Molchanov's technique for small-time heat kernel asymptotics at cut points

Robert Neel

Department of Mathematics
Lehigh University

November 20, 2018
Stochastic Analysis and Related Topics
Okayama University

Acknowledgments

I would again like to thank everyone involved in arranging such an interesting conference, and inviting me to it, especially Professors Kusuoka, Kawabi, and Aida.

This work is joint with Ugo Boscain (CNRS), Davide Barilari (Paris VII), and Grégoire Charlot (Grenoble).

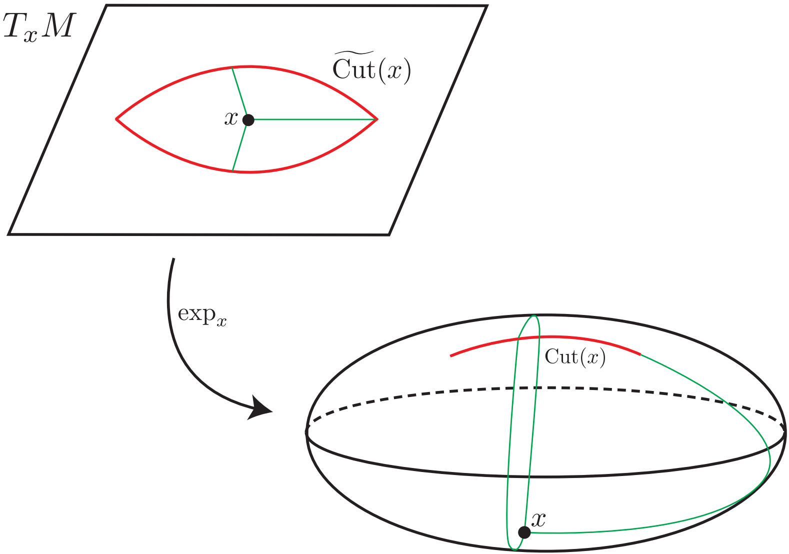
The cut locus

Let M be a complete, connected, smooth Riemannian manifold of dimension n.

For $x \in M$, $\mathrm{Cut}(x)$ 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 $\mathrm{dist}(x, \cdot)$ is not differentiable;
- the points where geodesics cease to minimize distance.

The picture



sub-Riemannian geometry

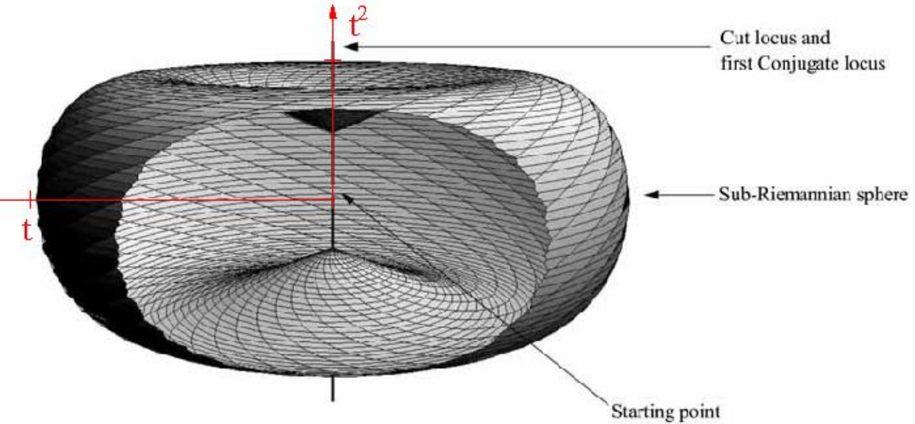
A sub-Riemannian manifold may admit abnormal minimizers in addition to (normal) geodesics. These are poorly understood, and we will avoid them. Note that in several important classes of sub-Riemannian manifolds, such as contact and CR geometry, abnormals do not arise.

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

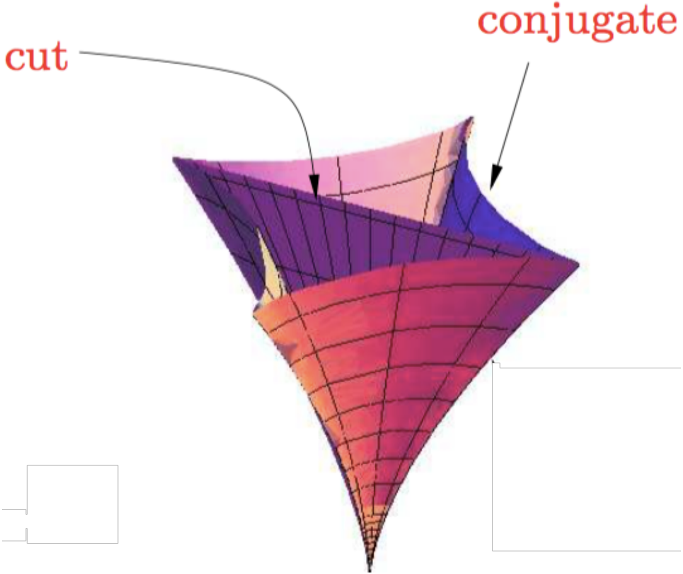
We equip our (smooth, connected, complete) sub-Riemannian manifold with a sub-Laplacian, which gives rise to a hypoelliptic diffusion, and a smooth volume, which serves as a reference measure for the associated heat kernel.

The Heisenberg group

Let $X = \partial_x - (y/2)\partial_z$ and $Y = \partial_y + (x/2)\partial_z$ be orthonormal in $\mathbb{R}^3$ (here $\Delta = X^2 + Y^2$ and the volume is the Euclidean one):



Perturbed: 3D contact case



The heat kernel

Let

- $E(x,y) = \frac{1}{2}\,\mathrm{dist}(x,y)^2$ be the energy function,
- Δ 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.)

As $t \searrow 0$,

- $-2t \log p_t(x,y) \to E(x,y)$ 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 $M \setminus \mathrm{Cut}(x)$ (or also minus x and any abnormals), due to Minakshisundaram and Pleijel (or Ben Arous).

Molchanov's technique

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. It is quite flexible, requiring 3 ingredients

- a "global" coarse estimate, like Varadhan/Leandre above
- a finer estimate off of the cut locus, like Minakshisundaram-Pleijel/Ben Arous above
- the Markov property/Chapman-Kolmogorov equation

Below, we develop this idea for the leading term in the Riemannian case and extend it to the sub-Riemannian case. (Further extensions are work in progress...)

The idea is to glue two copies of the expansion at Γ.

Other approaches

Integral representations of hypoelliptic heat kernels for left-invariant structures on Lie groups have been studied algebraically going back to Gaveau and Hulanicki (Heisenberg group, late 70s) and Beals-Gaveau-Greiner (higher-dimension extension of this, mid-90s). Asaad-Gordina (2016) gave a general treatment for nilpotent Lie groups via generalized Fourier transform.

The positively and negatively curved sub-Riemannian model spaces, de Sitter and anti-de Sitter, space also admit explicit integral representations for the heat kernel, as developed by Bonnefont, Badoin-Bonnefont, and Baudoin-Wang ('09-'12).

Recently, Inahama-Taniguchi (2017) used Watanabe's distributional Malliavin calculus to give a general approach to sub-Riemannian heat kernel asymptotics, and Ludewig (2018) gave similar asymptotics for Riemannian vector bundles via a path-integral-type approach. (Also Kusuoka-Stroock…)

Basic objects

Take $x, y \in M$, let Γ be the set of midpoints of minimal geodesics from x to y and let Γ_ϵ be an ϵ-neighborhood. For example, if M is the standard sphere and x, y the north and south poles, Γ is the equator.

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 Γ.
- For $z \in \Gamma$, $\nabla^2 h_{x,y}(z)$ is non-degenerate if and only if the geodesic from x to y through z is non-conjugate.

The computation

$$
\begin{aligned}
p_t(x,y) &= \int_M p_{t/2}(x,z)p_{t/2}(z,y)\,dz \\
&= \int_{\Gamma_\epsilon} p_{t/2}(x,z)p_{t/2}(z,y)\,dz + \int_{M\setminus\Gamma_\epsilon} \cdots \\
&\sim \int_{\Gamma_\epsilon} p_{t/2}(x,z)p_{t/2}(z,y)\,dz \\
&\sim \int_{\Gamma_\epsilon} \left[\left(\frac{1}{2\pi t}\right)^{n/2}\right]^2 e^{-E(x,z)/t}H_0(x,z)\cdot e^{-E(z,y)/t}H_0(z,y)\,dz \\
&= \left(\frac{1}{2\pi t}\right)^n \int_{\Gamma_\epsilon} H_0(x,z)H_0(z,y)e^{-h_{x,y}(z)/t}\,dz
\end{aligned}
$$

The rigorous sub-Riemannian version of this goes back to Barilari-Boscain-N. ('12).

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.

For example, in 1D, suppose, maybe after smooth change of coordinates, that $g(z) = g(0) + z^2$ on $(-\epsilon, \epsilon)$. Then

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

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

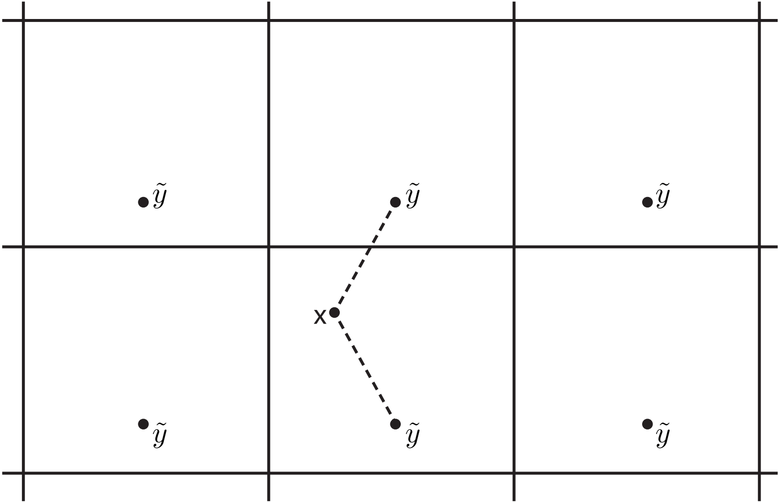
$$p_t(0, \theta) \sim c(\theta) \frac{1}{t^{1/2}} e^{-\theta^2/4t}.$$

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 $y \neq S$,

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

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

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

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

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

Note that in both of these cases, $h_{x,y}$ is Morse-Bott (the Hessian is not degenerate on the normal bundle to $T\Gamma$.)

The Morse-Bott case

Theorem (BBN in press at Ann. Fac. Sci. Toulouse Math. & '12, Kusuoka-Stroock 90s, Inahama-Taniguchi '17)

Let M be a Riemannian or sub-Riemannian manifold as above, with x and y distinct and every optimal geodesic joining x to y strongly normal. Define

$$\mathcal{O} := \{p \in T_x^* M \mid \operatorname{Exp}_x(p, d(x,y)) = y\}$$

Assume that:

(i) *$\mathcal{O}$ is a submanifold of $T_x^* M$ of dimension r.*

(ii) *for every $p \in \mathcal{O}$ we have* $\dim \ker D_{p,d(x,y)} \operatorname{Exp}_x = r$.

Then there exists a positive constant C such that

$$p_t(x,y) = \frac{C + O(t)}{t^{\frac{n+r}{2}}} e^{-d^2(x,y)/4t} \quad \textit{for small } t.$$

The role of conjugacy

The Taylor expansion/normal form of $h_{x,y}$ near its minima governs the power of $1/t$ appearing in the expansion of $p_t(x,y)$. The behavior of $h_{x,y}$, in turn is governed by the exponential map; a "more degenerate" Hessian corresponds to "more conjugacy."

Thus more conjugacy leads to a larger power of $1/t$.

Generic singularities

Up to (and including) dimension 5, the generic singularities of the Riemannian exponential map have normal forms from the following list, in the Arnold classification (Weinstein, Janeczko-Mostowski):

$A_2 : x \mapsto x^2$ 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 $\gamma(t) = \mathrm{Exp}_x(tv)$ for $0 \le t \le 1$ gives a minimizing conjugate geodesic from x to y. Then we say that γ is A_2-*conjugate* if at v, Exp_p has a normal form given by A_2. We define A_3-conjugacy, etc. in a similar way.

A_m-conjugacy

If γ is A_m-conjugate, then near the midpoint of γ, $h_{x,y}$ has the form

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

Note: This implies a minimizing geodesics can't be A_{2k}-conjugate.

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_ℓ . 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. '17)

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 γ from x to some y, γ 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* $p_t(x,y) = \frac{C+O(t)}{t^{\frac{n}{2}}} e^{-d^2(x,y)/4t}$,
- *If at least one minimizing geodesic from p to q is A_3-conjugate but none is A_5-conjugate,* $p_t(x,y) = \frac{C+O(t^{1/2})}{t^{\frac{n}{2}+\frac{1}{4}}} e^{-d^2(x,y)/4t}$,
- *If at least one minimizing geodesic from p to q is A_5-conjugate,* $p_t(x,y) = \frac{C+O(t^{1/3})}{t^{\frac{n}{2}+\frac{1}{6}}} e^{-d^2(x,y)/4t}$.

A sub-Riemannian case

Theorem (BBCN '17)

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-generic cases

Non-generically, there is much more variety.

Theorem (BBCN '17)

For any integer $\eta \geq 3$, any positive real α, and any real β, 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

$$p_t(x,y) = e^{-d^2(x,y)/4t} \frac{1}{t^{(3\eta-1)/2\eta}} \left\{ \alpha + t^{1/\eta}\beta + o\left(t^{1/\eta}\right) \right\}.$$

In particular, in contrast to some suggestions in older literature, the expansion need not proceed in integer powers of t.

More general statements

If we only have estimates on the exponential map, we get heat kernel estimates.

Theorem (BBCN '17)

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

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

General bounds

Theorem (BBN '12)

For x and y in a Riemannian or 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,*

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

More asymptotics

Assume M is Riemannian (and compact). Motivated by Varadhan's result, we define

$$E_t(x,y) = -2t \log p_t(x,y) \quad \text{so that}$$
$$E_t(x,y) \to E(x,y) \quad \text{as } t \searrow 0.$$

Malliavin and Stroock (probabilistically, '96) and Berline, Getzler, and Vergne (analytically, -'92) show that, away from the cut locus, spatial derivatives of $E_t(x,y)$ commute with the limit as $t \searrow 0$.

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."

Example: $\mathbb{S}^1$

Again let $M = \mathbb{S}^1 \equiv \mathbb{R}/2\pi\mathbb{Z}$.

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

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

$$\text{while} \quad \partial_\theta^2 E_t(0,\theta)\Big|_{\theta=\pi} \sim -\frac{\pi^2}{t}.$$

- 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 Γ. But because of the log-derivatives, we need the following one-parameter family of probability measures:

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

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

Relationship to pathspace

Fix x and y, and let P_t be the measure on path-space corresponding to the Brownian bridge from x to y at time t. Let ν_t be the push-forward of P_t under the map that takes each path to its position at time $t/2$. Then $\mu_t \to \mu$ (weakly as measures on M) if and only if "$\nu_t \to \mu$."

Let A be any vector in T_yM. Let $\theta_A(z)$ be the angle between A and the unit tangent to the geodesic from x to y through z, evaluated at y. Then

$$\nabla_A E(z,y) = \frac{1}{2}\operatorname{dist}(x,y)|A|\cos\theta_A(z),$$

which we consider as a function of $z \in \Gamma_\epsilon$.

The main formulas

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

Theorem (N. '08)

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

$$\begin{aligned}
\nabla_A E_t(x,y) &= \int_{\Gamma_\epsilon} \nabla_A E(z,y)\mu_t(dz) + O(t) \\
&= 2\mathbb{E}^{\mu_t}\left[\nabla_A E(\cdot,y)\right] + O(t)
\end{aligned}$$

$$\textit{and} \quad \nabla^2_{A,A} E_t(x,y) = -\frac{4}{t}\operatorname{Var}^{\mu_t}\left[\nabla_A E(\cdot,y)\right] + O(1).$$

These formulas are derived by extending the approach of Molchanov. They also require global estimates on log-derivatives, due to Stroock and Turetsky (late 90s), and Hsu ('02).

Example: $\mathbb{S}^n$

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

It is straightforward to compute that

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

as $t \searrow 0$.

The variance

Malliavan and Stroock previosuly 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 $t \searrow 0$.

Why the variance?

- $L(t) = \log \mathbb{E}[e^{tX}]$ is the moment generating function of the random variable X.
- Then $L''(0) = \mathrm{Var}(X)$.
- 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 Δ.

The Feynman picture

How do we interpret the "variance" of Δ?

- 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 $t \searrow 0$ this measure should be concentrating on the minimal geodesics joining x and y.
- Heuristically, we guess that, as $t \searrow 0$, $\nabla^2 \log p_t(x,y)$ should be the "variance" of minimal geodesics from x to y.

Taylor series of $h_{x,y}$

As before, the Taylor series of $h_{x,y}$ near its minima governs the asymptotics of μ_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 γ_1 non-conjugate and γ_2 and γ_3 each A_3-conjugate. Then $\mu_t \to \mu_0$ with μ_0 supported on the midpoints of γ_2 and γ_3.

Characterizing the cut locus

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

Theorem (N. '08)

Let M be a compact, smooth Riemannian manifold, and let x and y be any two distinct points of M. Then $y \notin \mathrm{Cut}(x)$ if and only if

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

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

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

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