1. Introduction

Fourier in 1822 was the first to derive the heat equation in the following context: assume \( M \subset \mathbb{R}^3 \) is a sufficiently homogeneous body. Then the temperature function \( u : (0, \infty) \times M \to [0, \infty) \) of \( M \), that is, \( u(t, x) \) is the temperature at the time \( t \) in \( x \in M \), satisfies the heat equation

\[
\partial_t u(t, x) = k \Delta_x u(t, x)
\]

if \( M \) has no sources or sinks of heat. Above, \( k > 0 \) is a material constant (heat conductivity constant), and \( \Delta = \sum_{j=1}^m \partial_j^2 \) is the Laplace operator.

The heat equation was also the basis for modern probability theory: in 1827 the botanist Brown was watching small test particles (pollen,...) in suspended in a fluid medium (water,...) in a body \( M \subset \mathbb{R}^3 \) and was shocked by the fact that the pollen is moving. Having started with pollen, his first conclusion was that pollen is alive, until he repeated the experiment with other test particles that were.

His observations were that the trajectory \( X \) of each test particle was random and independent of the trajectory of any other test particle (so wlog we can consider one test particle). This leads to the idea that \( X \) should be what we call today a stochastic process, that is, a map

\[
X : [0, \infty) \times (\Omega, \mathcal{F}, P) \to M,
\]

where \((\Omega, \mathcal{F}, P)\) is a probability space. Here, the set \( \Omega \) contains the random parameters and for each fixed \( \omega \in \Omega \), the map

\[
X(\omega) : [0, \infty) \to M
\]

is called a (random) path of the process. Then Brown observed that the expected displacement of the test particle was a decreasing function of its size and of viscosity of the medium, and increasing with the temperature of the medium.

Let \( u(\cdot, \cdot, y) : (0, \infty) \times M \to [0, \infty), (t, x, y) \mapsto u(t, x, y) \) denote the probability density of \( X \), assuming that \( X \) starts in some \( y \in M \). In other words, the probability of finding \( X \) in \( A \subset M \) at the time \( t \) is given by

\[
P\{X_t \in A\} = \int_A u(t, x, y)dy.
\]
It was then Einstein who derived in 1905 that 
\[
\frac{\partial}{\partial t} u(t, x, y) = D \Delta_x u(t, x, y),
\]
where the diffusion constant \(D > 0\) of the system is given by 
\[
kT \frac{6\pi \nu R}{\theta^2},
\]
where \(k\) is the Boltzmann constant, \(T\) the temperature of the medium, \(\nu\) its viscosity and \(R\) the radius of the test particle. Assuming that \(u(t, x, y)\) behaves like the Gauss kernel 
\[
p : (0, \infty) \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow [0, \infty),
p(t, x, y) := (4\pi t)^{3/2} e^{-\frac{|x-y|^2}{4t}},
\]
which is a solution of the heat equation in \((t, x)\), one easily derives the fundamental relation 
\[
\int_{\Omega} |X(t) - y|^2 dP \approx Dt,
\]
for the average square displacement, which justifies all observations of Brown. The stochastic process underlying the random trajectory of a test particle as above is nowadays called a Brownian motion.

Einstein’s conclusion was that the medium consists of very small particles (molecules), subject to some random kinematics, which bombard the larger test particles and lead to their random movement. The above fundamental relation (1) was confirmed in an experiment by Perrin in 1908 for which he received the Nobel price later. Note that all of this is roughly 20 years before quantum mechanics, and so these results can be thought of as a first confirmation of the atomic structure of matter.

Let us now take a closer look at the properties of the \(m\)-dimensional Gauss kernel 
\[
p : (0, \infty) \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow [0, \infty),
p(t, x, y) := (4\pi t)^{-m/2} e^{-\frac{|x-y|^2}{4t}},
\]
In the sequel, we are going to consider \(\mathbb{R}^m\) as a Riemannian manifold with its Euclidean metric \(g_{ij}(x) = \delta_{ij}\). Then \(\Delta\) is the underlying Laplace-Beltrami operator, the Lebesgue measure \(dx\) becomes the Riemannian volume measure and the Euclidean distance \(|x - y|\) the Riemannian distance. One can then prove the following facts for the Riemannian manifold \(\mathbb{R}^m\):

i) \((t, x, y) \mapsto p(t, x, y)\) is jointly smooth and \((t, x) \mapsto p(t, x, y)\) satisfies 
\[
\partial_t u(t, x) = \Delta_x u(t, x), \quad \lim_{t \to 0^+} u(t, \cdot) = \delta_y \quad \text{for all } y \in \mathbb{R}^m,
\]
ii) one has 
\[
\int_{\Omega} p(t, x, y) dy = 1 \quad \text{for all } t > 0, x \in \mathbb{R}^m,
\]
iii) one has \(p(t, x, y) = p(t, y, x)\) for all \(t > 0, x, y \in \mathbb{R}^m\),
iv) one has 
\[
p(t + s, x, y) = \int p(t, x, z)p(s, y, z)d^m z \quad t, s > 0, x \in \mathbb{R}^m
\]
v) one has $p > 0$,

and $p$ is the unique nonnegative function satisfying (2),

vi) there is exactly one self-adjoint realization $H \geq 0$ of $-\Delta$ in the Hilbert space $L^2(\mathbb{R}^m)$ and one has

$$e^{-tH} f(x) = \int p(t, x, y) f(y) dy \quad \text{for all } f \in L^2(\mathbb{R}^m),$$

where the heat semigroup on the left hand side is defined via the spectral theorem,

vii) if $m \leq 2$ we have

$$G(x, y) := \int_0^\infty p(t, x, y) dt = \infty \quad \text{for all } x, y \in \mathbb{R}^m,$$

while for $m \geq 3$ we have

$$G(x, y) < \infty \quad \text{for all } x, y \in \mathbb{R}^m \text{ with } x \neq y.$$

In this course we will attack the following problem: to what extend hold the above results on Riemannian manifolds? It is easy to convince oneself that some subtleties must appear: for example, even if we replace $\mathbb{R}^m$ above with an arbitrary bounded open subset $U$ of $\mathbb{R}^m$, then there exist at least two nonnegative solutions to (2) in $U$ and $\Delta$ has at least two self-adjoint realizations in $L^2(U)$: the Dirichlet realization and the Neumann realization, and the integral kernels of the corresponding heat semigroups both solve (2) in $U$.

Assume now $(M, g)$ is an arbitrary Riemannian manifold, let $\Delta_g$ denote the induced Laplace-Beltrami operator, let $\mu_g$ denote the Riemannian volume measure, and $d_g(x, y)$ the Riemannian distance. The first question is: what is the analogue of the Gauss kernel and of $H$ in this case? Firstly, we are going to show that $\Delta_g$ has a canonically given self-adjoint realization $H_g \geq 0$ in $L^2(M, g)$, its Friedrichs realization. Then one can define $p_g(t, x, y)$ as the integral kernel of the heat semigroup of $H_g$. These construction rely on some several deep results from functional analysis (which is why we are actually going to start with abstract functional analysis).

It will turn out that without any further assumptions on the geometry,

- the map

  $$p_g : (0, \infty) \times M \times M \to [0, \infty)$$

  is jointly smooth, and $(t, x) \mapsto p_g(t, x, y)$ satisfies

  $$\partial_t u(t, x) = \Delta_{g, x} u(t, x), \quad \lim_{t \to 0^+} u(t, \cdot) = \delta_y \quad \text{for all } y \in M,$$

  which is why one calls $p_g$ the heat kernel of $(M, g)$,

- one has

  $$\int p_g(t, x, y) d\mu_g(y) \leq 1 \quad \text{for all } t > 0, x \in M,$$

  one has the natural analogues of iii) and iv),

- and the analogue of vi) holds by definition.

Moreover, we are going to address some of the following facts:
In general, \( H_g \) need not be the unique self-adjoint realization of \(-\Delta_g\), but we are going to show that this is the case if \((M, d_g)\) is complete.

The property \( \int p_g(t, x, y) d\mu_g(y) = 1 \) turns out to be a property which is highly subtle for noncompact Riemannian manifolds and typically depends on the growth of the volume of metric balls. Riemannian manifolds having this property are called stochastically complete.

While one always has \( p_g \geq 0 \), the strict positivity \( p_g > 0 \) turns out to be related with the connectedness of \( M \).

The property \( G(x, y) = \int_0^\infty p_g(t, x, y) dt < \infty \) for \( x \neq y \) turns out to be subtle again: it implies the noncompactness of \( M \) and there are noncompact Riemannian manifolds of dimension \( \geq 3 \) which need not satisfy the above finiteness, which is called nonparabolicity.

Both, stochastic completeness and nonparabolicity are linked with probability theory: on every Riemannian manifold one can define Brownian motion, and stochastic completeness means that this process cannot explode in a finite time, while nonparabolicity means that the process eventually leaves every relatively compact subset.

the Gauss type behaviour of the heat kernel
\[
C_1 t^{\dim(M)/2} e^{-\frac{d_g(x,y)^2}{2t}} \leq p_g(t, x, y) \leq C_3 t^{\dim(M)/2} e^{-\frac{d_g(x,y)^2}{C_4 t}}
\]
depends sensitively on the geometry of \((M, g)\): in fact, it turns to be more natural in the above estimate to replace the factor \( t^{\dim(M)/2} \) by the volume of a Riemannian ball centered in \( x \) with radius \( \sqrt{t} \). These are the celebrated Li-Yau heat kernel estimates.

2. Linear operators in Banach and Hilbert spaces

2.1. Motivation. For the convenience of the reader, we collect some facts linear operators. For a detailed discussion of the (standard) results below, we refer the reader to [35, 28, 22].

This section is motivated by the following observations from linear algebra: Assume a linear operator \( T \) in a (say) complex finite dimensional Hilbert space \( \mathcal{H} \cong \mathbb{C}^l \) is given. Then for every \( \psi_0 \in \mathcal{H} \) there is a unique solution \( \Psi : [0, \infty) \to \mathcal{H} \) of the 'heat equation'
\[
(d/dt)\Psi(t) = -T\Psi(t), \quad \Psi(0) = \psi_0.
\]
In fact, we can simply set \( \Psi(t) = e^{-tT} \psi_0 \), with
\[
e^{-tT} = \sum_{j=0}^{\infty} (-tT)^j / j!
\]
the matrix exponential series. Now if \( \mathcal{H} \) is infinite dimensional (in our case this will the Hilbert space of square integrable functions on a Riemannian manifold), for \( T \)'s one is
interested in (in our case: the Laplace-Beltrami operator), the exponential series will never converge. The way out of this is provided by the following observation: assume in the above finite dimensional situation that $T$ is \textit{self-adjoint}. Then, as $T$ is diagonalizable, one has

$$T = \int \lambda \, dP_T(\lambda) := \sum_{\{\lambda \in \mathbb{R}: \lambda \text{ is an eigenvalue of } T\}} \lambda P_T(\lambda)$$

with $P_T(\lambda)$ the projection onto the eigenspace of $\lambda$. Given a function $f : \mathbb{R} \to \mathbb{C}$ (like $f(r) = e^{-tr}$!), the above formula suggests to define a linear operator $f(T)$ in $\mathcal{H}$ by setting

$$f(T) := \int f(\lambda) \, dP_T(\lambda) := \sum_{\{\lambda \in \mathbb{R}: \lambda \text{ is an eigenvalue of } T\}} f(\lambda) P_T(\lambda).$$

For $f(r) = e^{-tr}$ this definition is equivalent to using the matrix exponential.

By John von Neumann’s spectral theorem, it turns out that given any self-adjoint operator $T$ in a possibly infinite dimensional Hilbert space there exists a unique projection-valued measure $P_T$ such that one has

$$T = \int \lambda \, dP_T(\lambda),$$

and using the above observations this fact leads to satisfactory solution theory of the abstract heat equation induced by $T$. The purpose of this section is to explain these facts in detail. Before that, let us list some issues that are supposed to motivate some of the following definitions:

- a self-adjoint operator $T : \mathcal{H} \to \mathcal{H}$ in a Hilbert space is automatically continuous. However, the operators we will be interested in (like the Laplace-Beltrami operator) turn out to be never continuous. The way out of this is consider linear operators $T : \text{Dom}(T) \to \mathcal{H}$ that are defined on a (typically dense) subspace $\text{Dom}(T) \subset \mathcal{H}$, called the domain of definition of $T$. Thus: any self-adjoint operator $T$ in $\mathcal{H}$ with $\text{Dom}(T) = \mathcal{H}$ is automatically continuous, the self-adjoint operators of interest are not continuous (and so cannot be defined everywhere). Although self-adjoint operators are not continuous, they turn out to satisfy a weaker useful property, namely they are closed.

- In infinite dimensions, it is often easier to define a self-adjoint operator via symmetric sesquilinear forms. Note that in finite dimensions, given any symmetric sesquilinear form

$$Q : \mathcal{H} \times \mathcal{H} \longrightarrow \mathbb{C}$$

there exists a unique self-adjoint operator $T_Q : \mathcal{H} \to \mathcal{H}$ such that

$$Q(\Psi_1, \Psi_2) = \langle T_Q \Psi_1, \Psi_2 \rangle.$$ 

In the infinite dimensional case, again domain of definition questions arise and, in particular, one needs the sesquilinear form to be bounded from below in a certain sense in order that it induces a self-adjoint operator (which is then also bounded from below).
2.2. Facts about linear operators in Banach and Hilbert spaces. We understand all our normed spaces to be over \( \mathbb{C} \). As we have explained above, it is essential to require a linear operator \( T \) between Banach spaces \( B_1 \) and \( B_2 \) to be only defined on a subspace \( \text{Dom}(T) \subset B_1 \), called its domain of definition, so that \( T \) is actually a linear map \( T : \text{Dom}(T) \to B_2 \).

Its image range \( \text{Ran}(T) \subset B_2 \) is defined to be the linear space of all \( f_2 \in B_2 \) for which there exists \( f_1 \in \text{Dom}(T) \) with \( Tf_1 = f_2 \). Its kernel \( \text{Ker}(T) \) is given by all \( f \in \text{Dom}(T) \) with \( Tf = 0 \).

Such a linear operator \( T \) is called bounded, if there exists a constant \( C \geq 0 \) such that \( \|Tf\| \leq C \|f\| \) for all \( f \in \text{Dom}(T) \), and the smallest such \( C \) is called the operator norm of \( T \) and denoted by \( \|T\| \). Boundedness of \( T \) is equivalent to its continuity as a map between normed spaces (considered as metric and thus topological spaces in the usual way). If \( \text{Dom}(T) \) is dense, then \( T \) can be uniquely extended to a bounded linear map \( B_1 \to B_2 \), which will be denoted with the same symbol again. The linear space of bounded linear operators is denoted by \( L(B_1, B_2) \) and becomes a Banach itself with the above operator norm. One sets \( L(B_1) := L(B_1, B_1) \).

Theorem 2.1 (Closed graph theorem). A linear operator \( T \) from \( B_1 \) to \( B \) is bounded, if and only if its graph

\[
\{(f_1, f_2) \in \text{Dom}(T) \times B_2 : Tf_1 = f_2\} \subset B_1 \times B_2
\]

is closed.

We also record:

Theorem 2.2 (Uniform boundeness principle). For a subset \( A \subset L(B_1, B_2) \) the following conditions are equivalent:

- for all \( f \in B_1 \) there exists a constant \( C_f \geq 0 \) with \( \|Tf\| \leq C_f \) for all \( T \in A \).
- there exists a constant \( C \geq 0 \) with \( \|T\| \leq C \) for all \( T \in A \).

Let \( \mathcal{H} \) be a separable Hilbert space. The underlying scalar product, which is assumed to be antilinear in its first slot, will be simply denoted by \( \langle \cdot, \cdot \rangle \), and the induced norm (as well as the induced operator norm) is denoted by \( \|\cdot\| \).

Theorem 2.3 (Riesz-Fischer’s duality theorem). Assume \( T \in L(\mathcal{H}, \mathbb{C}) \), that is, \( T \) is a linear continuous functional on \( \mathcal{H} \). Then there exists a unique \( f_T \in \mathcal{H} \) such that for all \( h \in \mathcal{H} \) one has

\[
T(h) = \langle f_T, h \rangle.
\]

The map \( T \mapsto f_T \) induces an antilinear isometric isomorphism between \( L(\mathcal{H}, \mathbb{C}) \) and \( \mathcal{H} \).
If $\tilde{H}$ is another separable complex Hilbert space case and $R$ is a densely defined linear operator from $H$ to $\tilde{H}$, then the adjoint $R^*$ of $R$ is a linear operator from $\tilde{H}$ to $H$ which is defined as follows: $\text{Dom}(R^*)$ is given by all $f \in \tilde{H}$ for which there exists $f^* \in H$ such that

$$\langle f^*, h \rangle = \langle f, Rh \rangle$$

for all $h \in \text{Dom}(R)$, and then $R^* f := f^*$.

In the sequel, let $S$ and $T$ be arbitrary linear operators in $H$. Firstly, $T$ is called an extension of $S$ (symbolically $S \subset T$), if $\text{Dom}(S) \subset \text{Dom}(T)$ and $S f = T f$ for all $f \in \text{Dom}(S)$.

If $S$ is densely defined, then $S$ is called symmetric, if $S \subset S^*$ and self-adjoint if $S = S^*$.

Clearly, self-adjoint operators are symmetric. Note for the symmetry of only needs to check that it is densely defined

$$\langle S f_1, f_2 \rangle = \langle f_1, S f_2 \rangle$$

for all $f_1, f_2 \in \text{Dom}(S)$. Checking self-adjointness is a tricky business for unbounded operators, while checking symmetry is very easy:

**Example 2.4.** Assume $U \subset \mathbb{R}^m$ is open and the operator $S := -\Delta = -\sum_{j=1}^m \partial_j^2$ in the complex Hilbert space $L^2(U)$ is given the domain of definition $\text{Dom}(S) := C_c^\infty(U)$. Then $S$ is symmetric: for all $f_1, f_2 \in \text{Dom}(S) = C_c^\infty(U)$ by Stokes’ Theorem one has

$$\langle S f_1, f_2 \rangle = \int_U (-\Delta) f_1 f_2 dx$$

$$= \int_U (\nabla f_1, \nabla f_2) dx + \text{a boundary term that vanishes because } f_j \text{ is compactly supported in } U$$

$$= \int_U \overline{f_1} (-\Delta) f_2 dx = \langle f_1, S f_2 \rangle .$$

This operator is not closed and so surely not self-adjoint (in general it has many self-adjoint extensions; in case $U = \mathbb{R}^m$ it has precisely one self-adjoint extension).

The operator $S$ is called *semibounded* (from below), if there exists a constant $C \geq 0$ such that for all $f \in \text{Dom}(S)$ one has

$$\langle S f, f \rangle \geq -C \| f \|^2 ,$$

or in short: $S \geq -C$. Since $H$ is assumed to be complex, semibounded operators are automatically symmetric (by complex polarization).

$S$ is called *closed*, if whenever $(f_n) \subset \text{Dom}(S)$ is a sequence such that $f_n \to f$ for some $f \in H$ and $S f_n \to h$ for some $h \in H$, then one has $f \in \text{Dom}(S)$ and $S f = h$.

$S$ is called *closable*, if it has a closed extension. In this case, $S$ has a smallest closed extension $\overline{S}$, which is called the *closure of $S$*. The closure $\overline{S}$ is determined as follows: $\text{Dom}(\overline{S})$ is given by all $f \in H$ for which there exists a sequence $(f_n) \subset \text{Dom}(S)$ such that $f_n \to f$ and such that $(S f_n)$ converges, and then $\overline{S} f := \lim_n S f_n$.
Adjoints of densely defined operators are closed, so that that symmetric operators are closable; self-adjoint operators are closed. Bounded operators are always closed by the closed graph theorem. If $S$ is densely defined and closable, then $S^*$ is densely defined and $S^{**} = \overline{S}$. If $T$ is symmetric, then $T$ is called essentially self-adjoint, if $\overline{T}$ is self-adjoint. Then $\overline{T}$ is the unique self-adjoint extension of $T$.

We record:

**Theorem 2.5.** Assume that $S$ is semibounded (in particular symmetric) with $S \geq -C$ for some constant $C \geq 0$. Then $S$ is essentially self-adjoint, if and only if there exists $z \in \mathbb{C} \setminus [-C, \infty)$ such that $\ker((S - z)^*) = \{0\}$.

The resolvent set $\rho(S)$ is defined to be the set of all $z \in \mathbb{C}$ such that $S - z$ is invertible as a linear map $\text{Dom}(S) \to \mathcal{H}$ and is in addition bounded as a linear operator from $\mathcal{H}$ to $\mathcal{H}$. If $S$ is closed and $(S - z)^{-1}$ invertible, then $(S - z)^{-1}$ is automatically bounded by the closed graph theorem. The spectrum $\sigma(S)$ of $S$ is defined as the complement $\sigma(S) := \mathbb{C} \setminus \rho(S)$. Resolvent sets of closed operators are open, therefore spectra of closed operators are always closed.

A number $z \in \mathbb{C}$ is called an eigenvalue of $S$, if $\ker(S - z) \neq \{0\}$. In this case, $\dim \ker(S - z)$ is called the multiplicity of $z$, and each $f \in \ker(S - z) \setminus \{0\}$ is called an eigenvector of $S$ corresponding to $z$. Of course each eigenvalue is in the spectrum. The eigenvalues of a symmetric operator are real, and the eigenvectors corresponding to different eigenvalues of a symmetric operator are orthogonal. A simple result that reflects the subtlety of the notion of a “self-adjoint operator” when compared to that of a “symmetric operator” is the following: A symmetric operator in $\mathcal{H}$ is self-adjoint, if and only if its spectrum is real. If $S$ is self-adjoint, then $S \geq -C$ for a constant $C \geq 0$ is equivalent to $\sigma(S) \subset [-C, \infty)$ (cf. Satz 8.26 in [36]).

The essential spectrum $\sigma_{\text{ess}}(S) \subset \sigma(S)$ of $S$ is defined to be the set of all eigenvalues $\lambda$ of $S$ such that either $\lambda$ has an infinite multiplicity, or $\lambda$ is an accumulation point of $\sigma(S)$. Then the discrete spectrum $\sigma_{\text{dis}}(S) \subset \sigma(S)$ is defined as the complement

$$\sigma_{\text{dis}}(S) := \sigma(S) \setminus \sigma_{\text{ess}}(S).$$

As every isolated point in the spectrum of a self-adjoint operator is an eigenvalue (cf. Folgerung 3, p. 191 in [35]), it follows that in case of $S$ being self-adjoint, the set $\sigma_{\text{dis}}(S)$ is precisely the set of all isolated eigenvalues of $S$ that have a finite multiplicity.

Let $\tilde{\mathcal{H}}$ be another complex separable Hilbert space. We recall that given $q \in [1, \infty)$, some $K \in \mathcal{L}(\mathcal{H}, \tilde{\mathcal{H}})$ is called

- **compact**, if for every orthonormal sequence $(e_n)$ in $\mathcal{H}$ and every orthonormal sequence $(f_n)$ in $\tilde{\mathcal{H}}$ one has $\langle Ke_n, f_n \rangle \to 0$ as $n \to \infty$
• $q$-summable (or an element of the $q$-th Schatten class of operators $\mathcal{H} \to \tilde{\mathcal{H}}$), if for every $(e_n), (f_n)$ as above one has
\[
\sum_n |\langle Ke_n, f_n \rangle|^q < \infty.
\]

Let us denote the class of compact operators with $J^\infty(\mathcal{H}, \tilde{\mathcal{H}})$ and the $q$-th Schatten class with $J^q(\mathcal{H}, \tilde{\mathcal{H}})$, with the convention $J^\bullet(\mathcal{H}, \tilde{\mathcal{H}})$ := $J^\bullet(\mathcal{H}, \mathcal{H})$. These are linear spaces with $J^{q_1}(\mathcal{H}, \tilde{\mathcal{H}}) \subset J^{q_2}(\mathcal{H}, \tilde{\mathcal{H}})$ for all $q_2 \in [1, \infty]$, with $q_1 \leq q_2$, and one has inclusions of the type $J^q \circ \mathcal{L} \subset J^q, \mathcal{L} \circ J^q \subset J^q$ for all $q \in [1, \infty]$, and $J^{q_1} \circ J^{q_2} \subset J^{q_3}$ if $1/q_1 + 1/q_2 = 1/q_3$ with $q_3 \in [1, \infty)$.

For obvious reasons, $J^1$ is called the trace class, and moreover $J^2$ is called the Hilbert-Schmidt class.

**Example 2.6.** A bounded operator $K$ in $L^2(X, \mu)$-space is Hilbert-Schmidt, if (and only if) it is an integral operator with a square integrable integral kernel, that is, if
\[
Kf(x) = \int k(x, y)f(y)d\mu(y)
\]
for some $k \in L^2(X \times X, \mu \otimes \mu)$. This follows from evaluating
\[
\sum_n |\langle Ke_n, f_n \rangle|^2
\]
explicitly using Parseval’s identity.

Let us now turn towards the formulation of the spectral theorem:

**Definition 2.7.** A spectral resolution $P$ on $\mathcal{H}$ is a map $P : \mathbb{R} \to L(\mathcal{H})$ such that
- for every $\lambda \in \mathbb{R}$ one has $P(\lambda) = P(\lambda)^*$, $P(\lambda)^2 = P(\lambda)$ (that is, each $P(\lambda)$ is an orthogonal projection onto its image)
- $P$ is monotone in the sense that $\lambda_1 \leq \lambda_2$ implies $\text{Ran}(P(\lambda_1)) \subset \text{Ran}(P(\lambda_2))$
- $P$ is right-continuous in the strong topology\(^1\) of $L(\mathcal{H})$
- $\lim_{\lambda \to -\infty} P(\lambda) = 0$ and $\lim_{\lambda \to \infty} P(\lambda) = \text{id}_{\mathcal{H}}$, both in the strong sense.

It follows that for every $f \in \mathcal{H}$, the function
\[
\lambda \mapsto \langle P(\lambda)f, f \rangle = \|P(\lambda)f\|^2
\]
is right-continuous and increasing. Thus by the usual Stieltjes construction\(^2\) it induces a Borel measure on $\mathbb{R}$, which will be denoted by $\langle P(d\lambda)f, f \rangle$. This measure has the total mass
\[
\langle P(\mathbb{R})f, f \rangle = \|f\|^2.
\]

\(^1\) $T_n \to T$ strongly in $L(\mathcal{H})$ means that $T_n f \to Tf$ for all $f \in \mathcal{H}$; this is weaker than convergence in the operator norm topology, which means that $\|T_n - T\| \to 0$.

\(^2\) Given a right-continuous and increasing function $F : \mathbb{R} \to \mathbb{R}$ there exists precisely one measure $\mu_F$ on $\mathbb{R}$ such that for all $b > a$ one has $\mu_F((a, b]) = F(b) - F(a)$. 
Given such $P$ and a Borel function $\phi : \mathbb{R} \to \mathbb{C}$, the set

$$D_{P,\phi} := \left\{ f \in \mathcal{H} : \int_{\mathbb{R}} |\phi(\lambda)|^2 \langle P(d\lambda)f, f \rangle < \infty \right\}$$

is a dense linear subspace of $\mathcal{H}$ (cf. Satz 8.8 in [36]), and accordingly one can define a linear operator $\phi(P)$ with $\text{Dom}(\phi(P)) := D_{P,\phi}$ in $\mathcal{H}$ by mimicking the complex polarization identity,

$$\langle \phi(P)f_1, f_2 \rangle := (1/4) \int_{\mathbb{R}} \phi(\lambda) \langle P(d\lambda)(f_1 + f_2), f_1 + f_2 \rangle - (1/4) \int_{\mathbb{R}} \phi(\lambda) \langle P(d\lambda)(f_1 - f_2), f_1 - f_2 \rangle + (\sqrt{-1}/4) \int_{\mathbb{R}} \phi(\lambda) \langle P(d\lambda)(f_1 - \sqrt{-1}f_2), f_1 - \sqrt{-1}f_2 \rangle - (\sqrt{-1}/4) \int_{\mathbb{R}} \phi(\lambda) \langle P(d\lambda)(f_1 + \sqrt{-1}f_2), f_1 + \sqrt{-1}f_2 \rangle,$$

where $f_1, f_2 \in \text{Dom}(\phi(P))$. Every spectral measure induces the following “calculus”:

**Theorem 2.8.** Let $P$ be a spectral resolution on $\mathcal{H}$, and let $\phi : \mathbb{R} \to \mathbb{C}$ be a Borel function. Then:

(i) One has $\phi(P)^* = \overline{\phi(P)}$; in particular, $\phi(P)$ is self-adjoint, if and only if $\phi$ is real-valued.

(ii) One has $\|\phi(P)\| \leq \sup_{\mathbb{R}} |\phi| \in [0, \infty].$

(iii) If $\phi \geq -C$ for some constant $C \geq 0$, then one has $\phi(P) \geq -C$.

(iv) If $\phi' : \mathbb{R} \to \mathbb{C}$ is another Borel function, then

$$\phi(P) + \phi'(P) \subset (\phi + \phi')(P), \quad \text{Dom}(\phi(P) + \phi'(P)) = \text{Dom}((|\phi| + |\phi'|)(P))$$

and

$$\phi(P)\phi'(P) \subset (\phi\phi')(P), \quad \text{Dom}(\phi(P)\phi'(P)) = \text{Dom}((\phi\phi')(P)) \cap \text{Dom}(\phi');$$

in particular, if $\phi'$ is bounded, then

$$\phi(P) + \phi'(P) = (\phi + \phi')(P),$$

$$\phi(P)\phi'(P) = (\phi\phi')(P).$$

(v) For every $f \in \text{Dom}(\phi(P))$ one has

$$\|\phi(P)f\|^2 = \int_{\mathbb{R}} |\phi(\lambda)|^2 \langle P(d\lambda)f, f \rangle.$$
if \( \phi : \mathbb{R} \to \mathbb{R} \) is continuous, then \( \sigma(\phi(P_S)) = \overline{\phi(\sigma(S))} \)

- if \( \phi, \phi' : \mathbb{R} \to \mathbb{R} \) are Borel functions, then one has the transformation rule \( (\phi \circ \phi')(P_S) = \phi(P_{\phi'(P_S)}) \).

In view of these results, given a self-adjoint operator \( S \) in \( \mathcal{H} \), the calculus of Theorem 2.8 applied to \( P = P_S \) is usually referred to as the **spectral calculus of** \( S \). Likewise, given a Borel function \( \phi : \mathbb{R} \to \mathbb{C} \) one sets \( \phi(S) := \phi(P_S) \).

**Remark 2.10.** Let \( S \) be a self-adjoint operator in \( \mathcal{H} \).

1. The spectral calculus of \( S \) is compatible with all functions of \( S \) that can be defined "by hand". For example, for every \( z \in \mathbb{C} \setminus \mathbb{K} \) one has \( \phi(S) = (S - z)^{-1} \) with \( \phi(\lambda) := 1/(\lambda - z) \), or \( S^n = \phi(S) \) with \( \phi(\lambda) := \lambda^n \).

2. If \( S \) is a semibounded operator and \( z \in \mathbb{C} \) is such that \( \Re z < \min \sigma(S) \), then the spectral calculus (together with a well-known Laplace transformation formula for functions) shows that for every \( b > 0 \) one has the following formula for \( f_1, f_2 \in \mathcal{H} \):

\[
\langle (S - z)^{-b} f_1, f_2 \rangle = \frac{1}{\Gamma(b)} \int_0^\infty s^{b-1} \langle e^{zs} e^{-ss} f_1, f_2 \rangle \, ds.
\]

3. If \( S \geq -C \) for some constant \( C \geq 0 \) and if \( e^{-tS} \in \mathcal{F}^2(\mathcal{H}) \) for some \( t > 0 \), then \( S \) has a purely discrete spectrum (so the spectrum consists of countably many eigenvalues having a finite multiplicity) and if one enumerates the eigenvalues in an increasing way and counting multiplicity, \( (\lambda_n) \), then one has \( -C \leq \lambda_0 < \lambda_1 \rightarrow \infty \) if \( \mathcal{H} \) is infinite dimensional.

**Example 2.11.** Assume on a sigma-finite measurable space \( (X, \mu) \) we are given a measurable function \( \psi : X \to \mathbb{C} \). Then the associated maximally defined multiplication in \( L^2(X, \mu) \) is given by

\[
\text{Dom}(M_\psi) := \{ f \in L^2(X, \mu) : \psi f \in L^2(X, \mu) \}, \quad M_\psi f(x) := \psi(x) f(x).
\]

\( M_\psi \) is bounded from below, if and only if \( \psi \geq C \mu \text{-a.e.} \) for some \( C \in \mathbb{R} \) and bounded, if and only \( |\psi| \leq c \mu \text{-a.e.} \) for some \( c \geq 0 \). Moreover, \( M_\psi \) is always closed, and a point \( z \in \mathbb{C} \)
lies in the spectrum if and only there exists no $\epsilon > 0$ such that $|\psi - z| \geq \epsilon \mu$-a.e. and in the discrete spectrum if and only of

$$\mu\{|\psi - z|\} > 0.$$  

The operator $M_\psi$ is self-adjoint if only if $\psi(x) \in \mathbb{R}$ for $\mu$-a.e. $x \in X$. In the latter case, concerning the spectral calculus, one has $\phi(M_\psi) = M_{\phi \psi}$.

Using the spectral theorem one can show that every self-adjoint operator is unitarily equivalent to a self-adjoint multiplication operator on some finite measure space. Here, a linear operator $V$ between two Hilbert spaces is called \textit{unitary}, if it is bijective with $V^{-1} = V^*$ and two linear operators are called \textit{unitarily equivalent}, if there exits a unitary operator $V$ with $B = V^*AV$.

We now collect some basic facts about possibly unbounded sesquilinear forms on Hilbert spaces. Unless otherwise stated, all statements below can be found in section VI of T. Kato’s book [22].

Let again $\mathcal{H}$ be a complex separable Hilbert space. A \textit{sesquilinear form} $Q$ on $\mathcal{H}$ is understood to be a map $Q : \text{Dom}(Q) \times \text{Dom}(Q) \to \mathbb{C}$, where $\text{Dom}(Q) \subset \mathcal{H}$ is a linear subspace called the \textit{domain of definition of $Q$}, such that $Q$ is antilinear\footnote{We warn the reader, however, that in [22] the forms are assumed to be antilinear in their second slot; thus, if $Q(f_1, f_2)$ is a form in our sense, the theory from [22] has to be applied to the complex conjugate form $Q(f_1, f_2)^*$} in its first slot, and linear in its second slot. The \textit{quadratic form} induced by $Q$ is simply the map $Q : \text{Dom}(Q) \to \mathbb{C}$, $f \mapsto Q(f, f)$.

Let $Q$ and $Q'$ be sesquilinear forms on $\mathcal{H}$ in this section. $Q'$ is called an \textit{extension of $Q$}, symbolically $Q \subset Q'$, if $\text{Dom}(Q) \subset \text{Dom}(Q')$ and if both forms coincide on $\text{Dom}(Q)$.

$Q$ is called \textit{symmetric}, if $Q(f_1, f_2) = Q(f_2, f_1)^*$, and \textit{semibounded (from below)}, if its quadratic form is real-valued with and there exists a constant $C \geq 0$ such that

$$Q(f, f) \geq -C \|f\|^2 \quad \text{for all } f \in \text{Dom}(Q),$$

symbolically $Q \geq -C$. Again by complex polarization, every semibounded form is automatically symmetric (as the quadratic form is real-valued).

Following Kato, given a sequence $(f_n) \subset \text{Dom}(Q)$ and $f \in \mathcal{H}$ we write $f_n \to f$ as $n \to \infty$, if one has $f_n \to f$ in $\mathcal{H}$ and in addition

$$Q(f_n - f_m, f_n - f_m) \to 0 \quad \text{as } n, m \to \infty.$$
Then $Q$ is called \textit{closed}, if $f_n \xrightarrow{Q} f$ implies that $f \in \text{Dom}(Q)$. A semibounded $Q$ is closed, if and only if for some/every $C \geq 0$ with $Q \geq -C$ the scalar product on $\text{Dom}(Q)$ given by
\[
\langle f_1, f_2 \rangle_{Q,C} = (1 + C) \langle f_1, f_2 \rangle + Q(f_1, f_2)
\]
turns $\text{Dom}(Q)$ into a Hilbert space. Futhermore, for a semibounded $Q \geq -C$ its closedness is equivalent to the lower-semicontinuity of the function
\[
\mathcal{H} \rightarrow [-C, \infty], \quad f \mapsto \begin{cases} Q(f, f), & \text{if } f \in \text{Dom}(Q) \\ \infty & \text{else.} \end{cases}
\]
The form $Q$ is called \textit{closable}, if it has a closed extension. If $Q$ is semibounded and closable, then it has a smallest semibounded and closed extension $\overline{Q}$, which is (well-)defined as follows: $\text{Dom}(\overline{Q})$ is given by all $f \in \mathcal{H}$ that admit a sequence $(f_n) \subset \text{Dom}(Q)$ with $f_n \xrightarrow{Q} f$; then one has
\[
\overline{Q}(f, h) = \lim_n Q(f_n, h_n), \quad \text{where } f_n \xrightarrow{Q} f, \ h_n \xrightarrow{Q} h.
\]
If $Q$ is closed, then a linear subspace $D \subset \text{Dom}(Q)$ is called a \textit{core} of $Q$, if $\overline{Q}|_D = Q$.

Using the spectral calculus one defines:

**Definition 2.12.** Given a self-adjoint operator $S$ in $\mathcal{H}$, the (densely defined and symmetric) sesquilinear form $Q_S$ in $\mathcal{H}$ given by $\text{Dom}(Q_S) := \text{Dom}(\sqrt{|S|})$ and
\[
Q_S(f_1, f_2) := \left\langle \sqrt{|S|}f_1, \sqrt{|S|}f_2 \right\rangle
\]
is called the \textit{form associated with} $S$.

The following fundamental result links the world of densely defined, semibounded, closed forms with that of semibounded self-adjoint operators (cf. Theorem VIII.15 in [28] for this exact formulation):

**Theorem 2.13.** For every self-adjoint semibounded operator $S$ in $\mathcal{H}$, the form $Q_S$ is densely defined, closed and semibounded. Conversely, for every densely defined, closed and semibounded sesquilinear form $Q$ in $\mathcal{H}$, there exists precisely one self-adjoint semibounded operator $S_Q$ in $\mathcal{H}$ such that $Q = Q_{S_Q}$. The operator $S_Q$ will be called the operator \textit{associated} with $Q$.

The correspondence $S \mapsto Q_S$ has the following additional properties:

**Theorem 2.14.** Let $Q$ be densely defined, closed and semibounded. Then:
\begin{itemize}
  \item $S_Q$ is the uniquely determined self-adjoint and semibounded operator in $\mathcal{H}$ such that $\text{Dom}(S_Q) \subset \text{Dom}(Q)$ and
    \[
    \langle S_Q f_1, f_2 \rangle = Q(f_1, f_2) \quad \text{for all } f_1 \in \text{Dom}(S_Q), \ f_2 \in \text{Dom}(Q).
    \]
\end{itemize}
• \( \text{Dom}(S_Q) \) is a core of \( Q \); some \( f_1 \in \text{Dom}(Q) \) is in \( \text{Dom}(S_Q) \), if and only if there exists \( f_2 \in \mathcal{H} \) and a core \( D \) of \( Q \) with
\[
Q(f_1, f_3) = \langle f_2, f_3 \rangle \quad \text{for all } f_3 \in D,
\]
and then \( S_Q f_1 = f_2 \).

• One has
\[
\text{Dom}(Q) = \left\{ h \in \mathcal{H} : \lim_{t \to 0^+} \left\langle \frac{h - e^{-tS_Q}h}{t}, h \right\rangle < \infty \right\},
\]
\[
Q(h, h) = \lim_{t \to 0^+} \left\langle \frac{h - e^{-tS_Q}h}{t}, h \right\rangle.
\]

• One has the variational principle
\[
\min \sigma(S_Q) = \inf \{ Q(f, f) : f \in \text{Dom}(Q), \|f\| = 1 \}
\]
(8)
\[
= \inf \{ \langle S_Q f, f \rangle : f \in \text{Dom}(S_Q), \|f\| = 1 \}.
\]

**Notation 2.15.** If \( Q, Q' \) are symmetric, we write \( Q \geq Q' \), if and only if \( \text{Dom}(Q) \subset \text{Dom}(Q') \) and \( Q(f, f) \geq Q'(f, f) \) for all \( f \in \text{Dom}(Q) \).

The Friedrichs extension of a semibounded operator can be defined as follows:

**Example 2.16.** Let \( S \geq -C \) be a symmetric (in particular, a densely defined) and semi-bounded operator in \( \mathcal{H} \). Then the form \( (f_1, f_2) \mapsto \langle S f_1, f_2 \rangle \) with domain of definition \( \text{Dom}(S) \) is closable, and of course the closure \( \tilde{Q}_S \) of that form is densely defined and semibounded. The operator \( S_F \) associated with \( \tilde{Q}_S \) is called the Friedrichs realization of \( S \). The operator \( S_F \) can also be characterized as follows: \( S_F \) is the uniquely determined self-adjoint semibounded extension of \( S \) with domain of definition \( \subset \text{Dom}(\tilde{Q}_S) \). Let \( \mathcal{M}_C(S) \) denote the class of all self-adjoint extensions of \( S \) which are \( \geq -C \). Thus we have \( S_F \in \mathcal{M}_C(S) \), and in addition the following maximality property holds:
\[
T \in \mathcal{M}_C(S) \quad \Rightarrow \quad Q_T \leq \tilde{Q}_S.
\]

In particular, \( S_F \) has the smallest bottom of spectrum \( \min \sigma(S_F) \) among all operators in \( \mathcal{M}_C(S) \). This is Krein’s famous result on the characterization of semibounded extensions [1] [23].

Let us see how the Friedrichs construction can be used to define a self-adjoint realization of the Laplace-Beltrami operator \( -\Delta \) in \( L^2(U) \), where \( U \) is an arbitrary open subset of \( \mathbb{R}^m \): consider \( -\Delta \) as a linear operator in \( L^2(U) \), defined initially on \( C_c^\infty(U) \). We have seen above that \( -\Delta \) is symmetric; more precisely, for all \( f_1, f_2 \in C_c^\infty(U) \) one has
\[
\langle (-\Delta) f_1, f_2 \rangle_U = \int_U (\nabla f_1, \nabla f_2),
\]
so
\[
\langle (-\Delta) f_1, f_1 \rangle_U = \int_U |\nabla f_1|^2 \geq 0,
\]
and \( -\Delta \geq 0 \) in \( L^2(U) \). It follows from the previous example that \( -\Delta \) canonically induces a self-adjoint operator \( H_U \geq 0 \) in \( L^2(U) \), called the Dirichlet-Laplacian in \( U \). In terms of the Euclidean Sobolev spaces \( W^{k,p}(U) \) and \( W_0^{k,p}(U) \): one has

\[
\text{Dom}(H_U) = \{ f \in W_0^{1,2}(U) : \Delta f \in L^2(U) \}, \quad H_U f = -\Delta f,
\]

where \( \Delta f \) is understood in the sense of distributions. We will come to a detailed explanation of these facts in the more general context of Riemannian manifolds later on.

3. Basic facts on differential operators on Riemann manifolds

Let \( M \) be a manifold\(^4\) of dimension \( m \) and let \( E \to M, F \to M \) be vector bundles over \( M \) with rank \( \ell_0 \) and rank \( \ell_1 \), respectively. We understand all vector bundles over \( C \) (if not we can complexify; for example, a priori, the tangent bundle \( TM \to M \) is of course naturally given over \( R \)). We denote with \( \Gamma_\infty(M,E) \) the smooth sections of \( E \to N \), that is, the linear space (in fact \( C_\infty \) left module) of all smooth maps \( \psi : M \to E \) with \( \psi(x) \in E_x \) for all \( x \in N \). Likewise, smooth compactly supported sections will be denoted with \( \Gamma_{C_\infty}(M,E) \).

In case \( E = M \times C^l \to M \) is a trivial vector bundle, then each fiber \( E_x \) is given by \( \{x\} \times C^l \) and we can identify \( \Gamma_{C_\infty}(M,E) \) with \( C_\infty(M,C^l) \), where \( C_\infty(M) := C_\infty(M,C) \) for \( l = 1 \).

A map

\[
P : \Gamma_{C_\infty}(M,E) \to \Gamma_{C_\infty}(M,F)
\]

is called restrictable, if for all open \( U \subset X \) there exists a linear map

\[
P|_U : \Gamma_{C_\infty}(U,E) \to \Gamma_{C_\infty}(U,F)
\]

with \( P|_U \psi|_U = (P\psi)|_U \) for all \( \psi \in \Gamma_{C_\infty}(M,E) \).

**Definition 3.1.** A restrictable linear map

\[
P : \Gamma_{C_\infty}(M,E) \to \Gamma_{C_\infty}(M,F)
\]

is called a (smooth, linear) **partial differential operator of order \( \leq k \in \mathbb{N}_{\geq 0} \)**, if for any chart \( (x^1,\ldots,x^m),U \) of \( M \) which admits frames\(^5\) \( e_1,\ldots,e_{\ell_0} \in \Gamma_{C_\infty}(U,E), f_1,\ldots,f_{\ell_1} \in \Gamma_{C_\infty}(U,F) \), and any multi-index\(^6\) \( \alpha \in \mathbb{N}_k^m \), there are (necessarily uniquely determined) smooth functions

\[
P_\alpha : U \to \text{Mat}(C;\ell_0 \times \ell_1)
\]

such that for all \( (\phi^{(1)},\ldots,\phi^{(\ell_0)}) \in C_\infty(U,C^{\ell_0}) \) one has

\[
P|_U \sum_{i=1}^{\ell_0} \phi^{(i)} e_i = \sum_{j=1}^{\ell_1} \sum_{i=1}^{\ell_0} \sum_{\alpha \in \mathbb{N}_k^m} P_{\alpha ij} \frac{\partial^{|\alpha|} \phi^{(i)}}{\partial x^\alpha} f_j \quad \text{in } U.
\]

Any differential operator \( P \) satisfies \( \text{supp}(P\psi) \subset \text{supp}(\psi) \), that is, \( P \) is local.

---

\(^4\)We understand all our manifolds to be smooth and without boundary.

\(^5\)that is, 'frame' means that \( e_1(x),\ldots,e_{\ell_0}(x) \) is basis of \( E_x \) for all \( x \in U \)

\(^6\)\( \mathbb{N}_k^m \) denotes the set of multi-indices \( \alpha = (\alpha_1,\ldots,\alpha_m) \in (\mathbb{N}_{\geq 0})^m \) such that \( \alpha_1 + \cdots + \alpha_m \leq k \).
Definition 3.2. Let \( k \in \mathbb{N}_{\geq 0} \) and let
\[
P : \Gamma_{C^\infty}(M, E) \to \Gamma_{C^\infty}(M, F)
\]
be a differential operator of order \( \leq k \).

a) The (linear \( k \)-th order principal) symbol of \( P \) is the unique morphism
\[
symb^k_P : (T^*M)^\otimes k \otimes E \to F
\]
of vector bundles, where \( \otimes \) stands for the symmetric tensor product, such that for all \((x_1, \ldots, x_m, U), e_1, \ldots, e_{\ell_0}, f_1, \ldots, f_{\ell_1}\) as in Definition 3.1, and all real-valued \( \zeta^{(i)}_\alpha \in C^\infty(U) \) (where \( i \) runs through \( i = 1, \ldots, \ell_0 \) and \( \alpha \) runs through \( \alpha \in \mathbb{N}^m \) is such that \( \alpha_1 + \cdots + \alpha_m = k \)), one has
\[
symb^k_P \left( \sum_{\alpha \in \mathbb{N}^m : \alpha_1 + \cdots + \alpha_m = k} \sum_{i=1}^{\ell_0} \zeta^{(i)}_\alpha dx_\alpha \otimes e_i \right) = \sum_{\alpha \in \mathbb{N}^m : \alpha_1 + \cdots + \alpha_m = k} \sum_{i=1}^{\ell_0} \sum_{j=1}^{\ell_1} P_{\alpha ij} \zeta^{(i)}_\alpha f_j \text{ in } U.
\]

b) \( P \) is called elliptic, if for all \( x \in M, v \in T_x^*M \setminus \{0\} \), the linear map \( \text{symb}^k_{P,x}(v \otimes^k) : E_x \to F_x \) is invertible.

Remark 3.3. 1. Keep in mind that (at least locally) an operator \( P \) of order \( \leq k \) can also be considered as having order \( \leq l \) where \( l > k \) (set the higher order coefficients = 0), and then \( P \) can be elliptic as in the \( k \)-sense but not in the \( l \)-sense. Thus we always have to specify the order of \( P \) when we talk about ellipticity.

2. Ellipticity is a local question: it needs to be checked in some chart around \( x \) only.

A (smooth) metric \( h_E \) on \( E \to M \) is by definition a section \( h_E \in \Gamma_{C^\infty}(M, E^* \otimes E^*) \), such that \( h_E \) is fiberwise a scalar product. Then the datum \((E, h_E) \to M \) is referred to as a metric vector bundle. In other words, for every \( x \in M \) zuwe have a scalar product \( h_E(x) : E_x \times E_x \to \mathbb{C} \) and \( h_E(x) \) depends smoothly on \( x \). The trivial vector bundle \( M \times \mathbb{C}^l \to M \) is equipped with its canonical smooth metric which is induced by \((z, z') \mapsto \sum_{j=1}^l \overline{z}_j z_j \), where \( z, z' \in \mathbb{C}^l \).

Definition 3.4. A Riemannian metric on \( M \) is by definition a metric \( g \) on \( TM \to M \), and then the pair \((M, g) \) is called a (smooth) Riemannian manifold.

Proposition and definition 3.5. For any Riemannian metric \( g \) on \( TM \to M \), and then the pair \((M, g) \) is called a (smooth) Riemannian manifold.

Proof: Exercise.
The above measure \( \mu_g \) is called the \textit{Riemannian volume measure} on \((M, g)\). It is a Radon measure with a full topological support in the sense that \( \mu_g(U) > 0 \) for all open nonempty \( U \subset M \).

\textbf{Remark 3.6.} That two Borel sections are equal \( \mu_g \)-a.e. does not depend on a particular choice of \( g \). Thus, given \( k \in \mathbb{N}_{\geq 0}, q \in [1, \infty] \) we can define the local Sobolev space \( \Gamma^{k,q}_{\text{loc}}(M, E) \) to be the space of equivalence classes of Borel sections \( \psi \) of \( E \to M \) such that in every chart \( U \subset M \) in which \( E \to M \) admits a local frame \( e_j \) one has \( \psi^{(j)} \in \mathcal{W}^{k,q}_{\text{loc}}(U) \), if \( \psi = \sum_j \psi^{(j)} e_j \) in \( U \). In particular, we get the local \( L^q \)-spaces \( \Gamma^{0,q}_{\text{loc}}(M, E) := \Gamma^{0,q}_{\text{loc}}(M, E) \).

The fundamental lemma of distribution theory takes the following form:

\textbf{Lemma 3.7.} For all \( f_1, f_2 \in \Gamma^{1}_{\text{loc}}(M, E) \) one has \( f_1 = f_2 \) a.e., if and only if there exists a pair for all pairs of metrics \((g, h_E)\) with

\[ \int_M h_E(f_1, \psi) \mu_g = \int_M h_E(f_2, \psi) \mu_g \quad \text{for all } \psi \in \Gamma^{\infty}_{C^\infty}(M, E). \]

\textbf{Proof:} \( \Rightarrow \): Clear.

\( \Leftarrow \): Let \( U \subset M \) be a chart which admits an orthonormal frame \( e_1, \ldots, e_l \) for \((E, h_E) \to M \) (of course \( M \) be can covered with such \( U \)'s) and let \( \psi \) be an arbitrary smooth section with a compact support in \( U \). Then writing \( f_j = \sum_i f_j^i e_i, j = 1, 2, \) and \( \psi = \sum_i \psi^i e_i \) we have

\[ \int_U \sum_i \sqrt{\det(g)} \cdot f_1^i \psi^i dx = \int_M h_E(f_1, \psi) \mu_g = \int_M h_E(f_2, \psi) \mu_g = \int_U \sum_i \sqrt{\det(g)} \cdot f_2^i \psi^i dx, \]

so that by the Euclidean fundamental lemma of distribution theory we have

\[ \sqrt{\det(g)} \cdot f_1^i = \sqrt{\det(g(x))} \cdot f_2^i \]

in \( U \), for all \( i \), so \( f_1 = f_2 \) as \( \sqrt{\det(g)} > 0 \).

\textbf{Proposition and definition 3.8.} Assume that \( g \) is a Riemannian metric on \( M \) and that \((E, h_E) \to M \) and \((F, h_F) \to M \) are metric vector bundles. Then for any differential operator

\[ P : \Gamma^{C^\infty}(M, E) \to \Gamma^{C^\infty}(M, F) \]

of order \( \leq k \) there is a uniquely determined differential operator

\[ P^{g,h_E,h_F} : \Gamma^{C^\infty}(M, F) \to \Gamma^{C^\infty}(M, E) \]

of order \( \leq k \) which satisfies

\[ \int_M h_E(P^{g,h_E,h_F} \psi, \phi) \mu_g = \int_M h_F(\psi, P \phi) \mu_g \]
for all $\psi \in \Gamma_{C^\infty}(M,F)$, $\phi \in \Gamma_{C^\infty}(M,E)$ with either $\phi$ or $\psi$ compactly supported. The operator $P^{g,h_E,h_F}$ is called the formal adjoint of $P$ with respect to $(g,h_E,h_F)$. An explicit local formula for $P^{g,h_E,h_F}$ can be found in the proof.

Proof: Uniqueness follows from the fundamental lemma of distribution theory. As differential operators are local, it is sufficient to prove the local existence. To this end, in the situation of Definition 3.1, we assume that $e_i$ and $f_j$ are orthonormal with respect to $h_E$ and $h_F$, respectively. Then an integration by parts shows that

$$P^{g,h_E,h_F} \ell_1 \sum_{j=1}^{\ell_1} \psi^{(i)} f_j := \frac{1}{\sqrt{\det(g)}} \sum_{i=1}^{\ell_0} \sum_{j=1}^{\ell_1} \sum_{\alpha \in \mathbb{N}^n} \frac{(-1)^{|\alpha|}}{\alpha! \partial x^\alpha} \partial^{\alpha} \psi^{(j)} e_i \text{ in } U$$

does the job. ■

There is a way to define the action of differential operators on locally integrable functions:

**Proposition and definition 3.9.** Given $P$ as above, $f \in \Gamma_{L^1_{loc}}(M,E)$ and a subspace $A \subset \Gamma_{L^1_{loc}}(M,F)$ we write $Pf \in A$, if there exists $h \in A$, such that for all triples of metrics $(g,h_E,h_F)$ it holds that

$$\int_M h_E (P^{g,h_E,h_F} \psi, f) d\mu_g = \int_M h_F (\psi, h) d\mu_g \text{ for all } \psi \in \Gamma_{C^\infty}(M,F).$$

Then $h$ is uniquely determined and we set $Pf := h$. This property is equivalent to (11) being true for some triple $(g,h_E,h_F)$ of this kind (and is thus independent of the metrics).

Proof: Clearly $h$ is uniquely determined by the fundamental lemma of distribution theory. It remains to show that if (11) holds for some triple $(g,h_E,h_F)$ then it also holds for any other such triple. This is left as an exercise. ■

**Remark 3.10.** One says that given $f_n, f \in \Gamma_{L^1_{loc}}(M,E)$ that $f_n \to f$ in the sense of distributions, if for all $\psi \in \Gamma_{C^\infty}(M,E)$ and some pair of metrics $(g,h_E)$ one has

$$\int_M h_E (f_n - f, \psi) d\mu_g \to 0$$

as $n \to \infty$. Using that $\psi$ is compactly supported one easily checks that this automatically holds for all pairs of metrics $(g,h_E)$. Moreover, distributional limits are uniquely determined. Given $P$ as above, it is clear that $f_n \to f$ in the sense of distributions implies $Pf_n \to Pf$ in the sense of distributions, if $Pf \in \Gamma_{L^1_{loc}}(M,E)$ (as the action of $P$ is defined by duality).

**Lemma 3.11** (Local elliptic regularity). Assume

$$P : \Gamma_{C^\infty}(M,E) \to \Gamma_{C^\infty}(M,F)$$

is elliptic of order $\leq k$ and let $q \in [1,\infty)$. Then for all $f \in \Gamma_{L^1_{loc}}(M,E)$ with $Pf \in \Gamma_{L^1_{loc}}(M,F)$ one has $f \in \Gamma_{W^k,q_{loc}}(M,E)$ if $q > 1$ and $f \in \Gamma_{W^{k-1,1}_{loc}}(M,E)$ if $q = 1$. 

$\textbf{Proof:}$ The $q > 1$ is a classical fact by Nirenberg [26] and can be found in many textbooks such as [27]. The $q = 1$ case is nonstandard uses Besov spaces. Together with Guidetti and Pallara I have given a proof in [9].

Recall in this context that the local Sobolev embedding implies

$$\bigcap_{l \in \mathbb{N}} \Gamma_{W^{l,p}_{\text{loc}}}(M, E) \subset \Gamma_{C^\infty}(M, E) \quad \text{for all } p \in (1, \infty).$$

\textbf{Remark 3.12.} To give an idea of how ellipticity comes into play in such a result: Assume $M = \mathbb{R}^m$ and $E = F$ are the trivial line bundles $\mathbb{R}^m \times \mathbb{C} \to \mathbb{C}$ (so that $P$ acts on functions and has scalar coefficients). Assume further that $P = \sum_{|\alpha| \leq k} P_\alpha \partial^\alpha$ has constant coefficients. The global Sobolev spaces $W^{k,2}(\mathbb{R}^m)$, $k \in \mathbb{N}$, can be equivalently defined via Fourier transform

$$F : S'(\mathbb{R}^m) \to S'(\mathbb{R}^m)$$

mapping between Schwartz distributions:

$$W^{k,2}(\mathbb{R}^m) = \{ f \in L^2(\mathbb{R}^m) : \int |Ff(\zeta)|^2(1 + |\zeta|^2)^k d\zeta < \infty \}.$$

Then $P$ defines a continuous map

$$P : W^{k,2}(\mathbb{R}^m) \to L^2(\mathbb{R}^m),$$

which, using that $F^{-1}PF$ is nothing but multiplication by $\zeta \mapsto \sum_{|\alpha| \leq k} P_\alpha \zeta^\alpha$, can be shown to be bijective, if $\sum_{|\alpha| = k} P_\alpha \zeta^\alpha \neq 0$ for all $\zeta \in \mathbb{R}^m \setminus \{0\}$ (this requires some work). So $Pf = g \in L^2(\mathbb{R}^m)$ implies $f = P^{-1}g \in W^{k,2}(\mathbb{R}^m)$.

The local case of nonconstant coefficients can be deduced from this result by 'approximating' the coefficients and using some cut-off function machinery, and the local manifold case follows from this by a partition of unity argument.

From now on we fix once for all a connected Riemannian manifold $M = (M, g)$ with dimension $m$.

We are going to omit the dependence on $g$ in the notation whenever there is no danger of confusion. For example the Riemann volume measure is denoted by $\mu$. In addition, a metric vector bundle is simply depicted by $E \to M$, that is, the dependence on the fiber metrics will be omitted in the notation and the metric on $E \to M$ is simply denoted by $(\cdot, \cdot)$. For all $q \in [1, \infty]$ we get the Banach space $\Gamma_{L^q}(M, E)$ given by all equivalence classes of Borel sections $f$ of $E \to M$ such that

$$\|f\|_q < \infty,$$

where

$$\|f\|_q := \begin{cases} 
\inf \{ C \geq 0 : \|f\| \leq C \mu\text{-a.e.} \}, & \text{if } q < \infty \\
\left( \int_M |f|^q \mu \right)^{1/q}, & \text{else,}
\end{cases}$$

and

$$|f| := \sqrt{(f,f)}.$$
is the fiberwise norm. The space $\Gamma_{L^2}(M, E)$ becomes a Hilbert space via

$$\langle f_1, f_2 \rangle := \int_M (f_1, f_2) d\mu.$$  

With this convention, it makes sense to denote the formal adjoint of a differential operator

$$P : \Gamma_{C^\infty}(M, E) \to \Gamma_{C^\infty}(M, F)$$

acting between metric vector bundles simply by

$$P^\dagger : \Gamma_{C^\infty}(M, F) \to \Gamma_{C^\infty}(M, E).$$

We record:

**Lemma 3.13.** The space $\Gamma_{C^\infty_c}(M, E)$ is dense in $\Gamma_{L^q}(M, E)$ for all $q \in [1, \infty)$. In particular, $C^\infty_c(M)$ is dense in $L^q(M)$.

**Proof:**

**Step 1:** $A := \Gamma_{L^q}(M, E)$ is dense in $\Gamma_{L^q}(M, E)$.

**Proof of step 1:** Pick an exhaustion $K_n$ of $M$ with compact sets. Given $f \in \Gamma_{L^q}(M, E)$ set $f_n := 1_{K_n} f \in A$. Then we have

$$\lim_n \int |f_n - f|^q d\mu = \lim_n \int |(1_{K_n} - 1)|^q |f|^q d\mu = 0$$

by dominated convergence.

**Step 2:** $\Gamma_{C^\infty_c}(M, E)$ is dense in $A$.

**Proof of step 2:** Given $f \in A$ cover its support by finitely many charts $(U_n)$ for $M$ which admit an orthonormal frame. Pick a partition of unity $(\phi_n) \subset C^\infty_c(M)$ subordinate to $(U_n)$. Then $f_n := \phi_n f$ is compactly supported in $U_n$ and $L^q$ thereon. Given arbitrary $\epsilon > 0$, using Friedrichs mollifiers, for each $n$ we can pick $f_{n, \epsilon} \subset \Gamma_{C^\infty_c}(U_n, E)$ with

$$\|f_{n, \epsilon} - f_n\|_q < \epsilon / 2^{n+1}.$$  

Then $f_\epsilon := \sum_n f_{n, \epsilon} \in \Gamma_{C^\infty_c}(M, E)$ and

$$\|f_\epsilon - f\|_q = \left\| \sum_n f_{n, \epsilon} - \sum_n f_n \right\|_q \leq \sum_n \|f_{n, \epsilon} - f_n\|_q < \epsilon,$$

completing the proof. \qed

**4. The Friedrichs realization of the Laplace-Beltrami operator**

Since we have fixed $g$, the tangent bundle $TM \to M$ is by definition a metric bundle, using the isomorphism of vector bundles

$$\sharp : T^* M \to TM$$

induced by the fiberwise nondegeneracy of $g$, we get a metric $g^*$ on $T^* M \to M$ by setting

$$(\alpha, \beta) := (\sharp \alpha, \sharp \beta).$$

Let

$$d : C^\infty(M) \to \Omega^1_{C^\infty}(M) := \Gamma_{C^\infty}(M, T^* M)$$
denote the exterior differential. It is a first order differential operator (which does not
depend on \( g \)) given locally by \( df = \sum_i \partial_i f dx^i \).

**Definition 4.1.** The Laplace-Beltrami operator is the second order differential operator
given by

\[
\Delta := -d^\dagger d : C^\infty(M) \longrightarrow C^\infty(M).
\]

Locally one has

\[
d^\dagger \alpha = -\frac{1}{\sqrt{\det(g)}} \sum_k \partial_k \left( \sqrt{\det(g)} \sum_j g^{kj} \alpha_j \right)
\]

if \( \alpha = \sum_j \alpha_j dx^j \) and \( g^{kj} := (dx^k, dx^j) \). This formula shows

\[
\Delta = \frac{1}{\sqrt{\det(g)}} \sum_i \partial_i \left( \sqrt{\det(g)} \sum_j g^{ij} \partial_j \right),
\]

which can be worked out to give

\[
\Delta = \sum_{ij} g^{ij} \partial_i \partial_j + \text{lower order terms},
\]

in particular, in each chart \( U \), the symbol of \( \Delta \) (as an operator of order \( \leq 2 \)...) is given by \( g^{ij}(x)\zeta_i \zeta_j \), \( x \in U, \zeta U \). This implies that \( \Delta \) is elliptic (as \( g^{ij} \) is nondegenerate).

**Remark 4.2.** Local elliptic regularity shows: \( f \in L^2_{loc}(M), \Delta f \in L^2_{loc}(M) \) implies \( f \in W^{2,2}_{loc}(M) \), in particular, locally all weak partial derivatives of order \( \leq 2 \) of \( f \) are in \( L^2_{loc} \) (say in each chart). It is a more delicate question to investigate the following GLOBAL question: Does \( f \in L^2(M), \Delta f \in L^2(M) \) implies, say \( df \in \Omega^1_{L^2}(M) \). We will come back to this later (geodesic completeness!).

**Lemma 4.3.** a) One has

\[
(13) \quad d(f_1 f_2) = f_1 df_2 + f_2 df_1,
\]

\[
(14) \quad d^\dagger(f \alpha) = f d^\dagger \alpha - (df, \alpha),
\]

\[
(15) \quad \Delta(f_1 f_2) = f_1 \Delta f_2 + f_2 \Delta f_1 + 2 \Re(df_1, df_2),
\]

\[
(16) \quad \Delta(u \circ f) = (u'' \circ f) \cdot |df|^2 + (u' \circ f) \cdot \Delta f.
\]

**Proof:** Exercise. For example, one can use the above local formulae. ■

Consider now the densely defined, nonnegative, symmetric sesquilinear form \( Q' \) in \( L^2(M) \) given by

\[
\text{Dom}(Q') = C^\infty_c(M), \quad Q'(f_1, f_2) = \int (df_1, df_2) d\mu.
\]

It is induced by the symmetric nonnegative operator \(-\Delta \) (with \( \text{Dom}(-\Delta) = C^\infty_c(M) \)), as we have

\[
Q'(f_1, f_2) = \int -\Delta f_1 f_2 d\mu = \langle -\Delta f_1, f_2 \rangle.
\]
By Friedrichs’ theorem (cf. Example 2.16), it follows that $Q'$ is closable. Let us describe its closure. To this end, define the global Sobolev space

$$W^{1,2}(M) := \{ f \in L^2(M) : df \in \Omega^1_{L^2}(M) := \Gamma L^2(M, T^*M) \},$$

which is a Hilbert space with scalar product

$$\langle f_1, f_2 \rangle_{W^{1,2}} := \langle f_1, f_2 \rangle + \langle df_1, df_2 \rangle = \int \overline{f_1} f_2 d\mu + \int (df_1, df_2)d\mu.$$

Then we define

$$W^{1,2}_0(M) := \text{closure of } C^\infty_c(M) \text{ with respect to } \|\cdot\|_{W^{1,2}}.$$

**Remark 4.4.** If $M = \mathbb{R}^m$ (with its Euclidean metric) then one has $W^{1,2}_0(\mathbb{R}^m) = W^{1,2}(\mathbb{R}^m)$, while if $M$ is a bounded open subset $U$ of $\mathbb{R}^m$ then one has $W^{1,2}_0(U) \neq W^{1,2}(U)$. We will come to problems of this kind later on.

Now by Kato’s theory it follows that the closure $Q$ of $Q'$ is the closed nonnegative densely defined nonnegative symmetric sesquilinear form given by

$$\text{Dom}(Q) = W^{1,2}_0(M), \quad Q(f_1, f_2) = \int (df_1, df_2)d\mu.$$

By Kato’s theory (cf. Theorem 2.14) there exists a uniquely determined self-adjoint non-negative operator $H$ in $L^2(M)$ such that $\text{Dom}(H) \subset \text{Dom}(Q)$ and

$$\langle H f_1, f_2 \rangle = Q(f_1, f_2) \text{ for all } f_1 \in \text{Dom}(H), f_2 \in \text{Dom}(Q).$$

Moreover, some $f_1 \in \text{Dom}(Q)$ is in $\text{Dom}(H)$, if and only if there exists $f_2 \in L^2(M)$ with

$$Q(f_1, f_3) = \langle f_2, f_3 \rangle \text{ for all } f_3 \in C^\infty_c(M),$$

and then $H f_1 = f_2$. It follows now easily that

$$\text{Dom}(H) = \{ f \in W^{1,2}_0(M) : \Delta f \in L^2(M) \}, \quad H f = -\Delta f.$$

**5. Geodesic completeness and the essential self-adjointness of $-\Delta$**

This section deals with the following question: under which condition on the geometry of $M$, that is, on $g$, is $H$ the unique self-adjoint realization of $-\Delta$?

To this end, for all $x, y \in M$ we define $g(x, y)$ to be the infimum of all $\int_a^b |\dot{\gamma}(s)|ds$ such that $[a, b] \subset \mathbb{R}$ is a closed interval and $\gamma : [a, b] \to M$ is a piecewise smooth curve with $\gamma(a) = x, \gamma(b) = y$. Note that $\dot{\gamma}(s) \in T_{\gamma(s)}M$ and

$$\ell(\gamma) := \int_a^b |\dot{\gamma}(s)|ds$$

can be interpreted as the Riemannian length of the curve $\gamma$ (this notion is, as usual, justified by approximating with 'summing up the lengths of polygons approximating the curve' and taking the limit.
**Remark 5.1.** The main reason why we assume throughout that $M$ is connected is that otherwise the set whose infimum defines $\rho(x, y)$ could be empty, leading to $\rho(x, y) = \infty$.

The main properties of

$$
\rho : M \times M \to [0, \infty), \quad (x, y) \mapsto \rho(x, y)
$$

are collected in the following Theorem:

**Theorem 5.2.** a) $\rho$ is a distance on $M$ (the corresponding open balls will simply be denoted with $B(x, r) := \{y : \rho(x, y) < r\} \subset M$ in the sequel) and one has

$$
B(x, r) = \{y : \rho(x, y) \leq r\}.
$$

(17)

b) $\rho$ induces the original topology on $M$.

c) The following statements are equivalent:

i) $M$ is complete.

ii) All closed bounded subsets of $M$ are compact.

ii') All bounded subsets of $M$ are relatively compact.

iii) $M$ admits a sequence $(\chi_n) \subset C^\infty_c(M)$ of first order cut-off functions, that is, $(\chi_n)$ has the following properties:

(C1) $0 \leq \chi_n(x) \leq 1$ for all $n \in \mathbb{N}_{\geq 1}, x \in M$,

(C2) for all compact $K \subset M$, there is an $n_0(K) \in \mathbb{N}$ such that for all $n \geq n_0(K)$ one has $\chi_n |_K = 1$,

(C3) $\|d\chi_n\|_\infty \to 0$ as $n \to \infty$.

**Proof:** a) Clearly $\rho$ is nonnegative and $\rho(x, x) = 0$. To show the triangle inequality, fix $x, y, z \in M$ and pick a piecewise smooth path $\gamma_1$ from $x$ to $z$ and a piecewise smooth path $\gamma_2$ from $z$ to $y$. Let $\gamma$ be the path from $x$ to $y$ obtained as $\gamma = \gamma_2 \gamma_1$ in the obvious sense. Then one has

$$
\rho(x, y) \leq \ell(\gamma) = \ell(\gamma_2) + \ell(\gamma_1),
$$

so

$$
\rho(x, y) \leq \rho(x, z) + \rho(z, y)
$$

follows from minimizing in $\gamma$.

To see that $\rho$ is nondegenerate, we first prove:

**Claim:** for all $p \in M$ there exists a chart $p \in U \subset M$ and a constant $C$ such that

$$
C^{-1}|x - y| \leq \rho(x, y) \leq C|x - y|
$$

for all $x, y \in U$.

Proof of the claim: pick a chart $p \in W$ with coordinates $x^1, \ldots, x^m$ and pick a *Euclidean* ball $V \subset W$ of radius $r > 0$ around $p$ whose closure is included in $W$. For all $x \in \overline{V}$, $\zeta \in T_x M$ one has

$$
|\zeta|^2 := |\zeta|^2_g = \sum_{ij} g_{ij}(x)\zeta^i \zeta^j, \quad |\zeta|^2_e = \sum_j (\zeta^j)^2.
$$
Since \((g_{ij}(x))_{ij}\) is positive semidefinite and depends continuously on \(x\) we find \(C > 1\) such that for all \(x \in V\), \(\zeta \in T_xM\) one has
\[
C^{-2} \sum_j (\zeta^j)^2 \leq \sum_{ij} g_{ij}(x) \zeta^i \zeta^j \leq C^2 \sum_j (\zeta^j)^2,
\]
so
\[
C^{-1} |\zeta| \leq |\zeta_e| \leq C |\zeta|.
\]
For any piecewise smooth path \(\gamma\) which remains in \(V\) we get
\[
C^{-1} \ell_e(\gamma) \leq \ell(\gamma) \leq C \ell_e(\gamma).
\]
If \(x, y \in V\), then we get
\[
\varrho(x, y) \leq \ell(\gamma_{x, y}) \leq C |x - y|,
\]
where \(\gamma_{x, y}\) is the straight line from \(x\) to \(y\).

We are going to show that on \(U\) defined as the Euclidean ball in \(W\) around \(p\) of radius \(r/3\) one has the reverse inequality, so that \(U\) does the job.

Let \(x, y \in U\) and let \(\gamma\) be an arbitrary piecewise smooth curve in \(M\) from \(x\) to \(y\). If \(\gamma\) stays in \(V\) then \(\gamma\) stays in \(V\).

\[
\ell(\gamma) \geq \ell(\gamma_{x, y}) \geq C^{-1} |x - y|.
\]

Thus taking \(\inf_{\gamma}\) we get
\[
\varrho(x, y) \geq C^{-1} |x - y|,
\]
proving the claim.

In order to show that \(\varrho\) is nondegenerate, fix distinct \(p, x \in M\). Pick a chart \(U\) around \(p\) and \(C > 1\) as in the above claim. If \(x \in U\) then clearly \(\varrho(x, p) > 0\). If \(x \in M \setminus U\) pick \(r > 0\) small with \(B_e(p, r) \subset U\) (Euclidean ball). Then any curve \(\gamma\) from \(x\) to \(p\) must hit \(\partial B_e(p, r)\), and so \(\ell(\gamma) \geq C^{-1} r\) and by taking \(\inf_{\gamma}\) we arrive at \(\varrho(x, p) \geq C^{-1} r > 0\). This completes the proof that \(\varrho\) is a distance.

The proof of (17) is left as an exercise.

b) It is enough to show that for all \(p \in M\) there exists a chart \(U\) around \(p\) and \(R > 0\), \(C > 1\) such that for all \(r \in (0, R]\) one has
\[
B_e(p, C^{-1} r) \subset B(p, r) \subset B_e(p, C r) \subset U.
\]

\(^7\)Here we use that the length distance of \(x, y \in \mathbb{R}^m\) induced by the Euclidean Riemannian metric is precisely \(|x - y|\).
To this end pick $U$, $C$ as in the claim and $\epsilon > 0$ small with $B_e(p, \epsilon) \subset U$. Set $R := \epsilon/(2C)$ and let $0 < r \leq R$. If $x \in B_e(p, C^{-1}r)$ we have $x \in U$ and so $x \in B(p, r)$. If $x \notin U$ then and curve $\gamma$ from $x$ to $p$ hits a point $y \in U$ with $|y - p| = \epsilon/2$. Thus we obtain,

$$\ell(\gamma) \geq g(y, p) \geq C^{-1}|y - p| = \epsilon/(2C) \geq r,$$

and taking $\inf$, this shows $g(y, p) \geq r$ and so $x \notin B(p, r)$. This completes the proof.

c i) $\iff$ ii): Exercise (a proof which does not use exponential coordinates).

c ii) $\iff$ iii): this is trivial.

c i) $\iff$ iii): I sketch a proof: if $M = (M, g)$ is complete, then by a small generalization of Nash’s embedding theorem we can pick a smooth embedding $\iota : M \rightarrow \mathbb{R}^l$ such that $g$ is the pull-back of the Euclidean metric on $\mathbb{R}^l$ (thus an isometric embedding), where $l \geq m$ is large enough, and such that $\iota(M)$ is a closed subset of $\mathbb{R}^l$: note here that the original Nash embedding does not produce a closed image; to correct this, one constructs a new metric $\tilde{g}$ on $M$, embeds $(M, \tilde{g})$ into some $\mathbb{R}^l$ isometrically via some map $\Psi : M \rightarrow \mathbb{R}^l$ and constructs, using that closed balls are compact on $(M, g)$, a map $\phi : M \rightarrow \mathbb{R}$, such that

$$\iota := (\Psi, \psi) : M \rightarrow \mathbb{R}^l$$

is an isometric embedding of $(M, g)$, where $l := l' + 1$. A detailed explanation of the above construction of $\iota$ has been given by O. Mueller in [25].

From here the proof is straightforward: $\iota$ is proper, and therefore the composition

$$f : M \rightarrow \mathbb{R}, \quad f(x) := \log(1 + |\iota(x)|^2)$$

is a smooth proper function with $|df| \leq 1$, since

$$\tilde{f} : \mathbb{R}^l \rightarrow \mathbb{R}, \quad \tilde{f}(v) := \log(1 + |v|^2)$$

is a smooth proper function whose gradient is absolutely bounded by 1. Pick now a sequence $(\varphi_n) \subset C^\infty_c(\mathbb{R})$ of first order cut-off functions on the Euclidean space $\mathbb{R}$. (For example, let $\varphi : \mathbb{R} \rightarrow [0, 1]$ be smooth and compactly supported with $\varphi = 1$ near 0, and set $\varphi_n(r) := \varphi(r/n), r \in \mathbb{R}$.) Then $\chi_n(x) := \varphi_n(f(x))$ obviously has the desired properties, in view of the chain rule $d\chi_n(x) = \varphi'_n(f(x))df(x)$.

iii) $\iff$ ii'): Suppose that $M$ admits a sequence $(\chi_n) \subset C^\infty_c(M)$ of first order cut-off functions. Then given $\Theta \in M$, $r > 0$, we are going to show that there is a compact set $A_{\Theta, r} \subset M$ such that

$$g(x, \Theta) > r \quad \text{for all} \quad x \in M \setminus A_{\Theta, r},$$

which implies that any open geodesic ball is relatively compact. To see this, we define $A_{\Theta} := \{\Theta\}$, and a number $n_{\Theta, r} \in \mathbb{N}$ large enough such that $\chi_{n_{\Theta, r}} = 1$ on $A_{\Theta}$ and

$$\sup_{x \in M} |d\chi_{n_{\Theta, r}}(x)| \leq 1/(r + 1).$$

(19)

Now let $A_{\Theta, r} := \text{supp}(\chi_{n_{\Theta, r}})$, let $x \in M \setminus A_{\Theta, r}$, and let

$$\gamma : [a, b] \rightarrow M.$$
be a piecewise smooth curve with \( \gamma(a) = x, \gamma(b) = \mathcal{O} \). Then we have

\[
1 = \chi_{\mathcal{O},r}(\mathcal{O}) - \chi_{\mathcal{O},r}(x) = \chi_{\mathcal{O},r}(\gamma(b)) - \chi_{\mathcal{O},r}(\gamma(a)) = \int_a^b (d\chi_{\mathcal{O},r}(\gamma(s)), \gamma'(s)) \, ds,
\]

where we have used the chain rule. By using (19) and taking \( \inf_{\gamma} \cdots \), we arrive at

\[
g(x, \mathcal{O}) \geq r + 1 \quad \text{for all } x \in M \setminus A_{\mathcal{O},r},
\]

as claimed.

Now we can prove the following result, which has been first shown by Gaffney, 1954 (from my point of view: much ahead of his time!). We follows a proof given by Strichartz in 1983:

**Theorem 5.3.** Assume \( M \) is complete. Then the symmetric nonnegative operator \(-\Delta\) (defined on \( C_c^\infty(M)\)) is essentially self-adjoint in \( L^2(M) \). As a consequence, it has a unique self-adjoint extension which necessarily coincides with \( H \geq 0 \).

**Proof:** By the abstract functional analytic fact Theorem 2.5, it suffices to show that \( \text{Ker}((\Delta + 1)*) = \{0\} \). Let

\[
f \in \text{Ker}((\Delta + 1)*).
\]

Unpacking definitions one finds that this is equivalent to \( f \in L^2(M) \) and \(-\Delta f = -f\), in particular, \( f \) is smooth by local elliptic regularity. We pick a sequence \((\chi_n)\) of first order cut-off functions. Then by the product rule for \( d \) from Lemma 4.3 we have

\[
(d(\chi_n f), d(\chi_n f)) = (df, \chi_n f d\chi_n) + (df, \chi_n^2 df) + |f d\chi_n|^2 + (f \chi_n, \chi_n df),
\]

which, using

\[
(df, d(\chi_n^2 f)) = (df, \chi_n^2 df) + 2(df, f \chi_n d\chi_n),
\]

implies

\[
|d(\chi_n f)|^2 = (d(\chi_n f), d(\chi_n f)) = (df, d(\chi_n^2 f)) + |f d\chi_n|^2 - (df, f \chi_n d\chi_n) + (f d\chi_n, \chi_n df).
\]

This in turn implies (after adding the complex conjugate of the formula to itself)

\[
2|d(\chi_n f)|^2 = 2\Re(df, d(\chi_n^2 f)) + 2|f d\chi_n|^2.
\]

Integrating and then integrating by parts in the last equality, we get

\[
\int |d(\chi_n f)|^2 \, d\mu = \Re \int (\chi_n d^1 df, \chi_n f) \, d\mu + \int |f d\chi_n|^2 \, d\mu.
\]

Using \( d^1 df = -\Delta f = -f \) and

\[
\int |d(\chi_n f)|^2 \, d\mu \geq 0
\]

we see

\[
\int |\chi_n|^2 |f|^2 \, d\mu \leq \int |f d\chi_n|^2 \, d\mu,
\]
which implies \( \int |f|^2 d\mu = 0 \) and thus \( f = 0 \) by dominated convergence, using the properties of \((\chi_n)\).

Some remarks are in order:

**Remark 5.4.** 1. There are some interesting (though not many) incomplete Riemannian manifolds such that \(-\Delta\) is essentially self-adjoint.
2. We are going to prove in the exercises that even the Schrödinger operator \(-\Delta + V\) in \(L^2(M)\) is essentially self-adjoint, if \(M\) is complete and \(V : M \to \mathbb{R}\) is smooth and bounded from below. Note \(V\) has to be real-valued to get a symmetric operator.
3. The ultimate essential self-adjointness result on Riemann manifolds is the following one: assume \(M\) is complete and \(V \in L^2_{\text{loc}}(M)\) has a little more local regularity ('local Kato class' of \(M\)) such that \(-\Delta + V\) is bounded from below. Then \(-\Delta + V\) is essentially self-adjoint. This result can by applied to get that the Hamilton operator corresponding to a molecule is essentially self-adjoint (so there is no ambiguity concerning the quantum mechanics of matter).
4. Similar essential self-adjointness results hold for operators of the form \(\nabla^\dagger \nabla + V\) on metric vector bundles \(E \to M\), where \(\nabla\) is a metric connection on \(E \to M\) and \(V\) is a pointwise self-adjoint \(L^2_{\text{loc}}\)-section of \(\text{End}(E) \to M\) (Guneysu/Post; Braverman/Milatovic/Shubin; Lesch). These results are needed at least to deal with molecules in magnetic fields.

6. Some regularity results

**Lemma 6.1.** Assume \(f_1 \in W^{1,2}_0(M)\), \(f_2 \in W^{1,2}(M)\), \(\Delta f_2 \in L^2(M)\). Then one has the following integration by parts formula,

\[
\int f_1 \Delta f_2 d\mu = - \int (df_1, df_2) d\mu.
\]

**Proof:** If \(f_1\) is smooth and compactly supported, then the identity follows immediately from the definition of weak (= distributional) derivatives. It carries over to general \(f_1\)'s by a trivial density argument.

Note that every \(f_2 \in \text{Dom}(H)\) satisfies the above assumption. Often, this is used in the form \(f_2 = e^{-tH}h\) for some \(h \in L^2(M)\), \(t > 0\), as we know that for all \(t > 0\),

\[
\text{Ran}(e^{-tH}) \subset \bigcap_{n \in \mathbb{N}} \text{Dom}(H^n)
\]

by the spectral calculus.

**Lemma 6.2.** Given a sequence of smooth functions \(\psi_k : \mathbb{R} \to \mathbb{R}, k \in \mathbb{N}, \) with

\[
\psi_k(0) = 0, \quad \sup_{k \in \mathbb{N}} \sup_{t \in \mathbb{R}} |\psi'_k(t)| < \infty,
\]

and a pair of functions \(\psi : \mathbb{R} \to \mathbb{R}, \varphi : \mathbb{R} \to \mathbb{R}\) with

\[
\psi_k \to \psi, \psi'_k \to \varphi
\]
pointwise as $k \to \infty$.

a) For every real-valued $f \in W^{1,2}_0(M)$ one has $\psi \circ f \in W^{1,2}_0(M)$ and
\[ d(\psi \circ f) = (\varphi \circ f)df. \]

b) For every real-valued $f \in W^{1,2}(M)$ one has $\psi \circ f \in W^{1,2}(M)$ and
\[ d(\psi \circ f) = (\varphi \circ f)df. \]

If in addition $\varphi$ is continuous away from an at most countable set, then $f_n, f \in W^{1,2}(M)$, $f_n \to f$ in $W^{1,2}(M)$ implies $\psi \circ f_n \to \psi \circ f$ in $W^{1,2}(M)$, as $n \to \infty$.

c) For every real-valued $f \in W^{1,2}_{\text{loc}}(M)$ one has $\psi \circ f \in W^{1,2}_{\text{loc}}(M)$ and
\[ d(\psi \circ f) = (\varphi \circ f)df. \]

Proof: a) Lemma 5.2 in [7].
b) Theorem 5.7 in [7].
c) This follows from applying b) with $M$ replaced by a relatively compact chart of $M$. ■

Denote with $a_+ := \max(0, a) \in [0, \infty)$ the positive part of $a \in \mathbb{R}$ and with $a_- := a_+ - a \in [0, \infty)$ its negative part.

Example 6.3. Given $c \geq 0$ set $\psi(t) := (t - c)_+$,
\[ \phi(t) := \begin{cases} 0, & t \leq c, \\ 1, & t > c. \end{cases} \]

Then picking $\psi_1 : \mathbb{R} \to \mathbb{R}$ smooth with
\[ \psi_1(t) := \begin{cases} 0, & t - 1 \leq c, \\ 1, & t > c + 2, \end{cases} \]

the sequence $\psi_k(t) := k^{-1}\psi_1(kt)$ satisfies the assumptions of the previous lemma, yielding that for all real-valued $f \in W^{1,2}_0(M)$ (resp. $f \in W^{1,2}(M)$) one has $(f - c)_+ \in W^{1,2}_0(M)$ (resp. $(f - c)_+ \in W^{1,2}(M)$) and the formula
\[ d(f - c)_+ = \begin{cases} df, & \text{if } f > c \\ 0, & \text{else} \end{cases} \]

Moreover, $f_n \to f$ in $W^{1,2}(M)$ implies $(f_n - c)_+ \to (f - c)_+$ in $W^{1,2}(M)$.

Let $N \subseteq M$ be an arbitrary subset. Then a function $f : N \to \mathbb{R}$ on $M$ is called Lipschitz, if there exists a constant $C$ such that for all $x, y \in N$ one has
\[ |f(x) - f(y)| \leq C \varrho(x, y). \]

(20)

Lipschitz functions are continuous, restrictions of Lipschitz functions are again Lipschitz, and for a fixed $x_0 \in M$, the function
\[ M \ni x \mapsto \varrho(x, x_0) \]
is Lipschitz. Note also that if $U \subset M$ is open, then with an obvious notation one has
\[ g_U(x, y) \geq g(x, y) \quad \text{for all } x, y \in U, \]
so a Lipschitz function $f : U \to \mathbb{R}$ in the above sense is also a Lipschitz function with respect to the Riemannian manifold $(U, g_U)$.

**Remark 6.4.** The following assertions can be deduced in an elementary way and hold on every metric space: If $f, g$ are Lipschitz, then so is $f + g, \min(f, g), \max(f, g)$; the product $fg$ is Lipschitz, if in addition $f$ is bounded on the support of $g$.

A function $f : M \to \mathbb{R}$ is called **locally Lipschitz**, if for each compact $K \subset M$ there exists $C = C_K$ with (20) for all $x, y \in K$. The composition of a Lipschitz function with a Lipschitz function on $\mathbb{R}$ is Lipschitz; the composition of a locally Lipschitz function with a locally Lipschitz function on $\mathbb{R}$ is locally Lipschitz.

**Lemma 6.5.** a) If $f : M \to \mathbb{R}$ is a Lipschitz function, then $df$ exists as an element of $\Omega^1_{L^\infty}(M)$ and one has $\|df\|_\infty \leq C'$, where $C'$ is the smallest $C$ with (20). If $f : M \to \mathbb{R}$ is locally Lipschitz, then $df$ exists as an element of $\Omega^1_{L^\infty,\text{loc}}(M)$.

b) A $C^1$-function $f : M \to \mathbb{R}$ with $\|df\|_\infty$ is Lipschitz. In particular, $C^1$-functions are locally Lipschitz.

**Proof:** a) This follows from applying the corresponding Euclidean result (Rademacher’s theorem) in ‘nice charts’ like those appearing in the proof of Theorem 5.2, namely, by scaling the charts if necessary, one can find for each $p \in M$ a chart $U$ with $p \in U$ and
\[ (1/2)\delta_{ij} \leq g_{ij}(x) \leq 2\delta_{ij} \]
for all $x \in U$, as bilinear forms (the point is that the constant in this quasi-isometry, $C = 2$, is uniform in each chart). Rademacher’s theorem can either be deduced with methods of Analysis 1, by reducing to the $m = 1$ case with a covering argument (‘Vitali covering’), using that functions on an interval having a bounded variation are almost everywhere differentiable by Lebesgue’s theorem, or by using a Sobolev embedding theorem (cf. Theorem 3.1, resp. section 4.2 in [13]).

The local statement can be deduced as follows from the above: Assume $N \subset M$ is open and relatively compact and let $f : M \to \mathbb{R}$ be locally Lipschitz. Pick $\phi \in C_c^\infty(M)$ with $\phi = 1$ on $N$. Then $\phi f$ is globally Lipschitz and so $d(\phi f) \in \Omega^1_{L^\infty}(M)$. Since $f = \phi f$ on $N$ we thus get $df \in \Omega^1_{L^\infty}(N)$.

b) This follows applying the mean value theorem for differentiation in nice charts.

**Lemma 6.6.** One has $W^{1,2}_c(M) \subset W^{1,2}_0(M)$. In particular, for every compactly supported Lipschitz function $f : M \to \mathbb{R}$ one has $f \in W^{1,2}_0(M)$.

**Proof:** Let $h \in W^{1,2}_c(M)$. Covering the support of $h$ with finitely many nice charts, we can assume that $M$ is an open subset of the Euclidean $\mathbb{R}^m$. In this case the assertion follows from Friedrichs mollifiers.

For the second statement, note that $f$ is $L^2$ (continuous and compactly supported) and $df$ is $L^2$ (bounded by the previous Lemma and compactly supported), so $f \in W^{1,2}_c(M)$.
Lemma 6.7. One has the product rule \( d(f_1f_2) = f_1df_2 + f_2df_1 \) if \( f_1, f_2 : M \to \mathbb{R} \) are locally Lipschitz. If \( \psi : \mathbb{R} \to \mathbb{R} \) is \( C^1 \) and \( f : M \to \mathbb{R} \) is locally Lipschitz, then \( \psi \circ f \) is locally Lipschitz with the chain rule \( d(\psi \circ f) = (\psi' \circ f) df \).

Proof: In view of the cut-off function argument from the proof of Lemma 6.5 b) we can assume that \( f_1 \) and \( f_2 \) are compactly supported, so that \( f_1, f_2 \in W^{1,2}_0(M) \) by the previous lemma. Pick a sequence \((\phi_n) \subset C_0^\infty(M) \) with \( \phi_n \to f_1 \) in \( W^{1,2}(M) \). Since \( f_1 \) is bounded, we can assume for the proof that \( (\phi_n) \) is uniformly bounded in \( W^{1,2}(M) \) (and so \( \phi_n \) is uniformly bounded in \( L^2(M) \) and \( (d\phi_n) \) is uniformly bounded in \( \Omega_{L2}^1(M) \): indeed, let \( C := \|f_1\|_\infty \) and pick \( \psi : \mathbb{R} \to \mathbb{R} \) with \( \psi(t) = t \) for all \( t \) with \( |t| \leq C \) and with \( \psi' \) bounded. Then by Lemma 6.2 we have \( \psi \circ \phi_n \to f_1 \), and \( (\psi \circ \phi_n) \) is uniformly bounded in \( W^{1,2}(M) \).

Thus we can pick sequences \( \phi_n, \theta_n \in C_0^\infty(M) \) which are both uniformly bounded in \( W^{1,2}(M) \) and \( \phi_n \to f_1 \) and \( \theta_n \to f_2 \) in \( W^{1,2}(M) \). Then one easily checks that (a standard \( \epsilon/2 \) type argument which uses that \( f_2 \) are bounded) that \( \phi_n \theta_n \to f_1 f_2 \) in \( L^2(M) \), which using Cauchy-Schwarz implies \( \phi_n \theta_n \to f_1 f_2 \) in the sense of distributions, and so \( d(\phi_n \theta_n) \to d(f_1 f_2) \) in the sense of distributions (cf. Remark 3.10).

Similarly, as \( df_1 df_2 \in L^\infty(M) \) by Lemma 6.5 a) (since \( f_j \) are compactly supported), one can check that \( \phi_n d\theta_n \to f_1 df_2 \) and \( \theta_n d\phi_n \to f_2 df_1 \) in \( \Omega_{L2}^1(M) \), and so

\[
d(\phi_n \theta_n) = \theta_n(d\phi_n) + \phi_n d\theta_n \to f_2 df_1 + f_1 df_2
\]

in \( \Omega_{L2}^1(M) \), which using Cauchy-Schwarz implies \( \phi_n \theta_n \to f_2 df_1 + f_1 df_2 \) in the sense of distributions. This completes proof.

It remains to prove the asserted chain rule: since \( C^1 \)-functions are locally Lipschitz, it suffices to prove the formula in each open relatively compact subset of \( M \). In particular, we can assume that \( f \) is compactly supported. Furthermore, we can assume that \( \psi(0) = 0 \) (if not: consider \( \hat{\psi} := \psi - \psi(0) \)), and as \( f \) is bounded also that \( \psi \) is compactly supported (in particular, \( \psi' \) is bounded). Under these assumptions, the chain rule follows trivially from Lemma 6.2.

Lemma 6.8. Assume \( f_1 : M \to \mathbb{R} \) is bounded and Lipschitz and \( f_2 \in W^{1,2}_0(M) \). Then \( f_1 f_2 \in W^{1,2}_0(M) \) and one has the product rule \( d(f_1 f_2) = f_1 df_2 + f_2 df_1 \).

Proof: Assume first \( f_2 \in C_0^\infty(M) \). Then we have \( f_1 f_2 \) is compactly supported and Lipschitz, thus in \( W^{1,2}_0(M) \), and the product rule holds by the previous lemma. If \( f_2 \in W^{1,2}_0(M) \), then \( f_1 f_2 \in L^2(M) \) and \( f_1 df_2 + f_2 df_1 \in \Omega_{L2}^1(M) \), as \( f_1 \) and \( df_1 \) are bounded. This implies \( f_2 \in W^{1,2}(M) \). Pick a sequence \( \phi_n \in C_0^\infty(M) \) such that \( \phi_n \to f_2 \) in \( W^{1,2}(M) \). Then we have \( f_1 \phi_n \to W^{1,2}_0(M) \) and \( f_1 \phi_n \to f_1 f_2 \) in \( L^2(M) \), as \( f_1 \) is bounded. Applying the product rule to \( f_1 \phi_n \) and using that \( f, df \) are bounded, one easily finds that also

\[
d(f_1 \phi_n) \to f_1 df_2 + f_2 df_1.
\]

in \( \Omega_{L2}^1(M) \). It follows from these two convergences that \( f_1 \phi_n \to f_1 f_2 \) in \( W^{1,2}(M) \) and so \( f_1 f_2 \in W^{1,2}_0(M) \), as the latter is a closed subspace of \( W^{1,2}(M) \). Finally, \( d(f_1 \phi_n) \to f_1 df_2 + f_2 df_1 \) in \( \Omega_{L2}^1(M) \) implies the corresponding convergence in the sense of distributions (by Cauchy-Schwarz), \( f_1 \phi_n \to f_1 f_2 \) in \( L^2(M) \) implies the corresponding convergence in
the sense of distributions, and so by Remark 3.10 also \( d(f_1 \phi_n) \to d(f_1 f_2) \), which also establishes the product formula for \( f_1 f_2 \).

**Lemma 6.9.** Assume \( f_1 \in W^{1,2}_{\text{loc}}(M) \) and that \( f_2 : M \to \mathbb{R} \) is compactly supported and Lipschitz. Then one has \( f_1 f_2 \in W^{1,2}_0(M) \) and the product rule applies.

**Proof:** Multiplying \( f_1 \) with a smooth compactly supported function which is \( = 1 \) on the support of \( f_2 \) we can assume that \( f_1 \in W^{1,2,0}_{\text{loc}}(M) \) (cf. Lemma 6.6), in which case the statement follows from the previous lemma. ■

### 7. Basic properties of the heat kernel

The “heat semigroup”

\[
(e^{-tH})_{t \geq 0} \subset \mathcal{L}(L^2(M))
\]

is defined by the spectral calculus. It is a strongly continuous and self-adjoint semigroup with

\[
\|e^{-tH}\|_{2,2} \leq 1,
\]

where \( \|\cdot\|_{q_1,q_2} \) denotes the operator for linear operators from \( L^{q_1}(M) \) to \( L^{q_2}(M) \). Moreover, for every \( f \in L^2(M) \) the path

\[
[0, \infty) \ni t \mapsto e^{-tH} f \in L^2(M)
\]

is the uniquely determined continuous path

\[
[0, \infty) \to L^2(M)
\]

which is \( C^1 \) in \( (0, \infty) \) (in the norm topology) with values in \( \text{Dom}(H) \) thereon, and which satisfies the abstract “heat equation”

\[
(d/dt)e^{-tH} f = -He^{-tH} f, \quad t > 0,
\]

subject to the initial condition \( e^{-tH} f|_{t=0} = f \). All of the above facts follow from abstract functional analytic results and only rely on the fact that \( H \) is self-adjoint and nonnegative.

The aim of this section is to show that \( e^{-tH} \) is given by an integral kernel

\[
e^{-tH} f(x) = \int p(t, x, y) f(y) d\mu(y),
\]

such that for fixed \( x \), \( (t, y) \mapsto p(t, x, y) \) solves the heat equation

\[
\partial_t u(t, y) = \Delta_y u(t, y)
\]

with initial condition \( u(0, x) = \delta_x \).

**Theorem 7.1.** a) There is a unique smooth map

\[
(0, \infty) \times M \times M \ni (t, x, y) \mapsto p(t, x, y) \in [0, \infty),
\]

the heat kernel of \( H \), such that for all \( t > 0 \), \( f \in L^2(M) \), and \( \mu \)-a.e. \( x \in M \) one has

\[
e^{-tH} f(x) = \int p(t, x, y) f(y) d\mu(y).
\]

(21)
b) For all $s, t > 0$, $x, y \in M$ one has

\begin{align}
(22) & \quad \int p(t, x, y)^2 d\mu(y) < \infty, \\
(23) & \quad p(t, y, x) = p(t, x, y), \\
(24) & \quad p(t + s, x, y) = \int p(t, x, z)p(s, z, y) d\mu(z), \\
(25) & \quad \int p(t, x, z) d\mu(z) \leq 1.
\end{align}

c) For any $f \in L^2(M)$, the function

$$(0, \infty) \times M \ni (t, x) \mapsto P_t f(x) := \int p(t, x, y)f(y) d\mu(y) \in \mathbb{C}$$

is smooth and one has

$$\frac{\partial}{\partial t} P_t f(x) = \Delta_x P_t f(x) \quad \text{for all } (t, x) \in (0, \infty) \times M.$$ 

d) For all fixed $x \in M$, the function $(t, y) \mapsto p(t, x, y)$ solves the heat equation

$$\partial_t u(t, y) = \Delta_y u(t, y)$$

in $(0, \infty) \times M$, with initial condition $u(0, x) = \delta_x$, in the sense that

$$\lim_{t \to 0^+} \int p(t, x, y)\phi(y) d\mu(y) = \phi(x) \quad \text{for all } \phi \in C_c^\infty(M).$$

Proof: Before we come to the proof of the actual statements of Theorem 7.1, let us first establish some auxiliary results.

Step 1: For fixed $t > 0$, there exists a smooth version of $x \mapsto e^{-tH}f(x)$ (which from now on will always be taken).

Proof: To see this, note that for any $n \in \mathbb{N}_{\geq 1}$ one has

$$\text{Dom}(H^n) \subset W^{2n, 2}_{\text{loc}}(M),$$

by local elliptic regularity. By the spectral calculus and the local Sobolev embedding (12), this implies

$$\text{Ran}(e^{-tH}) \subset \bigcap_{n \in \mathbb{N}_{\geq 1}} \text{Dom}(H^n) \subset C^\infty(M) \text{ for any } t > 0.$$ 

Step 2: For any $t > 0$, $U \subset M$ open and relatively compact, the map

\begin{equation}
(26) \quad e^{-tH} : L^2(M) \to C_b(U)
\end{equation}

is a bounded linear operator between Banach spaces, where the space of bounded continuous functions $C_b(U)$ is equipped with its usual uniform norm.

Proof: A priori, this map is algebraically well-defined by step 1. The asserted boundedness follows from the closed graph theorem, noting that the $L^2(M)$-convergence of a sequence
implies the existence of a subsequence which converges $\mu$-a.e.

**Step 3:** For fixed $s > 0$, the map

$$L^2(M) \times M \ni (f, x) \mapsto e^{-sH} f(x) \in \mathbb{C}$$

is jointly continuous.

Proof: Let $U \subset M$ be an arbitrary open and relatively compact subset. Given a sequence

$$((f_n, x_n))_{n \in \mathbb{N} \geq 0} \subset L^2(M) \times U$$

which converges to

$$(f, x) \in L^2(M) \times U,$$

we have

$$
\left| e^{-sH} f_n(x_n) - e^{-sH} f(x) \right|
\leq \left| e^{-sH} [f_n - f](x_n) \right| + \left| e^{-sH} f(x) - e^{-sH} f(x_n) \right|
\leq \left\| e^{-sH} \right\|_{L^2(M), C_0(U)} \left\| f_n - f \right\|_2 + \left| e^{-s\tilde{P}} f(x) - e^{-s\tilde{P}} f(x_n) \right|
\to 0, \text{ as } n \to \infty,
$$

by step 2 and step 1.

**Step 4:** For fixed $\epsilon > 0$ and $f \in L^2(M)$, the map

$$\{ \Re > \epsilon \} \times M \ni (z, x) \mapsto e^{-zH} f(x)$$

is jointly continuous.

Proof: Indeed, this map is equal to the composition of the maps

$$\{ \Re > \epsilon \} \times M \xrightarrow{(z, x) \mapsto e^{-(z-\epsilon)H} f(x)} L^2(M) \times X \xrightarrow{(f, x) \mapsto e^{-\epsilon H} f(x)} \mathbb{C},$$

where the second map is continuous by Step 3. The first map is continuous, since the map

$$\{ \Re > 0 \} \ni z \mapsto e^{-zH} f \in L^2(M)$$

is holomorphic. Note that, a priori, (27) is a weakly holomorphic semigroup by the spectral calculus, which is then indeed (norm-) holomorphic by the weak-to-strong differentiability theorem.

**Step 5:** For any $f \in L^2(M)$, there exists a jointly smooth version $(t, x) \mapsto P_t f(x)$ of $(t, x) \mapsto e^{-tH} f(x)$, which satisfies

$$\frac{\partial}{\partial t} P_t f(x) = \Delta_x P_t f(x).$$

Proof: By Step 4, for arbitrary $f \in L^2(M)$, the map

$$\{ \Re > 0 \} \times M \ni (z, x) \mapsto e^{-zH} f(x) \in \mathbb{C}$$

is jointly continuous. It then follows from the holomorphy of (27) that for any open ball $B$ in the open right complex plane which has a nonempty intersection with $(0, \infty)$, for any
Let us now come to the actual proof of Theorem 7.1.

a) First of all, it is clear that any such heat kernel is uniquely determined (by the fundamental lemma of distribution theory). To see its existence, we start by remarking that for every \( x \in M, t > 0 \), the complex linear functional given by 
\[
L^2(M) \ni f \mapsto \langle p_{t,x}, f \rangle \in \mathbb{C}
\]
is bounded by Step 2. Thus by Riesz-Fischer’s representation theorem, there exists a unique function \( p_{t,x} \in L^2(M) \) such that for all \( f \in L^2(M) \) one has 
\[
(29) \quad P_t f(x) = \langle p_{t,x}, f \rangle.
\]

Claim 1: One has 
\[
(30) \quad p(t, x, y) := \langle p_{t/2,x}, p_{t/2,y} \rangle
\]
has the desired properties. Firstly, the smoothness of \( (t, x, y) \mapsto p(t, x, y) \) follows immediately from the norm smoothness of \( (t, x) \mapsto p_{t,x} \) and the smoothness of the Hilbertian pairing \( (f, g) \mapsto \langle f, g \rangle \).

Claim 1: One has 
\[
P_{t+s} f(x) = \int \langle p_{t,z}, p_{s,x} \rangle f(z) d\mu(z)
\]
Proof of Claim 1:

\[
P_{t+s}f(x) = PsP_tf(x)
= \langle ps, Pt f \rangle
= \int P_t ps(x) f(z) d\mu(z)
= \int \langle ps, x, Pt f \rangle f(z) d\mu(z).
\]

Claim 2: For all \(t > 0\), the scalar product \(\langle p_{s'}, z, p_{t-s'}, x \rangle\) does not depend on \(s' \in (0, t)\).

Proof of Claim 2: Let \(r \in (0, s')\). Then using Claim 1 with \(f = p_{r,x}\),

\[
\langle p_{s'}, z, p_{t-s'}, x \rangle = \langle ps', p_{t-s'}, p_{r,x} \rangle
= \int p_{r,x} \langle ps'-r, z, p_{t-s'}, y \rangle d\mu(z)
= \langle p_{r,y}, p_{r,x} \rangle = \langle p_{r,x}, p_{r-y} \rangle.
\]

Now it follows from Claim 1 that

\[
P_t f(x) = \int \langle p_{t/2,x}, p_{t/2,y} \rangle f(y) d\mu(y) = \int p(t, x, y) f(y) d\mu(y).
\]

It remains to show \(p(t, x, y) \geq 0\): It will be shown as an exercise (which relies on Lemma 6.2 and Example 6.3!) that \(f \leq 1\) implies \(P_t f \leq 1\). Thus if \(c > 0\) and \(f \leq c\) we have \(P_t f \leq c\). If \(f \geq 0\) we have \(-f \leq c\) for all \(c > 0\), so that we get \(P_t(-f) \leq c\) and taking \(c \to 0^+\) we have shown that \(f \geq 0\) implies \(P_t f \geq 0\). Thus writing

\[
p(t, x, y) = p(t, x, y)_+ - p(t, x, y)_-
\]

we get

\[
0 \leq P_t(p(t, x, \cdot)_-) = \langle p(t, x, \cdot), p(t, x, \cdot)_- \rangle
= \langle p(t, x, \cdot)_+, p(t, x, \cdot)_- \rangle - \langle p(t, x, \cdot)_-, p(t, x, \cdot)_- \rangle
= -\langle p(t, x, \cdot)_-, p(t, x, \cdot)_- \rangle,
\]

so \(\|p(t, x, \cdot)_-\|_2 = 0\) and the claim follows from continuity.

b) As by the fundamental lemma of distribution theory we have \(p_{t,x} = p(t, x, \cdot) \mu\text{-a.e.}, and \(p_{t,x} \in L^2(M)\) it is clear that

\[
\int p(t, x, y)^2 d\mu(y) < \infty.
\]

The symmetry \(p(t, y, x) = p(t, x, y)\) follows immediately from

\[
p(t, x, y) = \langle p_{t/2,x}, p_{t/2,y} \rangle.
\]

Next, for all \(0 < s' < t'\) one has

\[
p(t', x, y) = \langle p_{s',x}, p_{t'-s',y} \rangle,
\]
as the formula holds for $s' = t/2$ and the as the RHS does not depend on $s'$ by Claim 2. So
\[
\int p(t, x, z)p(s, z, y)d\mu(z) = \langle p(t, x, \cdot), p(s, y, \cdot) \rangle = \langle p_{s, x}, p_{s, y} \rangle = p(t + s, x, y).
\]
It remains to show
\[
\int p(t, x, y)d\mu(y) \leq 1.
\]
This follows by monotone convergence from $P_t f \leq 1$ for all $f \leq 1$, by letting $f$ run through $f = 1_{K_n}$ for $K_n$ some compact exhaustion of $M$.

c) = Step 5 and the proof of part a).

d) For fixed $s$ we set
\[
v(t, y) := p(t + s, x, y) = p(t + s, y, x) = \int p(t, y, z)p(s, z, x)d\mu(z) = P_t p(s, \cdot, x)(y),
\]
which by Step 5 solves the heat equation in $(t, y)$. It follows that $(t, y) \mapsto v(t - s, y) = p(t, x, y)$ solves the heat equation, too.

\section{8. Strong parabolic maximum principle and its applications}

From here on we will closely follow the presentation from Grigor’yan’s book [7]. The following result (and all its consequences) relies heavily on our standing assumption that $M$ is connected:

\textbf{Theorem 8.1.} i) The strong parabolic minimum principle holds: assume $I \subset \mathbb{R}$ is an open interval and $0 \leq u \in C^2(I \times M)$ solves
\[
\partial_t u \geq \Delta u.
\]
If there exists $(t', x') \in I \times M$ with $u(t', x') = 0$, then one has $u(t, x) = 0$ for all $x \in M$ and all $t \leq t'$.

ii) The strong parabolic maximum principle holds: assume $I \subset \mathbb{R}$ is an open interval and $0 \geq u \in C^2(I \times M)$ solves
\[
\partial_t u \leq \Delta u.
\]
If there exists $(t', x') \in I \times M$ with $u(t', x') = 0$, then one has $u(t, x) = 0$ for all $x \in M$ and all $t \leq t'$.

\textbf{Proof:} i) Step 1): Let $\Omega \subset \mathbb{R} \times M$ be nonempty, open, and relatively compact and assume\footnote{This means that $u$ is the restriction of $C^2$-function on $M$ to $\Omega$} $u \in C^2(\overline{\Omega})$ is such that
\[
\partial_t u \geq \Delta u \quad \text{in } \Omega.
\]
Then one has $\inf_{\overline{\Omega}} u = \inf_{\partial_p \Omega}$, where $\partial_p \Omega$ denotes the parabolic boundary of $\Omega$, which is defined as the complement
\[
\partial \Omega \setminus \partial_{\text{top}} \Omega,
\]
where $\partial_{\text{top}} \Omega$ denotes the set of all $(t, x) \in \partial \Omega$ which admit an open neighborhood $U \subset M$ of $x$ and $\epsilon > 0$ such that $(t - \epsilon) \times U \subset \Omega$. This is called the parabolic minimum principle.

Proof of step 1:
WLOG we can assume the strict inequality $\partial_t u > \Delta u$ in $\Omega$ (if this is not satisfied, one can replace $u$ by $u_\epsilon := u + \epsilon t$ and take $\epsilon \to 0^+$). Let
\[(t_0, x_0) := \min_{(s, y) \in \Omega} u(s, y).
\]
It suffices to show $(t_0, x_0) \in \partial_p \Omega$. Assume the contrary. Then either $(t_0, x_0) \in \Omega$ or $(t_0, x_0) \in \partial_{\text{top}} \Omega$. In both cases there exists a chart $U$ around $x_0$ and $\epsilon > 0$ such that $\Gamma := [t_0 - \epsilon, t_0] \times U \subset \Omega$.

By the definition of $(x_0, t_0)$, one has $t_0 = \min_{s \in [t_0 - \epsilon, t_0]} u(s, x_0)$, and so $\partial_t u(t_0, x_0) \leq 0$. By diagonalizing $g_{ij}(x_0)$ and making the induced coordinate transformation on $U$, we can assume that the coordinates $(x^1, \ldots, x^m)$ on $U$ satisfy, for some constants, $b_1, \ldots, b_m$,
\[
\Delta f(x_0) = \sum_{i=1}^m \frac{\partial^2}{(\partial x^i)^2} f(x_0) + \sum_{i=1}^m b_i \frac{\partial}{\partial x^i} f(x_0),
\]
for all $f \in C^2(U)$. Since $x_0 = \min_{y \in U} u(t_0, y)$, we have
\[
\frac{\partial}{\partial x^i} u(t_0, x_0) = 0, \quad \frac{\partial^2}{(\partial x^i)^2} u(t_0, x_0) \geq 0,
\]
and so $\Delta u(t_0, x_0) \geq 0$.

This estimate together with $\partial_t u(t_0, x_0) \leq 0$ contradicts $\partial_t u > \Delta u$ in $\Omega$ and proves the parabolic minimum principle.

Step 2): Let $V$ be a chart in $M$, let $x_0, x_1 \in V$ be such that the line connecting $x_0$ and $x_1$ lies in $V$ and assume $I \subset \mathbb{R}$ is an open interval and $0 \leq u \in C^2(I \times M)$ solves
\[
\partial_t u \geq \Delta u.
\]

Then for all $t_0, t_1 \in I$ with $t_1 > t_0$ and $u(t_0, x_0) > 0$ one has $u(t_1, x_1) > 0$.

Proof of step 2: Assume WLOG $t_0 = 0$ and that
- $V$ is relatively compact and its closure is contained in a chart,
- $r > 0$ is so small that the $2r$-neighborhood of the line connecting $x_0$ and $x_1$ lies in $V$, and that for $U := B_\epsilon(x_0, r)$ one has
\[
\inf_{x \in U} u(0, x) > 0.
\]
Set
\[ \zeta := \frac{1}{t_1}(x_1 - x_0). \]
Then for all \( t \in [0, t_1] \) one has \( U + t\zeta \in V \). Consider the open tilted cylinder (‘schiefer Zylinder’)
\[ \Gamma := \{(t, x) : t \in (0, t_1), x \in U + t\zeta \}. \]
We are going to show that \( u > 0 \) in \( \Gamma \) away from the lateral surface of \( \Gamma \) (‘Oberfläche von \( \Gamma \) ohne Deckel und Boden’). To this end, pick a function \( v \in C^2(\Gamma) \) such that
\[ \begin{align*}
\partial_t v &\leq \Delta v \quad \text{in } \Gamma, \\
v &= 0 \quad \text{on the lateral surface of } \Gamma \\
v &> 0 \quad \text{elsewhere on } \Gamma.
\end{align*} \]
Such a function \( v \) will be constructed in the exercises. Pick \( \epsilon > 0 \) such that
\[ \inf_{x \in U} u(0, x) \geq \epsilon \sup_{x \in U} v(0, x), \]
in particular, \( u \geq \epsilon v \) at the bottom \( U \) of \( \Gamma \). In particular, \( u \geq \epsilon \) on \( \partial_p \Gamma \). Since the function \( u - \epsilon v \) satisfies the assumptions of the parabolic minimum principle, one has \( u \geq \epsilon v \) in \( \Gamma \).

Step 3): The strong parabolic minimum principle holds:
Proof of step 3: Given \( u(t', x') = 0 \) for some \((t', x') \in I \times M\), it suffices to prove \( u(t, x) = 0 \) for all \((t, x) \in I \times M \) with \( t < t' \). Pick a finite sequence of points \( x_0, \ldots, x_k \) such such \( x_0 = x, x_k = x' \) and such that \( x_i \) and \( x_{i+1} \) lie in the same chart together with line connecting these two points, for all \( i = 0, \ldots, k \) (finally, here we use that \( M \) is connected!!). Picking a finite sequence of times \( t = t_0 < \ldots t_k = t' \) we can use step 2 \( k \)-times to deduce that if \( u(t_0, x_0) = u(t, x) > 0 \) then also \( u(t_1, x_1) > 0 \), and so \( u(t_2, x_2) > 0 \) and so on, yielding finally that \( u(t_k, x_k) = u(t', x') > 0 \), a contradiction. This completes the proof of the strong parabolic minimum principle.

ii) This follows from applying i) to \(-u\).

**Corollary 8.2.** One has \( p > 0 \).

**Proof:** Assume there exist \( t', x', y' \) with \( p(t', x', y') = 0 \). Then as \((t, y) \mapsto p(t, x', y)\) solves the heat equation one has \( p(t, x', y) = 0 \) for all \( y \in M \) all \( t \leq t' \). Pick \( \phi \) smooth compactly supported with \( \phi(x') = 1 \). Then we have
\[ \int p(t, x, y)\phi(y)d\mu(y) \to 0 \]
as \( t \to 0^+ \) by \( p(t, x', y) = 0 \) for all \( y \in M \) all \( t \leq t' \), while
\[ \int p(t, x, y)\phi(y)d\mu(y) \to 1 \]
as \( t \to 0^+ \) by Theorem 7.1 d) and \( \phi(x') = 1 \).

**Definition 8.3.** Given \( \alpha \in \mathbb{R} \), a real-valued function \( u \in C^2(M) \) is called
• $\alpha$-superharmonic, if $(-\Delta + \alpha)u \geq 0$,
• $\alpha$-subharmonic, if $(-\Delta + \alpha)u \leq 0$,
• $\alpha$-harmonic, if $(-\Delta + \alpha)u = 0$.

In the $\alpha$-harmonic case we can assume that $u$ is smooth by local elliptic regularity. If $\alpha = 0$, one simply says superharmonic (subharmonic) [harmonic], instead of 0-superharmonic, (0-subharmonic) [0-harmonic].

**Theorem 8.4** (Strong elliptic minimum/maximum principle). i) Assume $\alpha \in \mathbb{R}$ and that $u \geq 0$ is $\alpha$-superharmonic. If there exists $x_0$ with $u(x_0) = 0$, then one has $u \equiv 0$.

ii) Assume $\alpha \in \mathbb{R}$ and that $u \leq 0$ is $\alpha$-subharmonic. If there exists $x_0$ with $u(x_0) = 0$, then one has $u \equiv 0$.

**Proof:** i) Apply the strong parabolic minimum principle to $v(t, x) := e^{\alpha t}u(x)$.

ii) Apply i) to $-u$.

**Corollary 8.5.** i) If $u$ is superharmonic and if there exits $x_0$ with $u(x_0) = \inf u$, then $u \equiv \inf u$.

ii) If $u$ is subharmonic and if there exits $x_0$ with $u(x_0) = \sup u$, then $u \equiv \sup u$.

**Proof:** i) Apply the strong elliptic minumum principle to $\tilde{u} := u - \inf u$.

ii) Apply i) to $-u$.

**Example 8.6.** Let $N$ be a compact connected manifold (smooth without boundary). By picking a Riemannian metric on $N$, using the Hodge-Theorem and that continuous real-valued functions on a compact space attain their minimum and maximum, we get from the above Corollary

$$H^0(N) = \{ f : \Delta f = 0 \} = \{ \text{constant real-valued functions on } N \} = \mathbb{R}$$

for the zeroth homology group of $N$.

**Theorem 8.7** (Elliptic minimum/maximum principle). Let $V \subset M$ be open, relatively compact with $\partial V$ nonempty.

i) Assume $u \in C^2(V) \cap C(\overline{V})$ is superharmonic, then one has

$$\inf_{\overline{V}} u = \inf_{\partial V} u.$$

ii) Assume $u \in C^2(V) \cap C(\overline{V})$ is subharmonic, then one has

$$\sup_{\overline{V}} u = \sup_{\partial V} u.$$

**Proof:** i) set $r := \inf_{\overline{V}} u$ and

$$S := \{ x \in \overline{V} : u(x) = r \}.$$

It suffices to show that $S$ intersets $\partial V$. Assume not. Then one has $S \subset V$. We are going to show that the closed set $S$ is open, so $S = M$, a contradiction to $S \subset V \subset M \setminus \partial V$.

Let $x \in S \subset V$ and let $N \subset V$ be a connected open nbh of $x$. Then $u|_N$ attains its minimum in $x$, and so $u \equiv r$ by the above Corollary. Thus we have shown $N \subset S$, showing
that $S$ is open.

ii) Apply i) to $-u$.

9. Some spectral theory

In general, both parts of the spectrum (discrete spectrum and essential spectrum) of $H$ can be nonempty and the only thing we know for sure is $\sigma(H) \subset [0, \infty)$, as $H \geq 0$. The following simple result indicated that essential spectrum can only be nonempty on noncompactness $M$'s:

**Theorem 9.1.** Assume that for some $t > 0$ one has

$$\sup_{x \in X} p(t, x, x) < \infty,$$

and that $\mu(M) < \infty$. Then $H$ has a purely discrete spectrum (so the spectrum consists of eigenvalues having finite multiplicity), and if $(\lambda_n)$ denotes the increasing enumeration of the eigenvalues with each eigenvalue counted according to its multiplicity, then one has

$$0 \leq \lambda_n \uparrow \infty.$$

**Proof:** By abstract functional analysis it suffices to show that $P_t = e^{-tH}$ is Hilbert-Schmidt. But the latter is an integral operator, so it suffices to show

$$\int \int p(t, x, y)^2 d\mu(x) d\mu(y) < \infty.$$

Since

$$\int \int p(t, x, y)^2 d\mu(x) d\mu(y) = \int p(t, x, x) d\mu(x),$$

the claim follows from the assumptions.

The latter result clearly applies to compact $M$'s (so compact $M$'s have a purely discrete spectrum), but also to some noncompact $M$'s! For example, as we shall see later on (cf. Corollary 9.8 below), the result applies to open relatively compact subsets of an arbitrary Riemannian manifold (so those have a purely discrete spectrum, too). To prove the latter statement, we are going to show

$$p^U(t, x, y) \leq p(t, x, y),$$

where $U \subset M$ is an arbitrary open relatively compact subset and $p^U$ its heat kernel, that is, the heat kernel of the Riemannian manifold $(U, g|_U)$. To this end, we record:

**Lemma 9.2.** For all $0 \leq f \in W^{1,2}_0(M)$ there exists a sequence $0 \leq f_k \in C_\infty(M)$ with $f_k \to f$ as $k \to \infty$ in $W^{1,2}(M)$.

**Proof:** Pick a sequence $h_k \in C_\infty(M)$ with $h_k \to f$ in $W^{1,2}(M)$, and pick $\psi : \mathbb{R} \to [0, \infty)$ smooth with $\psi(0) = 0$ and $\sup_t |\psi'(t)| < \infty$. Then $0 \leq \psi \circ h_k \in C_\infty(M)$ and $\psi \circ h_k \to \psi \circ f$ in $W^{1,2}(M)$ by Lemma 6.2 (applied with a constant sequence). Thus it suffices to show that there exists a sequence $\psi_k : \mathbb{R} \to [0, \infty)$ smooth with $\psi_k(0) = 0$ and $\sup_{k,t} |\psi'_k(t)| < \infty$ such that $\psi_k \circ f \to f$ in $W^{1,2}(M)$. So this end, let $\phi(t) := 1_{(0,\infty)}(t)$, $\psi(t) := t_+$, $t \in \mathbb{R}$,
and pick $\psi_k$ as in Example 6.3. Then for some $C > 0$ we have $|\psi_k(t) - \psi(t)| \leq C|t|$ for all $t, k$, so that $\psi_k \circ f \to \psi \circ f = f$ in $L^2(M)$ by dominated convergence. Moreover, we have $d(\psi_k \circ f) = (\psi_k' \circ f)df$ by Lemma 6.2 which shows that $d(\psi_k \circ f) \to (\phi \circ f)df = df$ again by dominated convergence. This completes the proof. ■

Let $U \subset M$ be open and denote by $\tilde{\alpha}$ the trivial extension to $M$ by zero away from $U$ of a function or a 1-form on $U$. Then we consider $L^2(U)$ as a closed subspace of $L^2(M)$ via the embedding $f \mapsto \tilde{f}$, and likewise we have $\Omega^1_{L^2}(U) \subset \Omega^1_{L^2}(M)$.

**Lemma 9.3.** Let $U \subset M$ be open. Then for all $f \in W^{1,2}_0(U)$ one has $\tilde{f} \in W^{1,2}_0(M)$ and $d\tilde{f} = df$. In particular, $W^{1,2}_0(U) \subset W^{1,2}_0(M)$ is a closed subspace.

**Proof:** If $f \in C^\infty_c(U)$ then clearly $\tilde{f} \in C^\infty_c(M)$ with $d\tilde{f} = df$ by locality of differential operators.

Now let $f \in W^{1,2}_0(U)$ and pick a sequence $f_n \subset C^\infty_c(U)$ with $f_n \to f$ in $W^{1,2}(U)$. Then clearly $f_n \to f$ in $L^2(M)$. Moreover, $f_n$ is Cauchy in $W^{1,2}(M)$, and its limit must be $\tilde{f}$, because $\tilde{f}_n \to \tilde{f}$ in $L^2(M)$. In particular, $d\tilde{f}_n \to d\tilde{f}$ in $\Omega_{L^2}(M)$. On the other hand, we have $d\tilde{f}_n \to df$ in $\Omega_{L^2}(U)$ and so $d\tilde{f}_n \to df$ in $\Omega_{L^2}(M)$. In view of $d\tilde{f}_n = d\tilde{f}_n$, we have ultimately shown $df = df$. ■

The analogous result with $W^{1,2}_0$ replaced by $W^{1,2}$ is wrong: in $\mathbb{R}^m$, one has $1_{B(x,1)} = 1 \in W^{1,2}(B(x,1))$ (this is trivial), but $1_{B(x,1)} \notin W^{1,2}(\mathbb{R}^m)$ (this requires some work).

**Lemma 9.4.** Let $h \in W^{1,2}(M)$. Then there exists $v \in W^{1,2}_0(M)$ with $h \leq v$, if and only if one has $h \in W^{1,2}_0(M)$.

**Proof:** Exercise. ■

**Lemma 9.5.** Assume $0 \leq u \in W^{1,2}(M)$, $f \in L^2(M)$ is real-valued, $\lambda > 0$, and that one has $(-\Delta + \lambda)u \geq f$ weakly, meaning that

$$
\int u(-\Delta + \lambda)\phi d\mu \geq \int f \phi d\mu
$$

for all $0 \leq \phi \in C^\infty_c(M)$. Then one has $u \geq (H + \lambda)^{-1} f$.

**Proof:** Write $f = (H + \lambda)(H + \lambda)^{-1} f$ and set $v := (H + \lambda)^{-1} f \in \text{Dom}(H) \subset W^{1,2}_0(M)$. Then one has

$$
(-\Delta + \lambda)(v - u) \leq 0
$$

weakly, and so

$$
(-\Delta + \lambda)h \leq 0
$$

weakly, if we set $h := v - u \in W^{1,2}(M)$. Thus, integrating by parts,

$$
\int (dh, d\phi) d\mu + \lambda \int h\phi d\mu \leq 0
$$
for all $0 \leq \phi \in C_0^\infty(M)$. Since both sided are continuous in the $W^{1,2}$-norm, Lemma 9.2 shows that the the latter inequality holds for all $0 \leq \phi \in W^{1,2}_0(M)$. By Lemma 36 we have $0 \leq h_+ \in W^{1,2}_0(M)$, and so

$$
\int (dh, dh_+)d\mu + \lambda \int hh_+d\mu = \int (dh, dh_+)d\mu + \lambda \int h_+^2d\mu \leq 0.
$$

By Example 6.3 one has

$$
\int (dh, dh_+)d\mu = \int |dh_+|^2d\mu,
$$

and so

$$
\int |dh_+|^2d\mu + \lambda \int h_+^2d\mu \leq 0,
$$

and so $h_+ = 0$, as $\lambda > 0$. This shows $v - u \leq 0$, and so $v = (H + \lambda)^{-1}f \leq u$.

Lemma 9.6. For all $\lambda > 0$ one has $(H + \lambda)^{-1}f_1 \leq (H + \lambda)^{-1}f_2$, whenever $f_1, f_2 \in L^2(M)$ are such that $f_1 \leq f_2$.

Proof: By linearity we can assume $f_1 = 0$. Using the formula (Laplace-transforms)

$$
(r + \lambda)^{-1} = \int_0^\infty e^{-\lambda r}e^{-rs}dr,
$$

with $r = H$ (spectral calculus) implies

$$
(H + \lambda)^{-1}f_2 = \int_0^\infty e^{-\lambda s}P_sf_2ds \geq 0,
$$

where the integral converges in the $L^2$-sense, so the claim follows from $P_sf_2 \geq 0$ (which is a consequence of $p(s, x, y) \geq 0$). Note there that on any measure space $0 \leq h_n \to h$ in $L^p$ for some $p \in [1, \infty)$ implies $h \geq 0$.

Given $U \subset M$ open, we denote with $H^U$, $P^U$, $p^U$ the objects $H$, $P$, $p$ which are defined on the Riemannian manifold $(U, g|_U)$. Based on the above auxiliary results we can prove:

Theorem 9.7. For all open $U \subset M$, $t > 0$, $x, y \in U$ one has $p^U(t, x, y) \leq p(t, x, y)$.

Proof: It suffices to prove that for all $0 \leq f \in L^2(U)$, $x \in U$, one has

$$
\left(\int_U p^U(t, x, y)f(y)d\mu(y)\right) = P^U_t f(x) \leq P_f(x) = \left(\int_U p(t, x, y)f(y)d\mu(y)\right).
$$

Step 1: For all $\lambda > 0$ one has $(H^U + \lambda)^{-1}f \leq (H + \lambda)^{-1}f$.

Proof of step 1: We have $u := (H + \lambda)^{-1}f \in W^{1,2}_0(M) \subset W^{1,2}(M)$, and this function in $\geq 0$ by Lemma 9.6. Clearly, $u|_U \in W^{1,2}(U)$ and $(-\Delta + \lambda)u|_U = f|_U$. Thus, we have

$$
(H + \lambda)^{-1}f = u \geq (H^U + \lambda)^{-1}f.
$$

from Lemma 9.5.

Step 2: For all $\lambda > 0$, $k \in \mathbb{N}$, one has $(H^U + \lambda)^{-k}f \leq (H + \lambda)^{-k}f$. 

Proof of step 2: This follows from applying step 1 and Lemma 9.6 (using the latter for \(M\) and for \(U\)).

Step 3: One has (34).

Proof of step 3: By applying the formula \(e^{-tr} = \lim_k \left(\frac{k}{t}\right)^k (r + k/t)^{-k}\) for \(r = H, H^U\) (spectral calculus), we get the \(L^2\)-convergences

\[
P^U_t = e^{-tH^U} = \lim_k \left(\frac{k}{t}\right)^k (H^U + k/t)^{-k}, \quad P_t = e^{-tH} = \lim_k \left(\frac{k}{t}\right)^k (H + k/t)^{-k},
\]

so that the claim follows from Step 2.

Applying the last result with \(M\) replaced by \(V\), where \(V \subset M\) is open with \(U \subset V\), implies that for all \(0 \leq f \in L^2(M)\) one has \(P^U_t f|_U \leq P^V_t f|_V\), in particular if \((U_j)_{j \in \mathbb{N}}\) is an exhaustion of \(M\) with open subsets the limit \(\lim_n P^U_{t_n} f|_{U_n}\) exists pointwise. As one might guess, one has that for all \(t > 0\) (exercise)

\[
P^U_{t_n} f|_{U_n} \nearrow P_t f \quad \mu\text{-a.e.}
\]

With some more efforts, one can prove that the above relation actually holds pointwise (and even in the \(C^\infty\)), but we will need this stronger statement.

**Corollary 9.8.** For all open relatively compact \(U \subset M\) the operator \(H^U\) has a purely discrete spectrum.

**Proof:** Combine Theorem 9.7 with Theorem 9.1.

10. Integrated maximum principle

**Theorem 10.1 (Integrated Maximum Principle).** Let \(I \subset [0, \infty)\) be an interval and let \(\zeta : I \times M \to \mathbb{R}\) be continuous such that

i) for all \(t \in I\), the function \(\zeta(t, \cdot)\) is locally Lipschitz,

ii) \(\partial_t \zeta\) exists and is continuous on \(I \times M\),

iii) one has

\[
\partial_t \zeta + \frac{1}{2} |d\zeta|^2 \leq 0.
\]

Then with

\[
\lambda_{\min}(M) := \inf \sigma(H),
\]

for all \(f \in L^2(M)\) the function

\[
I \ni t \mapsto J(t) := \int |P_tf|^2(x)e^{\zeta(t,x)}d\mu(x) \in [0, \infty]
\]

satisfies

\[
J(t) \leq J(t_0)e^{-2\lambda_{\min}(M)(t-t_0)} \quad \text{for all } t, t_0 \in I \text{ with } t > t_0,
\]

in particular, \(J\) is nonincreasing.

**Remark 10.2.** We make no statement here on the finiteness of \(J(t)\)! 
Proof: Because of \(|P_I f|^2 = |P_{t-t_0}P_{t_0} f|^2 \leq (P_{t-t_0}|P_{t_0} f|)^2\) WLOG (check this!) we can and we will assume \(f \geq 0\). Then, using (35) in combination with monotone convergence, and that by the variational principle (8)

\[
\lambda_{\text{min}}(U) = \inf_{\Psi \in W_0^{1,2}(U)} \left( \int_U |d\Psi|^2 d\mu / \left( \int_U |\Psi|^2 d\mu \right) \right)
\]

in combination with \(W_0^{1,2}(U) \subset W_0^{1,2}(M)\) one has

\[
\lambda_{\text{min}}(U) \geq \lambda_{\text{min}}(M),
\]

it suffices to show that for all open relatively compact \(U \subset M\) one has

\[
J^U(t) := \int_U (P^U_t f)^2(x)e^{\zeta(t, x)}d\mu(x) \leq J^U(t_0)e^{-2\lambda_{\text{min}}(U)(t-t_0)},
\]

where here and in the sequel \(P^U_t f\) is to be understood in the obvious sense \(P^U_t f := P^U_t f|_U\). Note that \(J^U\) is finite and continuous on \(I\), for we have

\[
J^U = \langle P^U_t f, e^\zeta P^U_t f \rangle_U.
\]

Thus in order to show (36), it suffices to show that \(J^U\) is differentiable in \(I \setminus \{0\}\) with

\[
(d/dt)J^U(t) \leq -2\lambda_{\text{min}}(U)J^U(t)
\]

for all \(t \in I \setminus \{0\}\). The functions \(\zeta(t, \cdot)\) and \(\partial_t \zeta(t, \cdot)\) are in \(C_b(U)\), so that (exercise) \(\partial_t \zeta\) is equal to the strong derivative \((d/dt)\zeta\). The same remark applies to \(e^\zeta\), and so

\[
(d/dt)e^\zeta = \partial_t e^\zeta = e^\zeta \partial_t \zeta.
\]

On the other hand \(P^U_t f\) is strongly differentiable as an \(L^2(U)\)-valued map by the spectral calculus with we record

\[
(d/dt)P^U_t f = \Delta P^U_t f.
\]

Using that \(e^\zeta\) is strongly differentiable as an \(C_b(U)\)-valued (and thus as an \(L^\infty(U)\)-valued) map it follows that \(P^U_t f e^\zeta\) is a strongly differentiable \(L^2(U)\)-valued map and one has

\[
(d/dt)(P^U_t f e^\zeta) = [(d/dt)P^U_t f]e^\zeta + [(d/dt)e^\zeta]P^U_t f.
\]

The product rule applied above is as follows: if \(s \mapsto h_1(s)\) is a strongly differentiable map from an open subset \(A \subset \mathbb{R}\) to \(L^\infty\), and if \(s \mapsto h_2(s)\) is strongly differentiable map from \(A\) to \(L^2\), then \(s \mapsto h_1(s)h_2(s)\) is a strongly differentiable map from \(A\) to \(L^2\), and the product rule applies. Thus

\[
J^U = \langle P^U_t f, e^\zeta P^U_t f \rangle_U
\]

is differentiable with

\[
(d/dt)J^U = \langle (d/dt)P^U_t f, e^\zeta P^U_t f \rangle_U + \langle P^U_t f, (d/dt)[ue^\zeta] \rangle_U
\]

\[
= 2 \langle (d/dt)P^U_t f, e^\zeta P^U_t f \rangle_U + \langle (P^U_t f)^2, (d/dt)e^\zeta \rangle_U
\]

\[
= 2 \langle \Delta P^U_t f, e^\zeta P^U_t f \rangle_U + \langle (P^U_t f)^2, [\partial \zeta]e^\zeta \rangle_U.
\]
By the Chain rule from Lemma 6.7 we have that $e^{\zeta(t, \cdot)}$ is locally Lipschitz on $M$, so this function is Lipschitz (and bounded) on $U$. As by the spectral calculus we have $P_t f \in W^{1,2}_0(U)$ it follows that $e^{\zeta} P_t f \in W^{1,2}_0(U)$ by Lemma 6.8 and we may integrate by parts

$$2 \langle \Delta P_t f, e^{\zeta} P_t f \rangle_U = 2 \int_U \Delta P_t f \cdot e^{\zeta} P_t f d\mu = -2 \int_U (dP_t f, d(e^{\zeta} P_t f)) d\mu.$$ 

In addition, $P_t f$ and $e^{\zeta(t, \cdot)}$ are locally Lipschitz in $U$ so that using the product rule from Lemma 6.7 and the chain rule therein we get

$$d(e^{\zeta} P_t f) = P_t f \cdot de^{\zeta} + e^{\zeta} \cdot dP_t f = P_t f \cdot e^{\zeta} d\zeta + e^{\zeta} \cdot dP_t f$$

and so

$$2 \langle \Delta P_t f, e^{\zeta} P_t f \rangle_U = -2 \int_U \left( dP_t f, P_t f \cdot e^{\zeta} d\zeta \right) d\mu - 2 \int_U \left( dP_t f, e^{\zeta} \cdot dP_t f \right) d\mu.$$ 

Plugging this into (41) and using assumption ii) from the theorem we obtain

$$(d/dt) J_U = 2 \langle \Delta P_t f, e^{\zeta} P_t f \rangle_U + \langle (P_t f)^2, (\partial \zeta) e^{\zeta} \rangle_U$$

$$= -2 \int_U e^{\zeta} P_t f \left( dP_t f, d\zeta \right) d\mu - 2 \int_U e^{\zeta} \left( dP_t f, dP_t f \right) + \int_U (P_t f)^2 \cdot (\partial \zeta) \cdot e^{\zeta} d\mu$$

$$\leq -2 \int_U \left( e^{\zeta} P_t f \left( dP_t f, d\zeta \right) + \int_U e^{\zeta} \left( dP_t f, dP_t f \right) + \frac{1}{4} e^{\zeta} (P_t f)^2 |d\zeta|^2 \right) d\mu$$

$$= -2 \int_U \left| dP_t f + \frac{1}{2} P_t f d\zeta \right|^2 e^{\zeta} d\mu$$

$$= -2 \int_U \left| d(e^{\zeta/2} P_t f) \right|^2 d\mu \leq -2 \lambda_{\min}(U) \int_U |e^{\zeta/2} P_t f|^2 d\mu$$

where the last equality follows from

$$\left( dP_t f + \frac{1}{2} P_t f d\zeta \right) e^{\zeta/2} = d(e^{\zeta/2} P_t f)$$

and the last inequality from the variational principle, since $e^{\zeta/2} P_t f \in W^{1,2}_0(U)$. Noting that

$$\int_U |e^{\zeta/2} P_t f|^2 d\mu = J_U,$$

this completes the proof.

\[\square\]

11. \(L^2\)-mean-value-inequality (MVI)

**Definition 11.1.** Let $I \subset \mathbb{R}$ be an interval and $V \subset M$ be open. Then a $C^2$-function $u : I \times V \to \mathbb{R}$ is called a subsolution of the heat equation, if $\partial_t u \leq \Delta u$. 


Theorem 11.2 (L²-MVI). Assume $B(x, R) \subset M$ is relatively compact and that for some $a, n > 0$ one has the following Faber-Krahn inequality:

$$\lambda_{\min}(U) = \min \sigma(H^U) \geq a\mu(U)^{-2/n} \quad \text{for all open } U \subset B(x, R).$$

Then there exists a constant $C_n$, which only depends on $n$, such that for all $T > 0$ and all subsolutions $u$ of the heat equation in the cylinder $\mathcal{C} := (0, T] \times B(x, R)$ one has

$$u_+^2(T, x) \leq \frac{C_n a^{-n/2}}{\min(\sqrt{T}, R)^{n+1}} \int_\mathcal{C} u_+^2 d\nu,$$

where $d\mu := dt d\mu$ is the product of the Lebesgue measure on $\mathbb{R}$ and the Riemann volume measure on $M$.

The proof of the $L^2$-MVI requires two auxiliary results:

Lemma 11.3. Let $V \subset M$ be open, $0 \leq T_0 < T$ and let $\eta$ be a Lipschitz function on $\mathcal{C} := [T_0, T] \times V$ (considered as a Riemann manifold) such that for some compact $K \subset V$ one has $\text{supp}(\eta(t, \cdot)) \subset K$ for all $t \in [T_0, T]$. Let $u$ be a subsolution of the heat equation in $\mathcal{C}$ and set $v := (u - \theta)_+$ for some $\theta \geq 0$. Then one has

$$\frac{1}{2} \left( \int_V v^2(T, \cdot) \eta^2(T, \cdot) d\mu - \int_V v^2(T_0, \cdot) \eta^2(T_0, \cdot) d\mu \right) + \int_\mathcal{C} |d(v\eta)|^2 d\nu \leq \int_\mathcal{C} v^2(|d\eta|^2 + |\eta \partial_\tau \eta|) d\nu. \quad (44)$$

In particular, if $\eta(T_0, \cdot) = 0$, then the following two additional inequalities hold:

$$\int_V v^2(t, \cdot) \eta^2(t, \cdot) d\mu \leq 2 \int_\mathcal{C} v^2(|d\eta|^2 + |\eta \partial_\tau \eta|) d\nu \quad \text{for all } t \in [T_0, T], \quad (45)$$

and

$$\int_\mathcal{C} |d(v\eta)|^2 d\nu \leq \int_\mathcal{C} v^2(|d\eta|^2 + |\eta \partial_\tau \eta|) d\nu. \quad (46)$$

Proof: Note first that the second inequality follows by applying the first with $T$ replaced by $t$, and the last one follows trivially from the first one (by leaving away one positive summand on the LHS).

To prove the first inequality, since $u(t, \cdot) \in C^2(V) \subset W^{1,2}_{\text{loc}}(V)$ for all $t$, by Lemma 6.2 one has $v(t, \cdot) \in W^{1,2}_{\text{loc}}(V)$ with

$$dv = 1_{\{u > \theta\}} du = 1_{\{v \neq 0\}} du,$$

so

$$\langle dv, du \rangle = |dv|^2, \quad vdu = vdv. \quad (47)$$

Since $\eta(t, \cdot)$ is compactly supported and Lipschitz in $V$ (and so its square, too) one has $v(t, \cdot) \eta(t, \cdot)^2 \in W^{1,2}_{0}(V)$ by Lemma 6.9, with

$$d(v\eta)^2 = v d\eta^2 + \eta^2 dv = 2v \eta d\eta + \eta^2 dv, \quad (48)$$

where $d\mu := dt d\mu$ is the product of the Lebesgue measure on $\mathbb{R}$ and the Riemann volume measure on $M$.\]
thus

\[(du, d(v\eta^2)) = 2\eta(v dv, d\eta) + \eta^2|dv|^2].\]

If we multiply \(\partial_t u \leq \Delta\) with \(v\eta^2\) and perform \(\int_{\mathcal{C}} \cdots d\nu\), we get

\[
\int_{\mathcal{C}} (\partial_t u)v\eta^2 d\nu \leq \int_{T_0}^T \int_V (\Delta u)v\eta^2 d\mu dt
\]

\[
= -\int_{T_0}^T \int_V (du, d(v\eta^2))d\mu dt
\]

\[
= -\int_{T_0}^T \int_V \left(2\eta(v dv, d\eta) + \eta^2|dv|^2\right)d\mu dt
\]

\[
- \int_{T_0}^T \int_V \left(|d(v\eta)|^2 - \eta^2|d\eta|^2\right)d\mu dt,
\]

where have integrated by parts (Lemma 6.1; note that \(v(t, \cdot)\eta(t, \cdot)^2 \in W^{1,2}_0(V')\) on some open relatively compact neighbourhood \(V' \subset V\) of \(K\) and that \(u(t, \cdot), du(t, \cdot), \Delta u(t, \cdot)\) are square integrable on \(V', u(t, \cdot)\) is \(C^2\) on \(V\), and where we have used (50). Let us also record an application of Lemma 6.2 to the \(t\)-variable gives as above

\[(52)\]

\[v\partial_t u = v\partial_t v,\]

which shows the first identity in

\[
\int_{T_0}^T (\partial_t u)v\eta^2 dt = \frac{1}{2} \int_{T_0}^T (\partial_t v^2)\eta^2 dt
\]

\[= \frac{1}{2} \left( v^2(T, \cdot)\eta^2(T, \cdot) - v^2(T_0, \cdot)\eta^2(T_0, \cdot) \right) - \frac{1}{2} \int_{T_0}^T v^2\partial_t\eta^2 dt
\]

\[= \frac{1}{2} \left( v^2(T, \cdot)\eta^2(T, \cdot) - v^2(T_0, \cdot)\eta^2(T_0, \cdot) \right) - \int_{T_0}^T v^2\eta\partial_t\eta dt,
\]

where we have integrated by parts (noting that this can be done for Lipschitz functions, say by using Friedrichs mollifiers). If we perform \(\int_V \cdots d\mu\) in the last identity and use (51), the proof is complete. 

Lemma 11.4. In the situation of the \(L^2\)-MVI consider

\[\mathcal{C}_i := [T_i, T] \times B(x, R_i), \quad i = 0, 1,\]

where \(R_i, T_i\) are chosen with \(0 < R_1 < R_0 \leq R, 0 \leq T_0 < T_1 < T\). Chose \(\theta_1 > \theta_0 \geq 0\) and set

\[J_i := \int_{\mathcal{C}_i} (u - \theta_i)^2 d\nu, \quad i = 0, 1.\]

Then for some constant \(c_n\), which only depends on \(n\), one has

\[J_1 \leq \frac{c_n J_0^{1+2/n}}{a\delta^{1+2/n}(\theta_1 - \theta_0)^{4/n}},\]
where
\[ \delta := \min \left( T_1 - T_0, (R_0 - R_1)^2 \right). \]

**Proof:** WLOG \( \theta_0 = 0, \theta := \theta_1. \) Define \( \eta \) by \( \eta(t, y) := \phi(t)\psi(y), \) where

\[ \phi(t) := \min \left( \frac{t - T_0}{T_1 - T_0}, 1 \right), \]
\[ \psi(y) := \min \left( \frac{(R_{1/4} - \varrho(x, y))_+}{R_{1/4} - R_{1/2}}, 1 \right), \]

where \( R_\lambda := \lambda R_1 + (1 - \lambda)R_0, \lambda \in [0, 1]. \) Note that \( R_\lambda \leq R_\lambda', \) iff \( \lambda' \leq \lambda. \) Note also that by construction \( \eta(t, \cdot) \) is supported in the compact ball \( K := B(x, R_{1/4}). \) Applying the second inequality from the previous lemma to \( \mathcal{C}_0, \) \( v := u_+, \) with \( t \in [T_1, T] \subset [T_0, T] \) gives the bound

\[ \int_{B(x, R_{1/2})} u_+^2(t, \cdot) d\mu \leq \int_{B(x, R_0)} u_+^2(t, \cdot) d\mu \leq 2 \int_{\varepsilon_0} u_+^2(|d\eta|^2 + |\eta \partial \eta|) d\nu, \]

where we have used that \( \eta = 1 \) in \( [T_1, T] \times B(x, R_{1/2}). \) Clearly we have

\[ \eta \leq 1, \quad |d\eta|^2 \leq (R_{1/4} - R_{1/2})^{-2} = 16(R_0 - R_1)^{-2} \leq 16/\delta, \quad |\partial \eta| \leq (T_1 - T_0)^{-1} \leq 1/\delta, \]

and so

\[ \int_{B(x, R_{1/2})} u_+^2(t, \cdot) d\mu \leq 34 \delta^{-1} J_0. \]

Fix \( t \) as above and set

\[ U_t := \{ y \in B(x, R_{3/4}) : u(t, y) > \theta \}, \]

so that by the latter inequality we get

(53) \[ \mu(U_t) \leq \frac{1}{\theta^2} \int_{B(x, R_{3/4})} u_+^2(t, \cdot) d\mu \leq \frac{1}{\theta^2} \int_{B(x, R_{1/2})} u_+^2(t, \cdot) d\mu \leq \frac{34 J_0}{\theta^2 \delta}. \]

Define now

\[ \psi'(y) := \min \left( \frac{(R_{3/4} - \varrho(x, y))_+}{R_{3/4} - R_1}, 1 \right) \]

and \( \eta'(t, y) := \phi(t)\eta'(y). \) Applying the third inequality of the previous lemma to \( v' = (u - \theta)_+ \) in \( \mathcal{C}_0 \) gives with a similar reasoning as above the inequality

(54) \[ \int_{\varepsilon_0} |d(v'\eta')|^2 d\nu \leq \int_{\varepsilon_0} v'^2(|d\eta'|^2 + |\eta' \partial \eta'|) d\nu \leq \frac{17}{\delta} \int_{\varepsilon_0} v'^2 d\nu \leq \frac{17 J_0}{\delta}. \]

The function \( \eta'(t, \cdot)v'(t, \cdot) \) is supported in the compact set \( \overline{U_t}, \) thus an element of \( W^{1,2}_0(V) \) for all open \( V \subset M \) with \( U_t \subset V. \) Choose such a \( V \) such that in addition \( V \subset B(x, R_0) \)
and

\[
\mu(V) \leq 2\mu(U_t) \leq \frac{68J_0}{\theta^2\delta},
\]

where the second inequality follows from (53). Since \(\eta'(t, \cdot) v'(t, \cdot) \in W^{1,2}_0(V)\) we can use the variational principle for \(V\) to conclude (using the support properties of \(\eta'(t, \cdot) v'(t, \cdot)\))

\[
\int_{B(x, R_0)} |d(\eta'(t, \cdot) v'(t, \cdot))|^2 d\mu \\
= \int_V |d(\eta'(t, \cdot) v'(t, \cdot))|^2 d\mu \\
\geq \lambda_{\text{min}}(V) \int_V (\eta'(t, \cdot) v'(t, \cdot))^2 d\mu \\
= \lambda_{\text{min}}(V) \int_{B(x, R_0)} (\eta'(t, \cdot) v'(t, \cdot))^2 d\mu \\
\geq a\mu(V)^{-2/n} \int_{B(x, R_0)} (\eta'(t, \cdot) v'(t, \cdot))^2 d\mu \\
\geq a \left( \frac{\theta^2\delta}{68} \right)^{2/n} J_0^{-2/n} \int_{B(x, R_0)} (\eta'(t, \cdot) v'(t, \cdot))^2 d\mu,
\]

where we have used the Faber-Krahn inequality (an assumption) and that \(\eta' = 1\) in \([T_1, T] \times B(x, R_1)\). The lemma now immediately follows from integrating the latter inequality with respect to \(\int_{T_1} \cdots dt\) and using (54).

Now we can give the proof of the \(L^2\)-MVI: Assume for the moment that \(\theta \geq 0\) is arbitrary and define \(\delta_k = \delta_k(\theta) > 0\) by

\[
\delta_k := \left( \frac{16^{2/n} C_n 16^{-1-k/n} J_0^{2/n}}{a\theta^{4/n}} \right)^{n/(n+2)} = C_n' 16^{-k/(n+2)} J_0^{2/(n+2)}
\]

where \(C_n\) is as in the previous lemma, where \(C_n' := 16^{2/n} C_n\), and where

\[
J_0 := \int_{\phi} (u - \theta/2)^2 d\nu.
\]

Set

\[
T_k := \sum_{i=0}^{k-1} \delta_i, \quad R_k := R - \sum_{i=0}^{k-1} \sqrt{\delta_i}.
\]

\[\text{In fact } \mu(V) \leq 2\mu(U_t) \text{ is a little technical to prove, but replacing 2 with some } A = A_n \text{ follows straightforwardly from covering } B(x, R_0) \text{ with Euclidean balls and using that locally } g \text{ is equivalent to the Euclidean metric.}\]
Then there exists $C''_n > 0$ such that (geometric series) if
\[ \theta^2 := \frac{a^{-n/2}J_0}{C''_n \min(\sqrt{T}, R)^{n+2}}, \]
then
\[ T_k \leq \sum_{i=0}^{\infty} \delta_i \leq T/2, \quad R_k \geq R - \sum_{i=0}^{\infty} \sqrt{\delta_i} \leq R/2. \]
Define a sequence of cylinders by
\[ \mathcal{C}_k := [T_k, T] \times B(x, R_k) \]
and further
\[ \theta_k := (1 - 2^{-(k+1)})\theta, \quad J_k := \int_{\mathcal{C}_k} (u - \theta_k)^2 d\nu. \]
Then we have
\[ \mathcal{C}_0 = \mathcal{C}, \quad [T/2, T] \times B(x, R/2) \subset \mathcal{C}_{k+1} \subset \mathcal{C}_k. \]
We claim that $J_k \to 0$ as $k \to \infty$, which would complete the proof, for one has
\[ \int_{T/2}^{T} \int_{B(x,R/2)} (u - \theta)^2 d\nu \leq J_k \to 0, \]
so $u_+(T,x) \leq \theta$, which implies the assertion of the $L^2$-MVI.  
In order to show $J_k \to 0$, we apply the previous lemma to the pair $(\mathcal{C}_k, \mathcal{C}_{k+1})$, which using
\[ \delta_k = (R_k - R_{k+1})^2 = (T_{k+1} - T_k) = \min((R_k - R_{k+1})^2, (T_{k+1} - T_k)) \]
yields the inequality
\[ J_{k+1} \leq \frac{C_n J_k^{1+2/n}a^{1+2/n}(\theta_{k+1} - \theta_k)^{4/n}}{\delta_k^{1+2/n}(\theta_{k+1} - \theta_k)^{4/n}}. \]
Using this inequality and the defining formula for $\delta_k$ a simple induction on $k$ shows
\[ J_k \leq 16^{-k}J_0, \]
which completes the proof. \[\square\]

12. Equivalent characterizations of Li-Yau upper heat kernel bounds

**Definition 12.1.** Given $t, D > 0, x \in M$ define the weighted $L^2$-norm of the heat kernel by
\[ E_D(t,x) := \int_{M} p(t,x,y)^2 e^{-y} d\mu(y) \in [0, \infty). \]
Note that in $\mathbb{R}^m$ one has $E_D(t,x) = \infty$ for all $t, x$ of $D \leq 2$. 
Remark 12.2. Let \( s > 0 \) be arbitrary. As we have

\[
E_D(t, x) = \int_M (P_{t-s} f)^2 e^{\zeta(t, \cdot)} d\mu,
\]

for all \( t > s \), where

\[
f := p(s, x, \cdot), \quad \zeta(t, y) := \varrho(x, y)^2/(Dt),
\]

it follows from the integrated maximum principle that \( t \mapsto E_D(t, x) \) is nonincreasing on \((0, \infty)\), and

\[
E_D(t, x) \leq E_D(t_0, x) e^{-2\lambda_{\min}(t-t_0)}
\]

so long as \( 0 < t_0 \leq t \).

Theorem 12.3. Let \( B(x, r) \) be a relatively compact ball and assume the following Faber-Krahn inequality: there exist \( a, n > 0 \) such that for all open \( U \subset B(x, r) \) one has

\[
\lambda_{\min}(U) \geq a \mu(U)^{-2/n}.
\]

Then for all \( t > 0, D > 2 \) one has

\[
E_D(t, x) \leq \frac{C_n(a\delta)^{-n/2}}{\min(t, r^2)^{n/2}} \quad \text{where} \quad \delta := \min(D - 2, 1)
\]

Remark 12.4. On \( \mathbb{R}^m \) one has the Faber-Krahn inequality uniformly on any ball (a classical fact), in the sense that for some \( a = a_m > 0 \), and all \( U \subset \mathbb{R}^m \) open one has\(^{10}\)

\[
\lambda_{\min}(U) \geq a \mu(U)^{-2/m}.
\]

Furthermore, one can show that there exists a (Lipschitz continuous) function \( \hat{r} \rightarrow \mathbb{R}^+ \rightarrow (0, \infty) \), so that one has

\[
\lambda_{\min}(U) = \min \sigma(H_U) \geq a \mu(U)^{-2/m} \quad \text{for all open} \ U \subset B(x, \hat{r}(x)),
\]

where the constant \( a > 0 \) only depends on \( m = \dim(M) \). Thus for all \( x \in M \) there exists \( r > 0 \) such that one has (56), and so the above theorem shows that for all \( t > 0, D > 2, x \in M \) one has \( E_D(t, x) \), regardless of how bad the geometry of \( M \) is! This is a highly nontrivial fact.

We prepare the proof of Theorem 12.3 with

Lemma 12.5. Let \( B(x, r) \) be a relatively compact ball and assume the following Faber-Krahn inequality: there exist \( a, n > 0 \) such that for all open \( U \subset B(x, r) \) one has

\[
\lambda_{\min}(U) \geq a \mu(U)^{-2/n}.
\]

Set \( \varrho(y) := (\varrho(x, y) - r)_+ , y \in M \). Then for all \( t > 0 \) one has

\[
\int_M p(t, x, y)^2 e^{\varrho(y)^2} d\mu(y) \leq \frac{C_n a^{-n/2}}{\min(t, r^2)^{n/2}}
\]

\(^{10}\)Sketch of proof: One first shows that \( \lambda_{\min}(U) \geq \lambda_{\min}(U^*) \), where \( U^* \) is any ball with the same volume as \( U \) (this estimate uses the co-area formula). Then one shows \( \lambda_{\min}(U^*) = c_m/r^2 \), and uses \( \mu(U) = \mu(U^*) = c_m r^m \).
**Proof:** With the same argument as in Remark 12.2 one finds that the left hand side of the inequality is noncreasing in \(t\), so WLOG we can assume \(t \leq r^2\). For the moment let \(0 \leq f \in L^2(M)\) be arbitrary and set \(u := Pf\). The \(L^2\)-mean value inequality applied to the cylinder \([t/2, t] \times B(x, r)\) shows

\[
u(t, x)^2 \leq \frac{C_n a^{-n/2}}{t_1+2/n} \int_0^t \int_{B(x, r)} u(s, y)^2 d\mu(y) ds.
\]

Set \(\zeta(s, y) := \varrho(y)^2(2(s - t))^{-1}\) and note \(\text{supp}(\zeta(s, \cdot)) \subset M \setminus B(x, r)\), so that the latter inequality is gives

\[
u(t, x)^2 \leq \frac{C_n a^{-n/2}}{t_1+2/n} \int_0^t \int_{B(x, r)} u(s, y)^2 \zeta(s, y) d\mu(y) ds.
\]

Again the integrated maximum principle implies that

\[
J(s) := \int_M u(s, y)^2 e^{\zeta(s, y)} d\mu(y) ds
\]

is nonincreasing in \(s \in [0, t]\) and so the latter inequality implies

\[
u(t, x)^2 \leq \frac{C_n a^{-n/2}}{t_1+2/n} \int_0^{t/2} J(s) ds
\]

\[
\leq \frac{C_n a^{-n/2}}{t_1+2/n} J(0)
\]

\[
= C_n'(at)^{-n/2} \int f^2 e^{-\varrho^2/(2t^2)} d\mu.
\]

Pick \(\phi\) smooth compactly supported with \(0 \leq \phi \leq 1\) and applying the latter inequality with

\[
f := p(t, x, \cdot)e^{\varrho^2/(2t)} \phi
\]

to get

\[
\left( \int_M p(t, x, \cdot)^2 e^{\varrho^2/(2t)} \phi d\mu \right)^2 \leq C_n( at)^{-n/2} \int_M p(t, x, \cdot)^2 e^{\varrho^2/(2t)} \phi^2 d\mu
\]

\[
\leq C_n( at)^{-n/2} \int_M p(t, x, \cdot)^2 e^{\varrho^2/(2t)} \phi d\mu,
\]

which is equivalent to the assertion of the lemma (letting \(\phi \rightarrow 1\)).

**Proof of Theorem 12.3:** It suffices to prove the inequality for \(D \leq 3\) and \(t \leq r^2\) (as \(E_D(t, x)\) is decreasing \(t\) and \(D\)).

We have \(\sqrt{\delta t} \leq r\). Thus Faber-Krahn holds in \(B(x, \sqrt{\delta t})\) and applying the previous lemma on that ball we get

\[
\int_M p(t, x, y)^2 e^{\frac{(\varrho(x, y) - \sqrt{\delta t})^2}{2t}} d\mu(y) \leq \frac{C_n a^{-n/2}}{(\delta t)^{n/2}}
\]
using
\[ a^2/t_2 + b^2/t_1 \geq (a + b)^2/(t_1 + t_2), \]
valid for all \( a, b \in \mathbb{R}, t_1, t_1 > 0 \), we get
\[ \frac{(\varrho(x, y) - \sqrt{t\delta})^2}{2t} + \frac{\sqrt{t\delta}^2}{\delta t} \geq \frac{\varrho(x, y)^2}{(2 + \delta)t}, \]
thus
\[ \frac{(\varrho(x, y) - \sqrt{t\delta})^2}{2t} \geq \frac{\varrho(x, y)^2}{Dt} - 1, \]
which completes the proof. \( \blacksquare \)

The following lemma connects the weighted \( L^2 \)-norm with Gaussian heat kernel upper bounds:

**Lemma 12.6.** For all \( D > 0, t \geq t_0 > 0, x, y \in M \) one has
\[ p(t, x, y) \leq \sqrt{E_D(t_0/2, x)E_D(t_0/2, y)}e^{-\varrho(x, y)^2/(2Dt) - \lambda_{\min}(M)(t-t_0)}, \]
in particular,
\[ p(t, x, y) \leq \sqrt{E_D(t/2, x)E_D(t/2, y)}e^{-\varrho(x, y)^2/(2Dt)}. \]

**Proof:** In view of (55) it suffices to prove (59). Set
\[ a := \varrho(y, z), \quad b := \varrho(x, z), \quad c := \varrho(x, y). \]
Then \( \exp((a^2 + b^2 - c^2/2)/(Dt)) \geq 1 \) by the triangle inequality, and so
\[ p(t, x, y) = \int_M p(t/2, x, z)p(t/2, y, z)d\mu(z) \leq \exp(-\varrho(x, y)^2/(2Dt)) \int_M p(t/2, x, z)\exp(\varrho(x, z)^2/(Dt))p(t/2, y, z)\exp(\varrho(y, z)^2/(Dt))d\mu(z) \]
for which the claim follows using Cauchy-Schwarz. \( \blacksquare \)

**Theorem 12.7.** Assume \((B(x_i, r_i))_{i \in I}\) is a family of relatively compact balls such that there exists constant \( n > 0 \) and for each \( i \) a constant \( a_i > 0 \) with the following property: for each \( i \) and each open \( U \subset B(x_i, r_i) \) one has the Faber-Krahn inequality
\[ \lambda_{\min}(U) \geq a_i\mu(U)^{-2/n}. \]
Then for all \( i, j \), all \( x \in B(x_i, r_i/2), y \in B(x_j, r_j/2) \), all \( t \geq t_0 > 0 \) one has
\[ p(t, x, y) \leq \frac{C_n(1 + \varrho(x, y)^2/t)^{n/2}e^{-\varrho(x, y)^2/(4t) - \lambda_{\min}(t-t_0)}}{(a_i a_j \min(t_0, r_i^2) \min(t_0, r_j^2))^{n/4}}. \]

**Proof:** We have \( B(x, r_i/2) \subset B(x_i, r_i) \) so the Faber-Krahn inequality holds on \( B(x, r_i/2) \) and Theorem 12.3 gives with
\[ D := 2 + (1 + \varrho(x, y)^2/t)^{-1} \]
the bound
\[ E_D(t, x) \leq \frac{C_n(a_i \delta)^{-n/2}}{\min(t, r_i^2)^{n/2}} \]
where \( \delta = \min(D - 2, 1) = (1 + \varrho(x, y)^2/t)^{-1} \).
Likewise we have
\[ E_D(t, y) \leq \frac{C_n(a_j \delta)^{-n/2}}{\min(t, r_j^2)^{n/2}}. \]
and so the previous lemma gives
\[ p(t, x, y) \leq \frac{C_n(1 + \varrho(x, y)^2/t)^{n/2} e^{-\varrho(x,y)^2/(2Dt) - \lambda_{\min}(t-t_0)}}{(a_i a_j \min(t_0, r_i^2) \min(t_0, r_j^2))^{n/4}}. \]
In view of
\[ \varrho(x, y)^2/t = 1 - 1/\delta, \quad D = \delta + 2 \]
we get
\[ \frac{\varrho(x, y)^2}{4t} - \frac{\varrho(x, y)^2}{2Dt} = \frac{\delta \varrho(x, y)^2}{4t} = \frac{\delta(1 - \delta)}{4(\delta + 2)} < \frac{\delta}{4(\delta + 2)} < 1 \]
we have
\[ e^{-\varrho(x,y)^2/(2Dt)} e^{\varrho(x,y)^2/4t} e^{-\varrho(x,y)^2/4t} \leq ee^{-\varrho(x,y)^2/4t}, \]
this completes the proof.

As a first application of the above estimate in combination with Remark 12.4 (apply the above estimate with the family \( B(x, \tilde{r}(x))_{x \in M} \)), we obtain for all \( x, y \in M \) the estimate
\[ t \log p(t, x, y) \leq t \log(C(x, y)t^{-n/2}) + t \log(1 + \varrho(x, y)^2/2)^{n/2} - \varrho(x, y)^2/4, \]
and so (since log growths slower then any polynomial)
\[ \limsup_{t \to 0+} 4t \log p(t, x, y) \leq -\varrho(x, y)^2, \]
which is one half of Varadhan’s famous asymptotic formula [34]
\[ \lim_{t \to 0+} 4t \log p(t, x, y) = -\varrho(x, y)^2, \]
The latter formula states that the heat kernel captures a Euclidean behaviour for small times, regardless of the geometry. A very noneuclidean behaviour may occur for large times, though.

Here comes the main result of this lecture course:

**Theorem 12.8** (Grigor’yvan 1994). **Assume** \( M \) **is geodesically complete and noncompact.** Then the following statements are equivalent:

a) \( M \) satisfies a relative Faber-Krahn inequality, that is, there exist constants \( b, n' > 0 \) such that for all \( x \in M, r > 0 \) and all open relatively compact \( U \subset B(x, r) \) one has

\[ \lambda_{\min}(U) \geq \frac{b}{r^2} \left( \frac{\mu(x, r)}{\mu(U)} \right)^{2/n'} \]

(60)
b) $M$ is doubling, that is, there exists a constant $C > 0$ such that for all $x \in M$, $r > 0$ one has
\begin{equation}
\mu(x, 2r) := \mu(B(x, 2r)) \leq C \mu(x, r),
\end{equation}
and $M$ satisfies the following Li-Yau upper heat kernel bound: there exist a constants $n, C' > 0$ such that for all $t > 0$, $x, y \in M$ one has
\begin{equation}
p(t, x, y) \leq C' \left(1 + \varrho(x, y)^2 / t\right)^{n'/2} e^{-\varrho(x, y)^2 / (4t)} \sqrt{\mu(x, \sqrt{t}) \mu(y, \sqrt{t})}.
\end{equation}

c) $M$ is doubling and satisfies the following on-diagonal upper heat kernel bound: there exists a constant $C'' > 0$ such that for all $t > 0$, $x \in M$ one has
\begin{equation}
p(t, x, x) \leq C'' \mu(x, \sqrt{t}).
\end{equation}

Remark 12.9. 1. The noncompactness is only used in c) $\implies$ a).

2. The proof shows that the constants are related as follows:

a) implies b) with $n = n'$ and $C', C$ depending only on $n'$ and $b$.

b) implies c) with $C'' = C$ (trivial).

c) implies a) with $b$ depending only the doubling constant $C$ and $C''$, and $n' = \log_2(C)$.

3. Using that for all $r, \varepsilon > 0$ we can find a constant $C_{r, \varepsilon} > 0$ such that for all $\zeta \geq 0$ one has
\begin{equation}
(1 + \zeta)^r e^{-\zeta/4} \leq C_{r, \varepsilon} e^{-\zeta/(4 + \varepsilon)},
\end{equation}
one immediately gets that b) is equivalent to the following condition: $M$ is volume doubling and there exists a constant $n > 0$ and for all $\varepsilon > 0$ a constant $C_{\varepsilon, n} > 0$ such that for all $t > 0$, $x, y \in M$ one has
\begin{equation}
p(t, x, y) \leq C_{\varepsilon, n} e^{-\varrho(x, y)^2 / ((4+\varepsilon)t)} \sqrt{\mu(x, \sqrt{t}) \mu(y, \sqrt{t})}.
\end{equation}

4. As arguments from the proof entail, b) is also equivalent to the following condition: $M$ is volume doubling and there exists a constant $n > 0$ and for all $\varepsilon > 0$ a constant $A_{\varepsilon, n, C} > 0$, where $C$ is the doubling constant, such that for all $t > 0$, $x, y \in M$ one has
\begin{equation}
p(t, x, y) \leq A_{\varepsilon, n, C} e^{-\varrho(x, y)^2 / ((4+\varepsilon)t)} \mu(x, \sqrt{t}).
\end{equation}
To see this, apply estimate (64) below to estimate
\[
\mu(x, \sqrt{t}) / \mu(y, \sqrt{t}) \leq \mu(y, \sqrt{t} + \varrho(x, y))/\mu(y, \sqrt{t}) \leq A_C (1 + \varrho(x, y)^2 / t)^{n''/2},
\]
where $n'' = \log_2(C)$ and use again (63) with $r$ depending on $n$ and $n''$. 
5. Note that the Theorem entails that certain heat kernel upper bounds (+ doubling) is stable under quasi-isometry\textsuperscript{11}. This is very surprising, as the heat kernel depends on $\Delta$ which carries derivatives of the metric which can differ as dramatically as we wish for quasi-isometric metrics.

6. One can also prove that if $M$ geodesically complete and doubling with doubling constant $A > 0$, and if there exists a constant $B > 0$ such that for all $t > 0$,

$$p(t, x, x) \leq \frac{B}{\mu(x, \sqrt{t})},$$

then there exists a constant $C > 0$ (which only depends on $A$ and $B$) such that for all $t > 0$ one has the lower bound

$$p(t, x, x) \geq \frac{C}{\mu(x, \sqrt{t})}.$$

This is Theorem 16.6 in Grigor’yan’s book.

Proof of Theorem 12.8: a) $\Rightarrow$ b): Let $x \in M$, $r > 0$ be arbitrary. Then for all open $U \subset B(x, r)$ we have

$$\lambda_{\text{min}}(U) \geq a(x, r)\mu(U)^{-2/n'},$$

where

$$a(x, r) := br^{-2}\mu(x, r)^{2/n}.$$ 

Applying Theorem 12.7 to the family $B(x, \sqrt{t})$, $x \in M$, immediatly proves the heat kernel estimate. That relative Faber-Krahn implies doubling will be an exercise.

b) $\Rightarrow$ c): trivial.

c) $\Rightarrow$ a): Let $C$ be the doubling constant. Iterating the doubling inequality one gets that for all $0 < r \leq R$, $x \in M$ one has

$$\mu(x, R)/\mu(x, r) \leq C(R/r)^{n''},$$

where $n'' := \log_2(C)$. Indeed, $R \leq 2^N r$, where $N$ is the smallest natural number $\geq \log_2(R/r)$, and so $N \leq \log_2(R/r) + 1$ and

$$\mu(x, R)/\mu(x, r) \leq \mu(x, 2^N r)/\mu(x, r) \leq C^N \leq C^{1+\log_2(R/r)} = C(R/r)^{n''}.$$

Fix $x \in M$, $r > 0$, $U \subset B(x, r)$ open. Then for all $t > 0$ one has

$$e^{-\lambda_{\text{min}}(U)t} \leq \text{tr}(e^{-tH''}) = \int_U p^U(t, y, y)d\mu(y) \leq C'' \int_U \mu(y, \sqrt{t})^{-1}d\mu(y).$$

If $y \in U$, $t \leq r^2$, then by (64)

$$\mu(x, r)/\mu(y, \sqrt{t}) \leq \mu(y, 2r)/\mu(y, \sqrt{t}) \leq C(r/\sqrt{t})^{n''},$$

\textsuperscript{11}Two Riemann metrics $g, h$ are called quasi-isometric, if $C_1 h \leq g \leq C_2 h$, which easily shows that the volume measures and the quadratic forms $\int |f|^2d\mu$ are equivalent.
and so
\[ e^{-\lambda_{\min}(U)t} \leq C'' \int_U \mu(y, \sqrt{t})^{-1} d\mu(y) \leq \frac{C''C\mu(U)}{\mu(x, r)} \left( \frac{r}{\sqrt{t}} \right)^{n'} = \frac{C''\mu(U)}{\mu(x, r)} \left( \frac{r}{\sqrt{t}} \right)^{n'}, \]
yielding
\[ \lambda_{\min}(U) \geq -t^{-1} \log \left( \frac{C''\mu(U)}{\mu(x, r)} \left( \frac{r}{\sqrt{t}} \right)^{n'} \right). \]  

Case: \( \mu(U) \leq (C''e)^{-1}\mu(x, r) \). Define \( t \) by
\[ (r/\sqrt{t})^{n'} = \frac{\mu(x, r)}{C''e\mu(U)}, \]
so
\[ 1/t = \frac{1}{r^2} \left( \frac{\mu(x, r)}{C''e\mu(U)} \right)^{2/n'}. \]
Then we have \( t \leq r^2 \) and so by (68),
\[ \lambda_{\min}(U) \geq -t^{-1} \log (e^{-1}) = t^{-1} = \frac{1}{r^2} \left( \frac{\mu(x, r)}{C''e\mu(U)} \right)^{2/n'}. \]
Proving relative Faber-Krahn in this case.

Case \( \mu(U) > (C''e)^{-1}\mu(x, r) \): using that balls are relatively compact, \( M \) is connected and noncompact one proves (exercise) that doubling implies reverse doubling: there exists \( n'', c \) (which only depend on the doubling constant \( C \)) such that for all \( y \in M, 0 < s \leq S \) one has
\[ \frac{\mu(y, S)}{\mu(y, s)} \geq c(S/s)^{n''}. \]
This reverse doubling implies that we can pick a constant \( A > 1 \), which only depends on \( n'', c \), such that
\[ \frac{\mu(x, Ar)}{\mu(x, r)} \geq C''e. \]
Then we have \( U \subset B(x, Ar) \) and \( \mu(U) \leq (C''e)^{-1}\mu(x, Ar) \) and the previous case applied to \( Ar \) implies the first inequality in
\[ \lambda_{\min}(U) \geq \frac{1}{(Ar)^2} \left( \frac{\mu(x, Ar)}{C''e\mu(U)} \right)^{2/n'} \geq \frac{1}{(Ar)^2} \left( \frac{\mu(x, r)}{C''e\mu(U)} \right)^{2/n'}, \]
completing the proof.

Recall that the *Levi-Civita connection* on \( M \) is the uniquely determined smooth metric covariant derivative \( \nabla \) on \( TM \) which is torsion free, in the sense that
\[ \nabla_A B - \nabla_B A = [A, B] \] for all smooth vector fields smooth \( A, B \) on \( M \).
The Riemannian curvature tensor $\text{Riem}$ is defined to be the curvature of $\nabla$,

$$\text{Riem} := R_{\nabla} \in \Gamma_{C^\infty}(M, \wedge^2 T^* M \otimes \text{End}(TM)),$$

We recall that for smooth vector fields $A, B, C$ on $M$, this tensor is explicitly given by

$$\text{Riem}(A, B)C := \nabla_A \nabla_B C - \nabla_B \nabla_A C - \nabla_{[A, B]} C \in \mathfrak{X}_{C^\infty}(M).$$

Then the Ricci curvature

$$\text{Ric} \in \Gamma_{C^\infty}(M, T^* M \otimes T^* M)$$

is the field of symmetric bilinear forms on $TM$ given by the fiberwise $(g)$-trace

$$\text{Ric}(A, B) |_U = \sum_{j=1}^m (\text{Riem}(e_j, B)A, e_j),$$

where $e_1, \ldots, e_m \in \mathfrak{X}_{C^\infty}(U)$ is a local orthonormal frame, and $A, B$ are as above. The condition $\text{Ric} \geq \kappa$ for some constant $\kappa \in \mathbb{R}$ means that for all $A$ as above one has

$$\text{Ric}(A, A) \geq \kappa |A|^2 \quad \text{on } M.$$

Clearly the Ricci curvature of the Euclidean $\mathbb{R}^m$ is zero, and compact $M$'s have $\text{Ric} \geq \kappa$ for some $\kappa \in \mathbb{R}$. On the other hand, if $M$ is complete and has $\text{Ric} \geq \kappa > 0$, then $M$ is compact (by the Bonnet-Meyers theorem). The hyperbolic space $\mathbb{H}^m$ has constant Ricci curvature $-1$, in the sense that $\text{Ric}(A, B) = -\langle A, B \rangle$ for all $A, B$ as above.

**Theorem 12.10** (Grigor’yan 1986). If $M$ is geodesically complete with $\text{Ric} \geq 0$, the $M$ satisfies the relative Faber-Krahn inequality. In particular, Theorem 12.8 applies.

The ultimate reason for the above Theorem is that geod. compl. and nonnegative Ricci imply the Laplacian comparison theorem, which states that for all fixed $\mathcal{O} \in M$, one has

$$\Delta g(\cdot, \mathcal{O}) \leq (m - 1)/g(\cdot, \mathcal{O}),$$

wherever $g(\cdot, \mathcal{O})$ is smooth (away from union of $\{\mathcal{O}\}$ and the cut-locus of $\mathcal{O}$).

13. Wiener measure and Brownian motion on Riemannian manifolds

Roughly speaking, one would like to construct Brownian motion $X(x_0)$ on $M$, starting from $x_0 \in M$, as follows: It should be an $M$-valued process$^{12}$ with continuous paths

$$X(x_0) : [0, \infty) \times \Omega \longrightarrow M,$$

which is defined on some probability space $(\Omega, \mathbb{P}, \mathcal{F})$, and which has the transition probability densities given by $p(t, x, y)$. In other words, given $n \in \mathbb{N}$, a finite sequence of times $0 < t_1 < \cdots < t_n$ and Borel sets $A_1, \ldots, A_n \subset M$, setting $\delta_j := t_{j+1} - t_j$ with $t_0 := 0$, $t_n := \infty$, $t_{n+1} := 0$, $\sum_{j=1}^n \delta_j = 1$.

$^{12}$We recall that given two measurable spaces $\Omega_1$ and $\Omega_2$, a map

$$X : [0, \infty) \times \Omega_1 \longrightarrow \Omega_2, \quad (t, \omega) \longmapsto X_t(\omega)$$

is called an $\Omega_2$-valued process, if for all $t \geq 0$ the induced map $X_t : \Omega_1 \rightarrow \Omega_2$ is measurable. The maps $t \mapsto X_t(\omega)$, with fixed $\omega \in \Omega_1$, are referred to as the paths of $X$. 

we would like the probability of finding the Brownian particle simultaneously in $A_1$ at the time $t_1$, in $A_2$ at the time $t_2$, and so on, to be given by the quantity
\begin{equation}
\mathbb{P}\{X_{t_1}(x_0) \in A_1, \ldots, X_{t_n}(x_0) \in A_n\}
= \int \cdots \int 1_{A_1}(x_1)p(\delta_0, x_0, x_1) \cdots \\
\times 1_{A_n}(x_n)p(\delta_{n-1}, x_{n-1}, x_n)d\mu(x_1) \cdots d\mu(x_n),
\end{equation}
whenever the particle starts from $x_0$. Equivalently, one could say that a Brownian motion on $M$ with starting point $x_0$ is a process with continuous paths \eqref{eqn:process}, such that the finite-dimensional distributions of its law are given by the right-hand side of \eqref{eqn:probability}. In fact, such a path space measure is uniquely determined by its finite-dimensional distributions (cf. Remark 13.7 below). In particular, all Brownian motions should have the same law, which we will call the Wiener measure later on.

Ultimately, the above prescriptions indeed turn out to work perfectly well in terms of giving Brownian motion for the Euclidean $\mathbb{R}^m$ or for compact Riemannian manifolds. On the other hand, we see from \eqref{eqn:probability} that, in particular, it is required that for all $t > 0$,
\begin{equation}
\mathbb{P}\{X_t(x_0) \in M\} = \int_M p(t, x_0, y)d\mu(y),
\end{equation}
and already if $M$ is any open bounded subset of $\mathbb{R}^m$, it automatically happens that
\begin{equation}
\int_M p(t, x_0, y)d\mu(y) < 1 \text{ for some } (t, x_0) \in (0, \infty) \times M,
\end{equation}
This leads to the conceptual difficulty that the process can leave its space of states with a strictly positive probability and leads to:

**Definition 13.1.** $M$ is called *stochastically complete*, if for all $t > 0$, $x_0 \in M$ one has
\[ \int_M p(t, x_0, y)d\mu(y) = 1. \]

**Remark 13.2.** Stochastic completeness is unrelated to geodesic completeness. For example, $\mathbb{R}^m \setminus \{0\}$ is stochastically complete but geodesically incomplete, and there exist geodesically complete and but stochastically incomplete $M$’s. On the other hand, a celebrated result by Yau states that if $M$ is geodesically complete with a Ricci curvature bounded from below by a constant, then $M$ is stochastically complete. In particular, the Euclidean $\mathbb{R}^m$ is stochastically complete (of course this follows also simply from a calculation), and compact $M$’s are stochastically complete.

---

The law of $X(x_0)$ is by definition the probability measure on the space of continuous paths on $M$, which is defined as the pushforward of $\mathbb{P}$ under the induced map
\[ \Omega \rightarrow C([0, \infty), M), \omega \mapsto X_\bullet(x_0)(\omega). \]
Since we aim to work on arbitrary Riemannian manifolds, we need to solve the above conceptual problem of stochastic incompleteness. This is done by using the Alexandrov compactification of $M$. Since it does not cause much extra work, we start by explaining the corresponding constructions in the setting of an arbitrary Polish space, recalling that a topological space is called Polish, if it is separable and if it admits a complete metric which induces the original topology.

**Notation 13.3.** Given a locally compact Polish space $N$, we set

$$
\tilde{N} := \begin{cases} 
N, & \text{if } N \text{ is compact} \\
N \cup \{\infty_N\}, & \text{if } N \text{ is noncompact.}
\end{cases}
$$

We recall here that $\infty_N$ is any point $\notin N$, and that the topology on $N \cup \{\infty_N\}$ is defined as follows: $U \subset N \cup \{\infty_N\}$ is declared to be open, if and only if either $U$ is an open subset of $N$ or if there exists a compact set $K \subset N$ such that $U = (N \setminus K) \cup \{\infty_N\}$. This construction depends trivially on the choice of $\infty_N$, in the sense that for any other choice $\infty'_N \notin N$, the canonical bijection $N \cup \{\infty_N\} \to N \cup \{\infty'_N\}$ is a homeomorphism.

We consider the path space $\Omega_N := C([0, \infty), \tilde{N})$, and thereon we denote (with a slight abuse of notation) the canonically given coordinate process by

$$
X : [0, \infty) \times \Omega_N \longrightarrow \tilde{N}, \quad X_t(\gamma) := \gamma(t).
$$

We consider $\Omega_N$ a topological space with respect to the topology of uniform convergence on compact subsets, and we equip it with its Borel sigma-algebra $\mathcal{F}^N$.

We fix such a locally compact Polish space $N$ (e.g., a manifold) for the moment. It is well-known that $\Omega_N$ as defined above is Polish again. In fact, $\tilde{N}$ is Polish, and if we pick a bounded metric $\varrho_{\tilde{N}} : \tilde{N} \times \tilde{N} \to [0,1]$ which induces the original topology on $\tilde{N}$, then

$$
\varrho_{\Omega_N}(\gamma_1, \gamma_2) := \sum_{j=1}^{\infty} \max_{0 \leq t \leq j} \varrho_{\tilde{N}}(\gamma_1(t), \gamma_2(t))
$$

is a complete separable metric\(^{14}\) on $\Omega_N$ which induces the original topology (of local uniform convergence). Furthermore, since evaluation maps of the form

$$
X_1 \times C(X_1, X_2) \longrightarrow X_2, \quad (x, f) \longmapsto f(x)
$$

are always jointly continuous, if $X_1$ is locally compact and Hausdorff and if $C(X_1, X_2)$ is equipped with its topology of local uniform convergence, it follows that $X$ is in fact jointly continuous. In particular, $X$ is jointly (Borel) measurable.

\(^{14}\)In fact, it is easy to see that this is a complete metric which induces the original topology. On the other hand, the proof that this topology is separable is a little tricky. Although it is not so easy to find a precise reference, we believe that these results can be traced back to Kolmogorov.
**Notation 13.4.** Given a set $\Omega$ and a collection $\mathcal{C}$ of subsets of $\Omega$ or of maps with domain $\Omega$, the symbol $\langle \mathcal{C} \rangle$ stands for the smallest sigma-algebra on $\Omega$ which contains $\mathcal{C}$. Furthermore, whenever there is no danger of confusion, we will use notations such as

$$\{ f \in A \} := \{ y \in \Omega : f(y) \in A \} \subset \Omega,$$

where $f : \Omega \to \Omega'$ and $A \subset \Omega'$.

**Definition 13.5.**

1. A subset $C \subset \Omega_N$ is called a **Borel cylinder**, if there exist $n \in \mathbb{N}$, $0 < t_1 < \cdots < t_n$ and Borel sets $A_1, \ldots, A_n \subset \tilde{N}$, such that

$$C = \{ X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n \} = \bigcap_{j=1}^n X_{t_j}^{-1}(A_j).$$

The collection of all Borel cylinders in $\Omega_N$ will be denoted by $\mathcal{C}_N^N$.

2. Likewise, given $t \geq 0$, the collection $\mathcal{C}_t^N$ of **Borel cylinders in $\Omega_N$ up to the time $t$** is defined to be the collection of subsets $C \subset \Omega_N$ of the form

$$C = \{ X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n \} = \bigcap_{j=1}^n X_{t_j}^{-1}(A_j),$$

where $n \in \mathbb{N}$, $0 < t_1 < \cdots < t_n < t$, and where $A_1, \ldots, A_n \subset \tilde{N}$ are Borel sets.

It is easily checked inductively that both $\mathcal{C}_N^N$ and $\mathcal{C}_t^N$ are $\pi$-systems in $\Omega_N$, that is, both collections are (nonempty and) stable under taking finitely many intersections. The following fact makes $\mathcal{F}_N$ handy in applications:

**Lemma 13.6.** One has

$$\mathcal{F}_N = \langle \mathcal{C}_N^N \rangle = \langle (X_s : \Omega_N \to \tilde{N})_{s \geq 0} \rangle.$$

**Proof:** Since for every fixed $s \geq 0$ the map

$$X_s : \Omega_N \to \tilde{N}, \gamma \mapsto \gamma(s)$$

is $\mathcal{F}_N$-measurable, it is clear that $\mathcal{C}_N^N \subset \mathcal{F}_N$, and therefore

$$\langle \mathcal{C}_N^N \rangle \subset \mathcal{F}_N.$$

In order to see

$$\mathcal{F}_N \subset \langle \mathcal{C}_N^N \rangle,$$

pick a topology-defining metric $\varrho_{\tilde{N}}$ on $\tilde{N}$ and denote the corresponding closed balls by $B_{\tilde{N}}(x, r)$. Then, since the elements of $\Omega_N$ are continuous, for all $\gamma_0 \in \Omega_N$, $n \in \mathbb{N}$, $\epsilon > 0$
one has
\[
\left\{ \gamma : \max_{0 \leq t \leq n} \varrho_N(\gamma(t), \gamma_0(t)) \leq \epsilon \right\} = \bigcap_{0 \leq t \leq n, \ t \ \text{is rational}} \left\{ \gamma : \gamma(t) \in B_N(\gamma_0(t), \epsilon) \right\},
\]
\[
= \bigcap_{0 < t \leq n, \ t \ \text{is rational}} \left\{ \gamma : \gamma(t) \in B_N(\gamma_0(t), \epsilon) \right\}.
\]
Therefore, sets of the form
\[
\left\{ \gamma : \max_{0 \leq t \leq n} \varrho_N(\gamma(t), \gamma_0(t)) \leq \epsilon \right\}, \ \gamma_0 \in \Omega_N, \ n \in \mathbb{N}, \ \epsilon > 0
\]
are \(\langle \mathcal{C}^N \rangle\)-measurable. Since the collection of sets of the form (73) generates the topology of local uniform convergence\(^{15}\), it is clear that the induced Borel sigma-algebra \(\mathcal{F}^N\) satisfies \(\mathcal{F}^N \subset \langle \mathcal{C}^N \rangle\).

The inclusion
\[
\langle \mathcal{C}^N \rangle \subset \langle (X_s : \Omega_N \rightarrow \tilde{N})_{s \geq 0} \rangle
\]
is clear, since each set in \(\mathcal{C}^N\) is a finite intersection of sets of the form \(X_s^{-1}(A), \ s > 0, \ A \subset \tilde{N}\) Borel. To see
\[
\langle (X_s : \Omega_N \rightarrow \tilde{N})_{s \geq 0} \rangle \subset \langle \mathcal{C}^N \rangle,
\]
note that for every metric \(\varrho_N\) that generates the topology on \(\tilde{N}\), one has
\[
\langle (X_s : \Omega_N \rightarrow \tilde{N})_{s \geq 0} \rangle = \left\{ X_s^{-1}(B_N(x, r)) : x \in \tilde{N}, r > 0, s \geq 0 \right\}
\]
with the corresponding closed balls \(B_N(\ldots)\), so that it only remains to prove
\[
X_0^{-1}(B_N(x, r)) \in \langle \mathcal{C}^N \rangle
\]
for all \(x \in \tilde{N}, \ r > 0\). This, however, follows from
\[
X_0^{-1}(B_N(x, r)) = \left\{ \gamma : \lim_{n \to \infty} \varrho_N(\gamma(1/n), x) \leq r \right\},
\]
since clearly \(\gamma \mapsto \varrho_N(\gamma(1/n), x)\) is a \(\langle \mathcal{C}^N \rangle\)-measurable function on \(\Omega_N\) (the pre-image of an interval of the form \([0, R]\) under this map is the cylinder set \(X_{1/n}^{-1}(B_N(x, R))\)). This completes the proof. \(\blacksquare\)

\textbf{Remark 13.7.} By the above lemma, \(\mathcal{C}^N\) is a \(\pi\)-system that generates \(\mathcal{F}^N\). It then follows from an abstract measure theoretic result that every finite measure on \(\mathcal{F}^N\) is uniquely determined by its values on \(\mathcal{C}^N\).

\(^{15}\)To be precise, this collection forms a basis of neighbourhoods of this topology.
Definition 13.8. Setting

\[ \mathcal{F}_N^t := \mathcal{F}(\mathcal{X}_s: \Omega_N \to \tilde{N})_{0 \leq s \leq t} \]

for every \( t \geq 0 \), it follows from Lemma 13.6 that

\[ \mathcal{F}_N^N := \bigcup_{t \geq 0} \mathcal{F}_N^t \]

becomes a filtration of \( \mathcal{F}^N \). It is called the filtration generated by the coordinate process on \( \Omega_N \).

Precisely as for the second equality in (72), one proves

\[ \mathcal{F}_t^N = \langle \mathcal{C}_i^N \rangle \]

for all \( t \geq 0 \).

Particularly important \( \mathcal{F}_N^t \)-measurable sets are provided by exit times:

Definition 13.9. Given an arbitrary subset \( U \subset \tilde{N} \), we define

\[ \zeta_U : \Omega_N \to [0, \infty], \quad \zeta_U := \inf\{ t \geq 0 : \mathcal{X}_t \in \tilde{N} \setminus U \} \]

and call this map the the first exit time of \( \mathcal{X} \) from \( U \), with \( \inf\{...\} := \infty \) in case the set is empty.

There is the following result, which in a probabilistic language means that first exit times from open sets are \( \mathcal{F}_N^t \)-optional times:

Lemma 13.10. Assume that \( U \subset \tilde{N} \) is open with \( U \neq \tilde{N} \). Then one has

\[ \{ t < \zeta_U \} \in \mathcal{F}_t^N \]

for all \( t \geq 0 \).

Proof: The proof actually only uses that \( \mathcal{X} \) has continuous paths and that \( \tilde{N} \) is metrizable:

Pick a metric \( \varrho_{\tilde{N}} \) on \( \tilde{N} \) which induces the original topology. Then, since \( \tilde{N} \setminus U \) is closed and \( \mathcal{X} \) has continuous paths, we have

\[ \{ t < \zeta_U \} = \bigcup_{n \in \mathbb{N}} \bigcup_{0 \leq s \leq t, s \text{ is rational}} \{ \varrho_{\tilde{N}}(\mathcal{X}_s, \tilde{N} \setminus U) \geq 1/n \} \]

The set on the right-hand side clearly is \( \in \mathcal{F}_t^N \), since the distance function to a nonempty set is continuous and thus Borel. ■

14. The Wiener measure on Riemannian manifolds

We return to our Riemannian setting. In order to apply the above abstract machinery in this case, we have to extend some Riemannian data to the compactification of \( M \) (in the noncompact case):
**Notation 14.1.** Let \( \tilde{\mu} \) denote the Borel measure on \( \tilde{M} \) given by \( \mu \) if \( M \) is compact, and which is extended to \( \infty_M \) by setting \( \mu(\infty_M) = 1 \) in the noncompact case. Then we define a Borel function
\[
\tilde{p} : (0, \infty) \times \tilde{M} \times \tilde{M} \to [0, \infty)
\]
as follows: \( \tilde{p} := p \) if \( M \) is compact, and in case \( M \) is noncompact, then for \( t > 0, x, y \in M \) we set
\[
\tilde{p}(t, x, y) := p(t, x, y), \quad \tilde{p}(t, x, \infty_M) := 0, \quad \tilde{p}(t, \infty_M, \infty_M) := 1, \\
\tilde{p}(t, \infty_M, y) := 1 - \int_M p(t, y, z) d\mu(z).
\]
It is straightforward to check that the pair \((\tilde{p}, \tilde{\mu})\) satisfies the Chapman-Kolmogorov equations, that is, for all \( s, t > 0, x, y \in \tilde{M} \) one has
\[
\int_{\tilde{M}} \tilde{p}(t, x, z) \tilde{p}(s, y, z) d\tilde{\mu}(z) = \tilde{p}(s + t, x, y).
\]
Furthermore, one has
\[
\int_{\tilde{M}} \tilde{p}(t, x, y) d\tilde{\mu}(y) = 1 \quad \text{for all} \ x \in \tilde{M},
\]
in contrast to the possibility of \( \int_M p(t, x, y) d\mu(y) < 1 \) in case \( M \) is stochastically incomplete.

It is precisely the conservation of probability (77) which motivates the above Alexandrov machinery. Since there is no danger of confusion, the following abuse of notation will be very convenient in the sequel:

**Notation 14.2.** We write \( \zeta := \zeta_M \) for the first exist time of the coordinate process \( \mathbb{X} \) on \( \Omega_M \) from \( M \subset \tilde{M} \).

For obvious reasons, \( \zeta \) is also called the explosion time of \( \mathbb{X} \). Note also that one has \( \zeta > 0 \), and that by our previous conventions we have \( \zeta \equiv \infty \) if \( M \) is compact. The last fact is consistent with the fact that compact Riemannian manifolds are stochastically complete.

The following existence result will be central in the sequel:

**Proposition and definition 14.3.** The Wiener measure \( \mathbb{P}^{x_0} \) with initial point \( x_0 \in M \) is defined to be the unique probability measure on \((\Omega_M, \mathcal{F}^M)\) which satisfies
\[
\mathbb{P}^{x_0} \{ \mathbb{X}_{t_1} \in A_1, \ldots, \mathbb{X}_{t_n} \in A_n \}
= \int \cdots \int 1_{A_1}(x_1) \tilde{p}(\delta_0, x_0, x_1) \cdots \\
\times 1_{A_n}(x_n) \tilde{p}(\delta_{n-1}, x_{n-1}, x_n) \tilde{\mu}(x_1) \cdots \tilde{\mu}(x_n)
\]
for all \( n \in \mathbb{N} \), all finite sequences of times \( 0 < t_1 < \cdots < t_n \) and all Borel sets \( A_1, \ldots, A_n \subset \tilde{M} \), where \( \delta_j := t_{j+1} - t_j \) with \( t_0 := 0 \). It has the additional property that
\[
\mathbb{P}^{x_0} \left( \{ \zeta = \infty \} \cup \left\{ \zeta < \infty \text{ and } \mathbb{X}_t = \infty_M \text{ for all } t \in [\zeta, \infty) \} \right) = 1,
\]
in other words, the point at infinity $\infty_M$ is a “trap” for $\mathbb{P}^{x_0}$-a.e. path.\footnote{It is a trap in the sense that once a path touches $\infty_M$, it remains there for all times.}

Proof: Some remarks during class.

An obvious but nevertheless very important consequence of (78) is that for all $x_0 \in M$ one has
\begin{equation}
\mathbb{P}^{x_0}\{1_{\{t<\zeta\}} = 1_{\{X_t \in M\}}\} = 1.
\end{equation}

In the sequel, integration with respect to the Wiener measure will often be written as an expectation value,
\[ E^{x_0}[\Psi] := \int \Psi d\mathbb{P}^{x_0} := \int \Psi(\gamma) d\mathbb{P}^{x_0}(\gamma), \]
where $\Psi : \Omega_M \to \mathbb{C}$ is any appropriate (say, nonnegative or integrable) Borel function.

We remark that using monotone convergence, the defining relation of the Wiener measure implies that for all $n \in \mathbb{N}$, all finite sequences of times $0 < t_1 < \cdots < t_n$ and all Borel functions $f_1, \ldots, f_n : \tilde{M} \to [0, \infty)$, one has
\begin{align}
E^{x_0} &\left[ f_1(X_{t_1}) \cdots f_n(X_{t_n}) \right] \\
&= \int \cdots \int f_1(x_1) p(\delta_0, x_0, x_1) \cdots \\
&\quad \times f_n(x_n) p(\delta_{n-1}, x_{n-1}, x_n) d\mu(x_1) \cdots d\mu(x_n),
\end{align}
where $\delta_j := t_{j+1} - t_j$ with $t_0 := 0$. In particular, by the very construction of $\tilde{M}$ and $\bar{\mu}$, the above formula in combination with (79) implies
\begin{align}
E^{x_0} &\left[ 1_{\{t_1<\zeta\}} f_1(X_{t_1}) \cdots 1_{\{t_n<\zeta\}} f_n(X_{t_n}) \right] \\
&= E^{x_0} \left[ 1_{\{X_{t_1} \in M\}} f_1(X_{t_1}) \cdots 1_{\{X_{t_n} \in M\}} f_n(X_{t_n}) \right] \\
&= \int \cdots \int f_1(x_1) p(\delta_0, x_0, x_1) \cdots \\
&\quad \times f_n(x_n) p(\delta_{n-1}, x_{n-1}, x_n) d\mu(x_1) \cdots d\mu(x_n),
\end{align}
therefore quantities that are given by averaging over paths that remain on $M$ until any fixed time can be calculated by genuine Riemannian data on $M$, as it should be. In the sequel, we will also freely use the following facts:

Remark 14.4. 1. Each of the measures $\mathbb{P}^{x_0}$ is concentrated on the set of paths that start in $x_0$, meaning that
\[ \mathbb{P}^{x_0}\{X_0 = x_0\} = 1 \quad \text{for all } x_0 \in M, \]
as it should be. To see this, pick a metric $\tilde{\rho}$ on $\tilde{M}$ which induces the topology on $\tilde{M}$, and set
\[ \tilde{f} := \tilde{\rho}(\bullet, x_0) - \tilde{\rho}(\infty_M, x_0) \in C(\tilde{M}). \]
As $x_0 \in M$, the very definition of $(\tilde{p}, \tilde{\mu})$ implies that for all $t > 0$ one has
\[
\int_M \tilde{p}(t, x_0, y) \tilde{\varrho}(y, x_0) d\tilde{\mu}(y) = \int_M p(t, x_0, y) f|_M(y) d\mu(y) + \tilde{\varrho}(\infty, x_0),
\]
which, since $\tilde{f}|_M$ is a continuous bounded function on $M$, implies through (80) and the fact that for all $f \in C_b(M)$ one has
\[
P_t f \to f \text{ locally uniformly as } t \to 0^+,
\]
the $L^1$-convergence
\[
\mathbb{E}^{x_0}[\tilde{\varrho}(X_t, x_0)] = \int_M \tilde{p}(t, x_0, y) \tilde{\varrho}(y, x_0) d\tilde{\mu}(y) \to 0 \text{ as } t \to 0^+.
\]
Thus we can pick a sequence of strictly positive times $a_n$ with $a_n \to 0$ such that $\tilde{\varrho}(X_{a_n}, x) \to 0$ $\mathbb{P}^{x_0}$-a.e., and the claim follows from
\[
\tilde{\varrho}(x_0, x) \leq \tilde{\varrho}(X_0, X_{a_n}) + \tilde{\varrho}(X_{a_n}, x) \text{ for all } n \in \mathbb{N}
\]
and the continuity of the paths of $X$.

2. For every Borel set $N \subset M$ with $\mu(N) = 0$ and every $x \in M$, one has
\[
(84) \quad \int_0^\infty \int_{\Omega_M} 1_{\{s', \gamma': \gamma'(s') \in N\}}(s, \gamma) d\mathbb{P}^x(\gamma) ds = \int_0^\infty \int_N p(s, x, y) d\mu(y) ds = 0.
\]
This fact follows immediately from the defining relation of the Wiener measure. For the first identity in (84), one also needs Fubini’s Theorem, which can be used due to $X$ being jointly measurable.

3. For each fixed $A \in \mathcal{F}^M$, the map
\[
(85) \quad M \longrightarrow [0, 1], \ \ x \mapsto \mathbb{P}^x(A)
\]
is Borel measurable. In fact, this is obvious for $A \in \mathcal{G}^M$ by the defining relation of the Wiener measure, and it holds in general by the monotone class theorem, since $\mathcal{G}^M$ is a $\pi$-system which generates $\mathcal{F}^M$, and since the collection of sets
\[
\{ A : A \in \mathcal{F}^M, \ (85) \text{ is Borel measurable} \}
\]
forms a monotone Dynkin-system.

The following result is crucial:

**Lemma 14.5.** The family of Wiener measures satisfies the following Markov property: For all $x_0 \in M$, all times $t \geq 0$, all $\mathcal{F}^M$-measurable functions $\phi : \Omega_M \to [0, \infty)$, and all $\mathcal{F}^M$-measurable functions $\Psi : \Omega_M \to [0, \infty)$, one has
\[
(86) \quad \int \phi(\gamma) \Psi(\gamma(t + \bullet)) d\mathbb{P}^{x_0}(\gamma) = \int \phi(\gamma) \int \Psi(\omega) d\mathbb{P}^t(\omega) d\mathbb{P}^{x_0}(\gamma) \in [0, \infty].
\]
Proof: By monotone convergence, it is sufficient to consider the case \( \phi = 1_A, \Psi = 1_B \) with \( A \in \mathcal{F}_t^M, B \in \mathcal{F}_M \). Furthermore, for fixed \( A \in \mathcal{F}_t^M \), using a monotone class argument as in Remark 14.4.3, it follows that it is sufficient to prove the formula for \( B \in \mathcal{C}^M \). Using yet another monotone class argument, it follows that ultimately we have to check the formula only for \( \phi = 1_A, \Psi = 1_B \) with \( A \in \mathcal{C}_t^M, B \in \mathcal{C}_M \). So we pick \( k, l \in \mathbb{N} \), finite sequences of times \( 0 < r_1 < \cdots < r_k < t, 0 < s_1 < \cdots < s_l \), Borel sets

\[
A_1, \ldots, A_k, B_1, \ldots, B_l \subset \tilde{M}
\]

with

\[
A = \bigcap_{i=1}^k X^{-1}_{r_i}(A_i), \quad B = \bigcap_{i=1}^l X^{-1}_{s_i}(B_i),
\]

and \( s_0 := 0, r_0 := 0 \). Then by the defining relation of the Wiener measure we have

\[
\int 1_A(\gamma) \cdot 1_B(\gamma(\tau + \bullet))d\mathbb{P}^{x_0}(\gamma)
\]

\[
= \int 1_{\{x_1 \in A_1\}} \cdots 1_{\{x_k \in A_k\}} 1_{\{x_{s_1} \in B_1\}} \cdots 1_{\{x_{s_l} \in B_l\}}d\mathbb{P}^{x_0}
\]

\[
= \int \cdots \int 1_{A_1}(x_1)\eta(r_1 - r_0, x_0, x_1) \cdots 1_{A_k}(x_k)\eta(r_k - r_{k-1}, x_{k-1}, x_k)
\]

\[
\times 1_{B_1}(x_{k+1})\eta(s_1 + t - r_k, x_k, x_{k+1}) \cdots
\]

\[
\times 1_{B_l}(x_{l+1})\eta(s_l - s_{l-1}, x_{l+1}, x_{k+l})d\mu(x_1) \cdots d\mu(x_{l+1}).
\]

On the other hand, if for every \( y_0 \in \tilde{M} \) we set

\[
\Psi(y_0) := \int \cdots \int 1_{B_1}(y_1)\eta(s_1 - s_0, y_0, y_1) \cdots
\]

\[
\times 1_{B_l}(y_l)\eta(s_l - s_{l-1}, y_{l-1}, y_l)d\mu(y_1) \cdots d\mu(y_l),
\]

then by using the defining relation of the Wiener measure for the \( d\mathbb{P}^{t(\omega)}(\omega) \) integration and then using (80), we get

\[
\int 1_A(\gamma) \int 1_B(\omega)d\mathbb{P}^{t(\omega)}(\omega)d\mathbb{P}^{x_0}(\gamma)
\]

\[
= \int 1_{\{x_1 \in A_1\}}(\gamma) \cdots 1_{\{x_k \in A_k\}}(\gamma)\Psi(\gamma(t))d\mathbb{P}^{x_0}(\gamma)
\]

\[
= \int \cdots \int 1_{A_1}(z_1)\eta(r_1 - r_0, x_0, z_1) \cdots 1_{A_k}(z_k)\eta(r_k - r_{k-1}, x_{k-1}, z_k)
\]

\[
\times \eta(t - r_k, z_k, z)1_{B_1}(y_1)\eta(s_1 - s_0, z, y_1) \cdots 1_{B_l}(y_l)\eta(s_l - s_{l-1}, y_{l-1}, y_l)
\]

\[
\times d\mu(z_1) \cdots d\mu(z_l)d\mu(z)\mu(z),
\]

which is equal to the above expression for

\[
\int 1_A(\gamma) \cdot 1_B(\gamma(\tau + \bullet))d\mathbb{P}^{x_0}(\gamma),
\]
since by the Chapman-Kolomogorov equation and recalling \( s_0 = 0 \), we have
\[
\int \int \tilde{p}(t - r_k, z, y)1_{B_1}(y_1)\tilde{p}(s_1 - s_0, z, y_1)d\tilde{\mu}(z)d\tilde{\mu}(y_1)
= \int \tilde{p}(t - r_k + s_1, z, y_1)1_{B_1}(y_1)d\tilde{\mu}(y_1).
\]
This completes the proof. ■

Now we are in the position to define Brownian motion on an arbitrary Riemannian manifold:

**Definition 14.6.** 1. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space, \(x_0 \in M\), and let
\[
X(x_0) : [0, \infty) \times \Omega \rightarrow \tilde{M}, \quad (t, \omega) \mapsto X_t(x_0)(\omega)
\]
be a continuous process. Then the tuple \((\Omega, \mathcal{F}, \mathbb{P}, X(x_0))\) is called a *Brownian motion on M with starting point* \(x_0\), if the law of \(X(x_0)\) with respect to \(\mathbb{P}\) is equal to the Wiener measure \(\mathbb{P}^{x_0}\). Recall that this means the following: The pushforward of \(\mathbb{P}\) with respect to the \(\mathcal{F}/\mathcal{F}^M\) measurable\(^{18}\) map
\[
(87) \quad \Omega \rightarrow \Omega_M, \quad \omega \mapsto \left( t \mapsto X_t(x_0)(\omega) \right)
\]
is \(\mathbb{P}^{x_0}\).

2. Assume that \((\Omega, \mathcal{F}, \mathbb{P}, X(x_0))\) is a Brownian motion on \(M\) with starting point \(x_0\), and that \(\mathcal{F}_t := (\mathcal{F}_t)_{t \geq 0}\) is a filtration of \(\mathcal{F}\). Then the tuple \((\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}, X(x_0))\) is called an *adapted Brownian motion on M with starting point* \(x_0\), if \(X(x)\) is adapted to \(\mathcal{F}_t := (\mathcal{F}_t)_{t \geq 0}\) (that is, \(X_t(x_0) : \Omega \rightarrow \tilde{M}\) is \(\mathcal{F}_t\)-measurable for all \(t \geq 0\)) and if in addition the following Markov property holds: For all times \(t \geq 0\), all \(\mathcal{F}_t\) measurable functions \(\phi : \Omega \rightarrow [0, \infty)\), and all Borel functions \(\Psi : \Omega_M \rightarrow [0, \infty)\), one has
\[
\int \phi(\omega)\Psi(X_{t+\cdot}(x_0)(\omega))d\mathbb{P}(\omega) = \int \phi(\omega)\int \Psi(\gamma)d\mathbb{P}(\gamma)^{X_t(x_0)}(\omega)d\mathbb{P}(\omega).
\]
It follows from the above results that a canonical adapted Brownian motion with starting point \(x_0\) is given in terms of the Wiener measure by the datum
\[
(88) \quad (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}, X(x_0)) := (\Omega_M, \mathcal{F}^M, \mathcal{F}_t^M, \mathbb{P}^{x_0}, X).
\]
Having recorded the existence of Brownian motion, we can immediately record the following characterization of the stochastic completeness property that was previously defined by the “parabolic condition”
\[
\int_M p(t, x_0, y)d\mu(y) = 1 \quad \text{for all } (t, x_0) \in (0, \infty) \times M:
\]
Namely, \(M\) is stochastically complete, if and only if for every \(x_0 \in M\) and every Brownian motion \((\Omega, \mathcal{F}, \mathbb{P}, X(x_0))\) on \(M\) with starting point \(x_0\), one has
\[
\mathbb{P}\{X_t(x_0) \in M\} = 1 \quad \text{for all } t \geq 0,
\]
\(^{18}\)Note that by assumption \(X_t(x_0)\) is \(\mathcal{F}_t^M\)-measurable for all \(t \geq 0\), so that indeed (87) is automatically \(\mathcal{F}/\mathcal{F}^M\) measurable.
that is, if all Brownian motions remain on $M$ for all times. This observation follows immediately from the defining relation of the Wiener measure. The second part of Definition 14.6 is motivated by the fact that every Brownian motion has the required Markov property with respect to its own filtration:

**Lemma 14.7.** Every Brownian motion $(\Omega, \mathcal{F}, \mathbb{P}, X(x_0))$ on $M$ with starting point $x_0$ is automatically an $(\mathcal{F}^X_t)_{t \geq 0}$-Brownian motion, where

$$\mathcal{F}^X_t := \langle (X_s(x_0))_{0 \leq s \leq t} \rangle, \quad t \geq 0$$

denotes the filtration of $\mathcal{F}$ which is generated by $X(x_0)$.

**Proof:** We have to show that given $t \geq 0$, an $\mathcal{F}^X_t$-measurable function $\phi : \Omega \to [0, \infty)$, and a Borel function $\Psi : \Omega_M \to [0, \infty)$, one has

$$\int \phi(\omega)\Psi(X_{t+\bullet}(x_0)(\omega))d\mathbb{P}(\omega) = \int \phi(\omega) \int \Psi(\gamma)d\mathbb{P}^X(\omega)(\gamma)d\mathbb{P}(\omega).$$

Assume for the moment that we can pick an $\mathcal{F}^M_t$-measurable function $f : \Omega_M \to [0, \infty)$ such that $f(X'(x_0)) = \phi$, where

$$X'(x_0) : \Omega \longrightarrow \Omega_M$$

denotes the induced $\mathcal{F}/\mathcal{F}^M$ measurable map (87). Then, since the law of $X(x_0)$ is $\mathbb{P}^x_0$, we can use the Markov property from Lemma 14.5 to calculate

$$\int \phi(\omega)\Psi(X_{t+\bullet}(x_0)(\omega))d\mathbb{P}(\omega) = \int f(\omega')\Psi(\omega'(t + \bullet))d\mathbb{P}^x_0(\omega')$$

$$= \int f(\omega') \int \Psi(\gamma)d\mathbb{P}^x(\omega')(\gamma)d\mathbb{P}^x_0(\omega')$$

$$= \int f(X(x_0)(\omega)) \int \Psi(\gamma)d\mathbb{P}^X(x_0)(\omega)(\gamma)d\mathbb{P}(\omega)$$

$$= \int \phi(\omega) \int \Psi(\gamma)d\mathbb{P}^X(x_0)(\omega)(\gamma)d\mathbb{P}(\omega),$$

proving the claim in this case. It remains to prove that one can always “factor” $\phi$ in the above form. Somewhat simpler variants of such a statement are usually called Doob-Dynkin lemma in the literature. An important point here is that the factoring procedure can be chosen to be positivity preserving. We give a quick proof: Set $X := X(x_0)$, $X' := X'(x_0)$, and assume first that $\phi$ is a simple function, that is, $\phi$ is a finite sum $\phi = \sum_j c_j1_{A_j}$ with constants $c_j \geq 0$ and disjoint sets $A_j \in \mathcal{F}^X_t$. Then by the definition of this sigma-algebra, there exist times $0 \leq s_j \leq t$ and Borel sets $B_j \subset \tilde{M}$ with $A_j = X^{-1}(s_j)(B_j)$, such that with $C_j := X^{-1}(s_j)(B_j) \in \mathcal{F}^M_t$, the function $f := \sum_j c_j1_{C_j}$ on $\Omega_M$ is nonnegative, $\mathcal{F}^M_t$-measurable, and satisfies $f(X') = \phi$. In the general case, there exists an increasing sequence of nonnegative $\mathcal{F}^X_t$-measurable simple functions $\phi_n$ on $\Omega$ such that $\lim_n \phi_n = \phi$. Heat kernels on Riemannian manifolds 69
By the above, we can pick for each \( n \) an \( \mathscr{F}^M_t \)-measurable nonnegative function \( f_n \) on \( \Omega_M \) with \( f_n(X') = \phi_n \). The set 
\[ \Omega' := \{ f_n \text{ converges pointwise } \} \subset \Omega \]
clearly contains the image of \( X' \), and it is straightforwardly seen to be \( \mathscr{F}^M_t \)-measurable. Then \( f := \lim_n (f_n1_{\Omega'}) \) has the desired properties. Note that the above proof is entirely measure theoretic and does not use any particular (say, topological) properties of the involved quantities.

Without entering the details, we remark here that the importance of adapted Brownian motions stems from the fact that they are continuous \( M \)-valued semimartingales with respect to the given filtration (in the sense that their composition with arbitrary smooth functions are real-valued semimartingales) \([21, 10]\). Being a continuous semimartingale, the paths of an adapted Brownian motion can be almost surely horizontally lifted (in a natural sense that relies on Stratonovic stochastic integrals) to smooth principal bundles that come equipped with a smooth connection \([10]\). This is a very remarkable fact, since Brownian paths are almost surely nowhere differentiable \([10]\). Such lifts are the main ingredient of probabilistic formulae for the heat semigroups associated with operators of the \( \nabla^\dagger \nabla \) acting on \( L^2 \)-sections on a metric vector bundle over \( M \) \([21]\).

**Definition 14.8.** \( M \) is called nonparabolic, if and only if for all \( x, y \in M \) with \( x \neq y \) one has the finiteness of the Coulomb potential 
\[ G(x, y) := \int_0^\infty p(t, x, y)dt. \]

One can easily show that compact \( M \)'s are always parabolic, and that the Euclidean \( \mathbb{R}^m \) is nonparabolic if and only \( m \geq 3 \). Probabilistically, this property means \([8]\): 

**Theorem 14.9.** \( M \) is nonparabolic, if and only if every Brownian motion \( (\Omega, \mathcal{F}, \mathbb{P}, X(x_0)) \) on \( M \) with starting point \( x_0 \) is transient, in the sense that for every precompact set \( U \subset M \) one has 
\[ \mathbb{P}\{ \text{there exists } s > 0 \text{ such that for all } t > s \text{ one has } X_t(x_0) \notin U \} = 1, \]
that is, if and only if all Brownian motions on \( M \) eventually leave each precompact set almost surely.

One can show that if \( M \) is geodesically complete with \( \mathrm{Ric} \geq 0 \), then \( (M, \Psi) \) is nonparabolic, if and only if 
\[ \int_0^\infty \frac{t}{\mu(x, \sqrt{t})}dt < \infty \quad \text{for all } x \in M, \]

Finally, we are going to sketch a proof of the Feynman-Kac formula. The aim here is to derive a path integral formula for the semigroup \( P^w_t := e^{-tH^w} \in \mathcal{L}(L^2(M)) \) associated with a Schrödinger operator of the form \( H^w := -\Delta + w \), where \( w : M \to \mathbb{R} \) is a potential. In case \( w = 0 \) we simply have, 
\[ P_t f(x) = \mathbb{E}^x \left[ 1_{\{t < \zeta\}} f(X_t) \right], \]
which is a path integral formula, as
\[ \mathbb{E}^x \left[ 1_{\{t<\zeta\}} f(X_t) \right] = \int_{\{t<\zeta\}} f(\gamma(t)) d\mathbb{P}^x(\gamma). \]

In the general case, according to Richard Feynman’s thesis, we expect a formula of the form
\[ P^w_t f(x) = \int_{\{t<\zeta\}} e^{-\int_0^t w(\gamma(s)) ds} f(\gamma(t)) d\mathbb{P}^x(\gamma) = \mathbb{E}^x \left[ 1_{\{t<\zeta\}} e^{-\int_0^t w(X_s) ds} f(X_t) \right]. \]

Actually, in quantum physics, one is rather interested in the unitary group \( e^{-itHw} \in \mathcal{L}(L^2(M)) \), which with \( \Psi(t) := P^w_t \Psi, \Psi \in L^2(M) \), solves the Schrödinger equation
\[ \frac{d}{dt} \Psi(t) = -iHw \Psi(t), \quad \Psi(0) = \Psi. \]

Feynman then ‘showed’ that (without any mathematical rigour) that
\[ e^{-itHw} f(x) = \int_{\{t<\zeta\}} e^{-i\int_0^t w(\gamma(s)) ds} f(\gamma(t)) e^{-\int_0^t |\dot{\gamma}(s)|^2 ds} D^x(\gamma), \]
where \( D^x \) is some sort of Riemannian volume measure on the space of paths on \( M \) starting \( x \) and \( \int_0^t |\dot{\gamma}(s)|^2 ds \) is the energy of such a path \( \gamma \). Now one can prove that \( D^x \) does not exist, and of course many paths do not have a finite energy. On the other hand, switching from \( i \) to \( t \), although each factor is problematic, the product
\[ e^{-\int_0^t |\dot{\gamma}(s)|^2 ds} D^x(\gamma) \]
is well-defined and in fact one has
\[ e^{-\int_0^t |\dot{\gamma}(s)|^2 ds} D^x(\gamma) = d\mathbb{P}^x(\gamma) \]
in a sense that can be made precise. The point is that \( e^{-\int_0^t |\dot{\gamma}(s)|^2 ds} \) is damping and can absorb some of the infinities of \( D^x(\gamma) \), while \( e^{-\int_0^t |\dot{\gamma}(s)|^2 ds} \) was oscillating and could not do that.

The first issue that has to be attacked is which \( w \)’s can be dealt with in such a formula. In quantum physics, one has to deal with nonsmooth and unbounded \( w \)’s such as the Coulomb potential \( w(x) = -1/|x| \) for \( M = \mathbb{R}^3 \). Ultimately, the following class has turned out to be useful:

**Definition 14.10.** A Borel function \( w : M \to \mathbb{R} \) is said to be in the *Kato class* \( \mathcal{K}(M) \) of \( M \), if
\[ \lim_{t \to 0^+} \sup_{x \in M} \int_0^t \mathbb{E}^x \left[ 1_{\{s<\zeta\}} |w(X_s)| \right] ds = 0. \]

Obviously, \( \mathcal{K}(M) \) is a linear space. We are going to show in an exercise that \( \mathcal{K}(M) \subset L^1_{\text{loc}}(M) \) and that if \( M = \mathbb{R}^3 \) with its Euclidean metric then \( w(x) := 1/|x| \) is in \( \mathcal{K}(\mathbb{R}^3) \).
One can show with some efforts that for $w \in K(M)$ the symmetric densely defined quadratic form
\[
C^\infty_c(M) \times C^\infty_c(M) \ni (f_1, f_2) \mapsto \int (df_1, df_2) d\mu + \int w f_1 f_2 d\mu
\]
in $L^2(M)$ is closable and semibounded from below. The associated semibounded self-adjoint operator is denotes by $H_w$. Note that $H_w|_{w=0} = H$. Thus we are interested in a formula for
\[
P_t^w := e^{-tH_w} \in \mathcal{L}(L^2(M)).
\]
That (89) can hold at all relies on:

**Lemma 14.11.** Let $w \in K(M)$. Then:

a) One has
\[
\sup_{x \in M} \int_M \int_0^T p(s, x, y) |w(y)| ds \, d\mu(y) < \infty \text{ for all } T > 0.
\]

b) For all $x \in M$ one has
\[
\mathbb{P}^x \{ w(X_t) \in L^1_{\text{loc}}(0, \zeta) \} = 1.
\]

c) There are $c_j = c_j(w) > 0$, $j = 1, 2$, such that for all $t \geq 0$,
\[
(91) \quad \sup_{x \in M} \mathbb{E}^x \left[ e^{t^j w(X_t)} |ds_{1_{\{t<\zeta\}}} \right] \leq c_1 e^{tc_2} < \infty.
\]

**Proof:**

a) Take a $t > 0$ with
\[
\sup_{x \in M} \int_M \int_0^T p(s, x, y) |w(y)| ds \, d\mu(y) < \infty
\]
and pick $l \in \mathbb{N}$ with $T < lt$. Then we can estimate
\[
\begin{align*}
\sup_{x \in M} \int_M & \int_0^T p(s, x, y) |w(y)| ds \, d\mu(y) \\
\leq & \sup_{x \in M} \int_M \int_0^{lt} p(s, x, y) |w(y)| ds \, d\mu(y) \\
\leq & \sum_{k=1}^l \sup_{x \in M} \int_M \int_0^t p((k-1)t + s, x, y) |w(y)| ds \, d\mu(y) \\
= & \sum_{k=1}^l \sup_{x \in M} \int_0^t \int_M p((k-1)t, x, z) \int_M p(s, z, y) |w(y)| d\mu(y) d\mu(z) ds \\
\leq & \left( \sum_{k=1}^l \sup_{x \in M} \int_0^t p((k-1)t, x, z) d\mu(z) \right) \times \\
& \times \sup_{x \in M} \int_0^t \int_M p(s, z, y) |w(y)| d\mu(y) ds
\end{align*}
\]
\[ \leq l \sup_{z \in M} \int_0^1 \int_M p(s, z, y) |w(y)| d\mu(y) ds < \infty, \]

where we have used the Chapman-Kolomogorov identity and
\[ \int p(s', x', y') d\mu(y') \leq 1. \]

b) Pick a continuous function \( \rho : M \to [0, \infty) \) such that for all \( c \in [0, \infty) \) the level sets \( \{ \rho \in [c, \infty) \} \) are compact. Then the collection of subsets \( (U_n)_{n \in \mathbb{N}} \) of \( M \) given by

\[ U_n := \text{interior of } \{ \rho \in [1/n, \infty) \} \]

forms an exhaustion of \( M \) with open relatively compact subsets. For every \( n \in \mathbb{N} \), define the first exit times

\[ \zeta_{(1)}^n := \zeta_{U_n} : \Omega_M \to [0, \infty]. \]

Then the sequence \( \zeta_{(1)}^n \) announces \( \zeta \) with respect to \( P^x \) for every \( x \in M \) in the following sense: There exists a set \( \Omega_x \subset \Omega_M \) with \( P^x(\Omega_x) = 1 \), such that for all paths \( \gamma \in \Omega_x \) one has the following two properties:

- \( \zeta_{(1)}^n(\gamma) \nearrow \zeta(\gamma) \) as \( n \to \infty \),
- the implication \( \zeta(\gamma) < \infty \Rightarrow \zeta_{(1)}^n(\gamma) < \zeta(\gamma) \) holds true for all \( n \).

To see that \( \zeta \) is indeed announced by \( \zeta_{(1)}^n \) in the asserted form, one can simply set

\[ \Omega_x := \{ \gamma \in \Omega_M : \gamma(0) = x \}. \]

Then \( P^x(\Omega_x) = 1 \) and the asserted properties follow easily from continuity arguments, since \( \Omega_x \) is a set of continuous paths that start in \( x \). It follows immediately that \( \zeta_{(2)}^n := \min(\zeta_{(1)}^n, n) \) also announces \( \zeta \). As a consequence, we have

\[ \mathbb{P}^x \left\{ w(\mathcal{X}_s) \in L^1_{\text{loc}}[0, \zeta) \right\} = \mathbb{P}^x \left\{ \int_0^{\zeta_{(2)}^n} |h(\mathcal{X}_s)| ds < \infty \right\}. \]

Now we have

\[ \mathbb{E}^x \left[ \int_0^{\zeta_{(2)}^n} |w(\mathcal{X}_s)| ds \right] \leq \mathbb{E}^x \left[ \int_0^{\min(\zeta, n)} |w(\mathcal{X}_s)| ds \right], \]

and this number is finite for all \( n \) by a).

c) With \( \tilde{M} = M \cup \{ \infty_M \} \) the Alexandrov compactification of \( M \), we can canonically extend \( w \) to a Borel function \( \tilde{w} : \tilde{M} \to \mathbb{R} \) by setting \( \tilde{w}(\infty_M) = 0 \). Then one trivially has

\[ \mathbb{E}^x \left[ e^{\int_0^t |w(\mathcal{X}_s)| ds} \mathbf{1}_{\{ t < \zeta \}} \right] \leq \mathbb{E}^x \left[ e^{\int_0^t |\tilde{w}(\mathcal{X}_s)| ds} \right]. \]
2. (Khas’minskii’s lemma) For any $s \geq 0$, let

$$J(w, s) := \sup_{x \in M} \mathbb{E}^x \left[ e^{\int_0^s |\tilde{w}(X_r)| dr} \right] \in [0, \infty].$$

Then for every $s > 0$ with

$$D(w, s) := \sup_{x \in M} \mathbb{E}^x \left[ \int_0^s |w(X_r)| 1_{\{r < \zeta\}} dr \right] < 1,$$

it holds that

$$J(w, s) \leq \frac{1}{1 - D(w, s)}.$$  \hfill (93)

Proof: One has

$$D(w, s) = \sup_{x \in M} \mathbb{E}^x \left[ \int_0^s |\tilde{w}(X_r)| dr \right].$$

For any $n \in \mathbb{N}$, let

$$s\sigma_n := \{ q = (q_1, \ldots, q_n) : 0 < q_1 < \cdots < q_n < s \} \subset \mathbb{R}^n$$

denote the open scaled simplex. In the chain of equalities

$$\mathbb{E}^x \left[ e^{\int_0^s |\tilde{w}(X_r)| dr} \right] = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{[0,s]^n} \mathbb{E}^x \left[ |\tilde{w}(X_{q_1})| \cdots |\tilde{w}(X_{q_n})| \right] d^n q$$

$$= 1 + \sum_{n=1}^{\infty} \int_{s\sigma_n} \mathbb{E}^x \left[ |\tilde{w}(X_{q_1})| \cdots |\tilde{w}(X_{q_n})| \right] d^n q$$

$$= 1 + \sum_{n=1}^{\infty} \int_0^s \int_{q_1}^s \cdots \int_{q_{n-1}}^s \mathbb{E}^x \left[ |\tilde{w}(X_{q_1})| \cdots |\tilde{w}(X_{q_n})| \right] d^n q,$$

the first one follows from Fubini’s theorem, and the second one from combining the fact that the integrand is symmetric in the variables $q_j$ with the fact that the number of orderings of a real-valued tuple of length $n$ is $n!$. In particular, by comparison with a geometric series, it is sufficient to prove that for all natural $n \geq 2$, one has

$$J_n(w, s) := \sup_{x \in M} \int_0^s \int_{q_1}^s \cdots \int_{q_{n-1}}^s \mathbb{E}^x \left[ |\tilde{w}(X_{q_1})| \cdots |\tilde{w}(X_{q_n})| \right] d^n q$$

$$\leq D(w, s) J_{n-1}(w, s).$$  \hfill (94)

But the Markov property of the family of Wiener measures implies

$$J_n(w, s) = \sup_{x \in M} \int_0^s \int_{q_1}^s \cdots \int_{q_{n-1}}^s \mathbb{E}^x \left[ |\tilde{w}(\gamma(q_1))| \cdots |\tilde{w}(\gamma(q_{n-1}))| \right] \times$$

$$\times \int_{\Omega_M} \int_0^{s-q_{n-1}} |\tilde{w}(\omega(u))| d\mu^\gamma(\omega) d\mathbb{P}^{\gamma}(\gamma) d^{n-1} q$$

$$\leq D(w, s) J_{n-1}(w, s),$$  \hfill (95)

which proves Khas’minskii’s lemma.
3. Pick \( s > 0 \) with \( D(w, s) < 1 \). Then for any \( t > 0 \) one has
\[
J(w, t) \leq \frac{1}{1 - D(w, s)} e^{\frac{ts}{\log \left( \frac{1}{1 - D(w, s)} \right)}}.
\]

Proof: Pick a large \( n \in \mathbb{N} \) with \( t < (n + 1)s \). Then the Markov property of the family of Wiener measures and Khas’minskii’s lemma imply
\[
J(w, t) \leq J(w, (n + 1)s) \leq \sup_{x \in M} \int_{\Omega_M} e^{\int_0^{ns} |\tilde{\omega}(r)| dr} d\mathbb{P}^{\gamma(ns)}(\omega) d\mathbb{P}^x(\gamma)
\]
\[
\leq \cdots \ (n\text{-times})
\]
\[
\leq \frac{1}{1 - D(w, s)} \left( \frac{1}{1 - D(w, s)} \right)^n
\]
\[
\leq \frac{1}{1 - D(w, s)} e^{\frac{ts}{\log \left( \frac{1}{1 - D(w, s)} \right)}},
\]
which proves (91) in view of (92).

\[\square\]

**Definition 14.12.** An ordered pair \((\Xi, \tilde{\Xi})\) of functions
\[
\Xi : M \rightarrow (0, \infty], \quad \tilde{\Xi} : (0, \infty) \rightarrow (0, \infty)
\]
is called a **heat kernel control pair for the Riemannian manifold** \(M\), if the following assumptions are satisfied:

- \(\Xi\) is continuous with \(\inf \Xi > 0\), \(\tilde{\Xi}\) is Borel
- for all \(x \in M\), \(t > 0\) one has
  \[
  \sup_{y \in M} p(t, x, y) \leq \Xi(x) \tilde{\Xi}(t)
  \]
- for all \(q' \geq 1\) in the case of \(m = 1\), and for all \(q' > m/2\) in the case of \(m \geq 2\), one has
  \[
  \int_0^\infty \tilde{\Xi}^{1/q'}(t) e^{-At} dt < \infty \quad \text{for some } A > 0.
  \]

**Remark 14.13.** 1. Every Riemannian manifold admits a canonically given heat kernel control pair. Indeed, using an \(L^1\)-variant of the parabolic mean value inequality one can show that
\[
\Xi(x) = \frac{C}{\min(r_{\text{Eucl}}(x), 1)^m}, \quad \tilde{\Xi}(t) = t^{-m/2} + 1
\]
defines such a pair, where \( r_{\text{Eucl}}(x) \) is defined to be infimum of all \( r > 0 \) such that \( B(x, r) \) is relative compact and
\[
(1/2)\delta_{ij} \leq g_{ij} \leq 2\delta_{ij}
\]
thereon.

2. Assume that one has the ultracontractiveness
\[
\sup_{x, y \in M} p(t, x, y) \leq Ct^{-m/2} \quad \text{for all } 0 < t < T.
\]
Then,
\[
(\Xi(x), \tilde{\Xi}(t)) := (C, \min(t, T)^{-m/2})
\]
is a heat kernel control pair, which is constant in its first slot.

3. Assume that \( M \) is geodesically complete with \( \text{Ric} \geq -K \) for some constant \( K \geq 0 \). Then the (local in time) Li-Yau heat kernel bound shows that
\[
\Xi(x) := C\mu(B(x, 1))^{-1},
\]
\[
\tilde{\Xi}(t) := t^{-m/2} + 1
\]
is a heat kernel kontrol pair.

**Proposition 14.14.** Let \( w = w_1 + w_2 : M \to \mathbb{R} \) be a function which can be decomposed into Borel functions \( w_j : M \to \mathbb{R} \) satisfying the following two properties:
- \( w_2 \in L^\infty(M) \)
- there exists a real number \( q' < \infty \) such that \( q' \geq 1 \) if \( m = 1 \), and \( q' > m/2 \) if \( m \geq 2 \), and a heat kernel control pair \((\Xi, \tilde{\Xi})\), such that\(^{19}\) \( w_1 \in L^{q'}_\Xi(M) \).

Then for all \( u > 0 \) and all \( x \in M \), one has the bound
\[
\int_M p(u, x, y)|w(y)|d\mu(y) \leq \tilde{\Xi}(u)^{1/q'} \|w_1\|_{q', \Xi} + \|w_2\|_\infty.
\]
(96)

In particular, for any choice of \( q' \) and \((\Xi, \tilde{\Xi})\) as above one has
\[
L^{q'}_\Xi(M) + L^\infty(M) \subset K(M),
\]
where
\[
L^{q'}_\Xi(M) := L^{q'}(M, \Xi d\mu).
\]

**Proof:** Once we have proved
\[
\int_M p(u, x, y)|w(y)|d\mu(y) \leq \tilde{\Xi}(u)^{1/q'} \|w_1\|_{q', \Xi} + \|w_2\|_\infty,
\]
(97)
\[^{19}\text{Note that one automatically has } L^{q'}_\Xi(M) \subset L^{q'}(M), \text{ which is implied by } \inf \Xi > 0.\]
the inclusion \( w \in \mathcal{K}(M) \) clearly follows from

\[
\lim_{t \to 0^+} \sup_{x \in M} \int_0^t \int_M p(u, x, y) |w(y)| d\mu(y) du \leq C(w_1) \lim_{t \to 0^+} \int_0^t \tilde{\Xi}(u)^{1/q'} du + C(w_2) \lim_{t \to 0^+} t = 0.
\]

In order to derive (97), note first that the inequality

(98) \[
\int_M p(u, x, y) d\mu(y) \leq 1
\]

shows that we can assume \( w_2 = 0 \). Furthermore, the case \( q' = 1 \) (which is only allowed for \( m = 1 \)) is obvious, so let us assume \( m \geq 2 \) and \( q' > m/2 \). The essential trick to bound \[ \int_M p(u, x, y) |w_1(y)| d\mu(y) \]

is to factor the heat kernel appropriately: Indeed, with \( 1/q' + 1/q = 1 \), Hölder’s inequality and using (98) once more gives us the following estimate:

\[
\int_M p(u, x, y) |w_1(y)| d\mu(y) \leq \left( \int_M |w_1(y)|^{q'} p(u, x, y) d\mu(y) \right)^{1/q'} \leq \tilde{\Xi}(u)^{1/q'} \|w_1\|_{q', \Xi}.
\]

This completes the proof.

Example 14.15. We have \( w := 1/|\cdot| \in \mathcal{K}(\mathbb{R}^3) \). Indeed, this \( w = 1_{B(0,1)} + 1_{B(0,1)^c} w \) is in \( L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3) \).

Theorem 14.16. Let \( w \in \mathcal{K}(M) \). Then for all \( t \geq 0 \), \( f \in L^2(M) \), \( \mu \)-a.e. \( x \in M \), one has

(99) \[
e^{-tHw} f(x) = \mathbb{E}^x \left[ 1_{\{t < \zeta\}} e^{-\int_0^t w(X_s) ds} f(X_t) \right].
\]

Proof: Step 1: (99) holds in case \( w : M \to \mathbb{R} \) is continuous and bounded.

Proof: Decomposing

\[ f = f_1 - f_2 + f_3 - \sqrt{-1}f_4, \quad f_j \geq 0, \]

if necessary, we can and we will assume \( f \geq 0 \) for the proof. Since \( w \) is bounded, it simply acts as a bounded multiplication operator (that will be denoted by the same symbol again).
By (82), for every \( t > 0 \), \( n \in \mathbb{N} \) and \( \mu \text{-a.e. } x_0 \in M \), we have
\[
(e^{-t/n}H e^{-(t/n)w})^n f(x_0) = \\
\int \cdots \int \exp \left( -(t/n) \sum_{i=1}^{n} w(x_i) \right) p(t/n, x_0, x_1) \cdots p(t/n, x_{n-1}, x_n) f(x_n) \\
\times d\mu(x_1) \cdots d\mu(x_n) = \\
\mathbb{E}_x \left[ 1_{\{t<\zeta\}} \exp \left( (t/n) \sum_{i=1}^{n} w(X_{t/n}) \right) f(X_t) \right].
\]

Since \( w \) is continuous, for each fixed continuous path which remains on \( M \) until \( t \), the \( \exp(\cdots) \)-expression represents Riemann sums for \( -\int_0^t w(X_s) ds \). Furthermore, we have
\[
1_{\{t<\zeta\}} \exp \left( (t/n) \sum_{i=1}^{n} w(X_{t/n}) \right) f(X_t) \leq \exp(\|w\|_\infty) f(X_t),
\]
and clearly
\[
\mathbb{E}_x \left[ 1_{\{t<\zeta\}} f(X_t) \right] = \int_M p(t, x_0, y) f(y) d\mu(y) < \infty,
\]
therefore dominated convergence shows that for \( \mu \text{-a.e. } x_0 \in M \),
\[
\lim_{n \to \infty} (e^{-(t/n)H} e^{-(t/n)w})^n f(x_0) = \mathbb{E}_x \left[ 1_{\{t<\zeta\}} e^{-\int_0^t w(X_s) ds} f(X_t) \right].
\]

On the other hand, Trotter’s product formula
\[
A \text{ semibounded, } B \text{ bounded, } e^{-t(A+B)} = \lim_{n \to \infty} \left( e^{-t/n}A e^{-(t/n)B} \right)^n \text{ strongly,}
\]
gives (after picking a subsequence, if necessary, to turn the \( L^2 \)-convergence to a \( \mu \text{-a.e. } \) convergence)
\[
\lim_{n \to \infty} e^{-(t/n)H} e^{-(t/n)w} f(x_0) = e^{-tH} f(x_0) \text{ for } \mu \text{-a.e. } x_0,
\]
which proves the Feynman-Kac formula in this case.

Step 2: (99) holds in case \( w \) is a bounded potential.
Proof: We will use Friedrichs mollifiers to reduce everything to the continuous (in fact: smooth) bounded case from step 1. To this end, we pick an atlas \((U_l)_{l \in \mathbb{N}}\) for \( M \) such that each \( U_l \) is relatively compact. We also take a subordinate partition of unity \( \phi_l \in C_c(U_l) \).

Then
\[
w^{(l)} := \varphi_l w : U_l \to \mathbb{R}
\]
defines a bounded compactly supported function, and by Remark ???.ii) we can pick a sequence
\[
(w^{(l)}_n) \subset C_c(U_l)
\]
such that \( \mu \text{-a.e. } \) \( U_l \) we have
\[
|w^{(l)}_n| \leq \|w\|_\infty < \infty, \quad w^{(l)}_n \to w^{(l)} \text{ as } n \to \infty.
\]
Defining a sequence of smooth potentials
\[ w_n := \sum_l \varphi_l w_n^{(l)}, \]
one has
\[ |w_n| \leq \|w\|_\infty, \quad w_n \to w \mu\text{-a.e.} \tag{100} \]
It is clear from (100) and dominated convergence that
\[ \lim_{n \to \infty} H_{w_n} \psi = H_w \psi \text{ in } L^2(M) \]
for all
\[ \psi \in \text{Dom}(H_w) = \text{Dom}(H_{w_n}) = \text{Dom}(H). \]
Thus by an abstract convergence result for semigroups, which deduces the strong convergence of the semigroups
\[ A_n f \to Af \quad \text{for all } f \in \text{Dom}(A) = \text{Dom}(A_n), \quad A, A_n \text{ semibounded} \Rightarrow e^{-tA_n} \to e^{-tA} \text{ strongly} \]
we have
\[ \lim_{n \to \infty} e^{-tH_{w_n}} f = e^{-tH_w} f \text{ in } L^2(M). \]
In particular, passing to a subsequence if necessary, we can and we will assume
\[ \lim_{n \to \infty} e^{-tH_{w_n}} f(x) = e^{-tH_w} f(x) \quad \text{for } \mu\text{-a.e. } x, \tag{101} \]
so we find
\[ e^{-tH_w} f(x) = \lim_{n \to \infty} E^n_x \left[ 1_{\{t < \zeta\}} e^{-\int_0^t w_n(X_s) ds} f(X_t) \right] \quad \text{for } \mu\text{-a.e. } x, \tag{102} \]
by the already established validity of the covariant Feynman-Kac formula for \( e^{-tH_{w_n}} f \). It remains to show that the right-hand side of (102) is equal to
\[ E^n_x \left[ 1_{\{t < \zeta\}} e^{-\int_0^t w(X_s) ds} f(X_t) \right]. \]
To this end, applying (100) together with the elementary inequality
\[ |e^a - e^b| \leq 2 |a - b| e^{\max(a, b)}, \quad a, b \in \mathbb{R}, \tag{103} \]
since that one has
\[ 1_{\{t < \zeta\}} \left| e^{-\int_0^t w_n(X_s) ds} - e^{-\int_0^t w_n(X_s) ds} \right| \leq 2 \cdot 1_{\{t < \zeta\}} e^{\|w\|_\infty t} \int_0^t |w(X_s) - w_n(X_s)| ds \quad \mathbb{P}^x\text{-a.s.}, \]
so using (100) once more with dominated convergence, we find
\[ \lim_{n \to \infty} 1_{\{t < \zeta\}} \left| e^{-\int_0^t w_n(X_s) ds} - e^{-\int_0^t w(X_s) ds} \right| = 0 \quad \mathbb{P}^x\text{-a.s.} \tag{104} \]
Finally, we may use (104) and
\[ 1_{\{t < \zeta\}} \left| e^{-\int_0^t w_n(X_s) ds} \right| \leq e^{\|w\|_\infty t} \quad \mathbb{P}^x\text{-a.s.} \]
to deduce (99) from (102) and dominated convergence.

Step 3: (99) holds in case $w$ is bounded from below.

Proof: Set $w_n := \min(n, w)$, use step 2 and convergence results for semigroups for the LHS and convergence results for integrals for RHS.

Step 4: (99) holds for $w$ Kato.

Proof: Set $w_n := w_n := \max(-n, w)$, use step 2 and convergence results for semigroups for the LHS and convergence results for integrals for RHS.

References