J. Math. Soc. Japan
Vol. 67, No. 4 (2015) pp. 1785–1799
doi: 10.2969/jmsj/06741785

An exercise in Malliavin's calculus

By Daniel W. Stroock

Dedicated to the memory of Professor Kiyosi Itô

(Received Jan. 8, 2014)
(Revised June 8, 2014)

Abstract. This note has two goals. First, for those who have heard the term but do not know what it means, it provides a gentle introduction to Malliavin's calculus as it applies to degenerate parabolic partial differential equations. Second, it applies that theory to generalizations of Kolmogorov's example of a highly degenerate operator which is nonetheless hypoelliptic.

1. Background.

The regularity theory for the fundamental solution to the Cauchy intitial value problem of parabolic equations of the form

$$\partial_t u = \frac{1}{2}\sum_{i,j=1}^N a_{ij}\partial_{x_i}\partial_{x_j}u + \sum_{i=1}^N b_i\partial_{x_i}u \text{ on } (0,\infty)\times\mathbb{R}^N \quad \text{with } u(0,\,\cdot\,) = f \tag{1.1}$$

has been well developed. Assuming that the second order coefficients are continuous and non-degenerate in the sense that

$$\sum_{i,j=1}^N a_{ij}\xi_i\xi_j > 0 \quad \text{for } \xi = (\xi_1,\dots,\xi_N) \neq 0, \tag{1.2}$$

one knows (cf. Theorem 9.2.6 in [**10**]) that fundamental solution $p(t,x,y)$ to an appropriately generalized version of (1.1) exists and, under mild growth conditions, is integrable to all orders. When the coefficients are Hölder continuous and satisfy (1.2), $p(t,x,y)$ has two Hölder continuous spacial derivatives and gives classical solutions to (1.1), and, when the coefficients are smooth and satisfy (1.2), $p(t,x,y)$ is smooth. (See [**1**] for the Hölder continuous case and [**9**] for the smooth case.) The traditional approaches to studying solutions to (1.1) involved one or another form of perturbation theory and relied heavily on the non-degeneracy condition in (1.2). Outside the traditional framework was an example found by Kolmogorov of a highly degenerate equation for which the fundamental solution exists and is smooth. Namely, what Kolmogorov noticed is that

2010 *Mathematics Subject Classification.* Primary 35H10, 35J65, 60J65.
Key Words and Phrases. hypoelliptic equations.

$$p(t,x,y) = \frac{1}{2\pi\sqrt{\det(C(t))}} \exp\left(-\frac{1}{2}\big(y - m(t,x), C(t)^{-1}(y - m(t,x))\big)\right),$$
$$\text{where } C(t) = \begin{pmatrix} t & \frac{t^2}{2} \\ \frac{t^2}{2} & \frac{t^3}{3} \end{pmatrix} \text{ and } m(t,x) = \begin{pmatrix} x_1 \\ x_2 + tx_1 \end{pmatrix}$$

is a fundamental solution to highly degenerate equation

$$\partial_t u = \frac{1}{2}\partial_{x_1}^2 u + x_1 \partial_{x_2} u.$$

The easiest way to understand how he discovered this solution is to use probability theory. Namely, if $\{B(t) : t \geq 0\}$ is a standard, 1-dimensional Brownian motion, then the density of the distribution of

$$X(t,x) \equiv m(t,x) + \begin{pmatrix} B(t) \\ \int_0^t B(\tau)\, d\tau \end{pmatrix}$$

will be a fundamental solution. Furthermore, because all elements in the span of $\{B(t) : t \geq 0\}$ are centered (i.e., mean 0) Gaussian and $\mathbb{E}[B(s)B(t)] = s \wedge t$, $X(t,x)$ is an $\mathbb{R}^2$-valued, Gaussian random variable with mean value $m(t,x)$ and covariance $C(t)$.

Kolmogorov's example stood in isolation until 1967, when Hörmander [**2**] proved a general theorem that put it in context. To state Hörmander's result, one first has to rewrite[1] (1.1) in the form

$$\partial_t u = \frac{1}{2}\sum_{k=1}^{r} V_k^2 u + V_0 u,$$

where the V_k are vector fields. After writing (1.1) in this form, Hörmander's theorem shows that a smooth fundamental solution will exist if the Lie algebra generated by

$$\big\{V_1, \ldots, V_r, [V_0, V_1], \ldots, [V_0, V_r]\big\}$$

has rank N at each point. When $\{V_1, \ldots, V_r\}$ already generate a Lie algebra of rank N and $V_0 = 0$, Rothschield and Stein [**7**] gave a beautiful geometric interpretation of these equations. However in many ways, Hörmander's theorem is most intriguing and difficult when, as in Kolmogorov's example, V_0 plays an essential role.

Although Kolmogorov used probability theory to find his example, it was not until Malliavin's groundbreaking work [**5**] that anyone was able to use probability theory to understand the theory for even non-degenerate cases, much less for degenerate ones. Malliavin's program was carried to completion by Kusuoka and me in Section 8 of [**4**], where Hörmander's result is proved using Malliavin's ideas starting from Itô's representa-

[1] To do this, one has to find a smooth square root of the coefficient matrix $((a_{i,j}))_{1\leq i,j\leq N}$, and, in the degenerate case, such a square root need not exist. Based on ideas of J. J. Kohn, Olenik and Radekevich [**6**] later extended Hörmander's result to remove the need for a smooth square root.

tion of the diffusion process for which (1.1) is the backward equation. Just like the other approaches, the case when V_0 is needed gave us the greatest trouble, and our treatment of it relies on a trick that is rather unsatisfying. For this reason, I was pleased when I realized that natural extensions of Kolmogorov's example can be handled by relatively elementary and transparent techniques using Malliavin's ideas. Specifically, in this note, I will prove that the initial value problem for equations of the form

$$\partial_t u = \frac{1}{2}\partial_1^2 u + b(x_1)\partial_2 \tag{1.3}$$

admits a fundamental solution when b is a smooth function whose derivatives have at most polynomial growth and for which there exist a $\kappa > 0$, $\beta \in [0, 2)$, and $\delta \in (0, 1]$ such that

$$|b'(\eta)| \geq \kappa e^{-|\eta-\xi|^{-\beta}} \quad \text{for } (\xi, \eta) \in \mathbb{R}^2 \text{ with } |\eta - \xi| \leq 2\delta. \tag{1.4}$$

Although I will not prove that the solution is more than continuous, those familiar with Malliavin's ideas will be able to show from the results here that the solution is smooth. In addition, if, for some $\beta \in [0, \infty)$ and $\delta \in (0, 1]$,

$$|b'(\eta)| \geq \exp\bigl(-(\log^+ |\eta - \xi|^{-1})^\beta\bigr) \quad \text{for } (\xi, \eta) \in \mathbb{R}^2 \text{ with } |\eta - \xi| \leq 2\delta, \tag{1.5}$$

I will argue that the operator

$$\partial_t - \frac{1}{2}\partial_1^2 - b(x_1)\partial_2$$

is hypoelliptic. It may be of some interest to observe that (1.4) and even (1.5) can hold even though Hörmander's condition does not.

2. Elements of Malliavin's calculus.

Let

$$\Omega = \left\{\omega \in C([0, \infty); \mathbb{R}) : \omega(0) = 0 = \lim_{t\to\infty} \frac{|\omega(t)|}{1 + t}\right\},$$

and turn Ω into the separable Banach space with norm

$$\|\omega\|_\Omega \equiv \sup_{t\geq 0} \frac{|\omega(t)|}{1 + t}.$$

Next, set

$$H = \{h \in \Omega : h \text{ is absolutely continuous and } \dot{h} \in L^2([0, \infty); \mathbb{R})\},$$

and make H into a separable Hilbert space with inner product

$$(h_1, h_2)_H \equiv (\dot{h}_1, \dot{h}_2)_{L^2([0,\infty);\mathbb{R})}.$$

If $h \in H$ is smooth and $\dot{h}$ has compact support, define $\mathcal{I}(h) : \Omega \longrightarrow \mathbb{R}$ by

$$[\mathcal{I}(h)](\omega) = \int_0^\infty \dot{h}(t)\, d\omega(t), \tag{$*$}$$

where the integral is taken in the sense of Riemann–Stieltjes. Then Wiener measure $\mathcal{W}$ is the unique Borel measure on Ω with the property that

$$\mathbb{E}^{\mathcal{W}}\big[e^{i\mathcal{I}(h)}\big] = e^{-(1/2)\|h\|_H^2} \tag{2.1}$$

for all smooth $h \in H$ such that $\dot{h}$ has compact support. In particular, $\mathcal{I}(h)$ is a centered, $\mathbb{R}$-valued, Gaussian random variable with variance $\|h\|_H^2$, and so there exists a unique isometric extension of $\mathcal{I}$ as map from H into $L^2(\mathcal{W};\mathbb{R})$, and (2.1) continues to hold for this extension. From this one sees that, for any $h \in H$, the distribution of $\omega \longmapsto \omega + h$ under $\mathcal{W}$ is absolutely continuous with respect to $\mathcal{W}$ and has Radon–Nikodym derivative given by

$$R_h = \exp\left(\mathcal{I}(h) - \frac{\|h\|_H^2}{2}\right).$$

Given $h \in H$ and a smooth[2] function $F : \Omega \longrightarrow \mathbb{R}$, define $D_h F : \Omega \longrightarrow \mathbb{R}$ by

$$D_h F(\omega) = \frac{d}{dt} F(\omega + th)\bigg|_{t=0}.$$

Then, from the preceding, it is easy to derive the integration by parts formula

$$\mathbb{E}^{\mathcal{W}}[\Psi D_h \Phi] = \mathbb{E}^{\mathcal{W}}[\mathcal{I}(h)\Phi\Psi - \Phi D_h \Psi] \tag{2.2}$$

for smooth functions Φ and Ψ satisfying reasonable bounds.

To see how (2.2) gets applied in the proof of regularity results, let $n \geq 0$ and consider the Itô representation

$$X(t, x, \omega) = \begin{pmatrix} x_1 + \omega(t) \\ x_2 + \int_0^t \big(x_1 + \omega(\tau)\big)^{2n+1}\, d\tau \end{pmatrix}$$

of the diffusion whose backward equation is

[2]In general, one has to consider functions on Ω which are differentiable in the sense of Sobolev but not classically. However, for the application here, we need only deal with functions that have classical Fréchet derivatives.

$$\partial_t u = \frac{1}{2}\partial_1^2 u + x^{2n+1}\partial_2 u.$$

Then, for any $\varphi \in C_c^\infty(\mathbb{R}^2;\mathbb{R})$,

$$D_h\big(\varphi \circ X(t,x)\big) = h(t)\partial_1\varphi \circ X(t,x) + (2n+1)\int_0^t h(\tau)(x_1+\omega(\tau))\big)^{2n}\,d\tau\partial_2\varphi \circ X(t,x).$$

Now set $h_1(\tau) = (3(\tau\wedge t)^2/t^2) - (2(\tau\wedge t)/t)$ and $h_2(\tau) = \tau(t-(\tau\wedge t))$. Then

$$D_{h_2}\big(\varphi \circ X(t,x)\big) = (2n+1)\int_0^t h_2(\tau)\big(x_1+\omega(\tau)\big)^{2n}\,d\tau\partial_{x_2}\varphi \circ X(t,x)$$

and

$$\begin{aligned} D_{h_1}\big(\varphi \circ X(t,x)\big) &= \partial_1\varphi \circ X(t,x) + (2n+1)\int_0^t h_1(\tau)\big(x_1+\omega(\tau)\big)^{2n}\,d\tau\partial_{x_2}\varphi \circ X(t,x) \\ &= \partial_{x_1}\varphi \circ X(t,x) + \frac{\int_0^t h_1(\tau)(x_1+\omega(\tau))^{2n}\,d\tau}{\int_0^t h_2(\tau)(x_1+\omega(\tau))^{2n}\,d\tau} D_{h_2}\big(\varphi \circ X(t,x)\big). \end{aligned}$$

Hence

$$\partial_{x_1}\varphi \circ X(t,x) = D_{h_1}\big(\varphi \circ X(t,x)\big) - \frac{\Phi_1}{\Phi_2}D_{h_2}\big(\varphi \circ X(t,x)\big)$$

and

$$\partial_{x_2}\varphi \circ X(t,x) = \frac{1}{(2n+1)\Phi_2}D_{h_2}\big(\varphi \circ X(t,x)\big),$$

where

$$\Phi_1 = \int_0^t h_1(\tau)\big(x_1+\omega(\tau)\big)^{2n}\,d\tau \quad\text{and}\quad \Phi_2 = \int_0^t h_2(\tau)\big(x_1+\omega(\tau)\big)^{2n}\,d\tau.$$

After applying (1.2), we find that

$$\mathbb{E}^{\mathcal{W}}\big[\partial_{x_1}\varphi \circ X(t,x)\big] = \mathbb{E}^{\mathcal{W}}\left[\left(\mathcal{I}(h_1) - \frac{\mathcal{I}(h_2)\Phi_1}{\Phi_2} + D_{h_2}\frac{\Phi_1}{\Phi_2}\right)\varphi \circ X(t,x)\right]$$

and

$$\mathbb{E}^{\mathcal{W}}\big[\partial_{x_2}\varphi \circ X(t,x)\big] = \mathbb{E}^{\mathcal{W}}\left[\left(\mathcal{I}(h_2) - D_{h_2}\frac{1}{\Phi_2}\right)\varphi \circ X(t,x)\right].$$

Therefore, if $\psi_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ and $\psi_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ are determined by

$$\psi_1 \circ X(t,x) = \mathbb{E}^{\mathcal{W}}\left[\mathcal{I}(h_1) - \frac{\mathcal{I}(h_2)\Phi_1}{\Phi_2} + D_{h_2}\frac{\Phi_1}{\Phi_2}\,\middle|\, X(t,x)\right]$$

and

$$\psi_2 \circ X(t,x) = \mathbb{E}^{\mathcal{W}}\left[\mathcal{I}(h_2) - D_{h_2}\frac{1}{\Phi_2}\,\middle|\, X(t,x)\right],$$

and if $P(t,x,\cdot\,)$ denotes the distribution of $X(t,x)$ under $\mathcal{W}$, then

$$\int_{\mathbb{R}^2} \partial_{x_j}\varphi(y)\, P(t,x,dy)\mu = \int_{\mathbb{R}^2} \varphi(y)\psi_j(y)\, P(t,x,dy) \text{ for } j \in \{1,2\}. \tag{2.3}$$

Given (2.3), the proof that $P(t,x,\cdot\,)$ admits a continuous density relies on the following simple variant of Sobolev's embedding theorem. This result is Lemma 1.26 in [**3**] and is included here for the convenience of the reader.

LEMMA 1. *Let μ be a probability measure on $\mathbb{R}^N$ and $p \in (N,\infty)$. If, for each $1 \le j \le N$, there is a $\psi_j \in L^p(\mu;\mathbb{R})$ such that $\partial_{x_j}\mu = \psi_j\mu$, then there is an $f \in C_b(\mathbb{R}^N;[0,\infty))$ such that $\mu(dy) = f(y)\,dy$. In fact, there is a $C < \infty$, depending only on N and p, such that*

$$\|f\|_u \le C\Big(\max_{1\le j\le N} \|\psi_j\|_{L^p(\mu;\mathbb{R})}\Big)^N.$$

PROOF. For $\lambda > 0$, set $r_\lambda(x) = (4\pi)^{-N/2}\int_0^\infty t^{-N/2}e^{-\lambda t}e^{-|x|^2/4t}\,dt$. Then

$$r_\lambda(x) = \lambda^{(N/2)-1}r_1(\lambda^{1/2}x) \text{ and so } \nabla r_\lambda(x) = \lambda^{(N/2)-(1/2)}\nabla r_1(\lambda^{1/2}x).$$

Hence

$$\|\lambda r_\lambda\|_{L^q(\mathbb{R}^N;\mathbb{R})} = \lambda^{N/2q'}\|r_1\|_{L^q(\mathbb{R}^N;\mathbb{R})} \quad \text{and}$$
$$\|\nabla r_\lambda\|_{L^q(\mathbb{R}^N;\mathbb{R})} = \lambda^{(N/2q')-(1/2)}\|\nabla r_1\|_{L^q(\mathbb{R}^N;\mathbb{R})},$$

where $q' = q/(q-1)$ is the Hölder conjugate of q. When $N = 1$, r_1 is bounded, $\|r_1\|_{L^1(\mathbb{R}^N;\mathbb{R})} = 1$, and so $\|r_1\|_{L^q(\mathbb{R}^N;\mathbb{R})} < \infty$ for all $q \in [1,\infty]$. When $N \ge 2$, and $p > N/2$, choose $\theta \in (0,1)$ so that $p = (N - 2\theta + 2)/2$. Then

$$r_1(x)^{p'} = (4\pi)^{-Np'/2}\left(\int_0^\infty t^{-(N/2)+\theta}e^{-|x|^2/4t}t^{-\theta}e^{-t}\,dt\right)^{p'}$$
$$\le (4\pi)^{-Np'/2}\Gamma(1-\theta)^{p'-1}\int_0^\infty t^{-((N-2\theta)p'/2)-\theta}e^{-t}e^{-|x|^2p'/4t}\,dt,$$

and so there is an $A(p,\theta) < \infty$ such that

$$\|r_1\|_{L^{p'}(\mathbb{R}^N;\mathbb{R})}^{p'} \le A(p,\theta)^{p'} \int_0^\infty t^{-(p'-1)((N/2)-\theta)} e^{-t}\,dt < \infty,$$

since $(p'-1)((N/2)-\theta) = (N-2\theta)/(N-2\theta+2) < 1$. Hence, $\|r_1\|_{L^{p'}(\mathbb{R}^N;\mathbb{R})} < \infty$ for all $p > N/2$. Similarly, $\|\nabla r_1\|_{L^{p'}(\mathbb{R}^N;\mathbb{R})} < \infty$ for all $p > N$.

Next observe that $\mu = \mu * r_1 + \sum_{j=1}^N (\psi_j \mu) * (\partial_{x_1} r_1)$, and therefore that $\mu = f\lambda_{\mathbb{R}}^N$ for some non-negative $f \in L^1(\mathbb{R}^N;\mathbb{R})$ with integral 1. Set $g = f^{1/p}$, and note that $\partial_{x_j} g = (1/p)\psi_j g$. Since $\|g\|_{L^p(\mathbb{R}^N;\mathbb{R})} = 1$ and $\|\partial_{x_j} g\|_{L^p(\mathbb{R}^N;\mathbb{R})} = \|\psi_j\|_{L^p(\mu;\mathbb{R})}^{1/p}$, for any $\lambda > 0$, $g = \lambda g * r_\lambda + \sum_{j=1}^N (\psi_j g) * (\partial_{x_j} r_\lambda)$, and so g is continuous and

$$\|g\|_{\mathrm{u}} \le C(p,N)\left(\lambda^{N/2p} + \lambda^{(N/2p)-(1/2)} \sum_{j=1}^N \|\psi_j\|_{L^p(\mu;\mathbb{R})}^{1/p}\right)$$

for some $C(p,N) < \infty$. Now minimize with respect to λ and use the fact that $\|f\|_{\mathrm{u}} = \|g\|_{\mathrm{u}}^p$. □

In view of Lemma 1, we will know that $P(t,x,\cdot\,)$ admits a continuous density once we know that the ψ_1 and ψ_2 in (1.3) belong to $L^p(\mu;\mathbb{R})$ for all $p \in [1,\infty)$, and, because conditioning is a L^p-contraction, this comes down to showing that

$$\mathcal{I}(h_1) - \frac{\mathcal{I}(h_2)\Phi_1}{\Phi_2} + D_{h_2}\frac{\Phi_1}{\Phi_2} \quad \text{and} \quad \mathcal{I}(h_2) - D_{h_2}\frac{1}{\Phi_2}$$

are in $L^p(\mathcal{W};\mathbb{R})$ for all $p \in [1,\infty)$. Clearly, the only terms in doubt are

$$\begin{aligned} D_{h_2}\frac{\Phi_1}{\Phi_2} = {} & \frac{2n\int_0^t h_1(\tau)h_2(\tau)(x_1+\omega(\tau))^{2n-1}\,d\tau}{\int_0^t h_2(\tau)(x_1+\omega(\tau))^{2n}\,d\tau} \\ & - \frac{2n\int_0^t h_2(\tau)^2(x_1+\omega(\tau))^{2n-1}\,d\tau \int_0^t h_1(\tau)(x_1+\omega(\tau))^{2n}\,d\tau}{\left(\int_0^t h_2(\tau)(x_1+\omega(\tau))^{2n}\,d\tau\right)^2} \end{aligned}$$

and

$$D_{h_2}\frac{1}{\Phi_2} = -\frac{2n\int_0^t h_2(\tau)^2(x_1+\omega(\tau))^{2n-1}\,d\tau}{\left(\int_0^t h_2(\tau)(x_1+\omega(\tau))^{2n}\,d\tau\right)^2}.$$

Hence, all that we have to check is that

$$\mathbb{E}^{\mathcal{W}}\left[\left(\int_0^t h_2(\tau)(x_1+\omega(\tau))^{2n}\,d\tau\right)^{-p}\right] < \infty \quad \text{for all } p \in [1,\infty).$$

To this end, first observe that

$$\begin{aligned}\int_0^t \tau(t-\tau)\big(x_1+\omega(\tau)\big)^{2n}\,d\tau &= t^3\int_0^1 \tau(1-\tau)\big(x_1+\omega(t\tau)\big)^{2n}\,d\tau\\ &\ge 6^{n-1}t^3\bigg(\int_0^1 \tau(1-\tau)\big(x_1+\omega(t\tau)\big)^2\,d\tau\bigg)^n\\ &\ge \frac{3\cdot 6^{n-1}t^3}{16}\bigg(\int_{1/4}^{3/4}\big(x_1+\omega(t\tau)\big)^2\,d\tau\bigg)^n.\end{aligned}$$

By Brownian scaling, $\int_{1/4}^{3/4}(x_1+\omega(t\tau))^2\,d\tau$ and $t\int_{1/4}^{3/4}(t^{-1/2}x_1+\omega(\tau))^2\,d\tau$ have the same distribution under $\mathcal{W}$, and, by standard Brownian computations (cf. Section 10.3.36 in [**8**]), for all $\alpha>0$,

$$\begin{aligned}&\mathbb{E}^{\mathcal{W}}\bigg[\exp\bigg(-\alpha\int_{1/4}^{3/4}\big(t^{-1/2}x_1+\omega(\tau)\big)^2\,d\tau\bigg)\bigg]\\ &\quad=\int_{\mathbb{R}}\mathbb{E}^{\mathcal{W}}\bigg[\exp\bigg(-\alpha\int_0^{1/2}\big(\xi+\omega(\tau)\big)^2\,d\tau\bigg)\bigg]\gamma_{t^{-1/2}x_1,1/4}(d\xi)\le\big(\cosh\sqrt{2\alpha}\big)^{-1/2},\end{aligned}$$

where γ_{m,σ^2} is the Gaussian measure on $\mathbb{R}$ with mean m and variance σ^2. Starting from this, we know that

$$\mathcal{W}\bigg(\Delta\le\frac{1}{R}\bigg)\le e\mathbb{E}^{\mathcal{W}}\big[e^{-R\Delta}\big]\le e\big(\cosh\sqrt{2R}\big)^{-1/2},$$

and therefore that

$$\lim_{R\to\infty}R^M\mathcal{W}\bigg(\int_0^t h_2(\tau)\big(x_1+\omega(\tau)\big)^{2n}\,d\tau\le\frac{1}{R}\bigg)=0\quad\text{for all } M\ge 0.$$

3. The general case.

Because $b'\ge 0$ in the cases treated in Section 1, we were able to handle them by differentiating in only two directions, h_1 and h_2. However to handle the general case, or, for that matter, even the case $b(\xi)=\xi^2$, we will need to differentiate in an infinite number of directions.

Let $\{h_n : n\ge 1\}\subseteq C_{\mathrm{c}}^\infty\big([0,\infty);\mathbb{R}\big)$ be an orthonormal basis in H, and if $\Phi:\Omega\longrightarrow\mathbb{R}$ is a smooth function for which

$$\sum_{n=1}^\infty |D_{h_n}\Phi(\omega)|^2<\infty\quad\text{for all }\omega\in\Omega,$$

define $D\Phi:\Omega\longrightarrow H$ so that $(D\Phi(\omega),h)_H=D_h\Phi(\omega)$ for all $\omega\in\Omega$ and $h\in H$.

The Itô representation for the diffusion whose backward equation is (1.3) is

$$X(t,x) = \begin{pmatrix} x_1 + \omega(t) \\ x_2 + \int_0^t b\big(x_1 + \omega(\tau)\big)\,d\tau \end{pmatrix}. \tag{3.1}$$

Thus, what we want to show is that the distribution $P(t, x, \cdot\,)$ of $X(t, x)$ under $\mathcal{W}$ admits a bounded, continuous density. By trivial rescaling and translation, it is easy to see that it suffices to treat the case when $t = 1$ and $x = 0$. Therefore, set $X = X(1, 0)$. Clearly,

$$D_h X = \begin{pmatrix} h(1) \\ \int_0^1 h(\tau) b'\big(\omega(\tau)\big)\,d\tau \end{pmatrix},$$

and so

$$[DX(\omega)](t) \equiv \begin{pmatrix} [DX_1(\omega)](t) \\ [DX_2(\omega)](t) \end{pmatrix} = \begin{pmatrix} t \wedge 1 \\ \int_0^{t\wedge 1} \big(\int_s^1 b'(\omega(\tau))\,d\tau\big) ds \end{pmatrix}.$$

Now let $\varphi \in C_c^\infty(\mathbb{R}^2; \mathbb{R})$ be given. Then

$$D(\varphi \circ X) = (\partial_{x_1}\varphi \circ X) DX_1 + (\partial_{x_2}\varphi \circ X) DX_2.$$

Hence, if

$$A \equiv \big(((DX)_i, DX_j)_H\big)_{1\le i,j\le 2} = \begin{pmatrix} 1 & \int_0^1 \big(\int_s^1 b'\big(\omega(\tau)\big)\,d\tau\big)\,ds \\ \int_0^1 \big(\int_s^1 b'\big(\omega(\tau)\big)\,d\tau\big)\,ds & \int_0^1 \big(\int_s^1 b'\big(\omega(\tau)\big)\,d\tau\big)^2\,ds \end{pmatrix},$$

then

$$\big(D(\varphi \circ X), DX_i\big)_H = \sum_{j=1}^{2} A_{ij} \partial_{x_j}\varphi \circ X \quad \text{for } i \in \{1, 2\}. \tag{3.2}$$

In order to take the next step, we will need to know that the matrix A is $\mathcal{W}$-almost surely invertible. For this purpose, set

$$\begin{aligned} \Delta \equiv \det(A) &= \int_0^1 \left(\int_s^1 b'\big(\omega(\tau)\big)^2\,d\tau \right) ds - \left(\int_0^1 \left(\int_s^1 b'\big(\omega(\tau)\big)\,d\tau \right) ds \right)^2 \\ &= \iint_{0\le s<t\le 1} \left(\int_s^t b'\big(\omega(\tau)\big)\,d\tau \right)^2 ds dt. \end{aligned}$$

Lemma 2. *Referring to the preceding,* $\mathbb{E}^{\mathcal{W}}[\Delta^{-p}] < \infty$ *for all* $p \in [1, \infty)$.

Proof. Define $\{\sigma_n : n \ge 1\}$ and $\{\tau_n : n \ge 1\}$ by

$$\sigma_n(\omega) = 1 \wedge \inf\left\{ t \ge 0 : |\omega(t)| \ge \frac{\delta}{n} \right\} \text{ and}$$

$$\tau_n(\omega) = 1 \wedge \inf\left\{t \geq \sigma_n(\omega) : |\omega(t)| \leq \frac{\delta}{2n} \text{ or } |\omega(t)| \geq 2\delta\right\}.$$

Then, by (1.4), there is an $c > 0$ such that $\Delta \geq c(\tau_n - \sigma_n)^4 e^{-(2\delta^{-1}n)^\beta}$ for all $n \geq 1$, and therefore

$$\begin{aligned}
\mathbb{E}^{\mathcal{W}}\big[e^{-\alpha\Delta}\big] &\leq \mathbb{E}^{\mathcal{W}}\left[e^{-c\alpha e^{-(2\delta^{-1})^\beta}(\tau_1-\sigma_1)^4}, \sigma_1 \leq \frac{1}{2}\right] \\
&\quad + \sum_{n=2}^{\infty} \mathbb{E}^{\mathcal{W}}\left[e^{-c\alpha e^{-(2\delta^{-1}n)^\beta}(\tau_n-\sigma_n)^4}, \sigma_n \leq \frac{1}{2} < \sigma_{n-1}\right] \\
&\leq \mathbb{E}^{\mathcal{W}}\left[e^{-c\alpha e^{-(2\delta^{-1})^\beta}(\tau_1-\sigma_1)^4}, \sigma_1 \leq \frac{1}{2}\right] \\
&\quad + \sum_{n=2}^{\infty}\left(\mathbb{E}^{\mathcal{W}}\left[e^{-2c\alpha e^{-(2\delta^{-1}n)^\beta}(\tau_n-\sigma_n)^4}, \sigma_n \leq \frac{1}{2}\right]\right)^{1/2} \mathcal{W}\left(\sigma_{n-1} > \frac{1}{2}\right)^{1/2}.
\end{aligned}$$

Standard Brownian calculations (cf. Exercise 10.3.36 in [**8**]) show that

$$\mathcal{W}\left(\sigma_{n-1} > \frac{T}{2}\right)^{1/2} \leq Ce^{-\gamma T n^2} \quad \text{for some } C < \infty \text{ and } \gamma > 0.$$

Next, set $\zeta_r(\omega) = \inf\{t \geq 0 : \omega(t) \geq r\}$, and note that, by Brownian scaling, ζ_r has the same distribution under $\mathcal{W}$ as $r^2\zeta_1$. Hence, for any $\rho > 0$,

$$\begin{aligned}
&\mathbb{E}^{\mathcal{W}}\left[e^{-\rho(\tau_n-\sigma_n)^4}, \sigma_n \leq \frac{1}{2}\right] \\
&= \mathbb{E}^{\mathcal{W}}\left[e^{-\rho(\tau_n-\sigma_n)^4}, \sigma_n \leq \frac{1}{2} \;\&\; \tau_n > 1\right] + \mathbb{E}^{\mathcal{W}}\left[e^{-\rho(\tau_n-\sigma_n)^4}, \sigma_n \leq \frac{1}{2} \;\&\; \tau_n \leq 1\right] \\
&\leq e^{-\rho/16} + \mathbb{E}^{\mathcal{W}}\big[e^{-\rho(\zeta_{\delta/2n}\wedge\zeta_{(2-(1/n))\delta})^4}\big] \\
&\leq e^{-\rho/16} + \mathbb{E}^{\mathcal{W}}\big[e^{-\rho\zeta_{\delta/2n}}\big] + \mathbb{E}^{\mathcal{W}}\big[e^{-\rho\zeta^4_{(2-(1/n))\delta}}\big] \\
&\leq e^{-\rho/16} + 2\mathbb{E}^{\mathcal{W}}\big[e^{-(\rho\delta^2/4n^2)\zeta_1^4}\big].
\end{aligned}$$

At the same time, for $\lambda > 0$,

$$\begin{aligned}
\sqrt{\frac{\pi}{2}}\mathbb{E}\big[e^{-\lambda\zeta_1^4}\big] &= \int_0^\infty t^{-3/2}e^{-(1/2t)-\lambda t^4}\,dt \\
&\leq \int_0^{(4\lambda)^{-1/5}} t^{-3/2}e^{-1/2t}\,dt + e^{-4^{-4/5}\lambda^{1/5}}\int_{(4\lambda)^{-1/5}}^\infty t^{-3/2}e^{-1/2t}\,dt \\
&\leq 4\sqrt{\pi}e^{-4^{-4/5}\lambda^{1/5}}.
\end{aligned}$$

Thus, there exist a $C < \infty$ and an $\epsilon > 0$ such that

$$\mathbb{E}\big[e^{-\alpha\Delta}\big] \le C \sum_{n=1}^{\infty} \Big(\exp\big(-\epsilon\alpha n^{-8} e^{-(2\delta^{-1}n)^{\beta}}\big) + \exp\big(-\epsilon(\alpha n^{-8} e^{-(2\delta^{-1}n)^{\beta}})^{1/5}\big)\Big) e^{-\gamma n^2}.$$

By splitting the sum according to whether $n \le (\log(2^{-1}\delta\alpha^{1/2}))^{1/\beta}$ or $n > (\log(2^{-1}\delta\alpha)^{1/2})^{1/\beta}$, one sees that $\mathbb{E}[e^{-\alpha\Delta}]$ goes to 0 faster than α^{-p} for every $p \ge 1$ and therefore, since

$$\mathcal{W}\bigg(\Delta \le \frac{1}{R}\bigg) \le e\mathbb{E}^{\mathcal{W}}[e^{-R\Delta}],$$

that Δ^{-p} is integrable for all $p \in [1, \infty)$. □

For those adept with the techniques used in Malliavin's calculus, it is easy to pass from Lemma 2 to the conclusion that the distribution of X admits a smooth density. What follows is a somewhat formal summary of the steps that show the existence of a bounded, continuous density.

Knowing that A is $\mathcal{W}$-almost surely invertible, we can rewrite (3.2) as

$$\partial_{x_i}\varphi \circ X = \sum_{j=1}^{2} A_{ij}^{-1}(D\varphi \circ X, DX_j)_H.$$

By writing $(D(\varphi \circ X), DX_j)_H$ as $\sum_{n=1}^{\infty}(D_{h_n}(\varphi \circ X))(D_{h_n}X_j)$ and applying (1.2), we see that

$$\mathbb{E}^{\mathcal{W}}\big[A_{ij}^{-1}\big(D(\varphi \circ X), DX_j\big)_H\big] = \sum_{n=1}^{\infty} \mathbb{E}^{\mathcal{W}}\big[\big(A_{ij}^{-1}\mathcal{I}(h_n)D_{h_n}X_j - D_{h_n}(A_{ij}^{-1}D_{h_n}X_j)\big)\varphi \circ X\big].$$

If $\omega \in H$, one sees that

$$\sum_{n=1}^{\infty}[\mathcal{I}(h_n)]DX_j(h_n) = \big(\omega, DX_j\big)_H = \begin{cases} \omega(1) & \text{if } j = 1, \\ \int_0^1 \big(\int_t^1 b'\big(\omega(\tau)\big)\,d\tau\big)\,d\omega(t) & \text{if } j = 2, \end{cases}$$

and by a limiting procedure one can show the conclusion just drawn holds for $\mathcal{W}$-almost every $\omega \in \Omega$ when, for $j = 2$, the integral is interpreted in the sense of Riemann–Stieltjes. As for the other term, note that

$$D_{h_n}\big(A_{ij}^{-1}D_{h_n}X_j\big) = \big(D_{h_n}A_{ij}^{-1}\big)\big(D_{h_n}X_j\big) + A_{ij}^{-1}D_{h_n}^2X_j,$$

and therefore

$$\sum_{n=1}^{\infty} D_{h_n}\big(A_{ij}^{-1}D_{h_n}X_j\big) = \big(DA_{ij}^{-1}, DX_j\big)_H + A_{ij}^{-1}\sum_{n=1}^{\infty} D_{h_n}^2X_j.$$

Clearly, $D_{h_n}^2 X_1 = 0$, and, because $\sum_{n=1}^\infty h_n(s)^2 = \|h_s\|_H^2 = s$ where $h_s(\tau) = \tau \wedge s$,

$$\sum_{n=1}^\infty D_{h_n}^2 X_2 = \sum_{n=1}^\infty \int_0^1 h_n(s)^2 b''\big(\omega(s)\big)\,ds = \int_0^1 s b''\big(\omega(s)\big)\,ds.$$

Finally, for $i \neq j$,

$$\big(DA_{ii}^{-1}, DX_i\big)_H = \frac{1}{\Delta}\big(DA_{jj}, DX_i\big)_H - \frac{A_{jj}}{\Delta^2}\big(D\Delta, DX_i\big)_H$$

and

$$\big(DA_{ij}^{-1}, DX_j\big)_H = -\frac{1}{\Delta}\big(DA_{ij}, DX_j\big)_H + \frac{A_{ij}}{\Delta^2}\big(D\Delta, DX_j\big)_H.$$

Since $A_{11} = 1$, $(DA_{11}, DX_1)_H = 0$. Note that

$$A_{12} = A_{21} = \int_0^1 \tau b'\big(\omega(\tau)\big)\,d\tau,$$

and so

$$D_h A_{12} = D_h A_{21} = \int_0^1 \tau h(\tau) b''\big(\omega(\tau)\big)\,d\tau.$$

Hence, since $\sum_{n=1}^\infty h_n(s) h_n(t) = (h_s, h_t)_H = s \wedge t$,

$$(DA_{21}, DX_1)_H = \int_0^1 \tau^2 b''(\omega(\tau))\,d\tau$$

and

$$(DA_{12}, DX_2)_H = \iint_{[0,1]^2} (\tau_1 \wedge \tau_2)\tau_1 b''(\omega(\tau_1)) b'(\omega(\tau_2))\,d\tau_1 d\tau_2.$$

Finally,

$$\Delta = \iint_{[0,1]^2} (\tau_1 \vee \tau_2)(1 - \tau_1 \vee \tau_2) b'(\omega(\tau_1)) b'(\omega(\tau_2))\,d\tau_1 d\tau_2,$$

and so

$$D_h \Delta = 2 \iint_{[0,1]^2} (\tau_1 \wedge \tau_2)(1 - \tau_1 \vee \tau_2) h(\tau_1) b''(\omega(\tau_1)) b'(\omega(\tau_2))\,d\tau_1 d\tau_2,$$

from which it follows that

$$(D\Delta, DX_1) = 2 \iint\limits_{[0,1]^2} (\tau_1 \wedge \tau_2)(1 - \tau_1 \vee \tau_2)\tau_1 b''(\omega(\tau_1))b'(\omega(\tau_2))\, d\tau_1 d\tau_2$$

and

$$(D\Delta, DX_2) = 2 \iiint\limits_{[0,1]^3} (\tau_1 \wedge \tau_2)(1 - \tau_1 \vee \tau_2)(\tau_1 \wedge \tau_3) b''(\omega(\tau_1))b'(\omega(\tau_2))b'(\omega(\tau_3))\, d\tau_1 d\tau_2 d\tau_3.$$

Since all these are in $L^p(\mathcal{W};\mathbb{R})$ for all $p \in [1,\infty)$, we have now proved that $P(1,0,\,\cdot\,)$ admits a bounded, continuous density $p(1,0,\,\cdot\,)$, and, as we said above, from this it follows that $P(t,x,dy)$ admits a bounded, continuous density $p(t,x,\,\cdot\,)$ for all $(t,x) \in (0,\infty)\times\mathbb{R}^2$. Given Lemma 2, going further and proving that $(t,x,y) \rightsquigarrow p(t,x,y)$ is smooth requires essentially no new ideas and is basically a matter of book keeping. See [**4**] for details.

To prove that $\partial_t - (1/2)\partial_{x_1}^2 - b(x_1)\partial_{x_2}$ is hypoelliptic when (1.5) holds, set (cf. (3.1)) $X(T) = X(T,0)$, $A(T) = ((DX(T)_i, DX(T)_j)_H)_{1\le i,j\le 2}$, and $\Delta(T) = \det(A(T))$. By the results in Section 8 of [**4**], it suffices for us to show that

$$\lim_{T\searrow 0} e^{-\epsilon/T}\mathbb{E}^{\mathcal{W}}\big[\Delta(T)^{-p}\big] = 0 \quad \text{for all } \epsilon > 0 \text{ and } p \in [1,\infty). \tag{3.3}$$

Reasoning as before,

$$\Delta(T) = \iint\limits_{0\le s<t\le T} \left(\int_s^t b'\big(\omega(\tau)\big)\, d\tau\right)^2 ds dt,$$

and so, just as before,

$$\mathbb{E}^{\mathcal{W}}\big[e^{-\alpha\Delta(T)}\big] \le C\sum_{n=1}^{\infty}\Big(\exp\big(-\epsilon\alpha T^2 n^{-8} e^{-(\log 2n)^{\beta}}\big) + \exp\big(-\epsilon(\alpha n^{-8} e^{-(\log 2n)^{\beta}})^{1/5}\big)\Big)e^{-\gamma T n^2}.$$

Without loss in generality, we will assume that $\beta \ge 1$, in which case we have that

$$\mathbb{E}^{\mathcal{W}}\big[e^{-\alpha\Delta(T)}\big] \le C\sum_{n=1}^{\infty}\Big(\exp\big(-\epsilon\alpha T^2 e^{-\kappa(\log n)^{\beta}}\big) + \exp\big(-\epsilon(\alpha e^{-\kappa(\log n)^{\beta}})^{1/5}\big)\Big)e^{-\gamma T n^2}$$

for an appropriate choices of $\epsilon > 0$ and $\kappa < \infty$. Splitting the sum according to whether n is dominated by or dominates $\exp((\log^+ \alpha)/\kappa)^{1/\beta}$, one sees that the preceding is bounded above by

$$F(T,\alpha) \equiv \frac{C}{\sqrt{\gamma T}}\left(e^{-\epsilon T^2\alpha^{1/2}} + e^{-\epsilon\alpha^{1/10}} + 2\exp\left[-\gamma T \exp\left[2\left(\frac{\log^+\alpha}{\kappa}\right)^{1/\beta}\right]\right]\right).$$

Hence $\mathcal{W}(\Delta \le 1/R) \le eF(T,R)$, and so

$$\mathbb{E}^{\mathcal{W}}\left[\Delta^{-p}\right] \le pe\int_0^\infty R^{p-1}F(T,R)\,dR.$$

When $T \in (0,1]$, the contribution to the above by the first two terms in the expression for $F(T,R)$ is bounded by a constant times $T^{-4p-(1/2)}$. To handle the third term, we must estimate

$$\int_1^\infty R^{p-1}\exp\left[-\gamma T\exp\left[2\left(\frac{\log R}{\kappa}\right)^{1/\beta}\right]\right]dR = \int_0^\infty e^{pr}\exp\left[-\gamma T\exp\left[2\left(\frac{r}{\kappa}\right)^{1/\beta}\right]\right]dr.$$

To this end, note that

$$\gamma\exp\left[2\left(\frac{r}{\kappa}\right)^{1/\beta}\right] \ge \rho r^4 \quad \text{for some } \rho \in (0,\infty),$$

and therefore that the preceding integral is dominated by

$$\int_0^\infty e^{pr}e^{-\rho T r^4}\,dr = T^{-1/4}\int_0^\infty e^{prT^{-1/4}}e^{-\rho r^4}\,dr.$$

Finally, set $R(T) = (2p/\rho T^{1/4})^{1/\beta}$, and decompose the preceding integral into

$$\int_0^{R(T)} e^{prT^{-1/4}}e^{-\rho r^4}\,dr + \int_{R(T)}^\infty e^{prT^{-1/4}}e^{-\rho r^4}\,dr.$$

Since

$$\frac{pr}{T^{1/4}} - \rho r^4 \le \begin{cases} \frac{3}{4}\left(\frac{p^4}{4\rho T}\right)^{1/3} & \text{for all } r > 0, \\ -\frac{\rho r^4}{2} & \text{for } r \ge R(T), \end{cases}$$

it follows that, for each $p \in [1,\infty)$, there exists a $C_p < \infty$ such that

$$\mathbb{E}^{\mathcal{W}}\left[\Delta(T)^{-p}\right] \le C_p e^{C_p T^{-1/3}} \quad \text{for } T \in (0,1].$$

Hence, (3.3) holds.

References

[1] A. Friedman, Partial Differential Equations of the Parabilic Type, Prentice Hall, 1965.

[2] L. Hörmander, Hypoelliptic second order differential operator, Acta. Math., **121** (1967), 147–171.

[3] S. Kusuoka and D. Stroock, Applications of Malliavin's calculus, part I, In: Stochastic Analysis, Proceedings of the Taniguchi International Symposium on Stochastic Analysis at Katata and Kyoto, 1982, North Holland, 1982, pp. 271–386.

[4] S. Kusuoka and D. Stroock, Applications of Malliavin's calculus, part II, J. Fac. Sci. of Tokyo Univ., Sec IA, **32** (1985), 1–76.

[5] P. Malliavin, Stochastic calculus of variations and hypoelliptic operators, Proc. International Symposium on Stochastic Differential Equations, Kyoto, 1976, J. Wiley, 1973.

[6] O. A. Olenik and E. V. Radekevich, Second Order Equations with Non-Negative Characteristic Form, Plenum Press, 1973.

[7] L. Rothschield and E. Stein, Hypoelliptic differential operators and nilpotent Lie groups, Acta Math., **137** (1977), 247–320.

[8] D. Stroock, Probability Theory, An Analytic View, 2nd ed., Cambridge Univ. Press, 2011.

[9] D. Stroock, Partial Differential Equations for Probabilists, Cambridge Univ. Press, 2008.

[10] D. Stroock and S. R. S. Varadhan, Multidimensional Diffusion Processes, Springer-Verlag, 1997.

Daniel W. Stroock
M.I.T. Department of Mathematics
77 Mass. Ave., Cambridge
MA 02139 USA
E-mail: dws@math.mit.edu