde{A}^j(\xi_t) \frac{\partial f}{\partial x_k}(\xi_t) \delta_{kj} = \sum_{j=1}^{3} \tilde{A}^j(\xi_t) \frac{\partial f}{\partial x_j}(\xi_t) = (\tilde{A}f)(\xi_t)
\end{aligned}$$

as desired. □

Definition 2.20. For a vector field $\mathbf{A}$ acting on functions of $\mathscr{W}$, we will denote the adjoint of $\mathbf{A}$ in the $L^2(\mu)$ inner product by $\mathbf{A}^*$, which has domain in $L^2(\mu)$ consisting of functions G such that

$$|\mathbb{E}[(\mathbf{A}F)G]| \leq C\|F\|_{L^2(\mu)},$$

for all $F \in \mathcal{D}^{1,2}$, for some constant C. For functions G in the domain of $\mathbf{A}^*$,

$$\mathbb{E}[F(\mathbf{A}^*G)] = \mathbb{E}[(\mathbf{A}F)G],$$

for all $F \in \mathcal{D}^{1,2}$.

Note that for any $F \in \mathcal{D}^{1,2}$,

$$\mathbb{E}[\mathbf{A}F] = \mathbb{E}[(DF, \mathbf{A})_{\mathscr{H}}] = \mathbb{E}[FD^*\mathbf{A}].$$

Thus, we must have that $\mathbf{A}^*1 = D^*\mathbf{A}$ a.s. Recall that D^* is a continuous operator from $\mathcal{D}^\infty(\mathscr{H})$ into $\mathcal{D}^\infty$. Thus, for $\mathbf{A}$ a vector field on $\mathscr{W}$ as defined in Equation (2.24), Remark 2.18 implies that

$$D^*\mathbf{A} = \sum_{i,j=1}^{3} D^*(\sigma_{ij}^{-1}\tilde{A}^j(\xi_t)D\xi_t^i) \in \mathcal{D}^\infty.$$

Thus we have the following proposition.

Proposition 2.21. *Let $\tilde{A}$ be a left invariant vector field on G with lifted vector field $\mathbf{A}$ on $\mathscr{W}$ as defined by Equation (2.24). Then $\mathbf{A}^*1$, where $\mathbf{A}^*$ is the $L^2(\mu)$-adjoint of $\mathbf{A}$, is an element of $\mathcal{D}^\infty$.*

2.8 An L^p-type gradient estimate ($p > 1$) and a Poincaré inequality

Theorem 2.22. *For all $p > 1$,*

$$|\nabla P_t f|^p \leq K_p P_t |\nabla f|^p, \tag{2.25}$$

for all $f \in C_p^\infty(G)$ and $t > 0$, where

$$K_p := 2^{p/q} + 2^{p(\frac{1}{q}+\frac{1}{2})}\left[\|\mathbf{X}^*\xi_1^1\|^2_{L^q(\mu)} + \|\mathbf{X}^*\xi_1^2\|^2_{L^q(\mu)}\right]^{p/2} < \infty, \tag{2.26}$$

with $\mathbf{X}^*$ the adjoint of the lifted vector field $\mathbf{X}$ as in Equation (2.24) with $t = 1$, and $q = \frac{p}{p-1}$.

Proof. By Proposition 2.6, we know the constants K_p are independent of t. Also, Lemma 2.4 states that the inequality is translation invariant. Thus, the proof is reduced to verifying the inequality at the identity for $t = 1$; that is, we must find finite constants K_p such that

$$|\nabla P_1 f|^p(0) \leq K_p P_1 |\nabla f|^p(0), \tag{2.27}$$

for all $f \in C_p^\infty(G)$. So applying Remark 2.1 and Lemma 4.2, consider

$$\begin{aligned}\tilde{X} P_1 f(0) &= P_1 \hat{X} f(0) \\ &= P_1(\tilde{X} + y\tilde{Z}) f(0) = P_1(\tilde{X} f)(0) + P_1(y\tilde{Z} f)(0).\end{aligned}$$

Similarly,

$$\tilde{Y} P_1 f(0) = P_1(\tilde{Y} f)(0) - P_1(x\tilde{Z} f)(0).$$

Thus,

$$\begin{aligned}|\nabla P_1 f|^p(0) &= \left| P_1 \nabla f + P_1 \left(\begin{pmatrix} y \\ -x \end{pmatrix} \tilde{Z} f \right) \right|^p (0) \\ &\leq \left(|P_1 \nabla f| + \left| P_1 \left(\begin{pmatrix} y \\ -x \end{pmatrix} \tilde{Z} f \right) \right| \right)^p (0) \\ &\leq 2^{p/q} \left(|P_1 \nabla f|^p(0) + \left| P_1 \left(\begin{pmatrix} y \\ -x \end{pmatrix} \tilde{Z} f \right) \right|^p (0) \right), \end{aligned} \tag{2.28}$$

where

$$\left| P_1 \left(\begin{pmatrix} y \\ -x \end{pmatrix} \tilde{Z} f \right) \right|^p (0) = [|P_1(y\tilde{Z} f)|^2(0) + |P_1(x\tilde{Z} f)|^2(0)]^{p/2}$$

and $q = \frac{p}{p-1}$ is the conjugate exponent to p. Expanding $\tilde{Z}$ in terms of the Lie bracket gives $\tilde{Z} = \tilde{X}\tilde{Y} - \tilde{Y}\tilde{X}$. So let $F = (F^1, F^2, F^3) := \xi_1$, and using Equation (2.22), consider

$$\begin{aligned}P_1(y\tilde{Z} f)(0) &= P_1(y\tilde{X}\tilde{Y} f)(0) - P_1(y\tilde{Y}\tilde{X} f)(0) \\ &= \mathbb{E}[F^2(\tilde{X}\tilde{Y} f)(F)] - \mathbb{E}[F^2(\tilde{Y}\tilde{X} f)(F)] \\ &= \mathbb{E}[F^2 \mathbf{X}((\tilde{Y} f)(F))] - \mathbb{E}[F^2 \mathbf{Y}((\tilde{X} f)(F))] \\ &= \mathbb{E}[\mathbf{X}^* F^2 \cdot (\tilde{Y} f)(F)] - \mathbb{E}[\mathbf{Y}^* F^2 \cdot (\tilde{X} f)(F)], \end{aligned} \tag{2.29}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are the lifted vector fields of $\tilde{X}$ and $\tilde{Y}$, as in Equation (2.24), with $t = 1$. Hence,

$$\begin{aligned}|P_1(y\tilde{Z}f)|^2(0) &\leq (|\mathbb{E}[\mathbf{X}^*F^2(\tilde{Y}f)(F)]| + |\mathbb{E}[\mathbf{Y}^*F^2(\tilde{X}f)(F)]|)^2 \\ &\leq 2(|\mathbb{E}[\mathbf{X}^*F^2(\tilde{Y}f)(F)]|^2 + |\mathbb{E}[\mathbf{Y}^*F^2(\tilde{X}f)(F)]|^2) \\ &\leq 2[(\mathbb{E}|\mathbf{X}^*F^2|^q)^{2/q}(P_1|\tilde{Y}f|^p)^{2/p}(0) \\ &\qquad + (\mathbb{E}|\mathbf{Y}^*F^2|^q)^{2/q}(P_1|\tilde{X}f|^p)^{2/p}(0)]\end{aligned}$$

by Hölder's inequality. Similarly,

$$|P_1(x\tilde{Z}f)|^2(0) \leq 2[(\mathbb{E}|\mathbf{X}^*F^1|^q)^{2/q}(P_1|\tilde{Y}f|^p)^{2/p}(0) + (\mathbb{E}|\mathbf{Y}^*F^1|^q)^{2/q}(P_1|\tilde{X}f|^p)^{2/p}(0)].$$

Combining this with Equation (2.28), we have

$$\begin{aligned}|\nabla P_1 f|^p(0) &\leq 2^{p/q}\bigg(|P_1\nabla f|^p(0) + \Big[2(\mathbb{E}|\mathbf{X}^*F^2|^q)^{2/q}(P_1|\tilde{Y}f|^p)^{2/p}(0) \\ &\qquad + 2(\mathbb{E}|\mathbf{Y}^*F^2|^q)^{2/q}(P_1|\tilde{X}f|^p)^{2/p}(0) \\ &\qquad + 2(\mathbb{E}|\mathbf{X}^*F^1|^q)^{2/q}(P_1|\tilde{Y}f|^p)^{2/p}(0) \\ &\qquad + 2(\mathbb{E}|\mathbf{Y}^*F^1|^q)^{2/q}(P_1|\tilde{X}f|^p)^{2/p}(0)\Big]^{p/2}\bigg) \\ &\leq 2^{p/q}\bigg(P_1|\nabla f|^p(0) \\ &\qquad + 2^{p/2}\Big[(P_1|\tilde{X}f|^p)^{2/p}(0)[(\mathbb{E}|\mathbf{Y}^*F^1|^q)^{2/q} + (\mathbb{E}|\mathbf{Y}^*F^2|^q)^{2/q}] \\ &\qquad + (P_1|\tilde{Y}f|^p)^{2/p}(0)[(\mathbb{E}|\mathbf{X}^*F^1|^q)^{2/q} + (\mathbb{E}|\mathbf{X}^*F^2|^q)^{2/q}\Big]^{p/2}\bigg),\end{aligned}$$

where we use Hölder's inequality and that $p_1(g)\,dg$ is a probability measure to get

$$|P_1\nabla f|^p(0) \leq P_1|\nabla f|^p(0).$$

So let

$$C_p := (\mathbb{E}|\mathbf{X}^*F^1|^q)^{2/q} + (\mathbb{E}|\mathbf{X}^*F^2|^q)^{2/q}.$$

Note that C_p is a finite constant for all $p > 1$ by Hölder's inequality, Remark 2.14, and Proposition 2.21, since

$$\mathbf{A}^*F = D^*(F\mathbf{A})$$

for any vector field $\mathbf{A}$ on $\mathscr{W}$ and $F \in \mathcal{D}^\infty$. By symmetry,

$$C_p = (\mathbb{E}|\mathbf{Y}^* F^1|^q)^{2/q} + (\mathbb{E}|\mathbf{Y}^* F^2|^q)^{2/q}.$$

Thus,

$$\begin{aligned}|\nabla P_1 f|^p(0) &\le 2^{p/q} P_1 |\nabla f|^p(0) + (2C_p)^{p/2} [(P_1|\tilde{X} f|^p)^{2/p}(0) + (P_1|\tilde{Y} f|^p)^{2/p}(0)]^{p/2} \\ &\le \left(2^{p/q} + 2^{p(\frac{1}{q}+\frac{1}{2})} C_p^{p/2}\right) P_1 |\nabla f|^p(0),\end{aligned}$$

which proves Equation (2.27), and hence, the theorem. □

Theorem 2.23 (Poincaré Inequality). *Let K_2 be given by Equation (2.26) with $p = 2$; that is*

$$K_2 = 2 + 4 \left(||\mathbf{X}^* \xi_1^1||^2_{L^2(\mu)} + ||\mathbf{X}^* \xi_1^2||^2_{L^2(\mu)} \right) < \infty.$$

Let $p_t(g)\,dg$ be the hypoelliptic heat kernel. Then

$$\int_{\mathbb{R}^3} f^2(g) p_t(g)\,dg - \left(\int_{\mathbb{R}^3} f(g) p_t(g)\,dg \right)^2 \le K_2 t \int_{\mathbb{R}^3} |\nabla f|^2(g) p_t(g)\,dg, \tag{2.30}$$

for all $f \in C_p^\infty(G)$ and $t > 0$.

Proof. Let $F_t(g) = (P_t f)(g)$. Then

$$\frac{d}{dt} P_{t-s} F_s^2 = P_{t-s} \left(-\frac{1}{2} L F_s^2 + F_s L F_s \right) = -P_{t-s} |\nabla F_s|^2.$$

Integrating this equation on t implies that

$$\begin{aligned}P_t f^2 - (P_t f)^2 &= \int_0^t P_{t-s} |\nabla F_s|^2\,ds \\ &= \int_0^t P_{t-s} |\nabla P_s f|^2\,ds \\ &\le \int_0^t K_2 P_{t-s} P_s |\nabla f|^2\,ds = K_2 t P_t |\nabla f|^2,\end{aligned}$$

wherein we have made use of Theorem 2.22. Evaluating the above at 0 gives the desired result. □

2.9 The L^1-type gradient estimate

2.9.1 Method fails for the $p = 1$ case

In this section, we show that the argument in the proof of Theorem 2.22 can not be used to prove the inequality (2.25) for $p = 1$.

Proposition 2.24. *Let* $F = (F^1, F^2, F^3) := \xi_1$. *Then*

$$\|\mathbf{X}^* F^1\|_{L^\infty(\mu)} + \|\mathbf{X}^* F^2\|_{L^\infty(\mu)} = \infty.$$

Proof. Let $\sigma(F)$ denote the $\sigma-$ algebra generated by $F : \mathscr{W} \to G$ and $p_t(g)\,dg$ denote the Heisenberg group heat kernel. Then for $f \in C_c^1\left(\mathbb{R}^3\right)$

$$\begin{aligned}
\mathbb{E}[\mathbf{X}^* F^1 f(F)] &= \mathbb{E}[F^1 (\tilde{X} f)(F)] = P_1(x\tilde{X} f)(0) \\
&= \int_G x\tilde{X} f(g) p_1(g)\,dg \\
&= -\int_G f(g)\tilde{X}(x p_1(g))\,dg \\
&= -\int_G f(g)(1 + x\tilde{X}\ln p_1(g))p_1(g)\,dg = -\mathbb{E}[f(F)(1 + x\tilde{X}\ln p_1)(F)],
\end{aligned}$$

where in the third line we have applied standard integration by parts. Consequently, we have shown

$$\mathbb{E}[\mathbf{X}^* F^1 | \sigma(F)] = -(1 + x\tilde{X}\ln p_1)(F).$$

By a similar computation one also shows

$$\mathbb{E}[\mathbf{X}^* F^2 | \sigma(F)] = -(y\tilde{X}\ln p_1)(F).$$

Since conditional expectation is L^p-contractive and the law of F is absolutely continuous relative to Lebesgue measure, it now follows that

$$\begin{aligned}
\|\mathbf{X}^* F^1\|_{L^\infty(\mu)} + \|\mathbf{X}^* F^2\|_{L^\infty(\mu)} &\geq \|\mathbb{E}[\mathbf{X}^* F^1|\sigma(F)]\|_{L^\infty(\mu)} + \|\mathbb{E}[\mathbf{X}^* F^2|\sigma(F)]\|_{L^\infty(\mu)} \\
&= \|1 + x\tilde{X}\ln p_1\|_{L^\infty(\mathbb{R}^3, m)} + \left\|y\tilde{X}\ln p_1\right\|_{L^\infty(\mathbb{R}^3, m)},
\end{aligned}$$

where m is Lebesgue measure. Hence, it suffices to show that either $x\tilde{X}\ln p_1$ or $y\tilde{X}\ln p_1$ is unbounded. We will show $x\tilde{X}\ln p_1$ is unbounded by making use the formula for $p_t(g)$ in Equation (2.5). Letting $t = 1$ in Equation (2.5) and making the change of variables $w \mapsto 2w$, we have

$$p_1(g) = \frac{1}{2\pi^2}\int_{\mathbb{R}} \frac{w}{\sinh w} \exp\left(-\frac{1}{2}|\vec{x}|^2 w \coth w\right) e^{2iwz}\, dw.$$

Then applying $\tilde{X} = \partial_x - \frac{1}{2}y\partial_z$ yields

$$\tilde{X}p_1(g) = -\frac{1}{2\pi^2}\int_{\mathbb{R}} (xw\coth w + iyw)\frac{w}{\sinh w} \exp\left(-\frac{1}{2}|\vec{x}|^2 w \coth w\right) e^{2iwz}\, dw.$$

Setting $y = z = 0$, we then have

$$\tilde{X}\ln p_1(x, 0, 0) = -x\int_{\mathbb{R}} w\coth w d\nu_x(w),$$

where

$$d\nu_x(w) := \frac{1}{z_x}\frac{w}{\sinh w} \exp\left(-\frac{1}{2}x^2 w \coth w\right) dw \tag{2.31}$$

and z_x is the normalizing constant

$$z_x := \int_{\mathbb{R}} \frac{w}{\sinh w} \exp\left(-\frac{1}{2}x^2 w \coth w\right) dw.$$

By Lemma 2.25 below,

$$\lim_{x\to\infty}\int_{\mathbb{R}} w\coth w\, d\nu_x(w) = 1,$$

and so

$$\lim_{x\to\infty}\tilde{X}\ln p_1(x, 0, 0) = \lim_{x\to\infty}\left(-x\int_{\mathbb{R}} w\coth w\, d\nu_x(w)\right) = -\infty.$$

□

Lemma 2.25. *Let $\psi(w) = w\coth w - 1$ and ν_x be as in Equation (2.31). Then*

$$\lim_{x\to\infty}\int \psi\, d\nu_x = \psi(0) = 0. \tag{2.32}$$

Proof. Since $\psi(0) = 0$ and ψ is continuous, to prove Equation (2.32) it suffices to show by the usual approximation of δ – function arguments that

$$\lim_{x\to\infty}\int_{|w|\geq\epsilon}\psi(w)\,d\nu_x(w) = 0$$

holds for every $\epsilon > 0$. We begin by rewriting Equation (2.31) as

$$d\nu_x(w) = \frac{1}{Z_x}\frac{w}{\sinh w}\exp\left(-\frac{1}{2}x^2\psi(w)\right)dw$$

where

$$Z_x := \int_{\mathbb{R}}\frac{w}{\sinh w}\exp\left(-\frac{1}{2}x^2\psi(w)\right)\,dw.$$

A glance at the graph of ψ will convince the reader that there are constants $\alpha, \beta > 0$ (depending on $\epsilon > 0$) such that $\alpha|w| \leq \psi(w) \leq \beta|w|$ for all $|w| \geq \epsilon$. (In fact, one could take $\beta = 1$ independent of ϵ.) Thus

$$\begin{aligned}\int_{|w|\geq\epsilon}\psi(w)\frac{w}{\sinh w}\exp\left(-\frac{1}{2}x^2\psi(w)\right)\,dw &\leq 2\int_{w\geq\epsilon}\beta w e^{-\alpha x^2 w/2}\,dw\\ &= \frac{4\beta}{x^2\alpha}\left(\epsilon + \frac{2}{x^2\alpha}\right)e^{-\alpha x^2\epsilon/2},\end{aligned}$$

where in the inequality we have also used that $\frac{w}{\sinh w} \leq 1$.

Now consider the constant Z_x. We know that for w small, there exists a constant $\gamma > 0$ such that $\psi(w) \leq \gamma w^2$. So letting $\varphi(w) = \frac{w}{\sinh w}$, we have

$$\begin{aligned}Z_x &\geq \int_{|w|\leq\epsilon}\varphi(w)\exp\left(-\frac{1}{2}x^2\psi(w)\right)\,dw\\ &\geq \int_{-\epsilon}^{\epsilon}\varphi(w)e^{-\gamma x^2 w^2/2}\,dw = \frac{1}{x}\int_{-\epsilon x}^{\epsilon x}\varphi\left(\frac{w}{x}\right)e^{-\gamma w^2/2}\,dw,\end{aligned}$$

where we have made the change of variables $w \mapsto \frac{w}{x}$. So, by the dominated convergence theorem,

$$\liminf_{x\to\infty}(xZ_x) \geq \liminf_{x\to\infty}\int_{-\epsilon x}^{\epsilon x}\varphi\left(\frac{w}{x}\right)e^{-\gamma w^2/2}\,dw = \varphi(0)\int_{-\infty}^{\infty}e^{-\gamma w^2/2}\,dw = \sqrt{\frac{2\pi}{\gamma}}.$$

Thus, $Z_x \geq \frac{1}{2}\sqrt{\frac{2\pi}{\gamma}\frac{1}{x}}$ for x sufficiently large, and so

$$\begin{aligned}\lim_{x\to\infty}\int_{|w|\geq\epsilon}\psi(w)\,d\nu_x(w) &= \lim_{x\to\infty}\frac{1}{Z_x}\int_{|w|\geq\epsilon}\psi(w)\frac{w}{\sinh w}\exp\left(-\frac{1}{2}x^2\psi(w)\right)\,dw \\ &\leq 2\lim_{x\to\infty}\frac{\frac{4\beta}{x^2\alpha}\left(\epsilon+\frac{2}{x^2\alpha}\right)e^{-\alpha x^2\epsilon/2}}{\sqrt{\frac{2\pi}{\gamma}\frac{1}{x}}} = 0\end{aligned}$$

as desired. □

2.9.2 Function classes for which the $p = 1$ inequality holds

Although we have not been able to adapt this method of proof to show the $p = 1$ case, we have shown using ad hoc methods that several large classes of functions on G do satisfy the $p = 1$ inequality.

Even and odd functions in z.

Proposition 2.26. *Suppose $f \in C_p^\infty(G)$ is an even or odd function in z on the Heisenberg group G. Then*

$$|\nabla P_t f| \leq \sqrt{2}P_t|\nabla f|,$$

for all $t > 0$.

Proof. Suppose $f \in C_p^\infty(G)$ is an even function on G. Let $g = (x, y, z)$. Then

$$\begin{aligned}\tilde{X}P_tf(0) = P_t(\hat{X}f)(0) &= \int_G\left(f_x + \frac{1}{2}yf_z\right)(g)\,p_t(g)\,dg \\ &= \int_G\left(f_x - \frac{1}{2}yf_z\right)(g)\,p_t(g)\,dg = P_t(\tilde{X}f)(0),\end{aligned}$$

where the third equality follows from

$$\int_{\mathbb{R}} f_z(g)\,p_t(g)\,dz = 0,$$

since f_zp_t is an odd function in z. Thus,

$$|\tilde{X}P_tf|(0) = |P_t\tilde{X}f|(0) \leq P_t|\tilde{X}f|(0).$$

Similarly, it may be shown that $\tilde{Y}P_tf(0) = P_t\tilde{Y}f(0)$, which implies that

$$|\tilde{Y}P_tf|(0) \leq P_t|\tilde{Y}f|(0).$$

Thus, for all $f \in C_c^\infty(G)$ which are even in z,

$$\begin{aligned}|\nabla P_tf|(0) &= \left(|\tilde{X}P_tf|^2 + |\tilde{Y}P_tf|^2\right)^{1/2}(0)\\ &\leq \left((P_t|\tilde{X}f|)^2 + (P_t|\tilde{Y}f|)^2\right)^{1/2}(0) \leq P_t|\tilde{X}f|(0) + P_t|\tilde{Y}f|(0) \leq \sqrt{2}P_t|\nabla f|(0),\end{aligned}$$

since $(a+b)^{1/2} \leq a^{1/2} + b^{1/2}$.

Now suppose $f \in C_p^\infty(G)$ is odd in z. As in the argument above, write

$$\begin{aligned}\tilde{X}P_tf(0) = P_t(\hat{X}f)(0) &= \int_G \left(f_x + \frac{1}{2}yf_z\right)(g)\,p_t(g)\,dg\\ &= \int_G \left(-f_x + \frac{1}{2}yf_z\right)(g)\,p_t(g)\,dg = -P_t(\tilde{X}f)(0),\end{aligned}$$

where in the third equality we have used that

$$\int_{\mathbb{R}} f_x(g)\,p_t(g)\,dz = 0,$$

as f_x is odd in z and p_t is even in z. Thus,

$$|\tilde{X}P_tf|(0) = |P_t\tilde{X}f|(0) \leq P_t|\tilde{X}f|(0).$$

A parallel argument shows that $|\tilde{Y}P_tf|(0) \leq P_t|\tilde{Y}f|(0)$, and thus

$$|\nabla P_tf|(0) \leq \sqrt{2}P_t|\nabla f|(0)$$

for all $f \in C_c^\infty(G)$ which are odd functions in z. □

Functions $f(x,y,z) = e^{i\lambda z}h(x,y)$.

Proposition 2.27. *Suppose $f(x,y,z) = e^{i\lambda z}h(x,y)$ for some $h \in C_p^\infty(\mathbb{R}^2, \mathbb{C})$. Then*

$$|\nabla P_tf| \leq (\sqrt{2}+4)P_t|\nabla f|,$$

for all $t > 0$.

Proof. In the case $\lambda = 0$, $f = f(x,y)$ depends only on $\vec{x} = (x,y)$ and

$$P_t f = e^{tL/2} f = e^{t\Delta/2} f,$$

where Δ is the standard two-dimensional Laplacian, and this reduces to the Gaussian measure case, for which the inequality holds.

So suppose $\lambda \neq 0$, and let $g = (x,y,z)$. For any function $f \in C_p^\infty(G)$, the computations in (2.28) imply that

$$\begin{aligned} |\nabla P_t f|(0) &\leq |P_t \nabla f|(0) + \left| \int_G \begin{pmatrix} y \\ -x \end{pmatrix} (\tilde{Z} f)(g) p_t(g) dg \right| \\ &\leq |P_t \nabla f|(0) + \left| \int_G y (\tilde{Z} f)(g) p_t(g) dg \right| + \left| \int_G x (\tilde{Z} f)(g) p_t(g) dg \right|. \end{aligned} \tag{2.33}$$

For $f(x,y,z) = e^{i\lambda z} h(x,y)$, define the ratios

$$\Lambda_1(f) := \frac{\left| \int_G y \tilde{Z} f(g) p_2(g)\, dg \right|}{P_2 |\tilde{Y} f|(0)} \quad \text{and} \quad \Lambda_2(f) := \frac{\left| \int_G x \tilde{Z} f(g) p_2(g)\, dg \right|}{P_2 |\tilde{X} f|(0)}.$$

For $\lambda > 0$ and h a real-valued function,

$$\Lambda_1(f) = \frac{\left| \int_G y \lambda e^{i\lambda z} h(x,y) p_2(g)\, dg \right|}{\int_G \left| g_y(x,y) + \frac{i}{2} \lambda x h(x,y) \right| p_2(g)\, dg}.$$

The numerator of this expression may be written as

$$\begin{aligned} &\int_G y \lambda e^{i\lambda z} h(x,y) p_2(g)\, dg \\ &\quad = \frac{1}{8\pi^2} \int_{\mathbb{R}} dx \int_{\mathbb{R}} dy\, \lambda y h(x,y) \int_{\mathbb{R}} dz\, e^{i\lambda z} \int_{\mathbb{R}} dw\, \frac{w}{\sinh w} \exp\left(-|\vec{x}|^2 \frac{w \coth w}{4} \right) e^{iwz} \\ &\quad = \frac{1}{8\pi^2} \int_{\mathbb{R}^2} \lambda y h(x,y) \frac{\lambda}{\sinh \lambda} \exp\left(-|\vec{x}|^2 \frac{\lambda \coth \lambda}{4} \right) dx\, dy \\ &\quad = \frac{1}{8\pi^2} \frac{\lambda^2}{\sinh \lambda} \int_{\mathbb{R}^2} h(x,y) \frac{-2}{\lambda \coth \lambda} \partial_y \left[\exp\left(-|\vec{x}|^2 \frac{\lambda \coth \lambda}{4} \right) \right] dx\, dy \\ &\quad = \frac{1}{8\pi^2} \frac{-2\lambda}{\cosh \lambda} \int_{\mathbb{R}^2} h_y(x,y) \exp\left(-|\vec{x}|^2 \frac{\lambda \coth \lambda}{4} \right) dx\, dy, \end{aligned}$$

where in the second equality we have evaluated the Fourier transform in w and the inverse Fourier transform in $-z$. Thus,

$$\begin{aligned}
\Lambda_1(f) &= \frac{\frac{1}{8\pi^2}\left|\frac{2\lambda}{\cosh\lambda}\int_{\mathbb{R}^2} h_y(x,y)\exp\left(-|\vec{x}|^2\frac{\lambda\coth\lambda}{4}\right)dx\,dy\right|}{\int_G \left|h_y(x,y)+\frac{i\lambda}{2}xh(x,y)\right| p_2(g)\,dg} \\
&\le \left|\frac{2\lambda}{\cosh\lambda}\right| \frac{\frac{1}{8\pi^2}\left|\int_{\mathbb{R}^2} h_y(x,y)\exp\left(-|\vec{x}|^2\frac{\lambda\coth\lambda}{4}\right)dx\,dy\right|}{\int_G |h_y(x,y)|\, p_2(g)\,dg} \\
&= \left|\frac{2\lambda}{\cosh\lambda}\right| \frac{\left|\int_{\mathbb{R}^2} h_y(x,y)\exp\left(-|\vec{x}|^2\frac{\lambda\coth\lambda}{4}\right)dx\,dy\right|}{\int_{\mathbb{R}^2} |h_y(x,y)|\, e^{-\frac{|\vec{x}|^2}{4}}\,dx\,dy} \\
&\le \left|\frac{2\lambda}{\cosh\lambda}\right| \frac{\int_{\mathbb{R}^2} |h_y(x,y)|\exp\left(-|\vec{x}|^2\frac{\lambda\coth\lambda}{4}\right)dx\,dy}{\int_{\mathbb{R}^2} |h_y(x,y)|\, e^{-\frac{|\vec{x}|^2}{4}}\,dx\,dy} \le \left|\frac{2\lambda}{\cosh\lambda}\right| \le 2,
\end{aligned}$$

where in the second equality we have evaluated

$$\int_{\mathbb{R}} p_2(g)\,dz = \frac{1}{8\pi^2}\,e^{-\frac{|\vec{x}|^2}{4}},$$

and in the penultimate inequality we have used that $\lambda\coth\lambda \ge 1$ and

$$\left|\frac{\int_{\mathbb{R}^2} h(x,y)e^{k|\vec{x}|^2}\,dx\,dy}{\int_{\mathbb{R}^2} h(x,y)e^{|\vec{x}|^2}\,dx\,dy}\right| \le 1,$$

for any function $h \in C_p^\infty(\mathbb{R}^2)$ and $k \ge 1$. Thus it has been shown that

$$\left|\int_G y\tilde{Z}f(g)p_2(g)\,dg\right| \le 2P_2|\tilde{Y}f|(0).$$

Similarly, one may show that $\Lambda_2(f) \le 2$ by performing integration by parts in the x variable and comparing with the appropriate denominator term. Thus,

$$\left|\int_G x\tilde{Z}f(g)p_2(g)\,dg\right| \le 2P_2|\tilde{X}f|(0).$$

Combining these inequalities with (2.33),

$$\begin{aligned}
|\nabla P_2 f|(0) &\le |P_2\nabla f|(0) + 2(P_2|\tilde{Y}f|(0) + P_2|\tilde{X}f|(0)) \\
&\le P_2|\nabla f|(0) + 2\sqrt{2}P_2|\nabla f|(0) = (1+2\sqrt{2})P_2|\nabla f|(0),
\end{aligned}$$

for all functions $f(x,y,z) = e^{i\lambda z}g(x,y)$ where g is real valued. With obvious modifications to the statement and proof of Proposition 2.6, this is then sufficient proof that

$$|\nabla P_t f|(0) \le K P_t |\nabla f|(0),$$

with $K := 1 + 2\sqrt{2}$, for all $t > 0$.

Clearly the previous argument holds when g is strictly imaginary. So consider $g = u + iv$ for u, v real functions.

$$\begin{aligned} |\nabla P_t f|(x,y,z) &= |\nabla P_t e^{i\lambda z}(u(x,y) + iv(x,y))| \\ &\le K(P_t|\nabla e^{i\lambda z}u(x,y)| + P_t|\nabla e^{i\lambda z}v(x,y)|) \\ &\le \sqrt{2}KP_t|\nabla e^{i\lambda z}(u(x,y) + iv(x,y))| = \sqrt{2}KP_t|\nabla f|(x,y,z) \end{aligned}$$

as desired. □

Radial functions $f = f(r,z)$.

Proposition 2.28. *Suppose $f(x,y,z) = g(r,z)$ for some function $g \in C_c^\infty(\mathbb{R}^2)$, where $r^2 = x^2 + y^2$. Then*

$$|\hat{\nabla} f| = |\nabla f|,$$

where $\hat{\nabla} = (\hat{X}, \hat{Y})$ is the right invariant gradient on G. In particular, this implies that

$$|\nabla P_t f| \le P_t |\nabla f|,$$

for all $t > 0$.

Proof. For $f(x,y,z) = g(r,z)$,

$$\tilde{X}f = \frac{x}{r}g_r - \frac{1}{2}yg_z \text{ and } \tilde{Y}f = \frac{y}{r}g_r + \frac{1}{2}xg_z.$$

Thus,

$$|\nabla f|^2 = \left(\frac{x^2}{r^2}g_r^2 - \frac{xy}{r}g_r g_z + \frac{1}{4}y^2 g_z^2\right) + \left(\frac{y^2}{r^2}g_r^2 + \frac{xy}{r}g_r g_z + \frac{1}{4}x^2 g_z^2\right) = g_r^2 + \frac{1}{4}r^2 g_z^2.$$

Recall from Equation 2.4 that

$$\hat{X} = \partial_x + \frac{1}{2} y \partial_z \text{ and } \hat{Y} = \partial_y - \frac{1}{2} x \partial_z.$$

Therefore,

$$\hat{X} f = \frac{x}{r} g_r + \frac{1}{2} y g_z \text{ and } \hat{Y} f = \frac{y}{r} g_r - \frac{1}{2} x g_z,$$

implies that

$$|\hat{\nabla} f|^2 = \left(\frac{x^2}{r^2} g_r^2 + \frac{xy}{r} g_r g_z + \frac{1}{4} y^2 g_z^2 \right) + \left(\frac{y^2}{r^2} g_r^2 - \frac{xy}{r} g_r g_z + \frac{1}{4} x^2 g_z^2 \right) = g_r^2 + \frac{1}{4} r^2 g_z^2.$$

Hence, $|\hat{\nabla} f| = |\nabla f|$, and it follows that

$$|\nabla P_t f|(0) = |P_t \hat{\nabla} f|(0) \leq P_t |\hat{\nabla} f|(0) = P_t |\nabla f|(0).$$

□

2.9.3 A Green's function inequality

Recall that for a negative linear operator L on a Banach space, the set $\rho(L)$ of all $\lambda \in \mathbb{C}$ such that $\lambda I + L$ is invertible is called the *resolvent set* of L, and the family of bounded linear operators

$$\{R(\lambda) := (\lambda I + L)^{-1} : \lambda \in \rho(L)\}$$

is called the *resolvent* of L; see for example [25, 41].

Proposition 2.29. *If*

$$|\nabla P_t f| \leq K P_t |\nabla f|, \tag{2.34}$$

for all $f \in C_c^\infty(G)$ *and* $t > 0$, *then*

$$|\nabla R(\lambda)^m f| \leq K R(\operatorname{Re} \lambda)^m |\nabla f|,$$

for all $f \in C_c^\infty(G)$, $\operatorname{Re} \lambda > 0$, *and* $m \in \mathbb{N}$.

Proof. For all $m \in \mathbb{N}$ and $\lambda \in \rho(L)$,

$$R(\lambda)^m = \frac{1}{(m-1)!} \int_0^\infty t^{m-1} e^{-\lambda t} P_t f \, dt;$$

see [41]. Suppose (2.34) holds. Then,

$$\begin{aligned} |\nabla R(\lambda)^m f| &= \left| \nabla \frac{1}{(m-1)!} \int_0^\infty t^{m-1} e^{-\mathrm{Re}\,\lambda t} e^{-i\mathrm{Im}\,\lambda t} P_t f \, dt \right| \\ &\le \frac{1}{(m-1)!} \int_0^\infty t^{m-1} e^{-\mathrm{Re}\,\lambda t} |\nabla P_t f| \, dt \\ &\le K \int_0^\infty t^{m-1} e^{-\mathrm{Re}\,\lambda t} P_t |\nabla f| \, dt = R(\mathrm{Re}\,\lambda)^m |\nabla f|, \end{aligned}$$

where we have also used the inequality (2.34) to justify differentiating through the integral. □

It is shown in Folland [18] that, on the Heisenberg group, for $\lambda = 0$ and $n = 1$, $R(0)f = L^{-1}f = f * h$, where

$$h(x, y, z) := \frac{C}{\sqrt{(x^2 + y^2)^2 + 16z^2}},$$

for C is a positive constant; that is, $u = f * h$ is a solution to the equation $Lu = f$. Then the following result then is a mild indication that the Heisenberg L^1 gradient inequality might hold.

Proposition 2.30. *For any* $f \in C_c^\infty(G)$*,*

$$|\nabla(f * h)| \le 9 |\nabla f| * h.$$

Proof. By the translation invariance of Haar measure, it is again sufficient to verify the inequality at the identity. Considering $\nabla(f * g)(0)$ componentwise,

$$\begin{aligned} \tilde{X}(f * h)(0) &= \hat{X}(f * h)(0) = ((\hat{X} f) * h)(0) \\ &= (((\tilde{X} + \frac{1}{2} y \tilde{Z}) f) * h)(0) = ((\tilde{X} f) * h)(0) + \frac{1}{2}((y \tilde{Z} f) * h)(0). \end{aligned} \tag{2.35}$$

Writing $\tilde{Z}$ as the commutator of $\tilde{X}$ and $\tilde{Y}$ and letting $g = (x, y, z)$, the second term in Equation (2.35) may be expanded as

$$\begin{aligned}
\frac{1}{2}((y\tilde{Z}f) * h)(0) &= \frac{1}{2}\int_G y\tilde{Z}f(g)h(g)\,dg = \frac{1}{2}\int_G y[\tilde{X},\tilde{Y}]f(g)h(g)\,dg \\
&= \frac{1}{2}\int_G \tilde{X}\tilde{Y}f(g)yh(g)\,dg - \frac{1}{2}\int_G \tilde{Y}\tilde{X}f(g)yh(g)\,dg \\
&= -\frac{1}{2}\int_G \tilde{Y}f(g)\tilde{X}(yh(g))\,dg + \frac{1}{2}\int_G \tilde{X}f(g)\tilde{Y}(yh(g))\,dg \\
&= -\frac{1}{2}\int_G \tilde{Y}f(g)\left[\frac{-2xy(x^2+y^2)+8xyz}{(x^2+y^2)^2+16z^2}\right]h(g)\,dg \\
&\quad + \frac{1}{2}\int_G \tilde{X}f(g)\left[1 - \frac{2y^2(x^2+y^2)+8xyz}{(x^2+y^2)^2+16z^2}\right]h(g)\,dg.
\end{aligned}$$

The following inequalities are trivially true.

$$\left|\frac{2xy(x^2+y^2)}{(x^2+y^2)^2+16z^2}\right| \le 1, \left|\frac{8y^2z}{(x^2+y^2)^2+16z^2}\right| \le 1, \left|\frac{2y^2(x^2+y^2)}{(x^2+y^2)^2+16z^2}\right| \le 2, \text{ and}$$
$$\left|\frac{8xyz}{(x^2+y^2)^2+16z^2}\right| \le \frac{1}{2}.$$

Thus, $|(y\tilde{Z}f) * h|(0) \le 2(|\tilde{Y}f| * h)(0) + \frac{7}{2}(|\tilde{X}f| * h)(0)$. Combining this with Equation (2.35), we have

$$|\tilde{X}(f * h)|(0) \le 2(|\tilde{Y}f| * h)(0) + \frac{9}{2}(|\tilde{X}f| * h)(0).$$

In a similar manner, one may determine that

$$|\tilde{Y}(f * h)|(0) \le 2(|\tilde{X}f| * h)(0) + \frac{9}{2}(|\tilde{Y}f| * h)(0).$$

Hence,

$$\begin{aligned}
|\nabla(f * h)|^2(0) &= |\tilde{X}(f * h)|^2(0) + |\tilde{Y}(f * h)|^2(0) \\
&\le \left[\left(2|\tilde{Y}f| + \frac{9}{2}|\tilde{X}f|\right) * h\right]^2(0) + \left[\left(2|\tilde{X}f| + \frac{9}{2}|\tilde{Y}f|\right) * h\right]^2(0) \\
&\le 2\left[\frac{9}{2}(|\tilde{X}f| + |\tilde{Y}f|) * h\right]^2(0) \\
&\le 2\left[\frac{9}{2}\left(\sqrt{2}\sqrt{|\tilde{X}f|^2 + |\tilde{Y}\ f|^2}\right) * h\right]^2(0) = (9|\nabla f| * h)^2(0),
\end{aligned}$$

as desired. □

Chapter 3

Wiener calculus over G

We now return to the case of a general Lie group G with identity e and Lie algebra $\mathrm{Lie}(G) = \mathfrak{g}$, and suppose $\{X_i\}_{i=1}^k \subset \mathfrak{g}$ is a Lie generating set, in the sense of Equation (1.7). Recall that we have defined the inner product on $\mathfrak{g}$ such that $\{X_i\}_{i=1}^k$ is an orthonormal basis of the hypoelliptic subspace $\mathfrak{g}_0 = \mathrm{span}(\{X_i\}_{i=1}^k)$.

Notation 3.1. Let $\mathrm{Ad} : G \to \mathrm{End}(\mathfrak{g})$ denote the adjoint representation of G with differential $\mathrm{ad} := d(\mathrm{Ad}) : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$. That is, $\mathrm{Ad}(g) = \mathrm{Ad}_g = L_{g*}R_{g^{-1}*}$, for all $g \in G$, and $\mathrm{ad}(X) = \mathrm{ad}_X = [X, \cdot]$, for all $X \in \mathfrak{g}$. For any function $\varphi \in C^1(G)$, define $\hat{\nabla}\varphi, \tilde{\nabla}\varphi : G \to \mathfrak{g}$ such that, for any $g \in G$ and $X \in \mathfrak{g}$,

$$\left\langle \hat{\nabla}\varphi(g), X \right\rangle := \langle d\varphi(g), R_{g*}X \rangle = (\hat{X}\varphi)(g)$$

and

$$\left\langle \tilde{\nabla}\varphi(g), X \right\rangle := \langle d\varphi(g), L_{g*}X \rangle = (\tilde{X}\varphi)(g).$$

Then,

$$\left\langle \tilde{\nabla}\hat{\nabla}\varphi(g), X \otimes Y \right\rangle = \frac{d}{ds}\bigg|_0 \frac{d}{dt}\bigg|_0 \varphi\left(e^{sX} g e^{tY}\right).$$

for all $X, Y \in \mathfrak{g}$, and similarly for $\hat{\nabla}\tilde{\nabla}\varphi$.

We will use the following facts in the sequel:

$$\begin{aligned}\left\langle \hat{\nabla}\varphi(g), X \right\rangle &= \langle d\varphi(g), L_{g*}L_{g^{-1}*}R_{g*}X\rangle \\ &= \left\langle d\varphi(g), L_{g*}\operatorname{Ad}_{g^{-1}} X\right\rangle = \left\langle \tilde{\nabla}\varphi(g), \operatorname{Ad}_{g^{-1}} X\right\rangle\end{aligned} \tag{3.1}$$

and similarly

$$\left\langle \tilde{\nabla}\varphi(g), X \right\rangle = \left\langle \hat{\nabla}\varphi(g), \operatorname{Ad}_g X\right\rangle. \tag{3.2}$$

We will also use the notation developed in Section 2.5 for calculus on the path space $\mathscr{W}(\mathbb{R}^k)$.

3.1 The main theorems

3.1.1 The rolling map on G

Now suppose $\{b_t^1, \ldots, b_t^k\}$ are k independent real-valued Brownian motions. Then

$$\vec{b}_t := X_i b_t^i := \sum_{i=1}^k X_i b_t^i$$

is a $(\mathfrak{g}_0, \langle\cdot,\cdot\rangle)$ Brownian motion. In the sequel, we will always observe the convention of summing over repeated upper and lower indices. Let $\xi : [0,1]\times\mathscr{W}(\mathbb{R}^k) \to G$ denote the solution to the Stratonovich stochastic differential equation

$$d\xi_t = \xi_t \circ d\vec{b}_t := L_{\xi_t *}\circ d\vec{b}_t = L_{\xi_t *}X_i \circ db_t^i = \tilde{X}_i(\xi_t)\circ db_t^i, \text{ with } \xi_0 = e. \tag{3.3}$$

The solution ξ exists by the standard theory; see, for example, Theorem V-1.1 of [30]. Additionally, Remark V-10.3 of [30] implies that $P_t = e^{tL/2}$, with $L = \sum_{i=1}^k \tilde{X}_i^2$, is the associated Markov diffusion semigroup to ξ, where P_t is as defined in Definition 1.8; that is, $\nu_t := (\xi_t)_*\mu = p_t(g)\,dg$ is the density of the transition probability of the diffusion process ξ_t, where dg here denotes right Haar measure, and

$$(P_t f)(0) = \mathbb{E}[f(\xi_t)], \tag{3.4}$$

for any $f \in C_c^\infty(G)$, where the right hand side is expectation conditioned on $\xi_0 = e$.

Theorem 3.2. *For any $f \in C_c^\infty(G)$, $f(\xi_t) \in \mathcal{D}^\infty$ for all $t \in [0,1]$. In particular, $D[f(\xi_t)] \in \mathscr{H}(\mathbb{R}^k) \otimes \mathbb{R}^k$ and*

$$(D[f(\xi_t)])^i = \left\langle \hat{\nabla} f(\xi_t), \int_0^{\cdot\wedge t} \mathrm{Ad}_{\xi_\tau} X_i \, d\tau \right\rangle, \tag{3.5}$$

for $i = 1, \ldots, k$, componentwise in $\mathscr{H}(\mathbb{R}^k)$.

Notation 3.3. In the sequel, $\mathscr{W} = \mathscr{W}(\mathbb{R}^k)$ and $\mathscr{H} = \mathscr{H}(\mathbb{R}^k)$. For $h = (h^1, \ldots, h^k) \in \mathscr{H}$, let $\vec{h} := X_i h^i \in \mathfrak{g}_0$. We also let

$$\bigcap_{p>1} L^p(\mu) =: L^{\infty-}(\mu).$$

Proof. This proof goes through a series of convergence arguments for solutions to cutoff versions of Equation (3.3). Let $|g|$ denote the distance from a point $g \in G$ to e with respect to the right invariant metric. Let $\{\varphi_m\}_{m=1}^\infty \subset C_c^\infty(G, [0,1])$ be a sequence of functions, $\varphi_m \uparrow 1$, such that $\varphi_m(g) = 1$ when $|g| \leq m$ and $\sup_m \sup_{g \in G} |\tilde{\nabla}^k \varphi_m(g)| < \infty$ for $k = 0, 1, 2, \ldots$; see Lemma 3.6 of Driver and Gross [16]. Let $\psi \in C_c^\infty(\mathrm{End}(\mathfrak{g}), [0,1])$ such that $\psi = 1$ in a neighborhood of I, and $\psi(x) = 0$ if $|x| \geq 2$, where here $|\cdot|$ is the distance from I with respect to any metric on $\mathrm{End}(\mathfrak{g})$. Define $v(g) := \psi(\mathrm{Ad}_g)$ and $u_m(g) := \varphi_m(g) v(g)$. Let $\eta^m : [0,1] \times \mathscr{W} \to G$ denote the solution to the equation

$$d\eta^m = u_m(\eta^m)\eta^m \circ d\vec{b} = u_m(\eta^m)\tilde{X}_i(\eta^m) \circ db^i,$$

with $\eta_0^m = e$, and let $\eta : [0,1] \times \mathscr{W} \to G$ denote the solution to

$$d\eta = v(\eta)\eta \circ d\vec{b} = v(\eta)\tilde{X}_i(\eta) \circ db^i,$$

with $\eta_0 = e$. By Lemma 3.19, these solutions exist for all time t, and by Proposition 3.23, for any $f \in C_c^\infty(G)$,

$$\lim_{m\to\infty} \mathbb{E} \sup_{\tau \leq 1} |f(\eta_\tau^m) - f(\eta_\tau)|^p = 0, \tag{3.6}$$

for all $p \in (1, \infty)$.

Since the vector fields $u_m\tilde{X}_i$ have compact support, we may embed G as a Euclidean submanifold in such a way that the embedded vector fields will be bounded with bounded derivatives. Thus, by Theorem 2.1 of Taniguchi [47], for $f \in C_c^\infty(G)$ and $t \in [0,1]$, $f(\eta_t^m) \in \mathcal{D}^\infty$. Now let $\theta^m : [0,1] \times \mathscr{W} \to \mathfrak{g}$ be the solution to the equation

$$d\theta^m = \left\langle \hat{\nabla} u_m(\eta^m), \theta^m \right\rangle \operatorname{Ad}_{\eta^m} \circ d\vec{b} + u_m(\eta^m) \operatorname{Ad}_{\eta^m} d\vec{h},$$

with $\theta_0^m = 0$, and $\theta_t : \mathscr{W} \to \mathfrak{g}$ be the solution to

$$d\theta = \left\langle \hat{\nabla} v(\eta), \theta \right\rangle \operatorname{Ad}_\eta \circ d\vec{b} + v(\eta) \operatorname{Ad}_\eta d\vec{h},$$

with $\theta_0 = 0$. In Proposition 3.25, we show that

$$\lim_{m\to\infty} \mathbb{E} \sup_{\tau \le 1} |\theta_\tau^m - \theta_\tau|^p = 0,$$

for all $p \in (1,\infty)$, and Proposition 3.20 implies that, for any $h \in \mathscr{H}$,

$$\partial_h[f(\eta_t^m)] = \left\langle \hat{\nabla} f(\eta_t^m), \theta_t^m \right\rangle.$$

It is then easily shown in Proposition 3.26 that these two facts give

$$\lim_{m\to\infty} \mathbb{E} \sup_{\tau \le t} \left| \partial_h[f(\eta_\tau^m)] - \left\langle \hat{\nabla} f(\eta_\tau), \theta_\tau \right\rangle \right|^p = 0, \tag{3.7}$$

for all $p \in (1,\infty)$. Since ∂_h is a closable operator, Equations (3.6) and (3.7) imply that $f(\eta_t) \in \operatorname{Dom}(\partial_h)$, and $\partial_h[f(\eta_t)] = \left\langle \hat{\nabla} f(\eta_t), \theta_t \right\rangle \in L^{\infty-}(\mu)$. It then follows from Theorem 2.12 that $f(\eta_t) \in \mathcal{D}^{1,\infty}$.

Now set $\psi_n(x) := \psi(n^{-1}x)$ and $v_n(g) := \psi_n(\operatorname{Ad}_g) = \psi(n^{-1}\operatorname{Ad}_g)$. Let $\xi^n : [0,1] \times \mathscr{W} \to G$ denote the solution to

$$d\xi^n = v_n(\xi^n)\xi^n \circ d\vec{b} = v_n(\xi^n)\tilde{X}_i(\xi^n) \circ db^i,$$

with $\xi_0^n = e$, and let ξ denote the solution to Equation (3.3). By the previous argument, for any $f \in C_c^\infty(G)$ and $t \in [0,1]$, $f(\xi_t^n) \in \mathcal{D}^{1,\infty}$, and

$$\partial_h[f(\xi_t^n)] = \left\langle \hat{\nabla} f(\xi_t^n), \Theta_t^n \right\rangle,$$

where $\Theta_t^n : \mathscr{W} \to \mathfrak{g}$ satisfies the equation

$$d\Theta^n = \left\langle \hat{\nabla} v_n(\xi^n), \Theta^n \right\rangle \mathrm{Ad}_{\xi^n} \circ d\vec{b} + v_n(\xi^n) \, \mathrm{Ad}_{\xi^n} \, d\vec{h},$$

with $\Theta_0^n = 0$. In Theorem 3.28, we show that

$$\lim_{n\to\infty} \mathbb{E} \sup_{\tau\le 1} |f(\xi_\tau^n) - f(\xi_\tau)|^p = 0$$

and

$$\lim_{n\to\infty} \mathbb{E} \sup_{\tau\le 1} \left| \Theta_\tau^n - R_{\xi_\tau *} \int_0^\tau \mathrm{Ad}_\xi \, X_i \dot{h}^i \, ds \right|^p = 0,$$

for all $p \in (1,\infty)$. Thus, for any $h \in \mathscr{H}$,

$$\lim_{n\to\infty} \mathbb{E} \sup_{\tau\le 1} \left| \partial_h f(\xi_\tau^n) - \left\langle \hat{\nabla} f(\xi_\tau), \int_0^\tau \mathrm{Ad}_\xi \, X_i \dot{h}^i \, ds \right\rangle \right|^p = 0.$$

So $f(\xi_t) \in \mathrm{Dom}(\partial_h)$, and

$$\partial_h f(\xi_t) = \left\langle \hat{\nabla} f(\xi_t), \int_0^t \mathrm{Ad}_\xi \, X_i \dot{h}^i \, ds \right\rangle \in L^{\infty-}(\mu). \tag{3.8}$$

From this equality, we may then show that $\|f(\xi_t)\|_{G^{1,p}} < \infty$, for all $p \in (1,\infty)$. (The norm $\|\cdot\|_{G^{k,p}}$ was defined in Section 2.5.) It then follows from Theorem 2.12 that $f(\xi_t) \in \mathcal{D}^{1,\infty}$, and

$$D_s[f(\xi_t)]^i = \left\langle \hat{\nabla} f(\xi_t), \int_0^{s\wedge t} \mathrm{Ad}_{\xi_\tau} \, X_i d\tau \right\rangle,$$

componentwise in $\mathscr{H}$.

Finally, by Proposition 3.15, $\overline{W} = \int_0^\cdot \mathrm{Ad}_{\xi_\tau} \, d\tau \in \mathcal{D}^\infty(\mathscr{H}(\mathrm{End}(\mathfrak{g})))$. Since $\mathcal{D}^\infty$ is an algebra, this and Equation (3.5) are sufficient to show that $f(\xi_t) \in \mathcal{D}^\infty$, for all $f \in C_c^\infty(G)$ and $t \in [0,1]$. $\square$

The supporting propositions and theorems cited in the above proof may be found in Section 3.2.

3.1.2 Covariance matrix of ξ_t

The Malliavin covariance matrix of ξ is the matrix $\sigma_t(\omega) := \xi_t'(\omega)\xi_t'(\omega)^* : T_{\xi_t(\omega)}G \to T_{\xi_t(\omega)}G$, where $\xi_t'(\omega) : \mathscr{H} \to T_{\xi_t(\omega)}G$ is the Frechet derivative given by

$$\xi_t'(\omega)h := \frac{d}{d\epsilon}\Big|_0 \xi_t(\omega + \epsilon h),$$

for all $h \in \mathscr{H}$, and its adjoint $\xi_t'(\omega)^* : T_{\xi_t(\omega)} \to \mathscr{H}$ is computed relative to the Cameron-Martin inner product on $\mathscr{H}$ and the chosen metric on G.

Notation 3.4. In the following, let $\mathrm{Ad}_{\xi_t}^\dagger$ denote the adjoint of Ad_{ξ_t} as an operator on $\mathfrak{g}$, and let $P : \mathfrak{g} \to \mathfrak{g}_0$ be orthogonal projection onto the hypoelliptic subspace $\mathfrak{g}_0$.

Theorem 3.5. *The Malliavin covariance matrix of ξ is*

$$\sigma_t := \xi_t'(\omega)\xi_t'(\omega)^* = R_{\xi_t *}\left(\int_0^t \mathrm{Ad}_{\xi_s} P \,\mathrm{Ad}_{\xi_s}^\dagger \, ds\right) R_{\xi_t *}^{\mathrm{tr}}, \tag{3.9}$$

Let $\bar{\sigma}_t = \int_0^t \mathrm{Ad}_{\xi_s} P \ \mathrm{Ad}_{\xi_s}^\dagger \ ds$, and $\Delta_t := \det \bar{\sigma}_t$. Then $\Delta_t > 0$ a.e., and so $\bar{\sigma}_t$ is invertible a.e. for $t > 0$. Moreover,

$$\Delta_t^{-1} \in L^{\infty -}(\mu).$$

Proof. We now compute $\xi_t'(\omega)^* : T_{\xi_t(\omega)}G \to \mathscr{H}$, the adjoint in $\xi_t'(\omega)$ with respect to the Cameron-Martin inner product and the right invariant metric on TG. By Equation (3.8), for any $X \in \mathfrak{g}$,

$$\begin{aligned}\left(\xi_t'(\omega)^*(R_{\xi_t *}X), h\right)_{\mathscr{H}} &= \left\langle R_{\xi_t *}X, \xi_t'(\omega)h\right\rangle \\ &= \left\langle R_{\xi_t *}X, R_{\xi_t *}\int_0^t \mathrm{Ad}_{\xi_s} \dot{h}_s \, ds\right\rangle \\ &= \left\langle X, \int_0^t \mathrm{Ad}_{\xi_s} \dot{h}_s \, ds\right\rangle = \int_0^t \left\langle \mathrm{Ad}_{\xi_s}^\dagger X, X_i\right\rangle \dot{h}_s^i ds,\end{aligned}$$

where the penultimate equality follows from the right invariance of the metric on G. It then follows that

$$\dot{k}_s^i = \frac{d}{ds}\left[\xi_t'(\omega)^*(R_{\xi_t *}X)\right]_s^i = 1_{s \le t}\left\langle \mathrm{Ad}_{\xi_s}^\dagger X, X_i\right\rangle, \tag{3.10}$$

as an element $(k^1, \ldots, k^k)$ of the Cameron-Martin space $\mathscr{H}$. Combining Equations (3.8) and (3.10) for $k := X_i k^i$, we have

$$\begin{aligned} \xi_t'(\omega)\xi_t'(\omega)^*(R_{\xi_t *}X) &= \sum_{i=1}^{k} R_{\xi_t *} \int_0^t \operatorname{Ad}_{\xi_s} X_i \left\langle \operatorname{Ad}_{\xi_s}^{\dagger} X, X_i \right\rangle ds \\ &= R_{\xi_t *} \int_0^t \operatorname{Ad}_{\xi_s} P \operatorname{Ad}_{\xi_s}^{\dagger} X ds, \end{aligned}$$

and Equation (3.9) follows from the above.

The proof that $\Delta_t > 0$ and $\Delta_t^{-1} \in L^{\infty-}(\mu)$ is by now standard and relies on satisfaction of the Hörmander bracket condition, $\operatorname{Lie}(\{\tilde{X}_i\}_{i=1}^k) = \mathfrak{g}$; see, for example, the proof of Theorem 8.6 in Driver [15]. For completeness, we have included the adapted proof in the appendix. □

Remark 3.6. By the general theory, Theorem 3.5 implies $\nu_t = \operatorname{Law}(\xi_t)$ is a smooth measure; see for example Remark V-10.3 of [30].

3.1.3 Lifted vector fields and their L^2-adjoints

Throughout this section, $t \in [0, 1]$ will be fixed.

Definition 3.7. Given $X \in \mathfrak{g}$, let $\tilde{X}$ be the associated left invariant vector field on G. Define the "*lifted vector field*" $\mathbf{X}$ of $\tilde{X}$ as

$$\mathbf{X} = \mathbf{X}^t := \xi_t'(\omega)^* \left[\xi_t'(\omega)\xi_t'(\omega)^*\right]^{-1} \tilde{X}(\xi_t) = \xi_t'(\omega)^* \sigma_t^{-1} \tilde{X}(\xi_t) \in \mathscr{H}, \tag{3.11}$$

acting on functions $F \in \mathcal{D}^{1,2}$ by

$$\mathbf{X}F = (DF, \mathbf{X})_{\mathscr{H}}.$$

Proposition 3.8. *For any $X \in \mathfrak{g}$, $\mathbf{X} \in \mathcal{D}^{\infty}(\mathscr{H})$, and*

$$\mathbf{X}[f(\xi_t)] = (\tilde{X} f)(\xi_t),$$

for any $f \in C^{\infty}(G)$,

Proof. Combining Equations (3.9) and (3.10), we have that

$$\frac{d}{ds}\mathbf{X}_s^i = 1_{s\le t}\left\langle \mathrm{Ad}_{\xi_s}^{\dagger}\left(\int_0^t \mathrm{Ad}_{\xi_r} P\ \mathrm{Ad}_{\xi_r}^{\dagger}\ dr\right)^{-1} \mathrm{Ad}_{\xi_t} X, X_i\right\rangle.$$

Thus, we may rewrite Equation (3.11) explicitly as

$$\begin{aligned}\mathbf{X}^i &= \int_0^{\cdot\wedge t}\left\langle \mathrm{Ad}_{\xi_s}^{\dagger}\left(\int_0^t \mathrm{Ad}_{\xi_r} P\ \mathrm{Ad}_{\xi_r}^{\dagger}\ dr\right)^{-1} \mathrm{Ad}_{\xi_t} X, X_i\right\rangle ds \\ &= \left\langle \left(\int_0^{\cdot\wedge t} \mathrm{Ad}_{\xi_s}^{\dagger}\ ds\right)\left(\int_0^t \mathrm{Ad}_{\xi_s} P\ \mathrm{Ad}_{\xi_s}^{\dagger}\ ds\right)^{-1} \mathrm{Ad}_{\xi_t} X, X_i\right\rangle. \end{aligned} \tag{3.12}$$

By Proposition 3.15, $\overline{W} = \int_0^{\cdot} \mathrm{Ad}_{\xi_\tau}\ d\tau \in \mathcal{D}^\infty(\mathscr{H}(\mathrm{End}(\mathfrak{g})))$. Note that $W_t^\dagger = \mathrm{Ad}_{\xi_t}^\dagger : \mathscr{W} \to \mathrm{End}(\mathfrak{g})$ satisfies the differential equation

$$dW_t^\dagger = \mathrm{ad}_{X_i}^\dagger W_t^\dagger \circ db_t^i, \text{ with } W_0^\dagger = I,$$

which is linear with smooth coefficients. Then by a similar argument to that in Proposition 3.15, we may show that

$$\overline{W}^\dagger := \int_0^{\cdot} \mathrm{Ad}_{\xi_\tau}^\dagger\ d\tau \in \mathcal{D}^\infty(\mathscr{H}(\mathrm{End}(\mathfrak{g}))),$$

for all $t \in [0,1]$. Also, Theorem 3.5 implies that

$$\bar{\sigma}_t^{-1} = \left(\int_0^t \mathrm{Ad}_{\xi_r} P\ \mathrm{Ad}_{\xi_r}^\dagger\ dr\right)^{-1}$$

exists and is in $L^{\infty-}(\mu)$ componentwise. Thus the Equation (3.12 implies that $\mathbf{X} \in \mathcal{D}^\infty(\mathscr{H})$, since $\mathcal{D}^\infty(\mathscr{H})$ is an algebra.

For $f \in C^\infty(G)$ and $(h^1, \ldots, h^k) \in \mathscr{H}$, by Equation (3.53),

$$\partial_h[f(\xi_t)] = (D[f(\xi_t)], h)_{\mathscr{H}} = \left\langle \hat{\nabla} f(\xi_t), \int_0^t \mathrm{Ad}_{\xi_s} X_i \dot{h}_s^i\, ds\right\rangle,$$

and so

$$
\begin{aligned}
\mathbf{X}[f(\xi_t)] &= (D[f(\xi_t)], \mathbf{X})_{\mathscr{H}} \\
&= \left\langle \hat{\nabla} f(\xi_t), \int_0^t \operatorname{Ad}_{\xi_s} X_i \left\langle \operatorname{Ad}^\dagger_{\xi_s} \left(\int_0^t \operatorname{Ad}_{\xi_r} P \operatorname{Ad}^\dagger_{\xi_r} \, dr \right)^{-1} \operatorname{Ad}_{\xi_t} X, X_i \right\rangle ds \right\rangle \\
&= \left\langle \hat{\nabla} f(\xi_t), \int_0^t \operatorname{Ad}_{\xi_s} P \operatorname{Ad}^\dagger_{\xi_s} \left(\int_0^t \operatorname{Ad}_{\xi_r} P \operatorname{Ad}^\dagger_{\xi_r} \, dr \right)^{-1} \operatorname{Ad}_{\xi_t} X \, ds \right\rangle \\
&= \left\langle \hat{\nabla} f(\xi_t), \operatorname{Ad}_{\xi_t} X \right\rangle = \left\langle \tilde{\nabla} f(\xi_t), \operatorname{Ad}_{\xi_t^{-1}} \operatorname{Ad}_{\xi_t} X \right\rangle \\
&= (\tilde{X} f)(\xi_t),
\end{aligned}
$$

where we have used Equation (3.1) in the penultimate equality. □

Definition 3.9. For a vector field $\mathbf{X}$ acting on functions of $\mathscr{W}$, we will denote the adjoint of $\mathbf{X}$ in the $L^2(\mu)$ inner product by $\mathbf{X}^*$, which has domain in $L^2(\mu)$ consisting of functions G such that for all $F \in \mathcal{D}^{1,2}$,

$$\mathbb{E}[(\mathbf{X}F)G] \leq c\|F\|_{L^2(\mu)}$$

for some constant c. For functions G in the domain of $\mathbf{X}^*$,

$$\mathbb{E}[F(\mathbf{X}^* G)] = \mathbb{E}[(\mathbf{X}F)G] \tag{3.13}$$

for all $F \in \mathcal{D}^{1,2}$.

Note that for any lifted vector field $\mathbf{X}$ acting on function $F \in \mathcal{D}^{1,2}$ as defined in Definition 3.7,

$$\mathbb{E}[\mathbf{X}F] = \mathbb{E}[(DF, \mathbf{X})_{\mathscr{H}}] = \mathbb{E}[F D^* \mathbf{X}].$$

Thus, we must have that $\mathbf{X}^* = \mathbf{X}^* 1 = D^* \mathbf{X}$ a.s. Recall that D^* is a continuous operator from $\mathcal{D}^\infty(\mathscr{H})$ into $\mathcal{D}^\infty$; see for example Theorem V-8.1 and its corollary in [30]. Thus, for $\mathbf{X}$ a vector field on W as defined in Equation (3.11), Proposition 3.8 implies that $D^* \mathbf{X} \in \mathcal{D}^\infty$. We have then proven the following proposition.

Proposition 3.10. *Let $\tilde{X}$ be a left invariant vector field on G. Then for the vector field on $\mathscr{W}$ defined by*

$$\mathbf{X} = \xi_t'(\omega)^*[\xi_t'(\omega)\xi_t'(\omega)^*]^{-1}\tilde{X}(\xi_t(\omega)),$$

$\mathbf{X}^* \in \mathcal{D}^\infty$, where $\mathbf{X}^*$ is the $L^2(\mu)$-adjoint of $\mathbf{X}$.

This completes the main results of this chapter, and if desired the reader may now continue to Chapter 4 and skip the next technical section which gives details of the proof of Theorem 3.2.

3.2 Results needed in the proof of Theorem 3.2

To prove Theorem 3.2, we will repeatedly use the following standard proposition; see for example Driver [15].

Proposition 3.11. *Suppose $p \in [2, \infty)$, α_t is a predictable $\mathbb{R}^d$–valued process, A_t is a predictable $\mathrm{Hom}(\mathfrak{g}_0, \mathbb{R}^d)$–valued process, and*

$$Y_t := \int_0^t A_\tau \, d\vec{b}_\tau + \int_0^t \alpha_\tau \, d\tau = \int_0^t A_\tau X_i \, db_\tau^i + \int_0^t \alpha_\tau \, d\tau, \tag{3.14}$$

where $\{b^1, \ldots, b^k\}$ are k independent real Brownian motions. Then

$$\mathbb{E} \sup_{\tau \le t} |Y_\tau|^p \le C_p \left\{ \mathbb{E} \left(\int_0^t |A_\tau|^2 \, d\tau \right)^{p/2} + \mathbb{E} \left(\int_0^t |\alpha_\tau| \, d\tau \right)^p \right\}, \tag{3.15}$$

where

$$|A|^2 = \mathrm{tr}(AA^*) = \sum_{i=1}^n (AA^*)_{ii} = \sum_{i,j} A_{ij} A_{ij} = \mathrm{tr}(A^*A).$$

Notation 3.12. Here and in the sequel, we will let δ_n denote constants such that $\lim_{n\to\infty} \delta_n = 0$. Also, we will write $f \lesssim g$, if there is a positive constant C so that $f \le Cg$.

This section is divided into two parts. Section 3.2.1 gives convergence results in a matrix group setting which are necessary to resolve certain convergence issues on the Lie group. These Lie group issues are addressed in Section 3.2.2.

3.2.1 Matrix group computations

In this section, let $M = \mathrm{End}(\mathfrak{g})$ and $A_i = \mathrm{ad}_{X_i} \in M$, for $i = 1, \ldots, k$. Further, let $B := A_i\, b^i = \mathrm{ad}_{X_i}\, b^i = \mathrm{ad}_{\vec{b}}$, where $\{b^i\}_{i=1}^k$ is a set of independent real-valued Brownian motions, and let $H := A_i\, h^i = \mathrm{ad}_{X_i}\, h^i$, where $h = (h^1, \ldots, h^k)$ is a fixed element of $\mathscr{H}(\mathbb{R}^k)$.

Proposition 3.13. *Let $W : [0,1] \times \mathscr{W} \to M$ denote the solution to the Stratonovich differential equations*

$$dW = W \circ dB = W\, dB + \frac{1}{2} \sum_{i=1}^k W A_i^2\, dt, \text{ with } W_0 = I, \tag{3.16}$$

and let $W^s : [0,1]^2 \times \mathscr{W} \to M$ denote the solution to the equation

$$dW^s = W^s \left(\circ dB + s dH \right), \text{ with } W_0^s = I. \tag{3.17}$$

Then

$$\lim_{s \downarrow 0} \mathbb{E} \sup_{\tau \leq 1} |W_\tau^s - W_\tau|^p = 0,$$

for all $p \in (1, \infty)$.

Proof. Writing Equation (3.17) in Itô form gives

$$dW^s = W^s dB + \frac{1}{2} \sum_{i=1}^k W^s A_i^2 dt + s W^s dH.$$

Then by Proposition 3.11, for any $s \in [0,1]$,

$$\begin{aligned} \mathbb{E} \sup_{\tau \leq 1} |W_\tau^s|^p &\lesssim 1 + \mathbb{E} \left(\int_0^t |W_\tau^s|^2\, d\tau \right)^{p/2} + \mathbb{E} \left(\int_0^t |W_\tau^s|\, d\tau \right)^p \\ &\lesssim 1 + \mathbb{E} \int_0^t |W_\tau^s|^p\, d\tau, \end{aligned}$$

for all $t \in [0,1]$. An application of Gronwall's inequality then shows that

$$\mathbb{E} \sup_{\tau \leq 1} |W_\tau^s|^p \leq C e^C,$$

where these constants are independent of s; that is, there exists some finite constant C_p such that

$$\sup_{s\in[0,1]}\left(\mathbb{E}\sup_{\tau\le 1}|W^s_\tau|^p\right)\le C_p,$$

for all $p\in(1,\infty)$.

Now let $\varepsilon^s := W^s - W$. Then from Equation (3.16), ε^s satisfies

$$\begin{aligned} d\varepsilon^s &= (W^s - W)dB + \frac{1}{2}\sum_{i=1}^k (W^s - W)A_i^2 dt + sW^s dH \\ &= \varepsilon^s dB + \frac{1}{2}\sum_{i=1}^k \varepsilon^s A_i^2 dt + sW^s dH. \end{aligned}$$

Then applying Proposition 3.11, we have

$$\mathbb{E}\sup_{\tau\le 1}|\varepsilon^s_\tau|^p \lesssim \int_0^t |\varepsilon^s|^p d\tau + \delta_s,$$

for all $t\in[0,1]$, where

$$\delta_s = \sum_{i=1}^k \int_0^t s^p \left|W^s \dot{h}^i\right|^p d\tau \to 0,$$

as $s\downarrow 0$, by the dominated convergence theorem. Thus, by Gronwall's inequality,

$$\mathbb{E}\sup_{\tau\le 1}|W^s_\tau - W_\tau|^p = \mathbb{E}\sup_{\tau\le 1}|\varepsilon^s_\tau|^p \le \delta_s e^C \to 0,$$

as $s\downarrow 0$, for all $p\in(1,\infty)$. □

Proposition 3.14. *Let W be the solution to Equation (3.16) and W^s be the solution to Equation (3.17). Then $W_t\in\mathcal{D}^\infty(\mathrm{End}(\mathfrak{g}))$, for all $t\in[0,1]$, and $\partial_h W:[0,1]\times\mathscr{W}\to M$ solves the equation*

$$\partial_h W_t = \left(\int_0^t W_\tau \dot{H} W_\tau^{-1}\, d\tau\right) W_t. \tag{3.18}$$

Furthermore,

$$\lim_{s\downarrow 0}\mathbb{E}\sup_{\tau\le 1}\left|\frac{W^s_\tau - W_\tau}{s} - \partial_h W_\tau\right|^p = 0,$$

for all $p\in(1,\infty)$.

Proof. Note that $W = \mathrm{Ad}_\xi$ satisfies the Stratonovich stochastic differential equation

$$d\,\mathrm{Ad}_\xi = \mathrm{Ad}_\xi \circ \mathrm{ad}_{db} = \mathrm{Ad}_\xi\,\mathrm{ad}_{X_i} \circ db^i, \tag{3.19}$$

a linear differential equation with smooth coefficients. Then by Theorem V-10.1 of Ikeda and Watanabe [30], $W_t = \mathrm{Ad}_{\xi_t} \in \mathcal{D}^\infty(\mathrm{End}(\mathfrak{g}))$ componentwise with respect to some basis.

Now let $\Psi : [0,1] \times \mathscr{W} \to M$ denote the solution to the equation

$$d\Psi = \Psi W \circ dB + W dH = \Psi W dB + \frac{1}{2}\sum_{i=1}^{k} \Psi A_i^2 dt + W dH,$$

with $\Psi_0 = 0$. Let $\varepsilon^s := \left(\frac{W^s - W}{s} - \Psi\right)$. Then we have that

$$\begin{aligned} d\varepsilon^s &= \left(\frac{W^s - W}{s} - \Psi\right) dB + \frac{1}{2}\sum_{i=1}^{k}\left(\frac{W^s - W}{s} - \Psi\right) A_i^2 dt + (W^s - W)dH \\ &= \varepsilon^s dB + \frac{1}{2}\sum_{i=1}^{k} \varepsilon^s A_i^2 dt + (W^s - W)dH. \end{aligned}$$

By Proposition 3.11, this implies that

$$\mathbb{E}\sup_{\tau \le t} |\varepsilon^s_\tau|^p \lesssim \int_0^t |\varepsilon^s|^p d\tau + \delta_s,$$

for all $t \in [0,1]$, where

$$\delta_s = \sum_{i=1}^{k} \int_0^t \left|(W^s - W)\dot{h}^i\right|^p d\tau \to 0,$$

as $s \downarrow 0$, by Proposition 3.13 and the dominated convergence theorem. An application of Gronwall's inequality then gives

$$\mathbb{E}\sup_{\tau \le 1}\left|\frac{W^s_\tau - W_\tau}{s} - \Psi_\tau\right|^p = \mathbb{E}\sup_{\tau \le 1} |\varepsilon^s_\tau|^p \le \delta_s e^C \to 0,$$

as $s \downarrow 0$, for all $p \in (1, \infty)$.

By Theorem VIII.2B of Elworthy [17], there exists a modification of W_t^s so that the mapping $s \mapsto W_t^s$ is smooth. Let $F \in \mathcal{S}$ be a smooth cylinder function on $\mathscr{W}$. By the above convergence, we have that

$$\left.\frac{d}{ds}\right|_0 \mathbb{E}\left[W_t^s F\right] = \mathbb{E}[\Psi_t F].$$

Consider also

$$\begin{aligned}\left.\frac{d}{ds}\right|_0 \mathbb{E}\left[W_t^s F\right] &= \int \left.\frac{d}{ds}\right|_0 W_t(b+sh)F(b)\,d\mu(b) \\ &= \int W_t(b) \left.\frac{d}{ds}\right|_0 F(b-sh)\,d\mu(b-sh) \\ &= \int W_t(b)\left[-\partial_h F(b) + \left(\int_0^1 \dot{h}_s \cdot db_s\right) F(b)\right] d\mu(b) = \mathbb{E}[W_t \partial_h^* F];\end{aligned}$$

where the third equality follows from differentiating the shifted measure (see for example Theorem 8.1.1 of Hsu [27]), and the final equality follows from Proposition 2.10. This then implies that $\mathbb{E}[\Psi_t F] = \mathbb{E}[W_t \partial_h^* F]$, and so $\partial_h W_t = \Psi_t$. Thus, $\partial_h W$ satisfies the differential equation

$$d(\partial_h W) = \partial_h W \circ dB + W dH,$$

with $\partial_h W_0 = 0$. Equation (3.18) then follows from an application of Duhamel's principle.

□

Proposition 3.15. *Let W be the solution to Equation (3.16). For $\overline{W} : [0,1] \times \mathscr{W} \to \mathrm{End}(\mathfrak{g})$ given by*

$$\overline{W}_t := \int_0^t W_\tau \, d\tau,$$

we have $\overline{W} \in \mathcal{D}^\infty(\mathscr{H}(\mathrm{End}(\mathfrak{g})))$.

Proof. Let $V = W^{-1} : [0,1] \times \mathscr{W} \to \mathrm{End}(\mathfrak{g})$. By differentiating the identity $W_t W_t^{-1} = I$, one may verify that V satisfies the differential equation

$$dV = -\circ dB \;\; V = -A_i V \circ db^i, \text{ with } V_0 = I.$$

Let $V_t^s : [0,1]^2 \times \mathscr{W} \to \mathrm{End}(\mathfrak{g})$ denote the solution to the equation

$$dV_t^s = -(\circ dB + s dH)V_t^s = -A_i V_t^s \circ db_t^i - s A_i V_t^s \dot{h}_t^i dt, \text{ with } V_0^s = I. \tag{3.20}$$

By the same arguments as in Propositions 3.13 and 3.14, we may determine that

$$\lim_{s \downarrow 0} \mathbb{E} \sup_{\tau \le 1} |V_\tau^s - V_\tau|^p = 0, \tag{3.21}$$

$$\partial_h V_t = -V_t \int_0^t V_\tau^{-1} \dot{H} V_\tau \, d\tau = -V_t \int_0^t W_\tau \dot{H} V_\tau \, d\tau, \tag{3.22}$$

and

$$\lim_{s \downarrow 0} \mathbb{E} \sup_{\tau \le 1} \left| \frac{V_\tau^s - V_\tau}{s} - \partial_h V_\tau \right|^p = 0, \tag{3.23}$$

for all $p \in (1, \infty)$. From the proof of Proposition 3.13, we know that there exists a finite constant C_p such that

$$\sup_{s \in [0,1]} \left(\sup_{\tau \le 1} \mathbb{E} |W_\tau^s|^p \right) \le C_p,$$

for all $p \in (1, \infty)$. We may similarly show that

$$\sup_{s \in [0,1]} \left(\sup_{\tau \le 1} \mathbb{E} |V_\tau^s|^p \right) \le C_p,$$

for all $p \in (1, \infty)$.

Now let $W_{s,t} := V_s W_t$. Note that $W_{s,t}^{-1} = W_{t,s}$, and $W_{s,t} W_{t,u} = W_{s,u}$. By the above bounds on W and V, there exist finite constants C_p such that

$$\sup_{\tau_1, \tau_2 \le t} \mathbb{E} |W_{\tau_1, \tau_2}|^p \le C_p, \tag{3.24}$$

for all $p \in (1, \infty)$. Using this notation and Equation (3.18) we may write

$$\begin{aligned} \partial_h W_t &= \left(\int_0^t W_\tau A_i W_\tau^{-1} \dot{h}_\tau^i \, d\tau \right) W_t \\ &= \int_{\{0 \le \tau \le t\}} W_\tau A_i W_{\tau,t} \dot{h}_\tau^i \, d\tau, \end{aligned} \tag{3.25}$$

and so, for $t_1 \leq t$ and $\alpha_1 = 1 \ldots, k$,

$$[D_{t_1} W_t]^{\alpha_1} = \int_{\{0 \leq \tau \leq t_1\}} W_\tau A_{\alpha_1} W_{\tau,t}\, d\tau, \tag{3.26}$$

where, for $F \in \operatorname{End}(\mathfrak{g})$, $F = \sum_{i=1}^k F^i \otimes A_i$. For $t_1 > t$, $D_{t_1} W_t = 0$.

Let $r \in \mathbb{N}$. For $\{t_1, \ldots, t_r\} \subset [0,1]$ and multi-index $\alpha = (\alpha_1, \ldots, \alpha_r) \subset \{1, \ldots, k\}^r$, let $D^{r,\alpha}_{t_r,\ldots,t_1} W_t$ denote the α^{th} component of $D^r_{t_r,\ldots,t_1} W_t$; that is,

$$D^{r,\alpha}_{t_r,\ldots,t_1} W_t := [D_{t_r}[\cdots [D_{t_2}[D_{t_1} W_t]^{\alpha_1}]^{\alpha_2} \cdots]^{\alpha_r}.$$

We will now show that the following properties hold for any integer $r \geq 1$:

(P1) For any $p \in (1, \infty)$ and multi-index α, $W_t \in \mathcal{D}^{r,p}(\operatorname{End}(\mathfrak{g}))$ and

$$\sup_{\{t_1,\ldots,t_r\} \in [0,1]} \mathbb{E}|D^{r,\alpha}_{t_r,\ldots,t_1} W_t|^p < \infty.$$

(P2) Let $T := \min\{t_1, \ldots, t_r\}$. If $T \leq t$, then the r^{th} derivative of W_t satisfies the linear differential equation

$$D^{r,\alpha}_{t_r,\ldots,t_1} W_t = \int_{[0,T]^r} W_{T_1} A_{J_1} W_{T_1,T_2} A_{J_2} \cdots W_{T_{r-1},T_r} A_{J_r} W_{T_r,t}\, d\tau_r \cdots d\tau_1, \tag{3.27}$$

where T_i denotes the i^{th} smallest element of the set $\{\tau_1, \ldots, \tau_r\}$, and J_i denotes the index corresponding to T_i (that is, $J_i := \sum_{l=1}^r 1_{\{T_i = \tau_l\}} \alpha_l$). If $T > t$, then $D^r_{t_1,\ldots,t_r} W_t = 0$.

By Equation (3.26), the above holds for $r = 1$. Now assume that these properaties hold up to and including order r. Note that Equations (3.22) and (3.25) imply that

$$\begin{aligned}
\partial_h W_{s,t} &= (\partial_h V_s) W_t + V_s (\partial_h W_t) \\
&= -V_s \left(\int_0^s W_\tau A_i V_\tau \dot{h}^i_\tau\, d\tau \right) W_t + V_s \left(\int_0^t W_\tau A_i V_\tau \dot{h}^i_\tau\, d\tau \right) W_t \\
&= \int_{\{s \leq \tau \leq t\}} W_{s,\tau} A_i W_{\tau,t} \dot{h}^i_\tau\, d\tau.
\end{aligned}$$

Let W^s denote the solution to Equation (3.17), and V^s denote the solution to (3.20). Then, for $W^s_{\tau_1,\tau_2} = V^s_{\tau_1} W^s_{\tau_2}$,

$$\begin{aligned}
&\left| \frac{W^s_{\tau_1,\tau_2} - W_{\tau_1,\tau_2}}{s} - \partial_h W_{\tau_1,\tau_2} \right| \\
&\qquad = \left| V^s_{\tau_1} \frac{W^s_{\tau_2} - W_{\tau_2}}{s} - V_{\tau_1}(\partial_h W_{\tau_2}) + \frac{V^s_{\tau_1} - V_{\tau_1}}{s} W_{\tau_2} - (\partial_h V_{\tau_1}) W_{\tau_2} \right| \\
&\qquad \leq \left| V^s_{\tau_1} - V_{\tau_1} \right| \left| \frac{W^s_{\tau_2} - W_{\tau_2}}{s} \right| + \left| V_{\tau_1} \right| \left| \frac{W^s_{\tau_2} - W_{\tau_2}}{s} - (\partial_h W_{\tau_2}) \right| \\
&\qquad\quad + \left| \frac{V^s_{\tau_1} - V_{\tau_1}}{s} - (\partial_h V_{\tau_1}) \right| \left| W_{\tau_2} \right|,
\end{aligned}$$

and thus Proposition 3.14 and Equations (3.21) and (3.23) imply that

$$\lim_{s \downarrow 0} \sup_{\tau_1,\tau_2 \leq t} \mathbb{E} \left| \frac{W^s_{\tau_1,\tau_2} - W_{\tau_1,\tau_2}}{s} - \partial_h W_{\tau_1,\tau_2} \right|^p = 0.$$

So for $a_1, a_2, b_1, b_2 \in [0,1]$, we have

$$\begin{aligned}
&\lim_{s \downarrow 0} \mathbb{E} \left| \int_{a_2}^{b_2} \int_{a_1}^{b_1} \frac{W^s_{\tau_1,\tau_2} - W_{\tau_1,\tau_2}}{s} \, d\tau_1 \, d\tau_2 - \int_{a_2}^{b_2} \int_{a_1}^{b_1} \partial_h W_{\tau_1,\tau_2} \, d\tau_1 \, d\tau_2 \right|^p \\
&\qquad \leq \lim_{s \downarrow 0} \mathbb{E} \int_{a_2}^{b_2} \int_{a_1}^{b_1} \left| \frac{W^s_{\tau_1,\tau_2} - W_{\tau_1,\tau_2}}{s} - \partial_h W_{\tau_1,\tau_2} \right|^p \, d\tau_1 \, d\tau_2 \\
&\qquad = \mathbb{E} \int_{a_2}^{b_2} \int_{a_1}^{b_1} \lim_{s \downarrow 0} \left| \frac{W^s_{\tau_1,\tau_2} - W_{\tau_1,\tau_2}}{s} - \partial_h W_{\tau_1,\tau_2} \right|^p \, d\tau_1 \, d\tau_2 = 0,
\end{aligned}$$

by dominated convergence. Thus,

$$\partial_h \left(\int_{a_2}^{b_2} \int_{a_1}^{b_1} W_{\tau_1,\tau_2} \, d\tau_1 \, d\tau_2 \right) = \int_{a_2}^{b_2} \int_{a_1}^{b_1} (\partial_h W_{\tau_1,\tau_2}) \, d\tau_1 \, d\tau_2.$$

Performing similar estimates, we may then show that

$$\begin{aligned}
&\partial_h D^{r,\alpha}_{t_r,\ldots,t_1} \\
&\quad = \int_{[0,T]^r} \Big[(\partial_h W_{T_1}) A_{J_1} W_{T_1,T_2} A_{J_2} \cdots W_{T_{r-1},T_r} A_{J_r} W_{T_r,t} \\
&\qquad + W_{T_1} A_{J_1} (\partial_h W_{T_1,T_2}) A_{J_2} \cdots W_{T_{r-1},T_r} A_{J_r} W_{T_r,t} \\
&\qquad + \cdots + W_{T_1} A_{J_1} W_{T_1,T_2} A_{J_2} \cdots (\partial_h W_{T_{r-1},T_r}) A_{J_r} W_{T_r,t} \\
&\qquad + W_{T_1} A_{J_1} W_{T_1,T_2} A_{J_2} \cdots W_{T_{r-1},T_r} A_{J_r} (\partial_h W_{T_r,t})\Big]\, d\tau_r \cdots d\tau_1 \\
&\quad = \int_{[0,T]^r} \Big[\int_{\{0\le\tau_{r+1}\le T_1\}} W_{\tau_{r+1}} A_{\alpha_{r+1}} W_{\tau_{r+1},T_1} A_{J_1} W_{T_1,T_2} A_{J_2} \cdots W_{T_{r-1},T_r} A_{J_r} W_{T_r,t} \\
&\qquad + \int_{\{T_1\le\tau_{r+1}\le T_2\}} W_{T_1} A_{J_1} W_{T_1,\tau_{r+1}} A_{\alpha_{r+1}} W_{\tau_{r+1},T_2} A_{J_2} \cdots W_{T_{r-1},T_r} A_{J_r} W_{T_r,t} \\
&\qquad + \cdots + \int_{\{T_{r-1}\le\tau_{r+1}\le T_r\}} W_{T_1} A_{J_1} W_{T_1,T_2} A_{J_2} \cdots W_{T_{r-1},\tau_{r+1}} A_{\alpha_{r+1}} W_{\tau_{r+1},T_r} A_{J_r} W_{T_r,t} \\
&\qquad + \int_{\{T_r\le\tau_{r+1}\le t\}} W_{T_1} A_{J_1} W_{T_1,T_2} A_{J_2} \cdots W_{T_{r-1},T_r} A_{J_r} W_{T_r,\tau_{r+1}} A_{\alpha_{r+1}} W_{\tau_{r+1},t}\Big] \\
&\qquad\quad \times \dot h^{\alpha_{r+1}}_{\tau_{r+1}}\, d\tau_{r+1}\, d\tau_r \cdots d\tau_1,
\end{aligned}$$

which implies exactly that (P2) holds for $r+1$, since

$$\{0\le\tau_{r+1}\le T_1\}, \{T_1\le\tau_{r+1}\le T_2\}, \cdots, \{T_{r-1}\le\tau_{r+1}\le T_r\}, \{T_r\le\tau_{r+1}\le t\}$$

partitions the set $[0,t]$. Clearly, this also implies that, for all $t_{r+1}\in[0,1]$ and $\alpha_{r+1} = 1,\ldots,k$, $[D_{t_{r+1}} D^{r,\alpha}_{t_r,\ldots,t_1}]^{\alpha_{r+1}} \in L^p(\mu)$ for all $p\in(1,\infty)$, by Equation (3.24), and so $W_t \in \mathcal{D}^{r+1,p}(\mathrm{End}(\mathfrak{g}))$.

Now, for $\overline{W}_t = \int_0^t W_\tau\, d\tau$, the above arguments imply that

$$\partial_h \overline{W}_t = \int_0^t \partial_h W_\tau\, d\tau = \int_0^t \int_{\{0\le\tau_1\le\tau\}} W_{\tau_1} A_{\alpha_1} W_{\tau_1,\tau} \dot h^{\alpha_1}_{\tau_1}\, d\tau_1\, d\tau$$

and, for $t_1\le t$ and $\alpha_1 = 1,\ldots,k$,

$$[D_{t_1}\overline{W}_t]^{\alpha_1} = \int_0^t \int_{\{0\le\tau_1\le\tau\wedge t_1\}} W_{\tau_1} A_{\alpha_1} W_{\tau_1,\tau}\, d\tau_1\, d\tau.$$

Let $r \in \mathbb{N}$. For any $\{t_1, \ldots, t_r\} \subset [0,1]$ and multindex $\alpha = (\alpha_1, \ldots, \alpha_r)$, let $D^{r,\alpha}_{t_1,\ldots,t_r}\overline{W}_t$ be the α^{th} component of $D^r_{t_1,\ldots,t_r}\overline{W}_t$; that is,

$$D^{r,\alpha}_{t_r,\ldots,t_1}\overline{W}_t := [D_{t_r}[\cdots[D_{t_2}[D_{t_1}\overline{W}_t]^{\alpha_1}]^{\alpha_2}]\cdots]^{\alpha_r}.$$

Then working as above, we may show that

$$\begin{aligned} D^{r,\alpha}_{t_r,\ldots,t_1}\overline{W}_t &= \int_0^t \int_{[0,T\wedge\tau]^r} W_{T_1}A_{J_1}W_{T_1,T_2}A_{J_2}\cdots W_{T_{r-1},T_r}A_{J_r}W_{T_r,\tau}\,d\tau_r\cdots d\tau_1\,d\tau \\ &= \int_0^t D^{r,\alpha}_{t_r,\ldots,t_1}W_\tau\,d\tau, \end{aligned}$$

and since W_τ satisfies (P1) for all $\tau \in [0,t]$, we have $\overline{W} \in \mathcal{D}^{r,p}(\mathscr{H}(\mathrm{End}(\mathfrak{g})))$, for all $r \in \mathbb{N}$ and $p \in (1,\infty)$. □

Proposition 3.16. *Let $\psi \in C_c^\infty(M)$ such that $\psi = 1$ near I and $\psi(x) = 0$ if $|x| \geq 2$, where $|\cdot|$ is the distance from I with respect to any metric on M. For any $n \in \mathbb{N}$, define $\psi_n(x) := \psi(x/n)$, and, for any $A \in M$, define $\langle\psi'(x), A\rangle := \frac{d}{dt}\big|_0\psi(x+tA)$. Let W denote the solution to Equation (3.16), and let $W^n : [0,1]\times\mathscr{W} \to M$ denote the solution to the Stratonovich differential equation*

$$dW^n = \psi_n(W^n)W^n \circ dB, \text{ with } W_0 = I. \tag{3.28}$$

Then

$$\lim_{n\to\infty}\mathbb{E}\sup_{\tau\leq 1}|W^n_\tau - W_\tau|^p = 0,$$

for all $p \in (1,\infty)$.

Remark 3.17. Notice that $\psi_n'(x) = n^{-1}\psi'(x/n)$, and therefore

$$\left|\psi_n'(x)\right||x| \leq n^{-1}C2n = 2C,$$

where C is a bound on ψ'. Similarly, we may show

$$\left|\psi_n''(x)\right||x|^2 \leq C,$$

where C is determined by a bound on ψ''. These bounds will be used repeatedly in the sequel without further mention.

Proof. Equation (3.28) in Itô form is

$$\begin{aligned} dW^n &= \psi_n(W^n)W^n dB \qquad (3.29)\\ &\quad + \frac{1}{2}\left[\left\langle \psi_n'(W^n), \psi_n(W^n)W^n dB\right\rangle W^n + \psi_n(W^n)\psi_n(W^n)W^n\, dB\right]\, dB \\ &= \psi_n(W^n)W^n dB \qquad (3.30)\\ &\quad + \frac{1}{2}\sum_{i=1}^{k}\left[\psi_n(W^n)\left\langle \psi_n'(W^n), W^n A_i\right\rangle W^n A_i + \psi_n^2(W^n)W^n A_i^2\right] dt. \qquad (3.31) \end{aligned}$$

By Proposition 3.11, Equation (3.31) implies that

$$\begin{aligned} \mathbb{E}\sup_{\tau\le t}|W_\tau^n|^p &\lesssim 1 + \mathbb{E}\left(\int_0^t |W^n|^2\, d\tau\right)^{p/2} + \mathbb{E}\left(\int_0^t |W^n|\, d\tau\right)^p \\ &\lesssim 1 + \mathbb{E}\int_0^t |W^n|^p\ d\tau, \end{aligned}$$

for all $t \in [0,1]$. An application of Gronwall's inequality then shows that

$$\mathbb{E}\sup_{\tau\le 1}|W_\tau^n|^p \le Ce^C,$$

where these constants are independent of n. Thus, there exists some finite constants C_p so that

$$\sup_{n\in\mathbb{N}}\left(\mathbb{E}\sup_{\tau\le 1}|W_\tau^n|^p\right) \le C_p, \qquad (3.32)$$

for all $p \in (1,\infty)$.

Now let $\varepsilon^n := W^n - W$ (so $W^n = W + \varepsilon^n$). Then we have

$$\begin{aligned} d\varepsilon^n &= \left[\psi_n(W^n)W^n - W\right] dB \\ &\quad + \frac{1}{2}\sum_{i=1}^{k}\left[\psi_n(W^n)\left\langle \psi_n'(W^n), W^n A_i\right\rangle W^n A_i + \left(\psi_n^2(W^n)W^n - W\right) A_i^2\right]\ dt. \qquad (3.33) \end{aligned}$$

Applying Proposition 3.11 to each term in Equation (3.33), we may bound $\mathbb{E}|\varepsilon^n|^p$. For the first term,

$$\mathbb{E}\left|\int_0^t [\psi_n(W^n)W^n - W]\,dB\right|^p = \mathbb{E}\left|\int_0^t [\psi_n(W^n)(\varepsilon^n + W) - W]\ dB\right|^p$$

$$\lesssim \mathbb{E}\int_0^t |\psi_n(W^n)\varepsilon^n + (\psi_n(W^n) - 1)W|^p\ d\tau$$

$$\lesssim \mathbb{E}\int_0^t |\psi_n(W^n)|^p|\varepsilon^n|^p\,d\tau + \mathbb{E}\int_0^t |\psi_n(W^n) - 1|^p|W|^p\,d\tau.$$

Similarly, for the second term,

$$\mathbb{E}\left|\int_0^t \psi_n(W^n)\langle\psi_n'(W^n), W^nA_i\rangle W^nA_i\,d\tau\right|^p$$

$$= \mathbb{E}\left|\int_0^t \psi_n(W^n)\langle\psi_n'(W^n), W^nA_i\rangle(\varepsilon^n + W)A_i\,d\tau\right|^p$$

$$\lesssim \mathbb{E}\int_0^t \left[|\psi_n(W^n)||\psi_n'(W^n)||W^n||\varepsilon^n + W|\right]^p d\tau$$

$$\lesssim \mathbb{E}\int_0^t \left[|\psi_n'(W^n)||W^n|\right]^p|\varepsilon^n|^p\,d\tau + \mathbb{E}\int_0^t \left[|\psi_n'(W^n)||W^n||W|\right]^p\ d\tau.$$

And finally,

$$\mathbb{E}\left|\int_0^t [\psi_n^2(W^n)W^n - W]A_i^2\,d\tau\right|^p = \mathbb{E}\left|\int_0^t [\psi_n^2(W^n)(\varepsilon^n + W) - W]A_i^2\,d\tau\right|^p$$

$$\lesssim \mathbb{E}\int_0^t \left|\psi_n^2(W^n)(\varepsilon^n + W) - W\right|^p\ d\tau$$

$$\lesssim \mathbb{E}\int_0^t |\psi_n^2(W^n)|^p|\varepsilon^n|^p\,d\tau + \mathbb{E}\int_0^t |\psi_n^2(W^n) - 1|^p|W|^p\,d\tau.$$

Bringing all of this together together, we have

$$\mathbb{E}\sup_{\tau\le t}|\varepsilon_\tau^n|^p \le C\mathbb{E}\int_0^t |\varepsilon^n|^p\,d\tau + \delta_n,$$

for all $t \in [0,1]$, where (up to constant multiple)

$$\delta_n = \mathbb{E}\int_0^t |\psi_n(W^n) - 1|^p|W|^p\,d\tau$$
$$+ \mathbb{E}\int_0^t \left[|\psi_n'(W^n)||W^n||W|\right]^p\ d\tau + \mathbb{E}\int_0^t |\psi_n^2(W^n) - 1|^p|W|^p\,d\tau.$$

Since $|\psi_n'(W^n)||W^n|$ and $|W|$ remain bounded, the dominated convergence theorem implies that $\lim_{n\to\infty}\delta_n = 0$. Thus, by Gronwall again, we have

$$\mathbb{E}\sup_{\tau\le 1}|W_\tau^n - W_\tau|^p = \mathbb{E}\sup_{\tau\le 1}|\varepsilon_\tau^n|^p \le \delta_n e^C \to 0,$$

as $n\to\infty$. □

Proposition 3.18. *Let W be the solution to Equation (3.16), and let W^n be the solution to Equation (3.28). Then $W^n \in \mathrm{Dom}(\partial_h)$ and*

$$\lim_{n\to\infty}\mathbb{E}\sup_{\tau\le 1}|\partial_h W_\tau^n - \partial_h W_\tau|^p = 0,$$

for all $p\in(1,\infty)$.

Proof. As in Proposition 3.14, we may show that $\partial_h W^n$ satisfies the Itô equation

$$d(\partial_h W^n) = \left[\left\langle \psi_n'(W^n), \partial_h W^n\right\rangle W^n + \psi_n(W^n)(\partial_h W^n)\right] dB + \psi_n(W^n)W^n dH$$
$$+\frac{1}{2}\sum_{i=1}^k \begin{bmatrix} \langle \psi_n'(W^n), \partial_h W^n\rangle \langle \psi_n'(W^n), W^n A_i\rangle W^n A_i \\ +\psi_n(W^n)\langle \psi_n''(W^n), \partial_h W^n \otimes W^n A_i\rangle W^n A_i \\ +\psi_n(W^n)\langle \psi_n'(W^n), (\partial_h W^n)A_i\rangle W^n A_i \\ +\psi_n(W^n)\langle \psi_n'(W^n), W^n A_i\rangle (\partial_h W^n)A_i \\ +2\psi_n(W^n)\langle \psi_n'(W^n), \partial_h W^n\rangle W^n A_i^2 + \psi_n^2(W^n)(\partial_h W^n)A_i^2 \end{bmatrix} dt.$$

Recall also that

$$d(\partial_h W) = (\partial_h W)dB + W dH + \frac{1}{2}(\partial_h W)A_i^2\,dt.$$

Let $\varepsilon^n := \partial_h W^n - \partial_h W$. Then

$$d\varepsilon^n = \left[\left\langle \psi_n'(W^n), \partial_h W^n\right\rangle W^n + (\psi_n(W^n)(\partial_h W^n) - (\partial_h W))\right] dB$$
$$+ \left[\psi_n(W^n)W^n - W\right] dH$$
$$+\frac{1}{2}\sum_{i=1}^k \begin{bmatrix} \langle \psi_n'(W^n), \partial_h W^n\rangle \langle \psi_n'(W^n), W^n A_i\rangle W^n A_i \\ +\psi_n(W^n)\langle \psi_n''(W^n), \partial_h W^n \otimes W^n A_i\rangle W^n A_i \\ +\psi_n(W^n)\langle \psi_n'(W^n), (\partial_h W^n)A_i\rangle W^n A_i \\ +\psi_n(W^n)\langle \psi_n'(W^n), W^n A_i\rangle (\partial_h W^n)A_i \\ +2\psi_n(W^n)\langle \psi_n'(W^n), \partial_h W^n\rangle W^n A_i^2 + \left[\psi_n^2(W^n)\partial_h W^n - \partial_h W\right] A_i^2 \end{bmatrix} dt.$$

That is,

$$d\varepsilon^n = \left\langle \psi_n'(W^n), \partial_h W^n \right\rangle W^n dB + [\psi_n(W^n)(\partial_h W^n) - (\partial_h W)]dB + [\psi_n(W^n)W^n - W]dH$$
$$+ \frac{1}{2}\sum_{i=1}^{k} \left[\Gamma_n(\partial_h W^n) + [\psi_n^2(W^n)\partial_h W^n - \partial_h W]A_i^2\right] dt, \quad (3.34)$$

where

$$\Gamma_n(A) = \sum_{i=1}^{k} \left[\left\langle \psi_n'(W^n), A\right\rangle \left\langle \psi_n'(W^n), W^n A_i\right\rangle W^n A_i \right.$$
$$+ \psi_n(W^n)\left\langle \psi_n''(W^n), A \otimes W^n A_i\right\rangle W^n A_i + \psi_n(W^n)\left\langle \psi_n'(W^n), AA_i\right\rangle W^n A_i$$
$$\left. + \psi_n(W^n)\left\langle \psi_n'(W^n), W^n A_i\right\rangle AA_i + 2\psi_n(W^n)\left\langle \psi_n'(W^n), A\right\rangle W^n A_i^2\right]$$

satisfies the bound

$$|\Gamma_n| \lesssim |\psi_n''(W^n)||W^n|^2 + |\psi_n'(W^n)||W^n| + |\psi_n'(W^n)|^2|W^n|^2. \quad (3.35)$$

Again using Proposition (3.11), we may work through (3.34) term by term to bound $\mathbb{E}|\varepsilon^n|^p$:

For the first term in the sum,

$$\mathbb{E}\left|\int_0^t \left\langle \psi_n'(W^n), \partial_h W^n\right\rangle W^n \, dB\right|^p \lesssim \mathbb{E}\int_0^t \left|\left\langle \psi_n'(W^n), \varepsilon^n + \partial_h W\right\rangle W^n\right|^p d\tau$$
$$\lesssim \mathbb{E}\int_0^t |\varepsilon^n|^p d\tau + \mathbb{E}\int_0^t \left[|\psi_n'(W^n)||\partial_h W||W^n|\right]^p d\tau.$$

Considering the second term, we have

$$\mathbb{E}\left|\int_0^t [\psi_n(W^n)\partial_h W^n - \partial_h W] \, dB\right|^p \lesssim \mathbb{E}\int_0^t |\psi_n(W^n)(\varepsilon^n + \partial_h W) - \partial_h W|^p \, d\tau$$
$$\lesssim \mathbb{E}\int_0^t |\varepsilon^n|^p d\tau + \mathbb{E}\int_0^t |\psi_n(W^n) - 1|^p |\partial_h W|^p \, d\tau.$$

For the third term, note that

$$\mathbb{E}\left|\int_0^t [\psi_n(W^n)W^n - W] \, dH\right|^p = C\|H\|_{\mathscr{H}}^p \mathbb{E}\sup_{\tau \le 1} |\psi_n(W_\tau^n)W_\tau^n - W_\tau|^p$$

and

$$\mathbb{E}\sup_{\tau\le 1}|\psi_n(W^n_\tau)W^n_\tau - W_\tau|^p = \mathbb{E}\sup_{\tau\le 1}|\psi_n(W^n_t)(W_t + (W^n_t - W_t)) - W_t|^p$$
$$\le \mathbb{E}\sup_{\tau\le 1}\left(|\psi_n(W^n_\tau) - 1|^p|W_\tau|^p + |\psi_n(W^n_\tau)|^n|W^n_\tau - W_\tau|^p\right) \to 0,$$

as $n\to\infty$, by Proposition 3.16 and dominated convergence. Using the bound in (3.35) on the fourth term, we have

$$\mathbb{E}\left|\int_0^t \Gamma_n(\partial_h W^n)\,d\tau\right|^p = \mathbb{E}\left|\int_0^t \Gamma_n(\varepsilon^n + \partial_h W)\,d\tau\right|^p$$
$$\lesssim \mathbb{E}\int_0^t |\varepsilon^n|^p\,d\tau$$
$$+\,\mathbb{E}\int_0^t \left(|\psi_n''(W^n)||W^n|^2 + |\psi_n'(W^n)||W^n| + \left|\psi_n'(W^n)\right|^2|W^n|^2\right)^p |\partial_h W|^p\,d\tau.$$

Finally, for the last term,

$$\mathbb{E}\left|\int_0^t \left[\psi_n^2(W^n)\partial_h W^n - \partial_h W\right] A_i^2\,d\tau\right|^p \lesssim \mathbb{E}\int_0^t \left|\psi_n^2(W^n)(\varepsilon^n + \partial_h W) - \partial_h W\right|^p\,d\tau$$
$$\lesssim \mathbb{E}\int_0^t |\varepsilon^n|^p d\tau + \mathbb{E}\int_0^t |\psi_n^2(W^n) - 1|^p|\partial_h W|^p\,d\tau.$$

Putting this all together shows

$$\mathbb{E}\sup_{\tau\le t}|\varepsilon^n_t|^p \le C\mathbb{E}\int_0^t |\varepsilon^n|^p d\tau + \delta_n,$$

for all $t\in[0,1]$, where

$$\delta_n = \mathbb{E}\int_0^t |\psi_n'(W^n)|^p|\partial_h W|^p|W^n|^p d\tau + \mathbb{E}\int_0^t |\psi_n(W^n) - 1|^p|\partial_h W|^p d\tau$$
$$+\,\mathbb{E}\left|\int_0^t \left[\psi_n(W^n)W^n - W\right] dH\right|^p$$
$$+\,\mathbb{E}\int_0^t \left(|\psi_n''(W^n)||W^n|^2 + |\psi_n'(W^n)||W^n| + \left|\psi_n'(W^n)\right|^2|W^n|^2\right)^p |\partial_h W|^p\,d\tau$$
$$+\,\mathbb{E}\int_0^t |\psi_n^2(W^n) - 1|^p|\partial_h W|^p\,d\tau.$$

Again using Remark 3.17, $\lim_{n\to\infty} \delta_n = 0$, by the dominated convergence theorem. Thus, another application of Gronwall's inequality shows that

$$\mathbb{E}\sup_{\tau\le 1} |\partial_h W^n_\tau - \partial_h W_\tau|^p = \mathbb{E}\sup_{\tau\le 1} |\varepsilon^n_\tau|^p \le \delta_n e^C \to 0,$$

as $n \to \infty$. □

3.2.2 Lie group computations

Lemma 3.19. *Let $u \in C^1(G)$ such that u and $\tilde{X}u$ are bounded for all $X \in \mathfrak{g}_0$. Then the solution $\eta : [0,1] \times \mathscr{W}(\mathbb{R}^k) \to G$ to the stochastic differential equation*

$$d\eta = u(\eta)\eta \circ d\vec{b} := u(\eta)L_{\eta *} \circ d\vec{b} = u(\eta)\tilde{X}_i(\eta) \circ db^i, \text{ with } \eta_0 = e$$

exists for all time; that is, η has no explosion.

Proof. Let ζ be the life-time of η and $\varphi \in C_c^\infty(G)$. Then on $\{t < \zeta\}$,

$$\begin{aligned} d\varphi(\eta) &= u(\eta)\tilde{X}_i\varphi(\eta) \circ db^i = u(\eta)\left\langle \tilde{\nabla}\varphi(\eta), X_i\right\rangle \circ db^i = u(\eta)\left\langle \tilde{\nabla}\varphi(\eta), \circ d\vec{b}\right\rangle \\ &= u(\eta)\left\langle \tilde{\nabla}\varphi(\eta), d\vec{b}\right\rangle + \frac{1}{2} d\left[u(\eta)\left\langle \tilde{\nabla}\varphi(\eta), \cdot\right\rangle\right] d\vec{b} \\ &= u(\eta)\left\langle \tilde{\nabla}\varphi(\eta), d\vec{b}\right\rangle + \frac{1}{2}\left\langle \tilde{\nabla}\left[u(\eta)\left\langle \tilde{\nabla}\varphi(\eta), \cdot\right\rangle\right], d\vec{b}\right\rangle d\vec{b} \\ &= u(\eta)\left\langle \tilde{\nabla}\varphi(\eta), d\vec{b}\right\rangle + \frac{1}{2}\left[\left\langle \tilde{\nabla}u(\eta), d\vec{b}\right\rangle\left\langle \tilde{\nabla}\varphi(\eta), d\vec{b}\right\rangle + u(\eta)\left\langle \tilde{\nabla}^2\varphi(\eta), d\vec{b}\otimes d\vec{b}\right\rangle\right] \\ &= u(\eta)\left\langle \tilde{\nabla}\varphi(\eta), d\vec{b}\right\rangle \\ &\quad + \frac{1}{2}\sum_{i=1}^k \left[\left\langle \tilde{\nabla}u(\eta), X_i\right\rangle\left\langle \tilde{\nabla}\varphi(\eta), X_i\right\rangle + u(\eta)\left\langle \tilde{\nabla}^2\varphi(\eta), X_i\otimes X_i\right\rangle\right] dt. \end{aligned} \tag{3.36}$$

Let $|g|$ denote the distance from a point $g \in G$ to e with respect to the right invariant metric. Let $\{\varphi_m\}_{m=1}^\infty \subset C_c^\infty(G,[0,1])$ be a sequence of functions, $\varphi_m \uparrow 1$, such that $\varphi_m(g) = 1$ when $|g| \le m$ and $\sup_m \sup_{g\in G} |\tilde{\nabla}^k \varphi_m(g)| < \infty$ for $k = 0, 1, 2, \ldots$; see Lemma 3.6 of Driver and Gross [16]. Also, let $\{K_n\}_{n=1}^\infty$ be a sequence of nested compact sets in G, such that $K_n \uparrow G$, and take τ_n to be the exit time from K_n,

$$\tau_n := \inf\{t > 0 : \eta_t \notin K_n\}.$$

Let $\eta_t^n := \eta_{t\wedge\tau_n}$. Then from Equation (3.36) (using the convention that $\varphi_m(\eta_t) = 0$, $\tilde{\nabla}\varphi_m(\eta_t) = 0$, and $\tilde{\nabla}^2\varphi_m(\eta_t) = 0$ on $\{t > \zeta\}$) we have

$$\varphi_m(\eta_t^n) = 1 + \int_0^{t\wedge\tau_n} u(\eta) \left\langle \tilde{\nabla}\varphi_m(\eta), d\vec{b} \right\rangle$$
$$+ \frac{1}{2}\sum_{i=1}^k \int_0^{t\wedge\tau_n} \left[\left\langle \tilde{\nabla}u(\eta), X_i \right\rangle \left\langle \tilde{\nabla}\varphi_m(\eta), X_i \right\rangle + u(\eta) \left\langle \tilde{\nabla}^2\varphi_m(\eta), X_i \otimes X_i \right\rangle \right] d\tau.$$

Taking the expectation of this equation then gives

$$\mathbb{E}\left[\varphi_m(\eta_t^n)\right] = 1 + \frac{1}{2}\delta_m,$$

where

$$\delta_m := \sum_{i=1}^k \mathbb{E}\int_0^{t\wedge\tau_n} \left[\left\langle \tilde{\nabla}u(\eta), X_i \right\rangle \left\langle \tilde{\nabla}\varphi_m(\eta), X_i \right\rangle + u(\eta) \left\langle \tilde{\nabla}^2\varphi_m(\eta), X_i \otimes X_i \right\rangle \right] d\tau.$$

Now by construction of the φ_m and τ_m and the assumptions on u, there is a constant $M < \infty$ such that

$$1_{\{\tau\le t\wedge\tau_m\}} \left| \left\langle \tilde{\nabla}u(\eta_\tau), X_i \right\rangle \left\langle \tilde{\nabla}\varphi_m(\eta_\tau), X_i \right\rangle + u(\eta_\tau) \left\langle \tilde{\nabla}^2\varphi_m(\eta_\tau), X_i \otimes X_i \right\rangle \right| \le M.$$

Moreover, $\lim_{m\to\infty} \left|\tilde{\nabla}\varphi_m\right| = 0 = \lim_{m\to\infty} \left|\tilde{\nabla}^2\varphi_m\right|$. So it follows by the dominated convergence theorem that $\lim_{m\to\infty}\delta_m = 0$, and we have proved

$$1 = \lim_{m\to\infty} \mathbb{E}\left[\varphi_m(\eta_t^n)\right] = \mathbb{E}\left[\lim_{m\to\infty}\varphi_m(\eta_t^n)\right] = \mathbb{E}1_{\{t\wedge\tau_n<\zeta\}} = P(t\wedge\tau_n < \zeta),$$

for all $n = 1, 2, \ldots$. Thus, $1 = \lim_{n\to\infty} P(t\wedge\tau_n < \zeta) = P(t < \zeta)$, and so $\zeta = \infty$ μ-a.s. $\square$

Now let $u \in C_c^\infty(G)$, and suppose $\eta : [0,1]\times\mathscr{W}(\mathbb{R}^k) \to G$ is a solution to the stochastic differential equation

$$d\eta = u(\eta)\eta \circ d\vec{b} := u(\eta)L_{\eta*} \circ d\vec{b} = u(\eta)\tilde{X}_i(\eta) \circ db^i. \qquad (3.37)$$

Since the vector fields $u\tilde{X}_i$ have compact support, we may embed G as a Euclidean submanifold in a "nice" way so that the embedded vector fields are bounded with bounded derivatives. Then Theorem 2.1 of Taniguchi [47] implies that, for any $f \in C_c^\infty(G)$, we have $f(\eta_t) \in \mathcal{D}^\infty$.

Proposition 3.20. *Fix $h \in \mathscr{H}$, and let η be the solution to Equation (3.37). Then, for $f \in C_c^\infty(G)$,*

$$\partial_h f(\eta_t) = \left\langle \hat{\nabla} f(\eta_t), \theta_t \right\rangle,$$

where $\theta_t : \mathscr{W} \to \mathfrak{g}$ solves

$$\theta_t = \int_0^t \left(\left\langle \hat{\nabla} u(\eta), \theta \right\rangle \operatorname{Ad}_\eta \circ d\vec{b} + u(\eta) \operatorname{Ad}_\eta d\vec{h} \right), \tag{3.38}$$

where $d\vec{h} = X_i \dot{h}^i dt$; writing the above in Itô form gives

$$\theta_t = \int_0^t \left(\left\langle \hat{\nabla} u(\eta), \theta \right\rangle \operatorname{Ad}_\eta d\vec{b} + u(\eta) \operatorname{Ad}_\eta d\vec{h} + \frac{1}{2} \sum_{i=1}^k \left\langle \tilde{\nabla} \hat{\nabla} u(\eta), X_i \otimes \theta \right\rangle \operatorname{Ad}_\eta X_i \, d\tau \right). \tag{3.39}$$

Proof. Let $\eta_t^s := \eta_t(\cdot + sh) : [0,1]^2 \times \mathscr{W} \to G$. Then η_t^s satisfies the Stratonovich equation in t,

$$d\eta_t^s = u(\eta_t^s) L_{\eta_t^s *} \left(\circ d\vec{b}_t + s d\vec{h}_t \right). \tag{3.40}$$

By Corollary 4.3 of Driver [14], there exists a modification of η_t^s so that the mapping $s \mapsto \eta_t^s$ is smooth in the sense that, for any function $f \in C_c^\infty(G)$, $s \mapsto f(\eta_t^s)$ is smooth, and, furthermore, for any one-form ϑ acting on $T_{\eta_t^s} G$,

$$\left.\frac{\partial}{\partial s}\right|_0 \int_0^t \vartheta(d\eta_\tau^s) = \int_0^t d\vartheta \left(\left.\frac{\partial}{\partial s}\right|_0 \eta_\tau^s, d\eta_\tau \right) + \vartheta \left(\left.\frac{\partial}{\partial s}\right|_0 \eta_\tau^s \right) \Bigg|_{\tau=0}^t. \tag{3.41}$$

Let ϑ be the $\mathfrak{g}$-valued one-form such that $\vartheta(\hat{X}) = X$, for all $X \in \mathfrak{g}$. Since ϑ is right invariant, the two-form $d\vartheta$ satisfies the identity

$$d\vartheta = \vartheta \wedge \vartheta \tag{3.42}$$

where $\vartheta \wedge \vartheta(X,Y) := [\vartheta(X), \vartheta(Y)]$ for any $X, Y \in \mathfrak{g}$; see for example [1]. Let $\theta_t := \vartheta\left(\left.\frac{\partial}{\partial s}\right|_0 \eta_t^s \right)$, so that $\left.\frac{\partial}{\partial s}\right|_0 \eta_t^s = R_{\eta_t *} \theta_t$. Thus,

$$\left.\frac{\partial}{\partial s}\right|_0 \operatorname{Ad}_{\eta_t^s} = d \operatorname{Ad}(\theta_t) \operatorname{Ad}_{\eta_t} = \operatorname{ad}_{\theta_t} \operatorname{Ad}_{\eta_t}.$$

By Equation (3.41), we then have

$$\begin{aligned}
\theta_t = \vartheta\left(\frac{\partial}{\partial s}\bigg|_0 \eta_t^s\right) &= \frac{\partial}{\partial s}\bigg|_0 \int_0^t \vartheta(d\eta_\tau^s) - \int_0^t \left[\vartheta\left(\frac{\partial}{\partial s}\bigg|_0 \eta_t^s\right), \vartheta\left(d\eta_\tau\right)\right] \\
&= \frac{\partial}{\partial s}\bigg|_0 \int_0^t u(\eta_t^s)\operatorname{Ad}_{\eta_t^s}\left(\circ d\vec{b}_\tau + s d\vec{h}_\tau\right) - \int_0^t \left[\theta_\tau, u(\eta_t)\operatorname{Ad}_{\eta_t} \circ d\vec{b}_\tau\right] \\
&= \int_0^t \left(\left\langle \hat{\nabla} u(\eta_\tau), \theta_\tau\right\rangle \operatorname{Ad}_{\eta_\tau} \circ d\vec{b}_\tau + u(\eta_\tau)\operatorname{ad}_{\theta_\tau}\operatorname{Ad}_{\eta_\tau} \circ d\vec{b}_\tau + u(\eta_t)\operatorname{Ad}_{\eta_t} d\vec{h}_\tau\right) \\
&\quad - \int_0^t \left[\theta_\tau, u(\eta_t)\operatorname{Ad}_{\eta_t} \circ d\vec{b}_\tau\right] \\
&= \int_0^t \left(\left\langle \hat{\nabla} u(\eta_\tau), \theta_\tau\right\rangle \operatorname{Ad}_{\eta_\tau} \circ d\vec{b}_\tau + u(\eta_\tau)\operatorname{Ad}_{\eta_\tau} d\vec{h}_\tau\right),
\end{aligned}$$

and for any $f \in C_c^\infty(G)$,

$$\partial_h f(\eta_t) = \frac{\partial}{\partial s}\bigg|_0 f(\eta_t^s) = \left\langle \hat{\nabla} f(\eta_t), \vartheta\left(\frac{\partial}{\partial s}\bigg|_0 \eta_t^s\right)\right\rangle = \left\langle \hat{\nabla} f(\eta_t), \theta_t\right\rangle.$$

Now, to write this equation in Itô form, first note that

$$\begin{aligned}
d\operatorname{Ad}_\eta &= u(\eta)\operatorname{Ad}_\eta \circ \operatorname{ad}_{d\vec{b}} \\
&= u(\eta)\operatorname{Ad}_\eta \operatorname{ad}_{d\vec{b}} + \frac{1}{2}\left[\left\langle \tilde{\nabla} u(\eta), d\vec{b}\right\rangle \operatorname{Ad}_\eta + u^2(\eta)\operatorname{Ad}_\eta \operatorname{ad}_{d\vec{b}}\right] \cdot \operatorname{ad}_{d\vec{b}} \\
&= u(\eta)\operatorname{Ad}_\eta \operatorname{ad}_{d\vec{b}} + \frac{1}{2}\sum_{i=1}^k \left[\left\langle \tilde{\nabla} u(\eta), X_i\right\rangle \operatorname{Ad}_\eta \operatorname{ad}_{X_i} + u^2(\eta)\operatorname{Ad}_\eta \operatorname{ad}_{X_i}^2\right] dt,
\end{aligned} \tag{3.43}$$

where $\operatorname{ad}_{d\vec{b}} = \operatorname{ad}_{X_i} db^i$. This then implies

$$\begin{aligned}
d\left[\left\langle \hat{\nabla} u(\eta), \theta\right\rangle \operatorname{Ad}_\eta\right] \cdot d\vec{b} &= \left[\left\langle \tilde{\nabla}\hat{\nabla} u(\eta), d\vec{b} \otimes \theta\right\rangle \operatorname{Ad}_\eta + \left\langle \hat{\nabla} u(\eta), \theta\right\rangle u(\eta)\operatorname{Ad}_\eta \operatorname{ad}_{d\vec{b}}\right] d\vec{b} \\
&= \sum_{i=1}^k \left[\left\langle \tilde{\nabla}\hat{\nabla} u(\eta), X_i \otimes \theta\right\rangle \operatorname{Ad}_\eta + \left\langle \hat{\nabla} u(\eta), \theta\right\rangle u(\eta)\operatorname{Ad}_\eta \operatorname{ad}_{X_i}\right] X_i\, dt \\
&= \sum_{i=1}^k \left\langle \tilde{\nabla}\hat{\nabla} u(\eta), X_i \otimes \theta\right\rangle \operatorname{Ad}_\eta X_i\, dt,
\end{aligned}$$

wherein we have used $\operatorname{ad}_X X = [X, X] = 0$. Therefore,

$$\begin{aligned}\left\langle \hat{\nabla} u(\eta), \theta \right\rangle \operatorname{Ad}_\eta \circ d\vec{b} &= \left\langle \hat{\nabla} u(\eta), \theta \right\rangle \operatorname{Ad}_\eta d\vec{b} + \frac{1}{2} d \left[\left\langle \hat{\nabla} u(\eta), \theta \right\rangle \operatorname{Ad}_\eta \right] \cdot d\vec{b} \\ &= \left\langle \hat{\nabla} u(\eta), \theta \right\rangle \operatorname{Ad}_\eta d\vec{b} + \frac{1}{2} \sum_{i=1}^{k} \left\langle \tilde{\nabla} \hat{\nabla} u(\eta), X_i \otimes \theta \right\rangle \operatorname{Ad}_\eta X_i \, dt,\end{aligned}$$

and θ satisfies the Itô differential equation

$$d\theta = \left\langle \hat{\nabla} u(\eta), \theta \right\rangle \operatorname{Ad}_\eta d\vec{b} + u(\eta) \operatorname{Ad}_\eta d\vec{h} + \frac{1}{2} \sum_{i=1}^{k} \left\langle \tilde{\nabla} \hat{\nabla} v(\eta), X_i \otimes \theta \right\rangle \operatorname{Ad}_\eta X_i \, dt. \tag{3.44}$$

□

Notation 3.21. Let us now fix $\psi \in C_c^\infty(\operatorname{End}(\mathfrak{g}))$ with

$$\left\langle \psi'(x), A \right\rangle := \frac{d}{dt}\bigg|_0 \psi(x + tA),$$

for any $A \in \operatorname{End}(\mathfrak{g})$. Let $|g|$ denote the distance from a point $g \in G$ to e with respect to the right invariant metric, and let $\{\varphi_m\}_{m=1}^\infty \in C_c^\infty(G)$ be chosen so that $\varphi_m \uparrow 1$ as $m \to \infty$ such that $\varphi_m(g) = 1$ when $|g| \leq m$ and $\sup_m \sup_{g \in G} |\hat{\nabla}^k \varphi_m(g)| < \infty$ for $k = 0, 1, 2, \ldots$. Take $v(g) := \psi(\operatorname{Ad}_g)$ and $u_m(g) := \varphi_m(g) v(g)$.

Remark 3.22. For $u_m = \varphi_m v$, we have the derivative formulae

$$\hat{\nabla} u_m = \hat{\nabla} v \cdot \varphi_m + v \cdot \hat{\nabla} \varphi_m$$

and

$$\begin{aligned}\tilde{\nabla} \hat{\nabla} u_m &= \tilde{\nabla} \left[\hat{\nabla} v \cdot \varphi_m + v \cdot \hat{\nabla} \varphi_m \right] \\ &= \tilde{\nabla} \hat{\nabla} v \cdot \varphi_m + \tilde{\nabla} v \otimes \hat{\nabla} \varphi_m + \tilde{\nabla} \varphi_m \otimes \hat{\nabla} v + v \cdot \tilde{\nabla} \hat{\nabla} \varphi_m.\end{aligned}$$

Proposition 3.23. *Let $\eta^m : [0,1] \times \mathscr{W} \to G$ denote the solution to the equation*

$$\begin{aligned} d\eta^m &= u_m(\eta^m) \eta^m \circ d\vec{b} = \varphi_m(\eta^m) v(\eta^m) \eta^m \circ d\vec{b} \\ &= \varphi_m(\eta^m) \psi(\operatorname{Ad}_{\eta^m}) \eta^m \circ d\vec{b}, \ \text{with } \eta_0^m = e, \end{aligned} \tag{3.45}$$

and $\eta : [0,1] \times \mathscr{W} \to G$ *denote the solution to*

$$d\eta = v(\eta)\eta \circ d\vec{b} = \psi(\mathrm{Ad}_\eta)\eta \circ d\vec{b}, \text{ with } \eta_0 = e. \tag{3.46}$$

Then for all $f \in C_c^\infty(G)$*,*

$$\lim_{m\to\infty} \mathbb{E} \sup_{\tau \le 1} |f(\eta_\tau^m) - f(\eta_\tau)|^p = 0,$$

for all $p \in (1, \infty)$*.*

Proof. By Lemma 3.19, Equation (3.45) has a global solution. Notice also that $\tilde{X}v(\eta) = \langle \psi'(\mathrm{Ad}_\eta), \mathrm{ad}_X \rangle$ is bounded, so that Equation(3.46) also has a global solution. Then by Equation (3.36),

$$\begin{aligned} df(\eta^m) = {} & u_m(\eta^m) \left\langle \tilde{\nabla} f(\eta^m), d\vec{b} \right\rangle \\ & + \frac{1}{2} \sum_{i=1}^{k} \left[\left\langle \tilde{\nabla} u_m(\eta^m), X_i \right\rangle \left\langle \tilde{\nabla} f(\eta^m), X_i \right\rangle + u_m(\eta^m) \left\langle \tilde{\nabla}^2 f(\eta^m), X_i \otimes X_i \right\rangle \right] dt, \end{aligned}$$

and similarly,

$$\begin{aligned} df(\eta) = {} & v(\eta) \left\langle \tilde{\nabla} f(\eta), d\vec{b} \right\rangle \\ & + \frac{1}{2} \sum_{i=1}^{k} \left[\left\langle \tilde{\nabla} v(\eta), X_i \right\rangle \left\langle \tilde{\nabla} f(\eta), X_i \right\rangle + v(\eta) \left\langle \tilde{\nabla}^2 f(\eta), X_i \otimes X_i \right\rangle \right] dt. \end{aligned}$$

Thus,

$$\begin{aligned} d[f(\eta^m) - f(\eta)] = {} & u_m(\eta^m) \left\langle \tilde{\nabla} f(\eta^m), d\vec{b} \right\rangle - v(\eta) \left\langle \tilde{\nabla} f(\eta), d\vec{b} \right\rangle \\ & + \frac{1}{2} \sum_{i=1}^{k} \left[\left\langle \tilde{\nabla} u_m(\eta^m), X_i \right\rangle \left\langle \tilde{\nabla} f(\eta^m), X_i \right\rangle - \left\langle \tilde{\nabla} v(\eta), X_i \right\rangle \left\langle \tilde{\nabla} f(\eta), X_i \right\rangle \right. \\ & \left. + u_m(\eta^m) \left\langle \tilde{\nabla}^2 f(\eta^m), X_i \otimes X_i \right\rangle - v(\eta) \left\langle \tilde{\nabla}^2 f(\eta), X_i \otimes X_i \right\rangle \right] dt. \end{aligned}$$

We bound this expression by applying Proposition 3.11 to each term. For the first term, note that $u_m \to v$ boundedly, as $m \to \infty$, and this implies that

$$\mathbb{E}\left|\int_0^t \left[u_m(\eta^m)\langle\tilde{\nabla} f(\eta^m), d\vec{b}\rangle - v(\eta)\langle\tilde{\nabla} f(\eta), d\vec{b}\rangle\right]\right|^p$$

$$\lesssim \sum_{i=1}^k \mathbb{E}\int_0^t \left|u_m(\eta^m)\left\langle\tilde{\nabla} f(\eta^m), X_i\right\rangle - v(\eta)\left\langle\tilde{\nabla} f(\eta), X_i\right\rangle\right|^p d\tau$$

$$\lesssim \|\tilde{\nabla} f\|_\infty \mathbb{E}\int_0^t |u_m(\eta^m) - v(\eta)|^p \, d\tau \to 0,$$

as $m \to \infty$, by the dominated convergence theorem. Similarly, for the second term, $\tilde{\nabla} u_m \to \tilde{\nabla} v$ boundedly, as $m \to \infty$, implies that

$$\mathbb{E}\left|\int_0^t \left[\left\langle\tilde{\nabla} u_m(\eta^m), X_i\right\rangle\left\langle\tilde{\nabla} f(\eta^m), X_i\right\rangle - \left\langle\tilde{\nabla} v(\eta), X_i\right\rangle\left\langle\tilde{\nabla} f(\eta), X_i\right\rangle\right] d\tau\right|^p$$

$$\lesssim \mathbb{E}\int_0^t \left|\left\langle\tilde{\nabla} u_m(\eta^m), X_i\right\rangle\left\langle\tilde{\nabla} f(\eta^m), X_i\right\rangle - \left\langle\tilde{\nabla} v(\eta), X_i\right\rangle\left\langle\tilde{\nabla} f(\eta), X_i\right\rangle\right|^p d\tau$$

$$\lesssim \|\tilde{\nabla} f\|_\infty \mathbb{E}\int_0^t \left|\tilde{\nabla} u_m(\eta^m) - \tilde{\nabla} v(\eta)\right|^p d\tau \to 0,$$

as $m \to \infty$, by the dominated convergence theorem. Finally,

$$\mathbb{E}\left|\int_0^t \left[u_m(\eta^m)\left\langle\tilde{\nabla}^2 f(\eta^m), X_i \otimes X_i\right\rangle - v(\eta)\left\langle\tilde{\nabla}^2 f(\eta), X_i \otimes X_i\right\rangle\right] d\tau\right|^p$$

$$= \mathbb{E}\int_0^t \left|u_m(\eta^m)\left\langle\tilde{\nabla}^2 f(\eta^m), X_i \otimes X_i\right\rangle - v(\eta)\left\langle\tilde{\nabla}^2 f(\eta), X_i \otimes X_i\right\rangle\right|^p d\tau$$

$$\lesssim \|\tilde{\nabla}^2 f\|_\infty \mathbb{E}\int_0^t |u_m(\eta^m) - v(\eta)|^p \, d\tau \to 0,$$

as $m \to \infty$, again by dominated convergence. Thus,

$$\lim_{m\to\infty} \mathbb{E}\sup_{\tau\le 1} |f(\eta^m_\tau) - f(\eta_\tau)|^p = 0,$$

as desired. □

Proposition 3.24. *Let $U^m_t = \mathrm{Ad}_{\eta^m_t} : \mathscr{W} \to \mathrm{End}(\mathfrak{g})$ and $U_t = \mathrm{Ad}_{\eta_t} : \mathscr{W} \to \mathrm{End}(\mathfrak{g})$, which satisfy the stochastic differential equations*

$$dU^m = u_m(\eta^m)U^m \circ \mathrm{ad}_{d\vec{b}} = \varphi_m(\eta^m)\psi(U^m)U^m \circ \mathrm{ad}_{d\vec{b}}, \text{ with } U^m_0 = I, \tag{3.47}$$

and

$$dU = v(\eta)U \circ \mathrm{ad}_{d\vec{b}} = \psi(U)U \circ \mathrm{ad}_{d\vec{b}}, \text{ with } U_0 = I, \tag{3.48}$$

where $\mathrm{ad}_{d\vec{b}} = \mathrm{ad}_{X_i} db^i$. *Then*

$$\lim_{m\to\infty} \mathbb{E}\sup_{\tau\le 1} |U^m_\tau - U_\tau|^p = 0,$$

for all $p \in (1,\infty)$.

Proof. From Equation (3.43), we may rewrite(3.47) and (3.48) in Itô form as

$$\begin{aligned} dU^m &= u_m(\eta^m)U^m \mathrm{ad}_{d\vec{b}} + \frac{1}{2}\sum_{i=1}^k \left[\left\langle \tilde{\nabla} u_m(\eta^m), X_i\right\rangle U^m \mathrm{ad}_{X_i} + u_m^2(\eta^m)U^m \mathrm{ad}^2_{X_i}\right] dt \\ &= \varphi_m(\eta^m)\psi(U^m)U^m \mathrm{ad}_{d\vec{b}} + \frac{1}{2}\sum_{i=1}^k \Bigg[\left\langle \tilde{\nabla}\varphi_m(\eta^m), X_i\right\rangle \psi(U^m)U^m \mathrm{ad}_{X_i} \\ &\qquad + \varphi_m(\eta^m)\psi(U^m)\left\langle \psi'(U^m), U^m \mathrm{ad}_{X_i}\right\rangle U^m \mathrm{ad}_{X_i} + \varphi_m^2(\eta^m)\psi^2(U^m)U^m \mathrm{ad}^2_{X_i}\Bigg] dt \end{aligned}$$

and

$$\begin{aligned} dU &= v(\eta)U \mathrm{ad}_{d\vec{b}} + \frac{1}{2}\sum_{i=1}^k \left[\left\langle \tilde{\nabla} v(\eta), X_i\right\rangle U \mathrm{ad}_{X_i} + v^2(\eta)U \mathrm{ad}^2_{X_i}\right] dt \\ &= \psi(U)U \mathrm{ad}_{d\vec{b}} + \frac{1}{2}\sum_{i=1}^k \left[\psi(U)\left\langle \psi'(U), U \mathrm{ad}_{X_i}\right\rangle U \mathrm{ad}_{X_i} + \psi^2(U)U \mathrm{ad}^2_{X_i}\right] dt. \end{aligned}$$

Let $\varepsilon^m := U^m - U$. Then by the above,

$$\begin{aligned} d\varepsilon^m &= (\varphi_m(\eta^m)\psi(U^m)U^m - \psi(U)U)\,\mathrm{ad}_{d\vec{b}} + \frac{1}{2}\sum_{i=1}^k \Bigg[\left\langle \tilde{\nabla}\varphi_m(\eta^m), X_i\right\rangle \psi(U^m)U^m \mathrm{ad}_{X_i} \\ &\quad + \big(\varphi_m(\eta^m)\psi(U^m)\left\langle \psi'(U^m), U^m \mathrm{ad}_{X_i}\right\rangle U^m - \psi(U)\left\langle \psi'(U), U \mathrm{ad}_{X_i}\right\rangle U\big)\,\mathrm{ad}_{X_i} \\ &\quad + \big(\varphi_m^2(\eta^m)\psi^2(U^m)U^m - \psi^2(U)U\big)\,\mathrm{ad}^2_{X_i}\Bigg] dt. \end{aligned} \tag{3.49}$$

Again applying Proposition 3.11, we work term by term to bound the above expression. Note first that, since ψ has compact support, U and U^m always remain in a **fixed** compact subset of End($\mathfrak{g}$). Thus,

$$\mathbb{E}\left|\int_0^t [\varphi_m(\eta^m)\psi(U^m)U^m - \psi(U)U]\,\mathrm{ad}_{d\vec{b}}\right|^p$$
$$\lesssim \mathbb{E}\int_0^t |\varphi_m(\eta^m)\psi(U^m)U^m - \psi(U)U|^p\,d\tau$$
$$\lesssim \mathbb{E}\int_0^t |\psi(U^m)U^m - \psi(U)U|^p\,d\tau + \mathbb{E}\int_0^t |\varphi_m(\eta^m) - 1|^p|\psi(U^m)|^p\,d\tau$$
$$\lesssim \mathbb{E}\int_0^t |\varepsilon^m|^p\,d\tau + \delta_m,$$

wherein we have applied the mean value inequality to $x \mapsto \psi(x)x$ to learn

$$|\psi(U^m)U^m - \psi(U)U| \le C(\psi)|U^m - U| = C|\varepsilon^m|,$$

and

$$\delta_m = \mathbb{E}\int_0^t |\varphi_m(\eta^m) - 1|^p|\psi(U^m)|^p\,d\tau \to 0,$$

as $m \to \infty$, by the dominated convergence theorem. Similarly, for the last term of the sum in (3.49),

$$\mathbb{E}\left|\int_0^t \left(\varphi_m^2(g^n)\psi^2(U^m)U^m - \psi^2(U)U\right)\mathrm{ad}_{X_i}^2\;d\tau\right|^p$$
$$\lesssim \mathbb{E}\int_0^t \left|\psi^2(U^m)U^m - \psi^2(U)U\right|^p d\tau + \mathbb{E}\int_0^t |\varphi_m^2(\eta^m) - 1|^p|\psi(U^m)|^{2p}\,d\tau$$
$$\lesssim \mathbb{E}\int_0^t |\varepsilon^m|^p\,d\tau + \delta_m,$$

where the mean value inequality has now been applied to the function $x \mapsto \psi^2(x)x$, and

$$\delta_m = \mathbb{E}\int_0^t |\varphi_m^2(\eta^m) - 1|^p|\psi(U^m)|^{2p}\,d\tau \to 0,$$

as $m \to \infty$. For the second term,

$$\mathbb{E}\left|\int_0^t \left\langle\tilde{\nabla}\varphi_m(\eta^m), X_i\right\rangle \psi(U^m)U^m\,\mathrm{ad}_{X_i}\,d\tau\right|^p \lesssim \mathbb{E}\int_0^t \left|\left\langle\tilde{\nabla}\varphi_m(\eta^m), X_i\right\rangle \psi(U^m)U^m\right|^p\,d\tau$$
$$= \mathbb{E}\int_0^t \left|\left\langle\tilde{\nabla}\varphi_m(\eta^m), X_i\right\rangle \psi(U^m)(\varepsilon^m + U)\right|^p\,d\tau$$
$$\lesssim \mathbb{E}\int_0^t |\varepsilon^m|^p\,d\tau + \delta_m,$$

where

$$\delta_m = \mathbb{E}\int_0^t \left|\left\langle \tilde{\nabla}\varphi_m(\eta^m), X_i\right\rangle \psi(U^m)U\right|^p \, d\tau \to 0,$$

as $m \to \infty$, since $\lim_{m\to\infty} |\tilde{\nabla}\varphi_m| = 0$. Finally, for the the third term, note first that

$$\begin{aligned}
&\varphi_m(g^n)\langle \psi'(U^m), U^m \operatorname{ad}_{X_i}\rangle U^m \\
&\qquad = \varphi_m(\eta^m)\left\langle \psi'(U^m), U^m \operatorname{ad}_{X_i}\right\rangle (\varepsilon^m + U) \\
&\qquad = \varphi_m(\eta^m)\left\langle \psi'(U^m), U^m \operatorname{ad}_{X_i}\right\rangle \varepsilon^m + \varphi_m(\eta^m)\left\langle \psi'(U^m), (\varepsilon^m + U)\operatorname{ad}_{X_i}\right\rangle U \\
&\qquad = \varphi_m(\eta^m)\left\langle \psi'(U^m), U^m \operatorname{ad}_{X_i}\right\rangle \varepsilon^m + \varphi_m(\eta^m)\left\langle \psi'(U^m), \varepsilon^m \operatorname{ad}_{X_i}\right\rangle U \\
&\qquad\quad + \varphi_m(\eta^m)\left\langle \psi'(U^m), U \operatorname{ad}_{X_i}\right\rangle U.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathbb{E}\left|\int_0^t (\varphi_m(\eta^m)\psi(U^m)\left\langle \psi'(U^m), U^m \operatorname{ad}_{X_i}\right\rangle U^m - \psi(U)\left\langle \psi'(U), U \operatorname{ad}_{X_i}\right\rangle U)\operatorname{ad}_{X_i}\, d\tau\right|^p \\
&\qquad \lesssim \mathbb{E}\int_0^t \left|\varphi_m(\eta^m)\left\langle \psi'(U^m), U^m \operatorname{ad}_{X_i}\right\rangle \varepsilon^m + \varphi_m(\eta^m)\left\langle \psi'(U^m), \varepsilon^m \operatorname{ad}_{X_i}\right\rangle U \right. \\
&\qquad\qquad \left. + \varphi_m(\eta^m)\left\langle \psi'(U^m), U \operatorname{ad}_{X_i}\right\rangle U - \left\langle \psi'(U), U \operatorname{ad}_{X_i}\right\rangle U\right|^p \, d\tau \\
&\qquad \lesssim \mathbb{E}\int_0^t |\varepsilon^m|^p \, d\tau + \delta_m,
\end{aligned}$$

where

$$\delta_m = \mathbb{E}\int_0^t |\varphi_m(\eta^m)\left\langle \psi'(U^m), U \operatorname{ad}_{X_i}\right\rangle - \left\langle \psi'(U), U \operatorname{ad}_{X_i}\right\rangle|^p |U|^p \, d\tau \to 0,$$

as $m \to \infty$, since $\varphi_m(\eta^m)\psi'(U^m) \to \psi'(U)$ boundedly. These bounds then imply that

$$\mathbb{E}\sup_{\tau\le t} |\varepsilon_\tau^m|^p \le C\int_0^t |\varepsilon^m|^p \, d\tau + \delta_m,$$

for all $t \in [0,1]$. Thus, by Gronwall's inequality, we have

$$\mathbb{E}\sup_{\tau\le 1} |U_\tau^m - U_\tau|^p = \mathbb{E}\sup_{\tau\le 1} |\varepsilon_\tau^m|^p \le \delta_m e^C \to 0,$$

as $m \to \infty$. □

Proposition 3.25. *Let $\theta_t^m : \mathscr{W} \to \mathfrak{g}$ be as in Equation (3.38) with u replaced by u_m; that is,*

$$\theta_t^m = \int_0^t \left(\left\langle \hat{\nabla} u_m(\eta^m), \theta \right\rangle \operatorname{Ad}_{\eta^m} \circ d\vec{b} + u_m(\eta^m) \operatorname{Ad}_{\eta^m} d\vec{h} \right). \tag{3.50}$$

Then

$$\lim_{m\to\infty} \mathbb{E} \sup_{\tau \le 1} |\theta_\tau^m - \theta_\tau|^p = 0,$$

for all $p \in (1, \infty)$, where $\theta_t : \mathscr{W} \to \mathfrak{g}$ is the solution to

$$\theta_t = \int_0^t \left(\left\langle \hat{\nabla} v(\eta), \theta \right\rangle \operatorname{Ad}_\eta \circ d\vec{b} + v(\eta) \operatorname{Ad}_\eta \, d\vec{h} \right) \tag{3.51}$$

Proof. Let $U_t^m = \operatorname{Ad}_{\eta_t^m}$ and $U_t = \operatorname{Ad}_{\eta_t}$. We rewrite Equation (3.50) in Itô form as

$$d\theta^m = \left\langle \hat{\nabla} u_m(\eta^m), \theta^m \right\rangle U^m d\vec{b} + u_m(\eta^m) U^m d\vec{h} + \frac{1}{2} \sum_{i=1}^k \left\langle \tilde{\nabla} \hat{\nabla} u_m\left(\eta^m\right), X_i \otimes \theta^m \right\rangle U^m X_i \, dt.$$

Note that, formally, θ is the solution to Equation (3.38) with u replaced by v (although v is not a function with compact support), and we may rewrite Equation (3.51) in Itô form as

$$\begin{aligned} d\theta &= \left\langle \hat{\nabla} v(\eta), \theta \right\rangle \operatorname{Ad}_\eta \, d\vec{b} + v(\eta) \operatorname{Ad}_\eta \, d\vec{h} + \frac{1}{2} \sum_{i=1}^k \left\langle \tilde{\nabla} \hat{\nabla} v(\eta), X_i \otimes \theta \right\rangle \operatorname{Ad}_\eta X_i \, dt, \\ &= \left\langle \hat{\nabla} v(\eta), \theta \right\rangle U \, d\vec{b} + v(\eta) U \, d\vec{h} + \frac{1}{2} \sum_{i=1}^k \left\langle \tilde{\nabla} \hat{\nabla} v(\eta), X_i \otimes \theta \right\rangle U X_i \, dt. \end{aligned}$$

Let $\varepsilon^m := \theta^m - \theta$ (so that $\theta^m = \varepsilon^m + \theta$). Then

$$\begin{aligned} d\varepsilon^m = {} & \left[\left\langle \hat{\nabla} u_m(\eta^m), \theta^m \right\rangle U^m - \left\langle \hat{\nabla} v(\eta), \theta \right\rangle U \right] d\vec{b} + \left[u_m(\eta^m) U^m - v(\eta) U \right] d\vec{h} \\ & + \frac{1}{2} \sum_{i=1}^k \left[\left\langle \tilde{\nabla} \hat{\nabla} u_m(\eta^m), X_i \otimes \theta^m \right\rangle U^m - \left\langle \tilde{\nabla} \hat{\nabla} v(\eta), X_i \otimes \theta \right\rangle U \right] X_i \, dt. \end{aligned}$$

Considering the first term of this expression, we have

$$\begin{aligned} \left\langle \hat{\nabla} u_m(\eta^m), \theta^m \right\rangle U^m - & \left\langle \hat{\nabla} v(\eta), \theta \right\rangle U \\ &= \langle \hat{\nabla} u_m(\eta^m), \theta + \varepsilon^m \rangle U^m - \left\langle \hat{\nabla} v(\eta), \theta \right\rangle U \\ &= \left\langle \hat{\nabla} u_m(\eta^m), \varepsilon^m \right\rangle U^m + \left\langle \hat{\nabla} u_m(\eta^m), \theta \right\rangle U^m - \left\langle \hat{\nabla} v(\eta), \theta \right\rangle U. \end{aligned}$$

Using again the U and U^m remain in a fixed compact subset of $\mathrm{End}(\mathfrak{g})$ and the fact that $\hat{\nabla} u_(\eta^m) \to \hat{\nabla} v(\eta)$ boundedly, we have

$$\mathbb{E}\left|\int_0^t \left[\left\langle \hat{\nabla} u_m(\eta^m), \theta^m \right\rangle U^m - \left\langle \hat{\nabla} v(\eta), \theta \right\rangle U\right] d\vec{b}\right|^p \lesssim \mathbb{E}\int_0^t |\varepsilon^m|^p \, d\tau + \delta_m,$$

where

$$\delta_m = \mathbb{E}\int_0^t \left|\left\langle \hat{\nabla} u_m(\eta^m), \theta \right\rangle U^m - \left\langle \hat{\nabla} v(\eta), \theta \right\rangle U\right| d\tau \to 0,$$

as $m \to \infty$. The second term converges to 0 since

$$\mathbb{E}\left|\int_0^t \left[u_m(\eta^m)U^m - v(\eta)U\right] d\vec{h}\right|^p = \mathbb{E}\left|\int_0^t \left[\varphi_m(\eta^m)v(\eta^m)U^m - v(\eta)U\right] d\vec{h}\right|^p \to 0,$$

as $m \to \infty$, by the dominated convergence theorem. For the third term, note that

$$\begin{aligned}\Big|\Big[\Big\langle \tilde{\nabla}\hat{\nabla} u_m(\eta^m), X_i \otimes (\theta + \varepsilon^m) \Big\rangle U^m - \Big\langle \tilde{\nabla}\hat{\nabla} v(\eta), X_i \otimes \theta \Big\rangle U\Big] X_i\Big| \\ \lesssim |\varepsilon^m| + \left|\left[\left\langle \tilde{\nabla}\hat{\nabla} u_m(\eta^m), X_i \otimes \theta \right\rangle U^m - \left\langle \tilde{\nabla}\hat{\nabla} v(\eta), X_i \otimes \theta \right\rangle U\right] X_i\right|,\end{aligned}$$

and so

$$\begin{aligned}\mathbb{E}\left|\int_0^t \left[\left\langle \tilde{\nabla}\hat{\nabla} u_m(\eta^m), X_i \otimes \theta^m \right\rangle U^m - \left\langle \tilde{\nabla}\hat{\nabla} v(\eta), X_i \otimes \theta \right\rangle U\right] X_i \, d\tau\right|^p \\ \lesssim \mathbb{E}\int_0^t |\varepsilon^m|^p \, d\tau + \delta_m,\end{aligned}$$

where

$$\delta_m = \mathbb{E}\left|\int_0^t \left[\left\langle \tilde{\nabla}\hat{\nabla} u_m(\eta^m), X_i \otimes \theta \right\rangle U^m - \left\langle \tilde{\nabla}\hat{\nabla} v(\eta), X_i \otimes \theta \right\rangle U\right] X_i \, d\tau\right|^p \to 0,$$

as $m \to \infty$, where we have used that $\tilde{\nabla}\hat{\nabla} u_m(\eta^m) \to \tilde{\nabla}\hat{\nabla} v(\eta)$ boundedly to apply the dominated convergence theorem.

Putting these bounds together then shows

$$\mathbb{E}\sup_{\tau \le t} |\varepsilon^m_\tau|^p \le C\mathbb{E}\int_0^t |\varepsilon^m|^p \, d\tau + \delta_m,$$

for all $t \in [0,1]$, and again applying Gronwall's inequality gives

$$\mathbb{E}\sup_{\tau \le 1} |\theta^m_\tau - \theta_\tau|^p = \mathbb{E}\sup_{\tau \le 1} |\varepsilon^m_\tau|^p \le \delta_m e^C \to 0,$$

as $m \to \infty$, finishes the proof. □

Proposition 3.26. *Let η^m be the solution to Equation (3.45) and θ^m be the solution to Equation (3.50). Then, for any $h \in \mathscr{H}$, $f \in C_c^\infty(G)$, and $t \in [0,1]$,*

$$\partial_h f(\eta_t^m) = \left\langle \hat{\nabla} f(\eta_t^m), \theta_t^m \right\rangle.$$

Furthermore, for η the solution to (3.46) and θ the solution to (3.51),

$$\lim_{m\to\infty} \mathbb{E} \left| \partial_h f(\eta_t^m) - \left\langle \hat{\nabla} f(\eta_t), \theta_t \right\rangle \right|^p = 0,$$

for all $p \in (1, \infty)$.

Proof. The first claim follows immediately from Proposition 3.20. Now note that

$$\begin{aligned} \left|\partial_h f(\eta^m) - \left\langle \hat{\nabla} f(\eta), \theta \right\rangle\right| &= \left| \left\langle \hat{\nabla} f(\eta^m), \theta^m \right\rangle - \left\langle \hat{\nabla} f(\eta), \theta \right\rangle \right| \\ &\leq \left| \left\langle \hat{\nabla} f(\eta^m), \theta^m \right\rangle - \left\langle \hat{\nabla} f(\eta^m), \theta \right\rangle \right| + \left| \left\langle \hat{\nabla} f(\eta^m), \theta \right\rangle - \left\langle \hat{\nabla} f(\eta), \theta \right\rangle \right| \\ &\leq |\hat{\nabla} f||\theta^m - \theta| + |\hat{\nabla} f(\eta^m) - \hat{\nabla} f(\eta)||\theta|. \end{aligned}$$

Thus,

$$\lim_{m\to\infty} \mathbb{E} \left| \partial_h f(\eta^m) - \left\langle \hat{\nabla} f(\eta), \theta \right\rangle \right|^p \leq \lim_{m\to\infty} \mathbb{E} \left[|\hat{\nabla} f||\theta^m - \theta| + |\hat{\nabla} f(\eta^m) - \hat{\nabla} f(\eta)||\theta| \right]^p = 0,$$

by Propositions 3.23 and 3.25 and the dominated convergence theorem. □

Corollary 3.27. *For any $h \in \mathscr{H}$, $f \in C_c^\infty(G)$, and $t \in [0,1]$, $f(\eta_t) \in \mathrm{Dom}(\partial_h)$ and*

$$\partial_h f(\eta_t) = \left\langle \hat{\nabla} f(\eta_t), \theta_t \right\rangle \in L^{\infty-}(\mu).$$

This corollary follows from ∂_h being a closed operator (we are thinking here that $\partial_h = \overline{\partial}_h$). Now we are able to remove the cutoff functions and prove the primary result of this section.

Theorem 3.28. *Let $\xi : [0,1] \times \mathscr{W} \to G$ denote the solution to Equation (3.3), and let $\Theta : \mathscr{W} \to \mathfrak{g}$ be the solution to*

$$\Theta_t := \int_0^t \mathrm{Ad}_\xi \, d\vec{h} = \int_0^t \mathrm{Ad}_\xi \, X_i \dot{h}^i \, d\tau. \tag{3.52}$$

Then, for any $f \in C_c^\infty(G)$ *and* $t \in [0,1]$, $f(\xi_t) \in \mathcal{D}^{1,\infty}$, *and for any* $h \in \mathscr{H}$,

$$\partial_h[f(\xi_t)] = \int_0^t \left\langle \hat{\nabla} f(\xi_t), \operatorname{Ad}_{\xi_\tau} X_i \right\rangle \dot{h}_\tau^i \, d\tau. \tag{3.53}$$

Thus,

$$(D[f(\xi_t)])^i = \int_0^{\cdot\wedge t} \left\langle \hat{\nabla} f(\xi_t), \operatorname{Ad}_{\xi_\tau} X_i \right\rangle d\tau,$$

componentwise in $\mathscr{H}$.

Proof. Let $\psi \in C_c^\infty(\operatorname{End}(\mathfrak{g}), [0,1])$ be chosen so that $\psi = 1$ near I and $\psi(x) = 0$ if $|x| \geq 2$, where $|\cdot|$ is the distance from I with respect to any metric on $\operatorname{End}(\mathfrak{g})$. Define $v_n(x) := \psi(n^{-1} \operatorname{Ad}_x)$, and let $\xi^n : [0,1] \times \mathscr{W} \to G$ be the solution to the Stratonovich equation

$$d\xi^n = v_n(\xi^n)\xi^n \circ d\vec{b}, \text{ with } \xi_0^n = e.$$

By Lemma 3.19, ξ^n exists for all time $t \in [0,1]$. Noting that $v_n \to 1$ and $\tilde{\nabla} v_n \to 0$ boundedly as $n \to \infty$, by an argument identical to that in Proposition 3.23, we may show that, for all $f \in C_c^\infty(G)$,

$$\lim_{n\to\infty} \mathbb{E} \sup_{\tau \leq 1} |f(\xi_\tau^n) - f(\xi_\tau)|^p = 0,$$

for all $p \in (1, \infty)$. Note that this convergence implies that $f(\xi_t) \in L^{\infty-}(\mu)$.

Now let $\Theta_t^n : \mathscr{W} \to \mathfrak{g}$ denote the solution to the Itô equation

$$d\Theta^n = \left\langle \hat{\nabla} v_n(\xi^n), \Theta^n \right\rangle \operatorname{Ad}_{\xi^n} d\vec{b} + v_n(\xi^n) \operatorname{Ad}_{\xi^n} d\vec{h} + \frac{1}{2} \sum_{i=1}^k \left\langle \tilde{\nabla}\hat{\nabla} v_n(\xi^n), X_i \otimes \Theta^n \right\rangle \operatorname{Ad}_{\xi^n} X_i \, dt, \tag{3.54}$$

with $\Theta_0^n = 0$. We will now show that

$$\lim_{n\to\infty} \mathbb{E} \sup_{\tau \leq 1} |\Theta_\tau^n - \Theta_\tau|^p = 0, \tag{3.55}$$

for all $p \in (1, \infty)$. So let $W^n = \operatorname{Ad}_{\xi^n}$ and $W = \operatorname{Ad}_\xi$. Then $W^n, W : \mathscr{W} \to \operatorname{End}(\mathfrak{g})$ satisfy the equations

$$dW^n = v_n(\xi^n) W^n \circ \operatorname{ad}_{d\vec{b}}, \text{ with } W_0^n = I,$$

and

$$dW = W \circ \mathrm{ad}_{d\vec{b}}, \text{ with } W_0 = I.$$

Rewriting Equation (3.54), Θ^n solves

$$d\Theta^n = \left\langle \hat{\nabla} v_n \left(\xi^n\right), \Theta^n \right\rangle W^n d\vec{b} + v_n(\xi^n) W^n \, d\vec{h} + \frac{1}{2} \sum_{i=1}^{k} \left\langle \tilde{\nabla}\hat{\nabla} v_n(\xi^n), X_i \otimes \Theta^n \right\rangle W^n X_i \, dt,$$

and Equation (3.52) implies that

$$d\Theta = W \, d\vec{h}.$$

Thus, for $\varepsilon^n := \Theta^n - \Theta$,

$$d\varepsilon^n = \left\langle \hat{\nabla} v_n \left(\xi^n\right), \Theta^n \right\rangle W^n d\vec{b} + (v_n(\xi^n) W^n - W) d\vec{h} + \frac{1}{2} \sum_{i=1}^{k} \left\langle \tilde{\nabla}\hat{\nabla} v_n(\xi^n), X_i \otimes \Theta^n \right\rangle W^n X_i \, dt. \quad (3.56)$$

Since W and W^n are the solutions to Equations (3.16) and (3.28), which have smooth, bounded coefficients, Theorem V-10.1 of Ikeda and Watanabe [30] implies that W and W^n are (componentwise) in the domain $\mathcal{D}^\infty$. By Propositions 3.16 and 3.18,

$$\lim_{n\to\infty} \mathbb{E} \sup_{\tau \le 1} |W_\tau^n - W_\tau|^p \quad (3.57)$$

and

$$\lim_{n\to\infty} \mathbb{E} \sup_{\tau \le 1} |\partial_h W_\tau^n - \partial_h W_\tau|^p, \quad (3.58)$$

for all $p \in (1, \infty)$. Furthermore,

$$\partial_h W^n = \partial_h \mathrm{Ad}_{\xi^n} = \mathrm{ad}_{\Theta^n} \mathrm{Ad}_{\xi^n} = \mathrm{ad}_{\Theta^n} W^n.$$

This then implies that

$$\begin{aligned} \left\langle \hat{\nabla} v_n(\xi^n), \Theta^n \right\rangle &= \frac{d}{dt}\bigg|_0 \psi(n^{-1} \mathrm{Ad}_{e^{t\Theta^n}\xi^n}) \\ &= \frac{1}{n} \left\langle \psi'(n^{-1} \mathrm{Ad}_{\xi^n}), \mathrm{ad}_{\Theta^n} \mathrm{Ad}_{\xi^n} \right\rangle = \frac{1}{n} \left\langle \psi'(n^{-1} W^n), \partial_h W^n \right\rangle. \end{aligned}$$

Thus, for the first term of Equation (3.56),

$$\begin{aligned}
\mathbb{E}\left|\int_0^t \left\langle \hat{\nabla} v_n(\xi^n), \Theta^n \right\rangle W^n \, d\vec{b}\right|^p &\lesssim \mathbb{E}\int_0^t \frac{1}{n}\left|\left\langle \psi'(n^{-1}W^n), \partial_h W^n \right\rangle W^n\right|^p \, d\tau \\
&\lesssim \mathbb{E}\int_0^t \left(\frac{1}{n}\left|\psi'(n^{-1}W^n)\right| |W^n|\right)^p |\partial_h W^n|^p \, d\tau \\
&\lesssim \mathbb{E}\int_0^t \left(\frac{1}{n}\left|\psi'(n^{-1}W^n)\right| |W^n|\right)^p |\partial_h W|^p \, d\tau \\
&\quad + \mathbb{E}\int_0^t \left(\frac{1}{n}\left|\psi'(n^{-1}W^n)\right| |W^n|\right)^p |\partial_h W^n - \partial_h W|^p \, d\tau \to 0,
\end{aligned}$$

as $n \to \infty$, by the dominated convergence theorem. Similarly,

$$\begin{aligned}
\left\langle \tilde{\nabla}\hat{\nabla} v_n(\xi^n), X_i \otimes \Theta^n \right\rangle &= \frac{1}{n^2}\left\langle \psi''(n^{-1}W^n), W^n \operatorname{ad}_{X_i} \otimes \operatorname{ad}_{\Theta^n} \operatorname{Ad}_{\xi^n} \right\rangle \\
&= \frac{1}{n^2}\left\langle \psi''(n^{-1}W^n), W^n \operatorname{ad}_{X_i} \otimes \partial_h W^n \right\rangle,
\end{aligned}$$

so that

$$\begin{aligned}
&\mathbb{E}\left|\int_0^t \left\langle \tilde{\nabla}\hat{\nabla} v_n(\xi^n), X_i \otimes \Theta^n \right\rangle W^n X_i \, d\tau\right|^p \\
&\qquad = \mathbb{E}\left|\int_0^t \frac{1}{n^2}\left\langle \psi''(n^{-1}W^n), W^n \operatorname{ad}_{X_i} \otimes \partial_h W^n \right\rangle W^n X_i\right|^p d\tau \\
&\qquad \lesssim \mathbb{E}\int_0^t \left(\frac{1}{n^2}\left|\psi''(n^{-1}W^n)\right| |W^n|^2\right)^p |\partial_h W^n|^p \, d\tau \\
&\qquad \lesssim \mathbb{E}\int_0^t \left(\frac{1}{n^2}\left|\psi''(n^{-1}W^n)\right| |W^n|^2\right)^p |\partial_h W|^p \, d\tau \\
&\qquad\quad + \mathbb{E}\int_0^t \left(\frac{1}{n^2}\left|\psi''(n^{-1}W^n)\right| |W^n|^2\right)^p |\partial_h W^n - \partial_h W|^p \, d\tau \to 0,
\end{aligned}$$

as $n \to \infty$, again by dominated convergence. Finally,

$$\begin{aligned}
&\mathbb{E}\left|\int_0^t (v_n(\xi^n) W^n - W) \, d\vec{h}\right|^p \\
&\qquad \lesssim \mathbb{E}\int_0^t |v_n(\xi^n) - 1|^p |W|^p \, d\vec{h} + \mathbb{E}\int_0^t |v_n(\xi^n)|^p \, |W^n - W|^p \, d\vec{h} \to 0
\end{aligned}$$

as $n \to \infty$, by the dominated convergence theorem. Thus, Proposition 3.11 completes the proof of Equation (3.55).

Now, by Corollary 3.27, for any $f \in C_c^\infty(G)$ and $t \in [0,1]$, $\partial_h[f(\xi_t^n)] \in L^{\infty-}(\mu)$, and

$$\partial_h[f(\xi_t^n)] = \left\langle \hat{\nabla} f(\xi_t^n), \Theta_t^n \right\rangle.$$

Thus, by the same argument as in Proposition 3.26, Equation (3.55) implies that

$$\lim_{n\to\infty} \mathbb{E}\left|\partial_h f(\xi_t^n) - \left\langle \hat{\nabla} f(\xi_t), \Theta_t \right\rangle\right|^p = 0,$$

for all $p \in (1,\infty)$. Since ∂_h is a closed operator, this and $f(\xi_t) \in L^{\infty-}(\mu)$ imply that $f(\xi_t) \in \mathrm{Dom}(\partial_h)$ and

$$\partial_h[f(\xi_t)] = \left\langle \hat{\nabla} f(\xi_t), \Theta_t \right\rangle = \left\langle \hat{\nabla} f(\xi_t), \int_0^t \mathrm{Ad}_\xi\, X_i \dot{h}^i \, d\tau \right\rangle \in L^{\infty-}(\mu).$$

In particular, for any $h \in \mathscr{H}$ such that $\|h\|_{\mathscr{H}} = 1$,

$$\mathbb{E}|\partial_h[f(\xi_t)]|^p \leq \|\hat{\nabla} f\|_\infty \mathbb{E}\left|\int_0^t \mathrm{Ad}_\xi\, X_i \dot{h}^i \, d\tau\right|^p \leq \|\hat{\nabla} f\|_\infty \sum_{i=1}^k \mathbb{E}\int_0^t |\mathrm{Ad}_\xi\, X_i|^p \, d\tau.$$

By the proof of Proposition 3.13,

$$\mathbb{E}\sup_{\tau\leq 1} |\,\mathrm{Ad}_{\xi_\tau}|^p \leq C,$$

for some finite constant C, and thus $\|f(\xi_t)\|_{G^{1,p}} < \infty$ for all $p \in (1,\infty)$, where

$$\|F\|_{G^{1,p}} = \|F\|_{L^p(\mu)} + \left(\mathbb{E}\sup_{\|h\|_{\mathscr{H}}=1} |\partial_h F|^p\right)^{1/p};$$

see Section 2.5. Then by Theorem 2.12, $f(\xi_t) \in \mathcal{D}^{1,\infty}$, and

$$\mathbb{E}[f(\xi_t) D^* h] = \mathbb{E}[\partial_h[f(\xi_t)]] = \mathbb{E}\left[\int_0^t \left\langle \hat{\nabla} f(\xi_t), \mathrm{Ad}_\xi\, X_i \right\rangle \dot{h}^i \, d\tau\right],$$

implies that

$$(D_s[f(\xi_t)])^i = \left\langle \hat{\nabla} f(\xi_t), \int_0^{s\wedge t} \mathrm{Ad}_\xi\, X_i \, d\tau \right\rangle,$$

componentwise in $\mathscr{H}$. ∎

Chapter 4

Lie group inequalities

Again, let G be a Lie group with identity e and Lie algebra $\mathrm{Lie}(G) = \mathfrak{g}$, and suppose $\{X_i\}_{i=1}^k \subset \mathfrak{g}$ is a Lie generating set, in the sense of Equation (1.7). We have the gradient $\nabla = (\tilde{X}_1, \dots, \tilde{X}_k)$ and the subLaplacian $L = \sum_{i=1}^k \tilde{X}_i^2$ as operators on smooth functions of G with compact support. Let L also denote the self-adjoint extension of the subLaplacian and $P_t = e^{tL/2}$ be the heat semigroup as in Definition 1.8.

We recall the following lemmas, proved in Chapter 2 in the Heisenberg context (Lemmas 2.4 and 2.5). The proofs are identical in the general Lie group case.

Lemma 4.1. *By the left invariance of ∇ and P_t, the inequality (I_p) holds for all $g \in G$, $f \in C_c^\infty(G)$, and $t > 0$, if and only if,*

$$|\nabla P_t f|^p(e) \leq K_p(t) P_t |\nabla f|^p(e), \tag{4.1}$$

for all $f \in C_p^\infty(G)$ and $t > 0$, where $e \in G$ is the identity element.

Lemma 4.2. *For $X \in \mathfrak{g}$,*

$$\tilde{X} P_t f(e) = P_t \hat{X} f(e).$$

for all $f \in C_c^\infty(G)$. More generally,

$$\hat{X} P_t f = P_t \hat{X} f,$$

from which the previous equation follows, since $\hat{X} = \tilde{X}$ at e.

(The proof of this statement is slightly easier than its analogue Lemma 2.5, since working with functions with compact support – versus functions with polynomial growth – requires only the invariance of Haar measure to justify passing the derivative through the integral.)

4.1 An L^p-type gradient estimate $(p > 1)$

Notation 4.3. For each $r \in \{1, \ldots, m\}$, let $\Lambda^r = \Lambda^{k,r}$ be the set of multi-indices $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_r)$, with $\alpha_l \in \{1, \ldots, k\}$ for $l \in \{0, \ldots, r\}$. For any $\alpha \in \Lambda^r$, define

$$\alpha' := (\alpha_1, \ldots, \alpha_r) \text{ and}$$

$$\overline{\alpha} := (\alpha_r, \ldots, \alpha_0) = \alpha \text{ reversed .}$$

We define the order of α by $|\alpha| := r + 1$. Let

$$X_\alpha = [X_{\alpha_r}, [\cdots, [X_{\alpha_1}, X_{\alpha_0}] \cdots]] = \mathrm{ad}_{X_{\alpha_r}} \cdots \mathrm{ad}_{X_{\alpha_1}} X_{\alpha_0} \text{ and}$$

$$X^\alpha = X_{\alpha_r} \cdots X_{\alpha_0}.$$

When $|\alpha| = 1$, that is, $\alpha = (\alpha_0)$, then $X^\alpha = X_{\alpha_0} = X_\alpha$. For each $\alpha \in \Lambda^r$, there exist $\epsilon_{\beta,\alpha} \in \{-1, 0, 1\}$ such that

$$X_\alpha = \sum_{\beta \in \Lambda^r} \epsilon_{\beta,\alpha} X^\beta.$$

For example, returning the Heisenberg group case and letting $X_1 = X$, $X_2 = Y$, and $X_3 = Z$, with commutation relations given as in Section 2.1, we have

$$X_{(1,2)} = [X_1, X_2] = X_3, \; X_{\overline{(1,2)}} = X_{(2,1)} = [X_2, X_1] = -X_3, \text{ and } X^{(1,2)'} = X^{(2)} = X_2.$$

Proposition 4.4. *For each $i = 1, \ldots, k$, $\hat{X}_i$ may be written as a linear combination,*

$$\hat{X}_i = \sum_{r=0}^{m} \sum_{\alpha \in \Lambda^r} c_{i,\alpha} \tilde{X}^\alpha, \tag{4.2}$$

with coefficients $c_{i,\alpha} : G \to \mathbb{R}$ *(some of these are 0) such that* $c_{i,\alpha}(\xi_t) \in \mathcal{D}^\infty$, *for all* $t \in [0,1]$.

Proof. Recall from Notation 1.4 that we defined

$$\Sigma_r = \{[X_{i_1}, [\cdots, [X_{i_{r-1}}, X_{i_r}]\cdots] : i_l \in \{1,\ldots,k\}, l \in \{1,\ldots,r\}\} = \{X_\alpha : \alpha \in \Lambda^r\},$$

for $r = 0,\ldots,m$. Recall also from Notation 1.4 that we have an orthonormal basis $\{X_i, Y_j : i \in \{1,\ldots,k\}, j \in \{1,\ldots,d-k\}\}$ of $\mathfrak{g}$, where $d = \dim(G)$ and, for each $j \in \{1,\ldots,d-k\}$, $Y_j = X_{\alpha(j)} \in \Sigma_{r(j)}$ for some $\alpha(j) \in \Lambda^{r(j)}$, $r(j) \in \{1,\ldots,m\}$. Thus, for any $g \in G$ and $X \in \mathfrak{g}$,

$$\begin{aligned}
\hat{X}(g) &= R_{g*}X = L_{g*}L_{g^{-1}*}R_{g*}X = L_{g*}\operatorname{Ad}_{g^{-1}} X \\
&= L_{g*}\left(\sum_{i=1}^k \left\langle \operatorname{Ad}_{g^{-1}} X, X_i\right\rangle X_i + \sum_{j=1}^{d-k} \left\langle \operatorname{Ad}_{g^{-1}} X, Y_j\right\rangle Y_j\right) \\
&= L_{g*}\left(\sum_{i=1}^k \left\langle \operatorname{Ad}_{g^{-1}} X, X_i\right\rangle X_i + \sum_{j=1}^{d-k} \sum_{\alpha\in\Lambda^{r(j)}} \epsilon_{\alpha,\alpha(j)} \left\langle \operatorname{Ad}_{g^{-1}} X, Y_j\right\rangle X^\alpha\right) \\
&= \sum_{i=1}^k \left\langle \operatorname{Ad}_{g^{-1}} X, X_i\right\rangle \tilde{X}_i + \sum_{j=1}^{d-k} \sum_{\alpha\in\Lambda^{r(j)}} \epsilon_{\alpha,\alpha(j)} \left\langle \operatorname{Ad}_{g^{-1}} X, Y_j\right\rangle \tilde{X}^\alpha
\end{aligned}$$

where $\epsilon_{\alpha,\alpha(j)} \in \{-1,0,1\}$. So, for each $l = 1,\ldots,k$,

$$\begin{aligned}
\hat{X}_l(g) &= \sum_{i=1}^k \left\langle \operatorname{Ad}_{g^{-1}} X_l, X_i\right\rangle \tilde{X}_i(g) + \sum_{j=1}^{d-k} \sum_{\alpha\in\Lambda^{r(j)}} \epsilon_{\alpha,\alpha(j)} \left\langle \operatorname{Ad}_{g^{-1}} X_l, Y_j\right\rangle \tilde{X}^\alpha(g) \\
&= \sum_{r=0}^m \sum_{\alpha\in\Lambda^r} c_{l,\alpha}\tilde{X}^\alpha(g),
\end{aligned}$$

where

$$c_{l,\alpha}(g) = \left\langle \operatorname{Ad}_{g^{-1}} X_l, X_i\right\rangle, \quad \text{when } r = 0 \text{ and } \alpha = (i),$$

and

$$c_{l,\alpha}(g) = \epsilon\left\langle \operatorname{Ad}_{g^{-1}} X_l, Y_j\right\rangle, \epsilon \in \{-1,0,1\}, \quad \text{when } r \in \{1,\ldots,m\}.$$

Now recall that $\mathrm{Ad}_{\xi_t} \in \mathcal{D}^\infty(\mathrm{End}(\mathfrak{g}))$ by Proposition 3.14, and $u : \mathrm{End}(\mathfrak{g}) \to \mathbb{R}$ given by $u(W) = \langle WX, Y\rangle$ is a smooth function for any fixed $X, Y \in \mathfrak{g}$. Thus, $u(\mathrm{Ad}_{\xi_t}) \in \mathcal{D}^\infty$ for all $t \in [0,1]$. Since $c_{l,\alpha}(\xi_t) = \epsilon u(\mathrm{Ad}_{\xi_t})$, with $X = X_l$ and $Y = X_i$ or Y_j, this implies that $c_{l,\alpha}(\xi_t) \in \mathcal{D}^\infty$, for all $\alpha \in \Lambda^r$ and $l \in \{1, \ldots, k\}$. □

Theorem 4.5. *Let G be a Lie group with Lie algebra $\mathfrak{g}$, and suppose $\{X_1, \ldots, X_k\} \subset \mathfrak{g}$ is a Lie generating set. Then, for all $p \in (1,\infty)$, $K_p(t) < \infty$, where $K_p(t)$ are the functions defined in Notation 1.9; that is, $K_p(t)$ is the best function so that*

$$|\nabla P_t f|^p \leq K_p(t) P_t |\nabla f|^p,$$

for all $f \in C_c^\infty(G)$ and $t > 0$.

Proof. Lemma 4.1 implies that the inequality (I_p) is translation invariant on groups. Thus, we need only determine finite $K_p(t)$ such that the inequality holds at the identity. Using Lemma 4.2 and Equation (4.2), we have

$$|\nabla P_t f|^2(e) = \sum_{i=1}^k |\tilde{X}_i P_t f|^2(e) = \sum_{i=1}^k |P_t \hat{X}_i f|^2(e) \leq C \sum_{i=1}^k \sum_{r=0}^m \sum_{\alpha \in \Lambda^r} |P_t c_{i,\alpha} \tilde{X}^\alpha f|^2(e),$$

for a constant $C = C(k,m)$. Recall that Equation (3.4) implies that, for any $f \in C_c^\infty(G)$, $P_t f(e) = \mathbb{E}[f(\xi_t)]$, where ξ is the solution to the Stratonovich equation (3.3). Thus, for any $\alpha \in \Lambda^r$,

$$\begin{aligned}
|P_t c \tilde{X}^\alpha f|(e) &\leq \mathbb{E}|c(\xi_t)(\tilde{X}^\alpha f)(\xi_t)| = \mathbb{E}|c(\xi_t)\mathbf{X}^{\alpha'}[(\tilde{X}_{\alpha_0} f)(\xi_t)]| \\
&= \mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c(\xi_t)](\tilde{X}_{\alpha_0} f)(\xi_t)\right| \\
&\leq \left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c(\xi_t)]\right|^q\right)^{1/q} \left(\mathbb{E}|(\tilde{X}_{\alpha_0} f)(\xi_t)|^p\right)^{1/p} \\
&= \left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c(\xi_t)]\right|^q\right)^{1/q} \left(P_t|\tilde{X}_{\alpha_0} f|^p(e)\right)^{1/p},
\end{aligned}$$

by Hölder's inequality, where q is the conjugate exponent to p, $\mathbf{X}^\alpha$ is the lifted vector field on W of the vector field $\tilde{X}^\alpha$, as defined in Equation (3.11), and $(\mathbf{X}^\alpha)^* = \mathbf{X}^*_{\alpha_r} \cdots \mathbf{X}^*_{\alpha_0}$

(so $\left(\mathbf{X}^{\overline{\alpha'}}\right)^* = \mathbf{X}^*_{\alpha_1} \cdots \mathbf{X}^*_{\alpha_r}$). Thus,

$$\begin{aligned}|\nabla P_t f|^p(e) &\le C\left[\sum_{i=1}^k \sum_{r=0}^m \sum_{\alpha\in\Lambda^r} \left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c_{i,\alpha}(\xi_t)]\right|^q\right)^{2/q} \left(P_t|\tilde{X}_{\alpha_0} f|^p(e)\right)^{2/p}\right]^{p/2} \\ &\le C\left[\sum_{i=1}^k \sum_{r=0}^m \sum_{\alpha\in\Lambda^r} \left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c_{i,\alpha}(\xi_t)]\right|^q\right)^{p/q}\right] P_t|\nabla f|^p(e),\end{aligned}$$

where $C = C(k,m,p)$ and $q = \frac{p-1}{p}$. Thus, the inequality (I_p) holds with

$$C_p(t) = C(k,m,p) \sum_{i=1}^k \sum_{r=0}^m \sum_{\alpha\in\Lambda^r} \left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c_{i,\alpha}(\xi_t)]\right|^q\right)^{p/q}. \tag{4.3}$$

Propositions 3.10 and 4.4 imply that $\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c_{i,\alpha}(\xi_t)] \in L^{\infty-}(\mu)$, for all $i = 1,\ldots,k$ and $\alpha \in \Lambda^r$. Therefore, $K_p(t) \le C_p(t) < \infty$ for all $t > 0$ and $p \in (1,\infty)$. □

It is important to note that there is currently no good control over the behavior of the functions C_p in Equation (4.3) with respect to t. In fact, from certain scaling arguments, it is expected that $C_p(t) \to \infty$ and $t \to 0$; see for example [8, 33]. However, we do not claim these are the best constants for which the inequality holds.

4.2 A Poincaré inequality

The following result is a direct corollary to Theorem 4.5. The proof is completely analogous to the proof of Theorem 2.23.

Theorem 4.6 (Poincaré Inequality). *Let $K_2(t)$ be the best function for which Equation I_p holds for $p = 2$, and let $p_t(g)\,dg$ be the hypoelliptic heat kernel. Then*

$$\int_G f^2(g)p_t(g)\,dg - \left(\int_G f(g)p_t(g)\,dg\right)^2 \le \Lambda(t)\int_G |\nabla f|^2(g)p_t(g)\,dg, \tag{4.4}$$

for all $f \in C_c^\infty(G)$ and $t > 0$, where

$$\Lambda(t) = \int_0^t K_2(s)\,ds.$$

Proof. Let $F_t(g) = (P_t f)(g)$. Then

$$\frac{d}{dt}P_{t-s}F_s^2 = P_{t-s}\left(-\frac{1}{2}LF_s^2 + F_sLF_s\right) = -P_{t-s}|\nabla F_s|^2.$$

Integrating this equation on t implies that

$$\begin{aligned} P_tf^2 - (P_tf)^2 &= \int_0^t P_{t-s}|\nabla F_s|^2\,ds \\ &= \int_0^t P_{t-s}|\nabla P_sf|^2\,ds \\ &\leq \int_0^t K_2(s)P_{t-s}P_s|\nabla f|^2\,ds = \left(\int_0^t K_2(s)\,ds\right)\cdot P_t|\nabla f|^2 \end{aligned}$$

wherein we have made use of Theorem 4.5. Evaluating the above at $e \in G$ gives the desired result. □

This Poincaré inequality is less useful in the general Lie group case because nothing is known about the integrability of $K_p(t)$. However, in the next two sections, we will carry out a generalization of the argument that was made in Section 2.3. That is, we will show that when G is a nilpotent Lie group, $K_p(t)$ is a bounded function for all $p \in (1, \infty)$. In particular, when $p = 2$, this implies the Poincaré inequality holds with $\Lambda(t) < \infty$, for all $t > 0$.

4.2.1 Stratified nilpotent Lie groups

Definition 4.7. A Lie algebra $\mathfrak{g}$ is said to be *nilpotent* if ad_X is a nilpotent endomorphism of $\mathfrak{g}$ for all $X \in \mathfrak{g}$, that is, if there exists $m \in \mathbb{N}$ such that

$$\mathrm{ad}_{Y_1}\cdots\mathrm{ad}_{Y_{m-1}}Y_m = [Y_1, [\cdots, [Y_{m-1}, Y_m]\cdots] = 0,$$

for any $Y_1, \ldots, Y_m \in \mathfrak{g}$. We say that $\mathfrak{g}$ is nilpotent of step m. A Lie group G is nilpotent if $\mathfrak{g} = \mathrm{Lie}(G)$ is a nilpotent Lie algebra.

Definition 4.8. A family of *dilations* on a Lie algebra $\mathfrak{g}$ is a family of algebra automorphisms $\{\Phi_r\}_{r>0}$ on $\mathfrak{g}$ of the form $\Phi_r = \exp(W \log r)$, where W is a diagonalizable linear operator on $\mathfrak{g}$ with positive eigenvalues.

Definition 4.9. A *stratified* group G is a simply connected nilpotent group for which there exists a subset of the Lie algebra $V_1 \subset \mathfrak{g}$, such that $\mathfrak{g} = \oplus_{j=1}^{m} V_j$ with $V_{j+1} = [V_1, V_j]$, for $j = 1, \ldots, m-1$, and $V_{m+1} = [V_1, V_m] = \{0\}$.

For a general exposition on nilpotent Lie groups and dilations, see [19, 22] and references contained therein. If G is a stratified Lie group, a natural family of dilations may be defined on $\mathfrak{g}$ by setting $\Phi_r(X) = r^j X$, for all $X \in V_j$. The generator W of this dilation acts on parts of the vector space decomposition by $WV_j = jV_j$, for each $j = 1, \ldots, m$. The automorphism Φ_r induces a group dilation ϕ_r via the exponential maps, $\phi_r = \exp \circ \Phi_r \circ \exp^{-1}$. Since G is a simply connected nilpotent group, the exponential map is in fact a global diffeomorphism on $\mathfrak{g}$, and $\exp^{-1}$ exists everywhere on G; see for example Theorem 3.6.2 of Varadarajan [49]. Then for each $X \in V_1$,

$$\begin{aligned} \tilde{X}(f \circ \phi_r)(g) &= \frac{d}{d\epsilon}\bigg|_0 (f \circ \phi_r)(ge^{\epsilon X}) = \frac{d}{d\epsilon}\bigg|_0 f(\phi_r(g)\phi_r(e^{\epsilon X})) \\ &= \frac{d}{d\epsilon}\bigg|_0 f(\phi_r(g)e^{r\epsilon X}) = \frac{d}{d\epsilon}\bigg|_0 r f(\phi_r(g)e^{\epsilon X}) = r(\tilde{X} f \circ \phi_r)(g), \end{aligned} \tag{4.5}$$

for all $f \in C^1(G)$, where in the second equality we have used that ϕ_r is a homomorphism. Let $\{X_i\}_{i=1}^k \subset V_1$ be a basis of V_1, and consider the operators $\nabla = (\tilde{X}_1, \ldots, \tilde{X}_k)$ and $L = \sum_{i=1}^k$. Equation (4.5) implies that

$$\nabla(f \circ \phi_r) = r(\nabla f) \circ \phi_r,$$

and we have the following proposition.

Proposition 4.10. *Let L denote the self-adjoint extension of $\sum_{i=1}^k \tilde{X}_i^2$, and $P_t = e^{tL/2}$ be as in Definition 1.8. Then*

$$L(f \circ \phi_r) = r^2 (Lf) \circ \phi_r$$

and

$$P_t(f \circ \phi_r) = e^{tL/2}(f \circ \phi_r) = \left(e^{r^2 tL/2} f\right) \circ \phi_r = P_{r^2 t} \circ \phi_r,$$

for any $f \in C_c^\infty(G)$.

Proof. Let $\mathcal{E}^0(f,h) := \sum_{i=1}^k (\tilde{X}_i f, \tilde{X}_i h)_{L^2(G)}$ be a Dirichlet form associated to L. Recall from Chapter 1 that $\mathcal{E}^0$ has a closed extension $\mathcal{E}$. By definition,

$$f_1 \in C_c^\infty(G) \text{ and } Lf_1 = h \iff \mathcal{E}(f_1, f_2) = (h, f_2),\ \forall f_2 \in \mathrm{Dom}(\mathcal{E}).$$

Now note that

$$\begin{aligned}
\mathcal{E}^0(f\circ\phi_r, f\circ\phi_r) &= \sum_{i=1}^k \int_G |\tilde{X}_i(f\circ\phi_r)|^2(g)\,dg \\
&= \sum_{i=1}^k r^2 \int |(\tilde{X}_i f)\circ\phi_r|^2(g)\,dg \\
&= \sum_{i=1}^k r^2 \int |\tilde{X}_i f|^2(g) J(r^{-1})\,dg = r^2 J(r^{-1})\mathcal{E}^0(f,f),
\end{aligned}$$

where $J(r)$ is the Jacobian of the transformation ϕ_r,

$$J(r) = \prod_{j=1}^m (r^j)^{d_j}$$

with $d_j = \dim(V_j)$. Thus, $J(r^{-1}) = J(r)^{-1}$. So $f \in \mathrm{Dom}(\mathcal{E})$ implies that $f\circ\phi_r \in \mathrm{Dom}(\mathcal{E})$, and, in general, $\mathcal{E}(f\circ\phi_r, h\circ\phi_r) = r^2 J(r^{-1})\mathcal{E}(f,h)$, for $f, h \in \mathrm{Dom}(\mathcal{E})$. Replacing h here by $h\circ\phi_{r^{-1}}$ gives

$$\begin{aligned}
\mathcal{E}(f\circ\phi_r, h) &= r^2 J(r^{-1})\mathcal{E}(f, h\circ\phi_{r^{-1}}) \\
&= r^2 J(r^{-1})(Lf, h\circ\phi_{r^{-1}})_{L^2(G)} \\
&= r^2 J(r^{-1})J(r)(Lf\circ\phi_r, h)_{L^2(G)} = r^2(Lf\circ\phi_r, h)_{L^2(G)},
\end{aligned}$$

implies that if $f \in \mathrm{Dom}(L)$, then $f\circ\phi_r \in \mathrm{Dom}(L)$ and $L(f\circ\phi_r) = r^2 Lf\circ\phi_r$.

Now, for $r > 0$, let $U_r : L^2(G) \to L^2(G)$ be the unitary operator given by $U_r f = \frac{1}{\sqrt{J(r^{-1})}} f\circ\phi_r$. Then

$$LU_r = r^2 U_r L = U_r(r^2 L)$$

as operators, and thus $U_r^{-1} L U_r = r^2 L$. Then we have that

$$U_r^{-1} e^{tL/2} U_r = e^{tU_r^{-1}LU_r/2} = e^{r^2 tL/2},$$

from which it follows that

$$r^2 e^{tL/2}(f \circ \phi_r) = e^{tL/2} U_r f = U_r e^{r^2 tL/2} f = r^2 (e^{r^2 tL/2} f) \circ \phi_r.$$

□

Given these equations, an argument identical to the proof of Proposition 2.6 proves the following proposition.

Proposition 4.11. *Suppose G is a stratified Lie group with vector space decomposition $\oplus_{j=1}^m V_j$. Let $\{X_i\}_{i=1}^k \subset V_1$, ∇, and L be as above, and let $p \in (1, \infty)$. If K_p is the best constant such that*

$$|\nabla P_1 f|^p \leq K_p P_1 |\nabla f|^p,$$

for all $f \in C_c^\infty(G)$, then $K_p(t) = K_p$ for all $t > 0$, where $K_p(t)$ is the function defined in Notation 1.9.

4.2.2 Nilpotent Lie groups

Now let G be a general nilpotent Lie group. Because not all nilpotent Lie groups admit dilations, we can not show that the functions $K_p(t)$ should be scale invariant in this context. However, by covering G with a group which has a family of dilations adapted to its structure, we may show that there exists some constant $K_p < \infty$ for which $K_p(t) < K_p$ for all $t > 0$.

Definition 4.12. Let $\mathcal{L} = \mathcal{L}(k, m)$ be the *free nilpotent Lie algebra of step m with k generators* $\{e_i\}_{i=1}^k$. Then $\mathcal{L}$ is the unique (up to isomorphism) nilpotent Lie algebra of rank m such that, for every nilpotent Lie algebra $\mathfrak{g}$ of rank m and map $\tilde{\Pi} : \{e_1, \ldots, e_k\} \to \mathfrak{g}$, there exists a unique homomorphism $\Pi : \mathcal{L} \to \mathfrak{g}$ which extends $\tilde{\Pi}$. Let $\mathcal{N} = \mathcal{N}(k, m)$ be the *free nilpotent Lie group of rank m with k generators*, which is the simply connected group of $\mathcal{L}(k, m)$.

The Lie algebra $\mathcal{L}(k, m)$ admits a vector space decomposition by setting $V_1 = \operatorname{span}\{e_1, \ldots, e_k\}$. Thus, $\mathcal{N}$ is a stratified Lie group with Hörmander set $\{e_i\}_{i=1}^k \subset \mathcal{L}$;

for definitions and further details, see [51]. Let $\nabla_{\mathcal{L}} = (\tilde{e}_1, \ldots, \tilde{e}_k)$, $\mathscr{L} = \sum_{i=1}^k \tilde{e}_i^2$, and $\mathscr{P}_t = e^{t\mathscr{L}/2}$. Theorem 4.5 and Proposition 4.11 imply that, for all $p \in (1, \infty)$, there exist constants $K_p^{\mathcal{L}} < \infty$ such that

$$|\nabla_{\mathcal{L}} \mathscr{P}_t f|^p \leq K_p^{\mathcal{L}} \mathscr{P}_t |\nabla_{\mathcal{L}} f|^p, \tag{4.6}$$

for all $f \in C_c^\infty(\mathcal{N})$ and $t > 0$.

Proposition 4.13. *Let G be a nilpotent group of step m with Hörmander set $\{X_i\}_{i=1}^k$. Then $K_p(t) \leq K_p^{\mathcal{L}}$ for all $t > 0$, where $K_p(t)$ is the function defined in Notation 1.9.*

Proof. By definition of $\mathcal{L} = \mathcal{L}(k, m)$, there exists a unique Lie algebra homomorphism $\Pi : \mathcal{L} \to \mathfrak{g}$ such that $\Pi(e_i) = X_i$. Then Π induces a group homomorphism $\pi : \mathcal{N} \to G$ via the exponential maps,

$$\pi = \exp_G \circ \Pi \circ \exp_{\mathcal{N}}^{-1}.$$

Again, because $\mathcal{N}$ is a simply connected nilpotent Lie group, the exponential map on $\mathcal{L}$ is a global diffeomorphism. Note that $d\pi = \Pi$,

$$\begin{array}{ccc} \mathcal{L}(k,m) & \xrightarrow{\Pi} & \mathfrak{g} \\ {\scriptstyle \exp_{\mathcal{N}}}\big\downarrow & & \big\downarrow{\scriptstyle \exp_G} \\ \mathcal{N}(k,m) & \xrightarrow[\pi]{} & G \end{array}$$

and the vector fields $\tilde{X}_i$ and $\tilde{e}_i$ are π-related; that is,

$$\tilde{e}_\alpha (f \circ \pi) = (\tilde{X}_\alpha f) \circ \pi,$$

for any multi-index $\alpha \in \Lambda^r$ and $f \in C_c^\infty(G)$. Since π is a Lie group homomorphism, $f \circ \pi \in C_c^\infty(\mathcal{N})$, and thus,

$$|\nabla P_t f|^p(e) = |\nabla_{\mathcal{L}} \mathscr{P}_t (f \circ \pi)|^p(e_{\mathcal{N}}) \leq K_p^{\mathcal{L}} \mathscr{P}_t |\nabla_{\mathcal{L}} (f \circ \pi)|^p(e_{\mathcal{N}}) = K_p^{\mathcal{L}} P_t |\nabla f|^p(e),$$

where $e_{\mathcal{N}}$ is the identity element of $\mathcal{N}$. Since $K_p(t)$ is the best constant for which

$$|\nabla P_t f|^p(e) \leq K_p(t) P_t |\nabla f|^p(e)$$

holds, the above implies that $K_p(t) \leq K_p^{\mathcal{L}}$ for all $t > 0$. $\square$

This method of lifting the vector fields to a free nilpotent Lie algebra was learned from [50, 51]. A generalization of this procedure may be found in [43].

Remark 4.14. Note that the above argument is independent of the minimality of the Hörmander set $\{X_i\}_{i=1}^k$. So suppose that the collection $\{X_i\}_{i=1}^k$ spans the Lie algebra $\mathfrak{g}$. Since G is a nilpotent Lie group (and thus unimodular) it is then well known that the operator $L = \sum_{i=1}^k \tilde{X}_i^2$ is in fact the Laplace-Beltrami operator on the Riemannian manifold $(G, \langle\cdot,\cdot\rangle)$. Then by Theorem 1.2,

$$|\nabla P_t f|^p \leq e^{pkt} P_t |\nabla f|^p,$$

where $-k$ is a lower bound on the Ricci curvature. Proposition 4.13 improves this result for large t by implying that there exists a $K_p < \infty$ independent of t such that

$$|\nabla P_t f|^p \leq K_p P_t |\nabla f|^p,$$

for all $f \in C_p^\infty(G)$ and $t > 0$. Thus, we have the following corollary.

Corollary 4.15. *Let G be a nilpotent Lie group of step m and $\{X_i\}_{i=1}^k \subset \mathfrak{g}$ such that $\{X_i\}_{i=1}^k$ spans the Lie algebra $\mathfrak{g}$. Then, for $K_p(t)$ as in Notation 1.9,*

$$K_p(t) \leq \min\{K_p^{\mathcal{L}}, e^{pkt}\},$$

where $K_p^{\mathcal{L}}$ is the best constant so that (I_p) holds on $\mathcal{L}(k,m)$ and $-k$ is a lower bound on the Ricci curvature associated to the Riemannian metric determined by $L = \sum_{i=1}^k \tilde{X}_i^2$.

Corollary 4.16 (Poincaré Inequality for nilpotent Lie groups). *Suppose G is a nilpotent Lie group, and let K_2 be a finite constant for which the inequality (I_p) holds for $p = 2$. Then the inequality (4.4) holds with $\Lambda(t) = K_2 t$, for all $t > 0$.*

Chapter 5

Appendix

5.1 Proof of Theorem 1.2

Proof. Comparing Equations (1.1) and (1.2), it is clear that a lower bound on the Ricci curvature immediately gives

$$\Gamma_2(f) = |\nabla^2 f|^2 + (\mathrm{Ric}\nabla f, \nabla f) \geq (\mathrm{Ric}\nabla f, \nabla f) \geq -2k|\nabla f|^2 = \Gamma_1(f),$$

and so (1) implies (2). For the converse, let $m \in M$, $v \in T_m M$, and let $f \in C^\infty(M)$ such that $f(\exp_m(w)) = (v, w)$, for all $w \in T_m M$ in a neighborhood of $0 \in T_m M$. Since $f(\exp_m(tw)) = tv \cdot w$, it follows that

$$df(w_m) = \frac{d}{dt}\bigg|_0 f(\exp_m(tw)) = (v, w)$$

and

$$\nabla^2 f(w_m, w_m) = \frac{d^2}{dt^2}\bigg|_0 f(\exp_m(tw)) = 0.$$

Thus, $(\nabla^2 f)_m = 0$, $\nabla f(m) = v$, and so

$$-2k|v|^2 = -2k\Gamma_1(f) \leq \Gamma_2(f) = (\mathrm{Ric}\, v, v);$$

that is, (2) implies (1).

An application of the Cauchy-Schwarz inequality shows that (3) implies (4), and (5) is a trivial consequence of (4). Assuming (5) holds, differentiating the inequality

$$|\nabla P_t f|^2 \le K(t) P_t |\nabla f|^2$$

at $t = 0$, implies that

$$\begin{aligned} 2(\nabla f, \frac{1}{2}\nabla L f) &= \frac{d}{dt}\Big|_0 |\nabla P_t f|^2 \\ &\le \frac{d}{dt}\Big|_0 K(t) P_t |\nabla f|^2 = \dot{K}(0)|\nabla f|^2 + \frac{1}{2} L |\nabla f|^2. \end{aligned}$$

Hence,

$$\Gamma_2(f) = \frac{1}{2} L|\nabla f|^2 - (\nabla f, \nabla L f) \ge \dot{K}(0)|\nabla f|^2 = \dot{K}(0)\Gamma_1(f),$$

and thus (2) follows with $k = \frac{1}{2}\dot{K}(0)$.

Finally, to verify that (1) implies (3), we may use the Feynman-Kac formula,

$$e^{t\Delta/2} f(x) = \mathbb{E}[Q_t //_t^{-1} f(B_t^x)],$$

where B_t^x is Brownian motion on M such that $B_0^x = x$, $//_t$ is parallel translation, and Q_t satisfies

$$\frac{d}{dt} Q_t = -\mathrm{Ric}_{//_t} Q_t = 0 \text{ with } Q_0 = Id;$$

see for example Theorem 7.2.1 of [27]. Thus,

$$\nabla e^{t\Delta/2} f(x) = \mathbb{E}[Q_t //_t^{-1} \nabla f(B_t^x)].$$

So if $\mathrm{Ric} \ge -kg$, then $|Q_t| \le e^{kt}$, and hence

$$\begin{aligned} |\nabla e^{t\Delta/2} f|(x) &\le \mathbb{E}\left|Q_t //_t^{-1} \nabla f(B_t^x)\right| \\ &\le e^{kt}\mathbb{E}\left|//_t^{-1}\nabla f(B_t^x)\right| = e^{kt}\mathbb{E}\left|\nabla f(B_t^x)\right| = e^{kt} e^{t\Delta/2}|\nabla f|(x), \end{aligned}$$

and the proof is complete. □

5.2 Proof of Theorem 3.5

Notation 5.1. Let $S := \{v \in \mathfrak{g} : \langle v, v\rangle = 1\}$, that is, S is the unit sphere in $\mathfrak{g}$.

Proof. To show $\bar{\sigma}_t^{-1} \in L^{\infty-}(\mu)$, it suffices to show that

$$\mu\left(\inf_{v\in S}\langle\bar{\sigma}_t v, v\rangle < \varepsilon\right) = O(\varepsilon^{\infty-}).$$

To verify this claim, notice that $\lambda_0 := \inf_{v\in S}\langle\bar{\sigma}_t v, v\rangle$ is the smallest eigenvalue of $\bar{\sigma}_t$. Since $\det\bar{\sigma}_t$ is the product of the eigenvalues of $\bar{\sigma}_t$ it follows that $\Delta_t := \det\bar{\sigma}_t \geq \lambda_0^n$ and so $\{\det\bar{\sigma}_t < \varepsilon^n\} \subset \{\lambda_0 < \varepsilon\}$ and hence

$$\mu\left(\det\bar{\sigma}_t < \varepsilon^n\right) \leq \mu\left(\lambda_0 < \varepsilon\right) = O(\varepsilon^{\infty-}).$$

By replacing ε by $\varepsilon^{1/n}$ above this implies $\mu\left(\Delta_t < \varepsilon\right) = O(\varepsilon^{\infty-})$. From this estimate it then follows that

$$\begin{aligned}\mathbb{E}\left[\Delta_t^{-q}\right] = \mathbb{E}\int_{\Delta_t}^{\infty} q\tau^{-q-1}d\tau = q\mathbb{E}\int_0^{\infty} 1_{\Delta_t\leq\tau}\ \tau^{-q-1}d\tau \\ = q\int_0^{\infty}\mu(\Delta_t\leq\tau)\ \tau^{-q-1}d\tau = q\int_0^{\infty} O(\tau^p)\ \tau^{-q-1}d\tau\end{aligned}$$

which is seen to be finite by taking $p \geq q+1$.

More generally, if T is any stopping time with $T \leq t$, since $\langle\bar{\sigma}_T v, v\rangle \leq \langle\bar{\sigma}_t v, v\rangle$ for all $v \in S$, it suffices to prove

$$\mu\left(\inf_{v\in S}\langle\bar{\sigma}_T v, v\rangle < \varepsilon\right) = O(\varepsilon^{\infty-}). \tag{5.1}$$

According to Lemma 5.2 and Proposition 5.3 below, Equation (5.1) holds with

$$T = T_\delta := \inf\left\{t > 0 : \max\left\{\left|\mathrm{Ad}_{\xi_t} - I_{\mathfrak{g}}\right|, |\xi_t|\right\} > \delta\right\} \tag{5.2}$$

provided $\delta > 0$ is chosen sufficiently small. □

The rest of this section is now devoted to the proof of Lemma 5.2 and Proposition 5.3 below. In what follows we will make repeated use of the identity,

$$\langle\bar{\sigma}_T v, v\rangle = \sum_{i=1}^{k}\int_0^T \left\langle\mathrm{Ad}_{\xi_\tau} X_i, v\right\rangle^2 d\tau. \tag{5.3}$$

To prove this, recall that $\{X_i\}_{i=1}^{k}$ is an orthonormal basis of $\mathfrak{g}_0$. Thus,

$$P\,\mathrm{Ad}^{\dagger}_{\xi_\tau} v = \sum_{i=1}^{k} \left\langle \mathrm{Ad}^{\dagger}_{\xi_\tau} v, X_i \right\rangle X_i = \sum_{i=1}^{k} \left\langle \mathrm{Ad}_{\xi_\tau} X_i, v \right\rangle X_i$$

so that

$$\left\langle \mathrm{Ad}_{\xi_\tau} P\,\mathrm{Ad}^{\dagger}_{\xi_\tau} v, v \right\rangle = \sum_{i=1}^{k} \left\langle \mathrm{Ad}_{\xi_\tau} \left\langle \mathrm{Ad}_{\xi_\tau} X_i, v \right\rangle X_i, v \right\rangle = \sum_{i=1}^{k} \left\langle \mathrm{Ad}_{\xi_\tau} X_i, v \right\rangle^2$$

which upon integrating in τ gives Equation (5.3).

Lemma 5.2 (Compactness Argument). *Let T_δ be as in Equation (5.2), and suppose for all $v \in S$ there exists $i \in \{1, \ldots, k\}$ and an open neighborhood $N \subset_o S$ of v such that*

$$\sup_{u \in N} \mu \left(\int_0^{T_\delta} \left\langle \mathrm{Ad}_{\xi_\tau} X_i, u \right\rangle^2 d\tau < \varepsilon \right) = O\left(\varepsilon^{\infty-}\right), \tag{5.4}$$

then Equation (5.1) holds provided $\delta > 0$ is sufficiently small.

Proof. By compactness of S, it follows from Equation (5.4) that

$$\sup_{u \in S} \mu \left(\int_0^{T_\delta} \left\langle \mathrm{Ad}_{\xi_\tau} X_i, u \right\rangle^2 d\tau < \varepsilon \right) = O\left(\varepsilon^{\infty-}\right). \tag{5.5}$$

For $w \in \mathfrak{g}$, let ∂_w denote the directional derivative acting on functions $f(v)$ with $v \in \mathfrak{g}$. For all $v, w \in \mathbb{R}^n$ with $|v| \le 1$ and $|w| \le 1$ (using Equation (5.3)),

$$\begin{aligned} \left|\partial_w \left\langle \bar{\sigma}_{T_\delta} v, v \right\rangle\right| &\le 2 \sum_{i=1}^{k} \int_0^{T_\delta} \left| \left\langle \mathrm{Ad}_{\xi_\tau} X_i, v \right\rangle \left\langle \mathrm{Ad}_{\xi_\tau} X_i, w \right\rangle \right| d\tau \\ &\le 2 \sum_{i=1}^{k} \int_0^{T_\delta} \left| \mathrm{Ad}_{\xi_\tau} X_i \right|^2_{\mathrm{Hom}(\mathbb{R}^n, \mathfrak{g})} d\tau. \end{aligned}$$

Thus, by choosing $\delta > 0$ in Equation (5.2) sufficiently small, we may assume there is a non-random constant $\theta < \infty$ such that

$$\sup_{|v|,|w| \le 1} \left|\partial_w \left\langle \bar{\sigma}_{T_\delta} v, v \right\rangle\right| \le \theta < \infty.$$

With this choice of δ, if $v, w \in S$ satisfy $|v-w| < \theta/\varepsilon$, then

$$|\langle \bar{\sigma}_{T_\delta} v, v\rangle - \langle \bar{\sigma}_{T_\delta} w, w\rangle| < \varepsilon. \tag{5.6}$$

There exists $D < \infty$ such that ,for any $\varepsilon > 0$, there is an open cover of S with at most $D \cdot (\theta/\varepsilon)^n$ balls of the form $B(v_j, \varepsilon/\theta)$. From Equation (5.6), for any $v \in S$ there exists j such that $v \in B(v_j, \varepsilon/\theta) \cap S$ and

$$|\langle \bar{\sigma}_{T_\delta} v, v\rangle - \langle \bar{\sigma}_{T_\delta} v_j, v_j\rangle| < \varepsilon.$$

So if $\inf_{v\in S} \langle \bar{\sigma}_{T_\delta} v, v\rangle < \varepsilon$, then $\min_j \langle \bar{\sigma}_{T_\delta} v_j, v_j\rangle < 2\varepsilon$; that is,

$$\left\{\inf_{v\in S} \langle \bar{\sigma}_{T_\delta} v, v\rangle < \varepsilon\right\} \subset \left\{\min_j \langle \bar{\sigma}_{T_\delta} v_j, v_j\rangle < 2\varepsilon\right\} \subset \bigcup_j \{\langle \bar{\sigma}_{T_\delta} v_j, v_j\rangle < 2\varepsilon\}.$$

Therefore,

$$\begin{aligned}\mu\left(\inf_{v\in S} \langle \bar{\sigma}_{T_\delta} v, v\rangle < \varepsilon\right) &\le \sum_j \mu\left(\langle \bar{\sigma}_{T_\delta} v_j, v_j\rangle < 2\varepsilon\right) \\ &\le D \cdot (\theta/\varepsilon)^n \cdot \sup_{v\in S} \mu\left(\langle \bar{\sigma}_{T_\delta} v, v\rangle < 2\varepsilon\right) \le D \cdot (\theta/\varepsilon)^n O(\varepsilon^{\infty-}) = O(\varepsilon^{\infty-}).\end{aligned}$$

□

Recall from Equation (3.19), that for any $X \in \mathfrak{g}$,

$$d \operatorname{Ad}_{\xi_s} X = \operatorname{Ad}_{\xi_s} \circ \operatorname{ad}_{db_s} X = \operatorname{Ad}_{\xi_s} \operatorname{ad}_{X_i} X \circ db_s^i = \operatorname{Ad}_{\xi_s} [X_i, X] \circ db_s^i,$$

which may be rewritten in Itô form as

$$d[\operatorname{Ad}_{\xi_s} X] = \operatorname{Ad}_{\xi_s} [X_i, X]\, db_s^i + \frac{1}{2} \operatorname{Ad}_{\xi_s} \operatorname{ad}_{X_i}^2 X\, ds. \tag{5.7}$$

This gives the first hint that Hörmander's condition is relevant to showing $\Delta_t^{-1} \in L^{\infty-}(\mu)$, or equivalently that $\bar{\sigma}_t^{-1} \in L^{\infty-}(\mu)$.

Proposition 5.3. *Let T_δ be as in Equation (5.2). If Hörmander's restricted bracket condition holds and $v \in S$ is given, there exists $i \in \{1, \ldots, k\}$ and an open neighborhood $U \subset_o S$ of v such that*

$$\sup_{u\in U} \mu\left(\int_0^{T_\delta} \langle \operatorname{Ad}_{\xi_\tau} X_i, u\rangle^2\, d\tau \le \varepsilon\right) = O\left(\varepsilon^{\infty-}\right).$$

Proof. Hörmander's condition implies that there exist $m \in \mathbb{N}$ and $\beta > 0$ such that

$$\sum_{r=0}^{m} \frac{1}{|\Sigma_r|} \sum_{V \in \Sigma_r} V V^{\mathrm{tr}} \geq 3\beta I.$$

Equivalently, for all $v \in S$,

$$3\beta \leq \sum_{r=0}^{m} \frac{1}{|\Sigma_r|} \sum_{V \in \Sigma_r} \langle V, v \rangle^2 \leq \max_{\substack{V \in \Sigma_r \\ r \in \{0,\ldots,m\}}} \langle V, v \rangle^2 .$$

By choosing $\delta > 0$ in Equation (5.2) sufficiently small we may assume that

$$\max_{\substack{V \in \Sigma_r \\ r \in \{0,\ldots,m\}}} \inf_{\tau \leq T_\delta} \langle \mathrm{Ad}_{\xi_\tau} V, v \rangle^2 \geq 2\beta, \text{ for all } v \in S.$$

Fix a $v \in S$ and $V \in \cup_{r=0}^{m} \Sigma_r$ such that

$$\inf_{\tau \leq T_\delta} \langle \mathrm{Ad}_{\xi_\tau} V, v \rangle^2 \geq 2\beta,$$

and choose an open neighborhood $U \subset S$ of v such that

$$\inf_{\tau \leq T_\delta} \langle \mathrm{Ad}_{\xi_\tau} V, u \rangle^2 \geq \beta, \text{ for all } u \in U.$$

Then, by Equation (8.19) of Driver [15], $\mu(T_\delta \leq \tau) = O(\tau^{\infty-})$, and so

$$\sup_{u \in U} \mu \left(\int_0^{T_\delta} \langle \mathrm{Ad}_{\xi_\tau} V, u \rangle^2 \, d\tau \leq \varepsilon \right) \leq \mu \left(\int_0^{T_\delta} \beta dt \leq \varepsilon \right) = \mu \left(T_\delta \leq \varepsilon/\beta \right) = O\left(\varepsilon^{\infty-} \right). \tag{5.8}$$

Write $V = \mathrm{ad}_{X_{i_r}} \cdots \mathrm{ad}_{X_{i_2}} X_{i_1}$ with $r \leq m$. When $r = 1$, Equation (5.8) becomes

$$\sup_{u \in U} \mu \left(\langle \bar{\sigma}_{T_\delta} u, u \rangle \leq \varepsilon \right) \leq \sup_{u \in U} \mu \left(\int_0^{T_\delta} \langle \mathrm{Ad}_{\xi_\tau} X_{i_1}, u \rangle^2 \, dt \leq \varepsilon \right) = O\left(\varepsilon^{\infty-} \right),$$

and we are done. So now suppose $r > 1$. Set

$$V_j = \mathrm{ad}_{X_{i_j}} \ldots \mathrm{ad}_{X_{i_2}} X_{i_1} \text{ for } j = 1, 2, \ldots, r,$$

so that $V_r = V$. We will now show by (decreasing) induction on j that

$$\sup_{u \in U} \mu \left(\int_0^{T_\delta} \langle \mathrm{Ad}_{\xi_\tau} V_j, u \rangle^2 \, dt \leq \varepsilon \right) = O\left(\varepsilon^{\infty-} \right). \tag{5.9}$$

From Equation (5.7), we have

$$d\left[\mathrm{Ad}_{\xi_\tau} V_{j-1}\right] = \mathrm{Ad}_{\xi_\tau} \sum_{i=1}^{k} [X_i, V_{j-1}]\, db_\tau^i + \frac{1}{2} \mathrm{Ad}_{\xi_\tau} \sum_{i=1}^{k} \mathrm{ad}_{X_i}^2 V_{j-1}\, d\tau$$

which upon integrating in τ gives

$$\langle \mathrm{Ad}_{\xi_t} V_{j-1}, u\rangle = \langle V_{j-1}, u\rangle + \int_0^t \langle \mathrm{Ad}_{\xi_\tau}[X_i, V_{j-1}], u\rangle\, dB_\tau^i + \frac{1}{2}\int_0^t \left\langle \mathrm{Ad}_{\xi_\tau} \mathrm{ad}_{X_i}^2 V_{j-1}, u\right\rangle d\tau.$$

Applying Proposition 5.4 of the appendix with $T = T_{\delta,}$

$$\begin{aligned} Y_t &:= \langle \mathrm{Ad}_{\xi_t} V_{j-1}, u\rangle, y = \langle V_{j-1}, u\rangle, \\ M_t &:= \int_0^t \langle \mathrm{Ad}_{\xi_\tau}[X_i, V_{j-1}], u\rangle\, db_\tau^i, \text{ and} \\ A_t &:= \frac{1}{2}\int_0^t \left\langle \mathrm{Ad}_{\xi_\tau} \mathrm{ad}_{X_i}^2 V_{j-1}, u\right\rangle d\tau, \end{aligned}$$

implies that

$$\sup_{u\in U} \mu\left(\Omega_1(u) \cap \Omega_2(u)\right) = O\left(\varepsilon^{\infty-}\right), \tag{5.10}$$

where

$$\begin{aligned} \Omega_1(u) &:= \left\{ \int_0^{T_\delta} \langle \mathrm{Ad}_{\xi_\tau} V_{j-1}, u\rangle^2\, dt < \varepsilon^q \right\}, \\ \Omega_2(u) &:= \left\{ \int_0^{T_\delta} \sum_{i=1}^{k} \langle \mathrm{Ad}_{\xi_\tau}[X_i, V_{j-1}], u\rangle^2\, d\tau \geq \varepsilon \right\}, \end{aligned}$$

and $q > 4$. Since

$$\begin{aligned} \sup_{u\in U} \mu\left([\Omega_2(u)]^c\right) &= \sup_{u\in U} \mu\left(\int_0^{T_\delta} \sum_{i=1}^{k} \langle \mathrm{Ad}_{\xi_\tau}[X_i, V_{j-1}], u\rangle^2\, d\tau < \varepsilon \right) \\ &\leq \sup_{u\in U} \mu\left(\int_0^{T_\delta} \langle \mathrm{Ad}_{\xi_\tau} V_j, u\rangle^2\, d\tau < \varepsilon \right), \end{aligned}$$

we may apply the induction hypothesis to learn

$$\sup_{u\in U} \mu\left([\Omega_2(u)]^c\right) = O\left(\varepsilon^{\infty-}\right). \tag{5.11}$$

It now follows from Equations (5.10) and (5.11) that

$$\begin{aligned}\sup_{u\in U}\mu(\Omega_1(u)) &\le \sup_{u\in U}\mu(\Omega_1(u)\cap\Omega_2(u)) + \sup_{u\in U}\mu(\Omega_1(u)\cap[\Omega_2(u)]^c)\\ &\le \sup_{u\in U}\mu(\Omega_1(u)\cap\Omega_2(u)) + \sup_{u\in U}\mu([\Omega_2(u)]^c)\\ &= O\left(\varepsilon^{\infty-}\right) + O\left(\varepsilon^{\infty-}\right) = O\left(\varepsilon^{\infty-}\right),\end{aligned}$$

that is,

$$\sup_{u\in U}\mu\left(\int_0^{T_\delta}\langle \mathrm{Ad}_{\xi_\tau} V_{j-1}, u\rangle^2\, dt < \varepsilon^q\right) = O\left(\varepsilon^{\infty-}\right).$$

Replacing ε by $\varepsilon^{1/q}$ in the previous equation and using $O\left(\left(\varepsilon^{1/q}\right)^{\infty-}\right) = O\left(\varepsilon^{\infty-}\right)$, completes the induction argument and hence the proof. □

The following proposition is contained in the appendix to [15] and is included here for completeness.

Proposition 5.4 (A martingale inequality). *Let T be a stopping time bounded by $t_0 < \infty$, and let $Y = y + M + A$, where M is a continuous martingale and A is a process of bounded variation such that $M_0 = A_0 = 0$. Further assume that, on the set $\{t \le T\}$, $\langle M\rangle_t$ and $|A|_t$ are absolutely continuous functions and there exist finite positive constants c_1 and c_2 such that*

$$\frac{d\langle M\rangle_t}{dt} \le c_1 \text{ and } \frac{d\,|A|_t}{dt} \le c_2.$$

Then, for all $\nu > 0$ and $q > \nu + 4$, there exist constants $c = c(t_0, q, \nu, c_1, c_2) > 0$ and $\varepsilon_0 = \varepsilon_0(t_0, q, \nu, c_1, c_2) > 0$ such that

$$P\left(\int_0^T Y_t^2 dt < \varepsilon^q,\ \langle Y\rangle_T = \langle M\rangle_T \ge \varepsilon\right) \le 2\exp\left(-\frac{1}{2c\varepsilon^\nu}\right) = O\left(\varepsilon^{-\infty}\right) \qquad (5.12)$$

for all $\varepsilon \in (0, \varepsilon_0]$.

5.3 Spectral decomposition of the Heisenberg heat kernel

In this section, we choose our notation to correspond to that in Thangavelu [48]; any uncited assertions in this section may be found in this text. Let G denote the

3-dimensional Heisenberg Lie group, and let $g = (z,t) \in G$. Then for any $k = 0, 1, 2, \ldots$, let

$$L_k(t) := \frac{1}{k!}\left(\frac{\partial}{\partial x}\right)^k \Big|_{x=0} \frac{e^{-tx/(1-x)}}{1-x}$$

and

$$e_k^\lambda(z,t) := e^{i\lambda t} L_k\left(\frac{1}{2}|\lambda||z|^2\right) e^{-\frac{1}{4}|\lambda||z|^2}.$$

The functions L_k are Laguerre polynomials of type 0, and the following identity is easy to verify,

$$\sum_{k=0}^{\infty} L_k(t)x^k = \frac{e^{-tx/(1-x)}}{1-x}. \tag{5.13}$$

The functions e_k^λ are eigenfunctions of L with corresponding eigenvalues $(2k+1)|\lambda|$; since L is left invariant, this implies that for any $f \in C_c^\infty(G)$

$$L(f * e_k^\lambda) = f * Le_k^\lambda = (2k+1)|\lambda|(f * e_k^\lambda). \tag{5.14}$$

Using these definitions, we have the following proposition.

Proposition 5.5. *The Heisenberg heat kernel may be written as the eigenfunction expansion,*

$$p_\tau(z,t) = \sum_{k=0}^{\infty} \int_{\mathbb{R}} e^{-\tau(2k+1)|\lambda|} e_k^\lambda(z,t)\, d\mu(\lambda),$$

where $d\mu(\lambda) = (2\pi)^{-2}|\lambda| d\lambda$ *is Plancherel measure for* G*.*

Proof. Theorem 2.1.1 of Thangavelu [48] states that, for any $f \in L^2(G)$, we have the expansion

$$f(z,t) = \sum_{k=0}^{\infty} \int_{\mathbb{R}} f * e_k^\lambda(z,t)\, d\mu(\lambda). \tag{5.15}$$

For each $k = 0, 1, 2, \ldots$,

$$\int_{\mathbb{R}} f * e_k^\lambda(z,t)\, d\mu(\lambda) = f * \int_{\mathbb{R}} e_k^\lambda(z,t)\, d\mu(\lambda),$$

where the right hand side should be interpreted in the principal value sense; see Section 2.2 of [48]. Additionally, since $f, e_k^\lambda \in L^2(G)$ and e_k^λ is summable in k, then

$$\sum_{k=0}^{\infty} f * \int_{\mathbb{R}} e_k^\lambda \, d\mu(\lambda) = f * \sum_{k=0}^{\infty} \int_{\mathbb{R}} e_k^\lambda d\mu(\lambda).$$

Then, by Equation (5.14) and the fact that $\left\{e^{-\tau(2k+1)|\lambda|}\right\}_{k=0}^{\infty} \subset L^2(d\mu)$ is summable in k,

$$e^{-\tau L/2} f = e^{-\tau L/2} \left(f * \sum_{k=0}^{\infty} \int_{\mathbb{R}} e_k^\lambda \, d\mu(\lambda) \right) = f * \sum_{k=0}^{\infty} \int_{\mathbb{R}} e^{-\tau(2k+1)|\lambda|} e_k^\lambda \, d\mu(\lambda),$$

and thus we may write the heat kernel of $P_\tau = e^{\tau L/2}$ as

$$p_\tau(z,t) = \sum_{k=0}^{\infty} \int_{\mathbb{R}} e^{-\tau(2k+1)|\lambda|} e_k^\lambda(z,t) \, d\mu(\lambda),$$

as desired. □

We would like to reassure ourselves that this decomposition is equivalent to the expression we already have for the Heisenberg heat kernel in Equation (2.5). Thus, using the definitions of L_k and e_k^λ given at the beginning of this section, we have

$$\begin{aligned}
p_\tau(z,t) &= \sum_{k=0}^{\infty} \int_{\mathbb{R}} e^{-\tau(2k+1)|\lambda|} e_k^\lambda(z,t) \, d\mu(\lambda) \\
&= \frac{1}{4\pi^2} \sum_{k=0}^{\infty} \int_{\mathbb{R}} e^{i\lambda t} e^{-\tau(2k+1)|\lambda|} e^{-\frac{1}{4}|\lambda||z|^2} L_k\left(\frac{1}{2}|\lambda||z|^2\right) |\lambda| \, d\lambda \\
&= \frac{1}{4\pi^2} \int_{\mathbb{R}} e^{i\lambda t} e^{-\tau|\lambda|} e^{-\frac{1}{4}|\lambda||z|^2} |\lambda| \\
&\qquad \times \left[\sum_{k=0}^{\infty} \frac{\left(e^{-2\tau|\lambda|}\right)^k}{k!} \left(\frac{\partial}{\partial x}\right)^k \Big|_{x=0} \frac{e^{-\frac{1}{2}|\lambda||z|^2 x/(1-x)}}{1-x} \right] d\lambda \\
&= \frac{1}{4\pi^2} \int_{\mathbb{R}} e^{i\lambda t} e^{-\tau|\lambda|} e^{-\frac{1}{4}|\lambda||z|^2} |\lambda| \, \frac{\exp\left(-\frac{1}{2}|\lambda||z|^2 \frac{e^{-2\tau|\lambda|}}{1-e^{-2\tau|\lambda|}}\right)}{1-e^{-2\tau|\lambda|}} \, d\lambda \\
&= \frac{1}{4\pi^2} \int_{\mathbb{R}} e^{i\lambda t} \frac{|\lambda|}{e^{\tau|\lambda|} - e^{-\tau|\lambda|}} \exp\left(-\frac{1}{2}|\lambda||z|^2 \left(\frac{1}{2} + \frac{e^{-\tau|\lambda|}}{e^{\tau|\lambda|} - e^{-\tau|\lambda|}}\right)\right) d\lambda \\
&= \frac{1}{8\pi^2} \int_{\mathbb{R}} \frac{\lambda}{\sinh \tau\lambda} \exp\left(-\frac{1}{4}|\lambda||z|^2 \coth \tau\lambda\right) e^{i\lambda t} \, d\lambda,
\end{aligned}$$

where we have applied the identity in Equation (5.13). Thus the expressions are equivalent, and we are satisfied.

5.4 Covariance matrix of ξ_t on the Heisenberg group

In this section, we explicitly compute the Malliavin covariance matrix of ξ in the Heisenberg case,

$$\sigma_t := \xi_t'(b)\xi_t'(b)^* : \mathbb{R}^3 \to \mathbb{R}^3,$$

where the adjoint here is taken relative to the Cameron-Martin and Euclidean metrics. Using Equation (2.21), we calculate $\xi_t^{'}(b) : T_bW \cong H \to T_{\xi_t(b)}G \cong \mathbb{R}^3$ as follows.

$$\begin{aligned}
\xi_t'(b)h = \partial_h \xi_t(b) &= \left.\frac{d}{d\epsilon}\right|_0 \xi_t(b+\epsilon h) \\
&= \left.\frac{d}{d\epsilon}\right|_0 \begin{pmatrix} b_t^1 + \epsilon h_t^1 \\ b_t^2 + \epsilon h_t^2 \\ \frac{1}{2}\int_0^t \left[(b_s^1 + \epsilon h_s^1)(db_s^2 + \epsilon dh_s^2) - (b_s^2 + \epsilon h_s^2)(db_s^1 + \epsilon dh_s^1)\right] \end{pmatrix} \\
&= \begin{pmatrix} h_t^1 \\ h_t^2 \\ \frac{1}{2}\int_0^t \left[h_s^1\, db_s^2 + b_s^1\, dh_s^2 - h_s^2\, db_s^1 - b_s^2\, dh_s^1\right] \end{pmatrix} \\
&= \begin{pmatrix} \int_0^t \dot{h}_s^1\, ds \\ \int_0^t \dot{h}_s^2\, ds \\ \int_0^t \left[\left(\frac{1}{2}b_t^2 - b_s^2\right)\dot{h}_s^1 - \left(\frac{1}{2}b_t^1 - b_s^1\right)\dot{h}_s^2\right] ds \end{pmatrix},
\end{aligned} \tag{5.16}$$

by integration by parts. To determine the action of the adjoint $\xi_t'(b)^* : \mathbb{R}^3 \to H$ on an element $a \in \mathbb{R}^3$, fix $h \in H$, and consider the following computations.

$$\begin{aligned}
(\xi_t'(b)^* a, h)_H &= a \cdot \xi_t'(b)h \\
&= a_1 \int_0^t \dot{h}_s^1\, ds + a_2 \int_0^t \dot{h}_s^2\, ds + a_3 \int_0^t \left[\left(\frac{1}{2}b_t^2 - b_s^2\right)\dot{h}_s^1 - \left(\frac{1}{2}b_t^1 - b_s^1\right)\dot{h}_s^2\right] ds \\
&= \int_0^t \left[\left(a_1 + \frac{1}{2}a_3 b_t^2 - a_3 b_s^2\right)\dot{h}_s^1 + \left(a_2 - \frac{1}{2}a_3 b_t^1 + a_3 b_s^1\right)\dot{h}_s^2\right] ds.
\end{aligned} \tag{5.17}$$

By definition of the Cameron-Martin inner product,

$$(\xi_t'(b)^* a, h)_H = \int_0^1 \frac{d}{ds} \left[\xi_t'(w)^* a\right]_s \cdot \dot{h}_s \, ds,$$

and so Equation (5.17) then implies

$$\frac{d}{ds} \left[\xi_t'(b)^* a\right]_s = \begin{pmatrix} a_1 + \frac{1}{2} a_3 b_t^2 - a_3 b_s^2 \\ a_2 - \frac{1}{2} a_3 b_t^1 + a_3 b_s^1 \end{pmatrix}. \tag{5.18}$$

Use (5.16) to write the vector $\xi_t'(b)\xi_t'(b)^* a \in \mathbb{R}^3$ componentwise:

$$[\xi_t'(b)\xi_t'(b)^* a]_1 = \int_0^t \left(a_1 + \frac{1}{2} a_3 b_t^2 - a_3 b_s^2\right) ds = a_1 t + \frac{1}{2} a_3 t b_t^2 - a_3 \overline{b_t^2},$$

$$[\xi_t'(b)\xi_t'(b)^* a]_2 = \int_0^t \left(a_2 - \frac{1}{2} a_3 b_t^1 + a_3 b_s^1\right) ds = a_2 t - \frac{1}{2} a_3 t b_t^1 + a_3 \overline{b_t^1}, \text{ and}$$

$$\begin{aligned}
&[{\xi_t}'(b)\xi_t'(b)^* a]_3 \\
&\quad = \int_0^t \left[\left(\frac{1}{2} b_t^2 - b_s^2\right)\left(a_1 + \frac{1}{2} a_3 b_t^2 - a_3 b_s^2\right) - \left(\frac{1}{2} b_t^1 - b_s^1\right)\left(a_2 - \frac{1}{2} a_3 b_t^1 + a_3 b_s^1\right)\right] ds \\
&\quad = a_1 \left(\frac{1}{2} t b_t^2 - \overline{b_t^2}\right) - a_2 \left(\frac{1}{2} t b_t^1 - \overline{b_t^1}\right) \\
&\qquad + a_3 \left(\frac{1}{4} t \left((b_t^1)^2 + (b_t^2)^2\right) - b_t^1 \overline{b_t^1} - b_t^2 \overline{b_t^2} + \int_0^t \left((b_s^1)^2 + (b_s^2)^2\right) ds\right),
\end{aligned}$$

where $\overline{b^i} = \int_0^t b_s^i \, ds$ for $i = 1, 2$. So the Malliavin covariance matrix may be written as

$$\xi_t'(b)\xi_t'(b)^* = \begin{pmatrix} t & 0 & \frac{1}{2} t b_t^2 - \overline{b_t^2} \\ 0 & t & -\frac{1}{2} t b_t^1 + \overline{b_t^1} \\ \frac{1}{2} t b_t^2 - \overline{b_t^2} & -\frac{1}{2} t b_t^1 + \overline{b_t^1} & \Xi(t, b) \end{pmatrix}, \tag{5.19}$$

where

$$\Xi(t, b) = \frac{1}{4} t |b|^2 - b_t^1 \overline{b_t^1} - b_t^2 \overline{b_t^2} + \|b\|_2^2,$$

with $|b_t|^2 = (b_t^1)^2 + (b_t^2)^2$ and $\|b\|_2 = \left(\int_0^t |b_s|^2 \, ds\right)^{1/2}$ the norm in $L^2([0, t])$.

Alternately, we may use the more general formula (3.9) for the Malliavin covariance matrix to determine σ_t in the Heisenberg case. For $g = (x, y, z) \in G$ and $(a, b, c) \in \mathbb{R}^3$,

$$\begin{aligned}\mathrm{Ad}_g(a,b,c) &= (x,y,z)(a,b,c)(x,y,z)^{-1} \\ &= (x+a, y+b, z+c+\frac{1}{2}(xb-ay))(-x,-y,-z) = (a,b,c+xb-ay).\end{aligned}$$

Thus, for $X, Y, Z \in \mathfrak{g} \cong \mathbb{R}^3$,

$$\mathrm{Ad}_{\xi_t} X = (1,0,-b_t^2), \mathrm{Ad}_{\xi_t} Y = (0,1,b_t^1), \text{ and } \mathrm{Ad}_{\xi_t} Z = (0,0,1)$$

implies that

$$\mathrm{Ad}_{\xi_t} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -b_t^2 & b_t^1 & 1 \end{pmatrix}.$$

(Equivalently, $\mathrm{Ad}_g = R_{g^{-1}*} L_{g*}$, and so, from Equations (2.3) and (2.4),

$$\mathrm{Ad}_{\xi_t} = R_{\xi_t^{-1}*} L_{\xi_t *} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2}b_t^2 & \frac{1}{2}b_t^1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2}b_t^2 & \frac{1}{2}b_t^1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -b_t^2 & b_t^1 & 1 \end{pmatrix}.)$$

Thus,

$$\mathrm{Ad}_{\xi_s} P \,\mathrm{Ad}_{\xi_s}^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -b_s^2 & b_s^1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & -b_s^2 \\ 0 & 1 & b_s^1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -b_s^2 \\ 0 & 1 & b_s^1 \\ -b_s^2 & b_s^1 & |b_s|^2 \end{pmatrix},$$

and combining this with Equation (3.9) gives

$$\begin{aligned}
\xi_t'(b)\xi_t'(b)^* &= R_{\xi_t *}\left(\int_0^t \mathrm{Ad}_{\xi_s}\, P\, \mathrm{Ad}_{\xi_s}^{\dagger}\; ds\right) R_{\xi_t *}^{\mathrm{tr}} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2}b_t^2 & -\frac{1}{2}b_t^1 & 1 \end{pmatrix} \begin{pmatrix} t & 0 & -\overline{b_t^2} \\ 0 & t & \overline{b_t^1} \\ -\overline{b_t^2} & \overline{b_t^1} & \|b\|_2^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & \frac{1}{2}b_t^2 \\ 0 & 1 & -\frac{1}{2}b_t^1 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} t & 0 & \frac{1}{2}tb_t^2 - \overline{b_t^2} \\ 0 & t & -\frac{1}{2}tb_t^1 + \overline{b_t^1} \\ \frac{1}{2}tb_t^2 - \overline{b_t^2} & -\frac{1}{2}tb_t^1 + \overline{b_t^1} & \Xi(t,b) \end{pmatrix},
\end{aligned}$$

as in Equation (5.19).

The determinant of the covariance matrix may be determined from these computations,

$$\begin{aligned}
\det \sigma_t &= t^2\left[\frac{1}{4}t|b_t|^2 - \frac{1}{2}b_t^1\overline{b_t^1} - \frac{1}{2}b_t^1\overline{b_t^1} + \|b\|_2^2\right] - t\left(-\frac{1}{2}tb_t^1 + \overline{b_t^1}\right)^2 - t\left(-\frac{1}{2}tb_t^2 + \overline{b_t^2}\right)^2 \\
&= t^2\|b\|_2^2 - t(\overline{b_t^1})^2 - t(\overline{b_t^2})^2.
\end{aligned}$$

Writing $|\overline{b_t}|^2 = (\overline{b_t^1})^2 + (\overline{b_t^2})^2$ and $\Delta = \det\sigma$,

$$\Delta_t = t^2\,\|b\|_2^2 - t\left|\overline{b_t}\right|^2.$$

Proposition 5.6. *Let* $B_s = \frac{1}{\sqrt{t}}b_{ts}$. *Then*

$$\Delta_t = \det\sigma_t = t^4\left(\int_0^1 |B_s|^2\,ds - \left|\int_0^1 B_s\,ds\right|^2\right).$$

Proof. Note that

$$\|b\|_2^2 = \int_0^t |b_r|^2\; dr = t\int_0^1 |b_{ts}|^2\; ds = t^2\int_0^1 \left|\frac{1}{\sqrt{t}}b_{ts}\right|^2 ds = t^2\int_0^1 |B_s|^2\,ds,$$

and for $i = 1, 2$

$$\overline{b_t^i} = \int_0^t b_r^i\,dr = t\int_0^1 b_{ts}^i\,ds = t^{3/2}\int_0^1 \frac{1}{\sqrt{t}}b_{ts}^i\,ds = t^{3/2}\int_0^1 B_s^i\,ds.$$

Thus,

$$t^2 \left\| b \right\|_2^2 - t \left| \overline{b_t} \right|^2 = t^4 \left(\int_0^1 \left| B_s \right|^2 ds - \left| \int_0^1 B_s \, ds \right|^2 \right).$$

□

5.5 Heisenberg lifted vector fields

Here we compute the lift of the Heisenberg vector fields,

$$\tilde{X} = \partial_x - \frac{1}{2} y \partial_x, \quad \tilde{Y} = \partial_y + \frac{1}{2} x \partial_z, \quad \text{ and } \tilde{Z} = \partial_z.$$

For now, let $A = (a_1, a_2, a_3)$ be any element of $\mathfrak{g}$. Then

$$\tilde{A}(\xi_t) = L_{\xi_t *} A = \left(a_1, a_2, a_3 + \frac{1}{2}(a_2 b_t^1 - a_1 b_t^2) \right)$$

is an element of $T_{\xi_t} G \cong \mathbb{R}^3$. The inverse of the matrix σ may be computed directly from Equation (5.19), so that for $t > 0$

$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{11} = \Delta_t^{-1} \left(t \|b\|_2^2 - (\overline{b_t^1})^2 - t b_t^2 \overline{b_t^2} + \frac{1}{4} (b_t^2)^2 \right)$$
$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{12} = \Delta_t^{-1} \left(-\overline{b_t^1 b_t^2} + \frac{1}{2} t b_t^2 \overline{b_t^1} + \frac{1}{2} t b_t^1 \overline{b_t^2} - \frac{1}{4} t^2 b_t^1 b_t^2 \right)$$
$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{13} = \Delta_t^{-1} \left(t \overline{b_t^2} - \frac{1}{2} t^2 b_t^2 \right)$$
$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{21} = \Delta_t^{-1} \left(-\overline{b_t^1 b_t^2} + \frac{1}{2} t b_t^2 \overline{b_t^1} + \frac{1}{2} t b_t^1 \overline{b_t^2} - \frac{1}{4} t^2 b_t^1 b_t^2 \right)$$
$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{22} = \Delta_t^{-1} \left(t \|b\|_2^2 - (\overline{b_t^2})^2 - t b_t^1 \overline{b_t^1} + \frac{1}{4} (b_t^1)^2 \right)$$
$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{23} = \Delta_t^{-1} \left(-t \overline{b_t^1} + \frac{1}{2} t^2 b_t^1 \right)$$
$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{31} = \Delta_t^{-1} \left(t \overline{b_t^2} - \frac{1}{2} t^2 b_t^2 \right)$$
$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{32} = \Delta_t^{-1} \left(-t \overline{b_t^1} + \frac{1}{2} t^2 b_t^1 \right)$$
$$\left[\left(\xi_t'(b) \xi_t'(b)^* \right)^{-1} \right]_{33} = \Delta_t^{-1} t^2.$$

Thus, writing $\xi_t'(b)\xi_t'(b)^*\tilde{A}(\xi_t)$ componentwise,

$$
\begin{aligned}
&\big[(\xi_t'(b)\xi_t'(b)^*)^{-1}\tilde{A}(\xi_t)\big]_1 \\
&\quad = a_1\left(t\|b\|_2^2 - (\overline{b_t^1})^2 - tb_t^2\overline{b^2} + \frac{1}{4}(b_t^2)^2\right) + a_2\left(-\overline{b_t^1 b_t^2} + \frac{1}{2}tb_t^2\overline{b_t^1} + \frac{1}{2}tb_t^1\overline{b_t^2} - \frac{1}{4}t^2 b_t^1 b_t^2\right) \\
&\qquad + \left(a_3 + \frac{1}{2}(a_2 b_t^1 - a_1 b_t^2)\right)\left(t\overline{b_t^2} - \frac{1}{2}t^2 b_t^2\right) \\
&\quad = a_1\left(t\|b\|_2^2 - (\overline{b_t^1})^2 - \frac{3}{2}tb_t^2\overline{b_t^2} + \frac{1}{4}(1+t^2)(b_t^2)^2\right) \\
&\qquad + a_2\left(-\overline{b_t^1 b_t^2} + \frac{1}{2}tb_t^2\overline{b_t^1} + tb_t^1\overline{b^2} - \frac{1}{2}t^2 b_t^1 b_t^2\right) + a_3\left(t\overline{b_t^2} - \frac{1}{2}t^2 b_t^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&\big[(\xi_t{}'(b)\xi_t'(b)^*)^{-1}\tilde{A}(\xi_t)\big]_2 \\
&\quad = a_1\left(-\overline{b_t^1 b_t^2} + \frac{1}{2}tb_t^2\overline{b_t^1} + \frac{1}{2}tb_t^1\overline{b_t^2} - \frac{1}{4}t^2 b_t^1 b_t^2\right) + a_2\left(t\|b\|_2^2 - \overline{b}_2^2 - tb_t^1\overline{b_t^1} + \frac{1}{4}(b_t^1)^2\right) \\
&\qquad + \left(a_3 + \frac{1}{2}(a_2 b_t^1 - a_1 b_t^2)\right)\left(-t\overline{b_t^1} + \frac{1}{2}t^2 b_t^1\right) \\
&\quad = a_1\left(-\overline{b_t^1 b_t^2} + tb_t^2\overline{b_t^1} + \frac{1}{2}tb_t^1\overline{b_t^2} - \frac{1}{2}t^2 b_t^1 b_t^2\right) \\
&\qquad + a_2\left(t\|b\|_2^2 - (\overline{b_t^2})^2 - \frac{3}{2}tb_t^1\overline{b_t^1} + \frac{1}{4}(1+t^2)(b_t^1)^2\right) + a_3\left(-t\overline{b_t^1} + \frac{1}{2}t^2 b_t^1\right)
\end{aligned}
$$

$$
\begin{aligned}
&\big[(\xi_t'(b)\xi_t'(b)^*)^{-1}\tilde{A}(\xi_t)\big]_3 \\
&\qquad = a_1\left(t\overline{b_t^2} - \frac{1}{2}t^2 b_t^2\right) + a_2\left(-t\overline{b_t^1} + \frac{1}{2}t^2 b_t^1\right) + \left(a_3 + \frac{1}{2}(a_2 b_t^1 - a_1 b_t^2)\right)t^2 \\
&\qquad = a_1(t\overline{b_t^2} - t^2 b_t^2) - a_2(t\overline{b_t^1} - t^2 b_t^1) + a_3 t^2.
\end{aligned}
$$

Equation (5.18) implies that

$$
\big[\xi_t'(w)^* a\big]_s = \begin{pmatrix} a_1 s + \frac{1}{2}a_3 s b_t^2 - a_3\int_0^s b_r^2\,dr \\ a_2 s - \frac{1}{2}a_3 s b_t^1 + a_3\int_0^s b_r^1\,dr \end{pmatrix},
$$

for any $a \in \mathbb{R}^3$, and so

$$
[\xi_t'(b)^*]_s = \begin{pmatrix} s & 0 & \frac{1}{2}sb_t^2 - \overline{b_s^2} \\ 0 & s & -\frac{1}{2}sb_t^1 + \overline{b_s^1} \end{pmatrix}
$$

Thus, for $0 < s \leq t$,

$$\begin{aligned}
\mathbf{A}_s^1 &= \left[\xi_t'(b)^*(\xi_t'(b)\xi_t'(b)^*)^{-1}\tilde{A}(\xi_t)\right]_1 \\
&= \Delta_t^{-1}\left[a_1 s\left(t\|b\|_2^2 - (\overline{b_t^1})^2 - \frac{3}{2}tb_t^2\overline{b^2} + \frac{1}{4}(1+t^2)(b_t^2)^2\right)\right. \\
&\quad + a_2 s\left(-\overline{b_t^1 b_t^2} + \frac{1}{2}tb_t^2\overline{b_t^1} + tb_t^1\overline{b_t^2} - \frac{1}{2}t^2 b_t^1 b_t^2\right) + a_3 s\left(t\overline{b_t^2} - \frac{1}{2}t^2 b_t^2\right) \\
&\quad \left. + \left(\frac{1}{2}sb_t^2 - \overline{b_s^2}\right)\left(a_1\left(t\overline{b_t^2} - t^2 b_t^2\right) - a_2\left(t\overline{b^1} - t^2 b_t^1\right) + a_3 t^2\right)\right]
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_s^2 &= \left[\xi_t'(b)^*(\xi_t'(b)\xi_t'(b)^*)^{-1}\tilde{A}(\xi_t)\right]_2 \\
&= \Delta_t^{-1}\left[a_1 s\left(-\bar{b}_1\overline{b_t^2} + tb_t^2\overline{b_t^1} + \frac{1}{2}tb_t^1\overline{b_t^2} - \frac{1}{2}t^2 b_t^1 b_t^2\right)\right. \\
&\quad + a_2 s\left(t\|b\|_2^2 - (\overline{b_t^2})^2 - \frac{3}{2}tb_t^1\overline{b_t^1} + \frac{1}{4}(1+t^2)(b_t^1)^2\right) + a_3 s\left(-t\overline{b_t^1} + \frac{1}{2}t^2 b_t^1\right) \\
&\quad \left. + \left(-\frac{1}{2}sb_t^1 + \overline{b_s^1}\right)\left(a_1\left(t\overline{b_t^2} - t^2 b_t^2\right) - a_2\left(t\overline{b^1} - t^2 b_t^1\right) + a_3 t^2\right)\right].
\end{aligned}$$

Thus, $\mathbf{X}_s = (\mathbf{X}_s^1, \mathbf{X}_s^2)$, where $\mathbf{X}^1$ and $\mathbf{X}^2$ are the quadratic expressions in b

$$\mathbf{X}_s^1 = \Delta_t^{-1}\left[s\left(t\|b\|_2^2 - (\overline{b_t^1})^2 - \frac{3}{2}tb_t^2\overline{b_t^2} + \frac{1}{4}(1+t^2)(b_t^2)^2\right) + \left(\frac{1}{2}sb_t^2 - \overline{b_s^2}\right)\left(t\overline{b_t^2} - t^2 b_t^2\right)\right]$$

and

$$\mathbf{X}_s^2 = \Delta_t^{-1}\left[s\left(-\overline{b_t^1 b_t^2} + tb_t^2\overline{b_t^1} + \frac{1}{2}tb_t^1\overline{b^2} - \frac{1}{2}t^2 b_t^1 b_t^2\right) + \left(-\frac{1}{2}sb_t^1 + \overline{b_s^1}\right)\left(t\overline{b_t^2} - t^2 b_t^2\right)\right].$$

Also, $\mathbf{Y}_s = (\mathbf{Y}_s^1, \mathbf{Y}_s^2)$, where

$$\mathbf{Y}_s^1 = \Delta_t^{-1}\left[s\left(-\overline{b_t^1 b_t^2} + \frac{1}{2}tb_t^2\overline{b_t^1} + tb_t^1\overline{b}^2 - \frac{1}{2}t^2 b_t^1 b_t^2\right) + \left(\frac{1}{2}sb_t^2 - \overline{b_s^2}\right)\left(t\overline{b_t^1} - t^2 b_t^1\right)\right]$$

and

$$\begin{aligned}
\mathbf{Y}_s^2 = \Delta_t^{-1}\Bigg[s\left(t\|b\|_2^2 - (\overline{b_t^2})^2 - \frac{3}{2}tb_t^1\overline{b_t^1} + \frac{1}{4}(1+t^2)(b_t^1)^2\right)& \\
+ \left(-\frac{1}{2}sb_t^1 + \overline{b_s^1}\right)\left(t\overline{b_t^1} - t^2 b_t^1\right)\Bigg]&,
\end{aligned}$$

and $\mathbf{Z}_s = (\mathbf{Z}_s^1, \mathbf{Z}_s^2)$ where

$$\mathbf{Z}_s^1 = \Delta_t^{-1}\left[s\left(t\overline{b_t^2} - \frac{1}{2}t^2 b_t^2\right) + \left(\frac{1}{2}sb_t^2 - \overline{b_s^2}\right)t^2\right]$$

and

$$\mathbf{Z}_s^2 = \Delta_t^{-1}\left[s\left(-t\overline{b_t^1} + \frac{1}{2}t^2 b_t^1\right) + \left(-\frac{1}{2}sb_t^1 + \overline{b_s^1}\right)t^2\right].$$

Bibliography

[1] I. Agricola and T. Friedrich. *Global analysis*, volume 52 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002. Differential forms in analysis, geometry and physics, Translated from the 2001 German original by Andreas Nestke.

[2] G. Alexopoulos. An application of homogenization theory to harmonic analysis: Harnack inequalities and Riesz transforms on Lie groups of polynomial growth. *Canad. J. Math.*, 44(4):691–727, 1992.

[3] P. Auscher, T. Coulhon, X. Duong, and S. Hofmann. Riesz transform on manifolds and heat kernel regularity. *Preprint*, 2004.

[4] D. Bakry. Ricci curvature and dimension for diffusion semigroups. In *Stochastic processes and their applications in mathematics and physics (Bielefeld, 1985)*, volume 61 of *Math. Appl.*, pages 21–31. Kluwer Acad. Publ., Dordrecht, 1990.

[5] D. Bakry and M. Émery. Hypercontractivité de semi-groupes de diffusion. *C. R. Acad. Sci. Paris Sér. I Math.*, 299(15):775–778, 1984.

[6] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[7] D. R. Bell. *The Malliavin calculus*, volume 34 of *Pitman Monographs and Surveys in Pure and Applied Mathematics*. Longman Scientific & Technical, Harlow, 1987.

[8] D. R. Bell and S. E. A. Mohammed. The Malliavin calculus and stochastic delay equations. *J. Funct. Anal.*, 99(1):75–99, 1991.

[9] A. Bellaïche. Tangent space in sub-Riemannian geometry. In *Sub-Riemannian Geometry*, volume 144 of *Progress in Mathematics*, pages 1–78. Birkhäuser, Basel, 1996.

[10] V. I. Bogachev. *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998.

[11] T. Coulhon and X. T. Duong. Riesz transform and related inequalities on noncompact Riemannian manifolds. *Comm. Pure Appl. Math.*, 56(12):1728–1751, 2003.

[12] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.

[13] E. B. Davies. *One-parameter semigroups*, volume 15 of *London Mathematical Society Monographs*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1980.

[14] B. K. Driver. A Cameron-Martin type quasi-invariance theorem for Brownian motion on a compact Riemannian manifold. *J. Funct. Anal.*, 110(2):272–376, 1992.

[15] B. K. Driver. Curved wiener space analysis. *Preprint, to appear in "Real and Stochastic Analysis: New Perspectives," xxx.lanl.gov/list/math/0403*, 2004.

[16] B. K. Driver and L. Gross. Hilbert spaces of holomorphic functions on complex Lie groups. In *New trends in stochastic analysis (Charingworth, 1994)*, pages 76–106. World Sci. Publishing, River Edge, NJ, 1997.

[17] K. D. Elworthy. *Stochastic differential equations on manifolds*, volume 70 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1982.

[18] G. B. Folland. A fundamental solution for a subelliptic operator. *Bull. Amer. Math. Soc.*, 79:373–376, 1973.

[19] G. B. Folland and E. M. Stein. *Hardy Spaces on Homogeneous Groups*. Princeton University Press, Princeton, N.J., 1982.

[20] M. Fukushima. *Dirichlet forms and Markov processes*, volume 23 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1980.

[21] Ĭ. Ī. Gīhman and A. V. Skorohod. *Stochastic differential equations*. Springer-Verlag, New York, 1972. Translated from the Russian by Kenneth Wickwire, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 72.

[22] R. W. Goodman. *Nilpotent Lie groups: structure and applications to analysis*. Springer-Verlag, Berlin, 1976. Lecture Notes in Mathematics, Vol. 562.

[23] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[24] L. Gross. Logarithmic Sobolev inequalities and contractivity properties of semigroups. In *Dirichlet forms (Varenna, 1992)*, volume 1563 of *Lecture Notes in Math.*, pages 54–88. Springer, Berlin, 1993.

[25] E. Hille. *Functional Analysis and Semi-Groups.* American Mathematical Society Colloquium Publications, vol. 31. American Mathematical Society, New York, 1948.

[26] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[27] E. P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[28] N. Ikeda and S. Watanabe. An introduction to Malliavin's calculus. In *Stochastic analysis (Katata/Kyoto, 1982)*, volume 32 of *North-Holland Math. Library*, pages 1–52. North-Holland, Amsterdam, 1984.

[29] N. Ikeda and S. Watanabe. Malliavin calculus of Wiener functionals and its applications. In *From local times to global geometry, control and physics (Coventry, 1984/85)*, volume 150 of *Pitman Res. Notes Math. Ser.*, pages 132–178. Longman Sci. Tech., Harlow, 1986.

[30] 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.

[31] K. Itô, editor. *Proceedings of the International Symposium on Stochastic Differential Equations*. John Wiley & Sons, New York-Chichester-Brisbane, 1978. Held at the Research Institute for Mathematical Sciences, Kyoto Univ., Kyoto, July 9–14, 1976, A Wiley-Interscience Publication.

[32] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. I. In *Stochastic analysis (Katata/Kyoto, 1982)*, volume 32 of *North-Holland Math. Library*, pages 271–306. North-Holland, Amsterdam, 1984.

[33] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 32(1):1–76, 1985.

[34] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 34(2):391–442, 1987.

[35] M. Ledoux. The geometry of Markov diffusion generators. *Ann. Fac. Sci. Toulouse Math. (6)*, 9(2):305–366, 2000. Probability theory.

[36] P. Malliavin. C^k-hypoellipticity with degeneracy. In *Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978)*, pages 199–214. Academic Press, New York, 1978.

[37] P. Malliavin. C^k-hypoellipticity with degeneracy. II. In *Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978)*, pages 327–340. Academic Press, New York, 1978.

[38] P. Malliavin. Stochastic calculus of variation and hypoelliptic operators. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*, pages 195–263, New York, 1978. Wiley.

[39] J. Norris. Simplified Malliavin calculus. In *Séminaire de Probabilités, XX, 1984/85*, volume 1204 of *Lecture Notes in Math.*, pages 101–130. Springer, Berlin, 1986.

[40] D. Nualart. *The Malliavin calculus and related topics.* Probability and its Applications (New York). Springer-Verlag, New York, 1995.

[41] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1983.

[42] J. Picard. Gradient estimates for some diffusion semigroups. *Probab. Theory Related Fields*, 122(4):593–612, 2002.

[43] L. P. Rothschild and E. M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137(3-4):247–320, 1976.

[44] I. Shigekawa. Derivatives of Wiener functionals and absolute continuity of induced measures. *J. Math. Kyoto Univ.*, 20(2):263–289, 1980.

[45] D. W. Stroock. The Malliavin calculus and its application to second order parabolic differential equations. I. *Math. Systems Theory*, 14(1):25–65, 1981.

[46] D. W. Stroock. The Malliavin calculus and its application to second order parabolic differential equations. II. *Math. Systems Theory*, 14(2):141–171, 1981.

[47] S. Taniguchi. Malliavin's stochastic calculus of variations for manifold-valued Wiener functionals and its applications. *Z. Wahrsch. Verw. Gebiete*, 65(2):269–290, 1983.

[48] S. Thangavelu. *Harmonic analysis on the Heisenberg group*, volume 159 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1998.

[49] V. S. Varadarajan. *Lie groups, Lie algebras, and their representations*, volume 102 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984. Reprint of the 1974 edition.

[50] N. T. Varopoulos. Analysis on nilpotent groups. *J. Funct. Anal.*, 66(3):406–431, 1986.

[51] N. T. Varopoulos, L. Saloff-Coste, and T. Coulhon. *Analysis and geometry on groups*, volume 100 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1992.

[52] S. Watanabe. *Lectures on stochastic differential equations and Malliavin calculus*, volume 73 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Published for the Tata Institute of Fundamental Research, Bombay, 1984. Notes by M. Gopalan Nair and B. Rajeev.

[53] S. Watanabe. Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. *Ann. Probab.*, 15(1):1–39, 1987.

[54] S. Watanabe. Stochastic analysis on Wiener space. In *Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI, 1994)*, volume 52 of *Proc. Sympos. Appl. Math.*, pages 359–370, Providence, RI, 1997. Amer. Math. Soc.

[55] K. Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, sixth edition, 1980.