Construction of the Heat Kernel

First Step: Construction of the Parametrix

• We first construct a parametrix for the heat kernel, i.e. an approximate solution of the heat equation defined for \( x, y \) close.

Lemma 1. \( \Delta(fg) = (\Delta f)g - 2\langle df, dg \rangle + f\Delta g \).

Proof. Since the statement is a pointwise equality, we check it at a point \( x \) using Riemannian normal coordinates centered at \( x \). We arrange \( g_{ij}(x) = \delta_{ij} \) by choosing orthonormal coordinates in \( T_xM \), and we know that \( \partial_j g_{jk} = 0 \). Then

\[
\Delta = -\frac{1}{\sqrt{\det g}} \partial_x (\sqrt{\det g} g^{ij} \partial_x) = -\sum_{i=1}^n \frac{\partial^2}{(\partial x^i)^2},
\]

and

\[
\langle df, dg \rangle = \left\langle \frac{\partial f}{\partial x^i} dx^i, \frac{\partial g}{\partial x^j} dx^j \right\rangle = \sum_{i=1}^n \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^i} g^{ij}.
\]

• There exists \( \varepsilon > 0 \) such that for all \( x \in M \), the exponential map \( \exp_x \) takes \( B_{\varepsilon}(0) \subset T_xM \) diffeomorphically onto a neighborhood \( V_x \) of \( x \).

For \( y \in V_x \), set \( r(x,y) \) to be the length of the radial geodesic joining \( x \) to \( y \); note that \( r(x,y) < \varepsilon \).

We define a neighborhood of the diagonal in \( M \times M \) by

\[
U_\varepsilon = \{(x,y) \subset M \times M : y \in V_x, r(x,y) < \varepsilon \}.
\]

Then

\[
G(t,x,y) \equiv (4\pi t)^{-\frac{n}{2}} e^{-\frac{r^2(x,y)}{4t}} \in C^\infty(\mathbb{R}^+ \times U_\varepsilon).
\]

• We make the conjecture that the solution of the heat equation on \( U_\varepsilon \) is a modification of the Euclidean heat kernel.

Thus fix \( k \in \mathbb{Z}^+ \) and set

\[
S_k = S_k(y,x,y) = (4\pi t)^{-\frac{n}{2}} e^{-\frac{r^2(x,y)}{4t}} (u_0(x,y) + \cdots + u_k(x,y)t^k),
\]

for unknown functions \( u_i \in C^\infty(U_\varepsilon) \). We would like to have

\[
(\partial_t + \Delta_g)S_k = 0.
\]

Now

(1) \[
\frac{\partial S_k}{\partial t} = G \cdot \left( -\frac{n}{2t} + \frac{r^2}{4t^2} \right) (u_0 + \cdots + t^k u_k) + (u_1 + 2u_2 t + \cdots + ku_k t^{k-1})
\]

and, by Lemma 1,

(2) \[
\Delta_y S_k = (\Delta G)(u_0 + \cdots + u_k t^k) - 2\langle dG, d(u_0 + \cdots + u_k t^k) \rangle + G\Delta(u_0 + \cdots + u_k t^k),
\]

Also,

(3) \[
\Delta G = -\frac{\partial^2 G}{\partial r^2} - \frac{\partial G}{\partial r} \left( \frac{J'}{J} + \frac{n-1}{r} \right)
\]

\[
= \left( \frac{n}{2t} - \frac{r^2}{4t^2} \right) G + \frac{r}{2t} \frac{J'}{J} G,
\]

where \( J = \frac{\det(\Delta(r,\xi))}{r^n} \) and \( J' = \partial_r J \), and

Typeset by \texttt{AMS-\LaTeX}
\[ \langle dG, d(u_0 + \cdots + u_k t^k) \rangle = \left\langle \frac{\partial G}{\partial r} dr + \frac{\partial G}{\partial \theta} d\theta, d(u_0 + \cdots + u_k t^k) \right\rangle \]

\[ = \left\langle \frac{\partial G}{\partial r} dr, \frac{\partial u_k}{\partial r} dr + \cdots + t^k \frac{\partial u_k}{\partial r} dr + t^k \frac{\partial u_0}{\partial \theta} d\theta \right\rangle + t^k \frac{\partial u_0}{\partial \theta} d\theta \]

\[ = \frac{\partial G}{\partial r} \left( \frac{\partial u_0}{\partial r} + \cdots + t^k \frac{\partial u_k}{\partial r} \right) \]

\[ = -\frac{r}{2t} \left( \frac{\partial u_0}{\partial r} + \cdots + t^k \frac{\partial u_k}{\partial r} \right) G, \]

where \((\partial G/\partial \theta) d\theta\) has the obvious meaning. Combining (1)-(4), we obtain

\[ (\partial_t + \Delta_y) S_k = G \cdot \left( u_1 + 2u_2 t + \cdots + ku_k t^{k-1} + \frac{r}{2t} \frac{J'}{J} (u_0 + \cdots + t^k u_k) \right) \]

\[ + \frac{r}{t} \left( \frac{\partial u_0}{\partial r} + \cdots + t^k \frac{\partial u_k}{\partial r} \right) + \Delta_y u_0 + \cdots + t^k \Delta_y u_k. \]

While it is not possible in general to pick the \(u_i\) so the right hand side of (5) is zero, we can make this expression vanish up to the highest power of \(t\), i.e. we can solve

\[ (\partial_t + \Delta_y) S_k = (4\pi t)^{-n/2} e^{-r^2/(4t)} t^k \Delta_y u_k \]

by making all the terms containing \(t^{i-(n/2)-1}\) vanish, for \(i = 0, 1, \cdots, k\).

This leads to the following series of so-called transport equation:

The first equation is obtained by setting the coefficient of \(t^{i-(n/2)-1}\) equal to zero, and the second by doing the same for \(t^{i-(n/2)-1}\).

\[ r \frac{\partial u_0}{\partial r} + \frac{r}{2} \frac{J'}{J} u_0 = 0, \]

\[ r \frac{\partial u_i}{\partial r} + \left( \frac{r}{2} \frac{J'}{J} + i \right) u_i + \Delta_y u_{i-1} = 0, \quad i = 1, \cdots, k. \]

The equation (6) reduces to

\[ \frac{\partial \ln u_0}{\partial r} = -\frac{1}{2} \frac{\partial}{\partial r} \ln J, \]

so \(u_0 = k J^{-1/2}\), where \(k = k(\theta)\).

Since we want \(u_0\) to be defined at \(r = 0\), we must set \(k = \text{constant}\).

Setting \(k = 1\), we obtain

\[ u_0(x, y) = \frac{1}{\sqrt{J(\exp^{-1}(y))}} \]

and in particular

\[ u_0(x, y) = 1. \]
(II) To find the other $u_i$, we first solve a simpler version of (7),

$$r \frac{\partial u_i}{\partial r} + \left( \frac{r J'}{2} J + i \right) u_i = 0,$$

which has the solution $kr^{-i}J^{-1/2}$ with $k = k(\theta)$ arbitrary. This suggests us to assume that $u_i$ has the form

$$u_i = kr^{-i}J^{-1/2}, \text{ but with } k = k(r).$$

Plugging this expression into (7) gives

$$\frac{\partial k}{\partial r} = -J^{1/2}(\Delta u_i)rs^{-1}.$$

Let $x(s)$ be the unit speed geodesic from $x$ to $y$, $s \in [0, r]$. Thus $\Delta u_i$ is a function of $r$ along the geodesic, and the last equation can be solved by an $r$ integration.

Substituting this solution into the assumed form for $u_i$, we obtain

$$u_i(x, y) = -r^{-i}(x, y)J^{-\frac{1}{2}}(y) \int_0^r J^{\frac{1}{2}}(x(s)) [\Delta_y u_i(x(s), y)] s^{-1} ds. \quad (9)$$

In summary, for $u_i$ defined inductively by (8), (9), we have

$$(\partial_t + \Delta_y)S_k = \left(4\pi t \right)^{-\frac{n}{2}} e^{-\frac{r^2}{4t}} k \Delta u_k.$$
Lemma 4. $H_k$ is a parametrix if $k > \frac{n}{2}$. In fact,

$$(\partial_t + \Delta_y)H_k \in C^\ell((\mathbb{R}^+ \cup \{0\}) \times M \times M) \text{ if } k > \ell + \frac{n}{2}.$$ 

Proof. (i) First we show that $(\partial_t + \Delta_y)H_k$ extends to $t = 0$.

(a) On $\mathbb{R}^+ \times (M \times (M \setminus U_{\varepsilon}))$, $H_k \equiv 0$, and so $(\partial_t + \Delta_y)H_k$ trivially extends to $t = 0$.

(b) On $\mathbb{R}^+ \times U_{\varepsilon}/2$,

$$(\partial_t + \Delta_y)H_k = (\partial_t + \Delta_y)S_k = (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}t^k \Delta u_k \to 0, \text{ as } t \to 0.$$ 

Thus $(\partial_t + \Delta_y)H_k$ again extends by zero on this set.

(c) Finally, on $U_{\varepsilon} \setminus U_{\varepsilon}/2$,

$$(\partial_t + \Delta_y)H_k = \eta(\partial_t + \Delta_y)S_k - 2\langle d\eta, dS_k \rangle + (\Delta_y \eta)S_k = (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}t^k \phi(t, x, y),$$

for some function $\phi \in C^\infty(\mathbb{R}^+ \times M \times M)$ with at most a pole of order $t^{-1}$ at $t = 0$. Since $r > \frac{\varepsilon}{2}$, we may extend $(\partial_t + \Delta_y)H_k$ by zero.

(ii) To finish the proof, we must show that

$$\lim_{t \to 0} \int_M (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}\eta(x, y)(u_0(x, y) + \cdots + t^k u_k(x, y)) = f(x).$$

Now

$$\lim_{t \to 0} \int_M (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}\eta(x, y)u_i(x, y)f(y)dy$$

$$= \lim_{t \to 0} \int_{B_{\varepsilon}/2(x)} (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}\eta(x, y)u_i(x, y)f(y)$$

$$+ \lim_{t \to 0} \int_{M \setminus B_{\varepsilon}/2(x)} (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}\eta(x, y)u_i(x, y)f(y)dy.$$ 

(i) The second integral on the right hand side vanishes as $t \to 0$, since $r > \varepsilon/2$.

(ii) Using the exponential map as a coordinate chart, the first integral on the right-hand side becomes an ordinary integral over $\mathbb{R}^n \cong T_xM$:

$$\int_{B_{\varepsilon}/2(x)} (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}u_i(x, y)f(y)dy$$

$$= \int_{B_{\varepsilon}/2(x) \subset T_xM} (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}u_i(x, \exp_x v)f(\exp_x v) dv^1 \cdots dv^n$$

$$= \int_{T_xM} (4\pi t)^{-\frac{\ell}{2}}e^{-\frac{\ell^2}{4}}u_i(x, \exp_x v)f(\exp_x v) dv^1 \cdots dv^n,$$

with $u_i$ extended to be zero off $B_{\varepsilon}/2(x)$. 
\[ (4\pi t)^{-\frac{n}{2}} e^{-\frac{r^2}{4t}} \] is the ordinary kernel of \( \mathbb{R}^n \), as \( t \to 0 \), the last integral converges to

\[ u_i(x, \exp_x 0) f(\exp_x v) \mathcal{F}(0) = u_i(x, x) f(x). \]

Since \( u_0(x, x) = 1 \), we see that

\[ \lim_{t \to 0} \int_M (4\pi t)^{-\frac{n}{2}} e^{-\frac{r^2}{4t}} \eta(x, y) u_0(x, y) f(y) = f(x), \]

and

\[ \lim_{t \to 0} \int_{B_M(x)} (4\pi t)^{-\frac{n}{2}} e^{-\frac{r^2}{4t}} \eta(x, y) u_i(x, y) f(y) = 0, \quad \text{for } i > 1. \]

Final Step: Iterating Duhamel’s Formula

- Recall that we have a parametrix \( H_k(t, x, y) \) which is a good approximation to the heat kernel for \( x, y \) close and \( t \) small.
- To promote this approximation to a full heat kernel, we use the technique of iterating Duhamel’s formula.
- Let \( X, Y \) be operators on a Hilbert space of functions. We will assume that \( X, Y \) have well defined heat operators \( e^{-tX}, e^{-tY} \), i.e. a semigroup of bounded self-adjoint operators satisfying

\[ (\partial_t + X)e^{-tX}f = 0, \quad \lim_{t \to 0} e^{-tX}f = f, \]

and, similarly,

\[ (\partial_t + Y)e^{-tY}f = 0, \quad \lim_{t \to 0} e^{-tY}f = f. \]

In the following, we will denote expressions such as \( e^{-tX}f \) just by \( e^{-tX} \).

Duhamel’s Formula. If \( e^{-t(X+Y)} \) exists, we have

\[ e^{-t(X+Y)} = e^{-tX} - \int_0^t e^{-(t-s)(X+Y)}Ye^{-sX}ds. \]

Proof. (1.1) \( e^{-tX} \) is injective; indeed, if \( e^{-tX}f = 0 \), then

\[ 0 = (e^{-tX}f, f) = (e^{-\frac{t}{2}X}e^{-\frac{t}{2}X}f, f) = (e^{-\frac{t}{2}X}f, e^{-\frac{t}{2}X}f), \]

which implies \( e^{-\frac{t}{2}X}f = 0 \). Repeating this argument gives \( e^{-\frac{t}{2}X}f = 0 \), and hence

\[ f = \lim_{t \to 0} e^{-tX}f = 0. \]

(1.2) Also \( e^{-tX} \) is surjective, since it is injective and self-adjoint.

(1.3) Denote the inverse of \( e^{-tX} \) by \( e^{tX} \).

(2) Set \( B = e^{-t(X+Y)}e^{tX} \). Then

\[ \frac{dB}{dt} = e^{-t(X+Y)}(-(X+Y))e^{tX} + e^{-t(X+Y)}e^{tX}X = -e^{-t(X+Y)}Ye^{tX}, \]
since $e^{tX}X = Xe^{tX}$ by the spectral theory of unbounded operators. Thus
\[
e^{-t(X+Y)}e^{tX} - \text{Id} = - \int_0^t e^{-s(X+Y)}Ye^{sX}ds,
\]
and hence
\[
e^{-t(X+Y)} - e^{tX} = - \int_0^t e^{-s(X+Y)}Ye^{(s-t)X}ds
\]
\[= - \int_0^t e^{-(t-s)(X+Y)}Ye^{-sX}ds. \quad \Box
\]

- Given operators $A(t), B(t)$ on a Hilbert space, set
\[
A*B = \int_0^t A(t-s)B(s)ds.
\]

**Duhamel’s formula**. \(e^{-t(X+Y)} = e^{-tX} - e^{-t(X+Y)}* (Ye^{-tX})\).

- Denote the \(\lambda\)-fold product $A*\cdots A$ by $A^\lambda$ and set $A^1 = A$.

It is easy to check that $*$ is associative, so the notation is unambiguous.

**Corollary 5.** We have
\[
e^{-t(X+Y)} = e^{tX} + \sum_{j=1}^n (-1)^j b_j + (-1)^{n+1}r_{n+1},
\]
where
\[b_n = e^{-tX} * (Ye^{-tX})^n
\]
and
\[r_n = e^{-(X+Y)} * (Ye^{-tX})^n.
\]

**Proof.** The proof is by induction.

(i) For $n = 0$, the corollary reduces to Duhamel’s formula with $b_0 = 0$.

(ii) Assume the formula is true for $n - 1$, namely
\[
e^{-t(X+Y)} = e^{tX} + \sum_{j=1}^{n-1} (-1)^j b_j + (-1)^nr_n.
\]

To show that this is true for $n$, we apply Duhamel’s formula to the $e^{-t(X+Y)}$ term in $r_n$ to obtain
\[
r_n = (e^{-tX} - e^{-t(X+Y)}* (Ye^{-tX})) * (Ye^{-tX})^n
\]
\[= b_n - r_{n+1}. \quad \Box
\]
Given $X, Y$, we can construct the heat operator for $X+Y$ from the heat operator of $X$ by Corollary 5; namely,

\[(10) \quad e^{-t(X+Y)} = e^{-tX} + e^{-tX} \ast \sum_{\lambda=1}^{\infty} (-1)^\lambda (Ye^{-tX})^{*\lambda},\]

provided $r_{n+1} \to 0$ in the Hilbert space.

We now adopt this expression formally to our setup.

Let $A(t)$ be the operator on $L^2(M,g)$ with kernel $H(t, x, y)$. We may dangerously assume that $A(t)$ is like a heat operator, so that there exists an operator $X$ on $L^2$ such that $(\partial_t + X)A(t) = 0$. Thus

\[e^{-tX} = A(t) \quad \text{and} \quad (\partial_t + X)H = 0.\]

Set $Y = \Delta - X$ and $K = (\partial_t + \Delta)H$.

Note that the kernel of $YA(t)$ is just $YxH(t, x, y)$.

Extend the Hilbert space $H$ of functions to the space of time dependent functions $H \times \mathbb{R}$. Then $e^{-tX}$ acts on this space by

\[e^{-tX} \psi(t, x) = \int_0^t d\theta \int_M H(t - \theta, x, z)\psi(\theta, z) d\text{vol}(z).\]

Thus if $B(t)$ is any operator with kernel $B(t, x, y)$, $A \ast B$ has kernel

\[\int_0^t d\theta \int_M H(t - \theta, x, z)B(t - \theta, q, y) dq.\]

Interpreting (10) at the level of kernels gives

\[e(t, x, y) = H(t, x, y) + [H \ast \sum_{\lambda=1}^{\infty} (-1)^\lambda ((\Delta - X)H)^{*\lambda}](t, x, y)\]

\[= H(t, x, y) + [H \ast \sum_{\lambda=1}^{\infty} (-1)^\lambda ((\partial_t + \Delta)H)^{*\lambda}](t, x, y)\]

\[= H(t, x, y) + [H \ast \sum_{\lambda=1}^{\infty} (-1)^\lambda K^{*\lambda}](t, x, y).\]

This is our formal expression for the heat kernel.

It is remarkable that this formal procedure actually works, as we now show.

The hard work involves controlling the error terms $r_n$.

To begin, for $A, B \in C^0((\mathbb{R}^+ \cup \{0\}) \times M \times M)$, note that

\[(A \ast B)(t, x, y) = \int_0^t d\theta \int_M A(\theta, x, q)B(t - \theta, q, y) dq \in C^0((\mathbb{R}^+ \cup \{0\}) \times M \times M).\]
Lemma 6. Set $K_k = (\partial_t + \Delta_y) H_k$. Then

$$Q_k = \sum_{\lambda=1}^{\infty} (-1)^{\lambda + 1} K_k^{*\lambda}$$

exists and is in $C^\ell((\mathbb{R}^+ \cup \{0\}) \times M \times M)$ if $k > \ell + \frac{n}{2}$.

Moreover, given $T > 0$, $\exists C = C(T)$ so that $|Q_k(t, x, y)| \leq Ct^{k-\frac{n}{2}}$ for all $t \in [0, T]$.

Proof. Writing $K_k = (\partial_t + \Delta_y) (\eta S_k)$ and performing the differentiation, we obtain

$$|K_k| \leq A(T) t^{k-\frac{n}{2}} \leq A(T) T^{k-\frac{n}{2}} \overset{\text{def.}}{=} B,$$

for some constant $A = A(T)$.

Claim:

(11) $$|K_k^{*\lambda}(t, x, y)| \leq \frac{AB^{\lambda-1} V^{\lambda-1} t^{k-(n/2)+\lambda-1}}{(k-n/2 + 1)(k-n/2 + 2) \cdots (k-n/2 + \lambda - 1)},$$

where $V = \text{vol}(M)$.

(i) The case $\lambda = 1$ has just been done, provided we readjust the definition of $A$.

(ii) Assuming the claim for $\lambda - 1$, we have

$$|K_k^{*\lambda}| \leq \int_0^t \int_M |K_k^{*\lambda-1}(\theta, x, q)||K_k(t-\theta, q, y)| dq$$

$$\leq \int_0^t \int_M \frac{AB^{\lambda-2} V^{\lambda-2} t^{k-(n/2)+\lambda-2}}{(k-n/2 + 1)(k-n/2 + 2) \cdots (k-n/2 + \lambda - 2)} B$$

$$= \frac{AB^{\lambda-1} V^{\lambda-2} V}{(k-n/2 + 1)(k-n/2 + 2) \cdots (k-n/2 + \lambda - 2)} \int_0^t \theta^{k-(n/2)+\lambda-2}.$$ 

Evaluating the last integral finishes the claim.

— The right hand side of (11) is bounded by a constant times

(12) $$\frac{(BVt)^{\lambda-1} t^{k-(n/2)}}{\Gamma(k-n/2 + \lambda - 2)},$$

so the ratio test shows that $\sum_{\lambda=1}^{\infty} |K_k^{*\lambda}|$ converges if $k > \frac{n}{2}$.

— This implies that $\sum_{\lambda=1}^{\infty} (-1)^{\lambda + 1} K_k^{*\lambda}$ converges to a continuous function in $t, x, y$ if $k > \frac{n}{2}$.

⊙ The estimate (12) also implies $|Q_k| \leq C \cdot t^{k-(n/2)}$ for some constant $C$.

• Similar estimates on the derivatives of $K_k^{*\lambda}$ shows that $Q_k \in C^\ell$ if $k > \ell + \frac{n}{2}$. □
Lemma 7. (i) If \( P \in C^0(\mathbb{R}^+ \cup \{0\}) \times M \times M \), then
\[
P * H_k \in C^\ell(\mathbb{R}^+ \times M \times M), \quad \text{if } k > \ell + \frac{n}{2}.
\]
(ii) \((\partial_t + \Delta)(P * H_k) = P + P * K_k\) if \(k > 2 + \frac{n}{2}\).

Proof. (ii) By the Leibniz rule,
\[
(\partial_t + \Delta)(P * H_k) = \int_0^t d\theta \int_P(\theta, x, q)H_k(t - \theta, q, y)dq
\]
\[
= \lim_{s \to t} \int_0^t d\theta \int_P(\theta, x, q)H_k(t - s, q, y)dq
\]
\[
= P(t, x, y)
\]
\[
= \int_0^t d\theta \int_P(\theta, x, q) \cdot \partial_t H_k(t - \theta, q, y)dq
\]
\[
= P + P * K_k.
\]

Theorem 8. Set
\[
\epsilon(t, x, y) = H_k(t, x, y) - Q_k * H_k(t, x, y).
\]
Then \(\epsilon(t, x, y) \in C^\infty(\mathbb{R}^+ \times M \times M)\) is independent of \(k\) if \(k > 2 + \frac{n}{2}\) and is the heat kernel.

Proof. By Lemma 7,
\[
(\partial_t + \Delta_\gamma)\epsilon(t, x, y) = (\partial_t + \Delta_\gamma)(H_k - Q_k * H_k)
\]
\[
= K_k - Q_k - Q_k * K_k
\]
\[
= K_k - \sum_{\lambda=1}^\infty (-1)^{\lambda+1} K^\lambda_k - \sum_{\lambda=1}^\infty (-1)^{\lambda+1} K^\lambda_k * K_k
\]
\[
= 0.
\]
We also have
\[
\lim_{t \to 0} \int e(t, x, y)f(y)dy = \lim_{t \to 0} \left( \int_M H_k(t, x, y)f(y)dy - \int_M (Q_k * H_k)(t, x, y)f(y)dy \right)
\]
\[
= -f(x) - \lim_{t \to 0} \int_M (Q_k * H_k)(t, x, y)f(y)dy.
\]
Now \(R_k \equiv Q_k/t^{k-(n/2)}\) is bounded for all \(t\) in some finite interval, so by Lemma 4 for \(k > \frac{n}{2}\),
\[
\lim_{t \to 0} \int_M (Q_k * H_k)(t, x, y)f(y)dy = \lim_{t \to 0} t^{k-(n/2)} \int_M (R_k * H_k)(t, x, y)f(y)dy = 0.
\]
Thus
\[
\lim_{t \to 0} \int e(t, x, y)f(y)dy = f(x).
\]