ÉCOLE POLYTECHNIQUE
Programme International
JENDREJ Jacek

RAPPORT DE STAGE DE RECHERCHE

Small time heat kernel asymptotics at a cut-conjugate point on 2-spheres of revolution

NON CONFIDENTIEL

Département:	MATHÉMATIQUES APPLIQUÉES
Champ de l'option:	MAP593 Authomatique et recherche opérationelle
Directeur de l'option:	François GOLSE
Directeur de stage:	Davide BARILARI
Dates du stage:	10/04/2012 – 31/07/2012
Nom et adresse de l'organisme:	École Polytechnique Centre de Mathématiques Appliquées Route de Saclay, 91128 Palaiseau

Résumé

On examine l'asymptotique en temps petit sur le lieu de coupure d'une variété riemannienne. D'abord les développements asymptotiques du noyau de la chaleur sur le cercle, le cylindre 2-dimensionnel et la sphère 2-dimensionnelle sont calculés explicitement. Ces exemples montrent que sur le lieu de coupure le développement asymptotique a une forme différente que hors du lieu de coupure. De plus, l'ordre du premier terme de ce développement change si le point considéré est à la fois dans le lieu de coupure et dans le lieu conjugué.

Après on étudie les 2-sphères de révolution, qui est la classe la plus simple de variétés riemanniennes de dimension 2 ayant une forme générique du lieu de coupure et du lieu conjugué. On détermine l'ordre de dégénération de l'application exponentielle près de leur point commun et on présente des consequences de ce résultat pour l'asymptotique du noyau de la chaleur en temps petit en ce point.

Abstract

We investigate the small time heat kernel asymptotics on the cut locus of Riemannian manifolds. First, we compute explicitly the asymptotic expansion of the heat kernel on the cut locus of the circle, the 2-dimensional cylinder and the 2-sphere. These examples show that at the cut locus the asymptotic expansion has a different form than away from the cut locus and that the order of the leading term of this expansion depends on whether the considered point of the cut locus is also conjugate.

Then we study 2-spheres of revolution, which is the simplest class of 2-dimensional Riemannian manifolds with a generic shape of the cut-conjugate locus. We determine the degeneracy of the exponential map near a cut-conjugate point and present consequences of this result to the small time heat kernel asymptotics at this point.

1 Introductory material

1.1 Basic definitions

Let M be a smooth manifold (where "smooth" always means C^∞). We denote by TM the tangent bundle and by T^*M the cotangent bundle. Both canonical projections will be denoted π, so we have $\pi : TM \to M$ or $\pi : T^*M \to M$, depending on the context. The elements of TM are called *vectors*, whereas the elements of T^*M are called *covectors*. For $q \in M$ we have the fibers $T_qM := \{v \in TM : \pi(v) = q\}$ and $T_q^*M := \{\lambda \in T^*M : \pi(\lambda) = q\}$. When $v \in T_qM$ (respectively $\lambda \in T_q^*M$), q is called the *base point* of v (resp. λ). The natural pairing between T_qM and T_q^*M is denoted $\langle \cdot, \cdot \rangle$.

The space of differential k-forms on M is denoted $\Lambda^k M$ and the space of smooth vector fields on M is denoted $\mathrm{Vec}(M)$. For $X \in \mathrm{Vec}(M)$ the flow of X after time $t \in \mathbb{R}$ is denoted $e^{tX} : M \to M$. We say that $(X_1, \dots, X_n)$ is a *local orthonormal frame* if there exists an open subset $U \subset M$ such that $X_1, \dots, X_n \in \mathrm{Vec}(U)$ and $(X_1(q), \dots, X_n(q))$ is an orthonormal base of T_qM for all $q \in U$. If $v \in T_qM$, $q \in U \subset M$ and $f \in C^\infty(U)$, the directional derivative of f in the direction v is denoted vf. For $V \in \mathrm{Vec}(U)$ and $f \in C^\infty(U)$ we define $C^\infty(U) \ni Vf := q \mapsto V(q)f$.

Let M, N be smooth manifolds. A diffeomorphism $\Phi : M \to N$ induces a linear isomorphism $\Phi_* : \mathrm{Vec}(M) \to \mathrm{Vec}(N)$ defined by the formula $\Phi_*(X)(q) := D\Phi(\Phi^{-1}(q))X(\Phi^{-1}(q))$ for $X \in \mathrm{Vec}(M)$, $q \in N$. A smooth map $\Psi : M \to N$ induces a linear operator $\Psi^* : \Lambda^k N \to \Lambda^k M$ defined by the formula $\Psi^*(\omega)(p)(v_1, \dots, v_k) := \omega(D\Psi(p)v_1, \dots, D\Psi(p)v_k)$ for $\omega \in \Lambda^k M$, $p \in M$, $v_i \in T_pM$.

Definition 1.1. Let M be a smooth manifold, $X, Y \in \mathrm{Vec}(M)$ and $\omega \in \Lambda^k M$. We define the *Lie derivative* of Y with respect to X by the formula

$$\mathrm{Vec}(M) \ni L_X Y := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(e^{-tX}\right)_* Y.$$

We define the *Lie derivative* of ω with respect to X by the formula

$$\Lambda^k M \ni L_X \omega := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(e^{tX}\right)^* \omega.$$

It can be shown that for $f \in C^\infty(M)$ we have $(L_X Y)f = [X, Y]f = X(Yf) - Y(Xf)$.

Definition 1.2. Let $f \in C^\infty(U)$, $U \subset M$, and let x be a critical point of f. The *Hessian* of f is a bilinear form $\mathrm{Hess}\, f(x) : T_xM \times T_xM \to \mathbb{R}$ defined by

$$\mathrm{Hess}\, f(x)(v, w) = W(Vf)(x),$$

where $V, W \in \mathrm{Vec}(U)$ are arbitrary vector fields such that $V(x) = v$ and $W(x) = w$.

It is necessary to check that this definition does not depend on the choice of V and W. It is clear that $W(Vf)(x)$ does not depend on the choice of W. Moreover, $W(Vf)(x) =$

$V(Wf)(x) - [V,W]f(x) = V(Wf)(x)$, because x is a critical point of f. Thus $W(Vf)(x)$ does not depend on the choice of V. Also, we see that $\operatorname{Hess} f(x)$ is a symmetric form.

If the Hessian of f at a point $x \in M$ is a degenerate form, x is called a *degenerate critical point* of f.

Let g be a Riemannian metric on M, so that M becomes a Riemannian manifold. If $v \in T_qM$, we say that $\lambda \in T_q^*M$ is *associated* to v if $\lambda(w) = g(v,w)$ for every $w \in T_qM$. This defines an isomorphism between T_qM and T_q^*M. For a smooth function $f : M \supset V \to \mathbb{R}$ the *gradient* of f at a point $q \in V$ is defined as the vector in T_qM associated with $\mathrm{d}f(q) \in T_q^*M$. It will be denoted $\nabla f(q)$.

A *regular curve* on M is a C^1 function $c : [a,b] \to M$ such that $\dot{c}(t) \neq 0$ for $t \in [a,b]$. A *curve* on M is a function $c : [a,b] \to M$ for which there exists a finite sequence $a = t_0 < t_1 < \cdots < t_n = b$ such that $c|_{[t_i,t_{i+1}]}$ is a regular curve. A smooth curve is a regular curve which is C^∞.

Definition 1.3. Let $c : [a,b] \to M$ be a curve.

The *length* of c is defined by

$$L(c) := \int_a^b |\dot{c}(t)| \, \mathrm{d}t.$$

The *energy* of c is defined by

$$E(c) := \int_a^b |\dot{c}(t)|^2 \, \mathrm{d}t.$$

If $|\dot{c}(t)| = 1$ for almost every t, we say that the curve c is *parametrized by arc length.* It is immediate that L is invariant by reparametrizations, whereas E is not – it attains its minimum precisely when c is parametrized proportional to arc length.

Definition 1.4. Let $p, q \in M$. Let $\mathcal{C}_{p,q}$ be the set of piecewise C^1 curves $c : [a,b] \to M$ such that $c(a) = p$ and $c(b) = q$. The *Riemannian distance* on M is defined as

$$\operatorname{dist}(p,q) := \inf_{c \in \mathcal{C}_{p,q}} L(c).$$

It can be shown that (M, dist) is a metric space (the part $p \neq q \Rightarrow \operatorname{dist}(p,q) > 0$ is nontrivial) and the topology induced by dist coïncides with the manifold topology [9, Definition-Proposition 2.91].

1.2 Geodesics – Hamiltonian approach

The construction of a Hamiltonian vector field presented here is taken from [1, Chapter 4].

Definition 1.5. The *tautological* 1*-form* on T^*M is the form $s \in \Lambda^1(T^*M)$ defined by

$$\langle s(\lambda), w \rangle := \langle \lambda, \pi_* w \rangle$$

for $w \in T_\lambda(T^*M)$.

Definition 1.6. The 2-form $\sigma := \mathrm{d}s$ is called the *canonical symplectic structure* on T^*M.

Let $(x_1, \ldots, x_n, \xi_1, \ldots, \xi_n)$ be a canonical system of coordinates on T^*M, which means that if $q = (x_1, \ldots, x_n)$ and $\lambda = \sum_i \xi_i \,\mathrm{d}x_i \in T_q^*M$, then λ has the coordinate expression $(x_1, \ldots, x_n, \xi_1, \ldots, \xi_n)$. Let $\lambda = \sum_i \xi_i \,\mathrm{d}x_i \in T_q^*M$. It is easily checked that in canonical coordinates we have $s(\lambda) = \sum_i \xi_i \,\mathrm{d}x_i \in T_\lambda(T^*M)$. It is the same expression as for λ, but now $\mathrm{d}x_i$ is an element of $\Lambda^1(T^*M)$. Thus, by differentiation, the expression of the canonical symplectic structure on T^*M is

$$\sigma = \sum_{i=1}^{n} \mathrm{d}\xi_i \wedge \mathrm{d}x_i. \tag{1}$$

In particular σ is a non-degenerate form, which justifies the next definition.

Definition 1.7. Let $h \in C^\infty(T^*M)$. The *Hamiltonian vector field* associated with h is the unique vector field $\vec{h} \in \mathrm{Vec}(T^*M)$ satisfying the identity

$$\sigma(\cdot, \vec{h}) = \mathrm{d}h(\cdot).$$

Let $\vec{h} = \sum_i a_i \partial_{x_i} + \alpha_i \partial_{\xi_i}$ be the expression of the Hamiltonian vector field in canonical coordinates. We have $a_i = \sigma(\partial_{\xi_i}, \vec{h}) = \mathrm{d}h(\partial_{\xi_i}) = \frac{\partial h}{\partial \xi_i}$ and $\alpha_i = -\sigma(\partial_{x_i}, \vec{h}) = -\mathrm{d}h(\partial_{x_i}) = -\frac{\partial h}{\partial x_i}$, which leads to the formula

$$\vec{h} = \sum_{i=1}^{n} \frac{\partial h}{\partial \xi_i} \partial_{x_i} - \frac{\partial h}{\partial x_i} \partial_{\xi_i}. \tag{2}$$

For $\lambda \in T_q^*M$ we define the norm

$$\|\lambda\| := \sup_{\substack{v \in T_qM \\ g(v,v)=1}} |\langle \lambda, v \rangle|.$$

We define $H : T^*M \to \mathbb{R}$ by

$$H(\lambda) := \frac{1}{2}\|\lambda\|^2. \tag{3}$$

It is clear that $H \in C^\infty(T^*M)$.

Definition 1.8. A regular curve $\gamma : (a, b) \to M$ is called a *geodesic* if $\gamma(t) = \pi \circ \lambda(t)$, where $\lambda : (a, b) \to T^*M$ is a solution of the Hamiltonian equation $\dot{\lambda}(t) = \vec{H}(\lambda(t))$.

Notice that by the theorem of local existence and uniqueness of solutions of ordinary differential equations, for every $\lambda_0 \in T^*M$ there exists a unique maximal solution $\lambda : (a, b) \to T^*M$ of the Hamiltonian equation such that $\lambda(0) = \lambda_0$.

A curve $c : [a, b] \to M$ is called *minimal* if $L(c) = \mathrm{dist}(c(a), c(b))$. It is immediate that a minimal curve is locally minimal. Notice also that minimal curves joining two given points do not always exist (cf. Theorem 1.17).

Proposition 1.9. *Let $c : [0, T] \to M$ be a minimal curve from $c(0) = q_0$ to $c(T) = q_1$. Then c is a geodesic.*

This is a consequence of the following Pontryagin Maximum Principle for optimal control problems [2, Theorem 12.3].

Theorem 1.10 (PMP)**.** *Let M be a manifold and consider a control system*

$$\begin{aligned} &\dot{q}(t) = f_u(q(t)), && q(t) \in M, \quad u(t) \in U \subset \mathbb{R}^n, \\ &q(0) = q_0, q(T) = q_1, && q_0, q_1 \in M, \quad T \in \mathbb{R}_+ \end{aligned} \tag{4}$$

with the cost functional

$$J(u) = \int_0^T \phi(q_u(t), u(t))\,\mathrm{d}t, \tag{5}$$

where q_u denotes the trajectory associated with the control u.

For $\lambda \in T_q^*M, u \in U, \mu \in \mathbb{R}$ let

$$\mathcal{H}_u^\mu(\lambda) = \langle \lambda, f_u \rangle + \mu\phi(q, u). \tag{6}$$

Let $\widetilde{u} : [0, T] \to U$ be an optimal control and let $\widetilde{q} : [0, T] \to M$ be the associated trajectory. Then there exist $\mu \leq 0$ and $\lambda : [0, T] \to T^*M$ such that $(\mu, \lambda(t)) \neq 0$ for all t and

$$\dot{\lambda}(t) = \vec{\mathcal{H}}^\mu_{\widetilde{u}(t)}(\lambda(t)), \tag{7}$$

$$\mathcal{H}^\mu_{\widetilde{u}(t)}(\lambda(t)) = \max_{v \in U} \mathcal{H}^\mu_v(\lambda(t)). \tag{8}$$

Remark 1.11. If $U = \mathbb{R}^n$, condition (8) can be replaced by a weaker condition

$$\frac{\partial}{\partial u}\mathcal{H}^\mu_{\widetilde{u}(t)}(\lambda(t)) = 0. \tag{9}$$

This is called the Weak Pontryagin Maximum Principle.

Proof of Proposition 1.9. Let $(X_1, \ldots, X_n)$ be a local orthonormal frame on M. The problem of finding minimal curves can be stated (locally) as an optimal control problem $\dot{c}(t) = \sum_{i=1}^n u_i(t) X_i(c(t))$ with a control $u(t) \in \mathbb{R}^n$ and a cost functional $J(u) = \int_0^T |u(t)|^2\,\mathrm{d}t$.

In this setting we have

$$\mathcal{H}^\mu_u(\lambda) = \sum_{i=1}^n u_i \langle \lambda, X_i \rangle + \mu |u|^2.$$

Now apply the PMP. Suppose that $\mu = 0$. Equation (9) gives $\langle \lambda(t), X_i(t) \rangle = -2\mu u_i(t) = 0$, so $\lambda(t) = 0$, which is a contradiction with $(\mu, \lambda(t)) \neq 0$. Thus $\mu < 0$ and we can normalize in such a way that $\mu = -\frac{1}{2}$, which gives

$$u_i(t) = \langle \lambda(t), X_i(t) \rangle\,. \tag{10}$$

Notice that for $\lambda \in T_q^*M$ and $Y \in T_qM$ we have

$$\langle \lambda, Y \rangle = \sum_{i=1}^n g(Y, X_i) \langle \lambda, X_i \rangle \leq \left(g(Y, Y) \sum_{i=1}^n \langle \lambda, X_i \rangle^2 \right)^{1/2},$$

and equality holds for $Y = \sum_i \langle \lambda, X_i \rangle X_i$. Thus $\|\lambda\|^2 = \sum \langle \lambda, X_i \rangle^2$ and we get

$$\mathcal{H}^{\mu}_{u(t)}(\lambda(t)) = \sum_{i=1}^{n} \langle \lambda(t), X_i(t) \rangle^2 - \frac{1}{2} \langle \lambda(t), X_i(t) \rangle^2 = \frac{1}{2} \sum_{i=1}^{n} \langle \lambda(t), X_i(t) \rangle^2 = \frac{1}{2} \|\lambda(t)\|^2,$$

so equation (7) states exactly that $\lambda(t)$ is the required lift of $c(t)$. □

Remark 1.12. From equation (10) we obtain $g(\dot{c}(t), X_i) = u_i(t) = \langle \lambda(t), X_i \rangle$. Thus $\lambda(t)$ is the covector associated with $\dot{c}(t)$. Thus for Riemannian manifolds parameterizing geodesics by covectors is equivalent to parameterizing by vectors. In particular for every $q \in M$ and $v \in T_q M$ there exists exactly one maximal geodesic $\gamma : (a, b) \to M$ such that $\gamma(0) = q$ and $\gamma'(0) = v$. This geodesic will be denoted γ_v.

However, in sub-Riemannian geometry there is no canonical isomorphism between $T_q M$ and $T_q^* M$ and the natural approach is parameterizing geodesics by covectors (see [1]).

We will show that Definition 1.8 is equivalent to the classical definition of a geodesic as a regular curve $\gamma : (a, b) \to M$ satisfying the equation $D_{\dot{\gamma}(t)} \dot{\gamma}(t) = 0$, where D is the covariant derivative associated with the Riemannian metric g.

Fix on M a system of local coordinates $(x_1, \dots, x_n)$. Let $G = (g_{ij})$ be the matrix of the first fundamental form in these coordinates, i.e. $g_{ij} = g(\partial_{x_i}, \partial_{x_j})$. Let $G^{-1} = (g^{ij})$. The *Christoffel symbols* are defined by the formula

$$\Gamma^i_{jk} := \frac{1}{2} \sum_l g^{il} (\partial_{x_j} g_{kl} + \partial_{x_k} g_{lj} - \partial_{x_l} g_{jk}), \qquad i, j, k \in \{1, \dots, n\}.$$

Proposition 1.13. *A regular curve $\gamma : (a, b) \to M, \gamma(t) = (x_1(t), \dots, x_n(t))$ is a geodesic if and only if it satisfies the system of differential equations*

$$\ddot{x}_i + \sum_{j,k=1}^{n} \Gamma^i_{jk} \dot{x}_j \dot{x}_k = 0, \qquad i = 1, \dots, n. \tag{11}$$

Proof. Consider the canonical coordinates $(x_1, \dots, x_n, \xi_1, \dots, \xi_n)$ on T^*M. Let $\lambda = (x, \xi)$ be a covector, where $x = (x_1, \dots, x_n)$ and $\xi = (\xi_1, \dots, \xi_n)$ are considered as column vectors. Let B be a positive-definite matrix such that $G = B^2$. We have

$$\|\lambda\| = \sup_{|Bx|=1} (B^{-1}\xi)^T (Bx) = |B^{-1}\xi| = (\xi^T G^{-1} \xi)^{1/2}.$$

Hence $H(\lambda) = \frac{1}{2} \xi^T G^{-1} \xi$.

We will denote $G_{x_i} := \partial_{x_i} G$. Notice that $\partial_{x_i}(G^{-1}) = -G^{-1} G_{x_i} G^{-1}$. Thus the Hamiltonian equations can be written as

$$\dot{x} = \frac{\partial H}{\partial \xi} = G^{-1} \xi \tag{12}$$

$$\dot{\xi}_i = -\frac{\partial H}{\partial x_i} = -\frac{1}{2} \xi^T G^{-1} G_{x_i} G^{-1} \xi = -\frac{1}{2} \dot{x}^T G_{x_i} \dot{x}, \tag{13}$$

where the third equality in (13) follows from (12).

Differentiating (12) gives

$$\ddot{x} = G^{-1}\dot{\xi} - G^{-1}\dot{G}G^{-1}\xi = \frac{1}{2}G^{-1}\dot{\xi} - G^{-1}\sum_{j=1}^{n}\dot{x}_j G_{x_j}\dot{x}.$$

Now we have

$$\left(G^{-1}\dot{\xi}\right)_i = \sum_{l=1}^{n} g^{il}\dot{x}^T G_{x_l}\dot{x} = \sum_{j,k=1}^{n}\sum_{l=1}^{n} g^{il}\partial_{x_l}g_{jk}\dot{x}_j\dot{x}_k,$$

$$\left(G^{-1}\dot{x}_j G_{x_j}\dot{x}\right)_i = \dot{x}_j\sum_{l=1}^{n} g^{il}\sum_{k=1}^{n}\partial_{x_j}g_{kl}\dot{x}_k = \sum_{k=1}^{n}\sum_{l=1}^{n} g^{il}\partial_{x_j}g_{kl}\dot{x}_j\dot{x}_k.$$

Notice that

$$\sum_{j,k=1}^{n}\sum_{l=1}^{n} g^{il}\partial_{x_j}g_{kl}\dot{x}_j\dot{x}_k = \frac{1}{2}\sum_{j,k=1}^{n}\sum_{l=1}^{n} g^{il}(\partial_{x_j}g_{kl} + \partial_{x_k}g_{lj})\dot{x}_j\dot{x}_k.$$

Hence, using the definition of the Christoffel symbols, we see that (13) and (11) is in fact the same equation. □

If for each $v \in TM$ the corresponding geodesic γ_v is defined on whole $\mathbb{R}$, the manifold M is called *complete*.

Definition 1.14. Let (M, g) be a complete Riemannian manifold. The *exponential map* $\exp : TM \to M$ is defined by

$$\exp(v) := \gamma_v(1).$$

If $p \in M$, $\exp_p : T_pM \to M$ is the restriction of $\exp : TM \to M$ to T_pM.

Remark 1.15. In sub-Riemannian geometry exp is defined on T^*M (cf. Remark 1.12).

Theorem 1.16. *The map* $\exp : TM \to M$ *is smooth. For* $v \in TM$ *define* $\Phi(v) = (\pi(v), \exp(v)) \in M \times M$. *Then for every* $p \in M$ *the map* Φ *is a diffeomorphism from a neighborhood of* $0_p \in TM$ *onto a neighborhood of* $(p, p) \in M \times M$.

Proof. The first part follows from the theorem on smooth dependence of solutions of ordinary differential equations on the initial data.

The second part follows from the inverse function theorem, cf. [9, Proposition 2.88]. □

In particular for $p \in M$ there exists an $\epsilon > 0$ such that $\exp_p : T_pM \supset B(0, \epsilon) \to B' \subset M$ is a diffeomorphism. This defines a local system of coordinates called *local geodesic coordinates*. Using these coordinates one can prove that geodesics are locally minimal [9, Theorem 2.92], which together with Proposition 1.9 provides an alternative definition of geodesics as locally minimal curves.

The following two results, called Hopf-Rinow Theorems, build a connection between completeness and completeness as a metric space.

Theorem 1.17 (Hopf-Rinow [12]). *Any two points of a complete connected manifold can be joined by a minimal geodesic.*

Theorem 1.18 (Hopf-Rinow [12]). *A Riemannian manifold M is complete if and only if it is complete as a metric space with the Riemannian distance.*

Each of these theorems is a relatively simple consequence of the other. Theorem 1.18 follows from Theorem 1.17 and the criterion for a finite time blowup of an ODE. Theorem 1.17 follows from Theorem 1.18 together with, for example, Filippov Theorem on the existence of optimal controls [1, Theorem 3.25]. A short proof of Theorem 1.17, using Proposition 1.9, was given by de Rham in [6, p. 341–343].

In the sequel we always assume that a manifold M is complete.

Definition 1.19. Let $p \in M$ and $v \in T_pM$. If $\mathrm{d}\exp_p(v) : T_v(T_pM) \to T_pM$ is singular, the point $\exp_p(v)$ is called *conjugate* to p along γ_v. We define the *conjugate locus* as

$$\mathrm{Conj}(p) := \{q \in M : q \text{ is conjugate to } p \text{ along some geodesic}\}.$$

1.3 Cut locus

In general geodesics fail to be globally optimal, which leads to the notion of the *cut locus*.

Recall that (M, g) is assumed to be complete.

Definition 1.20. Let $v \in TM$ we define the *cut time* of the geodesic $\gamma_v : \mathbb{R} \to M$ as

$$t_c(v) := \sup\{t \in \mathbb{R} : \gamma|_{[0,t]} \text{ is minimal}\}.$$

Observe that if $w = av$, where $a \in \mathbb{R}$, then $t_c(w) = \frac{1}{|a|}v$. It can be proved that $t_c : T_pM \to \overline{\mathbb{R}}_+$ is a continuous function [16, Theorem 7.3].

For $p \in M$ we denote $U_p := \{v \in T_pM : t_c(v) > 1\}$. It follows from the continuity of t_c that $T_pM \supset U_p$ is an open neighborhood of the origin in T_pM.

Definition 1.21. Let $p \in M$. The *cut locus* of p is

$$\mathrm{Cut}(p) := \exp(\partial U_p).$$

The cut locus is thus the set of points where geodesics emanating from p lose global optimality.

Denote $V_p := \exp(U_p)$. In the following theorem we combine Proposition 2.113 and Corollary 3.77 from [9].

Theorem 1.22. *The manifold M is a disjoint union of V_p and* $\mathrm{Cut}(p)$. *The map* $\exp_p : U_p \to V_p$ *is a diffeomorphism.*

This theorem explains the name "cut locus". It turns out that after "cutting" the manifold along $\mathrm{Cut}(p)$, it can be diffeomorphically mapped onto a flat star-like open set $U_p \subset T_pM$.

The following proposition is a direct consequence of the definition of U_p.

Proposition 1.23. *Let $p, q \in M$ and let $\gamma : [0,1] \to M$ be a minimal geodesic from $\gamma(0) = p$ to $\gamma(1) = q$. Then $\gamma([0,1)) \subset V_p$.*

Remark 1.24. Notice that here we do not assume that $q \in V_p$.

Proof. Let $v \in T_pM$ be such that $\gamma = \gamma_v$. We assumed that $t_c(v) \geq 1$, so $av \in U_p$ for $0 \leq a < 1$. Hence $\gamma([0,1)) = \{\exp av : a \in [0,1)\} \subset \exp U_p = V_p$. □

Finally we have the following easy consequence of Theorem 1.22.

Corollary 1.25. *Let $p \in M$ and $q \in V_p$. Then there exists exactly one unit-speed minimal geodesic joining p and q.*

Proof. Let $v \in T_pM$ be such that $\gamma_v : [0,1] \to M$ is a minimal geodesic joining $\gamma(0) = p$ and $\gamma(1) = q$. Of course $t_c(v) \geq 1$. By assumption $\gamma(1) \notin \mathrm{Cut}(p)$, so $t_c(v) \neq 1$. Thus $t_c(v) > 1$, which means by definition that $v \in U_p$. Hence $v = \left(\exp|_{U_p}\right)^{-1}(q)$ is unique. □

1.4 Energy and hinged energy

Definition 1.26. Let $p \in M$. We define the *energy function* $E_p : M \to \mathbb{R}$ by

$$E_p(q) := \frac{1}{2}\mathrm{dist}(p,q)^2,$$

where dist is the Riemannian distance.

Proposition 1.27. *Let $q \in V_p$ and let $v = \left(\exp|_{U_p}\right)^{-1}(q)$. Then $\nabla E_p(q) = \gamma_v'(1) \in T_pM$.*

Proof. Let $w \in T_v(T_pM)$. Then we have [7, p. 367, Lemma 2]

$$g(v,w) = g\left(\mathrm{d}\exp_p(v)v, \mathrm{d}\exp_p(v)w\right), \tag{14}$$

where we identify $T_pM \sim T_v(T_pM)$.

Notice that $\exp_p(v + tv) = \exp_p((1+t)v) = \gamma_v(1+t)$. Hence, $\mathrm{d}\exp_p(v)v = \gamma_v'(1)$.

Let $\widetilde{q} \in V_p$ and let $\widetilde{v} = \left(\exp|_{U_p}\right)^{-1}(\widetilde{q})$. Then $E_p(\widetilde{q}) = \frac{1}{2}g(\widetilde{v},\widetilde{v})$. Hence, for $u \in T_qM$ we obtain, by the rule of differentiating quadratic functions,

$$\langle \mathrm{d}E_p(q), u\rangle = g\left(v, \mathrm{d}\left(\exp|_{U_p}\right)^{-1}(q)u\right).$$

Thus, substituting in (14) $w = \mathrm{d}\left(\exp|_{U_p}\right)^{-1}(q)u$, we get

$$\langle \mathrm{d}E_p(q), u\rangle = g\left(\gamma_v'(1), u\right).$$

□

Remark 1.28. From the formula $E_p(\widetilde{q}) = \frac{1}{2}g(\widetilde{v},\widetilde{v})$ and the fact that $\exp|_{U_p}$ is smooth it follows that $E_p : M \to \mathbb{R}$ is smooth on V_p.

Definition 1.29. Let $p, q \in M$. We define the *hinged energy function* $h_{p,q} : M \to \mathbb{R}$ by the formula

$$h_{p,q}(z) := E_p(z) + E_q(z).$$

Observe that $h_{p,q}$ is smooth on $V_p \cap V_q$.

Proposition 1.30. *Let $p, q \in M$ and let $\gamma : [0, 2] \to M$ be a minimal geodesic from $\gamma(0) = p$ to $\gamma(2) = q$. Then $\min h_{p,q} = \frac{1}{4}\mathrm{dist}(p, q)^2$ and this minimum is attained at the point $x := \gamma(1)$. Every global minimum of $h_{p,q}$ is a midpoint of some minimal geodesic joining p and q.*

Proof. Let $z \in M$, $a := \mathrm{dist}(p, x)$, $b := \mathrm{dist}(x, q)$, $\alpha := \mathrm{dist}(p, z)$, $\beta := \mathrm{dist}(z, q)$. Then $a = b = \frac{1}{2}\mathrm{dist}(p, q)$, so $h_{p,q}(x) = \frac{1}{2}a^2 + \frac{1}{2}b^2 = \frac{1}{4}\mathrm{dist}(p, q)^2$.

From the triangle inequality we have $\alpha + \beta \geq \mathrm{dist}(p, q)$, so

$$h_{p,q}(z) = \frac{\alpha^2 + \beta^2}{2} \geq \left(\frac{\alpha + \beta}{2}\right)^2 \geq \frac{1}{4}\mathrm{dist}(p, q)^2.$$

Suppose that the equality holds. From $\frac{\alpha^2+\beta^2}{2} = \left(\frac{\alpha+\beta}{2}\right)^2$ we infer that $\alpha = \beta$. From $\alpha + \beta = \mathrm{dist}(p, q)$ we deduce that z lies on a minimal curve from p to q. This curve is a geodesic by Proposition 1.9, and z is its midpoint. □

Remark 1.31. Notice that the midpoint of a minimal geodesic lies in the "good" set $V_p \cap V_q$ even if $q \in \mathrm{Cut}(p)$.

Remark 1.32. Take in particular $z := \gamma(1 + t)$. Then $h_{p,q}(z) = \frac{1}{2}a^2((1 - t)^2 + (1 + t)^2) = \frac{1}{4}\mathrm{dist}(p, q)^2(1 + t^2)$. We obtain that $\mathrm{Hess}\, h_{p,q}(z)$ is non-degenerate in the direction $\dot\gamma(1)$.

The reason to study the hinged energy function here is that the degeneracy of $\exp_p$ near q is reflected by the behavior of h near the midpoint of a geodesic joining them, which is assumed to lie in the "good" region. This relationship is given in the next proposition.

Proposition 1.33. *Let $p, q \in M$ and let $\alpha : (-\epsilon, \epsilon) \to V_p \cap V_q$ be a smooth curve. Assume that $x = \alpha(0)$ is a critical point of $h_{p,q}$. Suppose that $k \in \mathbb{N}$ is such that $\nabla h_{p,q}(\alpha(t)) = t^k u(t)$, where u is a smooth vector field along the curve α with $u(0) \neq 0$.*

Let $\gamma_{v(t)} : \mathbb{R} \to M$ be the minimal geodesic from $\gamma_{v(t)}(0) = p$ to $\gamma_{v(t)}(1) = \alpha(t)$. Let $\beta(t) := \gamma_{v(t)}(2)$.

Then $\beta : (-\epsilon, \epsilon) \to M$ is a smooth function, $\beta(0) = q$ and

$$\lim_{t \to 0} \frac{\dot\beta(t)}{t^{k-1}} = k \, \mathrm{d}\exp_x\left(\dot\gamma_{v(0)}(1)\right) u(0),$$

where in the last formula we identify $T_{\alpha(0)}M$ with $T_{\dot\gamma_{v(0)}(1)}(T_{\alpha(0)}M)$.

Remark 1.34. Observe that under our assumptions $v(t)$ is a smooth curve $(-\epsilon, \epsilon) \to T_pM$ and every such curve $v(t)$ gives rise to the corresponding curve $\alpha(t) := \gamma_{v(t)}(1)$.

Proof. Most of this argument is taken from [4, proof of Theorem 21.]

First part is obvious, since β is a composition of smooth functions.

Let $\gamma_{w(t)} : \mathbb{R} \to M$ be the optimal geodesic from $\gamma_{w(t)}(0) = q$ to $\gamma_{w(t)}(1) = \alpha(t)$. Denote $X(t) := \nabla h_{p,q}(\alpha(t)) \in T_{\alpha(t)}M$. It follows from Proposition 1.27 that

$$\dot\gamma_{v(t)}(1) = -\dot\gamma_{w(t)}(1) + X(t). \tag{15}$$

We want to use the Taylor expansion of exp around $\dot\gamma_{v(0)}(1) = -\dot\gamma_{w(0)}(1)$.

To this end we introduce on TM local canonical coordinates (x_1, x_2, v_1, v_2) around $\alpha(0)$ and (q_1, q_2, w_1, w_2) around q. In these coordinates we write

$$\begin{aligned}
\dot\gamma_{v(0)}(1) = -\dot\gamma_{w(0)}(1) &= (0, 0, c_1, c_2),\\
\dot\gamma_{v(t)}(1) &= (x_1(t), x_2(t), c_1 + a_1(t), c_2 + a_2(t)),\\
-\dot\gamma_{w(t)}(1) &= (x_1(t), x_2(t), c_1 + b_1(t), c_2 + b_2(t)),\\
u(t) &= (x_1(t), x_2(t), u_1(t), u_2(t)),\\
\beta(t) &= (q_1(t), q_2(t)).
\end{aligned}$$

Let

$$\begin{aligned}
\exp(q_1, q_2, c_1 + h_1, c_2 + h_2) = \mathrm{d}\exp_{(0,0,c_1,c_2)}&(q_1, q_2, h_1, h_2)^t\\
&+ \phi_2\left((q_1, q_2, h_1, h_2)^t\right)^2 + \cdots +\\
&+ \phi_k\left((q_1, q_2, h_1, h_2)^t\right)^k + O(\|(q_1, q_2, h_1, h_2)\|^{k+1})
\end{aligned}$$

be the Taylor expansion of exp near $\dot\gamma_{v(0)}(1) = (0, 0, c_1, c_2)$ (here ϕ_j is some j-linear map for $j = 2, \ldots, k$).

Observe that $\exp\left(-\dot\gamma_{w(t)}(1)\right) = q = (0, 0)$ and $\exp\left(\dot\gamma_{v(t)}(1)\right) = \beta(t) = (q_1(t), q_2(t))$. Also, for $j = 2, \ldots, k$ we have

$$\left\|\phi_j\left((q_1(t), q_2(t), a_1(t), a_2(t))^t\right)^j - \phi_j\left((q_1(t), q_2(t), b_1(t), b_2(t))^t\right)^j\right\| = O(t^{k+1}).$$

Hence

$$(q_1(t), q_2(t)) = \mathrm{d}\exp_{(0,0,c_1,c_2)}(0, 0, a_1(t) - b_1(t), a_2(t) - b_2(t)) + O(t^{k+1}).$$

By (15) we have $a_i(t) - b_i(t) = t^k u_i(t)$. Thus

$$(q_1(t), q_2(t)) = t^k \mathrm{d}\exp_{(0,0,c_1,c_2)}(0, 0, u_1(t), u_2(t)) + O(t^{k+1}),$$

so in coordinates (w_1, w_2) the derivative $\dot\beta(t)$ is expressed as

$$\begin{aligned}
(\dot q_1(t), \dot q_2(t)) &= kt^{k-1}\mathrm{d}\exp_{(0,0,c_1,c_2)}(0, 0, u_1(t), u_2(t)) + O(t^k)\\
&= kt^{k-1}\mathrm{d}\exp_{(0,0,c_1,c_2)}(0, 0, u_1(0), u_2(0)) + O(t^k)\\
&= kt^{k-1}\mathrm{d}\exp_{\dot\gamma_{v(0)}(1)}(u(0)) + O(t^k),
\end{aligned}$$

and the conclusion follows. $\square$

Corollary 1.35. *Under the assumptions above $\alpha(0)$ is the midpoint of some (not necessarily minimal) geodesic joining p and q.*

Corollary 1.36. *Let $p, q \in M$ and let $x \in M$ be a critical point of $h_{p,q}$. Assume additionally that $x \in V_p \cap V_q$. Then q is conjugate to p along some geodesic if and only if x is a degenerate critical point of $h_{p,q}$.*

Proof. Suppose first that q is conjugate to p along $\gamma_{v(0)}$. Then there exists a smooth curve $v : (-\epsilon, \epsilon) \to T_pM$ such that $\dot{\beta}(0) = 0$. Let $\alpha : (-\epsilon, \epsilon) \to M$ be the corresponding curve of midpoints (see Remark 1.34). Then in the last proposition we must have $k \geq 2$. Thus $\operatorname{Hess} h_{p,q}(\alpha(0))(v, \dot{\alpha}(0)) = 0$ for any v.

Conversely, if $\operatorname{Hess} h_{p,q}(x)(v, w) = 0$ for all v, it suffices to take an arbitrary regular curve $\alpha : (-\epsilon, \epsilon) \to V_p \cap V_q$ such that $\alpha(0) = x$ and $\dot{\alpha}(0) = w$ to obtain $\dot{\beta}(0) = 0$. $\square$

Note. In the rest of this section we assume additionally that M has dimension 2.

Let $p \in M$ and let $\gamma : \mathbb{R} \to M$ be a geodesic with $\gamma(0) = p$. Suppose that $\gamma(2) = q$ is a conjugate point of p along γ and that $\gamma|_{[0,2]}$ is minimal (such a point is called a *cut-conjugate point*).

It follows from Proposition 1.30 and Corollary 1.36 that $x := \gamma(1)$ is a global minimum of $h_{p,q}$ and a degenerate critical point. Notice however that, according to Remark 1.32, the degeneracy is only in one direction. The following result, called the Splitting Lemma or the Refined Morse Lemma, is a special case of [11, Lemma 1].

Lemma 1.37. *Let $U \subset \mathbb{R}^n$ be open and let $x \in U$ be a local minimum of a smooth function $f : U \to \mathbb{R}$. Suppose that the critical point x is degenerate in one direction. Then there exists a smooth local coordinate system $(z_1, \ldots, z_{n-1}, z_n)$ such that in some neighborhood of x we have $f((z_1, \ldots, z_{n-1}, z_n)) = f(x) + z_1^2 + \cdots + z_{n-1}^2 + g(z_n)$, where g is a smooth function $\mathbb{R} \to \mathbb{R}_+$ and $g(z) = O(z^4)$.*

Remark 1.38. The function g is not unique, but its order of vanishing at $z = 0$ is – this is the maximal order of vanishing at $z = 0$ of functions $z \mapsto f(\alpha(z)) - f(x)$ for a smooth curve α such that $\alpha(0) = x$.

Applying this lemma to $f = h_{p,q}$ we obtain that there exists a smooth local coordinate system (z_1, z_2) near x such that $h_{p,q}(z_1, z_2) = h_{p,q}(x) + z_1^2 + g(z_2)$, where $g : \mathbb{R} \to \mathbb{R}_+$ is a smooth function and $g(z) = O(z^4)$. Let $k + 1 \in \mathbb{N}$ be the order of vanishing of g at $z = 0$ (thus $k \geq 3$).

We define in these coordinates the smooth curve $\alpha(t) := (0, t)$. It is immediate that $\nabla h_{p,q}(\alpha(t)) = t^k u(t)$ for some smooth vector field u with $u(0) \neq 0$. Combining this with Proposition 1.33 and Remark 1.34 we obtain the following result.

Corollary 1.39. *Let $p, q \in M$ and assume that q is a cut-conjugate point of p along $\gamma : [0, 2] \to M$. Then there exists a smooth curve $v : (-\epsilon, \epsilon) \to T_pM$ such that $\gamma_{v(0)} = \gamma$ and*

$$\operatorname{dist}(\gamma_{v(t)}(2), q) = O(t^k),$$

where $k + 1$ is the order of vanishing at 0 of the function g described above.

It is natural to expect that for a generic metric the order of degeneracy is lowest possible, that is $k+1=4$. The last corollary gives an upper bound on k once we prove that geodesics close to the minimal one are not "too close" to the cut-conjugate point. We expect that in a generic situation these two bounds coïncide, hence determining the value of k (which would be equal 3).

In the next section we will see that the order of vanishing of the function g is related to the small-time asymptotics of the heat equation on M.

1.5 The heat kernel and its asymptotics

Let (M,g) be a complete orientable n-dimensional Riemannian manifold. We denote dvol the volume form associated with g and compatible with the orientation (which means it equals 1 on a positively oriented orthonormal frame).

Definition 1.40. Let $X \in \mathrm{Vec}(M)$. We define the *divergence* $\mathrm{div} X$ as the unique function satisfying

$$L_X \,\mathrm{dvol} = \mathrm{div} X \,\mathrm{dvol}.$$

The Laplace operator on smooth functions on M is defined as $\Delta f := \mathrm{div}\,(\nabla f)$ (sometimes, especially in differential geometry, Δ is defined as $-\mathrm{div}\nabla$).

Definition 1.41. A *heat kernel* on M is a function $e(t,p,q) \in C^\infty(\mathbb{R}_+ \times M \times M)$ satisfying the following conditions.

$$(\partial_t - \Delta_p) e(t,p,q) = 0, \tag{16}$$

$$\lim_{t\to 0} \int_M e(t,p,q) f(q) \,\mathrm{dvol}(q) = f(p), \ \forall f \in C^\infty(M). \tag{17}$$

Theorem 1.42. *A heat kernel exists and is unique. It satisfies $e(t,p,q) = e(t,q,p)$.*

Proof. Cf. [10, Theorem 7.13]. □

Definition 1.43. Let $e(t,p,q)$ be a heat kernel. The *heat operator* $e^{t\Delta} : L^2(M) \to L^2(M)$ is defined by the formula

$$e^{t\Delta} f(p) = \int_M e(t,p,q) f(q) \,\mathrm{dvol}(q).$$

Remark 1.44. Notice that $\psi(t,p) = e^{t\Delta} f(p)$ is the solution of the Cauchy problem

$$\begin{aligned} \psi(0,p) &= f(p) \\ \partial_t \psi(t,p) &= \Delta_p \psi(t,p). \end{aligned}$$

These are classical results of the Hodge theory that if M is compact then the operators Δ and $e^{t\Delta}$ are simultaneously diagonalizable and the proper vectors are smooth functions. Their spectra determine the long-time behavior of the heat flow and contain topological information.

Here we investigate the short-time behavior, which appears to be connected to the geometry of the manifold. The short-time short-distance asymptotics is related to approximating $e(t, p, p)$ for $t \to 0$. This is quite well understood.

Definition 1.45. Let $\beta : \mathbb{R}_+ \to \mathbb{R}$ be a function. If for all $N \geq k_0$ we have

$$\lim_{t\to 0} \frac{\beta(t) - \sum_{k=k_0}^{N} b_k t^k}{t^N} = 0,$$

then we say that β has an asymptotic expansion $\beta(t) \sim \sum_{k=k_0}^{\infty} b_k t^k$.

It is a classical result (cf. [5, Corollaire E.III.9], [21, Proposition 3.23, Lemma 3.26]) that

$$e(t, p, p) \sim (4\pi t)^{-n/2} \sum_{k=0}^{\infty} u_k(p) t^k,$$

where $u_0(p) = 1$ and the other coefficients can be expressed as universal (i.e. independent of the manifold except for its dimension) polynomials of the Riemann curvature and its derivatives.

Later short-time long-distance asymptotic expansions away from the cut locus were obtained (cf. [3, Theorem (3.1)]).

Theorem 1.46. *Let $\Sigma = \{(p, q) : q \in V_p\}$. If $(p, q) \in \Sigma$, then we have an asymptotic expansion*

$$e(t, p, q) \sim \frac{1}{t^{n/2}} \exp\left(-\frac{\mathrm{dist}(p, q)^2}{4t}\right) \left(\sum_{k=0}^{\infty} c_k(p, q) t^k\right),$$

uniformly on every compact subset of Σ, where the functions $c_i : \Sigma \to \mathbb{R}$ are smooth and $c_0(p, q) > 0$.

For $(p, q) \in M \times M$ we still have

$$\lim_{t\to 0} -4t \log e(t, p, q) = \mathrm{dist}(p, q)^2$$

uniformly on compact subsets of $M \times M$, even on the cut locus (see [23], [17], [13, Theorem 5.2.1]).

However, the following result shows that the order of the leading term of the asymptotic expansion can be different on the cut locus, depending on the behavior of the hinged energy function.

Theorem 1.47. *Let $p, q \in M$ and assume that there exists only one minimal geodesic joining p and q. Let x be the midpoint of this geodesic and assume that there exists a coordinate system $(z_1, \ldots, z_n)$ near x such that*

$$h_{p,q}(z) = h_{p,q}(x) + z_1^{2m_1} + \cdots + z_n^{2m_n} + o(|z_1|^{2m_1} + \cdots + |z_n|^{2m_n})$$

for some integers $1 \leq m_1 \leq \cdots \leq m_n$. Then there exists a constant C (depending on M, p and q) such that

$$e(t, p, q) = \frac{C + o(1)}{t^{n - \sum_i \frac{1}{2m_i}}} \exp\left(-\frac{\operatorname{dist}(p, q)^2}{4t}\right).$$

A full proof can be found in [19], whereas the mains ideas are already in [18]. The result was extended to sub-Riemannian manifolds in [4].

Remark 1.48. As explained in [4, Remark 2], if there exists a one-parameter family of minimal geodesics joining p and q, the theorem is still valid, but it should be understood that some m_i is infinite. For example let $M = S^2$ be a sphere of radius ρ and let p, q be two opposite poles. Then we have $m_1 = 1, m_2 = \infty$, so the asymptotics is

$$e(t, p, q) = \frac{C + o(1)}{t^{3/2}} \exp\left(-\frac{\rho^2 \pi^2}{4t}\right).$$

We will check this result in Section 2.

In particular, let us return to the situation of Corollary 1.39. We obtain the following result.

Corollary 1.49. *Let M be a 2-dimensional orientable Riemannian manifold. Let $p, q \in M$ and assume that there exists only one minimal geodesic joining p and q. Assume further that there exists no smooth curve $v : (-\epsilon, \epsilon) \to T_pM$ such that $\gamma_{v(0)} : [0, 2] \to M$ is the minimal geodesic from p to q and*

$$\operatorname{dist}(\gamma_{v(t)}(2), q) = o(t^3).$$

Then there exists a constant C (depending on M, p and q) such that

$$e(t, p, q) = \frac{C + o(1)}{t^{5/4}} \exp\left(-\frac{\operatorname{dist}(p, q)^2}{4t}\right).$$

The purpose of Section 3 is to provide a class of manifolds satisfying the assumptions of this corollary.

2 First examples

We analyze the circle S^1 to show that the expansion in Theorem 1.46 is not (in general) valid on the cut locus. Then, using the formulas obtained for S^1, we study the heat kernel on simple 2-dimensional manifolds – the cylinder $S^1 \times \mathbb{R}$ and the sphere S^2.

2.1 The heat kernel on the circle

We follow here Section 1.1.2 in [21]

Let $(M, g) = (S^1, \mathrm{d}\theta)$. It is easily seen that for $X(\theta) = a(\theta)\partial_\theta$ we have $\mathrm{div}X(\theta) = \partial_\theta a(\theta)$, so the Laplace operator on smooth functions is given by

$$\Delta f = \frac{\mathrm{d}^2 f}{\mathrm{d}\theta^2}.$$

The heat equation

$$\partial_t f(t,\theta) = \partial_\theta^2 f(t,\theta) \tag{18}$$

can be solved explicitly using Fourier series. Let

$$f_0(\theta) = f(0,\theta) = \sum_{n\in\mathbb{Z}} a_n e^{in\theta}$$

be the initial data. We search a solution of the form

$$f(t,\theta) = \sum_{n\in\mathbb{Z}} a_n(t) e^{in\theta}.$$

Substituting into (18) we obtain

$$\sum_{n\in\mathbb{Z}} \left(a_n'(t) + n^2 a_n(t)\right) e^{in\theta} = 0,$$

which gives $a_n(t) = e^{-n^2 t} a_n$ and the desired solution

$$f(t,\theta) = \sum_{n\in\mathbb{Z}} a_n e^{-n^2 t} e^{in\theta}. \tag{19}$$

Notice that $a_n = \frac{1}{2\pi}\int_{S^1} e^{-in\psi} f_0(\psi)\,\mathrm{d}\psi$, so (19) can be written in the form

$$e^{t\Delta} f_0(\theta) = f(t,\theta) = \frac{1}{2\pi}\int_{S^1} \sum_{n\in\mathbb{Z}} e^{-n^2 t} e^{in\theta} e^{-in\psi} f_0(\psi)\,\mathrm{d}\psi.$$

Hence the heat kernel on $(S^1, \mathrm{d}\theta)$ is given by the formula

$$e_{S^1}(t,\theta,\psi) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} e^{-n^2 t} e^{in(\theta-\psi)}. \tag{20}$$

It is well known that the heat kernel on the real line is given by the formula

$$e_{\mathbb{R}}(t,x,y) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{(x-y)^2}{4t}}.$$

Consider the covering map $\mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z} \simeq S^1$. Intuitively the heat goes from one point on the circle to another by turning left or right an arbitrary number of times, which would lead to a formula

$$e_{S^1}(t,\theta,\psi) = \sum_{n\in\mathbb{Z}} e_{\mathbb{R}}(t,\theta,\psi + 2n\pi). \tag{21}$$

It is indeed the case.

Proposition 2.1. *Let* $\widetilde{e}_{S^1} : \mathbb{R}_+ \times S^1 \times S^1 \to \mathbb{R}$ *be defined by*

$$\widetilde{e}_{S^1}(t, \theta, \psi) = \sum_{n \in \mathbb{Z}} e_{\mathbb{R}}(t, \theta, \psi + 2n\pi).$$

Then $\widetilde{e}_{S^1}$ *is well defined and* $\widetilde{e}_{S^1} = e_{S^1}$.

Proof. It is evident that for $k, l \in \mathbb{Z}$ we have $\widetilde{e}_{S^1}(t, \theta + 2k\pi, \psi + 2l\pi) = \widetilde{e}_{S^1}(t, \theta, \psi)$, so the value of $\widetilde{e}_{S^1}(t, \theta, \psi)$ does not depend on the choice of representatives.

To prove that $\widetilde{e}_{S^1} = e_{S^1}$ it is sufficient to show that $\widetilde{e}_{S^1}$ is a heat kernel. The conclusion will follow by the uniqueness of the heat kernel.

From $(\partial_t - \partial_\theta^2) e_{\mathbb{R}} = 0$ it follows immediately that $(\partial_t - \partial_\theta^2)\widetilde{e}_{S^1} = 0$.

Now let $f \in C^\infty(S^1)$, $\theta \in S^1$. We want to prove that

$$\lim_{t \to 0} \int_{S^1} \widetilde{e}_{S^1}(t, \theta, \psi) f(\psi) \, \mathrm{d}\psi = f(\theta).$$

The left hand side equals $\lim_{t \to 0} \int_{\mathbb{R}} e_{\mathbb{R}}(t, \theta, \psi) \widetilde{f}(\psi) \, \mathrm{d}\psi$, where $\widetilde{f} : \mathbb{R} \to \mathbb{R}$ is the periodic extension of f. The function $\widetilde{f}$ is bounded and C^∞, so it is a standard exercise in real analysis that

$$\lim_{t \to 0} \int_{\mathbb{R}} e_{\mathbb{R}}(t, \theta, \psi) \widetilde{f}(\psi) \, \mathrm{d}\psi = \widetilde{f}(\theta) = f(\theta).$$

□

Remark 2.2. Putting $\theta = \psi = 0$ in (21) we get for example

$$\sum_{n \in \mathbb{Z}} e^{-n^2 t} = \sqrt{\frac{\pi}{t}} \sum_{n \in \mathbb{Z}} e^{-\frac{n^2 \pi^2}{t}}.$$

Proving directly an identity like this seems to be a hard task. The proof above is simple but it relies on the (nontrivial) uniqueness of the heat kernel.

Proposition 2.3. *Let* $\Sigma = \{(\theta, \psi) \in S^1 \times S^1 : \theta$ *and* ψ *are not antipodal*$\}$. *Then we have an asymptotic expansion*

$$e_{S^1}(t, \theta, \psi) \sim \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{\mathrm{dist}(\theta, \psi)^2}{4t}\right).$$

uniformly on compact subsets of Σ.

If θ *and* $\hat{\theta}$ *are antipodal, then we have an asymptotic expansion*

$$e_{S^1}(t, \theta, \hat{\theta}) \sim \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{\mathrm{dist}(\theta, \hat{\theta})^2}{4t}\right).$$

Proof. Notice that if $c > 0$, then for any $N \in \mathbb{Z}$ we have $\lim_{t\to 0} t^{-N} e^{-\frac{c}{t}} = 0$ and the convergence is uniform on compact subsets of $\{c \in \mathbb{R} : c > 0\}$. Thus all the terms on the right hand side of (21) such that $(\theta - \psi - 2n\pi)^2$ is not minimal are negligible. If the two points are not antipodal, only one term is left. If they are antipodal, there are two (equal) terms that count. $\square$

Remark 2.4. The antipodal point is the unique cut point. It turns out that the asymptotic expansion from Theorem 1.46 cannot be extended to this point, because $c_0(p, q)$ would have to be discontinuous for q antipodal to p.

2.2 The heat kernel on the 2-dimensional cylinder

Let (M_1, g_1) and (M_2, g_2) be orientable Riemannian manifolds and consider the product manifold $(M_1 \times M_2, g_1 \otimes g_2)$. Let e_{M_1} be the heat kernel on M_1 and let e_{M_2} be the heat kernel on M_2.

Proposition 2.5. *The heat kernel on $M_1 \times M_2$ satisfies*

$$e_{M_1\times M_2}(t, (x_1, x_2), (y_1, y_2)) = e_{M_1}(t, x_1, y_1) e_{M_2}(t, x_2, y_2). \tag{22}$$

Proof. Let $X_1 \in \mathrm{Vec}(M_1)$, $X_2 \in \mathrm{Vec}(M_2)$, $f_1 \in C^\infty(M_1)$, $f_2 \in C^\infty(M_2)$. Then $(X_1, X_2) \in \mathrm{Vec}(M_1 \times M_2)$ and $f_1 \otimes f_2 \in C^\infty(M_1 \times M_2)$, where $(f_1 \otimes f_2)(x_1, x_2) := f_1(x_1) f_2(x_2)$. It is easily checked from the definition that

$$\mathrm{div}_{M_1\times M_2}(X_1, X_2) = \mathrm{div}_{M_1} X_1 + \mathrm{div}_{M_2} X_2$$

and

$$\nabla_{M_1\times M_2}(f_1 \otimes f_2)(x_1, x_2) = (f_2(x_2)\nabla_{M_1} f_1(x_1), f_1(x_1)\nabla_{M_2} f_2(x_2)).$$

Thus $\Delta_{M_1\times M_2}(f_1 \otimes f_2)(x_1, x_2) = f_2(x_2)\Delta_{M_1} f_1(x_1) + f_1(x_1)\Delta_{M_2} f_2(x_2)$, which means that the right hand side of (22) satisfies condition (16). Condition (17) follows by iterated integration. $\square$

Consider the 2-dimensional cylinder $\mathrm{Cyl} := S^1 \times \mathbb{R}$. The cut locus of a point $(\theta, z_0) \in S^1 \times \mathbb{R}$ is the line $\{(\hat{\theta}, z) : z \in \mathbb{R}\}$, where $\hat{\theta}$ is the point of S^1 antipodal to θ.

Proposition 2.6. *Let $\Sigma = \{((\theta, z), (\psi, w)) \in \mathrm{Cyl} \times \mathrm{Cyl} : \theta$ and ψ are not antipodal$\}$. Then we have an asymptotic expansion*

$$e_{S^1\times\mathbb{R}}(t, (\theta, z), (\psi, w)) \sim \frac{1}{4\pi t} \exp\left(-\frac{\mathrm{dist}((\theta, z), (\psi, w))^2}{4t}\right).$$

uniformly on compact subsets of Σ.

If θ and $\hat{\theta}$ are antipodal, then we have an asymptotic expansion

$$e_{S^1}(t, \theta, \hat{\theta}) \sim \frac{1}{2\pi t} \exp\left(-\frac{\mathrm{dist}((\theta, z), (\hat{\theta}, w))^2}{4t}\right).$$

Proof. Notice that $\mathrm{dist}_{\mathrm{Cyl}}((\theta, z), (\psi, w))^2 = \mathrm{dist}_{S^1}(\theta, \psi)^2 + \mathrm{dist}_{\mathbb{R}}(z, w)^2$. Thus the conclusion follows from Proposition 2.3, Proposition 2.5 and the formula for $e_{\mathbb{R}}$. $\square$

2.3 The heat kernel on the 2-sphere

The heat kernel on the 2-sphere was first computed in [8]. It is done by considering S^2 as a homogeneous space $SU(2)/U(1)$ and studying the representations of $SU(2)$ (which slightly resembles the methods used in the case of S^1). In order to be able to use this result, we have to recall briefly the action of $SU(2)$ on S^2.

Let $z \in \overline{\mathbb{C}} \simeq \mathbb{PC}^1 \simeq S^2$ (the *Riemann sphere*) and let $\begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \in SU(2)$ (we have $|\alpha|^2 + |\beta|^2 = 1$). The required group action is given by the following *Möbius transformation*:

$$\begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} z = \frac{\alpha z - \overline{\beta}}{\beta z + \overline{\alpha}}.$$

It is a transitive group action with isotropy group $U(1) \subset SU(2)$ given by diagonal matrices. The group $SU(2)$ is a Riemannian manifold with the metric induced from $\mathbb{C}^4$. The quotient (homogeneous) space $SU(2)/U(1)$ is the standard sphere of radius $\sqrt{2}$.

After this short introduction we can state the result of [8].

Theorem 2.7. *Let* $\mathrm{Id} \in S^2$ *be represented by* $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in SU(2)$ *and let* $q \in S^2$ *be represented by* $\begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \in SU(2)$. *Then*

$$e(t, \mathrm{Id}, q) = \frac{1}{8\pi} \sum_{n=0}^{\infty} (2n+1) e^{-\frac{n(n+1)}{2}t} P_n(2|\alpha|^2 - 1), \tag{23}$$

where P_n *is the Legendre polynomial.*

Remark 2.8. Notice that $|\alpha|^2$ does not depend on the choice of a representative of q.

The point antipodal to Id, which we denote i, is represented by a matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ with $\alpha = 0$. From the recursion formula for Legendre polynomials

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$$

it follows immediately that $P_n(-1) = (-1)^n$. Thus we obtain

$$e(t, \mathrm{Id}, i) = \frac{1}{8\pi} \sum_{n=0}^{\infty} (-1)^n (2n+1) e^{-\frac{n(n+1)}{2}t}.$$

We will compute an asymptotic expansion of this expression.

To this end first put $\psi = 0$ in (21). This gives

$$1 + 2\sum_{n=0}^{\infty} e^{-n^2 t} \cos(n\theta) = \sqrt{\frac{\pi}{t}} \sum_{n \in \mathbb{Z}} e^{-\frac{(\theta - 2n\pi)^2}{4t}}.$$

For fixed $t > 0$, both sides can be differentiated with respect to θ term by term. This leads to

$$\begin{aligned} 2\sum_{n=0}^{\infty} ne^{-n^2 t}\sin(n\theta) &= \sqrt{\frac{\pi}{t}}\sum_{n\in\mathbb{Z}}\frac{\theta-2n\pi}{2t}e^{-\frac{(\theta-2n\pi)^2}{4t}} \\ &= \frac{\sqrt{\pi}}{2t^{3/2}}\sum_{n\in\mathbb{Z}}(\theta-2n\pi)e^{-\frac{(\theta-2n\pi)^2}{4t}}. \end{aligned}$$

Put $\theta = \frac{\pi}{2}$. On the left hand side we have

$$\begin{aligned} \sum_{n=0}^{\infty} ne^{-n^2 t}\sin\left(\frac{n\pi}{2}\right) &= \sum_{n=0}^{\infty}(-1)^n(2n+1)e^{-(2n+1)^2 t} \\ &= e^{-t}\sum_{n=0}^{\infty}(-1)^n(2n+1)e^{-\frac{n^2+n}{2}\cdot 8t}. \end{aligned}$$

On the right hand side all the terms except $n = 0$ are negligible for an asymptotic expansion. Hence

$$\sum_{n=0}^{\infty}(-1)^n(2n+1)e^{-\frac{n^2+n}{2}\cdot 8t} \sim \frac{\pi^{3/2}}{8t^{3/2}}e^{-\frac{\pi^2}{16t}}e^t$$

or, writing t instead of $8t$,

$$e(t,\mathrm{Id},i) = \frac{1}{8\pi}\sum_{n=0}^{\infty}(-1)^n(2n+1)e^{-\frac{n^2+n}{2}t} \sim \frac{\sqrt{\pi}}{(2t)^{3/2}}e^{-\frac{2\pi^2}{4t}}e^{t/8}.$$

If we recall that $\mathrm{dist}(\mathrm{Id}, i) = \sqrt{2}\pi$, we see that the asymptotic expansion confirms what was said in Remark 1.48. We see again that Theorem 1.46 fails on the cut locus. As the cut point is also conjugate, in the asymptotic expansion we obtain the power of t different from $t^{-n/2} = t^{-1}$.

3 Oblate ellipsoids

3.1 Cut and conjugate loci on ellipsoids

The study of cut and conjugate loci on 2-dimensional manifolds dates back to the works of Poincaré [20] and Jacobi [15].

It is worth mentioning that the structure of the cut and conjugate loci on general ellipsoids was determined quite recently [14], although it was already predicted in [15]. It turns out that every non-umbilical point on an ellipsoid with distinct axis has exactly two cut-conjugate points and the shape of the cut and conjugate loci in their neighborhood is as presented on Figure 1.

To perform explicit computations, it is more convenient to investigate ellipsoids of revolution. The shape of the cut and conjugate loci on 2-spheres of revolution (of which an

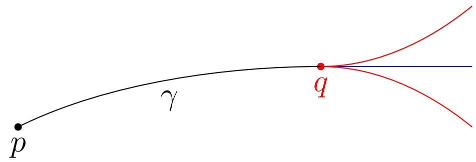

Figure 1: Generic shape of the cut and conjugate loci.

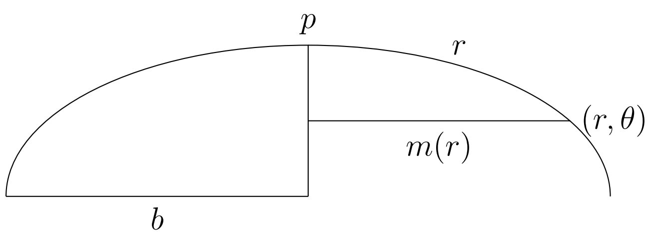

Figure 2: Notational convention for a 2-sphere of revolution.

ellipsoid of revolution is an example) were found in [22]. In paragraphs 3.2, 3.3 and 3.4 we use these results to compute the order of degeneracy of the exponential map starting from a point on the equator of an oblate ellipsoid near its cut-conjugate point. In paragraph 3.5 these results are expressed in terms of the geometry of ellipsoids.

3.2 2-spheres of revolution

Definition 3.1. Let (M, g) be a compact Riemannian manifold homeomorphic to a 2-sphere. M is called a 2-*sphere of revolution* if there exists a point $p \in M$ called a *pole* such that for any $q_1, q_2 \in M$ satisfying $\mathrm{dist}(p, q_1) = \mathrm{dist}(p, q_2)$ there exists an isometry $f : M \to M$ satisfying $f(q_1) = q_2, f(p) = p$.

Remark 3.2. A model example of a 2-sphere of revolution is a closed rotational surface in $\mathbb{R}^3$.

Let (M, g) be a 2-sphere of revolution and let p be a pole. It can be proved that p has a unique cut point q which is also a pole [22, Lemma 2.1]. Therefore $M \setminus \{p, q\}$ can be parametrized by geodesic polar coordinates around p, which we denote (r, θ). Let $M' := M \setminus \{p, q\}$. This allows to express the Riemannian metric on M' as $g = \mathrm{d}r^2 + m(r)^2\,\mathrm{d}\theta^2$ with $\lim_{r\to 0} m(r) = 0$ [7, p. 287, Proposition 3]. We choose m to be positive. It can be thought of as the distance from the rotational axis (see also Figure 2, which shows a section of M by a plane containing the rotational axis).

We denote $a := \frac{1}{2}\mathrm{dist}(p, q)$ and $b := m(a)$. The set $\{(r, \theta) \in M : r = a\}$ is called the *equator*. Each set $\{(r, \theta) \in M : r = \mathrm{const}\}$ is called a *parallel*. Each set $\{(r, \theta) \in M : \theta = \mathrm{const}\}$ is called a *meridian*. In what follows we make the following assumptions:

(A1) $m(2a - r) = m(r)$, i.e. the metric is invariant by the reflection flipping the poles.

(A2) The Gaussian curvature is monotone increasing along a meridian.

(A3) $m(r)$ is analytic in some neighborhood of $r = a$.

The first two assumptions are taken directly from [22]. The third assumption allows to perform formal computations on power series, which will be the method used in what follows.

Fix a point p_0 on the equator and assume $\theta(p_0) = 0$. Recall that M is compact, so geodesically complete by Theorem 1.18. We need a classical result on surfaces of revolution.

Proposition 3.3. *Let $\gamma(s) = (r(s), \theta(s))$ be a unit speed geodesic on M'. There exists a constant ν, called the* Clairaut constant *of γ, such that*

$$m(r(s))^2\dot{\theta}(s) = \nu. \tag{24}$$

Here (and later) $\dot{\theta}(s)$ should be understood as $\dot{\widetilde{\theta}}(s)$, where $\widetilde{\theta}(s) \in \mathbb{R}$ is a lift of $\theta(s) \in \mathbb{R}/2\pi\mathbb{R}$.

Proof. In canonical coordinates $\lambda = (r, \theta, \xi, \zeta)$ the Hamiltonian has the form

$$H(\lambda) = \xi^2 + \frac{1}{m(r)^2}\zeta^2.$$

We have $\dot{\zeta} = \partial_\theta H = 0$, so ζ is a constant. Hence $m(r(s))^2\dot{\theta}(s) = m(r(s))^2\partial_\theta H = 2\zeta$ is a constant. □

Remark 3.4. Observe that $m(r(s))\dot{\theta}(s) = \cos\eta(s)$, where $\eta(s)$ is the angle between $\dot{\gamma}(s)$ and $\frac{\partial}{\partial\theta}$. The constant ν can be viewed as the angular momentum of the geodesic.

Proposition 3.5. *Let $\nu \in \mathbb{R}$. Then we have the following.*

i) If $|\nu| > b$, then no geodesic on M' emanating from p_0 satisfies (24).

ii) If $|\nu| = b$, then there is exactly one geodesic $\gamma : \mathbb{R} \to M'$ emanating from p_0 satisfying (24)*. Its image is the equator.*

iii) If $0 < |\nu| < b$, then there is exactly one geodesic $\gamma : \mathbb{R} \to M'$ emanating from p_0 satisfying (24) *and $r'(0) < 0$.*

This unique geodesic will be denoted by c_ν.

Proof. We assumed $\gamma(s)$ to be unit speed, that is $r'(s)^2 + m(r(s))^2\theta'(s)^2 = 1$. Thus (24) is equivalent to

$$r'(s)^2 = \frac{m(r(s))^2 - \nu^2}{m(r(s))^2}, \tag{25}$$

in particular

$$r'(0)^2 = \frac{b^2 - \nu^2}{b^2}. \tag{26}$$

i) If $|\nu| > b$ we clearly get a contradiction.

ii) If $|\nu| = b$ we get $r'(0) = 0$ and $\theta'(0) = \frac{\nu}{b^2}$. There is a unique geodesic with initial tangent vector $(r'(0), \theta'(0)) = \left(0, \frac{\nu}{b^2}\right)$. By symmetry this geodesic will not quit the equator. By (24) $\theta'(s)$ is constant, so the geodesic covers the whole equator.

iii) If $|\nu| < b$ we obtain $r'(0) = -\sqrt{\frac{b^2-\nu^2}{b^2}}$ (because we chose $r'(0)$ negative) and $\theta'(0) = \frac{\nu}{b^2}$. This determines the unique geodesic with Clairaut constant ν. We have $|m(r(s))\theta'(s)| = |\cos\eta(s)| \leq 1$, so $m(r(s)) \geq |\nu| > 0$. This assures that the geodesic does not meet the poles.

□

From assumption (A2) and Gauss-Bonnet Theorem it follows that the Gaussian curvature on the equator is strictly positive. Thus from Lemmas 2.2 and 2.3 in [22] one can obtain that m is strictly increasing on $(0, a]$. For ν, $0 < |\nu| \leq b$, we will denote by $R = R(\nu)$ the unique $R \in (0, a]$ such that $m(R) = |\nu|$. It is the minimal geodesic distance of $c_\nu(s)$ from the pole p.

Theorem 4.1 in [22] states that the cut locus of p_0 is a subset of the equator. For $0 < |\nu| < a$ the cut point along c_ν is the first point of intersection of c_ν with the equator. This point is given by the formula [22, p. 385] $(r, \theta) = (a, \varphi(\nu))$, where

$$\varphi(\nu) := 2\int_R^a \frac{\nu\,\mathrm{d}r}{m(r)\sqrt{m(r)^2 - \nu^2}}. \tag{27}$$

From now on we will assume that $\nu \geq 0$ (the case $\nu \leq 0$ is symmetric). It can be shown that $\varphi(\nu)$ is non-increasing for $\nu \in (0, b)$ [22, Lemma 4.2]. Thus the cut-conjugate point of p_0 has coordinates $(r, \theta) = (a, \lim_{\nu\to b^-} \varphi(\nu))$. We note $\theta_{\mathrm{cut}} := \lim_{\nu\to b^-} \varphi(\nu)$ and $t_{\mathrm{cut}} := b\theta_{\mathrm{cut}}$, so that $c_b(t_{\mathrm{cut}})$ is the cut-conjugate point.

Remark 3.6. The geodesic c_ν starting from p_0 is determined by each of the parameters ν, R or η. We recall the relationships between these parameters – we have $\nu = m(R)$ and $\cos\eta = \frac{\nu}{b}$. In particular $b - \nu$ is of order η^2 as $\eta \to 0$.

3.3 Cut point of geodesics close to the equator

We will now derive the power expansion of $\varphi(\nu)$ for ν close to b from the power expansion of $m(r)$ near $r = a$. By assumption (A1) $m(a + \epsilon) = m(a - \epsilon)$, so in the power expansion of m near $r = a$ odd powers do not appear. Let

$$m(r) = b - \alpha(a-r)^2 + \beta(a-r)^4 + O\left((a-r)^6\right)$$

be the beginning of this expansion.

Remark 3.7. From the formula for the Gaussian curvature of a surface of revolution [7, p. 162] we obtain that the Gaussian curvature of M on the equator equals $G = -\frac{m''(a)}{m(a)} = \frac{2\alpha}{b}$. In particular $\alpha > 0$.

To justify our computation we need a simple lemma.

Lemma 3.8. *Let $f(x,y) = a_{00} + a_{10}x + a_{01}y + \cdots$ be a real analytic function of two variables in some neighborhood of $(x,y) = (0,0)$. Then the function*

$$F(y) = \int_0^y \frac{f(x,y)\,\mathrm{d}x}{\sqrt{y^2 - x^2}}$$

is real analytic in a neighborhood of $y = 0$ and

$$F(y) = b_0 + b_1 y + \cdots \qquad \text{where } b_k = \sum_{j=0}^{k} a_{j,k-j} \int_0^1 \frac{u^j\,\mathrm{d}u}{\sqrt{1-u^2}}. \tag{28}$$

Proof. Let $f_n(x,y) := \sum_{i+j\le n} a_{ij}x^i y^j$ and $F_n(y) := \sum_{k\le n} b_k y^k$. Substituting $u = \frac{x}{y}$ we get

$$\begin{aligned}\int_0^y \frac{f_n(x,y)\,\mathrm{d}x}{\sqrt{y^2-x^2}} &= \sum_{i+j\le n} a_{ij} \int_0^y \frac{x^i y^j\,\mathrm{d}x}{\sqrt{y^2-x^2}} \\ &= \sum_{k=0}^{n} y^k \sum_{i+j=k} a_{ij} \int_0^1 \frac{u^i\,\mathrm{d}u}{\sqrt{1-u^2}} = F_n(y).\end{aligned} \tag{29}$$

Observe that in some neighborhood of $(x,y) = (0,0)$ the sequence $f_n(x,y)$ converges uniformly to $f(x,y)$ as $n \to \infty$, so in (29) the left hand side converges to $\int_0^y \frac{f(x,y)\,\mathrm{d}x}{\sqrt{y^2-x^2}} = F(y)$.

Notice that $\int_0^1 \frac{u^j\,\mathrm{d}u}{\sqrt{1-u^2}} \le \frac{\pi}{2}$. Thus $|b_k| \le \frac{\pi}{2}\sum_{j=0}^{k} |a_{j,k-j}|$ and it is clear that the series $\sum_k b_k y^k$ converges in some neighborhood of $y = 0$. Let $\hat{F}(y) := \sum_k b_k y^k$. Hence the right hand side of (29) converges to $\hat{F}(y)$ as $n \to \infty$. It follows that $F(y) = \hat{F}(y)$. □

Remark 3.9. If in the expansion of $f(x,y)$ only even powers appear, the same is true for $F(y)$.

Recall that $\nu = m(R)$ and $R \to a^-$ as $\nu \to b^-$. Let $y = a - R$. We have

$$\begin{aligned}\varphi(\nu) &= \int_R^a \frac{2m(R)}{m(r)\sqrt{m(r)^2 - m(R)^2}} \\ &= \int_0^y \frac{2m(a-y)\,\mathrm{d}x}{m(a-x)\sqrt{m(a-x)+m(a-y)}\sqrt{m(a-x)-m(a-y)}}.\end{aligned} \tag{30}$$

From $b > 0$ it is clear that

$$(x,y) \mapsto \frac{2m(a-y)}{m(a-x)\sqrt{m(a-x)+m(a-y)}}$$

is an analytic function of two variables in a neighborhood of $(x, y) = 0$ with even powers only. As for $(m(a-x) - m(a-y))^{-1/2}$, it can be written as

$$\frac{1}{\sqrt{y^2 - x^2}} \frac{1}{\sqrt{\alpha - \beta(y^2 + x^2) + \cdots}}.$$

We have $\alpha > 0$, so the second factor is an analytic function in a neighborhood of $(x, y) = 0$. Only even powers appear in its expansion.

Summing up, the right hand side of (30) has the form required in Lemma 3.8. Performing explicitly the computation described above (it is long but straightforward) we get the expression

$$\phi(\nu) = \int_0^a \frac{f(x, y)\,\mathrm{d}x}{\sqrt{y^2 - x^2}}$$

with

$$f(x, y) = \frac{\sqrt{2}}{\sqrt{b\alpha}} + \frac{5\alpha^2 + 2b\beta}{2\sqrt{2}(b\alpha)^{3/2}} x^2 + \frac{-3\alpha^2 + 2b\beta}{2\sqrt{2}(b\alpha)^{3/2}} y^2 + O(x^4 + y^4),$$

and in the power expansion of the function f only even powers appear.

Using the fact that $\int_0^1 \frac{\mathrm{d}u}{\sqrt{1-u^2}} = \frac{\pi}{2}$ and $\int_0^1 \frac{u^2\,\mathrm{d}u}{\sqrt{1-u^2}} = \frac{\pi}{4}$ we obtain

$$b_0 = \frac{\pi}{2} a_{00} = \frac{\pi}{\sqrt{2b\alpha}}$$

and

$$b_2 = \frac{\pi}{2} a_{02} + \frac{\pi}{4} a_{20} = \frac{(6b\beta - \alpha^2)\,\pi}{8\sqrt{2}(b\alpha)^{3/2}},$$

so

$$\varphi(\nu) = \frac{\pi}{\sqrt{2b\alpha}} + \frac{(6b\beta - \alpha^2)\,\pi}{8\sqrt{2}(b\alpha)^{3/2}} (a - R)^2 + O\left((a - R)^4\right), \tag{31}$$

where the expansion contains only even powers of $a - R$.

Remark 3.10. From this formula it is clear that the case $6b\beta = \alpha^2$ is going to be singular. It is easy to compute that this is equivalent to $G''(a) = 0$, where $G(s)$ is the Gaussian curvature on the parallel $\{r = s\}$.

It follows from the fact that $\alpha > 0$ and a well-known theorem on analyticity of inverse functions of analytic functions that $(a - R)^2$ is an analytic function of $\nu = m(R)$ in a neighborhood of $\nu = b$. An explicit computation gives

$$(a - R)^2 = \frac{b - \nu}{\alpha} + \frac{\beta(b - \nu)^2}{\alpha^3} + O\left((b - \nu)^3\right),$$

which together with (31) leads to the following conclusion.

Proposition 3.11. *The function $\varphi(\nu)$ is analytic near $\nu = b$ and we have the expansion*

$$\varphi(\nu) = \frac{\pi}{\sqrt{2\alpha b}} + \frac{(6b\beta - \alpha^2)\,\pi}{8\sqrt{2}b^{3/2}\alpha^{5/2}}(b-\nu) + O\left((b-\nu)^2\right). \tag{32}$$

In particular $\theta_{\mathrm{cut}} = \frac{\pi}{\sqrt{2b\alpha}}$.

Remark 3.12. In our case the Gaussian curvature on the equator is constant and equal $G = \frac{2\alpha}{b}$. This permits to calculate explicitly an appropriate Jacobi field and in that way find the first conjugate point on the equator.

More precisely, let $\gamma : \mathbb{R} \to M$ be a geodesic parametrized by arc length and let $G(s)$ denote the Gaussian curvature at the point $\gamma(s)$. It is a general fact that the first conjugate point along γ is $\gamma(s_{\mathrm{conj}})$, where s_{conj} is the smallest positive solution of the differential equation

$$u''(s) + G(s)u(s) = 0$$

with the initial conditions $u(0) = 0$, $u'(0) = 1$ [7, p. 368, Exercise 1b].

In our case $G(s) = \frac{2\alpha}{b}$, so the solution of this equation is given by

$$u(s) = \sqrt{\frac{b}{2\alpha}} \sin\left(\sqrt{\frac{2\alpha}{b}} s\right)$$

and its first positive zero is $t_{\mathrm{cut}} = s_{\mathrm{conj}} = \frac{\pi b}{\sqrt{2b\alpha}}$. On the equator $\mathrm{d}s = b\,\mathrm{d}\theta$, so we get indeed $\theta_{\mathrm{cut}} = \frac{t_{\mathrm{cut}}}{b} = \frac{\pi}{\sqrt{2b\alpha}}$.

3.4 Degeneracy of $\nu \mapsto c_\nu(t_{\mathrm{cut}})$ and related questions

Recall that the geodesic $c_b(s)$ follows the equator and reaches the cut-conjugate point $(r,\theta) = (a, \theta_{\mathrm{cut}})$ for $s = t_{\mathrm{cut}}$.

We are now interested in the map $\nu \mapsto c_\nu(t_{\mathrm{cut}})$ near $\nu = b$. Recall that $\eta = \eta(0) = \arccos\frac{\nu}{b}$ is the angle between $c'_\nu(0)$ and $\left(\frac{\partial}{\partial\theta}\right)_{(a,0)}$. Observe also that $\arccos\left(1 - \frac{x^2}{2}\right) = x + O(x^3)$. For $x = \sqrt{\frac{2(b-\nu)}{b}}$ this gives

$$\eta = \sqrt{\frac{2\alpha}{b}}(a-R) + O((a-R)^3).$$

The goal is to prove the following result.

Proposition 3.13.

$$\operatorname{dist}\left(c_\nu(t_{\mathrm{cut}}), \left(a - \frac{(6\beta b - \alpha^2)\sqrt{b}\pi}{16\sqrt{2}\alpha^{5/2}}\eta^3, \theta_{\mathrm{cut}}\right)\right) = O(\eta^4). \tag{33}$$

First we will determine the asymptotics of the point of intersection of geodesics with the "critical" meridian.

Lemma 3.14. *For $0 < \nu < b$ let $(r_\nu, \theta_{\mathrm{cut}})$ be the first point of intersection of c_ν with the meridian $P := \{\theta = \theta_{\mathrm{cut}}\}$. Then $r_\nu = a - \frac{(6b\beta - \alpha^2)\sqrt{b}\pi}{16\sqrt{2}\alpha^{5/2}}\eta^3 + O(\eta^5)$.*

Proof. We assume that $c_\nu(s)$ reaches $r = R$ before it reaches $\theta = \theta_{\mathrm{cut}}$, which means it goes down after intersecting with P (this will be true for η small enough). This means that $r'(s) \geq 0$ for $\theta \in [\theta_{\mathrm{cut}}, \varphi(\nu)]$. Thus from (25) we get

$$r'(s) = \frac{\sqrt{m(r(s))^2 - \nu^2}}{m(r(s))}.$$

Together with the Clairaut relation

$$\theta'(s) = \frac{\nu}{m(r(s))^2}$$

this permits to compute

$$\begin{aligned}
\frac{\mathrm{d}r}{\mathrm{d}\theta} &= \frac{r'(s)}{\theta'(s)} = \frac{m(r(s))\sqrt{m(r(s))^2 - \nu^2}}{\nu}, \\
r''(s) &= \frac{\nu^2 m'(r(s))}{m(r(s))^3}, \\
\theta''(s) &= -\frac{2\nu\sqrt{m(r(s))^2 - \nu^2}\, m'(r(s))}{m(r(s))^4}, \\
\frac{\mathrm{d}^2 r}{\mathrm{d}\theta^2} &= \frac{r''(s)\theta'(s) - r'(s)\theta''(s)}{(\theta'(s))^3} = \frac{(2m(r(s))^2 - \nu^2)\, m(r(s)) m'(r(s))}{\nu^2}.
\end{aligned}$$

In particular

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)_{\theta = \varphi(\nu)} = \frac{b\sqrt{b^2 - \nu^2}}{\nu} = b\tan(\eta)$$

and

$$\frac{\mathrm{d}^2 r}{\mathrm{d}\theta^2} = O(R) = O(\eta).$$

Thus

$$\begin{aligned}
a - r_\nu &= \int_{\theta_{\mathrm{cut}}}^{\varphi(\nu)} \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)_{\theta_1} \mathrm{d}\theta_1 = \int_{\theta_{\mathrm{cut}}}^{\varphi(\nu)} \left(\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)_{\varphi(\nu)} - \int_{\theta_1}^{\varphi(\nu)} \left(\frac{\mathrm{d}^2 r}{\mathrm{d}\theta^2}\right) \mathrm{d}\theta \right) \mathrm{d}\theta_1 \\
&= b\tan(\eta)(\varphi(\nu) - \theta_{\mathrm{cut}}) + \int_{\theta_{\mathrm{cut}}}^{\varphi(\nu)} \int_{\theta_1}^{\varphi(\nu)} O(\eta)\, \mathrm{d}\theta\, \mathrm{d}\theta_1.
\end{aligned}$$

Recall that $b - \nu = b(1 - \cos\eta) = \frac{b\eta^2}{2} + O(\eta^4)$. The second term is $O(\eta^5)$ because we integrate twice on intervals of length $O(b - \nu) = O(\eta^2)$. The conclusion follows by substituting $b - \nu = \frac{b\eta^2}{2} + O(\eta^4)$ in Proposition 3.11. □

Proof of Proposition 3.13. Let $t(\nu)$ denote the cut time of c_ν, in other words $c_\nu(t(\nu)) = (a, \varphi(\nu))$. One can prove [22, p. 390] that $t'(\nu) = \nu\varphi'(\nu)$, so $t(\nu)$ is also analytic near $\nu = b$ and

$$t(\nu) = t_{\text{cut}} + \frac{(6\beta b - \alpha^2)\,\pi}{8\sqrt{2}\alpha^{5/2}b^{1/2}}(b-\nu) + O\left((b-\nu)^2\right). \tag{34}$$

Let $\tau(\nu)$ denote the distance between $(r_\nu, \theta_{\text{cut}})$ and $(a, \varphi(\nu))$ along c_ν. From formula (2.5) in [22] we obtain

$$\tau(\nu) = \nu(\varphi(\nu) - \theta_{\text{cut}}) + \int_{r_\nu}^{a} \frac{\sqrt{m(r)^2 - \nu^2}\,\mathrm{d}r}{m(r)} = \frac{(6\beta b - \alpha^2)\,\pi}{8\sqrt{2}\alpha^{5/2}b^{1/2}}(b-\nu) + O(\eta^4),$$

where in the last step we use the fact that the expression under the integral is $O(\eta)$ and that (from the last lemma) $a - r_\nu = O(\eta^3)$.

Comparing with (34) we see that $t_{\text{cut}} = t(\nu) - \tau(\nu) + O(\eta^4)$. This means that the distance from $(r_\nu, \theta_{\text{cut}})$ to $c_\nu(t_{\text{cut}})$ along c_ν is $O(\eta^4)$.

Hence $\operatorname{dist}(c_\nu(t_{\text{cut}}), (r_\nu, \theta_{\text{cut}})) = O(\eta^4)$, so the conclusion follows from Lemma 3.14. $\square$

For the proof of the next proposition we need one simple general result about 2-spheres of revolution.

Lemma 3.15. *Let M be any 2-sphere of revolution and let $0 < r_1 < r_2 < 2a$. The Riemannian distance between the parallels $\{r = r_1\}$ and $\{r = r_2\}$ is equal to $r_2 - r_1$.*

Proof. This follows from the fact that the meridians are minimal curves. $\square$

Proposition 3.16. *There exists a constant $A > 0$ such that the distance between (a, θ_{cut}) and the geodesic segment $c_\nu([t_{\text{cut}} - A; t_{\text{cut}} + A])$ equals $\frac{(6\beta b - \alpha^2)\sqrt{b}\pi}{16\sqrt{2}\alpha^{5/2}}\eta^3 + O(\eta^4)$.*

Proof. The role of the constant A is only to forbid the geodesic to do a full turn. Let $\widetilde{q}$ be a point on the geodesic segment under consideration. If $\widetilde{q}$ is after $(a, \varphi(\nu))$, the distance is at least $\varphi(\nu) - \theta_{\text{cut}} \gg \frac{(6\beta b - \alpha^2)\sqrt{b}\pi}{16\sqrt{2}\alpha^{5/2}}\eta^3$.

Suppose that $\widetilde{q} = (r_{\widetilde{q}}, \theta_{\widetilde{q}})$ lies on the geodesic segment $c_\nu([0, t(\nu)])$ and $\operatorname{dist}(\widetilde{q}, (a, \theta_{\text{cut}})) = O(\eta^3)$. From Lemma 3.15 we obtain $a - r_{\widetilde{q}} = O(\eta^3)$. Hence, by the triangle inequality, $|\theta_{\widetilde{q}} - \theta_{\text{cut}}| = O(\eta^3)$. Repeating the computation from Lemma 3.14 with $\theta_{\widetilde{q}}$ instead of θ_{cut} gives

$$a - r_{\widetilde{q}} = \frac{(6\beta b - \alpha^2)\sqrt{b}\pi}{16\sqrt{2}\alpha^{5/2}}\eta^3 + O(\eta^4).$$

Using again Lemma 3.15 we get

$$\operatorname{dist}(\widetilde{q}, (a, \theta_{\text{cut}})) \geq \frac{(6\beta b - \alpha^2)\sqrt{b}\pi}{16\sqrt{2}\alpha^{5/2}}\eta^3 + O(\eta^4).$$

$\square$

Remark 3.17. The idea behind this proof is that moving "vertically" is one order slower than "horizontally" (this intuition is used all the time throughout this section). If we can pass from a "vertical" estimate to a "horizontal" one without losing the order of precision (for example by the triangle inequality), the estimates auto-improve.

We can now prove that a 2-sphere of revolution satisfying assumptions (A1), (A2) and (A3) which is nonsingular (that is such that $G''(a) \neq 0$) satisfies the assumptions of Corollary 1.49.

Let $p := (a, 0)$ and $q := (a, \theta_{\mathrm{cut}})$. Suppose $v : (-\epsilon, \epsilon) \to T_pM$ is a smooth curve such that $\gamma_{v(0)}$ is the optimal geodesic (along the equator) and

$$\operatorname{dist}(\gamma_{v(t)}(2), q) = o(t^3).$$

In particular $(\mathrm{d}\exp_p)_{2v(0)}v'(0) = 0$, which by the Gauss Lemma [7, p. 367, Lemma 2] gives $\langle v(0), v'(0)\rangle = 0$ (we are using the usual identification $T_pM \sim T_{2v(0)}(T_pM)$ here). This means that $\eta(t) \sim t$, where $\eta(t)$ is the angle between $v(0)$ and $v(t)$. Hence

$$\operatorname{dist}(\gamma_{v(t)}(2), q) = o(\eta(t)^3),$$

which is in contradiction with Proposition 3.16, because $\alpha^2 \neq 6b\beta$.

Remark 3.18. The singular case $G''(a) = 0$ can be treated similarly. One obtains that if k is the order of the second nonzero term in the expansion (32), then there exists a smooth curve $v : (-\epsilon, \epsilon) \to T_pM$ such that $\gamma_{v(0)}$ is the optimal geodesic from p to q and $\operatorname{dist}(\gamma_{v(t)}(2), q) = O(t^{2k+1})$, but there exists no such curve with $\operatorname{dist}(\gamma_{v(t)}(2), q) = o(t^{2k+1})$.

3.5 Oblate ellipsoid as a 2-sphere of revolution

Let M be an ellipsoid with semi-axes b, b, c, where $b \geq c$. We denote p, q its northern and southern pole respectively. Clearly M is a 2-sphere of revolution.

It is a well-known fact (cf. for example [7, p. 173]) that the Gaussian curvature of such an ellipsoid is non-decreasing along a meridian, so assumption (A2) is satisfied.

Assumption (A1) is obviously true.

Let a be the distance from the equator to a pole. Let x be a point in the northern half of M and let ν be its distance from the rotational axis. Let R be the geodesic distance from x to p. Then $m(R) = \nu$.

On the other hand,

$$a - R = \int_\nu^b \sqrt{\frac{b^2 - c^2}{b^2} + \frac{c^2}{b^2 - t^2}}\,\mathrm{d}t.$$

Let $Z = \sqrt{b - \nu}$. After substitution $t = b - z^2$ and some operations on power series the integral above transforms into

$$\int_0^Z \frac{c\sqrt{2}}{\sqrt{b}} + \frac{(4b^2 - 3c^2)\, z^2}{2\sqrt{2}b^{3/2}c} + O(z^4)\,\mathrm{d}z,$$

which results in the expansion

$$a - R = \frac{\sqrt{2}c}{\sqrt{b}} Z + \frac{(4b^2 - 3c^2)}{6\sqrt{2}b^{3/2}c} Z^3 + O(Z^5).$$

By the theorem on locally invertible analytic functions we have that $\sqrt{b-\nu} = Z$ is an analytic function of R in a neighborhood of $R = a$. Explicitly, we obtain

$$\sqrt{b-\nu} = Z = \frac{\sqrt{b}}{\sqrt{2}c}(a-R) - \frac{\sqrt{b}\,(4b^2 - 3c^2)}{24\sqrt{2}c^5}(a-R)^3 + O\left((a-R)^5\right),$$

so finally, squaring both sides,

$$m(R) = \nu = b - \frac{b(a-R)^2}{2c^2} + \frac{b\,(4b^2 - 3c^2)}{24c^6}(a-R)^4 + O\left((a-R)^6\right).$$

We obtained the following

Proposition 3.19. *An oblate ellipsoid with axes b, b, c, where $b \geq c$, is a 2-sphere of revolution which satisfies assumptions (A1), (A2) and (A3). The function m has an expansion $m(r) = b - \alpha(a-r)^2 + \beta(a-r)^4 + O\left((a-r)^6\right)$ with $\alpha = \frac{b}{2c^2}$ and $\beta = \frac{b(4b^2-3c^2)}{24c^6}$.*

It is easy to find the singularity condition:

$$0 = 6b\beta - \alpha^2 = \frac{b^2(b^2 - c^2)}{c^6} \qquad \text{if and only if} \qquad b = c,$$

so spheres are the only singular oblate ellipsoids of revolution. Every other oblate ellipsoid is an example of a surface for which the assumptions of Corollary 1.49 hold.

References

[1] Andrei Agrachev, Davide Barilari, and Ugo Boscain. *Introduction to Riemannian and sub-Riemannian geometry.* `http://people.sissa.it/agrachev/agrachev_files/notes.html`.

[2] Andrei Agrachev and Yuri Sachkov. *Control Theory from the Geometric Viewpoint.* Encyclopædia of Mathematical Sciences. Cambridge University Press, 2004.

[3] Gérard Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Ann. Sci. Ecole Norm. Sup.*, (21):307–331, 1988.

[4] Davide Barilari, Ugo Boscain, and Robert W. Neel. Small time heat kernel asymptotics at the sub-Riemannian cut locus. *J. Differential Geometry.* To appear.

[5] Marcel Berger, Paul Gauduchon, and Edmond Mazet. *Le spectre d'une variété riemannienne*, volume 194 of *Lecture Notes in Mathematics.* Springer, 1971.

[6] Georges de Rham. Sur la réductibilité d'un espace de Riemann. *Comment. Math. Helv.*, (26):328–344, 1952.

[7] Manfredo P. do Carmo. *Differential geometry of curves and surfaces.* Prentice-Hall, 1976.

[8] Hans R. Fischer, Jerry J. Jungster, and Floyd L. Williams. The heat kernel on the two-sphere. *Advances in Mathematics*, (54):226–232, 1984.

[9] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian geometry.* Universitext. Springer, third edition, 2004.

[10] Alexander Grigor'yan. *Heat Kernel and Analysis on Manifolds*, volume 47 of *Studies in Advanced Mathematics.* AMS/IP, 2009.

[11] Detlef Gromoll and Wolfgang Meyer. On differentiable functions with isolated critical points. *Topology*, (8):361–369, 1969.

[12] Heinz Hopf and Willi Rinow. Über den Begriff der vollständigen differentialgeometrischen Fläche. *Comment. Math. Helv.*, (3):209–225, 1931.

[13] Elton P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics.* American Mathematical Society, 2002.

[14] Jin-ichi Itoh and Kauyoshi Kiyohara. The cut loci and the conjugate loci on ellipsoids. *Manuscripta Math.*, (114):247–264, 2004.

[15] Carl G. J. Jacobi. *Vorlesungen über Dynamik.* C. G. J. Jacobi's Gesammelte Werke. Berlin: Georg Reimer, second edition, 1884.

[16] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of Differential Geometry. Volume II.* Wiley Classics Library. Wiley, 1996.

[17] Rémi Léandre. Majoration en temps petit de la densité d'une diffusion dégénérée. *Probab. Theory Related Fields*, (74):289–294, 1987.

[18] S. A. Molchanov. Diffusion processes and Riemannian geometry. *Russian Mathematical Surveys*, 1, 1975.

[19] Robert Neel. The small-time asymptotics of the heat kernel at the cut locus. *Comm. Anal. Geom.*, (15):845–890, 2007.

[20] Henri Poincaré. Sur les lignes géodésiques des surfaces convexes. *Trans. Amer. Math. Soc.*, (6):237–274, 1905.

[21] Steve Rosenberg. *The Laplacian on a Riemannian Manifold.* Student Texts. Cambridge University Press, 1997.

[22] Robert Sinclair and Minoru Tanaka. The cut locus of a two-sphere of revolution and Topogonov's comparison theorem. *Tohoku Math. J.*, (59):379–399, 2007.

[23] S. R. S. Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.*, 20(2):431–455, 1967.