Degenerate Hypoelliptic Operators and the Gaussian Kernel

Denis Bell

University of North Florida

1. Hörmanders theorem

Definition. A differential operator G is *hypoelliptic* if, whenever Gu is smooth, for some distribution u defined on an open subset of the domain of G, then u is smooth.

Suppose $X_0, \ldots, X_n$ are bounded smooth vector fields on $\mathbf{R}^d$ with bounded derivatives of all orders. We consider these as first-order differential operators and write

$$X_i = \sum_{j=1}^{d} X_{ij} \frac{\partial}{\partial x_j},\ i = 1, \ldots, n.$$

Let L denote the second-order differential operator

$$L = \frac{1}{2} \sum_{i=1}^{n} X_i^2 + X_0$$

Theorem (Hörmander, 1967). *Suppose the Lie algebra generated by* $X_0, \ldots, X_n$ *has full rank in an open set* $U \subset \mathbf{R}^d$*, i.e. the vectors*

$$\{X_i, [X_i, X_j], [[X_i, X_j], X_k], \ldots, i, j, k, \cdots = 0, \ldots n\}$$

span $\mathbf{R}^d$ *at all points in* U*. Then* L *is hypoelliptic in* U*.*

The hypothesis above is known as *Hörmanders Lie algebra condition* (HC).

2. A probabilistic formulation of the problem

Let $(w_1, \ldots, w_n)$ denote a standard Wiener process in $\mathbf{R}^n$. Consider the SDE

$$\xi_t^x = x + \sum_{i=1}^{n} \int_0^t X_i(\xi_s^x) dw_i(s) + \int_0^t X_0(\xi_s^x) ds.$$

The solution process ξ is a time-homogeneous Markov process. The transition probabilities

$$p(t, x, dy) \equiv p(\xi_t^x \in dy)$$

satisfy the following partial differential equations (*Kolmogorov backward and forward equations*) in the weak sense

$$\frac{\partial p}{\partial t} = L_x p$$

$$\frac{\partial p}{\partial t} = L_y^* p.$$

Suppose the vector fields $X_0, \dots X_n$ satisfy the following parabolic version of HC at each point in $\mathbf{R}^d$

$$\{X_i, [X_j, X_k], [[X_j, X_k], X_l], \dots$$
$$i \le 1 \le n,\ 0 \le j, k, l \dots \le n\}.$$

Then it follows from Hörmander s theorem that the operators $\frac{\partial}{\partial t} - L$ and $\frac{\partial}{\partial t} - L^*$ are hypoelliptic.

Thus the transition probabilities $p(t, x, dy)$ for the SDE

$$d\xi_t = x + \sum_{i=1}^{n} X_i(\xi_t^x) dw_i + X_0(\xi_t) dt$$

admit densities $p(t, x, y)$ that are smooth in t, x and y.

In the opposite direction, if one can establish by *direct probabilistic methods* that under HC, the above SDE admits smooth transition probabilities, then this can be used to give a probabilistic proof of HT.

3. Basic Malliavin calculus

Let γ denote the Wiener measure on the space of paths

$$C_0 = \{\sigma : [0,1] \mapsto \mathbf{R}^n : \sigma(0) = 0\}.$$

Let $\phi : C_0 \mapsto \mathbf{R}$ be a *cylindrical* function, i.e.

$$\phi(w) = F\Big(w(t_1), \ldots, w(t_d)\Big)$$

where $F : \mathbf{R}^{nd} \mapsto \mathbf{R}$ is C^∞. Then we define an operation D_t by

$$D_t \phi = \sum_{i=1}^{d} \frac{\partial F}{\partial x_i}\Big(w(t_1), \ldots, w(t_d)\Big) I_{[0,t_i]}(t).$$

Let D_1^2 denote the closure of the set of cylindrical functions under the norm

$$||\phi||_{D_1^2} = \left(E\left[|\phi|^2 + \int_0^1 |D_t\phi|^2 dt\right]\right)^{1/2}.$$

Define D_r^p analogously and

$$D^\infty = \bigcap_{r,p \geq 1} D_r^p.$$

Finally, denote the extension of D_t to D^∞ by the same symbol.

For $g : C_0 \mapsto \mathbf{R}^d$ such that each $g_i \in D^\infty$, define the *Malliavin covariance matrix* C by

$$C_{ij} = \int_0^1 (D_t g_i) \cdot (D_t g_j) dt.$$

Motivation. Introduce the *Cameron-Martin* space H, the Hilbert subspace of C_0 consisting of absolutely continuous paths h with finite *energy*

$$\int_0^1 |h'_t|^2.$$

If ϕ is a cylindrical function then

$$D_t \phi = \frac{d}{dt} D_H \phi(w)$$

(note that $D_H \phi(w) \in H^* \sim H$). Furthermore

$$C = D_H g(w) D_H g(w)^*.$$

Theorem (Malliavin). *Suppose* $C \in GL(d)$ *a.s. and*

$$(det\, C)^{-1} \in L^p,\ \forall p \geq 1 \qquad (*)$$

Then the random variable $g(w)$ *is absolutely continuous and has a smooth density.*

Theorem. *Consider the SDE*

$$\xi_t^x = x + \sum_{i=1}^{n} \int_0^t X_i(\xi_s^x) dw_i(s) + \int_0^t X_0(\xi_s^x) ds.$$

Then the map $w \mapsto \xi_t^x$ *lies in* D^∞. *The MCM* C *is given by*

$$C = Y_t \int_0^t Z_s A(\xi_s^x) A(\xi_s^x)^* Z_s^* ds Y_t^*$$

where $A = [X_1 \ldots X_n]$, Y_t *is the derivative of the stochastic flow* $x \mapsto \xi_t^x$, *and* $Z_t = Y_t^{-1}$.

Theorem (Kusuoka-Stroock) *If the vector fields* $X_0, \ldots, X_n$ *satisfy the parabolic HC at* x *then (*) holds. Hence* ξ_t^x *admits a smooth density.*

4. Exponentially degenerate hypoelliptic operators

The hypothesis of Hörmander s theorem (HC) is known to be *necessary* for hypoellipticity of L if the coefficients of L are analytic. This is not the case in the smooth non-analytic category.

Theorem (Kusuoka-Stroock, 1987). *Consider the class of differential operators on* $\mathbf{R}^3$ *of the form*

$$L_p \equiv \frac{\partial^2}{\partial x^2} + \exp(-|x|^p)\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},\ p < 0$$

Then L_p *is hypoelliptic if and only if* $p \in (-1, 0)$*.*

In particular, if $p \in (-1, 0)$ then L_p is hypoelliptic on $\mathbf{R}^3$ but fails to satisfy HC on the hyperplane $\{x = 0\}$.

We describe an extension of HT that encompasses operators with degeneracy of exponential order and includes the K-S operators. The statement of the result requires the following notation.

For $k \geq 0$, define $X^{(k)}$ to be a matrix with columns $X_0, \ldots, X_n$, and all vector fields obtained from $X_0, \ldots, X_n$ by forming iterated Lie brackets up to order k. Define

$$\lambda^{(k)} \equiv \text{ smallest eigenvalue of } X^{(k)} X^{(k)*}.$$

Let H^c denote the set of points in D where L fails to satisfy HC. Then

$$H^c = \{x \in D : \lambda^{(k)}(x) = 0, \ \forall k\}.$$

A C^1 hypersurface $S \subset \mathbf{R}^d$ is said to be *non-characteristic* (with respect to L) at $x \in S$ if at least one of the vector fields $X_1, \ldots, X_n$ is non-tangential to S at x.

Theorem (B-Mohammed). *Suppose the non-Hörmander set H^c of L is contained in a C^2 hypersurface S. Let U be any open subset of the domain of L and assume that for all $x \in H^c \cap U$*

(i) S is non-characteristic at x.

(ii) There exists an open neighborhood V of x, an integer $k \geq 0$, and $p \in (-1, 0)$ such that

$$\lambda^{(k)}(y) \geq \exp\{-[d(y, S)]^p\}, \forall y \in V.$$

Then L is hypoelliptic on U.

Concerning the hypotheses:

Condition (i) is known to be necessary for the hypoellipticity of L.

Condition (ii) controls the rate at which HC fails at points in H^c as we approach S. The non-hypoellipticity of the Kusuoka-Stroock operator L_{-1} shows that an assumption of this type is necessary and the allowed range of p in the theorem is optimal.

5. Outline of the proof

We use the following result to prove a parabolic version of the theorem.

Lemma (Kusuoka-Stroock). *Let* Δ *denote det C where, as before* C *is the MCM*

$$C = Y_t \int_0^t Z_s A(\xi_s^x) A(\xi_s^x)^* Z_s^* ds Y_t^*. \qquad (1)$$

Suppose that for all $q \geq 1$ *and* $x \in D$*, there exists a neighborhood* $V \subset D$ *such that*

$$\lim_{t \to O^+} t \log \left\{ \sup_{y \in V} E\Big[|\Delta(t,y)^{-q}| \Big] \right\} = 0. \qquad (2)$$

Then $L + \partial/\partial t$ *is hypoelliptic on* $\mathbf{R} \times D$*.*

Establishing condition (2) under the hypotheses of the theorem requires analyzing the interaction between the diffusion process ξ and the non-Hörmander surface of L. This constitutes the majority of the proof.

The strategy is as follows:

(i) We express the surface S locally in the form

$$S = \{x \in \mathbf{R}^d / \, \phi(x) = 0\}$$

and translate the hypotheses of the theorem into conditions on ϕ.

(ii) Probabilistic lower bounds are obtained on the L^p-norms of the process $y_t \equiv \phi(\xi_t)$ for arbitrarily large values of p.

(iii) We study how the estimates in (ii) are degraded under exponential-type degeneracy. This yields lower bounds on the integrand in the MCM that are shown to imply the hypothsesis of the K-S lemma.

Our proof characterizes the class of degenerate hypoelliptic operators of Hörmander type in terms of properties of the Wiener process.

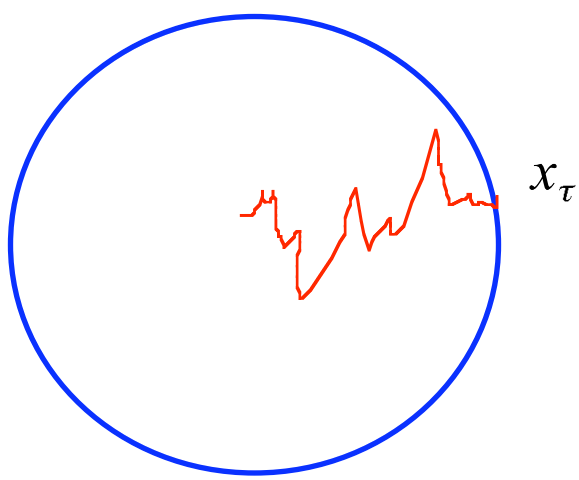
The *space-time scaling* property of the Wiener process plays an essential role in the proof. By means of this result we establish a precise relationship between the maximal class of hypoelliptic operators and the form of the Gaussian kernel $e^{-||x||^2}$.

Definition. A random time τ is *exponentially positive* if there exist positive constants a and b (the *characteristics* of τ) such that

$$P(\tau < \epsilon) \leq \exp(-b/\epsilon)$$

for all $\epsilon < a$.

Example The exit time τ of a diffusion process with bounded coefficients from a ball of fixed radius is exponentially positive.



Lemma 1. *Let y be an Ito process of the form*

$$dy(t) = \sum_{i=1}^{n} a_i(t)dw(t) + b(t)dt$$

where $a_1, \ldots, a_n$ and b are adapted processes. Suppose (i) at least one of $a_1(0), \ldots, a_n(0)$ is non-zero.
(ii) There exists a deterministic constant c such that

$$\sum_{i=1}^{n} |a_i(t)| + |b(t)| \leq c, \forall t \in [0, T].$$

Let τ be an exponentially positive stopping time. Then for every $p \in (-1, 0)$, there exists a positive constant β and $q > 1$ such that

$$P\left(\int_0^{t\wedge\tau} \exp(-|y(u)|^p)du < \epsilon \right) \leq \exp(-\beta|\log \epsilon|^q)$$

for all $\epsilon < \exp(-t^{-1/q})$. The constants β and q depend only on p, $a_1(0), \ldots, a_n(0)$, c, and the characteristics of τ.

Define $\lambda^{(k)}$ to be the smallest eigenvalue of $X^{(k)}X^{(k)*}$ where $X^{(k)}$ is a matrix with columns $X_1, \ldots, X_n$ and vector fields consisting of iterated Lie brackets of $X_0, \ldots, X_n$ up to order k.

Lemma 2. *Under the hypotheses of the Theorem, for each $x \in H^c$, there exists a neighborhood U of x, a C^2 map $\phi : U \mapsto \mathbf{R}^d$ and $p \in (-1, 0)$ such that*
(i) $\phi(x) = 0$, $\nabla\phi(x).X_i(x) \neq 0$ for at least one $i = 1, \ldots, n$.
(ii) For some $k \geq 1$

$$\lambda^{(k)}(y) \geq \exp\Big\{-|\phi(y)|^p\Big\},\ \forall y \in U. \qquad (3)$$

Remark ϕ is a function whose vanishing set locally defines the surface S containing H^c.

We are trying to establish the estimate

$$\lim_{t\to O^+} t \log\left\{\sup_{x\in V} E\Big[|\Delta(t,x)^{-q}|\Big]\right\} = 0. \qquad (4)$$

where $\Delta(t,x)$ is the determinant of the Malliavin covariance matrix corresponding to the map $w \mapsto \xi_t^x$.

We initially prove that

$$E\Big[|\Delta(t,x)^{-q}|\Big]\Big\} \le c \sum_{j=1}^{\infty} P\Big(Q(t,x) \le j^{-1/(dq)}\Big)$$

where

$$Q(t,x) \equiv \inf\left\{ \sum_{i=1}^{n} \int_0^t < Z^x(u)X_i(\xi_u^x), h >^2 du,\ |h| = 1\right\}.$$

Since Z^x satisfies strong stochastic lower bounds, one can (effectively) replace $Q(t,x)$ above by

$$\int_0^{t\wedge\tau} \lambda^{(1)}(\xi_u^x) du$$

where τ is some exponentially positive stopping time.

Thus

$$E\Big[|\Delta(t,x)^{-q}|\Big]\Big\} \leq$$

$$c\sum_{j=1}^{\infty} P\Big(\int_0^{t\wedge\tau} \lambda^{(1)}(\xi_u^x)du \leq j^{-1/(dq)}\Big) \qquad (5)$$

In order to simplify the exposition, we now assume the hypothesis of the Theorem holds at x with $k = 1$. Hence conclusion (3) in Lemma 2 (ii) holds with $k = 1$:

$$\lambda^{(1)}(y) \geq \exp\Big\{-|\phi(y)|^p\Big\}.$$

Using this in (5), we have

$$E\Big[|\Delta(t,x)^{-q}|\Big]\Big\} \leq$$

$$c\sum_{j=1}^{\infty} P\Big(\int_0^{t\wedge\tau} \exp\Big\{-|\phi(\xi_u^x)|^p\Big\}du \leq j^{-1/(dq)}\Big) \qquad (6)$$

By Ito s formula the process $\phi(\xi_t^x)$ satisfies

$$d\phi(\xi_t^x) = \sum_{i=1}^{n} \nabla\phi((\xi_t^x) \cdot X_i(\xi_t^x) dw_i(t) + G(t)dt$$

for some function G. By Lemma 2 (i), the process $y_t \equiv |\phi(\xi_t^x)|$ satisfies the hypotheses of Lemma 1.

Applying Lemma 1 with this choice of y_t gives: there exists $r > 1$ such that

$$P\Big(\int_0^{t\wedge\tau} \text{exp}\Big\{ -|\phi(\xi_u^x)|^p \Big\} du \leq j^{-1/(dq)} \Big)$$
$$\leq \text{exp}\{-\beta(\log j)^r\} \tag{7}$$

for j satisfying $j^{-1/(dq)} \leq \text{exp}(-t^{-1/r})$, i.e.

$j \geq \text{exp}(\gamma t^{-1/r})$ where $\gamma = dq$.

Substituting (7) into (6), we deduce

$$E\Big[|\Delta(t,x)^{-q}|\Big]\Big\} \leq$$

$$c\Big(\exp(\gamma t^{-1/r}) + \sum_{j=1}^{\infty} \exp\{-\beta(\log j)^r\}\Big).$$

The constants can be chosen to uniform be in x in a small enough neighborhood $V \subset D$.

This implies the criterion in the Kusuoka-Stroock lemma:

$$\lim_{t \to O^+} t \log \Big\{ \sup_{y \in V} ||E\Big[|\Delta(t,y)^{-q}|\Big]\Big\}\Big\} = 0.$$

and it follows that the operator $L + \partial/\partial t$ is hypoelliptic.

This is a parabolic form of the main result. The full result can be deduced from this.