TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 368, Number 2, February 2016, Pages 989–1022
http://dx.doi.org/10.1090/tran/6461
Article electronically published on June 11, 2015

HYPOELLIPTIC HEAT KERNELS ON INFINITE-DIMENSIONAL HEISENBERG GROUPS

BRUCE K. DRIVER, NATHANIEL ELDREDGE, AND TAI MELCHER

Abstract. We study the law of a hypoelliptic Brownian motion on an infinite-dimensional Heisenberg group based on an abstract Wiener space. We show that the endpoint distribution, which can be seen as a heat kernel measure, is absolutely continuous with respect to a certain product of Gaussian and Lebesgue measures, that the heat kernel is quasi-invariant under translation by the Cameron–Martin subgroup, and that the Radon–Nikodym derivative is Malliavin smooth.

Contents

1. Introduction 989
2. Quadratic Brownian integrals in finite dimensions 995
3. Quadratic Brownian integrals in infinite dimensions 997
4. Heisenberg heat kernels 1000
5. Estimates on ρ_T 1002
6. A Fernique-type theorem 1006
7. Quasi-invariance of hypoelliptic heat kernel measures 1007
8. Smoothness properties of hypoelliptic heat kernel measures 1010
9. Conclusion 1016
Appendix A. Another proof of Theorem 2.1 1017
Acknowledgments 1020
References 1021

1. Introduction

In [12] (also see [13, 14]), M. Gordina and the first author began studying the properties of elliptic heat kernel measures on certain infinite-dimensional Heisenberg-like groups. There it was shown that these heat kernel measures enjoyed many of the quasi-invariance and other smoothness properties found in finite-dimensional settings and in commutative abstract Wiener space examples. Later Baudoin, Gordina, and the third author [5] proved similar results for hypoelliptic heat kernel

Received by the editors December 12, 2013.
2010 *Mathematics Subject Classification.* Primary 58J35; Secondary 58J65, 60B15.
Key words and phrases. Heisenberg group, hypoelliptic, heat kernel, smooth measures.
The first author's research was supported in part by NSF Grant DMS-1106270.
The second author's research was supported in part by NSF Grant DMS-0739164.
The third author's research was supported in part by NSF Grants DMS-0907293 and DMS-1255574.

©2015 American Mathematical Society

measures on infinite-dimensional Heisenberg-like groups. The quasi-invariance result proved there relied on detailed curvature-dimension inequalities first suggested by Baudoin, Bonnefont, and Garofalo [3, 4].

The aim of the current paper is to revisit the hypoelliptic setting and to show that pure stochastic calculus techniques may be used to re-prove and actually strengthen the main results for heat kernel measures in [12] and [5]. This is done by developing a concrete representation for the heat kernel measure, which allows us to show that, in fact, these measures satisfy a strong definition of smoothness, as given for example in [11]. Typically such smoothness results have been unavailable in the infinite-dimensional subelliptic context and alternative interpretations must be made (for example, as smoothness of all appropriate finite-dimensional projections of the measure; see [6, 29, 31]). To the authors' knowledge, the smoothness results in the present paper are the first of their type in the infinite-dimensional subelliptic setting.

1.1. **Heat kernel measures on finite-dimensional Lie groups.** As motivation, let us briefly recall the finite-dimensional situation. Let G be a finite-dimensional simply connected Lie group, and let $\mathfrak{g}$ be the Lie algebra of G, identified with the set of left-invariant vector fields on G. Suppose $\{X_j\}_{j=1}^k \subset \mathfrak{g}$ is such that $\mathrm{Lie}(\{X_j\}_{j=1}^k) = \mathfrak{g}$. Then $\{X_j\}_{j=1}^k$ satisfies the "bracket generating" hypothesis of Hörmander's Theorem [18], which then asserts that the "sub-Laplacian" operator $L := X_1^2 + \cdots + X_k^2$ is **hypoelliptic**: if ϕ is a distribution such that $L\phi \in C^\infty(G)$, then in fact $\phi \in C^\infty(G)$. Similarly, the fundamental solution or heat kernel $p_t(x, y)$ of the heat equation $u_t - \frac{1}{2}Lu = 0$ is C^∞. In some examples, one can write down an explicit integral formula for p_t from which its smoothness is apparent. These formulae have been derived many times in the literature; as a small sample, we mention [1, 17, 28, 33].

Furthermore, the heat kernel $p_t(x, y)$ is strictly positive for all $x, y \in G$ and $t > 0$. In particular, the heat kernel measure $p_t(e, \cdot)\, dm$, where m is Haar measure, is quasi-invariant under the action of G on itself by left or right translation. This strict positivity is typically not immediately apparent even in the settings where an explicit formula for the heat kernel is available, as these formulae often involve oscillatory integrals. (However, see [8] for an elementary proof of strict positivity for the heat kernel on the real three-dimensional Heisenberg group.) More generally, strict positivity can be shown to hold through deeper means, for instance, by the use of parabolic Harnack inequalities; see for example [34] and the references therein. On the other hand, for the formula we derive in this paper (Corollary 4.7), which applies to step-two nilpotent Lie groups (of potentially infinite dimension), strict positivity is obvious by inspection (see Corollary 4.8).

In the finite-dimensional nilpotent setting, one can obtain rather precise information about the integrability of p_t and its derivatives; in particular, p_t has Gaussian decay at infinity, with respect to the Carnot–Carathéodory distance on G. This is proved in a general setting in [34]; in the special case of H-type groups, sharper estimates were obtained in [15]. In the infinite-dimensional setting of the present paper, we are not able to obtain such precise results, but we are able to derive a Fernique-type theorem (Proposition 6.1), which together with the L^p-integrability of derivatives (Corollary 8.11) points roughly in the same direction.

One may think of p_t more probabilistically as the endpoint distribution of a Brownian motion on G. That is, define a Brownian motion on G to be the unique

process g_t starting at the identity of G and solving the Stratonovich stochastic differential equation

$$\delta g_t = \sum_i X_i(g_t)\,\delta B_t^{(i)}$$

where $B_t = (B_t^{(1)}, \ldots, B_t^{(k)})$ is a standard k-dimensional Brownian motion. Then L is the generator of the Markov process g_t, and p_t is the transition density of g_t. In particular, for each t, the law of g_t is absolutely continuous with respect to Haar measure, and its density $p_t(e,\cdot)$ is strictly positive and C^∞. Intuitively, despite being driven by a Brownian motion whose dimension is in general smaller than that of G, the process g_t is still able to wander throughout the group G.

The purpose of this paper is to obtain, by fairly elementary methods, analogous results for a class of infinite-dimensional Heisenberg-like groups modeled on an abstract Wiener space. A key idea is to replace Haar measure, which no longer exists on our infinite-dimensional groups, by Gaussian measures on the relevant abstract Wiener space.

We end this section with some standard notation that we will use throughout the rest of the paper. If $(\Omega, \mathcal{F}, \mu)$ is a probability space and $X : \Omega \to \mathbb{R}$ is a random variable we will denote $\int_\Omega X d\mu$ by either $\mathbb{E}X$, $\mathbb{E}_\mu X$, or simply by $\mu(X)$.

1.2. **Heat kernel measures on infinite-dimensional Heisenberg-like groups.** Here we recall infinite-dimensional Heisenberg-like groups as first constructed in [12]. We also define hypoelliptic Brownian motion and heat kernel measure on these spaces. Infinite-dimensional Heisenberg-like groups are central extensions of an abstract Wiener space, for which we record here a standard definition.

Definition 1.1. Let (W, H, μ) be a real **abstract Wiener space**, that is, W is a separable Banach space, H is a Hilbert space (known as the **Cameron–Martin subspace**) which embeds continuously into W, and μ is the Gaussian Borel measure on W determined by

$$\mathbb{E}_\mu\left[e^{i\varphi}\right] = \exp\left(-\frac{1}{2}\left\|\varphi|_H\right\|_{H^*}^2\right) \text{ for all } \varphi \in W^*.$$

Thus, the covariance of μ is determined by the inner product of H. We shall assume for simplicity that (W, H, μ) is non-degenerate: H is dense in W, and (equivalently) the support of the measure μ is all of W.

Notation 1.2. The adjoint of the continuous dense embedding $H \to W$ is a continuous dense embedding $W^* \to H$. We denote by W_* the image of W^* in H under this embedding. Equivalently, W_* consists of those $h \in H$ such that the linear functional $\langle \cdot, h\rangle \in H^*$ extends continuously to W. Recall, though, that even for $h \in H \setminus W_*$, the "functional" $W \ni x \mapsto \langle x, h\rangle_H$ makes sense as an element of $L^2(W, \mu)$.

Notation 1.3. For $T > 0$, let μ_T be the dilation of μ defined by $\mu_T(A) = \mu(T^{-1/2}A)$. ($\mu_0$ is then a point mass at 0.)

For further background on abstract Wiener space and Gaussian measures, see for example [7, 21].

Definition 1.4. Let C be a finite-dimensional real Hilbert space of dimension d, whose inner product is denoted by " $\cdot$ ", and suppose that $\omega : W \times W \to C$ is a

continuous, skew-symmetric bilinear form. As explained in [12, Proposition 3.14], we may make $W \times C$ into a Banach Lie group G using the group multiplication law

$$(w_1, c_1)(w_2, c_2) = \left(w_1 + w_2, c_1 + c_2 + \frac{1}{2}\omega(\omega_1, w_2) \right).$$

We shall call such a group **Heisenberg-like**, by analogy with the classical Heisenberg group (see Example 1.5 below). The identity of G is $(0, 0)$, which we denote as $\mathbf{e}$, and the inverse operation is given by $(w, c)^{-1} = (-w, -c)$. The subset $G_{CM} := H \times C$ is a subgroup of G, which we call the **Cameron–Martin subgroup**.

The Lie algebra $\mathfrak{g}$ of G may be identified with $W \times C$ equipped with the Lie bracket defined by

$$[(w_1, c_1), (w_2, c_2)] = (0, \omega(w_1, w_2))$$

for all $w_1, w_2 \in W$ and $c_1, c_2 \in C$. Then $\mathfrak{g}$ is a nilpotent Banach Lie algebra of step two. Under this identification, the exponential map is the identity. The subset $\mathfrak{g}_{CM} := H \times C$ is a Lie subalgebra of $\mathfrak{g}$, which we call the **Cameron–Martin subalgebra**.

Throughout this paper, we shall assume that ω is surjective, or in other words that the Lie algebra $\mathfrak{g}$ is generated by its subspace $W \times \{0\}$. This is the analogue of Hörmander's bracket generating condition.

Example 1.5. If $W = \mathbb{R}^2$, $C = \mathbb{R}$, and ω is defined by

$$\omega((x_1, y_1), (x_2, y_2)) = x_1 y_2 - x_2 y_1,$$

then the group G constructed above is (isomorphic to) the classical three-dimensional Heisenberg group $\mathbb{H}_3$.

Definition 1.6. A **smooth cylinder function** on G is a function $F : G \to \mathbb{R}$ of the form

$$F(x, c) = \psi(f_1(x), \ldots, f_n(x), c)$$

for some $n \geq 0$, $f_1, \ldots, f_n \in W^*$, and $\psi \in C_c^\infty(\mathbb{R}^n \times C)$.

Notation 1.7. Given $h \in H$ and $z \in C$, we can compute the partial derivative of a smooth cylinder function F in the (h, z) direction as

$$\partial_{(h,z)} F(x, c) = \sum_{i=1}^{n} \partial_i \psi(f_1(x), \ldots, f_n(x), c) \langle f_i, h \rangle_H + \partial_z F(x, c).$$

Note that in the inner product $\langle f_i, h \rangle_H$ we are identifying $f_i \in W^*$ with its image in $W_* \subset H$ as explained in Notation 1.2. We observe that $\partial_{(h,z)} F$ is another smooth cylinder function.

Notation 1.8. We may identify each $X = (h, z) \in \mathfrak{g}_{CM}$ with a left-invariant vector field $\tilde{X}$, or first-order differential operator, acting on the cylinder functions on G via

$$\tilde{X} F(g) = \partial_{(h,z)}(F \circ L_g)(\mathbf{e})$$

where L_g is left translation by $g \in G$, that is, $L_g(k) = gk$. A simple computation [12, Proposition 3.7] shows that

$$\tilde{X} F(x, c) = \left(\partial_{\left(h, z + \frac{1}{2}\omega(x,h)\right)} F \right)(x, c). \tag{1.1}$$

We may now define a group Brownian motion on G and the associated heat kernel measure.

Definition 1.9. Let $\{B_t\}_{t\geq 0}$ be a W**-valued standard Brownian motion**. That is, $\{B_t\}$ is a continuous, adapted, W-valued stochastic process defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, such that for all $0 \leq s \leq t$, we have that $B_t - B_s$ is independent of $\mathcal{F}_s$ and has μ_{t-s} as its distribution, where μ_t is as in Notation 1.3.

See [21] for more information about Brownian motion on abstract Wiener space.

Definition 1.10. A **hypoelliptic Brownian motion on** G is the G-valued stochastic process $\{g_t\}_{t\geq 0}$ which is the solution to the stochastic differential equation,

$$dg_t = (L_{g_t})_* \delta B_t, \quad g_0 = e$$

where $(L_x)_*$ denotes the differential of left translation by $x \in G$. (This stochastic differential equation can be interpreted in either the sense of Itô or Stratonovich; the solutions in the step-two nilpotent setting are the same.) The solution may be written formally as

$$g_t = \left(B_t, \frac{1}{2} \int_0^t \omega \left(B_s, dB_s \right) \right). \tag{1.2}$$

(This formal expression will be made precise in Section 4 below.) For fixed $t > 0$, the measure $\nu_t = \text{Law}(g_t)$ will be called the **hypoelliptic heat kernel measure on** G (at time t).

Example 1.11. Let us return to Example 1.5, where $G = \mathbb{H}_3$ is the classical three-dimensional Heisenberg group. If we write the two-dimensional Brownian motion B_t on $W = \mathbb{R}^2$ in terms of its components as $B_t = (B_t^{(1)}, B_t^{(2)})$, then we have

$$g_t = \left(B_t, \frac{1}{2} \int_0^t B_s^{(1)} \, dB_s^{(2)} - B_s^{(2)} \, dB_s^{(1)} \right)$$

so that g_t is just the two-dimensional Brownian motion B_t coupled with its stochastic Lévy area process.

Let Λ be an orthonormal basis for H, and for F a smooth cylinder function as in Definition 1.6 let

$$LF := \sum_{h\in\Lambda} \widetilde{(h,0)}^2 F, \quad \Delta_H F := \sum_{h\in\Lambda} \partial^2_{(h,0)} F$$

where $\partial_{(h,0)}$ and $\widetilde{(h,0)}$ are as defined in Notations 1.7 and 1.8, respectively. A computation analogous to that yielding (1.1) (see [12, Proposition 3.29]) shows that

$$(LF)(x,c) = (\Delta_H F)(x,c) + \sum_{h\in\Lambda} \left(\left(\frac{1}{4} \partial^2_{(0,\omega(x,h))} + \partial_{(0,\omega(x,h))} \partial_{(h,0)} \right) F \right)(x,c).$$

Furthermore, from [12, Corollary 4.5], ν_t weakly solves the heat equation

$$\partial_t \nu_t = \frac{1}{2} L \nu_t \text{ with } \lim_{t\downarrow 0} \nu_t = \delta_{\mathbf{e}},$$

that is, for all smooth cylinder functions F as in Definition 1.6,

$$\nu_T(F) = F(\mathbf{e}) + \frac{1}{2} \int_0^T \nu_t (LF) \, dt, \quad \text{for all } T > 0.$$

1.3. **Summary of results.** Here we briefly describe the main results proved in this paper. A key ingredient in our results is the identity of Theorem 3.9, proved in certain finite-dimensional cases by M. Yor (see Section 2): if B_t is a Brownian motion on W and $A : H \to H$ is Hilbert–Schmidt, then for any $T \geq 0$ and any bounded measurable $f : W \to \mathbb{R}$, we have

$$\mathbb{E}\left[f(B_T)\, e^{i\int_0^T \langle AB_t, dB_t\rangle_H}\right] = \mathbb{E}\left[f(B_T)\, e^{-\frac{1}{2}\int_0^T \|AB_t\|_H^2\, dt}\right]. \tag{1.3}$$

(The process AB_t and the stochastic integral on the left side are defined precisely in Section 3; see Propositions 3.6 and 3.8.) Using the identity (1.3), in Section 4 we derive a formula for the heat kernel ν_T. For $\lambda \in C$, define the Hilbert–Schmidt operator $\Omega_\lambda : H \to H$ by $\langle \Omega_\lambda h, k\rangle_H = \omega(h,k)\cdot\lambda$ for $h, k \in H$, and let ρ_T be the random linear transformation on C defined by

$$\rho_T \lambda \cdot \lambda' = \frac{1}{4}\int_0^T \langle \Omega_\lambda B_t, \Omega_{\lambda'} B_t\rangle_H\, dt.$$

In this notation, our Corollary 4.7 states that

$$\nu_T(dx, dc) = \gamma_T(x,c)\mu_T(dx)m(dc)$$

where m is Lebesgue measure on C and

$$\gamma_T(x,c) := \mathbb{E}\left[\frac{\exp\left(-\frac{1}{2}\rho_T^{-1}c\cdot c\right)}{\sqrt{(2\pi)^d \det\rho}}\,\middle|\, B_T = x\right]. \tag{1.4}$$

In particular, ν_T is absolutely continuous with respect to product measure $\mu_T \otimes m$ on $G = W \times C$. To the best of our knowledge, the formula (1.4) is new even in the finite-dimensional case.

As a first application of (1.4), we prove in Section 6 a Fernique-type theorem (see Proposition 6.1) on the integrability of ν_T: there exists an $\varepsilon > 0$ such that

$$\int e^{\frac{\varepsilon}{T}(\|x\|_W^2 + |c|_C)}\, \nu_T(dx, dc) < \infty.$$

Necessary ingredients include estimates on the integrability of ρ_T^{-1}, which are the subject of Section 5.

By further analysis of the formula (1.4), and use of the estimates of Section 5, we show in Sections 7 and 8 that the heat kernel is quasi-invariant under translations by G_{CM} and is infinitely differentiable in those directions. In particular, for $\tilde{X}$ the left-invariant vector field on G associated to $X \in \mathfrak{g}_{CM}$,

$$\tilde{X}^* = -\tilde{X} + \psi_T^X$$

where $\psi_T^X : G \to \mathbb{R}$, essentially the first logarithmic derivative of ν_T, is a "Malliavin smooth" function in the sense that ψ_T^X is in $L^{\infty-}(\nu_T) = \bigcap_{p\in[1,\infty)} L^p(\nu_T)$ and all of its derivatives exist and are also in $L^{\infty-}(\nu_T)$; see Lemma 8.7 and Corollaries 8.10 and 8.11. Similar techniques were used in [9] to study elliptic heat kernel measures on G.

1.4. **Measurable group actions.** It was first observed in [12] that the assumption that ω be a continuous operator on $W \times W$ is not strictly necessary. A standard fact about abstract Wiener spaces is that, if K is any Hilbert space and $T : W \to K$ is a continuous linear operator, then the restriction $T|_H : H \to K$ is Hilbert–Schmidt [21, Corollary 1.4.4]. It follows that, if $\omega : W \times W \to C$ is continuous, then its

restriction $\omega|_{H\times H} : H \times H \to C$ is Hilbert–Schmidt, in the sense that for any orthonormal basis Λ of H, we have $\|\omega\|_{HS}^2 := \sum_{h,k\in\Lambda} |\omega(h,k)|_C^2 < \infty$.

Conversely, suppose that we are merely given a skew-symmetric bilinear form $\omega : H \times H \to C$ which is Hilbert–Schmidt. We will show in Section 3 that the formula (1.2) still makes sense, *even if ω does not extend continuously to $W \times W$*. Indeed, the results of this paper (which are all expressed in terms of the process g_t) will be proved under this weaker assumption. In this case, there is no canonical way to extend ω to $W \times W$, and thus it does not really make sense to speak of G as a group. However, the Cameron–Martin group G_{CM} is still perfectly well-defined, and we obtain a measurable left and right action of G_{CM} on the measurable space G. In particular, fix $h \in H$ and let $T_h : H \to C$ be the linear map given by $T_h k := \omega(h,k)$ for any $k \in H$. Then, for any orthonormal basis $\{e_j\}_{j=1}^d$ of C, we have that the linear functional $\ell_j : H \to \mathbb{R}$ defined by $\ell_j(k) := T_h k \cdot e_j = \omega(h,k)\cdot e_j$ extends to a μ-measurable linear functional $\hat{\ell}_j$ on W; see for example [7, Theorem 2.10.11]. Thus, we may define

$$\omega(h,w) := \sum_{j=1}^{d} \hat{\ell}_j(w) e_j, \tag{1.5}$$

and hence

$$(w,c) \mapsto (h,z)(w,c) := \left(h+w, z+c+\frac{1}{2}\omega(h,w)\right) \tag{1.6}$$

is a measurable transformation on G. In fact, under the assumption that ω is Hilbert–Schmidt, this does not depend on C being finite-dimensional. The adjoint $T_h^* : C \to H$ is also Hilbert–Schmidt, and we may write $\ell_j(k) = \langle k, T_h^* e_j\rangle_H$. Then $\|\hat{\ell}_j\|_{L^2(\mu)} = \|T_h^* e_j\|_H$, which implies that

$$\sum_{j=1}^{\infty} \|\hat{\ell}_j\|_{L^2(\mu)}^2 < \infty,$$

and so $\sum_{j=1}^{\infty} \hat{\ell}_j(w)^2 < \infty$ μ-a.s and equation (1.5) makes sense with $d = \infty$. Thus, for example, the quasi-invariance results of Section 7 can be interpreted as statements about how the measure ν_t on G behaves under the left action by elements $(h,z) \in G_{CM}$ as in (1.6), and the analogously defined right action. Further discussion can be found in [12, Section 9].

For concreteness, however, we encourage the reader to continue to think of the case when ω does have a continuous extension to $W \times W$, in which G is an honest group.

2. Quadratic Brownian integrals in finite dimensions

This section is devoted to the discussion of the identity (1.3) in the case where W is finite-dimensional.

Theorem 2.1. *Let $\{B_t\}_{t\geq 0}$ be an N-dimensional Brownian motion, A be an $N\times N$ skew-symmetric matrix, and $T > 0$. Then, for any measurable $f : \mathbb{R}^N \to \mathbb{C}$ such that $\mathbb{E}\,|f(B_T)| < \infty$,*

$$\mathbb{E}\left[f(B_T)\, e^{i\int_0^T AB_t\cdot dB_t}\right] = \mathbb{E}\left[f(B_T)\, e^{-\frac{1}{2}\int_0^T |AB_t|^2 dt}\right]. \tag{2.1}$$

Remark 2.2. This result was proved in the case of the three-dimensional Heisenberg group $\mathbb{H}_3$ (in which $N = 2$, $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, and $\int_0^T AB_t \cdot dB_t$ is the Lévy area process) by M. Yor [36]. An alternative derivation can be found in [30, Section 2.1]. The relationship between the integrals $\int_0^T AB_t \cdot dB_t$ and $\int_0^T |AB_t|^2\, dt$ was studied by P. Lévy [24, 25] but to the best of our understanding the identity (2.1) is not contained in his work. A similar computation appears in [20] which effectively obtains (2.1) in the case $f = 1$. Here we provide a proof based on Yor's result for the stochastic Lévy area. In Appendix A, we also provide another self-contained proof of Theorem 2.1, based on analysis of the infinitesimal generator of g_t.

Proof of Theorem 2.1. We include here a sketch of Yor's argument for $\mathbb{H}_3$; we will then show how the general case follows. Suppose $N = 2$ and $A = \left(\begin{smallmatrix} 0 & a \\ -a & 0 \end{smallmatrix}\right)$ for some $a \in \mathbb{R}$. By the rotational invariance of the Brownian motion B and the fact that A commutes with rotations, it is sufficient to establish (2.1) with $f(B_T)$ replaced by $g(|B_T|)$; that is, to show that

$$\mathbb{E}\left[e^{i \int_0^T AB_t \cdot dB_t} \,\middle|\, |B_T| \right] = \mathbb{E}\left[e^{-\frac{1}{2}\int_0^T |AB_t|^2\, dt} \,\middle|\, |B_T| \right]. \tag{2.2}$$

Now we observe that

$$\beta_t := \int_0^t \frac{B_s}{|B_s|} \cdot dB_s, \quad \gamma_t := \int_0^t \frac{AB_s}{|AB_s|} \cdot dB_s$$

are two independent one-dimensional Brownian motions. (The integrals are well-defined because, almost surely, $B_t \neq 0$ for almost every t. They can be seen to be Brownian motions by Lévy's characterization; each is a continuous martingale with quadratic variation t. Finally, since A is skew-symmetric, it follows from Itô's isometry that β, γ are uncorrelated and hence independent.) We note that since $|AB_s| = a|B_s|$, we can write

$$Z_t = \int_0^t AB_s \cdot dB_s = a \int_0^t |B_s|\, d\gamma_s. \tag{2.3}$$

Now if we let $S_t = |B_t|^2$, by Itô's formula we have

$$S_t = 2\int_0^t B_s \cdot dB_s + 2t = 2\int_0^t \sqrt{S_t}\, d\beta_t + 2t.$$

In particular, S is $\sigma(\beta)$-measurable, and hence S (and also $|B| = \sqrt{S}$) is independent of the process γ. (For details, see [35] and the references therein.) Thus, (2.3) implies that, conditioned on $|B|$, Z_T is Gaussian with variance given by $a^2 \int_0^T |B_t|^2\, dt = \int_0^T |AB_t|^2\, dt$. Hence by the Gaussian Fourier transform, we have

$$\mathbb{E}\left[e^{iZ_T} \,\middle|\, |B| \right] = e^{-\frac{1}{2}\int_0^T |AB_t|^2\, dt}.$$

Conditioning on $|B_T|$ we have (2.2).

For arbitrary N, we begin with the case that A is quasi-diagonal, that is, block diagonal with its non-zero blocks of the form $\left(\begin{smallmatrix} 0 & a_i \\ -a_i & 0 \end{smallmatrix}\right)$. If f is of the form $f(x_1, \ldots, x_N) = f_1(x_1) \ldots f_N(x_N)$ with $f_1, \ldots, f_n$ bounded and measurable, then (2.1) follows immediately from the $N = 2$ case by using independence. Then the case of general bounded measurable f follows from the multiplicative system theorem [19, Appendix A, p. 309], and for general f we can use a truncation argument.

Finally, an arbitrary skew-symmetric A can be written $A = U^* \tilde{A} U$, where U is orthogonal and $\tilde{A}$ is quasi-diagonal. (U can be taken to have rows given by the real

and imaginary parts of the eigenvectors of the Hermitian complex matrix iA.) We have shown that (2.1) holds for $\tilde{A}$; if we replace B_t by the Brownian motion UB_t and f by $f \circ U^*$, we have the desired result for A. □

3. Quadratic Brownian integrals in infinite dimensions

It is now fairly easy to generalize Theorem 2.1 to the infinite-dimensional setting using a finite-dimensional approximation argument. Before doing this we need to construct the infinite-dimensional stochastic processes involved. We give a self-contained construction that suffices for our purposes, but for a more general view of Hilbert space stochastic calculus, see [32]. The Itô integral relative to Brownian motion in an abstract Wiener space is discussed in Kuo [21, pp. 188-207, especially Theorem 5.1], [22, p. 5], and the appendix in [10].

Notation 3.1. Let $HS = HS(H)$ denote the Hilbert space of Hilbert–Schmidt operators on H, with the usual norm $\|A\|_{HS}^2 := \sum_{h\in\Lambda} \|Ah_i\|_H^2$, where Λ is any orthonormal basis for H. Let $HS_0 = HS_0(H)$ denote the subspace consisting of those operators which extend continuously to W, and whose range is finite-dimensional and contained in W_*.

Lemma 3.2. *A Hilbert–Schmidt operator A is in HS_0 iff there exists a finite rank orthogonal projection $P : H \to H$ such that $A = PAP$, $\mathrm{Ran}(P) \subset W_*$, and P extends continuously to W.*

Proof. If $A = PAP$ with P being a projection as in the statement of the lemma, then it is clear that $A \in HS_0$. Conversely, if $A \in HS_0$, let $\{u_i\}_{i=1}^n \subset W_*$ be an orthonormal basis for $\mathrm{Ran}(A)$. Then for all $h \in H$

$$Ah = \sum_{i=1}^n \langle Ah, u_i\rangle\, u_i = \sum_{i=1}^n \langle h, A^*u_i\rangle\, u_i = \sum_{i=1}^n \langle h, v_i\rangle\, u_i$$

where $v_i := A^*u_i \in W_*$ for all i. We now define P to be orthogonal projection onto

$$\mathrm{span}\{u_i, v_i : 1 \le i \le n\} = \mathrm{Ran}(A) + \mathrm{Ran}(A^*).$$

It is now a simple matter to check that $AP = A = PA$ on H and $A = PAP$ on W. □

Lemma 3.3. *If $A \in HS$ and $\{S_n\}_{n=1}^\infty$ and $\{T_n\}_{n=1}^\infty$ are bounded operators on H such that $S_n \xrightarrow{s} I$ and $T_n \xrightarrow{s} I$ (strong convergence), then $\|S_nAT_n^* - A\|_{HS} \to 0$ as $n \to \infty$. (Note that $T_n \xrightarrow{s} I$ does not necessarily imply $T_n^* \xrightarrow{s} I$.)*

We would like to thank Martín Argerami [2] for suggesting the following proof.

Proof. Let $s_n := S_n - I$ and $t_n := T_n - I$ so that $S_n = I + s_n$ and $T_n = I + t_n$ with $s_n, t_n \xrightarrow{s} 0$ as $n \to \infty$. By the uniform boundedness principle we know $C := \sup_n \left(\|S_n\|_{op} \vee \|T_n\|_{op}\right) < \infty$. Then

$$\begin{aligned}\|S_nAT_n^* - A\|_{HS} &= \|S_nA(I + t_n^*) - A\|_{HS}\\ &= \|s_nA + S_nAt_n^*\|_{HS}\\ &\le \|s_nA\|_{HS} + C\,\|At_n^*\|_{HS} = \|s_nA\|_{HS}\, C\,\|t_nA^*\|_{HS}.\end{aligned}$$

By the dominated convergence theorem we have, for any orthonormal basis Λ,

$$\lim_{n\to\infty} \|s_n A\|_{HS}^2 = \lim_{n\to\infty} \sum_{h\in\Lambda} \|s_n Ah\|_H^2 = \sum_{h\in\Lambda} \lim_{n\to\infty} \|s_n Ah\|_H^2 = \sum_{h\in\Lambda} 0 = 0$$

and similarly $\lim_{n\to\infty} \|t_n A^*\|_{HS}^2 = 0$. □

Corollary 3.4. *HS_0 is dense in HS.*

Proof. Since W_* is dense in H, we can choose an orthonormal basis $\{h_i\}_{i=1}^\infty$ for H with $\Lambda \subset W_*$. Set $P_n h = \sum_{i=1}^n \langle h, h_i\rangle_H h_i$ to be the orthogonal projection onto the span of $\{h_1, \dots, h_n\}$. Note that P_n is self-adjoint and $P_n \to I$ strongly. Then for any $A \in HS$, it is simple to verify that $P_n A P_n \in HS_0$, and taking $S_n = T_n = T_n^* = P_n$ in Lemma 3.3, we have $P_n A P_n \to A$ in HS-norm. □

Notation 3.5. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space on which there is defined a W-valued standard Brownian motion $\{B_t\}_{t\geq 0}$ as in Definition 1.9. Fix $T > 0$ and let $\mathcal{M}_T(H)$ denote the vector space of continuous, square-integrable, H-valued martingales up to time T defined on Ω. We equip $\mathcal{M}_T(H)$ with the Banach norm

$$\|M\|_{\mathcal{M}_T(H)}^2 := \mathbb{E}\left[\sup_{t\in[0,T]} \|M_t\|_H^2\right].$$

We likewise define $\mathcal{M}_T = \mathcal{M}_T(\mathbb{R})$ as the space of scalar-valued continuous square-integrable martingales with the analogous Banach norm.

Proposition 3.6. *The linear map*

$$HS_0 \ni A \mapsto \{AB_t\}_{0\leq t\leq T} \in \mathcal{M}_T(H)$$

is bounded. Hence it extends continuously to a map $HS \to \mathcal{M}_T(H)$, which we will still denote AB_t.

Proof. For $A \in HS_0$, there exists by Lemma 3.2 a finite rank projection P with $\mathrm{Ran}\,(P) \subset W_*$ such that $A = PAP$. Then $b_t := PB_t$ is a standard $\mathrm{Ran}\,(P)$-valued Brownian motion and therefore $M_t := AB_t = PAPB_t = PAb_t$ is a $\mathrm{Ran}\,(P)$-valued continuous martingale. Now $\{\|M_t\|_H\}_{t\geq 0}$ is a continuous submartingale and so Doob's maximal inequality gives

$$\|M\|_{\mathcal{M}_T(H)}^2 \leq 4\mathbb{E}\|M_T\|_H^2.$$

A simple computation shows that $\mathbb{E}\|M_T\|_H^2 = \mathbb{E}\|Ab_T\|_H^2 = T\|A\|_{HS}^2$, which completes the proof. □

Remark 3.7. One can replace the space $\mathcal{M}_T(H)$ of square-integrable martingales with spaces $\mathcal{M}_T^p(H)$ of L^p martingales, $1 \leq p < \infty$, with an analogous norm. The corresponding statements still hold, showing for instance that AB_t is an L^p martingale, upon replacing Doob's maximal inequality with the Burkholder–Davis–Gundy inequality.

For $A = PAP \in HS_0$ as in Lemma 3.2, we may interpret

$$\int_0^t \langle AB_s, dB_s\rangle_H = \int_0^t \langle PAPB_s, dB_s\rangle_H = \int_0^t \langle APB_s, dPB_s\rangle_H = \int_0^t \langle Ab_s, db_s\rangle_H$$

where $b_t := PB_t$ is a standard $\mathrm{Ran}\,(P)$-valued Brownian motion. Hence we are dealing solely with finite-dimensional stochastic calculus.

Proposition 3.8. *The linear map*

$$HS_0 \ni A \mapsto \int_0^{\cdot} \langle AB_s, dB_s\rangle_H \in \mathcal{M}_T$$

is bounded. Hence it extends continuously to a map $HS \to \mathcal{M}_T$ which we shall still denote by $\int_0^{\cdot}\langle AB_s, dB_s\rangle_H$.

In particular, for fixed T, we have a continuous linear map

$$HS \ni A \mapsto \int_0^T \langle AB_s, dB_s\rangle \in L^2(\mathbb{P}).$$

Proof. Let $A \in HS_0$ and set $M_t = \int_0^t \langle AB_s, dB_s\rangle_H = \int_0^t \langle Ab_s, db_s\rangle_H$ as above, where $b_t = PB_t$ is a standard $\mathrm{Ran}\,(P)$-valued Brownian motion. By Doob's maximal inequality and Itô's isometry, we have

$$\|M\|^2_{\mathcal{M}_T} \le 4\mathbb{E}|M_T|^2 = 4\int_0^T \mathbb{E}\|Ab_t\|_H^2\, dt = 2T\|A\|^2_{HS}$$

which completes the proof. □

Now a simple limiting argument shows that Theorem 2.1 still holds in this infinite-dimensional setting.

Theorem 3.9. *Let $A \in HS$ be skew-adjoint (i.e. $A^* = -A$) and $T > 0$. Then, for any bounded measurable $f : W \to \mathbb{R}$,*

$$\mathbb{E}\left[f(B_T)\, e^{i\int_0^T \langle AB_t, dB_t\rangle_H}\right] = \mathbb{E}\left[f(B_T)\, e^{-\frac{1}{2}\int_0^T \|AB_t\|_H^2 dt}\right]. \tag{3.1}$$

Proof. Suppose first that f is a bounded cylinder function, that is,

$$f(x) = \psi(\langle h_1, x\rangle_H, \ldots, \langle h_N, x\rangle_H)$$

for some $h_1, \ldots, H_N \in W_*$. Extend $\{h_1, \ldots, h_N\}$ to an orthonormal basis $\{h_n\} \subset W_*$ for H, and use this basis to define P_n as in Corollary 3.4. In particular, $P_nAP_n \in HS_0$ (and is also skew-adjoint), and $P_nAP_n \to A$ in HS norm. Now by Theorem 2.1, we have

$$\mathbb{E}\left[f(P_nB_T)\, e^{i\int_0^T \langle P_nAP_nB_t, dB_t\rangle_H}\right] = \mathbb{E}\left[f(P_nB_T)\, e^{-\frac{1}{2}\int_0^T \|P_nAP_nB_t\|_H^2 dt}\right].$$

Now we pass to the limit. For $n \ge N$ we have $f(P_nB_T) = f(B_T)$. Next, the continuity of the map in Proposition 3.8 shows that $\int_0^T \langle P_nAP_nB_t, dB_t\rangle_H \to \int_0^T \langle AB_t, dB_t\rangle_H$ in $L^2(\mathbb{P})$.

Finally, Proposition 3.6 tells us that $P_nAP_nB_t$ converges to AB_t in $\mathcal{M}_T(H)$; that is, as random elements of $C([0,T];H)$, they converge in $L^2(\mathbb{P})$. The map $\mathbf{x} \mapsto \int_0^T \|\mathbf{x}(t)\|_H^2\, dt$ is continuous on $C([0,T];H)$, so by continuous mapping we have $\int_0^T \|P_nAP_nB_t\|_H^2\, dt \to \int_0^T \|AB_t\|_H^2\, dt$ in probability.

Putting this all together and using continuous mapping again, we have

$$f(P_nB_T)\, e^{i\int_0^T \langle P_nAP_nB_t, dB_t\rangle_H} \to f(B_T)\, e^{i\int_0^T \langle AB_t, dB_t\rangle_H} \qquad \text{i.p.},$$

$$f(P_nB_T)\, e^{-\frac{1}{2}\int_0^T \|P_nAP_nB_t\|_H^2 dt} \to f(B_T)\, e^{-\frac{1}{2}\int_0^T \|AB_t\|_H^2 dt} \qquad \text{i.p.}$$

Everything in sight is bounded, so the dominated convergence theorem gives us the conclusion, still assuming that f is a cylinder function. An application of the multiplicative system theorem then covers the case that f is merely bounded and measurable. □

4. Heisenberg heat kernels

As in previous sections, (W, H, μ) is an abstract Wiener space, B_t is a Brownian motion on W, C is a finite-dimensional Hilbert with inner product $\cdot$, and $\omega : H \times H \to C$ is a skew-symmetric Hilbert–Schmidt bilinear form which is surjective.

Notation 4.1. For $\lambda \in C$, let $\Omega_\lambda \in HS$ be the Hilbert–Schmidt operator defined by $\langle \Omega_\lambda h, k\rangle = \omega(h,k)\cdot\lambda$ for all $h, k \in H$.

Definition 4.2. A **hypoelliptic Brownian motion** on G is the G-valued process $g_t = (B_t, Z_t)$, where $Z_t = \frac{1}{2}\int_0^t \omega(B_s, dB_s)$. To be precise, Z_t is defined by $Z_t \cdot \lambda = \frac{1}{2}\int_0^t \langle \Omega_\lambda B_t, dB_t\rangle_H$, where the stochastic integral is defined as in Proposition 3.8. For $T > 0$, let $\nu_T = \mathrm{Law}(g_T)$ denote the **hypoelliptic heat kernel measure** at time T on G.

We note the scaling relation

$$g_{ct} \overset{d}{=} (\sqrt{c}B_t, cZ_t) \quad \text{in law.} \tag{4.1}$$

Alternatively, following the development in Section 3, we could also define Z_t as the limit of the continuous C-valued processes $\frac{1}{2}\int_0^t \omega(P_n B_s, dP_n B_s)$, for P_n as in Corollary 3.4.

Notation 4.3. Let $\mathrm{End}(C)$ denote the space of linear transformations of the finite-dimensional Hilbert space C. We will consider $\mathrm{End}(C)$ as a finite-dimensional Banach space equipped with the operator norm, which we denote by $\|\cdot\|_{op}$. Also, let $\mathrm{End}_+(C) \subset \mathrm{End}(C)$ denote the closed cone of self-adjoint, non-negative definite transformations.

Notation 4.4. For $\mathbf{x}, \mathbf{y} \in C([0,T]; H)$, let $\rho_T(\mathbf{x}, \mathbf{y}) \in \mathrm{End}(C)$ be defined by

$$\rho_T(\mathbf{x}, \mathbf{y})\lambda \cdot \lambda' := \frac{1}{4}\int_0^T \langle \Omega_\lambda \mathbf{x}(s), \Omega_{\lambda'}\mathbf{y}(s)\rangle_H \, ds.$$

As usual, we will let $\rho_T(\mathbf{x}) = \rho_T(\mathbf{x}, \mathbf{x})$, and note that $\rho_T(\mathbf{x}) \in \mathrm{End}_+(C)$.

By Proposition 3.6, $\rho_T(\mathbf{x}, \mathbf{y})$ also makes sense (as a random linear transformation) if one or both of $\mathbf{x}, \mathbf{y}$ is replaced by a W-valued Brownian motion B. Henceforth ρ_T by itself will denote the random linear transformation $\rho_T(B)$, so that

$$\rho_T\lambda\cdot\lambda' = \rho_T(B)\lambda\cdot\lambda' = \frac{1}{4}\int_0^T \langle \Omega_\lambda B_s, \Omega_{\lambda'} B_s\rangle_H \, ds. \tag{4.2}$$

In this notation, (3.1) reads

$$\mathbb{E}\left[f(B_T)e^{i\lambda\cdot Z_T}\right] = \mathbb{E}\left[f(B_T)e^{-\frac{1}{2}\rho_T\lambda\cdot\lambda}\right]. \tag{4.3}$$

By making the change of variables $s = Ts'$ in (4.2), we get the scaling relation[1]

$$\rho_T \overset{d}{=} T^2\rho_1 \quad \text{in law.} \tag{4.4}$$

[1]Here we have made use of the fact that $B_{T(\cdot)} \overset{d}{=} \sqrt{T}B_{(\cdot)}$.

The following essential fact will be proved (in a stronger form) in the next section; see Corollary 5.9.

Proposition 4.5. *Almost surely, the random linear transformation ρ_T is strictly positive definite.*

In particular, ρ_T^{-1} exists almost surely and is also strictly positive definite. Given this, we can derive a formula for the heat kernel ν_T.

Theorem 4.6. *For any bounded measurable function $F : G \to \mathbb{R}$, we have*

$$\mathbb{E}[F(g_T)] = \mathbb{E}\int_C F(B_T, c) J_T^0(B, c)\, dc \tag{4.5}$$

where dc denotes Lebesgue measure on C, and

$$J_T^0(B, c) := \frac{\exp\left(-\frac{1}{2}\rho_T(B)^{-1}c \cdot c\right)}{\sqrt{\det(2\pi\rho_T(B))}}.$$

Moreover, $J_T^0(B, \cdot) > 0$, $\mathbb{P} \otimes m$ – a.e. where m denotes Lebesgue measure on C.

Proof. The fact that $J_T^0(B, \cdot) > 0$, $\mathbb{P} \otimes m$ – a.e. follows by Fubini's theorem along with Proposition 4.5. If F is of the form $F(x, c) = f(x)e^{i\lambda \cdot c}$, for some bounded measurable $f : W \to \mathbb{R}$ and some $\lambda \in C$, then by (4.3) and the Gaussian Fourier transform formula,

$$\begin{aligned}
\mathbb{E}\left[f(B_T)e^{i\lambda \cdot Z_T}\right] &= \mathbb{E}\left[f(B_T)e^{-\frac{1}{2}\rho_T \lambda \cdot \lambda}\right] \\
&= \mathbb{E}\left[f(B_T)\int_C e^{i\lambda \cdot c}\frac{\exp\left(-\frac{1}{2}\rho_T^{-1}c \cdot c\right)}{\sqrt{\det(2\pi\rho_T)}}\right] \\
&= \mathbb{E}\int_C f(B_T)e^{i\lambda \cdot c} J_T^0(B, c)\, dc \qquad (4.6)
\end{aligned}$$

as desired. Taking $f = 1$, $\lambda = 0$ in (4.6) shows that $\mathbb{E}\int_C J_T^0(B, c)\, dc = 1$.

The proof is now easily completed with the help of the multiplicative system theorem. Indeed, the set of all such functions $F(x, c) = f(x)e^{i\lambda \cdot c}$ is a multiplicative system, and it is standard to show that it generates the Borel σ-algebra of G. The set of functions F for which (4.5) holds is a vector space. Therefore if F_n is a sequence of functions satisfying (4.5) and $F_n \to F$ boundedly, the dominated convergence theorem shows that F also satisfies (4.5). Having verified the hypotheses of the multiplicative system theorem, we conclude that (4.5) holds for all bounded measurable F. □

Corollary 4.7. *The heat kernel measure ν_T is absolutely continuous to the product measure $d\mu_T \otimes dc$, and the Radon–Nikodym derivative is given by*

$$\gamma_T(x, c) := \mathbb{E}\left[J_T^0(B, c) \mid B_T = x\right].$$

(That is, for each c, $\gamma_T(\cdot, c)$ is a measurable function on W such that $\gamma_T(B_T, c) = \mathbb{E}\left[J_T^0(B, c) \mid B_T\right]$ almost surely; this function is unique up to μ_T-null sets.)

A few properties are immediately apparent from this formula.

Corollary 4.8. *We have $\gamma_T > 0$, $d\mu_T \otimes dc$-almost everywhere on G.*

Proof. For each $c \in C$, $J_T^0(B, c) > 0$ $\mathbb{P}$-almost surely, hence its conditional expectation $\gamma_T(B_T, c)$ is also strictly positive $\mathbb{P}$-almost surely. In other words, for each $c \in C$, we have $\gamma_T(x, c) > 0$ for μ_T-almost every $x \in W$. The conclusion follows by Fubini's theorem. □

Corollary 4.9. *For any $T > 0$, the heat kernel measure ν_T is invariant under the inversion map $g \mapsto g^{-1}$; that is,*

$$\mathbb{E}[F(g_T)] = \int_G F(g)\, d\nu_T(g) = \int_G F(g^{-1})\, d\nu_T(g) = \mathbb{E}[F(g_T^{-1})].$$

In other words, g_T and g_T^{-1} have the same law.

Proof. This follows immediately from the observation that $J_T^0(-B, -c) = J_T^0(B, c)$ together with the symmetry of the law of Brownian motion. □

This fact can also be extracted from finite-dimensional approximations; see [12, Corollary 4.9]. It is worth noting that, in contrast to flat Brownian motion, the *processes* $\{g_t\}_{t\geq 0}$ and $\{g_t^{-1}\}_{t\geq 0}$ generally do *not* have the same law.

5. Estimates on ρ_T

In this section, we derive technical estimates on the random linear transformation ρ_T, which were used in the previous section to define the heat kernel and will be needed in the sequel for further development of the smoothness properties of the heat kernel. In particular, we need to show that ρ_T is almost surely invertible, and that its inverse is unlikely to be large. Throughout this section, $T > 0$ is fixed, and, for any $A : H \to H$, $\|A\|_{op}$ denotes the standard operator norm of A on H.

5.1. **Small ball estimates.** We will need the following "small ball" result, which essentially says that a Brownian motion is unlikely to stay close to the origin. See [26, Lemma 2.3] for a proof (of a more general statement) as well as historical notes.

Theorem 5.1. *Let b_t be a one-dimensional Brownian motion. Then*

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{P}\left(\int_0^1 b_t^2\, dt < \varepsilon \right) = -\frac{1}{8}.$$

In particular, there is a positive constant K_0 such that for all $\varepsilon > 0$,

$$\mathbb{P}\left(\int_0^1 b_t^2\, dt \leq \varepsilon \right) \leq K_0 e^{-1/(4\varepsilon)}.$$

Lemma 5.2. *Let $A \in HS$. Then for all $\varepsilon > 0$*

$$\mathbb{P}\left(\int_0^T \|AB_t\|_H^2\, dt < \varepsilon \right) \leq K_0 \exp\left(-\frac{\|A\|_{op}^2 T^2}{4\varepsilon} \right)$$

where K_0 is the constant from Theorem 5.1.

Proof. By rescaling, it is sufficient to consider the case $T = 1$.

By replacing A by $\sqrt{A^*A}$, we can assume that A is self-adjoint and non-negative definite. In particular, $\lambda := \|A\|_{op}$ is an eigenvalue of A; let u be a corresponding unit eigenvector. Then $\|AB_t\|_H^2 \geq |\langle AB_t, u\rangle|^2 = \lambda^2 |\langle B_t, u\rangle|^2$. Here we interpret $\langle B_t, u\rangle$ in the sense of Proposition 3.6, viewing $\langle \cdot, u\rangle$ as a rank-one Hilbert–Schmidt

operator on H. Then it is easy to verify that $\langle B_t, u\rangle$ is a one-dimensional standard Brownian motion. As such,

$$\mathbb{P}\left(\int_0^1 \|AB_t\|_H^2\,dt < \varepsilon\right) \le \mathbb{P}\left(\int_0^1 b_t^2\,dt < \frac{\varepsilon}{\lambda^2}\right) \le K_0 \exp\left(-\frac{\lambda^2}{4\varepsilon}\right).$$

As $\lambda = \|A\|_{op}$ the proof is complete. □

We need an analogous statement when the Brownian motion is perturbed by a one-dimensional drift, under the assumption that A is skew-adjoint. This will follow from the following lemma from linear algebra.

Lemma 5.3. *Suppose $A \in HS$ is skew-adjoint, and P is orthogonal projection onto a subspace of H with codimension 1. Then $\|PA\|_{op} = \|AP\|_{op} = \|A\|_{op}$.*

Proof. The first equality is just the fact that $PA = -(AP)^*$. Moreover, the inequality $\|AP\|_{op} \le \|A\|_{op}$ is obvious.

Since $\sqrt{A^*A}$ is compact, self-adjoint and non-negative definite, it has $\lambda := \|\sqrt{A^*A}\|_{op} = \|A\|_{op}$ as an eigenvalue. Let u be a unit eigenvector of $\sqrt{A^*A}$ with eigenvalue λ. Then u is also an eigenvector of $A^*A = -A^2$ with eigenvalue λ^2. Since $Au \neq 0$, set $v = Au/\|Au\|$; then we have $\langle u, v\rangle = 0$ and $-A^2 v = \lambda^2 v$.

Let h_0 be a unit vector in the kernel of P, so that $Ph = h - \langle h, h_0\rangle h_0$. Now choose a unit vector $w \in \operatorname{span}\{u, v\}$ with $\langle w, h_0\rangle = 0$. (If $\langle u, h_0\rangle = 0$, take $w = u$; else take $w = \langle u, h_0\rangle v - \langle v, h_0\rangle u$, appropriately rescaled.) Then $Pw = w$, so $\|APw\|^2 = \|Aw\|^2 = \langle A^*Aw, w\rangle = \lambda^2$. We thus have shown $\|AP\|_{op} \ge \lambda = \|A\|_{op}$. □

Lemma 5.4. *Suppose $A \in HS$ is skew-adjoint. For any fixed $h_0 \in H$,*

$$\mathbb{P}\left(\inf_{\gamma \in C([0,T];H)} \int_0^T \|A(B_t + \gamma(t)h_0)\|_H^2\,dt < \varepsilon\right) \le K_0 \exp\left(-\frac{\|A\|_{op}^2 T^2}{4\varepsilon}\right)$$

where K_0 is the constant from Theorem 5.1.

Proof. Let $Ph = h - \|Ah_0\|_H^{-1}\langle h, Ah_0\rangle h_0$ be orthogonal projection onto the orthogonal complement of $\{Ah_0\}$. (If $Ah_0 = 0$, then take $P = I$.) For any γ, we have

$$\|A(B_t + \gamma(t)h_0)\|_H^2 \ge \|PA(B_t + \gamma(t)h_0)\|_H^2 = \|PAB_t\|_H^2.$$

Thus by Lemma 5.2 and Lemma 5.3, we have

$$\begin{aligned}\mathbb{P}\left(\inf_{\gamma \in C([0,T];H)} \int_0^T \|A(B_t + \gamma(t)h_0)\|_H^2\,dt < \varepsilon\right) &\le \mathbb{P}\left(\int_0^T \|PAB_t\|_H^2\,dt < \varepsilon\right) \\ &\le K_0 \exp\left(-\frac{\|PA\|_{op}^2 T^2}{4\varepsilon}\right) \\ &= K_0 \exp\left(-\frac{\|A\|_{op}^2 T^2}{4\varepsilon}\right).\end{aligned}$$

□

5.2. **Large deviations for $\|\rho_T\|_{op}$.** We will need the following large deviations result for Wiener chaos random variables, which can be found in [23, p. 6] together with the relevant definitions, background, and further references.

Theorem 5.5. *Let $(\tilde{W}, \tilde{H}, \tilde{\mu})$ be an abstract Wiener space, let B be a real separable Banach space, and let $f \in L^2(\tilde{\mu}; B)$ be a random variable that is in $\mathcal{H}^{(d)}(\tilde{\mu}; B)$, the B-valued homogeneous Wiener chaos of degree d. Then*

$$\lim_{r\to\infty} \frac{1}{r^{2/d}} \log \tilde{\mu}\left(x : \|f(x)\|_B > r\right) = -\frac{1}{2}\left(\sup_{h\in\tilde{H}, \|h\|_{\tilde{H}}\le 1} \left\| \int_{\tilde{W}} f(x+h)\, \tilde{\mu}(dx) \right\|_B \right)^{-2/d}.$$

In particular, as in [23, p. 3], by the Cameron–Martin theorem and the Cauchy–Schwarz inequality we have the bound

$$\limsup_{r\to\infty} \frac{1}{r^{2/d}} \log \tilde{\mu}\left(x : \|f(x)\|_B > r\right) \le -\frac{e^{-1/d}}{2}\left(\int_{\tilde{W}} \|f(x)\|_B^2\, \tilde{\mu}(dx)\right)^{-2/d}. \tag{5.1}$$

This bound can be applied to ρ_T, as the following lemma shows.

Lemma 5.6. *There is a constant $k > 0$, depending only on ω, such that*

$$\limsup_{r\to\infty} \frac{1}{r} \log \mathbb{P}(\|\rho_T\|_{op} > r) \le -\frac{k}{T^2}$$

where $\|\cdot\|_{op}$ is the operator norm on $\mathrm{End}(C)$ as in Notation 4.3. In particular, for any $k_1 < k$ there is a $K_1 > 0$ (depending on k_1 and ω) such that

$$\mathbb{P}(\|\rho_T\|_{op} > r) \le K_1 \exp\left(-\frac{k_1 r}{T^2}\right).$$

Proof. By (4.4), we have $\rho_T \overset{d}{=} T^2\rho_1$ in law, so it suffices to consider $T = 1$. We will write ρ for ρ_1.

Take $\tilde{W} = C([0,1]; W)$ to be the path space over W, and $\tilde{\mu}$ to be the law of Brownian motion $\{B_t\}$ on W, which is a Gaussian measure on $\tilde{W}$. Then $(\tilde{W}, \tilde{H}, \tilde{\mu})$ is an abstract Wiener space, where $\tilde{H}$ is the space of finite-energy H-valued paths in $\tilde{W}$. If we take $E = \mathrm{End}(C)$ with the operator norm $\|\cdot\|_{op}$, then we can consider ρ as an E-valued random variable on $\tilde{W}$. It is not hard to show that $\rho - \mathbb{E}\rho \in \mathcal{H}^{(2)}(\tilde{W}; E)$. So applying (5.1) with $d = 2$ and adjusting notation, we obtain

$$\limsup_{r\to\infty} \frac{1}{r} \log \mathbb{P}\left(\|\rho - \mathbb{E}\rho\|_{op} > r\right) \le -\frac{e^{-1/2}}{2}\left(\mathbb{E}\,\|\rho - \mathbb{E}\rho\|_{op}^2\right)^{-1}.$$

We can drop the constant $\mathbb{E}\rho$ from the left side without changing the limit. Setting

$$k := \frac{e^{-1/2}}{2}\left(\mathbb{E}\,\|\rho - \mathbb{E}\rho\|_{op}^2\right)^{-1}$$

which only depends on ω, we have the conclusion. □

5.3. **Estimates on ρ_T.**

Notation 5.7. Given $h \in H$, let $\mathbf{h}(t) = \frac{t}{T}h$, so that $\mathbf{h}$ is a finite-energy path in $H \subset W$ with $\mathbf{h}(0) = 0$ and $\mathbf{h}(T) = h$.

Lemma 5.8. *Fix $T > 0$, $\alpha_0 \geq 0$, and $h \in H$. There are constants K, k, depending on ω, α_0, $\|h\|_H$, and T, so that*

$$\mathbb{P}\left(\sup_{|\alpha| \leq \alpha_0} \|\rho_T(B + \alpha \mathbf{h})^{-1}\|_{op} > r \right) \leq K e^{-kr}.$$

Proof. Let S denote the unit sphere of C. For an arbitrary $\delta > 0$, we may cover S with a finite number n of balls of radius at most δ; let $\{\lambda_i\}_{i=1}^n$ be their centers. We can choose $n \leq M\delta^{-d}$, where $d = \dim C$ and M is some universal constant.[2]

For any $Q \in \mathrm{End}_+(C)$ and any $\lambda \in C$, we can choose a λ_i with $|\lambda - \lambda_i| < \delta$. Then the mean value theorem gives us $|Q\lambda \cdot \lambda - Q\lambda_i \cdot \lambda_i| \leq 2\|Q\|_{op}\delta$. Thus

$$\|Q^{-1}\|_{op}^{-1} = \min_{\lambda \in S} Q\lambda \cdot \lambda \geq \min_i Q\lambda_i \cdot \lambda_i - 2\delta \|Q\|_{op}.$$

Thus, for any $r > 0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{|\alpha| \leq \alpha_0} \|\rho_T(B + \alpha \mathbf{h})^{-1}\|_{op} > r \right) = \mathbb{P}\left(\inf_{|\alpha| \leq \alpha_0} \inf_{\lambda \in S} \rho_T(B + \alpha \mathbf{h})\lambda \cdot \lambda < r^{-1} \right) \\
&\leq \mathbb{P}\left(\inf_{|\alpha| \leq \alpha_0} \min_i \rho_T(B + \alpha \mathbf{h})\lambda_i \cdot \lambda_i - 2\delta \sup_{|\alpha| \leq \alpha_0} \|\rho_T(B + \alpha \mathbf{h})\|_{op} < r^{-1} \right) \\
&\leq \mathbb{P}\left(\inf_{|\alpha| \leq \alpha_0} \min_i \rho_T(B + \alpha \mathbf{h})\lambda_i \cdot \lambda_i < 2r^{-1} \right) \\
&\quad + \mathbb{P}\left(2\delta \sup_{|\alpha| \leq \alpha_0} \|\rho_T(B + \alpha \mathbf{h})\|_{op} > r^{-1} \right) \\
&=: P_1 + P_2.
\end{aligned}$$

To analyze the second term P_2, let us choose $\delta = 1/r^2$, so we have

$$P_2 = P_2(r) = \mathbb{P}\left(2 \sup_{|\alpha| \leq \alpha_0} \|\rho_T(B + \alpha \mathbf{h})\|_{op} > r \right).$$

We note that for $|\alpha| \leq \alpha_0$, we have

$$\|\rho_T(B + \alpha \mathbf{h})\|_{op} \leq \|\rho_T(B)\|_{op} + 2\alpha_0 \|\rho_T(B, \mathbf{h})\|_{op} + \alpha_0^2 \|\rho_T(\mathbf{h})\|_{op},$$

so for any $\varepsilon > 0$,

$$\begin{aligned}
P_2(r) &\leq \mathbb{P}(2\|\rho_T(B)\|_{op} > r/3) + \mathbb{P}(4\alpha_0 \|\rho_T(B, \mathbf{h})\|_{op} > r/3) \\
&\quad + \mathbb{P}(2\alpha_0^2 \|\rho_T(\mathbf{h})\|_{op} > r/3) \\
&=: P_{2,1}(r) + P_{2,2}(r) + P_{2,3}(r).
\end{aligned}$$

Now $P_{2,1}(r)$ is controlled by Lemma 5.6. For $P_{2,2}$, we note that $\rho_T(B, \mathbf{h})$ is linear in B and hence Gaussian. Thus by Fernique's theorem [16] (see also [21, Theorem 3.1]), we have $P_{2,2}(r) \leq K' e^{-k' r^2}$ for some K', k'. And since $\rho_T(\mathbf{h})$ is deterministic, $P_{2,3}(r)$ vanishes for all sufficiently large r. Thus we have $P_2(r) \leq K e^{-kr}$ for suitable K, k.

[2]This crude bound is easily proved when the ℓ^2 norm on C is replaced by an ℓ^∞ norm. In this case balls are cubes and it is just a question of dividing the unit cube into order $(1/\delta)^{\dim C}$ sub-cubes. Since the ℓ^2 and ℓ^∞ norms are equivalent in finite dimensions it follows that the same "entropy" estimates hold for round balls.

Next we estimate $P_1(r) = \mathbb{P}\left(\inf_{|\alpha|\leq\alpha_0} \min_i \rho_T(B+\alpha\mathbf{h})\lambda_i \cdot \lambda_i < 2r^{-1}\right)$. Here i ranges from 1 to $n = n(r) \leq Mr^{2d}$, since we have chosen $\delta = r^{-2}$. Thus by a union bound we have

$$\begin{aligned} P_1(r) &\leq \sum_{i=1}^{n(r)} \mathbb{P}\left(\inf_{|\alpha|\leq\alpha_0} \rho_T(B+\alpha\mathbf{h})\lambda_i \cdot \lambda_i < 2r^{-1}\right) \\ &\leq n(r) \max_i \mathbb{P}\left(\inf_{|\alpha|\leq\alpha_0} \rho_T(B+\alpha\mathbf{h})\lambda_i \cdot \lambda_i < 2r^{-1}\right) \\ &\leq Mr^{2d} \max_i K_0 \exp\left(-\frac{r\|\Omega_{\lambda_i}\|_{op}}{2}\right) \end{aligned}$$

applying Lemma 5.4 with $A = \Omega_{\lambda_i}$, $h_0 = h$, $\gamma(t) = \alpha\frac{t}{T}$, and $\varepsilon = 2r^{-1}$.

Now we note that since ω is assumed to be surjective, we have $\Omega_\lambda \neq 0$ for every $\lambda \neq 0$. Then since S is compact and the map $C \ni \lambda \to \Omega_\lambda \in B(H)$ is continuous, we have $\inf_{\lambda\in S} \|\Omega_\lambda\|_{op} > 0$. Thus we have

$$P_1(r) \leq Ke^{-kr}$$

for some (new) constants K, k, which have been adjusted so as to absorb the polynomial factor Mr^{2d}.

Combining the estimates on $P_1(r), P_2(r)$ gives the result. □

The previous results and proofs give the following immediate corollary.

Corollary 5.9. *We have*

(1) $\sup_{|\alpha|\leq\alpha_0} \|\rho_T(B+\alpha\mathbf{h})^{-1}\|_{op} \in L^{\infty-}(\mathbb{P})$. *(Here $L^{\infty-} := \bigcap_{1\leq p<\infty} L^p$.)*
(2) $(\det \rho_T)^{-1} = \det(\rho_T^{-1}) \leq \|\rho_T^{-1}\|_{op}^d \in L^{\infty-}$.
(3) *Almost surely, ρ_T is invertible and hence strictly positive definite.*

This was the missing piece in the proof of Theorem 4.6 and its corollaries.

Remark 5.10. There is a simpler argument to see ρ_T is invertible almost surely when ω is continuous on W. Increments of a W-valued Brownian motion contain, after scaling, an i.i.d. sequence distributed according to the Gaussian measure μ, which by assumption has full support. Hence the image of the Brownian motion over times $[0, T]$ is total, almost surely, so it cannot live in any proper closed subspace of W, such as the kernel of any non-zero continuous operator. We would like to thank George Lowther [27] and Clinton Conley for suggesting this argument.

6. A Fernique-type theorem

As an application of the formula in Theorem 4.6, we give a proof of a Fernique-type theorem (compare [16]) giving square-exponential integrability for the hypoelliptic heat kernel measure. This result was previously obtained in [12, Theorem 4.16] via finite-dimensional projections. (Note that [12, Theorem 4.16] actually handles an elliptic heat kernel measure, but the same proof works in the hypoelliptic case by simply omitting the B_0 term.)

Proposition 6.1 (Fernique-type theorem)**.** *There exists $\varepsilon > 0$ sufficiently small that, for any $T > 0$,*

$$\int_G e^{\frac{\varepsilon}{T}(\|x\|_W^2 + |c|_C)}\, \nu_T(dx, dc) = \mathbb{E}\left[e^{\frac{\varepsilon}{T}(\|B_T\|_W^2 + |Z_T|_C)}\right] < \infty.$$

Proof. By the scaling relation (4.1), we see that it is sufficient to show the result when $T = 1$; then the same ε will work for every other $T > 0$. As before, we write ρ for ρ_1.

From Theorem 4.6, we have

$$\begin{aligned}\mathbb{E}\left[e^{\varepsilon(\|B_1\|_W^2+|Z_1|_C)}\right] &= \mathbb{E}\left[e^{\varepsilon\|B_1\|_W^2}\int_C \frac{\exp\left(\varepsilon|c| - \frac{1}{2}\rho^{-1}c\cdot c\right)}{\sqrt{\det(2\pi\rho)}}\,dc\right] \\ &\le \mathbb{E}\left[e^{\varepsilon\|B_1\|_W^2}\int_C \frac{\exp\left(\varepsilon\sqrt{\|\rho\|_{op}}|c| - \frac{1}{2}|c|^2\right)}{\sqrt{(2\pi)^d}}\,dc\right]\end{aligned}$$

where we made the change of variables $c \to \rho^{1/2}c$ and used the inequality $|\rho^{1/2}c| \le \sqrt{\|\rho\|_{op}}|c|$. Now letting $a = \varepsilon\sqrt{\|\rho\|_{op}}$ and writing the dc integral in polar coordinates gives us the bound

$$\begin{aligned}&\int_C \exp\left(a|c| - \frac{1}{2}|c|^2\right)\,dc \\ &\quad= \omega_{d-1}\int_0^\infty e^{ar-\frac{1}{2}r^2}r^{d-1}\,dr \\ &\quad\le \omega_{d-1}\left(\int_0^{4a} e^{ar-\frac{1}{2}r^2}r^{d-1}\,dr + \int_{4a}^\infty e^{-\frac{1}{4}r^2}r^{d-1}\,dr\right) \\ &\quad\le \omega_{d-1}\left(\left(\sup_{r\ge 0} e^{-\frac{1}{2}r^2}r^{d-1}\right)\int_0^{4a} e^{ar}\,dr + \int_0^\infty e^{-\frac{1}{4}r^2}r^{d-1}\,dr\right) \\ &\quad\le Ke^{4a^2}\end{aligned}$$

for a suitable constant K not depending on a. (Here ω_{d-1} is the volume of the $(d-1)$-dimensional unit sphere of C.) Thus

$$\begin{aligned}\mathbb{E}\left[e^{\varepsilon(\|B_1\|_W^2+|Z_1|_C)}\right] &\le K\mathbb{E}\left[e^{\varepsilon\|B_1\|_W^2}e^{4\varepsilon^2\|\rho\|_{op}}\right] \\ &\le K\left(\mathbb{E}e^{2\varepsilon\|B_1\|_W^2}\right)^{1/2}\left(\mathbb{E}e^{8\varepsilon^2\|\rho\|_{op}}\right)^{1/2}\end{aligned}$$

by the Cauchy–Schwarz inequality. Fernique's theorem asserts that the first factor is finite for small enough ε. For the second factor, Lemma 5.8 shows that it is finite as soon as $\varepsilon < k$. □

7. Quasi-invariance of hypoelliptic heat kernel measures

The strict positivity of γ combined with the standard Cameron–Martin theorem implies quasi-invariance for ν_T under translations by Cameron–Martin subgroup elements.

Proposition 7.1 (Quasi-invariance under right translations I)**.** *For any $T > 0$, the heat kernel measure ν_T is quasi-invariant under right translation by elements of the Cameron–Martin subgroup G_{CM}. In particular, for any bounded measurable $F : G \to \mathbb{R}$,*

$$\mathbb{E}[F(g_T \cdot g)] = \mathbb{E}[F(g_T)\mathcal{J}_g(g_T)]$$

where

$$\mathcal{J}_g(x,c) = \bar{J}_h(x)\gamma_T\left(x - h, c - z - \frac{1}{2}\omega(x,h)\right)\gamma_T(x,c)^{-1}$$

with

$$\bar{J}_h(x) := \exp\left(\frac{1}{T}\langle h, x\rangle_H - \frac{1}{2T}\|h\|_H^2\right).$$

Proof. Let $g = (h, z) \in G_{CM}$. Then by the translation invariance of Lebesgue measure and the standard Cameron–Martin theorem for (W, H, μ_T), we have that

$$\begin{aligned}
&\int_G F((x,c)\cdot(h,z))\,d\nu_T(x,c) \\
&\qquad = \int_C \int_W F\left(x + h, c + z + \frac{1}{2}\omega(x,h)\right)\gamma_T(x,c)\,\mu_T(dx)\,dc \\
&\qquad = \int_C \int_W F\left(x, c + z + \frac{1}{2}\omega(x-h,h)\right)\gamma_T(x-h,c)\bar{J}_h(x)\,\mu_T(dx)\,dc \\
&\qquad = \int_C \int_W F(x,c)\gamma_T\left(x-h, c-z-\frac{1}{2}\omega(x,h)\right)\bar{J}_h(x)\,\mu_T(dx)\,dc.
\end{aligned}$$

Since $\gamma_T > 0$ $\mu_T \otimes dm_C$ a.e., this shows that

$$\begin{aligned}
&\int_G F((x,c)\cdot(h,z))\,d\nu_T(x,c) \\
&\qquad = \int_G F(x,c)\gamma_T\left(x-h, c-z-\frac{1}{2}\omega(x,h)\right)\bar{J}_h(x)\gamma_T(x,c)^{-1}\,d\nu_T(x,c).
\end{aligned}$$

□

Alternatively, given the expression of the density γ from Corollary 4.7, we may reformulate this quasi-invariance result as follows, which will be more useful in proving subsequent integration by parts formulae.

Proposition 7.2 (Quasi-invariance under right translations II)**.** *Let $g = (h, z) \in G_{CM}$. Then, for any measurable $F : G \to [0, \infty]$ or measurable $F : G \to \mathbb{R}$ satisfying $\mathbb{E}|F(g_T \cdot g)| < \infty$, we have that*

$$\mathbb{E}[F(g_T \cdot g)] = \int_C dc\,\mathbb{E}\left[F(B_T, c)J_g(B, c)\bar{J}_h(B_T)\right]$$

where, for $\mathbf{h}(t) = \frac{t}{T}h$ as in Notation 5.7,

$$J_g(B, c) := J_T^0\left(B - \mathbf{h}, c - z - \frac{1}{2}\omega(B_T, h)\right),$$

and $\bar{J}_h$ is as given in Proposition 7.1.

Proof. By Theorem 4.6, Fubini's theorem, and the translation invariance of Lebesgue measure, we have that

$$\begin{aligned}
\mathbb{E}[F(g_T \cdot g)] &= \mathbb{E}\left[F\left(B_T + h, Z_T + z + \frac{1}{2}\omega(B_T, h)\right)\right] \\
&= \int_C dc\,\mathbb{E}\left[F\left(B_T + h, c + z + \frac{1}{2}\omega(B_T, h)\right)\frac{\exp\left(-\frac{1}{2}\rho_T^{-1}c\cdot c\right)}{\sqrt{\det(2\pi\rho_T)}}\right] \\
&= \int_C dc\,\mathbb{E}\left[F(B_T+h, c)\frac{\exp\left(-\frac{1}{2}\rho_T^{-1}(c - z - \frac{1}{2}\omega(B_T,h))\cdot(c - z - \frac{1}{2}\omega(B_T,h))\right)}{\sqrt{\det(2\pi\rho_T)}}\right].
\end{aligned}$$

Now taking the given finite-energy path $\mathbf{h}$ and translating $B \mapsto B - \mathbf{h}$, the standard Cameron–Martin theorem on $C([0,T];W)$ for $\mathrm{Law}(B)$ states that

$$\mathbb{E}\left[F(B_T+h,c)\frac{\exp\left(-\frac{1}{2}\rho_T^{-1}(c-z-\frac{1}{2}\omega(B_T,h))\cdot(c-z-\frac{1}{2}\omega(B_T,h))\right)}{\sqrt{\det(2\pi\rho_T)}}\right]$$
$$= \mathbb{E}\left[F(B_T,c)\frac{\exp\left(-\frac{1}{2}\rho_T(B-\mathbf{h})^{-1}(c-z-\frac{1}{2}\omega(B_T,h))\cdot(c-z-\frac{1}{2}\omega(B_T,h))\right)}{\sqrt{\det(2\pi\rho_T(B-\mathbf{h}))}}\bar{\mathbf{J}}_{\mathbf{h}}\right]$$

where

$$\bar{\mathbf{J}}_{\mathbf{h}} = \bar{\mathbf{J}}_{\mathbf{h}}(B) = \exp\left(\int_0^T \langle \dot{\mathbf{h}}(t), dB_t\rangle_H - \frac{1}{2}\int_0^T \|\dot{\mathbf{h}}(t)\|_H^2\, dt\right).$$

Noting that for the given path $\mathbf{h}$, $\bar{\mathbf{J}}_{\mathbf{h}}$ simplifies to $\bar{J}_h$, completes the proof. □

Now, consistent with the notation J_g and $\bar{J}_h$ defined in Proposition 7.2, we set the following notation for the sequel.

Notation 7.3. For any $F = F(B,c)$ and $g = (h,z) \in G_{CM}$, we will write

$$F_g(B,c) := F\left(B-\mathbf{h}, c-z-\frac{1}{2}\omega(B_T,h)\right)$$

where $\mathbf{h}(t) = \frac{t}{T}h$ as in Notation 5.7. Also, without further comment, we will make the standard identification between $\mathfrak{g}_{CM}$ and G_{CM}, and for $X \in \mathfrak{g}_{CM}$ we will write F_X to mean the analogous expression to that given above for F_g. Furthermore, we will define

$$(\tilde{\mathbf{X}}F)(B,c) := \frac{d}{d\varepsilon}\bigg|_0 F_{\varepsilon X}(B,c).$$

Note that when $F = F(B_T,c)$, then $F_X(B_T,c) = F((B_T,c)\cdot -X)$ and $(\tilde{\mathbf{X}}F)(B_T,c) = -(\tilde{X}F)(B_T,c)$.

Although there is a natural group structure on the path space over G, note that the vector field $\tilde{\mathbf{X}}$ is not invariant with respect to this structure.

With this notation in place, we record the following statement which may be observed by following the proof of Proposition 7.2.

Corollary 7.4. *Let $g = (h,z) \in G_{CM}$. Then, for any measurable $F : G \to [0,\infty]$ and $\Psi : C([0,T];W)\times C \to [0,\infty]$ we have that*

$$\int_C dc\,\mathbb{E}[F((B_T,c)\cdot g)\Psi(B,c)J_T^0(B,c)]$$
$$= \int_C dc\,\mathbb{E}\left[F(B_T,c)\Psi_g(B,c)J_g(B,c)\bar{J}_h(B_T)\right]$$

for $\mathbf{h}(t) = \frac{t}{T}h$, J_g, and $\bar{J}_h$ as given in Proposition 7.2. Moreover, this equality holds for any measurable $F : G \to \mathbb{R}$ and $\Psi : C([0,T];W)\times C \to \mathbb{R}$ satisfying

$$\int_C dc\,\mathbb{E}[|F((B_T,c)\cdot g)\Psi(B,c)|J_T^0(B,c)] < \infty.$$

One can directly prove quasi-invariance under left translations in a similar manner to Propositions 7.1 or 7.2, but it also follows from quasi-invariance under right translations combined with the invariance of ν_T under inversions.

Corollary 7.5 (Quasi-invariance under left translations)**.** *Let $g \in G_{CM}$. Then, for any measurable $F\colon G \to [0,\infty]$ or measurable $F\colon G \to \mathbb{R}$ satisfying $\mathbb{E}|F(g_T \cdot g)| < \infty$,*

$$\mathbb{E}[F(g \cdot g_T)] = \mathbb{E}[F(g_T)\tilde{\mathcal{J}}_g(g_T)]$$

where

$$\tilde{\mathcal{J}}_g(g') = \mathcal{J}_{g^{-1}}(g'^{-1})$$

for all $g' \in G$ and $\mathcal{J}_g$ as given in Proposition 7.1.

Proof. Let $u(g) := F(g^{-1})$. Then repeatedly applying Corollary 4.9 (the invariance of ν_T under inversion) gives

$$\begin{aligned}\mathbb{E}[F(g \cdot g_T)] &= \mathbb{E}[F(g \cdot g_T^{-1})] = \mathbb{E}[F((g_T \cdot g^{-1})^{-1})] \\ &= \mathbb{E}[u(g_T \cdot g^{-1})] = \mathbb{E}[u(g_T)\mathcal{J}_{g^{-1}}(g_T)] \\ &= \mathbb{E}[u(g_T^{-1})\mathcal{J}_{g^{-1}}(g_T^{-1})] = \mathbb{E}[F(g_T)\mathcal{J}_{g^{-1}}(g_T^{-1})].\end{aligned}$$

□

8. Smoothness properties of hypoelliptic heat kernel measures

In this section, we expand on the quasi-invariance results in the previous section, and study the integrability and smoothness of the corresponding Radon–Nikodym derivatives. Similar techniques were used to handle the elliptic case in [9]. Again for this section, we fix $T > 0$.

First we recall some known results for the standard Cameron–Martin Radon–Nikodym derivative.

Proposition 8.1. *For $h \in H$, we have $\bar{J}_h(B_T) \in L^{\infty-}$ and*

$$\sup_{|\varepsilon|\le 1}\left|\frac{d}{d\varepsilon}\bar{J}_{\varepsilon h}(B_T)\right| = \sup_{|\varepsilon|\le 1}\left\{\bar{J}_{\varepsilon h}(B_T)\cdot\left|\frac{1}{T}\langle h, B_T\rangle - \frac{\varepsilon}{T}\|h\|_H^2\right|\right\} \in L^{\infty-}.$$

These estimates are easily proved, since for example, for all $p \in [1,\infty)$,

$$\mathbb{E}[\bar{J}_h(B_T)^p] = \mathbb{E}\left[\exp\left(\frac{p}{T}\langle h, B_T\rangle - \frac{p}{2T}\|h\|_H^2\right)\right] = \exp\left(\frac{1}{2T}(p^2-p)\|h\|_H^2\right) < \infty.$$

Now we need to prove analogous results for J_g and its derivatives.

Proposition 8.2. *For any $p \in [1,\infty)$ and $g \in G_{CM}$,*

$$\mathbb{E}\left[\int_{\mathbf{C}} dc \sup_{|\varepsilon|\le 1} J_{\varepsilon g}(B,c)^p\right] < \infty.$$

Proof. For $g = (h, z)$, we have that

$$\begin{aligned}&\sup_{|\varepsilon|\le 1} J_{\varepsilon g}(B,c) \\ &\le \frac{\exp\left(-\frac{1}{2}\inf_{|\varepsilon|\le 1}\rho_T(B-\varepsilon\mathbf{h})^{-1}\left(c-\varepsilon z-\frac{1}{2}\varepsilon\omega(B_T,h)\right)\cdot\left(c-\varepsilon z-\frac{1}{2}\varepsilon\omega(B_T,h)\right)\right)}{\inf_{|\varepsilon|\le 1}\sqrt{\det(2\pi\rho_T(B-\varepsilon\mathbf{h}))}}.\end{aligned}$$

Lemma 5.8 gives the integrability of $\left(\inf_{|\varepsilon|\le 1}\det(2\pi\rho_T(B-\varepsilon\mathbf{h}))\right)^{-p/2}$, so we need now only deal with the exponential term.

For brevity, set $V = z - \frac{1}{2}\omega(B_T, h)$ (as a random element of C); we must show that

$$\exp\left(-\frac{p}{2}\inf_{|\varepsilon|\le 1}\rho_T(B-\varepsilon\mathbf{h})^{-1}(c-\varepsilon V)\cdot(c-\varepsilon V)\right) \in L^1(\mathbb{P}\times dc).$$

Observe the following elementary inequality from linear algebra: if $x \in C$ and $A \in \mathrm{End}(C)$ is a symmetric positive-definite linear transformation, then

$$|x|^2 = |A^{1/2}A^{-1/2}x|^2 \le \|A^{1/2}\|_{op}^2|A^{-1/2}x|^2 = \|A\|_{op}(A^{-1}x\cdot x).$$

Thus if we set $Y = \sup_{|\varepsilon|\le 1}\|\rho_T(B-\varepsilon\mathbf{h})\|_{op}$, we have

$$\rho_T(B-\varepsilon\mathbf{h})^{-1}(c-\varepsilon V)\cdot(c-\varepsilon V) \ge \frac{1}{Y}|c-\varepsilon V|^2.$$

For $|\varepsilon|\le 1$, we have $|c-\varepsilon V| \ge |c| - \varepsilon|V| \ge |c| - |V|$, and thus

$$|c-\varepsilon V| \ge \begin{cases} 0, & |c| < 2|V|, \\ \frac{1}{2}|c|, & |c| \ge 2|V|. \end{cases}$$

Putting all of this together, we have

$$\begin{aligned}
&\mathbb{E}\left[\int_C \exp\left(-\frac{p}{2}\inf_{|\varepsilon|\le 1}\rho_T(B-\varepsilon\mathbf{h})(c-\varepsilon V)\cdot(c-\varepsilon V)\right)\,dc\right] \\
&\le \mathbb{E}\left[\int_{|c|\le 2|V|} 1\,dc + \int_{|c|\ge 2|V|}\exp\left(-\frac{p}{8Y}|c|^2\right)\,dc\right] \\
&\le \mathbb{E}\left[m(B(0,2|V|)) + \left(\frac{8\pi Y}{p}\right)^{d/2}\right] \le C_p\mathbb{E}\left[|V|^d + Y^{d/2}\right]
\end{aligned}$$

where m denotes Lebesgue measure on C. To complete the proof, it suffices to show that $|V|, Y \in L^{\infty-}(\mathbb{P})$. This is straightforward since

$$\|\rho_T(B-\varepsilon\mathbf{h})\|_{op} \le \|\rho_T(B)\|_{op} + 2\|\rho_T(B,\mathbf{h})\|_{op} + \|\rho_T(\mathbf{h})\|_{op}$$

for $|\varepsilon|\le 1$, $\rho_T(B)$ is in the second homogeneous Wiener chaos, and $\omega(B_T,h)$ and $\rho_T(B,\mathbf{h})$ are linear in B and hence Gaussian. □

Definition 8.3. A **polynomial in** $A_1,\dots,A_k \in \mathrm{End}(C)$ **and** $c_1,\dots,c_\ell \in C$ is a function which may be written as sums of products of factors of the form

$$A_{i_1}\cdots A_{i_r}c_i\cdot c_j \quad \text{and} \quad \mathrm{tr}(A_{i_1}\cdots A_{i_r})$$

for some $i_1,\dots,i_r \in \{1,\dots,k\}$ and $i,j \in \{1,\dots,\ell\}$.

Lemma 8.4. *Given $h_1,\dots,h_m \in H$ and $X = (h,z) \in \mathfrak{g}_{CM}$, suppose that $F = F(B,c)$ is polynomial in the matrices $\rho_T(B)^{-1}$ and $\rho_T(B,\mathbf{h}_i)$ and vectors c and $\omega(B_T,h_i)$. Then for any $p\in[1,\infty)$*

$$\mathbb{E}\left[\int_C dc\,\sup_{|\varepsilon|\le 1}\left\{|F_{\varepsilon X}(B,c)|^p\,J_{\varepsilon X}(B,c)\right\}\right] < \infty.$$

Proof. There exist $K, M < \infty$ such that

$$\begin{aligned}
|F_{\varepsilon X}(B,c)| \le K\Bigg(1 + \|\rho_T(B-\varepsilon\mathbf{h})^{-1}\|_{op} + \sum_{i=1}^m\Bigg(&\|\rho_T(B-\varepsilon\mathbf{h},\mathbf{h}_i)\|_{op} \\
&+ |\omega(B_T-\varepsilon h,h_i)|_C + \left|c-\varepsilon z - \frac{1}{2}\omega(B_T-\varepsilon h,h_i)\right|_C\Bigg)\Bigg)^M.
\end{aligned}$$

Thus,

$$\sup_{|\varepsilon|\le 1}|F_{\varepsilon X}(B,c)| \le C(h_i,X)\Bigg(1+\sup_{|\varepsilon|\le 1}\|\rho_T(B-\varepsilon\mathbf{h})^{-1}\|_{op} + \sum_{i=1}^{m}(\|\rho_T(B,\mathbf{h}_i)\|_{op}+|\omega(B_T,h_i)|_C)+|c|_C\Bigg)^M$$

and the result follows from each $\omega(B_T,h_i)$ being Gaussian, Fernique's theorem, Lemma 5.8, Proposition 8.2, and the fact that

$$\mathbb{E}\left[\int_C dc\,|c|_C^M \sup_{|\varepsilon|\le 1} J_{\varepsilon X}(B,c)\right]<\infty$$

by computations analogous to those in the proof of Proposition 8.2. □

Proposition 8.5. *For any $p\in[1,\infty)$ and $X=(h,z)\in G_{CM}$,*

$$\mathbb{E}\left[\int_{\mathbf{C}} dc\,\sup_{|\varepsilon|\le 1}\left|\frac{d}{d\varepsilon}J_{\varepsilon X}(B,c)\right|^p\right]<\infty.$$

Proof. First note that for any $\mathbf{x},\mathbf{h}\in C([0,T],H)$,

$$\frac{d}{d\varepsilon}\rho_T(\mathbf{x}-\varepsilon\mathbf{h})^{-1} = 2\rho_T(\mathbf{x}-\varepsilon\mathbf{h})^{-1}\rho_T(\mathbf{x}-\varepsilon\mathbf{h},\mathbf{h})\rho_T(\mathbf{x}-\varepsilon\mathbf{h})^{-1} \tag{8.1}$$

and

$$\frac{d}{d\varepsilon}\det(\rho_T(\mathbf{x}-\varepsilon\mathbf{h}))^{-1/2} = \frac{1}{\sqrt{\det(\rho_T(\mathbf{x}-\varepsilon\mathbf{h}))}}\cdot\operatorname{tr}\left(\rho_T(\mathbf{x}-\varepsilon\mathbf{h})^{-1}\rho_T(\mathbf{x}-\varepsilon\mathbf{h},\mathbf{h})\right). \tag{8.2}$$

Thus

$$\frac{d}{d\varepsilon}J_{\varepsilon X}(B,c) = J_{\varepsilon X}(B,c)\cdot\left\{\varphi_\varepsilon(B,c)+\operatorname{tr}\left(\rho_T(B-\varepsilon\mathbf{h})^{-1}\rho_T(B-\varepsilon\mathbf{h},\mathbf{h})\right)\right\}$$

where

$$\begin{aligned}\varphi_\varepsilon(B,c) &= \frac{d}{d\varepsilon}\left\{-\frac{1}{2}\rho_T(B-\varepsilon\mathbf{h})^{-1}(c-\varepsilon V)\cdot(c-\varepsilon V)\right\}\\ &= -\rho_T(B-\varepsilon\mathbf{h})^{-1}\rho_T(B-\varepsilon\mathbf{h},\mathbf{h})\rho_T(B-\varepsilon\mathbf{h})^{-1}(c-\varepsilon V)\cdot(c-\varepsilon V)\\ &\quad+\rho_T(B-\varepsilon\mathbf{h})^{-1}V\cdot(c-\varepsilon V);\end{aligned}$$

here we again have taken $V=z+\frac{1}{2}\omega(B_T,h)$. We have also used the symmetry of $\rho_T(\mathbf{x})$ to simplify the above expression. Thus, the proof follows by Proposition 8.2 and Lemma 8.4. □

Remark 8.6. From computations like (8.1) and (8.2), it is clear that, for $X=(h,z)\in\mathfrak{g}_{CM}$ and $F=F(B,c)$ a polynomial as in Lemma 8.4, the function $\tilde{\mathbf{X}}F$ is again a polynomial in $\rho_T(B)^{-1},\rho_T(B,\mathbf{h}_i),\rho_T(B,\mathbf{h}),\omega(B_T,h_i),\omega(B_T,h),\omega(h,h_i)$,

z, and c. Thus, Lemma 8.4 is sufficient to show that under the same assumptions

$$\mathbb{E}\left[\int_C dc \sup_{|\varepsilon|\le 1}\left\{\left|\frac{d}{d\varepsilon}F_{\varepsilon X}(B,c)\right|^p J_{\varepsilon X}(B,c)\right\}\right] = \mathbb{E}\left[\int_C dc \sup_{|\varepsilon|\le 1}\left\{\left|(\tilde{\mathbf{X}}F)_{\varepsilon X}(B,c)\right|^p J_{\varepsilon X}(B,c)\right\}\right] < \infty.$$

Lemma 8.7. *For any $m \in \mathbb{N}$ and $X_1, \dots, X_m \in \mathfrak{g}_{CM}$, $\tilde{\mathbf{X}}_m \cdots \tilde{\mathbf{X}}_1 \log J_T^0(B,c)$ is a polynomial as in Lemma* 8.4. *In particular, we have that $\tilde{\mathbf{X}}_m \cdots \tilde{\mathbf{X}}_1 \log J_T^0(B,c) \in L^{\infty-}(J_T^0(B,c)\, d\mathbb{P}\, dc)$.*

Proof. For $m = 1$, this is essentially Proposition 8.5. In particular, the computations in the proof of Proposition 8.5 imply that, for $X = (h,z) \in \mathfrak{g}_{CM}$,

$$\tilde{\mathbf{X}} \log J_T^0(B,c) = -\rho_T(B)^{-1}\rho_T(B,\mathbf{h})\rho_T(B)^{-1}c \cdot c + \rho_T(B)^{-1}c \cdot \left(z + \frac{1}{2}\omega(B_T,h)\right) + \operatorname{tr}\left(\rho_T(B)^{-1}\rho_T(B,\mathbf{h})\right).$$

In particular, $\tilde{\mathbf{X}} \log J_T^0(B,c)$ is a polynomial in $\rho_T(B)^{-1}$, $\rho_T(B,h)$ and $c, z, \omega(B_T,h)$. Then the result follows from Lemma 8.4 and Remark 8.6. □

Definition 8.8. For all $(x,c) \in G$, let $\|(x,c)\|_{\mathfrak{g}} := \|x\|_W + |c|_C$. A function $F : G \to \mathbb{R}$ is *polynomially bounded* if there exist constants $K, M < \infty$ such that

$$|F(g)| \le K\left(1 + \|g\|_{\mathfrak{g}}\right)^M$$

for all $g \in G$. Given $X \in \mathfrak{g}_{CM}$, we say F *is left X-differentiable* if

$$(\tilde{X}F)(g) := \frac{d}{d\varepsilon}\Big|_0 F(g \cdot \varepsilon X)$$

exists for all $g \in G$. We will say that F is *smooth* if $(\tilde{X}_1 \cdots \tilde{X}_m F)(g)$ exists for all $m \in \mathbb{N}$, $X_1, \dots, X_m \in \mathfrak{g}_{CM}$, and $g \in G$.

Theorem 8.9. *Let $X = (h,z) \in \mathfrak{g}_{CM}$, $F : G \to \mathbb{R}$ be left X-differentiable such that F and $\tilde{X}F$ are polynomially bounded, and $\Psi = \Psi(B,c)$ be a polynomial as in Lemma* 8.4. *Then*

$$\int_C dc\, \mathbb{E}[(\tilde{X}F)(B_T,c)\Psi(B,c)J_T^0(B,c)] = \int_C \mathbb{E}[F(B_T,c)(\tilde{\mathbf{X}}^*\Psi)(B,c)J_T^0(B,c)] \tag{8.3}$$

where

$$\tilde{\mathbf{X}}^*(B,c) := \tilde{\mathbf{X}}(B,c) + \tilde{\mathbf{X}} \log J_T^0(B,c) + \langle h, B_T\rangle$$

for $\tilde{\mathbf{X}}$ as in Notation 7.3. *Furthermore, for any $X_1, \dots, X_m \in \mathfrak{g}_{CM}$,*

$$\mathbb{E}[(\tilde{X}_1 \cdots \tilde{X}_m F)(g_T)] = \int_C \mathbb{E}[F(B_T,c)(\tilde{\mathbf{X}}_1^* \cdots \tilde{\mathbf{X}}_m^* 1)(B,c)J_T^0(B,c)] \tag{8.4}$$

where $(\tilde{\mathbf{X}}_1^* \cdots \tilde{\mathbf{X}}_m^* 1)(B,c) \in L^{\infty-}(J_T^0(B,c)\, d\mathbb{P}\, dc)$.

Proof. Since $\tilde{X}F$ is polynomially bounded, there exist $K, M < \infty$ such that

$$\begin{aligned}\sup_{|\varepsilon|\leq 1}\left|\frac{d}{d\varepsilon}F((B_T,c)\cdot\varepsilon X)\right| &= \sup_{|\varepsilon|\leq 1}\left|(\tilde{X}F)((B_T,c)\cdot\varepsilon X)\right| \\ &\leq \sup_{|\varepsilon|\leq 1} K\left(1+\|(B_T,c)\cdot\varepsilon X\|_{\mathfrak{g}}\right)^M \\ &\leq C(X)\left(1+\|(B_T,c)\|_{\mathfrak{g}}\right)^M\end{aligned}$$

where this last expression is in $L^{\infty-}(J_T^0(B,c)\,d\mathbb{P}\,dc)$ again by Fernique's theorem and arguments similar to those in the proof of Proposition 8.2. Then this implies that

$$\begin{aligned}&\int_C dc\,\mathbb{E}\Big[(\tilde{X}F)(B_T,c)\Psi(B,c)J_T^0(B,c)\Big] \\ &\qquad = \int_C dc\,\mathbb{E}\left[\frac{d}{d\varepsilon}\Big|_0 F((B_T,c)\cdot\varepsilon X)\Psi(B,c)J_T^0(B,c)\right] \\ &\qquad = \frac{d}{d\varepsilon}\Big|_0\int_C dc\,\mathbb{E}\left[F((B_T,c)\cdot\varepsilon X)\Psi(B,c)J_T^0(B,c)\right].\end{aligned}$$

Now applying Corollary 7.4 gives that

$$\begin{aligned}&\int_C dc\,\mathbb{E}\Big[(\tilde{X}F)(B_T,c)\Psi(B,c)J_T^0(B,c)\Big] \\ &\quad = \frac{d}{d\varepsilon}\Big|_0\int_C dc\,\mathbb{E}\left[F(B_T,c)\Psi_{\varepsilon X}(B,c)J_{\varepsilon X}(B,c)\bar{J}_{\varepsilon h}(B_T)\right] \\ &\quad = \int_C dc\,\mathbb{E}\Bigg[F(B_T,c)\Bigg((\tilde{\mathbf{X}}\Psi)(B,c)J_T^0(B,c) \\ &\qquad\qquad + \Psi(B,c)\left(\tilde{\mathbf{X}}J_T^0(B,c) + J_T^0(B,c)\frac{d}{d\varepsilon}\Big|_0\bar{J}_{\varepsilon h}(B_T)\right)\Bigg)\Bigg] \\ &\quad = \int_C dc\,\mathbb{E}\left[F(B_T,c)\left(\tilde{\mathbf{X}}+\tilde{\mathbf{X}}\log J_T^0(B,c)+\langle h,B_T\rangle\right)\Psi(B,c)J_T^0(B,c)\right]\end{aligned}$$

where this second interchange of differentiation and integration is justified by Propositions 8.1 and 8.5, Lemma 8.4, and Remark 8.6. This completes the proof of (8.3).

Now, equation (8.3) implies that

$$\mathbb{E}[(\tilde{X}_1\cdots\tilde{X}_mF)(g_T)] = \int_C dc\,\mathbb{E}\left[(\tilde{X}_2\cdots\tilde{X}_mF)(B_T,c)(\tilde{\mathbf{X}}_1^*1)(B,c)J_T^0(B,c)\right]$$

where $(\tilde{\mathbf{X}}_1^*1)(B,c) = \tilde{\mathbf{X}}_1\log J_T^0(B,c) + \langle h_1,B_T\rangle$. By Lemma 8.7, $\tilde{\mathbf{X}}_1\log J_T^0(B,c)$ is a polynomial as in Lemma 8.4. We also have that $\langle h_1,B_T\rangle$ is Gaussian, and for $X=(h,z)\in\mathfrak{g}_{CM}$, $\tilde{X}\langle h_1,x\rangle = \partial_h\langle h_1,x\rangle = \langle h_1,h\rangle$. Thus, we may again use (8.3) and Proposition 8.1, along with Lemmas 8.4 and 8.7 and Remark 8.6, and iterative applications of these gives the desired result. The integrability follows from Proposition 8.1 and Lemma 8.7. □

The previous theorem and its proof give the following.

Corollary 8.10 (Integration by parts)**.** *Let $X = (h, z) \in \mathfrak{g}_{CM}$ and $F_1, F_2 : G \to \mathbb{R}$ be left X-differentiable such that $F_1, F_2, \tilde{X}F_1, \tilde{X}F_2$ are polynomially bounded. Then*

$$\mathbb{E}[(\tilde{X}F_1)(g_T)F_2(g_T)] = \mathbb{E}[F_1(g_T)(\tilde{X}^*F_2)(g_T)]$$

where

$$\tilde{X}^*(x,c) := -\tilde{X}(x,c) + \mathbb{E}[(\tilde{\mathbf{X}}\log J_T^0)(B,c)J_T^0(B,c) \,|\, B_T = x]\gamma_T^{-1}(x,c) + \langle h, x\rangle.$$

Proof. By Theorems 4.6 and 8.9, we have that

$$\begin{aligned}
\mathbb{E}[(\tilde{X}F_1)(g_T)F_2(g_T)] &= \int_C dc\, \mathbb{E}[(\tilde{X}F_1)(B_T,c)F_2(B_T,c)J_T^0(B,c)] \\
&= \int_C dc\, \mathbb{E}[F_1(B_T,c)(\tilde{\mathbf{X}}^*F_2)(B_T,c)J_T^0(B,c)] \\
&= \int_C dc\, \mathbb{E}\Big[F_1(B_T,c)\Big(-(\tilde{X}F_2)(B_T,c) \\
&\qquad + F_2(g_T)\left((\tilde{\mathbf{X}}\log J_T^0)(B,c) + \langle h, B_T\rangle\right)\Big)J_T^0(B,c)\Big].
\end{aligned}$$

□

Theorem 8.9, in particular, equation (8.4), also immediately gives the following result for higher order derivatives.

Corollary 8.11. *Let $m \in \mathbb{N}$ and $X_1, \dots, X_m \in \mathfrak{g}_{CM}$. Then, for any smooth $F : G \to \mathbb{R}$ such that F and all of its derivatives are polynomially bounded,*

$$\mathbb{E}[(\tilde{X}_1 \cdots \tilde{X}_m F)(g_T)] = \mathbb{E}[F(g_T)\psi^{X_m,\dots,X_1}(g_T)] \tag{8.5}$$

where

$$\psi^{X_m,\dots,X_1}(x,c) = \mathbb{E}\left[(\tilde{\mathbf{X}}_m^* \cdots \tilde{\mathbf{X}}_1^* 1)(B,c)J_T^0(B,c) \,\middle|\, B_T = x\right]\gamma_T^{-1}(x,c)$$

and $\psi^{X_m,\dots,X_1} \in L^{\infty-}(\nu_T)$.

Proof. Equation (8.5) and the expression given for $\psi^{X_m,\dots,X_1}$ are a direct consequence of Theorem 8.9. Now note that the integrability of $\tilde{\mathbf{X}}_m^* \cdots \tilde{\mathbf{X}}_1^* 1$ implies that $\psi^{X_m,\dots,X_1} \in L^1(\nu_T)$ and $F\psi^{X_m,\dots,X_1} \in L^1(\nu_T)$ for any such F. More particularly, for any fixed $p \in (1,\infty)$, we have that

$$\begin{aligned}
&\mathbb{E}|F(g_T)\psi^{X_m,\dots,X_1}(g_T)| \\
&\quad = \int_C dc\, \mathbb{E}\left[\left|F(B_T,c)\mathbb{E}\left[(\tilde{\mathbf{X}}_m^* \cdots \tilde{\mathbf{X}}_1^* 1)(B,c)J_T^0(B,c) \,\middle|\, B_T\right]\right| \gamma_T^{-1}(B_T,c)J_T^0(B,c)\right] \\
&\quad \le \int_C dc\, \mathbb{E}\left[\left|F(B_T,c)(\tilde{\mathbf{X}}_m^* \cdots \tilde{\mathbf{X}}_1^* 1)(B,c)\right| J_T^0(B,c)\right] \\
&\quad \le \|F\|_{L^p(\nu_T)}\|\tilde{\mathbf{X}}_m^* \cdots \tilde{\mathbf{X}}_1^* 1\|_{L^q(J_T^0(B,c)\, d\mathbb{P}\, dc)}
\end{aligned}$$

where q is the conjugate exponent to p. As this bound holds, for example, for all smooth cylinder functions F and these comprise a dense subspace of $L^p(\nu_T)$, this implies that $\psi^{X_m,\dots,X_1}$ represents a bounded linear operator on $L^p(\nu_T)$ and thus $\psi^{X_m,\dots,X_1} \in (L^p(\nu_T))^* \cong L^q(\nu_T)$. As this holds for any $p \in (1,\infty)$, we have that $\psi^{X_m,\dots,X_1} \in \bigcap_{q\in[1,\infty)} L^q(\nu_T)$. □

Right integration by parts formulae may be proved directly, but also follow from the left integration by parts formulae combined with the invariance of the heat kernel measure under inversions. For the following corollary, let $\hat{X}$ denote the right-invariant vector field given by

$$\hat{X}F(g) = \partial_{(h,z)}(F \circ R_g)(\mathbf{e})$$

where R_g is right translation by $g \in G$. A straightforward computation shows that

$$\hat{X}F(x,c) = \partial_{(h,z-\frac{1}{2}\omega(x,h))}F(x,c)$$

(compare with Notation 1.8 for left-invariant vector fields).

Corollary 8.12. *Under the hypotheses of Corollary* 8.11*,*

$$\mathbb{E}[(\hat{X}_1 \cdots \hat{X}_m F)(g_T)] = \mathbb{E}[F(g_T)\hat{\psi}^{X_m,\dots,X_1}(g_T)]$$

where

$$\hat{\psi}^{X_m,\dots,X_1}(g) := (-1)^m \psi^{X_m,\dots,X_1}(g^{-1}).$$

Proof. Take $u(g) := F(g^{-1})$. We proceed by induction. Note first that, for any $g \in G$ and $X \in \mathfrak{g}_{CM}$,

$$(\hat{X}F)(g) = \frac{d}{d\varepsilon}\bigg|_0 F(\varepsilon X \cdot g) = \frac{d}{d\varepsilon}\bigg|_0 u(g^{-1} \cdot -\varepsilon X) = -(\tilde{X}u)(g^{-1}). \tag{8.6}$$

This equation and repeated applications of Corollary 4.9 give that

$$\begin{aligned}
\mathbb{E}[(\hat{X}F)(g_T)] &= -\mathbb{E}[(\tilde{X}u)(g_T^{-1})] = -\mathbb{E}[(\tilde{X}u)(g_T)] \\
&= -\mathbb{E}[u(g_T)\psi^X(g_T)] = -\mathbb{E}[F(g_T^{-1})\psi^X(g_T)] \\
&= -\mathbb{E}[F(g_T)\psi^X(g_T^{-1})]
\end{aligned}$$

where we have applied Corollary 8.11 in the third equality. Now assuming the formula for m and again using equation (8.6), Corollary 4.9, and Corollary 8.11 gives

$$\begin{aligned}
\mathbb{E}\left[(\hat{X}_1 \cdots \hat{X}_{m+1}F)(g_T)\right] &= (-1)^m \mathbb{E}\left[(\hat{X}_{m+1}F)(g_T)\psi^{X_m,\dots,X_1}(g_T^{-1})\right] \\
&= (-1)^{m+1}\mathbb{E}\left[(\tilde{X}_{m+1}u)(g_T^{-1})\psi^{X_m,\dots,X_1}(g_T^{-1})\right] \\
&= (-1)^{m+1}\mathbb{E}\left[(\tilde{X}_{m+1}u)(g_T)\psi^{X_m,\dots,X_1}(g_T)\right] \\
&= (-1)^{m+1}\mathbb{E}\left[u(g_T)\psi^{X_{m+1},\dots,X_1}(g_T)\right] \\
&= (-1)^{m+1}\mathbb{E}\left[F(g_T^{-1})\psi^{X_{m+1},\dots,X_1}(g_T)\right] \\
&= (-1)^{m+1}\mathbb{E}\left[F(g_T)\psi^{X_{m+1},\dots,X_1}(g_T^{-1})\right].
\end{aligned}$$

□

9. Conclusion

We have shown that the hypoelliptic heat kernel measure ν_T on G is smooth, in a sense that naturally extends the well-known smoothness results in finite dimensions; namely, it is quasi-invariant under left and right translations by elements of the Cameron–Martin subgroup G_{CM}.

In flat abstract Wiener space (W, H, μ), the smoothness of Gaussian measure under translation by H (established by the Cameron–Martin theorem) is the starting point for defining the gradient operator on $L^2(W,\mu)$, the associated Sobolev spaces,

chaos decompositions, the Skorohod integral, and many other developments. Similar results should be possible in our hypoelliptic setting, and we hope in the future to explore some of this territory.

In this paper, we have considered only groups G whose center C is finite-dimensional, and our argument makes essential use of this assumption in several places. It would be interesting to relax this assumption, to allow for infinite-dimensional centers. For example, the definition of G makes sense if C is replaced by a separable Hilbert space. However, the lack of a natural reference measure on C seems to be a significant obstruction to proving analogous results in this case.

In another direction, it would be interesting to consider a more general class of groups; for example, nilpotent Lie groups of step three or higher. Unfortunately, in such examples, our argument for Theorem 2.1 no longer applies; and in the approach used in Appendix A, the commutation of terms analogous to S and L may fail. It appears that new ideas may be required to proceed beyond step two.

Appendix A. Another proof of Theorem 2.1

In this section, we provide another self-contained proof of Theorem 2.1, which is based on the analysis of the infinitesimal generator of g_t. We will begin with a slightly informal version of the proof and then fill in the missing technical points.

Recall that $\{B_t\}_{t\geq 0}$ is an N-dimensional Brownian motion, A is an $N\times N$ skew-symmetric matrix, and we wish to show for any measurable $f:\mathbb{R}^N\to\mathbb{C}$ such that $\mathbb{E}\left|f\left(B_T\right)\right|<\infty$, we have

$$\mathbb{E}\left[f\left(B_T\right)e^{i\int_0^T AB_t\cdot dB_t}\right]=\mathbb{E}\left[f\left(B_T\right)e^{-\frac{1}{2}\int_0^T|AB_t|^2dt}\right]. \tag{A.1}$$

Suppose that $F:\mathbb{R}\times\mathbb{R}^N\to\mathbb{R}$ is a C^2-function such that F and its derivatives up to order 2 have at most polynomial growth, and set

$$Y_t:=e^{i\int_0^t AB_\tau\cdot dB_\tau}.$$

Then by Itô's formula

$$dY_t=Y_t\left(iAB_t\cdot dB_t-\frac{1}{2}\left|AB_t\right|^2dt\right)$$

and

$$\begin{aligned}d\left[F\left(t,B_t\right)Y_t\right]&=Y_t\left(\nabla F_t\left(t,B_t\right)+iF\left(t,B_t\right)AB_t\right)\cdot dB_t\\&\quad+Y_t\left(F_t\left(t,B_t\right)+\frac{1}{2}\Delta F\left(t,B_t\right)-\frac{1}{2}\left|AB_t\right|^2F\left(t,B_t\right)+iAB_t\cdot\nabla F_t\left(t,B_t\right)\right)dt.\end{aligned}$$

(Here $\nabla F(t,x)$ denotes the gradient with respect to the x variable only.) By our assumptions on F it is easily verified that

$$\mathbb{E}\left[\int_0^T\left|Y_t(\nabla F_t\left(t,B_t\right)+iF\left(t,B_t\right)AB_t)\right|^2dt\right]<\infty,$$

and hence $\int_0^{\cdot} Y_t \left(\nabla F_t\left(t, B_t\right) + iF\left(t, B_t\right) AB_t\right) \cdot dB_t$ is a square integrable martingale. Therefore it follows that

(A.2)

$$\begin{aligned}
&\frac{d}{dt}\mathbb{E}\left[F\left(t, B_t\right) Y_t\right] \\
&= \mathbb{E}\left[Y_t\left(F_t\left(t, B_t\right) + \frac{1}{2}\Delta F\left(t, B_t\right) - \frac{1}{2}\left|AB_t\right|^2 F\left(t, B_t\right) + iAB_t \cdot \nabla F\left(t, B_t\right)\right)\right].
\end{aligned}$$

Now suppose that $f : \mathbb{R}^N \to \mathbb{R}$ and $T > 0$ are given such that there exists a function F as above, with the additional properties that $F\left(T, x\right) = f\left(x\right)$ and

$$F_t\left(t, x\right) + \frac{1}{2}\Delta F\left(t, x\right) - \frac{1}{2}\left|Ax\right|^2 F\left(t, x\right) + iAx \cdot \nabla F\left(t, x\right) = 0$$

for all (t, x). It then follows from (A.2) that $\frac{d}{dt}\mathbb{E}\left[F\left(t, B_t\right) Y_t\right] = 0$ and, in particular,

$$\mathbb{E}\left[f\left(B_T\right) Y_T\right] = \mathbb{E}\left[F\left(T, B_T\right) Y_T\right] = \mathbb{E}\left[F\left(0, B_0\right) Y_0\right] = F\left(0, 0\right). \tag{A.3}$$

Formally, the solution to (A.2) is given by

$$F\left(t, x\right) = \left(e^{(T-t)(L+S)} f\right)\left(x\right)$$

where

$$\begin{aligned}
\left(Lf\right)\left(x\right) &:= \frac{1}{2}\Delta f\left(x\right) - \frac{1}{2}\left|Ax\right|^2 f\left(x\right) \text{ and} \\
\left(Sf\right)\left(x\right) &:= iAx \cdot \nabla f\left(x\right).
\end{aligned}$$

With this notation we may summarize (A.3) as

$$\mathbb{E}\left[f\left(B_T\right) Y_T\right] = \left(e^{T(L+S)} f\right)\left(0\right).$$

Since $e^{-itS} f\left(x\right) = f\left(e^{tA} x\right)$ where e^{tA} is a rotation and Δ is invariant under rotations, it follows that

$$\left(e^{-itS}\Delta f\right)\left(x\right) = \left(\Delta f\right)\left(e^{tA} x\right) = \Delta\left(f \circ e^{tA}\right)\left(x\right) = \left(\Delta e^{-itS} f\right)\left(x\right).$$

Differentiating this equation in t then shows $\left[S, \Delta\right] = 0$. Also,

$$\begin{aligned}
\left(\left|Ay\right|^2 f\left(y\right)\right)\Big|_{y=e^{tA}x} &= \left|Ae^{tA}x\right|^2 f\left(e^{tA}x\right) = \left|e^{tA}Ax\right|^2 f\left(e^{tA}x\right) \\
&= \left|Ax\right|^2 f\left(e^{tA}x\right).
\end{aligned}$$

Equivalently, for M_g multiplication by g, $e^{-itS} M_{|A(\cdot)|^2} = M_{|A(\cdot)|^2} e^{-itS}$, and differentiating this at $t = 0$ implies

$$\left[S, M_{|A(\cdot)|^2}\right] = 0.$$

Combining this with $\left[S, \Delta\right] = 0$ allows us to conclude $\left[S, L\right] = 0$ and therefore, formally,

$$\mathbb{E}\left[f\left(B_T\right) Y_T\right] = \left(e^{T(L+S)} f\right)\left(0\right) = \left(e^{TS} e^{TL} f\right)\left(0\right) = \left(e^{TL} f\right)\left(0\right)$$

wherein the last equality we have used the fact that $Sf\left(0\right) = 0$ for all functions f. Hence, by an application of the Feynman–Kac formula, we conclude that (A.1) holds.

In order to make the above argument rigorous, it is helpful to find finite-dimensional subspaces of functions which are invariant under the actions of L and

S. So suppose for the moment that A is non-degenerate, in which case L is a harmonic oscillator Hamiltonian. If $\psi \in L^2(\mathbb{R}^n)$ is an eigenvector of L with eigenvalue λ, then

$$LS\psi = SL\psi = S\lambda\psi = \lambda S\psi.$$

Given any $\Lambda \in (0,\infty)$, we let K_Λ be the linear combination of eigenfunctions of L with eigenvalues $\lambda \leq \Lambda$. Then K_Λ is a finite-dimensional subspace of $\mathcal{S}(\mathbb{R}^N) \subset L^2(\mathbb{R}^N)$ which is invariant under the actions of L and S. For $f \in K_\Lambda$ all of the manipulations in the previous paragraph are justified and therefore (A.1) holds for all $f \in \bigcup_{\Lambda<\infty} K_\Lambda$, which is dense in $\mathcal{S}(\mathbb{R}^N)$. The full result for non-degenerate A then follows by density arguments.

We now have a couple of choices for how to proceed when A is degenerate. The first is to decompose $\mathbb{R}^N$ in $\mathrm{Nul}(A) \oplus \mathrm{Nul}(A)^\perp$ and then decompose the Brownian motion accordingly. With this decomposition it basically then suffices to prove (A.1) in two cases corresponding to $A = 0$ and to A being non-degenerate. As the case where $A = 0$ is a triviality, the argument is essentially complete. An alternative is to modify the method of the previous paragraph so as to work for general skew-symmetric A, and this is what we do now.

We start by finding the "ground state" for L in the form $\Phi(x) = \exp\left(-\frac{1}{2}\Sigma x \cdot x\right)$ for some symmetric non-negative $N \times N$ matrix Σ. A simple computation shows

$$2L\Phi = \Phi \cdot \left(\nabla \cdot (-\Sigma x) + |\Sigma x|^2 - |Ax|^2\right) = \Phi \cdot \left(-\operatorname{tr}(\Sigma) + |\Sigma x|^2 - |Ax|^2\right).$$

Thus taking $\Sigma := \sqrt{A^*A} = \sqrt{-A^2}$ then implies $L\Phi = -\frac{1}{2}\operatorname{tr}(\Sigma)\Phi$. Further observe that

$$S\Phi = iAx \cdot \nabla\Phi = -i(Ax \cdot \Sigma x)\Phi = 0$$

wherein we have used that $[A,\Sigma] = 0$ implies that

$$Ax \cdot \Sigma x = -x \cdot A\Sigma x = -x \cdot \Sigma Ax = -\Sigma x \cdot Ax,$$

and hence $Ax \cdot \Sigma x = 0$. Now let $\mathcal{P}_m$ denote the space of polynomial functions on $\mathbb{R}^N$ with degrees less than or equal to m, and let $\Phi\mathcal{P}_m := \{\Phi p : p \in \mathcal{P}_m\}$. It then follows that[3] $\Phi\mathcal{P}_m$ is a finite-dimensional subspace of functions on $\mathbb{R}^N$ which are invariant under the action of both L and S. Indeed, for any $g \in C^\infty(\mathbb{R}^N)$ we have

$$\begin{aligned} L(\Phi g) &= (L\Phi)g + \Phi \cdot \frac{1}{2}\Delta g + \nabla\Phi \cdot \nabla g \\ &= \Phi \cdot \left(-\frac{1}{2}\operatorname{tr}(\Sigma)\, g + \frac{1}{2}\Delta g + \nabla \ln\Phi \cdot \nabla g\right) \\ &= \Phi \cdot \left(\frac{1}{2}\Delta - \Sigma x \cdot \nabla - \frac{1}{2}\operatorname{tr}(\Sigma)\right) g \end{aligned}$$

and

$$S(\Phi g) = \Phi \cdot Sg.$$

Thus, for any $f \in \Phi\mathcal{P}_m$, we have

$$\mathbb{E}[f(B_T)\, Y_T] = \left(e^{TL} f\right)(0) = \mathbb{E}\left[f(B_T)\, e^{-\frac{1}{2}\int_0^T |AB_t|^2 dt}\right]$$

where the second equality follows from the fact that

$$d\left[e^{-\frac{1}{2}\int_0^t |AB_\tau|^2 d\tau}\left(e^{(T-t)L} f\right)(B_t)\right] = e^{-\frac{1}{2}\int_0^t |AB_\tau|^2 d\tau}\left(\nabla e^{(T-t)L} f\right)(B_t) \cdot dB_t$$

[3]The space $\Phi\mathcal{P}_m$ is a spectral subspace as described above in the case that A is non-degenerate.

by Itô's lemma, and thus

$$\mathbb{E}\left[e^{-\frac{1}{2}\int_0^t |AB_\tau|^2 d\tau}\left(e^{(T-t)L}f\right)(B_t)\right]$$

is constant in t; comparing the values of this expression at $t = T$ and $t = 0$ then gives the desired equality.

For any $z \in \mathbb{C}^N$,

$$\begin{aligned}\mathbb{E}\left|e^{z\cdot B_T} - \sum_{n=0}^{K}\frac{(z\cdot B_T)^n}{n!}\right| &= \mathbb{E}\left|\sum_{n=K+1}^{\infty}\frac{(z\cdot B_T)^n}{n!}\right| \\ &\leq \mathbb{E}\left[\sum_{n=K+1}^{\infty}\frac{|z|^n\cdot|B_T|^n}{n!}\right] \to 0 \text{ as } K\to\infty\end{aligned}$$

by the dominated convergence theorem because the last integrand goes to 0 as $K \to \infty$ and satisfies the estimate

$$\sum_{n=K+1}^{\infty}\frac{|z|^n\cdot|B_T|^n}{n!} \leq e^{|z||B_T|} \in L^1(P).$$

Using the previous observation we may now conclude

$$\begin{aligned}\mathbb{E}\left[e^{z\cdot B_T}\Phi(B_T)Y_T\right] &= \lim_{K\to\infty}\mathbb{E}\left[\sum_{n=0}^{K}\frac{(z\cdot B_T)^n}{n!}\Phi(B_T)Y_T\right] \\ &= \lim_{K\to\infty}\mathbb{E}\left[\sum_{n=0}^{K}\frac{(z\cdot B_T)^n}{n!}\Phi(B_T)e^{-\frac{1}{2}\int_0^T|AB_t|^2dt}\right] \\ &= \mathbb{E}\left[e^{z\cdot B_T}\Phi(B_T)e^{-\frac{1}{2}\int_0^T|AB_t|^2dt}\right]\end{aligned}$$

holds for all $z \in \mathbb{C}^N$. Restricting z to be in $i\mathbb{R}^N$, we may apply Dynkin's multiplicative system theorem [19, Appendix A, p. 309] in order to show

$$\mathbb{E}\left[u(B_T)\Phi(B_T)Y_T\right] = \mathbb{E}\left[u(B_T)\Phi(B_T)e^{-\frac{1}{2}\int_0^T|AB_t|^2dt}\right]$$

for all bounded measurable functions u on $\mathbb{R}^N$. Given $f : \mathbb{R}^N \to \mathbb{C}$ such that $\mathbb{E}|f(B_T)| < \infty$ and $m \in \mathbb{N}$, apply the previous formula with $u(x) = u_m(x) = \Phi^{-1}f\cdot 1_{|\Phi^{-1}f|\leq m}$ to learn

$$\mathbb{E}\left[f(B_T)\cdot 1_{|\Phi^{-1}f|(B_T)\leq m}Y_T\right] = \mathbb{E}\left[f(B_T)\cdot 1_{|\Phi^{-1}f|(B_T)\leq m}e^{-\frac{1}{2}\int_0^T|AB_t|^2dt}\right].$$

Now use the dominated convergence theorem to let $m \to \infty$ and finish the proof.

Acknowledgments

The authors would like to thank Clinton Conley and Leonard Gross for helpful discussions, and Fabrice Baudoin for pointing us to the works of M. Yor and others cited in Section 2.

References

[1] Andrei Agrachev, Ugo Boscain, Jean-Paul Gauthier, and Francesco Rossi, *The intrinsic hypoelliptic Laplacian and its heat kernel on unimodular Lie groups*, J. Funct. Anal. **256** (2009), no. 8, 2621–2655, DOI 10.1016/j.jfa.2009.01.006. MR2502528 (2010c:58042)

[2] Martín Argerami, *Approximating a Hilbert-Schmidt operator*, Mathematics Stack Exchange, URL:http://math.stackexchange.com/q/373117 (version: 2013-04-26).

[3] Fabrice Baudoin, Michel Bonnefont, and Nicola Garofalo, *A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality*, Math. Ann. **358** (2014), no. 3-4, 833–860, DOI 10.1007/s00208-013-0961-y. MR3175142

[4] Fabrice Baudoin and Nicola Garofalo, *Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries*, ArXiv e-prints (2011).

[5] Fabrice Baudoin, Maria Gordina, and Tai Melcher, *Quasi-invariance for heat kernel measures on sub-Riemannian infinite-dimensional Heisenberg groups*, Trans. Amer. Math. Soc. **365** (2013), no. 8, 4313–4350, DOI 10.1090/S0002-9947-2012-05778-3. MR3055697

[6] Fabrice Baudoin and Josef Teichmann, *Hypoellipticity in infinite dimensions and an application in interest rate theory*, Ann. Appl. Probab. **15** (2005), no. 3, 1765–1777, DOI 10.1214/105051605000000214. MR2152244 (2006g:60080)

[7] Vladimir I. Bogachev, *Gaussian measures*, Mathematical Surveys and Monographs, vol. 62, American Mathematical Society, Providence, RI, 1998. MR1642391 (2000a:60004)

[8] Chin-Huei Chang, Der-Chen Chang, Peter Greiner, and Hsuan-Pei Lee, *The positivity of the heat kernel on Heisenberg group*, Anal. Appl. (Singap.) **11** (2013), no. 5, 1350019, 15, DOI 10.1142/S021953051350019X. MR3104104

[9] Daniel Dobbs and Tai Melcher, *Smoothness of heat kernel measures on infinite-dimensional Heisenberg-like groups*, J. Funct. Anal. **264** (2013), no. 9, 2206–2223, DOI 10.1016/j.jfa.2013.02.013. MR3029152

[10] Bruce K. Driver, *Integration by parts and quasi-invariance for heat kernel measures on loop groups*, J. Funct. Anal. **149** (1997), no. 2, 470–547, DOI 10.1006/jfan.1997.3103. MR1472366 (99a:60054a)

[11] Bruce K. Driver, *Heat kernels measures and infinite dimensional analysis*, Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), Contemp. Math., vol. 338, Amer. Math. Soc., Providence, RI, 2003, pp. 101–141, DOI 10.1090/conm/338/06072. MR2039953 (2005e:58058)

[12] Bruce K. Driver and Maria Gordina, *Heat kernel analysis on infinite-dimensional Heisenberg groups*, J. Funct. Anal. **255** (2008), no. 9, 2395–2461, DOI 10.1016/j.jfa.2008.06.021. MR2473262 (2010f:60010)

[13] Bruce K. Driver and Maria Gordina, *Integrated Harnack inequalities on Lie groups*, J. Differential Geom. **83** (2009), no. 3, 501–550. MR2581356 (2011d:58086)

[14] Bruce K. Driver and Maria Gordina, *Square integrable holomorphic functions on infinite-dimensional Heisenberg type groups*, Probab. Theory Related Fields **147** (2010), no. 3-4, 481–528, DOI 10.1007/s00440-009-0213-y. MR2639713 (2011k:58052)

[15] Nathaniel Eldredge, *Precise estimates for the subelliptic heat kernel on H-type groups* (English, with English and French summaries), J. Math. Pures Appl. (9) **92** (2009), no. 1, 52–85, DOI 10.1016/j.matpur.2009.04.011. MR2541147 (2010h:35399)

[16] Xavier Fernique, *Intégrabilité des vecteurs gaussiens* (French), C. R. Acad. Sci. Paris Sér. A-B **270** (1970), A1698–A1699. MR0266263 (42 #1170)

[17] Bernard Gaveau, *Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents*, Acta Math. **139** (1977), no. 1-2, 95–153. MR0461589 (57 #1574)

[18] Lars Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171. MR0222474 (36 #5526)

[19] Svante Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997. MR1474726 (99f:60082)

[20] James Kuelbs and Wenbo Li, *A functional LIL for stochastic integrals and the Lévy area process*, J. Theoret. Probab. **18** (2005), no. 2, 261–290, DOI 10.1007/s10959-003-2604-9. MR2137043 (2006c:60040)

[21] Hui Hsiung Kuo, *Gaussian measures in Banach spaces*, Lecture Notes in Mathematics, Vol. 463, Springer-Verlag, Berlin-New York, 1975. MR0461643 (57 #1628)

[22] Shigeo Kusuoka and Daniel W. Stroock, *Precise asymptotics of certain Wiener functionals*, J. Funct. Anal. **99** (1991), no. 1, 1–74, DOI 10.1016/0022-1236(91)90051-6. MR1120913 (93a:60085)

[23] Michel Ledoux, *A note on large deviations for Wiener chaos*, Séminaire de Probabilités, XXIV, 1988/89, Lecture Notes in Math., vol. 1426, Springer, Berlin, 1990, pp. 1–14, DOI 10.1007/BFb0083753. MR1071528 (91i:60017)

[24] Paul Lévy, *Le mouvement brownien plan* (French), Amer. J. Math. **62** (1940), 487–550. MR0002734 (2,107g)

[25] Paul Lévy, *Wiener's random function, and other Laplacian random functions*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, University of California Press, Berkeley and Los Angeles, 1951, pp. 171–187. MR0044774 (13,476b)

[26] Wenbo V. Li, *Small ball probabilities for Gaussian Markov processes under the L_p-norm*, Stochastic Process. Appl. **92** (2001), no. 1, 87–102, DOI 10.1016/S0304-4149(00)00072-7. MR1815180 (2001m:60090)

[27] George Lowther, *Does infinite-dimensional Brownian motion live in hyperplanes?*, MathOverflow, URL:http://mathoverflow.net/q/102963 (version: 2012-07-24).

[28] Françoise Lust-Piquard, *A simple-minded computation of heat kernels on Heisenberg groups*, Colloq. Math. **97** (2003), no. 2, 233–249, DOI 10.4064/cm97-2-9. MR2031850 (2005g:22006)

[29] Paul Malliavin, *Hypoellipticity in infinite dimensions*, Diffusion processes and related problems in analysis, Vol. I (Evanston, IL, 1989), Progr. Probab., vol. 22, Birkhäuser Boston, Boston, MA, 1990, pp. 17–31. MR1110154 (93b:60132)

[30] Roger Mansuy and Marc Yor, *Aspects of Brownian motion*, Universitext, Springer-Verlag, Berlin, 2008. MR2454984 (2010a:60278)

[31] Jonathan C. Mattingly and Étienne Pardoux, *Malliavin calculus for the stochastic 2D Navier-Stokes equation*, Comm. Pure Appl. Math. **59** (2006), no. 12, 1742–1790, DOI 10.1002/cpa.20136. MR2257860 (2007j:60082)

[32] Michel Métivier, *Semimartingales*, de Gruyter Studies in Mathematics, vol. 2, Walter de Gruyter & Co., Berlin-New York, 1982. A course on stochastic processes. MR688144 (84i:60002)

[33] Jennifer Randall, *The heat kernel for generalized Heisenberg groups*, J. Geom. Anal. **6** (1996), no. 2, 287–316, DOI 10.1007/BF02921603. MR1469125 (99b:22018)

[34] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992. MR1218884 (95f:43008)

[35] Marc Yor, *Les filtrations de certaines martingales du mouvement brownien dans* $\mathbf{R}^n$ (French), Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78), Lecture Notes in Math., vol. 721, Springer, Berlin, 1979, pp. 427–440. MR544812 (81b:60045)

[36] Marc Yor, *Remarques sur une formule de Paul Lévy*, Seminar on Probability, XIV (Paris, 1978/1979) (French), Lecture Notes in Math., vol. 784, Springer, Berlin, 1980, pp. 343–346. MR580140 (82c:60144)

Department of Mathematics, University of California, San Diego, La Jolla, California 92093

E-mail address: bdriver@math.ucsd.edu

Department of Mathematics, Cornell University, Ithaca, New York 14853 – and – School of Mathematical Sciences, University of Northern Colorado, Greeley, Colorado 80639

E-mail address: neldredge@unco.edu

Department of Mathematics, University of Virginia, Charlottesville, Virginia 22904

E-mail address: melcher@virginia.edu