Heat kernel asymptotics at the cut locus

Robert Neel

Department of Mathematics
Lehigh University

October 3, 2014
Institute Henri Poincaré
Workshop on geometric analysis on sub-Riemannian manifolds
I would like to thank the organizers for the invitation to speak at this workshop and for arranging this trimester.

The work on logarithmic derivatives of the heat kernel began with work done for my thesis, written under the direction of Dan Stroock (MIT).

Subsequent work largely in the sub-Riemannian context (the majority of the talk) is joint with Ugo Boscain (CNRS and École Polytechnique), Davide Barilari (Paris VII), and Grégoire Charlot (Grenoble).
Let $M$ be a complete, connected, smooth Riemannian manifold of dimension $n$. For $x \in M$, $\text{Cut}(x)$ is:

- the set of $y \in M$ such that there is more than one minimal geodesic from $x$ to $y$, or there is a minimal geodesic from $x$ to $y$ which is conjugate (or both);
- the closure of the set where $\text{dist}(x, \cdot)$ is not differentiable;
- the points where geodesics cease to minimize distance.
The picture
A sub-Riemannian manifold may admit abnormal minimizers in addition to (normal) geodesics. These are poorly understood, and we will avoid them.

Away from abnormalities, the exponential map and cut and conjugate loci are largely analogous to the Riemannian case, although note that $\text{Cut}(x)$ is adjacent to $x$.

We equip our (complete) sub-Riemannian manifold with a smooth volume (say, the Popp volume) and associated sub-Laplacian, which generates a heat flow/hypoelliptic diffusion.
The Heisenberg group

Let $\partial_x - (y/2)\partial_z$ and $\partial_y + (x/2)\partial_z$ be orthonormal in $\mathbb{R}^3$:
Perturbed: 3D contact case

cut

conjugate

generic case

geodesics

conjugate locus

cut locus

Sphere

Front

1

3
The heat kernel

Let

- \( E(x, y) = \frac{1}{2} \text{dist}(x, y)^2 \) be the energy function,
- \( \Delta \) the (sub-)Laplacian on \( M \),
- \( p_t(x, y) \) the heat kernel (the fundamental solution to \( \partial_t u_t(x) = \Delta u_t(x) \)). (We try to stick to the analysts’ normalization, but it’s possible some factors of \( 1/2 \) will be off when discussing the probabilistic side of things.)

As \( t \downarrow 0 \),

- \(-2t \log p_t(x, y) \to E(x, y)\) uniformly on compacts, due to Varadhan (or Leandre).
- \( p_t(x, y) \sim \left(\frac{1}{4\pi t}\right)^{n/2} e^{-d^2(x,y)/4t} \sum_{i=0}^{\infty} H_i(x, y)t^i \) on \( M \setminus \text{Cut}(x) \) (or also minus \( x \) and any abnormalities), due to Minakshisundaram and Pleijel (or Ben Arous).
In the 70’s, Molchanov discussed a method (later formalized by Hsu) to get an expansion similar to that of Minakshisundaram and Pleijel at the cut locus in the Riemannian case. Take \( x, y \in M \), let \( \Gamma \) be the set of midpoints of minimal geodesics from \( x \) to \( y \) and let \( \Gamma_\epsilon \) be an \( \epsilon \)-neighborhood.

Let \( h_{x,y}(z) = E(x, z) + E(z, y) \) be the *hinged energy function*. Note \( h_{x,y}(z) \) achieves its minimum (of \( d^2(x, y)/4 \)) exactly on the set \( \Gamma \).

The idea is to glue two copies of the earlier expansion at \( \Gamma \); it is fairly elementary and broadly applicable.
The computation

\[ p_t(x, y) = \int_M p_{t/2}(x, z)p_{t/2}(z, y) \, dz \]

\[ = \int_{\Gamma_\epsilon} p_{t/2}(x, z)p_{t/2}(z, y) \, dz + \int_{M\setminus\Gamma_\epsilon} \cdots \]

\[ \sim \int_{\Gamma_\epsilon} p_{t/2}(x, z)p_{t/2}(z, y) \, dz \]

\[ \sim \int_{\Gamma_\epsilon} \left( \frac{1}{2\pi t} \right)^{n/2} e^{-E(x,z)/t} H_0(x, z) \]

\[ \times \left( \frac{1}{2\pi t} \right)^{n/2} e^{-E(z,y)/t} H_0(z, y) \, dz \]

\[ = \left( \frac{1}{2\pi t} \right)^n \int_{\Gamma_\epsilon} H_0(x, z)H_0(z, y)e^{-h_{x,y}(z)/t} \, dz \]
Laplace integrals

This leads us to study integrals of the form

$$\int \varphi(z)e^{-g(z)/t} \, dz$$

as $t \searrow 0$, for non-negative $g$.

- This is an established subject, called Laplace asymptotics.
- Kanwal and Estrada give a complete expansion if $g$ is diagonalizable.
- Arnold, Vasiliev, Varchenko, Gusein-Sade give a more general description of the leading term, using resolution of singularities.
We consider the one-dimensional case. Suppose, maybe after smooth change of coordinates, that \( g(z) = g(0) + z^2 \) on \((-\epsilon, \epsilon)\). Then

\[
\int_{|z| \leq \epsilon} \varphi(z) e^{-g(z)/t} \, dz \sim (\varphi(0) \sqrt{\pi}) \, t^{1/2} e^{-g(0)/t}.
\]

If \( g(z) = g(0) + z^4 \) on \((-\epsilon, \epsilon)\), then

\[
\int_{|z| \leq \epsilon} \varphi(z) e^{-g(z)/t} \, dz \sim \left( \varphi(0) \frac{\Gamma(1/4)}{2} \right) \, t^{1/4} e^{-g(0)/t}.
\]
For the heat kernel

Let $M = S^1 \equiv \mathbb{R}/2\pi\mathbb{Z}$. For $\theta \in (0, \pi)$, i.e. not the cut locus

$$p_t(0, \theta) \sim c(\theta) \frac{1}{t^{1/2}} e^{-\theta^2/4t}.$$

On the cut locus ($\theta = \pi$),

$$p_t(0, \pi) \sim 2 \cdot c(\pi) \frac{1}{t^{1/2}} e^{-\pi^2/4t}.$$

Here $c(\theta)$ is continuous.
General role of universal cover
Examples: $\mathbb{S}^2$ and the Heisenberg group

Let $N$ and $S$ be the North and South poles of $\mathbb{S}^2$. For $y \neq S$,

$$p_t(N, y) \sim \text{const.} \frac{1}{t} e^{-\text{dist}^2(N, y)/4t}.$$ 

Now let $y = S$. Then $\Gamma$ is the equator, and integrating over $\Gamma$ gives

$$p_t(N, S) \sim \text{const.} \frac{1}{t^{3/2}} e^{-\text{dist}^2(N, S)/4t}.$$ 

Similarly, for the Heisenberg group, if $y \neq x$ then

$$y \notin \text{Cut}(x) \Rightarrow p_t(x, y) \sim \text{const.} \frac{1}{t^{3/2}} e^{-\text{dist}^2(x, y)/4t},$$

$$y \in \text{Cut}(x) \Rightarrow p_t(x, y) \sim \text{const.} \frac{1}{t^2} e^{-\text{dist}^2(x, y)/4t}.$$
The role of conjugacy

As we’ve seen, the Taylor expansion of $h_{x,y}$ near its minima governs the power of $1/t$ appearing in the expansion of $p_t(x,y)$.

This Taylor series encodes information about how conjugate the geodesic is. For example, at the midpoint of a geodesic, $h_{x,y}$ has degenerate Hessian if and only if the geodesic is conjugate, and “more degeneracy” corresponds to “more conjugacy.”

Thus more conjugacy leads to a larger power of $1/t$. 
Theorem (Barilari, Boscain, N.)

For \( x \) and \( y \) in a (smooth, complete) Riemannian manifold or a sub-Riemannian manifold (with \( x \neq y \) and every minimizer from \( x \) to \( y \) is a strongly normal geodesic), we have:

- If \( x \) and \( y \) are conjugate along any minimal geodesic,
  \[
  \frac{C_1}{t^{(n/2)+(1/4)}} e^{-d^2(x,y)/4t} \leq p_t(x,y) \leq \frac{C_2}{t^{n-(1/2)}} e^{-d^2(x,y)/4t}
  \]
  for all small enough \( t \).

- If \( x \) and \( y \) are not conjugate along any minimal geodesic,
  \[
  p_t(x,y) = \frac{C + O(t)}{t^{n/2}} e^{-d^2(x,y)/4t}.
  \]
Theorem (Barilari, Boscain, Charlot, N.)

For $x$ and $y$ in a (smooth, complete) Riemannian manifold or a sub-Riemannian manifold (with $x \neq y$) suppose there is a unique minimizing strongly normal geodesic from $x$ to $y$ (which we denote $\text{Exp}_x(t\lambda)$ for $0 \leq t \leq 1$, and $\lambda$ a covector). Then if $D_\lambda \text{Exp}_x$ has rank $n - r$ for some $r \in \{0, 1, 2, \ldots, n-1\}$, then for all small enough $t$

$$\frac{C_1}{t^{n + \frac{r}{4}}} e^{-d^2(x,y)/4t} \leq p_t(x,y) \leq \frac{C_2}{t^{n + \frac{r}{2}}} e^{-d^2(x,y)/4t}.$$ 

The exact power of $t$ depends on the precise behavior of the exponential map in the conjugate directions.
Up to (and including ) dimension 5, the generic singularities of the exponential map have normal forms from the following list:

\( A_2 : x \mapsto x^2 \) or a suspension,
\( A_3 : (x, y) \mapsto (x^3 + xy, y) \) or a suspension,
\( A_4 : (x, y, z) \mapsto (x^4 + x^2y + xz, y, z) \) or a suspension,
\( A_5 : (x, y, z, t) \mapsto (x^5 + x^3y + x^2z + xt, y, z, t) \) or a suspension,
\( A_6, D_4^+, D_4^-, D_5^+, D_5^-, D_6^+, D_6^-, E_6^+, \) or \( E_6^- \).

Let \( M \) be a Riemannian manifold, \( x, y \in M \) such that \( \gamma(t) = \text{Exp}_x(tv) \) for \( 0 \leq t \leq 1 \) gives a minimizing conjugate geodesic from \( x \) to \( y \). Then we say that \( \gamma \) is \( A_2 \)-conjugate if at \( v, \text{Exp}_p \) has a normal form given by \( A_2 \). We define \( A_3 \)-conjugacy, etc. in a similar way.
If $\gamma$ is $A_m$-conjugate, then near the midpoint of $\gamma$, $h_{x,y}$ has the form

$$h_{x,y}(z) = \frac{1}{4}d^2(x, y) + z_1^2 + \ldots + z_{n-1}^2 + z_{m+1}^n.$$

Suppose that, for some $\ell \in \{3, 5, 7, \ldots\}$ every minimizing geodesic from $x$ to $y$ is non-conjugate or $A_m$-conjugate for some $3 \leq m \leq \ell$, and at least one is $A_\ell$. Then there exists $C > 0$ such that

$$p_t(x, y) = \frac{C + O\left(t^{\frac{2}{\ell+1}}\right)}{t^{\frac{n+1}{2} - \frac{1}{\ell+1}}} e^{-d^2(x,y)/4t}.$$
Generic minimizing singularities

Theorem (Barilari, Boscain, Charlot, N.)

Let $M$ be a smooth manifold, $\dim M = n \leq 5$, and $x \in M$. For a generic Riemannian metric on $M$ and any minimizing geodesic $\gamma$ from $x$ to some $y$, $\gamma$ is either non-conjugate, $A_3$-conjugate, or $A_5$-conjugate.

The only possible heat kernel asymptotics are (here $C > 0$ is some constant which can differ from line to line):

- If no minimizing geodesic from $x$ to $y$ is conjugate, then
  \[ p_t(x, y) = \frac{C+O(t)}{t^{n/2}} e^{-d^2(x,y)/4t}, \]

- If at least one minimizing geodesic from $p$ to $q$ is $A_3$-conjugate but none is $A_5$-conjugate, $p_t(x, y) = \frac{C+O(t^{1/2})}{t^{n/2+\frac{1}{4}}} e^{-d^2(x,y)/4t},$

- If at least one minimizing geodesic from $p$ to $q$ is $A_5$-conjugate, $p_t(x, y) = \frac{C+O(t^{1/3})}{t^{n/2+\frac{1}{6}}} e^{-d^2(x,y)/4t}.$
A sub-Riemannian case

**Theorem (Barilari, Boscain, Charlot, N.)**

Let $M$ be a smooth manifold of dimension 3. Then for a generic 3D contact sub-Riemannian metric on $M$, every $x$, and every $y$ close enough to $x$:

(i) If no minimizing geodesic from $x$ to $y$ is conjugate then

$$p_t(x, y) = \frac{C + O(t)}{t^{3/2}} e^{-d^2(x,y)/4t},$$

(ii) If at least one minimizing geodesic from $x$ to $y$ is conjugate it is $A_3$-conjugate and

$$p_t(x, y) = \frac{C + O(t^{1/2})}{t^{7/4}} e^{-d^2(x,y)/4t}.$$ 

Moreover, there are points $y$ arbitrarily close to $x$ such that case (ii) occurs.
Non-generically, there is much more variety.

**Theorem (Barilari, Boscain, Charlot, N.)**

For any integer $\eta \geq 3$, any positive real $\alpha$, and any real $\beta$, there exists a smooth metric on the sphere $\mathbb{S}^2$ and (distinct) points $x$ and $y$ such that the heat kernel has the small-time asymptotic expansion

$$p_t(x, y) = e^{-d^2(x,y)/4t} \frac{1}{t^{(3\eta-1)/2\eta}} \left\{ \alpha + t^{1/\eta} \beta + o \left( t^{1/\eta} \right) \right\}.$$
Assume $M$ is Riemannian (and compact). Motivated by Varadhan’s result, we define

$$E_t(x, y) = -2t \log p_t(x, y)$$

so that

$$E_t(x, y) \to E(x, y) \text{ as } t \downarrow 0.$$ 

Malliavin and Stroock (probabilistically) and Berline, Getzler, and Vergne (analytically) show that, away from the cut locus, spatial derivatives of $E_t(x, y)$ commute with the limit as $t \downarrow 0$.

The lack of differentiability of $E(x, y)$ at the cut locus means that something else must be occurring there; we will describe this “something else.”
Again let $M = S^1 \equiv \mathbb{R}/2\pi \mathbb{Z}$.

On the cut locus ($\theta = \pi$),

$$\lim_{t \searrow 0} \partial_\theta E_t(0, \theta)|_{\theta=\pi} = 0,$$

while

$$\partial^2_\theta E_t(0, \theta)|_{\theta=\pi} \sim -\frac{\pi^2}{t}.$$

- Hessian blows up like $1/t$.
- This blow-up is in the negative direction.
- Unsurprising, since $\nabla^2_{A,A} E(x, y)$, thought of as a distribution, has as singular part a non-positive measure.
The measure

As before, we’re concerned with $h_{x,y}$ near $\Gamma$. But because of the log-derivatives, we need the following one-parameter family of probability measures:

$$
\mu_t(dz) = \frac{1_{\Gamma_\epsilon}(z)}{Z_t} H_0(x, z) H_0(y, z) \exp \left( - \frac{h_{x,y}(z)}{t} \right) \, dz
$$

where

$$
Z_t = \int_{\Gamma_\epsilon} H_0(x, z) H_0(y, z) \exp \left( - \frac{h_{x,y}(z)}{t} \right) \, dz.
$$
The main formulas

Let $A$ be a smooth vector field on $M$. Our covariant derivatives act on the $y$-variable.

Theorem (N.)

For a smooth, compact, connected (Riemannian) manifold $M$, let $x$ and $y$ be any distinct points. Then, with the above notation, we have

$$\nabla_A E_t(x, y) = \int_{\Gamma_{\epsilon}} \nabla_A E(z, y) \mu_t(dz) + O(t)$$

$$= 2E^{\mu t} [\nabla A E(\cdot, y)] + O(t)$$

and

$$\nabla_{A,A}^2 E_t(x, y) = -\frac{4}{t} \text{Var}^{\mu t} [\nabla A E(\cdot, y)] + O(1).$$

These formulas are derived by extending the approach of Molchanov. They also require global estimates on log-derivatives, due to Stroock and Turetsky, and Hsu.
Malliavan and Stroock previously used path space integration to show that, if the set of minimal geodesics connecting $x$ and $y$ is sufficiently nice, then $\nabla^2 E_t(x, y)$ is asymptotic to $-1/t$ times the variance of some random variable on path space as $t \downarrow 0$.

Why the variance?

- $L(t) = \log \mathbb{E}[e^{tX}]$ is the moment generating function of the random variable $X$.
- Then $L''(0) = \text{Var}(X)$.
- The heat semigroup is $e^{t\Delta}$.
- If the heat kernel is the expectation of this semigroup, then the Hessian of the log of the heat kernel at time zero should be the “variance” of $\Delta$. 
How do we interpret the “variance” of $\Delta$?

- Think of “variance” of Brownian motion.
- Under the Feynman picture, distribution of $(\sqrt{2}$-dilated) BM has “density” on pathspace proportional to

$$\exp \left( -\frac{1}{4t} \int_0^1 |w'(\tau)|^2 \, d\tau \right).$$

- For paths from $x$ to $y$ in time $t$, as $t \downarrow 0$ this measure should be concentrating on the minimal geodesics joining $x$ and $y$.
- Heuristically, we guess that, as $t \downarrow 0$, $\nabla^2 \log p_t(x, y)$ should be the “variance” of minimal geodesics from $x$ to $y$. 
Example: $S^n$

Let $N$ and $S$ be the North and South poles of $S^n$. Then $\Gamma$ is the equatorial sphere $S^{n-1}(1)$. By symmetry, $\mu_t$ converges to the uniform probability measure on the equatorial sphere. Next, let $A$ be any vector in $T_SM$.

It is straightforward to compute that

$$\nabla^2_{A,A}E_t(N, S) \sim -\frac{\pi^2|A|^2}{nt}$$

as $t \searrow 0$. 

Robert Neel (Lehigh University)
As before, the Taylor series of $h_{x,y}$ near its minima governs the asymptotics of $\mu_t$. The more conjugate a geodesic is, the more degenerate the Hessian of $h_{x,y}$ is, and the more the mass desires to concentrate on that geodesic.

To be concrete, suppose that there are three minimal geodesics from $x$ to $y$, with $\gamma_1$ non-conjugate and $\gamma_2$ and $\gamma_3$ each $A_3$-conjugate. Then $\mu_t \to \mu_0$ with $\mu_0$ supported on the midpoints of $\gamma_2$ and $\gamma_3$. 
Characterizing the cut locus

Instead of understanding classes of examples, we can give a general result.

**Theorem (N.)**

Let $M$ be a compact, smooth Riemannian manifold, and let $x$ and $y$ be any two distinct points of $M$. Then $y \not\in \text{Cut}(x)$ if and only if

$$\lim_{t \searrow 0} \nabla^2 E_t(x, y) = \nabla^2 E(x, y)$$

and $y \in \text{Cut}(x)$ if and only if

$$\limsup_{t \searrow 0} \left\| \nabla^2 E_t(x, y) \right\| = \infty$$

where $\left\| \nabla^2 E_t(x, y) \right\|$ is the operator norm. Further, if $M$ is real-analytic, the limit supremum can be replaced with the limit (and the proof considerably simplified).