Contents

Chapter 1

Laplace operator on a Riemannian manifold

18.10.16

We introduce in this Chapter the notions of smooth and Riemannian manifolds, Riemannian measure, and the Riemannian Laplace operator.

1.1 Smooth manifolds

Recall that a *topological space* is a couple $(M, \mathcal{O})$ where M is any set and $\mathcal{O}$ is a collection of subsets of M that are called *open* and satisfy the following axioms:

- $\emptyset$ and M are open;
- union of any family of open sets is open;
- intersection of two open sets is open.

A subset F of M is called closed if its complement $F^c := M \setminus F$ is open. A subset K of M is called compact if any open covering $\{\Omega_\alpha\}$ of K contains a finite subcover. It is useful to observe that any closed set, that is a subset of a compact set, is also compact.

A topological space M is called *Hausdorff* if, for any two disjoint points $x, y \in M$, there exist two disjoint open sets $U, V \subset M$ containing x and y, respectively. In a Hausdorff space M, any compact subset K of M is closed[1].

We say that M has a *countable base* if there exists a countable family $\{B_j\}_{j=1}^\infty$ of open sets in M such that any other open set is a union of some sets B_j. The family $\{B_j\}$ is called a base of the topology of M.

Any metric space (M, d) has a standard topology: a subset $\Omega \subset M$ is called open if for any $x \in \Omega$ there is a metric ball $B(x, r)$ with radius $r > 0$ that is a subset of Ω. The topology of a metric space is automatically Hausdorff. A metric space has a countable base if and only if it is separable, that is, contains a countable dense subset D. Indeed, by all balls of rational radii centered at the points of D form a countable base.

[1] Indeed, the Hausdorff property implies that, for any $x \in K^c$ there is an open set U_x containing x and disjoint from K. Since $K^c = \bigcup_{x \in K^c} U_x$, it follows that K^c is open.

For example, $\mathbb{R}^n$ as a metric space with the Euclidean distance is an example of a Hausdorff topological space with a countable base.

Definition. A n-dimensional *chart* on a topological space M is any couple (U, φ) where U is an open subset of M and φ is a homeomorphism of U onto an open subset of $\mathbb{R}^n$ (which is called the image of the chart).

Any chart (U, φ) on M gives rise to the *local coordinate system* $x^1, x^2, ..., x^n$ in U by taking the φ-pullback of the Cartesian coordinate system in $\mathbb{R}^n$. Hence, we can say that a chart is an open set $U \subset M$ with a local coordinate system. Normally, we will identify U with its image so that the coordinates $x^1, x^2, ..., x^n$ can be regarded as the Cartesian coordinates in a region in $\mathbb{R}^n$.

Definition. A *C-manifold* of dimension n is a Hausdorff topological space M with a countable base such that any point of M belongs to a n-dimensional chart. The collection of all n-dimensional charts on M is called an *atlas*.

For example, $\mathbb{R}^n$ is a C-manifold and $U = \mathbb{R}^n$ is a single n-dimensional chart that covers $\mathbb{R}^n$. If U is any open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^k$ is a continuous function then its graph

$$\Gamma = \left\{(x, y) \in \mathbb{R}^{n+k} : x \in U \, , y = f\,(x)\right\}$$

is a C-manifold because it is covered by a single n-dimensional chart (Γ, φ) with $\varphi\,(x, y) = x$.

A *hypersurface* M in $\mathbb{R}^{n+1}$ is a subset of $\mathbb{R}^n$ such that for any point $x \in M$, there exists an open set $\Omega \subset \mathbb{R}^{n+1}$ containing x such that $\Omega \cap M$ is a graph of a continuous function defined on an open subset of $\mathbb{R}^n$, with respect to one of the coordinates $x^1, ..., x^{n+1}$. Clearly, any hypersurface is a C-manifold. For example, the unit sphere

$$\mathbb{S}^n := \left\{x \in \mathbb{R}^{n+1} : \|x\| = 1\right\}$$

is a C-manifold of dimension n.

If (U, φ) and (V, ψ) are two charts on a C-manifold M then in the intersection $U \cap V$ two coordinate systems are defined, say $x^1, x^2, ..., x^n$ and $y^1, y^2, ..., y^n$. The change of the coordinates is given then by *continuous* functions $y^i = y^i(x^1, ..., x^n)$ and $x^i = x^i\,(y^1, ..., y^n)$, because the functions $y^i(x^1, ..., x^n)$ are the components of the mapping $\psi \circ \varphi^{-1}$ and the functions $x^i\,(y^1, ..., y^n)$ are the components of the mapping $\varphi \circ \psi^{-1}$ (see Fig. 1.1).

Definition. A family $\mathcal{A}$ of charts on a C-manifold is called a *C^k-atlas* (where k is a positive integer or $+\infty$) if the charts from $\mathcal{A}$ covers all M and the change of coordinates in the intersection of any two charts from $\mathcal{A}$ is given by C^k-functions. Two C^k-atlases are said to be compatible if their union is again a C^k-atlas. The union of all compatible C^k-atlases determines a *C^k-structure* on M.

Definition. A C^k-manifold is a C-manifold endowed with a C^k-structure. A smooth manifold is a C^∞-manifold.

Alternatively, one can say that a C^k-manifold is a couple $(M, \mathcal{A})$, where M is a C-manifold and $\mathcal{A}$ is a C^k-atlas on M. However, if the two C^k-atlases $\mathcal{A}$ and $\mathcal{A}'$ are compatible then $(M, \mathcal{A})$ and $(M, \mathcal{A}')$ determine the same C^k-manifold.

In this course we are going to consider mostly smooth manifolds. By default, the term "manifold" will be used as a synonymous of "smooth manifold". By a chart on a

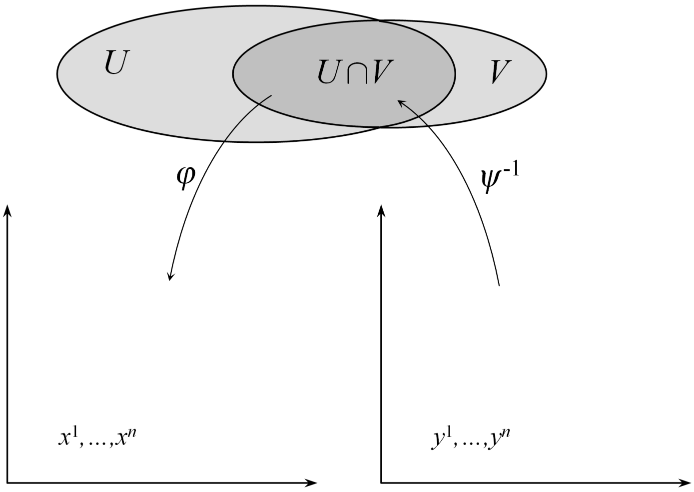

Figure 1.1: The mapping $\varphi \circ \psi^{-1}$

smooth manifold we will always mean a chart from its C^∞-structure, that is, any chart compatible with the defining atlas $\mathcal{A}$.

A trivial example of a smooth manifold is $\mathbb{R}^n$ with the C^∞-atlas consisting of a single chart $(\mathbb{R}^n, \mathrm{id})$. Also, the graph of any C^∞-function $f : U \to \mathbb{R}^k$ (where U is an open subset of $\mathbb{R}^n$) is a smooth manifold. Any C^∞-hypersurface (that is locally a graph of a C^∞-function) is a smooth manifold. In particular, the unit sphere $\mathbb{S}^n$ is a smooth manifold.

If f is a (real valued) function on a smooth manifold M and k is a non-negative integer or ∞ then we write $f \in C^k(M)$ (or $f \in C^k$) if the restriction of f to any chart is a C^k function of the local coordinates $x^1, x^2, ..., x^n$. The set $C^k(M)$ is a linear space over $\mathbb{R}$ with respect to the usual addition of functions and multiplication by constant.

For any function $f \in C(M)$, its support is defined by

$$\operatorname{supp} f = \overline{\{x \in M : f(x) \neq 0\}},$$

where the bar stands for the closure of the set in M. Denote by $C_0^k(M)$ the subspace of $C^k(M)$, which consists of functions whose support is compact.

It follows from the definition of $\operatorname{supp} f$ that if f vanishes outside a closed set $F \subset M$ then $\operatorname{supp} f \subset F$.

If Ω is an open subset of M then Ω naturally inherits all the above structures of M and becomes a smooth manifold if M is so. Indeed, the open sets in Ω are defined as the intersections of open sets in M with Ω, and in the same way one defines charts and atlases in Ω.

The hypothesis of a countable base will be mostly used via the next simple lemma.

Lemma 1.1 *For any manifold M, there is a countable family $\{U_i\}_{i=1}^\infty$ of relatively compact charts covering all M and such that the closure $\overline{U_i}$ is contained in a chart.*

Before the proof, let us clarify some topological issues. If $\Omega \subset M$ is an open set and $E \subset M$ then the relation $E \Subset \Omega$ (*compact inclusion*) means that the closure $\overline{E}$

of E in M is compact and $\overline{E} \subset \Omega$. The compact inclusion will be frequently used but it may become ambiguous if Ω is a chart on M because in this case $E \Subset \Omega$ can be understood also in the sense of the topology of $\mathbb{R}^n$, when Ω is identified as a subset of $\mathbb{R}^n$. Let us show that the two meanings of $E \Subset \Omega$ are identical. Assume $E \subset \Omega$ and denote by $\widetilde{E}$ the closure of E in $\mathbb{R}^n$. If $E \Subset \Omega$ in the topology of $\mathbb{R}^n$ then $\widetilde{E}$ is compact in $\mathbb{R}^n$ and, hence, its pullback to M (also denoted by $\widetilde{E}$) is compact in M. Since M is Hausdorff, $\widetilde{E}$ is also closed in M. Since $E \subset \widetilde{E} \subset \Omega$, it follows that $\overline{E} \subset \widetilde{E}$ and, hence, $\overline{E}$ is compact and is a subset of Ω also in M. The converse statement is proved in the same way.

Proof of Lemma 1.1. Any point $x \in M$ is contained in a chart, say V_x. Choose $U_x \Subset V_x$ to be a small open ball around x so that U_x is also a chart. By definition, manifold M has a countable base, say $\{B_j\}_{j=1}^{\infty}$. Let us mark each set B_j which is contained in some set U_x. Since U_x is open, it is a union of some marked sets B_j. It follows that all marked B_j cover M. Select for each marked B_j exactly *one* set U_x containing B_j. Thus, we obtain a countable family of sets U_x covering M, which finishes the proof. ∎

In particular, we see that a manifold M is a *locally compact* topological space.

1.2 Partition of unity

20.10.16

We say that a function $\varphi \in C_0^\infty(\mathbb{R}^n)$ is a *mollifier* if $\operatorname{supp}\varphi \subset B_1(0)$, $\varphi \geq 0$, and

$$\int_{\mathbb{R}^n} \varphi d\mu = 1. \tag{1.1}$$

For example, the following function

$$\varphi(x) = \begin{cases} c \exp\left(-\frac{1}{\left(\frac{1}{4}-|x|^2\right)^2}\right), & |x| < 1/2 \\ 0, & |x| \geq 1/2 \end{cases} \tag{1.2}$$

is a mollifier, for a suitable normalizing constant $c > 0$ (see Fig. 1.2).

If φ is a mollifier then, for any $0 < \varepsilon < 1$, the function

$$\varphi_\varepsilon := \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right)$$

is also a mollifier, and $\operatorname{supp}\varphi_\varepsilon \subset B_\varepsilon(0)$.

Theorem 1.2 (Partition of unity) *Let K be a compact subset of $\mathbb{R}^n$ and $\{U_j\}_{j=1}^k$ be a finite family of open sets covering K. Then there exist non-negative functions $\varphi_j \in C_0^\infty(U_j)$ such that $\sum_{j=1}^k \varphi_j \equiv 1$ in an open neighbourhood of K and $\sum_{j=1}^k \varphi_j \leq 1$ in $\mathbb{R}^n$.*

Such a family of functions φ_j is called a *partition of unity* at K subordinate to the covering $\{U_j\}$.

Proof. Consider first the case $k = 1$, that is, when the family $\{U_j\}$ consists of a single set U covering K. Then we will construct a function $\psi \in C_0^\infty(U)$ such that $0 \leq \psi \leq 1$

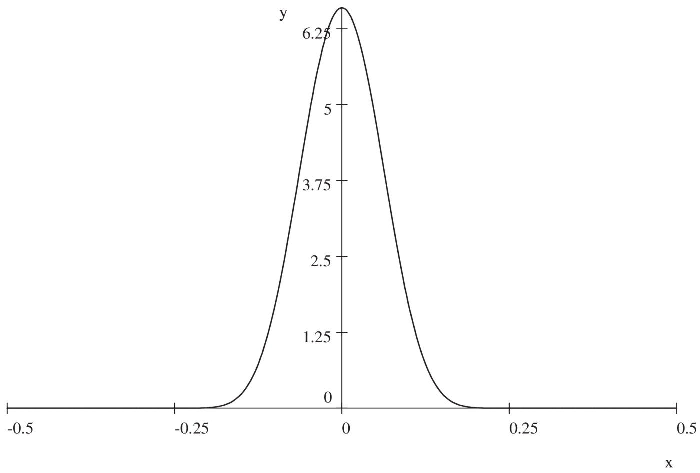

Figure 1.2: The mollifier (1.2) in $\mathbb{R}$.

and $\psi \equiv 1$ in an open neighbourhood of K. Such a function ψ is called a *cutoff function* of K in U.

Let V be an open neighborhood of K such that $V \Subset U$, and set $f = 1_V$. Fix a mollifier φ, $\varepsilon > 0$ and consider the convolution

$$f * \varphi_\varepsilon (x) = \int_{\mathbb{R}^n} f(x-y)\,\varphi_\varepsilon(y)\,dy = \int_{B_\varepsilon(x)} f(z)\,\varphi_\varepsilon(x-z)\,dz.$$

Since $f \in L^1(\mathbb{R}^n)$, we have $f * \varphi_\varepsilon \in C^\infty(\mathbb{R}^n)$. Clearly, $f * \varphi_\varepsilon \geq 0$ and

$$f * \varphi_\varepsilon (x) \leq \sup|f| \int_{\mathbb{R}^n} \varphi_\varepsilon(y)\,dy = \sup|f| = 1.$$

If ε is small enough then $f * \varphi_\varepsilon$ is supported in U so that $f * \varphi_\varepsilon \in C_0^\infty(U)$. Besides, for small enough ε and for any $x \in K$, we have $B_\varepsilon(x) \subset V$, whence $f|_{B_\varepsilon(x)} = 1$ and

$$f * \varphi_\varepsilon (x) = \int_{B_\varepsilon(x)} f(z)\,\varphi_\varepsilon(x-z)\,dz = \int_{B_\varepsilon(x)} \varphi_\varepsilon(x-z)\,dz = 1.$$

Hence, the function $\psi = f * \varphi_\varepsilon$ is a cutoff function of K in U, provided ε is small enough.

Consider now the general case of an arbitrary finite family $\{U_j\}_{j=1}^k$. Any point $x \in K$ belongs to a set U_j. Hence, there is a ball B_x centered at x and such that $B_x \Subset U_j$. The family of balls $\{B_x\}_{x \in K}$ obviously covers K. Since K is compact, there is a finite subfamily $\{B_i\}_{i=1}^m$ covering K. For any $j = 1, ..., k$, consider the set

$$V_j := \bigcup_{\{i : B_i \Subset U_j\}} B_i$$

(see Fig. 1.3).

By construction, the set V_j is open, $V_j \Subset U_j$, and the union of all sets V_j covers K. By the first part of the proof, there exists a cutoff function ψ_j of V_j in U_j. Define now functions φ_j, $j = 1, ..., k$, by

$$\varphi_j = \psi_j (1 - \psi_1) \ldots (1 - \psi_{j-1}), \tag{1.3}$$

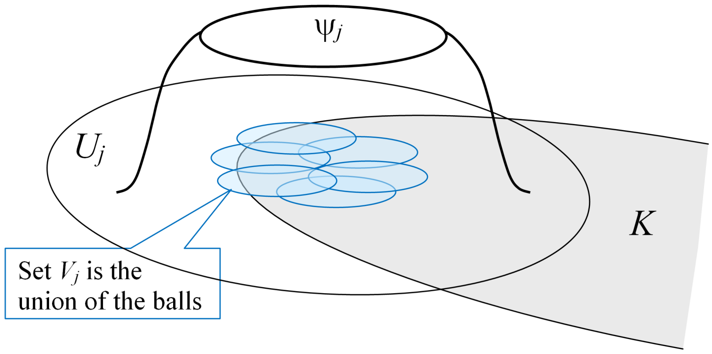

Figure 1.3: Function ψ_j is a cutoff function of V_j in U_j.

that is,

$$\varphi_1 = \psi_1,\ \varphi_2 = \psi_2 (1 - \psi_1),\ ...,\ \varphi_k = \psi_k (1 - \psi_1) ... \left(1 - \psi_{k-1}\right).$$

Obviously, $\varphi_j \in C_0^\infty (U_j)$ and $\varphi_j \geq 0$. It is easy to check by induction in k the following identity

$$1 - \sum_{j=1}^{k} \varphi_j = (1 - \psi_1) ... (1 - \psi_k). \tag{1.4}$$

Indeed, for $k = 1$ it is trivial. If it is true for some k, then

$$\begin{aligned} 1 - \sum_{j=1}^{k+1} \varphi_j &= (1 - \psi_1) ... (1 - \psi_k) - \varphi_{k+1} \\ &= (1 - \psi_1) ... (1 - \psi_k) - \psi_{k+1} (1 - \psi_1) ... (1 - \psi_k) \\ &= (1 - \psi_1) ... (1 - \psi_k) \left(1 - \psi_{k+1}\right), \end{aligned}$$

which proves the induction step. It follows from (1.4) that $\sum \varphi_j \leq 1$. Since $1 - \psi_j = 0$ on V_j, (1.4) implies also that $\sum_j \varphi_j \equiv 1$ on the union of sets V_j and, in particular, on K, which was to be proved. ∎

The following statement extends Theorem 1.2 and provides a convenient vehicle for transporting the local properties of $\mathbb{R}^n$ to manifolds.

Theorem 1.3 *Let K be a compact subset of a smooth manifold M and $\{U_j\}_{j=1}^k$ be a finite family of open sets covering K. Then there exist non-negative functions $\varphi_j \in C_0^\infty (U_j)$ such that $\sum_{j=1}^k \varphi_j \equiv 1$ in an open neighbourhood of K and $\sum_{j=1}^k \varphi_j \leq 1$ in M.*

A sequence of functions $\left\{\varphi_j\right\}$ as in Theorem 1.3 is called a *partition of unity* at K subordinate to the cover $\{U_j\}$.

A particular case of Theorem 1.3 with $k = 1$ says that, for any compact K and any open set $U \supset K$, there exists a function $\varphi \in C_0^\infty (U)$ such that $\varphi \equiv 1$ in a neighborhood of K and, besides, $0 \leq \varphi \leq 1$. Such a function φ is called a *cutoff function* of K in U.

Proof. Assume first that each set U_j is a chart. In this case the proof of Theorem 1.2 goes through unchanged. Indeed, as in the proof of Theorem 1.2, we construct first open subsets $V_j \Subset U_j$ such that $\{V_j\}$ is a covering of K and then use a cutoff function ψ_j of V_j in U_j. Note that most essentially the properties of $\mathbb{R}^n$ (like convolution etc) were used only in order to prove the existence of a cutoff function. As U_j is a chart, we can identify U_j with an open subset of $\mathbb{R}^n$, construct the cutoff function $\psi_j \in C_0^\infty(U_j)$ in $\mathbb{R}^n$ and then transplant ψ_j to M by extending ψ_j to the whole M by setting $\psi_j = 0$ in $M \setminus U_j$ (recall that the meaning of the relation $\Subset$ inside U_j is the same regardless of whether U_j is regarded as a subset of M or that of $\mathbb{R}^n$). After we have obtained the cutoff functions $\{\psi_j\}$, the partition of unity $\{\varphi_j\}$ is constructed by means of the same formula (1.3).

Consider now the general case, when U_j are arbitrary open subsets of M. For any point $x \in K$, there is a chart W_x containing x. Since x is also covered by one of the sets U_j, by reducing W_x we can assume that $W_x \subset U_j$ for some j. Since the family $\{W_x\}_{x\in K}$ covers K, there exists a finite subfamily $\{W_i\}_{i=1}^m$ also covering K. Since each W_i is a chart, by the first part of the proof there exists a partition of unity $\{\psi_i\}_{i=1}^m$ at K subordinate to $\{W_i\}$. Now define the sequence $\{\varphi_j\}_{j=1}^k$ as follows:

$$\varphi_1 = \sum_{\{i:\operatorname{supp}\psi_i \subset U_1\}} \psi_i,$$

$$\varphi_2 = \sum_{\{i:\operatorname{supp}\psi_i \subset U_2,\ \operatorname{supp}\psi_i \not\subset U_1\}} \psi_i,$$

$$\cdots$$

$$\varphi_k = \sum_{\{i:\operatorname{supp}\psi_i \subset U_k,\ \operatorname{supp}\psi_i \not\subset U_l\ \forall l<k.\}} \psi_i.$$

Clearly, each φ_j is non-negative and belongs to $C_0^\infty(U_j)$. Since W_i is covered by some U_j, each ψ_i is supported in some U_j and, hence, each ψ_i has been used in the above construction exactly once. It follows that

$$\sum_j \varphi_j \equiv \sum_i \psi_i,$$

which implies that $\{\varphi_j\}$ is a partition of unity at K subordinate to $\{U_j\}$. ■

Corollary 1.4 *Let $\{\Omega_\alpha\}$ be an arbitrary covering of M by open sets. Then, for any function $f \in C_0^\infty(M)$, there exists a finite sequence $\{f_j\}_{j=1}^k$ of functions from $C_0^\infty(M)$ such that each f_j is supported in one of the sets Ω_α and*

$$f = f_1 + ... + f_k \quad \text{on } M. \tag{1.5}$$

Proof. Let $K = \operatorname{supp} f$ and let $\Omega_1, ..., \Omega_k$ be a finite subfamily of $\{\Omega_\alpha\}$ that covers K. By Theorem 1.3, there exists a partition of unity $\{\varphi_j\}_{j=1}^k$ at K subordinate to $\{\Omega_j\}_{j=1}^k$. Set $f_j = f\varphi_j$ so that $f_j \in C_0^\infty(\Omega_j)$. Then we have

$$\sum_{j=1}^k f_j = f \quad \text{on } M,$$

because on K we have $\sum_j \varphi_j = 1$, while outside K all the functions f and f_j vanish. ■

1.3 Tangent vectors

25.10.16

Let M be a smooth manifold and x_0 be a point on M.

Definition. A mapping $\xi : C^\infty(M) \to \mathbb{R}$ is called an $\mathbb{R}$-*differentiation* at $x_0 \in M$ if

- ξ is linear;
- ξ satisfies the *product rule* in the following form:

$$\xi(fg) = \xi(f)\, g(x_0) + \xi(g)\, f(x_0),$$

  for all $f, g \in C^\infty$.

The set of all $\mathbb{R}$-differentiations at x_0 is denoted by $T_{x_0}M$. For any $\xi, \eta \in T_{x_0}M$ one defines the sum $\xi + \eta$ as the sum of two functions on C^∞, and similarly one defined $\lambda\xi$ for any $\lambda \in \mathbb{R}$. It is easy to check that both $\xi + \eta$ and $\lambda\xi$ are again $\mathbb{R}$-differentiations, so that $T_{x_0}M$ is a linear space over $\mathbb{R}$.

Definition. The linear space $T_{x_0}M$ is called the *tangent space* of M at x_0, and its elements (that is, $\mathbb{R}$-differentiations) are also called *tangent vectors* at x_0.

In $\mathbb{R}^n$ we have the following example of $\mathbb{R}$-differentiation:

$$\xi(f) = \frac{\partial f}{\partial x^i}(x_0),$$

that is clearly linear and satisfies the product rule. In particular, $T_{x_0}\mathbb{R}^n$ contains n linearly independent $\mathbb{R}$-differentiations $\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}$. Moreover, for any vector $v \in \mathbb{R}^n$, the directional derivative $\frac{\partial f}{\partial v}(x_0)$ is also a $\mathbb{R}$-differentiation, which allows us to identify $\mathbb{R}^n$ as a subspace of $T_{x_0}\mathbb{R}^n$. Since

$$\frac{\partial f}{\partial v} = v^i \frac{\partial f}{\partial x^i}$$

(where we assume summation over the repeated index i), it follows that

$$\frac{\partial}{\partial v} = v^i \frac{\partial}{\partial x^i}.$$

Theorem 1.5 *If M is a smooth manifold of dimension n then the tangent space $T_{x_0}M$ is a linear space of the same dimension n.*

Consequently, $\dim T_{x_0}\mathbb{R}^n = n$, which implies that every $\mathbb{R}$-differentiation in $\mathbb{R}^n$ has the form $\frac{\partial}{\partial v}$ for some $v \in \mathbb{R}^n$.

We will prove Theorem 1.5 after a series of claims.

Claim 1. *Let $U \subset M$ be an open set and $U_0 \Subset U$ be its open subset. Then, for any function $f \in C^\infty(U)$, there exists a function $F \in C^\infty(M)$ such that $f \equiv F$ in U_0.*

Proof. Indeed, let ψ be a cutoff function of U_0 in U (see Theorem 1.3). Then define function F by

$$\begin{cases} F = \psi f & \text{in } U, \\ F = 0 & \text{in } M \setminus U, \end{cases}$$

which clearly satisfies all the requirements. ■

Claim 2. *Let $f \in C^\infty(M)$ and let $f \equiv 0$ in an open neighbourhood U of a point $x_0 \in M$. Then $\xi(f) = 0$ for any $\xi \in T_{x_0}M$. Consequently, if f_1 and f_2 are smooth functions on M such that $f_1 \equiv f_2$ in an open neighbourhood of a point $x_0 \in M$ then $\xi(f_1) = \xi(f_2)$ for any $\xi \in T_{x_0}M$.*

Proof. Let U_0 be a neighborhood of x_0 such that $U_0 \Subset U$ and let ψ be a cutoff function of U_0 in U. Then we have $f\psi \equiv 0$ on M, which implies $\xi(f\psi) = 0$. On the other hand, we have by the product rule

$$\xi(f\psi) = \xi(f)\,\psi(x_0) + \xi(\psi)\,f(x_0) = \xi(f),$$

because $\psi(x_0) = 1$ and $f(x_0) = 0$. Hence, $\xi(f) = 0$. The second part follows from the first one applied to the function $f = f_1 - f_2$. ■

Remark. Originally a tangent vector $\xi \in T_{x_0}M$ is defined as a functional on $C^\infty(M)$. The results of Claims 1 and 2 imply that ξ can be regarded as a functional on $C^\infty(U)$ where U is any neighbourhood of x_0. Indeed, by Claim 1, for any $f \in C^\infty(U)$ there exists a function $F \in C^\infty(M)$ such that $f = F$ in a neighborhood U_0 of x_0; hence, set $\xi(f) := \xi(F)$. By Claim 2, this definition of $\xi(f)$ does not depend on the choice of F.

Claim 3. *Let f be a smooth function in a ball $B = B_R(o)$ in $\mathbb{R}^n$ where o is the origin of $\mathbb{R}^n$. Then there exist smooth functions $h_1, h_2, ..., h_n$ in B such that, for any $x \in B$,*

$$f(x) = f(o) + x^i h_i(x), \tag{1.6}$$

where we assume summation over the repeated index i. Also, we have

$$h_i(o) = \frac{\partial f}{\partial x^i}(o). \tag{1.7}$$

Proof. By the fundamental theorem of calculus applied to the function $t \mapsto f(tx)$ on the interval $t \in [0, 1]$, we have

$$f(x) = f(o) + \int_0^1 \frac{d}{dt} f(tx)\,dt, \tag{1.8}$$

whence it follows

$$f(x) = f(o) + \int_0^1 x^i \frac{\partial f}{\partial x^i}(tx)dt.$$

Setting

$$h_i(x) = \int_0^1 \frac{\partial f}{\partial x^i}(tx)dt$$

we obtain (1.6). Clearly, $h_i \in C^\infty(B)$. The identity (1.7) follows from the line above by substitution $x = o$. ■

Claim 4. *Under the hypothesis of* Claim 3, *there exist smooth functions h_{ij} in B (where $i,j=1,2,...,n$) such that, for any $x\in B$,*

$$f(x)=f(o)+x^i\frac{\partial f}{\partial x^i}(o)+x^ix^jh_{ij}(x). \tag{1.9}$$

Proof. Applying (1.6) to the function h_j instead of f we obtain that there exist smooth functions h_{ij} in B, such that

$$h_j(x)=h_j(o)+x^ih_{ij}(x).$$

Substituting this into the representation (1.6) for f and using $h_j(o)=\frac{\partial f}{\partial x^j}(o)$ we obtain

$$f(x)=f(o)+x^ih_i(x)=f(o)+x^i\frac{\partial f}{\partial x^i}(o)+x^ix^jh_{ij}(x).$$

■

Now we can prove Theorem 1.5.

Proof of Theorem 1.5. Let $x^1,x^2,...,x^n$ be local coordinates in a chart U containing x_0. All the partial derivatives $\frac{\partial}{\partial x^i}$ evaluated at x_0 are $\mathbb{R}$-differentiations at x_0, and they are linearly independent. We will prove that any tangent vector $\xi\in T_{x_0}M$ can be represented in the form

$$\xi=\xi^i\frac{\partial}{\partial x^i}\quad\text{where}\quad\xi^i=\xi\left(x^i\right). \tag{1.10}$$

Note that, by the above Remark, the $\mathbb{R}$-differentiation ξ applies also to smooth functions defined in a neighborhood of x_0; in particular, $\xi\left(x^i\right)$ is well-defined. The identity (1.10) implies that $\left\{\frac{\partial}{\partial x^i}\right\}_{i=1}^n$ is a basis in the linear space $T_{x_0}M$ and, hence, $\dim T_{x_0}M=n$.

Without loss of generality, we can assume that x_0 is the origin o of the chart U. For any smooth function f on M, we have by (1.9) the following representation in a ball $B\subset U$ centred at o:

$$f(x)=f(o)+x^i\frac{\partial f}{\partial x^i}(o)+x^ix^jh_{ij}(x),$$

where h_{ij} are some smooth functions in B. Using the linearity of ξ, we obtain

$$\xi(f)=\xi(1)f(o)+\xi\left(x^i\right)\frac{\partial f}{\partial x^i}(o)+\xi\left(x^ix^jh_{ij}\right). \tag{1.11}$$

By the product rule, we have

$$\xi(1)=\xi(1\cdot 1)=\xi(1)1+\xi(1)1=2\xi(1),$$

whence $\xi(1)=0$. Set $u_i=x^jh_{ij}$. By the linearity and the product rule, we have

$$\xi\left(x^iu_i\right)=\xi\left(x^i\right)u_i(o)+\xi(u_i)x^i(o)=0,$$

because $x^i(o) = 0$ and $u_i(o) = x^j(o)\, h_{ij}(o) = 0$. Hence, in the right hand side of (1.11), the first and the third term vanish. Setting $\xi^i = \xi(x^i)$, we obtain

$$\xi(f) = \xi^i \frac{\partial f}{\partial x^i}, \tag{1.12}$$

which is equivalent to (1.10). ■

The numbers ξ^i are referred to as the *components* of the vector ξ in the coordinate system $x^1, ..., x^n$. One often uses the following alternative notation for $\xi(f)$:

$$\xi(f) \equiv \frac{\partial f}{\partial \xi}.$$

Then the identity (1.12) takes the form

$$\boxed{\frac{\partial f}{\partial \xi} = \xi^i \frac{\partial f}{\partial x^i}}, \tag{1.13}$$

which allows to think of ξ as a *direction* at x_0 and to interpret $\frac{\partial f}{\partial \xi}$ as a directional derivative.

As any other finite dimensional linear space, T_xM possesses the *dual space* T_x^*M that consists of all linear functionals on T_xM. That is, any element $\omega \in T_x^*M$ is defined as a linear mapping $\omega : T_xM \to \mathbb{R}$. The value $\omega(\xi)$ for $\xi \in T_xM$ will be also denoted by $\langle \omega, \xi \rangle$ and referred to as the *pairing* of ω and ξ. It is known from linear algebra that the dual space is also a linear space of the same dimension; hence, $\dim T_x^*M = n$.

Definition. The linear space T_x^*M is referred to as the *cotangent space* of M at x. The elements of T_x^*M are called *tangent covectors*.

Note that the dual space to T_x^*M is T_xM, that is, every vector $\xi \in T_xM$ can be regarded as a linear functional on covectors given by $\xi(\omega) = \langle \omega, \xi \rangle$, and all linear functionals on T_x^*M have this form.

Fix a point $x \in M$ and let f be a smooth function in a neighborhood of x.

Definition. Define the notion of the *differential* df at x as follows: df is a tangent covector given by its values on tangent vectors as follows:

$$\langle df, \xi \rangle := \xi(f) = \frac{\partial f}{\partial \xi} \text{ for any } \xi \in T_xM. \tag{1.14}$$

27.10.16

Given the local coordinates $x^1, ..., x^n$, we can consider each x^i as a function in the chart. In particular, dx^i is a tangent covector. Let us verify that $\{dx^i\}$ is a basis in T_x^*M. Indeed, any basis $\{e_1, ..., e_n\}$ in T_xM has a dual basis $\{e^1, ..., e^n\}$ in the dual space T_x^*M that is defined by

$$\langle e^i, e_j \rangle = \delta_j^i := \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases}$$

Since $\left\{\frac{\partial}{\partial x^i}\right\}$ is a basis in T_xM and

$$\langle dx^i, \frac{\partial}{\partial x^j} \rangle = \frac{\partial}{\partial x^j} x^i = \delta_j^i,$$

it follows that $\{dx^i\}$ is the dual basis in T_x^*M. Consequently, any tangent covector $\omega \in T_x^*M$ has an expansion in this basis:

$$\omega = \omega_i dx^i,$$

where ω_i are called the components of ω. Hence, for any $\xi \in T_xM$, we obtain

$$\langle \omega, \xi \rangle = \omega_i \xi^i.$$

In particular, applying this with $\xi = \frac{\partial}{\partial x^i}$, we obtain

$$\omega_i = \langle \omega, \frac{\partial}{\partial x^i} \rangle.$$

For example, for the covector df we obtain from (1.14) that

$$(df)_i = \frac{\partial f}{\partial x^i}$$

and, hence,

$$\boxed{df = \frac{\partial f}{\partial x^i} dx^i.}$$

A *vector field* on a smooth manifold M is a family $\{\xi(x)\}_{x \in M}$ of tangent vectors such that $\xi(x) \in T_xM$ for any $x \in M$. In the local coordinates $x^1, ..., x^n$, it can be represented in the form

$$\xi(x) = \xi^i(x) \frac{\partial}{\partial x^i}.$$

The vector field $\xi(x)$ is called *smooth* if all the functions $\xi^i(x)$ are smooth in any chart. Similarly one defined a covector field.

1.4 Riemannian metric

Let M be a smooth n-dimensional manifold. A *Riemannian metric* (or a *metric tensor*) on M is a family $\mathbf{g} = \{\mathbf{g}(x)\}_{x \in M}$ such that, for any $x \in M$, $\mathbf{g}(x)$ is a symmetric, positive definite, bilinear form on the tangent space T_xM, smoothly depending on $x \in M$.

Using the metric tensor, one defines an inner product $\langle \cdot, \cdot \rangle_{\mathbf{g}}$ in any tangent space T_xM by

$$\langle \xi, \eta \rangle_{\mathbf{g}} := \mathbf{g}(x)(\xi, \eta),$$

for all tangent vectors $\xi, \eta \in T_xM$. Hence, T_xM becomes a Euclidean space. In the local coordinates $x^1, ..., x^n$, we have

$$\langle \xi, \eta \rangle_{\mathbf{g}} = \left\langle \xi^i \frac{\partial}{\partial x^i}, \eta^j \frac{\partial}{\partial x^j} \right\rangle_{\mathbf{g}} = g_{ij}(x) \xi^i \eta^j$$

where

$$\boxed{g_{ij}(x) = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle_{\mathbf{g}}}. \tag{1.15}$$

Clearly, $(g_{ij}(x))_{i,j=1}^n$ is a symmetric positive definite $n \times n$ matrix. The functions $g_{ij}(x)$ are called the *components* of the metric tensor $\mathbf{g}$ in the coordinates $x^1, ..., x^n$. The condition that $\mathbf{g}(x)$ smoothly depends on x means that all the components $g_{ij}(x)$ are C^∞-functions in the all charts.

The metric tensor can be represented in the following form:

$$\boxed{\mathbf{g} = g_{ij} dx^i dx^j}, \tag{1.16}$$

where $dx^i dx^j$ stands for the *tensor product* of the covectors dx^i and dx^j sometimes also denoted by $dx^i \otimes dx^j$; the latter is a bilinear functional on $T_x M$ defined by

$$dx^i dx^j (\xi, \eta) = \langle dx^i, \xi \rangle \langle dx^j, \eta \rangle \quad \forall \xi, \eta \in T_x M,$$

where $\langle \cdot, \cdot \rangle$ is the pairing of covectors and vectors (note that the tensor product is not commutative). Indeed, since

$$\langle dx^i, \xi \rangle = \xi \left(x^i \right) = \xi^i,$$

we obtain

$$g_{ij} dx^i dx^j (\xi, \eta) = g_{ij} \xi^i \eta^j = \mathbf{g}(\xi, \eta),$$

which proves (1.16).

Definition. A *Riemannian manifold* is a couple $(M, \mathbf{g})$ where $\mathbf{g}$ is a Riemannian metric on a smooth manifold M.

A trivial example of a Riemannian manifold is $\mathbb{R}^n$ with the *canonical Euclidean metric* $\mathbf{g}_{\mathbb{R}^n}$ defined in the Cartesian coordinates $x^1, ..., x^n$ by

$$\mathbf{g}_{\mathbb{R}^n} = \left(dx^1 \right)^2 + ... + (dx^n)^2 .$$

For this metric, we have $(g_{ij}) = \mathrm{id}$.

Let $(M, \mathbf{g})$ be a Riemannian manifold. The metric tensor $\mathbf{g}$ can be regarded as a linear mapping from $T_x M$ to $T_x^* M$. Indeed, for any vector $\xi \in T_x M$, define $\mathbf{g}(x) \xi \in T_x^* M$ by the identity

$$\langle \mathbf{g}(x) \xi, \eta \rangle = \langle \xi, \eta \rangle_{\mathbf{g}} \text{ for all } \eta \in T_x M, \tag{1.17}$$

Observe that if $\xi \neq 0$ then also $\mathbf{g}(x) \xi \neq 0$ because $\langle \mathbf{g}(x) \xi, \xi \rangle > 0$.Therefore, the mapping

$$\mathbf{g}(x) : T_x M \to T_x^* M \tag{1.18}$$

is injective and, hence, also bijective.

Rewriting (1.17) in the local coordinates, we obtain

$$(\mathbf{g}(x) \xi)_j \eta^j = g_{ij} \xi^i \eta^j .$$

which implies

$$(\mathbf{g}(x) \xi)_j = g_{ij} \xi^i .$$

In particular, the components of the linear operator $\mathbf{g}(x)$ are g_{ij} – the same as the components of the metric tensor.

If the Riemannian metric $\mathbf{g}$ is fixed then it is customary to drop $\mathbf{g}$ from all the notations. For example, the notation of the inner product of two tangent vectors ξ, η becomes $\langle \xi, \eta \rangle$. Moreover, the notation for the covector $\mathbf{g}(x)\xi$ becomes just ξ; that is, the same as for the vector. However, the notation ξ^i is still used to denote the components of the vector ξ in the basis $\left\{\frac{\partial}{\partial x^i}\right\}$, while ξ_j will be used to denote the components of the covector ξ in the basis $\{dx^j\}$. The relation between the vector components ξ^i and the covector components ξ_j is given then by

$$\xi_j := (\mathbf{g}(x)\xi)_j = g_{ij}\xi^i.$$

The operation of passing from ξ^i to ξ_j is called *lowering* the index.

Since the mapping (1.18) is a bijection, it has the inverse mapping

$$\mathbf{g}^{-1}(x) : T_x^*M \to T_xM.$$

Since $\mathbf{g}^{-1}(x)$ is linear, it has in the coordinates the following form: for any covector $\omega \in T_x^*M$,

$$\left(\mathbf{g}^{-1}(x)\omega\right)^i = g^{ij}\omega_j, \tag{1.19}$$

where the coefficients g^{ij} are called the components of $\mathbf{g}^{-1}(x)$. Since $\mathbf{g}^{-1}$ is inverse to $\mathbf{g}$, we see that the matrix (g^{ij}) is inverse to (g_{ij}), that is

$$\boxed{(g^{ij}) = (g_{ij})^{-1}}.$$

Denoting the vector $\mathbf{g}^{-1}(x)\omega$ also by ω, we obtain the following relation between the vector and covector components of ω:

$$\omega^i := \left(\mathbf{g}^{-1}(x)\omega\right)^i = g^{ij}\omega_j.$$

The operation of passing from ω_j to ω^i is called *raising* the index.

The operator $\mathbf{g}^{-1}(x)$ determines an inner product in T_x^*M as follows: for all $\upsilon, \omega \in T_x^*M$, set

$$\langle \upsilon, \omega \rangle_{\mathbf{g}^{-1}} := \langle \mathbf{g}^{-1}(x)\upsilon, \mathbf{g}^{-1}(x)\omega \rangle_{\mathbf{g}} = \langle \upsilon, \mathbf{g}^{-1}(x)\omega \rangle. \tag{1.20}$$

It follows that, in the local coordinates,

$$\langle \upsilon, \omega \rangle_{\mathbf{g}^{-1}} = \upsilon_i \left(\mathbf{g}^{-1}(x)\omega\right)^i = g^{ij}\upsilon_i\omega_j.$$

For any smooth function f on M, define its *gradient* $\nabla f(x)$ at a point $x \in M$ by

$$\nabla f(x) = \mathbf{g}^{-1}(x)\, df(x) \tag{1.21}$$

that is, $\nabla f(x)$ is a vector that is obtained from the covector $df(x)$ by raising the index. Applying (1.17) with $\xi = \nabla f(x)$, we obtain, for any $\eta \in T_xM$,

$$\boxed{\langle \nabla f, \eta \rangle_{\mathbf{g}} = \langle df, \eta \rangle = \frac{\partial f}{\partial \eta}}, \tag{1.22}$$

which can be considered as an alternative definition of the gradient. In the local coordinates $x^1, ..., x^n$, we obtain by (1.19) and (1.21)

$$\boxed{(\nabla f)^i = g^{ij}\frac{\partial f}{\partial x^j}}. \tag{1.23}$$

If h is another smooth function on M then we obtain from (1.20)

$$\langle \nabla f, \nabla h \rangle_{\mathbf{g}} = \langle \mathbf{g}^{-1}(x)\, df, \mathbf{g}^{-1}(x)\, dh \rangle_{\mathbf{g}} = \langle df, dh \rangle_{\mathbf{g}^{-1}}. \tag{1.24}$$

1.5 Submanifolds

If M is a smooth manifold then any open subset $\Omega \subset M$ trivially becomes a smooth manifold by restricting all charts to Ω. Also, if $\mathbf{g}$ is a Riemannian metric on M then $\mathbf{g}|_\Omega$ is a Riemannian metric on Ω. Hence, any open subset Ω of M can be considered as a (Riemannian) submanifold of a (Riemannian) manifold M of the same dimension.

Consider a more interesting notion of a submanifold of smaller dimension. Any subset S of a smooth manifold M can be regarded as a topological space with induced topology. It is easy to see that S inherits from M the properties of being Hausdorff and having a countable base.

Definition. A set $S \subset M$ is called an (embedded) *submanifold* of dimension m if, for any point $x_0 \in S$, there is a chart (U, φ) on M covering x_0 such that the intersection $S \cap U$ is given in U by the system of equations

$$x^{m+1} = x^{m+2} = ... = x^n = 0,$$

where $x^1, x^2, ..., x^n$ are the local coordinates in U (see Fig. 1.4).

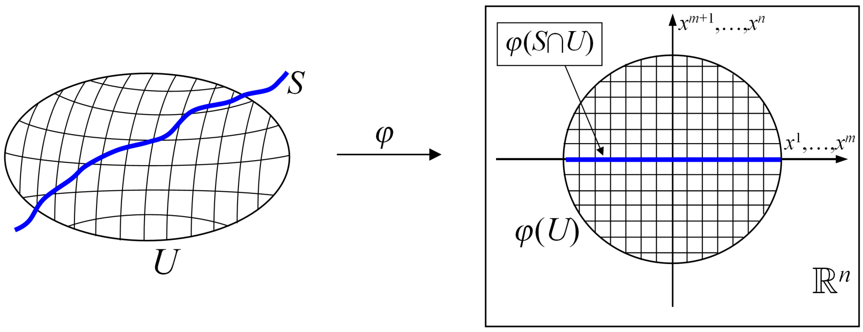

Figure 1.4: The image $\varphi(S \cap U)$

More precisely, this means the following:

$$\begin{aligned} \varphi(S \cap U) &= \left\{x \in \varphi(U) : x^{m+1} = ... = x^n = 0\right\} \\ &= \varphi(U) \cap \mathbb{R}^m, \end{aligned}$$

where we identify $\mathbb{R}^m$ with a subspace of $\mathbb{R}^n$ as follows:

$$\mathbb{R}^m = \left\{x \in \mathbb{R}^n : x^{m+1} = ... = x^n = 0\right\}.$$

Hence, $\varphi|_{S\cap U}$ can be considered as a mapping from $S \cap U$ to $\mathbb{R}^m$, and this mapping is an homeomorphism between $S \cap U$ and the open set $\varphi(U) \cap \mathbb{R}^m$. Hence, $(S \cap U, \varphi|_{S\cap U})$ is a m-dimensional chart on S, with the local coordinates $x^1, x^2, ..., x^m$. With the atlas consisting of all such charts, *the submanifold S becomes a smooth m-dimensional manifold.*

Lemma 1.6 *Let M be a smooth manifold of dimension n and $F : M \to \mathbb{R}$ be a smooth function on M. Consider the null set of F, that is*

$$S = \{x \in M : F(x) = 0\}.$$

If
$$dF \neq 0 \text{ on } S \tag{1.25}$$
then S is a submanifold of dimension $n-1$.

Proof. For any point $x_0 \in S$, there is a chart U on M containing x_0 and such that $dF \neq 0$ in U. The latter means that the row-vector $\left(\frac{\partial F}{\partial x^i}\right)$ does not vanish in U. Without loss of generality we can assume that the last component $\frac{\partial F}{\partial x^n}$ does not vanish in U. By the implicit function theorem, there exists an open set $V \subset U$, containing x_0, such that the equation $F(x) = 0$ in V can be resolved with respect to the coordinate x^n; that is, the equation $F(x) = 0$ is equivalent in V to
$$x^n = f\left(x^1, ..., x^{n-1}\right),$$
where f is a smooth function. After the change of coordinates
$$\begin{aligned} y^1 &= x^1, \\ &\dots \\ y^{n-1} &= x^{n-1}, \\ y^n &= x^n - f\left(x^1, ..., x^{n-1}\right), \end{aligned}$$
the equation of S in V becomes $y^n = 0$ and, hence, S is a $(n-1)$-dimensional submanifold. ■

Let S be a submanifold of M of dimension m and ξ be an $\mathbb{R}$-differentiation on S at a point $x_0 \in S$. For any smooth function f on M, its restriction $f|_S$ is a smooth function on S. Hence, setting
$$\xi(f) := \xi(f|_S), \tag{1.26}$$
we see that ξ can be extended to an $\mathbb{R}$-differentiation on M at the same point x_0. In other words, (1.26) defines a linear mapping
$$T_{x_0}S \to T_{x_0}M. \tag{1.27}$$
Observe that this mapping is injective. Indeed, if $\xi \in T_{x_0}S$ is non-zero then there exists a smooth function $h \in C^\infty(S)$ such that $\xi(h) \neq 0$. In the coordinate system $x^1, ..., x^n$ that is used in the definition of a submanifold, the function h depends on $x^1, ..., x^m$. Setting
$$f\left(x^1, ...x^m, ..., x^n\right) = h\left(x^1, ..., x^m\right),$$
we obtain a smooth function f in a neighborhood of x_0 in M, such that $f|_S = h$. Therefore, for the extension of ξ to $T_{x_0}M$ we have
$$\xi(f) = \xi(f|_S) = \xi(h) \neq 0,$$
that is, ξ is non-zero as element of $T_{x_0}M$. Therefore, the mapping (1.27) is injective, and (1.26) provides a natural identification of $T_{x_0}S$ as a subspace of $T_{x_0}M$.

Let $x^1, ..., x^n$ be local coordinates in M and $y^1, ..., y^m$ be local coordinates on S. Assume that in the intersection of the domains of these coordinate systems we have the relations
$$x^i = x^i\left(y^1, ..., y^m\right), \quad i = 1, ..., n.$$

Let x_0 be a point on S that lies in the intersection of the domains of the both coordinate systems. For any smooth function f in a neighborhood of x_0, we have by the chain rule

$$\frac{\partial f}{\partial y^k} = \frac{\partial x^i}{\partial y^k}\frac{\partial f}{\partial x^i}$$

that is

$$\boxed{\frac{\partial}{\partial y^k} = \frac{\partial x^i}{\partial y^k}\frac{\partial}{\partial x^i}}. \tag{1.28}$$

Note that $\left\{\frac{\partial}{\partial x^i}\right\}$ is a basis in $T_{x_0}M$ and $\left\{\frac{\partial}{\partial y^k}\right\}$ is a basis in $T_{x_0}S$, so that (1.28) identifies explicitly $T_{x_0}S$ as a subspace of $T_{x_0}M$.

Any tangent covector $\omega \in T^*_{x_0}M$ as a linear functional on $T_{x_0}M$ can be restricted to the subspace $T_{x_0}S$ thus yielding an element of $T^*_{x_0}S$ that also will be denoted by ω. Let us compute $dx^i|_{T_{x_0}S}$ in the basis dy^j. Since by (1.28)

$$\langle dx^i, \frac{\partial}{\partial y^j}\rangle = \langle dx^i, \frac{\partial x^l}{\partial y^j}\frac{\partial}{\partial x^l}\rangle = \frac{\partial x^l}{\partial y^j}\delta^i_l = \frac{\partial x^i}{\partial y^j},$$

it follows that the restriction of dx^i to $T_{x_0}S$ is given by

$$\boxed{dx^i = \frac{\partial x^i}{\partial y^j}dy^j}. \tag{1.29}$$

Let $\mathbf{g}$ be a Riemannian metric on M. As any point $x \in S$, we can restrict $\mathbf{g}(x)$ to a bilinear functional on T_xS thus obtaining a Riemannian metric $\mathbf{g}_S$ on S. The metric $\mathbf{g}_S$ is called the *induced metric.*

In the local coordinates as above, by restricting $\mathbf{g} = g_{kl}dx^k dx^l$ to $T_{x_0}S$, we obtain by (1.29)

$$\mathbf{g}_S = g_{kl}dx^k dx^l = \left(\frac{\partial x^k}{\partial y^i}dy^i\right)\left(\frac{\partial x^l}{\partial y^j}dy^j\right) = g_{kl}\frac{\partial x^k}{\partial y^i}\frac{\partial x^l}{\partial y^j}dy^i dy^j$$

that is,

$$\mathbf{g}_S = (g_S)_{ij}\, dy^i dy^j$$

where

$$\boxed{(g_S)_{ij} = g_{kl}\frac{\partial x^k}{\partial y^i}\frac{\partial x^l}{\partial y^j}}. \tag{1.30}$$

Alternatively, (1.30) can be proved using (1.28) as follows: 03.11.16

$$(g_S)_{ij} = \left\langle \frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}\right\rangle_{\mathbf{g}} = \left\langle \frac{\partial x^k}{\partial y^i}\frac{\partial}{\partial x^k}, \frac{\partial x^l}{\partial y^j}\frac{\partial}{\partial x^l}\right\rangle_{\mathbf{g}} = \frac{\partial x^k}{\partial y^i}\frac{\partial x^l}{\partial y^j}\left\langle \frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\right\rangle_{\mathbf{g}} = \frac{\partial x^k}{\partial y^i}\frac{\partial x^l}{\partial y^j}g_{kl}.$$

The identity (1.30) can also be presented in the matrix form as follows. Consider the Jacobi matrix of the change $x = x(y)$

$$J = (J_{ki}) = \left(\frac{\partial x^k}{\partial y^i}\right), \tag{1.31}$$

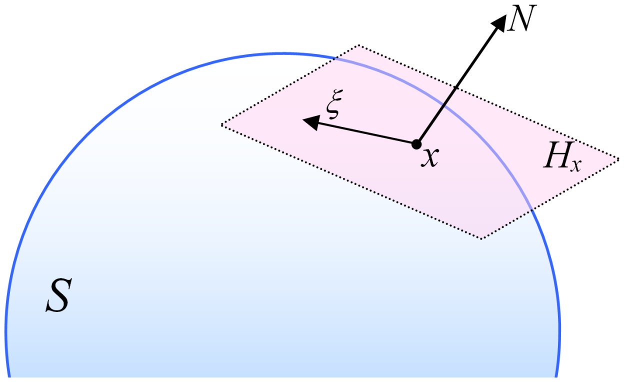

Figure 1.5: The normal N to the hypersurface S and the tangent hyperplane $H_x \cong T_x S$

where $k = 1, ..., n$ is the row index and $i = 1, ..., m$ is the column index, so that J is an $n \times m$ matrix. Then (1.30) is equivalent to

$$(g_S)_{ij} = J_{ki} g_{kl} J_{lj} = J^T_{ik} g_{kl} J_{lj} = \left(J^T g J\right)_{ij}$$

that is,

$$\boxed{g_S = J^T g J.} \tag{1.32}$$

Note that the right hand side of (1.32) is the product of the three matrices of the following dimensions: $m \times n$, $n \times n$, $n \times m$.

In a particular case $m = n$, the formula (1.32) was stated in Exercise 3. Indeed, in this case S is an open subset of M and the induced metric $\mathbf{g}_S$ coincides with the original metric $\mathbf{g}$, so that (1.32) provides simply the relation between the matrices of $\mathbf{g}$ in the two coordinate systems.

Example. Consider a smooth function $F : \mathbb{R}^n \to \mathbb{R}$ and its null set

$$S = \{x \in \mathbb{R}^n : F(x) = 0\}.$$

As in Lemma 1.6, assume that F is non-singular on S, that is, $dF \neq 0$ on S. Since $(dF)_i = \frac{\partial F}{\partial x^i}$, the latter condition means that at any point $x \in S$ at least one of the values $\frac{\partial F}{\partial x^i}(x)$, $i = 1, ..., n$, does not vanish. Fix a point $x \in S$. By Exercise 9, the tangent vector $\xi \in T_x\mathbb{R}^n$ belongs to $T_x S$ if and only if

$$\langle dF, \xi \rangle = 0,$$

that is, if at x

$$\frac{\partial F}{\partial x^1}\xi^1 + ... + \frac{\partial F}{\partial x^n}\xi^n = 0.$$

Geometrically this means that ξ is orthogonal to the vector $N = \left(\frac{\partial F}{\partial x^1}, ..., \frac{\partial F}{\partial x^n}\right)$. On the other hand, the vector N is known to be the normal of the tangent hyperplane H_x to S at x. Hence, ξ is identified as an element of H_x and the tangent space $T_x S$ is naturally identified with H_x (see Fig. 1.5).

Example. Consider in $\mathbb{R}^{n+1}$ the following equation

$$\left(x^1\right)^2 + ... + \left(x^{n+1}\right)^2 = 1,$$

which defines the unit sphere $\mathbb{S}^n$. Since $\mathbb{S}^n$ is the null set of the function

$$F(x) = \left(x^1\right)^2 + ... + \left(x^{n+1}\right)^2 - 1,$$

whose differential $dF = (2x^1, ..., 2x^{n+1})$ does not vanish on $\mathbb{S}^n$, we conclude that $\mathbb{S}^n$ is a submanifold of $\mathbb{R}^{n+1}$ of dimension n. Furthermore, considering $\mathbb{R}^{n+1}$ as a Riemannian manifold with the canonical Euclidean metric $\mathbf{g}_{\mathbb{R}^{n+1}}$, we see that $\mathbb{S}^n$ can be regarded as Riemannian manifold with the induced metric that is called the *canonical spherical metric* and is denoted by $\mathbf{g}_{\mathbb{S}^n}$.

Let us compute $\mathbf{g}_{\mathbb{S}^1}$ using the following chart on $\mathbb{S}^1$ (see also Exercise 11). The upper semi-circle $U := \mathbb{S}^1 \cap \{x^2 > 0\}$ is the graph of a function $f(t) = \sqrt{1-t^2}$ on the interval $(-1,1)$. The mapping $\varphi : U \to (-1,1)$ given by $\varphi(t, f(t)) = t$ provides homeomorphism of U onto the open interval $(-1,1)$, which means that (U, φ) is a chart[2] on $\mathbb{S}^1$. This chart has just one coordinate $y^1 = t$, and the metric $\mathbf{g}_{\mathbb{S}^1}$ has the form

$$\mathbf{g}_{\mathbb{S}^1} = (g_{\mathbb{S}^1})_{11} \left(dy^1\right)^2,$$

where by (1.30)

$$(g_{\mathbb{S}^1})_{11} = (g_{\mathbb{R}^2})_{kl} \frac{\partial x^k}{\partial y^1} \frac{\partial x^l}{\partial y^1} = \left(\frac{\partial x^1}{\partial y^1}\right)^2 + \left(\frac{\partial x^2}{\partial y^1}\right)^2.$$

Since $x^1 = y^1$ and $x^2 = \sqrt{1 - (y^1)^2}$, it follows that

$$(g_{\mathbb{S}^1})_{11} = 1 + \frac{(y^1)^2}{1-(y^1)^2} = \frac{1}{1-(y^1)^2}.$$

Hence,

$$\mathbf{g}_{\mathbb{S}^1} = \frac{(dy^1)^2}{1-(y^1)^2}.$$

1.6 Riemannian measure

Let us recall the definition of the notion of measure. Let X be an arbitrary set. A σ-algebra $\mathcal{A}$ on X is a subset of 2^X such that $\mathcal{A}$ contains $\emptyset$, M and $\mathcal{A}$ is closed under taking complement and countable unions (hence, also intersections). A measure μ on a σ-algebra $\mathcal{A}$ is a mapping $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(\emptyset) = 0$ and μ is σ-additive, that is, $\mu\left(\bigsqcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu(A_i)$ for all $A_i \in \mathcal{A}$. Given a measure μ, one can define the notion of the integral $\int_X f d\mu$ for a class of measurable functions.

The most famous example of a measure is the Lebesgue measure λ defined on the σ-algebra $\mathcal{L}(\mathbb{R}^n)$ of Lebesgue measurable subsets of $\mathbb{R}^n$. Recall that the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ is defined as the minimal σ-algebra containing all open subsets of $\mathbb{R}^n$, and the elements of $\mathcal{B}(\mathbb{R}^n)$ are called Borel sets. It is known that any Lebesgue measurable set is a union of a Borel set and a null set (=a set of measure zero).

[2]By rotating the Cartesian coordinate system in $\mathbb{R}^2$, we obtain that any semi-circle is a chart, and such charts cover all $\mathbb{S}^1$.

Let M be a smooth manifold of dimension n. Denote by $\mathcal{B}(M)$ the smallest σ-algebra containing all open sets in M. The elements of $\mathcal{B}(M)$ are called *Borel sets*. We say that a set $E \subset M$ is *measurable* if, for any chart U, the intersection $E \cap U$ is Lebesgue measurable in U. Obviously, the family of all measurable sets in M forms a σ-algebra, that will be denoted by $\mathcal{L}(M)$. Since any open subset of M is measurable, it follows that also all Borel sets are measurable, that is, $\mathcal{B}(M) \subset \mathcal{L}(M)$.

The purpose of this section is to show that any Riemannian manifold $(M, \mathbf{g})$ features a canonical measure ν that is defined on $\mathcal{L}(M)$ and that is called the *Riemannian measure (or volume)*. This measure is defined by means of the following theorem.

Theorem 1.7 *For any Riemannian manifold* $(M, \mathbf{g})$, *there exists a unique measure* ν *on* $\mathcal{L}(M)$ *such that, in any chart* U,

$$d\nu = \sqrt{\det g}\, d\lambda, \tag{1.33}$$

where $g = (g_{ij})$ *is the matrix of the Riemannian metric* $\mathbf{g}$ *in* U *and* λ *is the Lebesgue measure in* U.

Furthermore, the measure ν *has the following properties:* $\nu(K) < \infty$ *for any compact set* $K \subset M$ *and* $\nu(\Omega) > 0$ *for any non-empty open set* $\Omega \subset M$.

Note that $\det g > 0$ by the positive definiteness of g. The condition (1.33) implies that, for any non-negative measurable function f on U,

$$\boxed{\int_U f\, d\nu = \int_U f \sqrt{\det g}\, d\lambda}.$$

Proof. We need to construct measure ν with the domain $\mathcal{L}(M)$ such that, for any chart U and for any measurable set $A \subset U$,

$$\nu(A) = \int_A \sqrt{\det g}\, d\lambda. \tag{1.34}$$

Let us use (1.34) as definition of ν on the σ-algebra $\mathcal{L}(U)$ of Lebesgue measurable sets in U. Our task is to show that the measure ν defined by (1.34) in each chart, can be extended to $\mathcal{L}(M)$ and, moreover, this extension is unique. However, before that, we need to ensure that the measures in different charts agree on their intersection.

Claim. *If* U *and* V *are two charts on* M *and* A *is a measurable set in* $W := U \cap V$ *then* $\nu(A)$ *defined by* (1.34) *has the same values in the both charts.*

Let $x^1, ..., x^n$ and $y^1, ..., y^n$ be the local coordinate systems in U and V, respectively. Denote by g^x and g^y the matrices of $\mathbf{g}$ in the coordinates $x^1, ..., x^n$ and $y^1, ..., y^n$, respectively. We need to show that, for any measurable set $A \subset W$,

$$\int_A \sqrt{\det g^x} dx = \int_A \sqrt{\det g^y} dy,$$

where dx and dy stand for the Lebesgue measures in U and V, respectively.

Let J be the Jacobi matrix of the change $x = x(y)$, that is, $J = \left(\frac{\partial x^k}{\partial y^i}\right)$ (cf. (1.31)). By (1.32) we have

$$g^y = J^T g^x J,$$

which implies

$$\det g^y = \det J^T \det g^x \det J = \det g^x (\det J)^2 . \tag{1.35}$$

Next, let us use the following formula for change of variables in the Lebesgue integral in $\mathbb{R}^n$: if f is a non-negative measurable function in W then

$$\int_W f(x)\, dx = \int_W f(x(y)) |\det J|\, dy. \tag{1.36}$$

Applying this for $f = 1_A \sqrt{\det g^x}$ and using (1.35), we obtain

$$\int_A \sqrt{\det g^x} dx = \int_A \sqrt{\det g^x}\, |\det J|\, dy = \int_A \sqrt{\det g^x (\det J)^2} dy = \int_A \sqrt{\det g^y} dy,$$

which proves the claim. 08.11.16

Now let us prove the uniqueness of measure ν on $\mathcal{L}(M)$ that satisfies (1.33) in all charts. By Lemma 1.1, there is a countable family $\{U_i\}_{i=1}^\infty$ of relatively compact charts covering M and such that $\overline{U}_i$ is contained in a chart. For any measurable set A on M, define the sequence of sets $A_i \subset U_i$ by

$$A_1 = A \cap U_1,\ A_2 = A \cap U_2 \setminus U_1, ...,\ A_i = A \cap U_i \setminus U_1 \setminus ... \setminus U_{i-1}, ... \tag{1.37}$$

(see Fig. 1.6). Then we have $A_i \in \mathcal{L}(U_i)$.

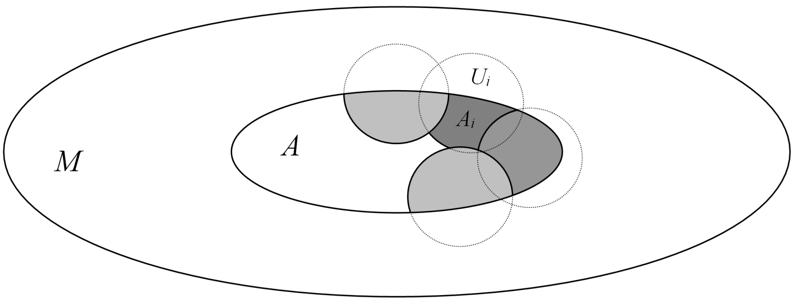

Figure 1.6: Splitting A into disjoint sets A_i.

Clearly,

$$A = \bigsqcup_i A_i,$$

where the sign $\sqcup$ means "disjoint union". Therefore, for any measure ν, we should have

$$\nu(A) = \sum_i \nu(A_i). \tag{1.38}$$

However, the value $\nu(A_i)$ is uniquely defined by (1.33) because A_i is contained in the chart U_i. Hence, $\nu(A)$ is also uniquely defined, which was to be proved.

To prove the existence of ν, we use the same construction: for any measurable set A, define $\nu(A)$ by (1.34), using the fact that $\nu(A_i)$ is already defined. Let us show that ν is a measure, that is, ν is σ-additive. Let $\left\{A^k\right\}_{k=1}^{\infty}$ be a sequence of disjoint measurable sets in M and let

$$A=\bigsqcup_k A^k.$$

Defining the sets A_i^k similarly to (1.37), that is,

$$A_i^k=A^k\cap U_i\setminus U_1\setminus ...\setminus U_{i-1}$$

we obtain that

$$A^k=\bigsqcup_i A_i^k$$

and

$$A_i=\bigsqcup_k A_i^k,$$

where $A_i^k\in\mathcal{L}\left(U_i\right)$. Since ν is σ-additive in each chart U_i, we obtain

$$\nu\left(A_i\right)=\sum_k \nu(A_i^k).$$

Adding up in i and interchanging the summation in i and k, we obtain

$$\nu\left(A\right)=\sum_i \nu\left(A_i\right)=\sum_i\sum_k \nu(A_i^k)=\sum_k\sum_i \nu(A_i^k)=\sum_k \nu\left(A^k\right),$$

which proves (1.38).

Any compact set $K\subset M$ can covered by a finite number of charts U_i. Applying (1.34) in a chart containing $\overline{U}_i$ and noticing $\sqrt{\det g}$ is bounded on $\overline{U}_i$, we obtain $\nu\left(U_i\right)<\infty$, which implies $\nu\left(K\right)<\infty$.

Any non-empty open set $\Omega\subset M$ contains some non-empty chart U, whence it follows from (1.34) that

$$\nu\left(\Omega\right)\geq\nu\left(U\right)=\int_U\sqrt{\det g}d\lambda>0.$$

■

The extension of measure ν from the charts to the whole manifold can also be done using the Carathéodory extension of measures. Consider the following family of subsets of M:

$$S=\left\{A\subset M: A\text{ is a relatively compact measurable set and }\overline{A}\text{ is contained in a chart}\right\}.$$

Observe that S is a semi-ring and, by the above Claim, ν is defined as a measure on S. Hence, the Carathéodory extension of ν exists and is a complete measure on M. It is not difficult to check that the domain of this measure is exactly $\mathcal{L}\left(M\right)$. Since the union of sets U_i from Lemma 1.1 is M and $\nu\left(U_i\right)<\infty$, the measure ν on S is σ-finite and, hence, its extension to $\mathcal{L}\left(M\right)$ is unique.

Since the Riemannian measure ν is finite on compact sets, any continuous function with compact support is integrable against ν. Let us record the following simple property of measure ν, which will be used in the next section.

Lemma 1.8 *If $f\in C\left(M\right)$ and*

$$\int_M f\varphi d\nu=0 \tag{1.39}$$

for all $\varphi\in C_0^\infty\left(M\right)$ then $f\equiv 0$.

Proof. See Exercise 8. ■

1.7 Divergence theorem

Recall that the divergence of a smooth vector field $v(x)$ in $\mathbb{R}^n$ (or in a domain in $\mathbb{R}^n$) is a function defined by

$$\operatorname{div} v(x) = \sum_{i=1}^{n} \frac{\partial v^i}{\partial x^i}.$$

The divergence theorem of Gauss says, in particular, that is supp v is compact then

$$\int_{\mathbb{R}^n} \operatorname{div} v \, dx = 0.$$

In fact, in this simple form, this theorem is a consequence of the integration by part formula and Fubini's theorem. For a general smooth vector field v and a smooth scalar function u with compact support in $\mathbb{R}^n$, we can apply this theorem to the vector field uv and obtain that

$$\int_{\mathbb{R}^n} (\operatorname{div} v)\, u \, dx = -\int_{\mathbb{R}^n} v \cdot \nabla u \, dx,$$

For any smooth vector field $v(x)$ on a Riemannian manifold $(M, \mathbf{g})$, its *divergence* $\operatorname{div} v(x)$ is a smooth function on M, defined by means of the following statement.

Theorem 1.9 (The divergence theorem) *For any smooth vector field $v(x)$ on a Riemannian manifold $(M, \mathbf{g})$, there exists a unique smooth function on M, denoted by $\operatorname{div} v$, such that the following identity holds*

$$\int_M (\operatorname{div} v)\ u \, d\nu = -\int_M \langle v, \nabla u\rangle d\nu, \tag{1.40}$$

for all $u \in C_0^\infty(M)$.

Both gradient ∇ and divergence div depend on the metric $\mathbf{g}$. In the cases when this dependence should be emphasized, we will use the extended notations $\nabla_{\mathbf{g}}$ and $\operatorname{div}_{\mathbf{g}}$.

The expression $\langle v, \nabla u\rangle = \langle v, \nabla u\rangle_{\mathbf{g}}$ is the inner product of the tangent vectors v and ∇u. By (1.22), we have

$$\langle v, \nabla u\rangle_{\mathbf{g}} = \langle \nabla u, v\rangle_{\mathbf{g}} = \langle du, v\rangle = \frac{\partial u}{\partial x^i} v^i,$$

where $\langle du, v\rangle$ is the pairing of the tangent covector du and vector v.

Proof. The uniqueness of $\operatorname{div} v$ is simple: if there are two candidates for $\operatorname{div} v$, say $(\operatorname{div} v)'$ and $(\operatorname{div} v)''$ then, for all $u \in C_0^\infty(M)$,

$$\int_M (\operatorname{div} v)' \, u \, d\nu = \int_M (\operatorname{div} v)'' \, u \, d\nu,$$

which implies $(\operatorname{div} v)' = (\operatorname{div} v)''$ by Lemma 1.8.

To prove the existence of $\operatorname{div} v$, let us first show that $\operatorname{div} v$ exists in any chart. Namely, if U is a chart on M with the coordinates $x^1, ..., x^n$ then, using (1.22), (1.33), and the integration-by-parts formula in U as a subset of $\mathbb{R}^n$, we obtain, for any $u \in C_0^\infty(U)$,

$$\begin{aligned}\int_U \langle v, \nabla u\rangle d\nu &= \int_U \langle du, v\rangle d\nu \\ &= \int_U \frac{\partial u}{\partial x^i} v^i \sqrt{\det g} d\lambda \\ &= -\int_U \frac{\partial}{\partial x^i}\left(v^i\sqrt{\det g}\right) u\, d\lambda \\ &= -\int_U \frac{1}{\sqrt{\det g}}\frac{\partial}{\partial x^i}\left(v^i\sqrt{\det g}\right) u\, d\nu. \end{aligned} \tag{1.41}$$

Comparing with (1.40) we see that the divergence in U can be defined by

$$\boxed{\operatorname{div} v = \frac{1}{\sqrt{\det g}}\frac{\partial}{\partial x^i}\left(\sqrt{\det g} v^i\right)}. \tag{1.42}$$

If U and V are two charts then (1.42) defines the divergences in U and in V, which agree in $U \cap V$ by the uniqueness statement. Hence, (1.42) defines $\operatorname{div} v$ as a function on the entire manifold M. Moreover, the divergence defined in this way satisfies the identity (1.40) for all test functions u compactly supported in one of the charts.

We are left to extend the identity (1.40) to all functions $u \in C_0^\infty(M)$. Let $\{\Omega_\alpha\}$ be any family of charts covering M. By Corollary 1.4, any function $u \in C_0^\infty(M)$ can be represented as a finite sum $u_1 + ... + u_m$, where each u_j is smooth and compactly supported in one of Ω_α. Hence, (1.40) holds for each of the functions u_j. By adding up all such identities, we obtain (1.40) for the function u. ■

It follows from (1.42) that

$$\operatorname{div} v = \frac{\partial v^i}{\partial x^i} + v^i \frac{\partial}{\partial x^i} \ln \sqrt{\det g}.$$

In particular, if $\det g \equiv 1$ then we obtain the same formula as in $\mathbb{R}^n$: $\operatorname{div} v = \frac{\partial v^i}{\partial x^i}$.

Corollary 1.10 *The identity (1.40) holds also if $u(x)$ is any smooth function on M and $v(x)$ is a compactly supported smooth vector field on M.*

Proof. Let $K = \operatorname{supp} v$. By Theorem 1.3, there exists a cutoff function of K, that is, a function $\varphi \in C_0^\infty(M)$ such that $\varphi \equiv 1$ in a neighbourhood of K. Then $u\varphi \in C_0^\infty(M)$, and we obtain by Theorem 1.9

$$\int_M \operatorname{div} v\, u\, d\nu = \int_M \operatorname{div} v\ (u\varphi)\ d\nu = -\int_M \langle v, \nabla(u\varphi)\rangle d\nu = -\int_M \langle v, \nabla u\rangle d\nu.$$

■

Alternative definition of divergence. Let us define the divergence $\operatorname{div} v$ in any chart by

$$\operatorname{div} v = \frac{1}{\sqrt{\det g}}\frac{\partial}{\partial x^i}\left(\sqrt{\det g} v^i\right), \tag{1.43}$$

and show by a direct computation that, in the intersection of any two charts, (1.43) defines the same function. This approach allows to avoid integration in the definition of divergence but it is more technically involved (besides, we need integration and Theorem 1.9 anyway).

We will use the following formula: if $a = (a_j^i)$ is a non-singular $n \times n$ matrix smoothly depending on a real parameter t and (b_j^i) is its inverse (where i is the row index and j is the column index) then

$$\frac{\partial}{\partial t} \ln \det a = b_k^l \frac{\partial a_l^k}{\partial t}. \tag{1.44}$$

In the common domain of two coordinate systems $x^1, ..., x^n$ and $y^1, ..., y^n$, set

$$J_i^k = \frac{\partial y^k}{\partial x^i} \text{ and } I_k^i = \frac{\partial x^i}{\partial y^k},$$

so that the matrices I and J are mutually inverse. Let g be the matrix of the tensor $\mathbf{g}$ and v^i be the components of the vector v in coordinates $x^1, ..., x^n$, and let $\widetilde{g}$ be the matrix of $\mathbf{g}$ and $\widetilde{v}^k$ be the components of the vector v in coordinates $y^1, ..., y^n$. Then we have

$$v = v^i \frac{\partial}{\partial x^i} = v^i \frac{\partial y^k}{\partial x^i} \frac{\partial}{\partial y^k} = v^i J_i^k \frac{\partial}{\partial y^k}$$

so that

$$\widetilde{v}^k = v^i J_i^k.$$

Since by (1.35)

$$\sqrt{\det \widetilde{g}} = \sqrt{\det g} \, |\det J|^{-1},$$

the divergence of v in the coordinates $y^1, ..., y^n$ is given by

$$\begin{aligned} \operatorname{div} v &= \frac{1}{\sqrt{\det \widetilde{g}}} \frac{\partial}{\partial y^k} \left(\sqrt{\det \widetilde{g}} \widetilde{v}^k \right) = \frac{\det J}{\sqrt{\det g}} I_k^j \frac{\partial}{\partial x^j} \left(\sqrt{\det g} v^i \left(\det J\right)^{-1} J_i^k \right) \\ &= \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^j} \left(\sqrt{\det g} v^i \right) I_k^j J_i^k + v^i I_k^j J_i^k \det J \frac{\partial}{\partial x^j} \left(\det J\right)^{-1} + v^i I_k^j \frac{\partial J_i^k}{\partial x^j} \\ &= \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^i} \left(\sqrt{\det g} v^i \right) - v^i \frac{\partial}{\partial x^i} \ln \det J + v^i I_k^j \frac{\partial J_i^k}{\partial x^j}, \end{aligned}$$

where we have used the fact that the matrices J and I are mutually inverse and, hence, $I_k^j J_i^k = \delta_i^j$. To finish the proof, it suffices to show that, for any index i,

$$-\frac{\partial}{\partial x^i} \ln \det J + I_k^j \frac{\partial J_i^k}{\partial x^j} = 0. \tag{1.45}$$

By (1.44), we have

$$\frac{\partial}{\partial x^i} \ln \det J = I_k^j \frac{\partial J_j^k}{\partial x^i}.$$

Noticing that

$$\frac{\partial J_j^k}{\partial x^i} = \frac{\partial^2 y^k}{\partial x^j \partial x^i} = \frac{\partial^2 y^k}{\partial x^i \partial x^j} = \frac{\partial J_i^k}{\partial x^j},$$

we obtain (1.45).

1.8 Laplace-Beltrami operator

Recall that the Laplace operator in $\mathbb{R}^n$ is given by

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\left(\partial x^i\right)^2}. \tag{1.46}$$

It is also easy to see that

$$\Delta f = \sum_{i=1}^{n} \frac{\partial}{\partial x^i}\left(\frac{\partial f}{\partial x^i}\right) = \operatorname{div}(\nabla f).$$

Having defined gradient and divergence, we can now define the *Laplace-Beltrami operator* (frequently referred to simply as the Laplace operator) on any Riemannian manifold $(M, \mathbf{g})$ as follows:

$$\boxed{\Delta = \operatorname{div} \circ \nabla}.$$

Strictly speaking, one should use the notations $\Delta_{\mathbf{g}}$, $\operatorname{div}_{\mathbf{g}}$ and $\nabla_{\mathbf{g}}$ but the index $\mathbf{g}$ is usually skipped when there is no danger of confusion.

Hence, for any smooth function f on M, we have

$$\Delta f = \operatorname{div}(\nabla f), \tag{1.47}$$

so that Δf is also a smooth function on M. In local coordinates, we have

$$(\nabla f)^i = g^{ij}\frac{\partial f}{\partial x^j},$$

where $(g^{ij}) = (g_{ij})^{-1}$, which yields

$$\boxed{\Delta f = \frac{1}{\sqrt{\det g}}\frac{\partial}{\partial x^i}\left(\sqrt{\det g}\, g^{ij}\frac{\partial f}{\partial x^j}\right)}, \tag{1.48}$$

For example, if $(g_{ij}) \equiv \mathrm{id}$ then also $(g^{ij}) \equiv \mathrm{id}$, and (1.48) takes the form (1.46). Hence, the classical Laplace operator in $\mathbb{R}^n$ is a particular case of the Laplace-Beltrami operator. Since the matrix (g^{ij}) is symmetric and positive definite, the operator Δ in
10.11.16 (1.48) is an elliptic second order operator in the divergence form.

Theorem 1.11 (The Green formula) *If u and v are smooth functions on a Riemannian manifold M and one of them has a compact support then*

$$\int_M u\Delta v\, d\nu = -\int_M \langle \nabla u, \nabla v\rangle d\nu = \int_M v\Delta u\, d\nu. \tag{1.49}$$

Proof. Consider the vector field ∇v. Clearly, $\operatorname{supp}\nabla v \subset \operatorname{supp} v$ so that either $\operatorname{supp} u$ or $\operatorname{supp}\nabla v$ is compact. By Theorem 1.9, Corollary 1.10, and (1.47), we obtain

$$\int_M u\Delta v\, d\nu = \int_M u \operatorname{div}(\nabla v)\, d\nu = -\int_M \langle \nabla u, \nabla v\rangle d\nu.$$

The second identity in (1.49) is proved similarly. ∎

1.9 Weighted manifolds

Any smooth positive function $D(x)$ on a Riemannian manifold $(M, \mathbf{g})$ gives rise to a measure μ on M given by $d\mu = Dd\nu$ and defined on the σ-algebra $\mathcal{L}(M)$. The function D is called the *density function* of the measure μ. For example, the density function of the Riemannian measure ν is 1.

Definition. A triple $(M, \mathbf{g}, \mu)$ is called a *weighted manifold* (or manifold with density) if $(M, \mathbf{g})$ is a Riemannian manifold and μ is a measure on M with a smooth positive density function.

The definition of gradient on a weighted manifold $(M, \mathbf{g}, \mu)$ is the same as on $(M, \mathbf{g})$, but the definition of divergence changes. For any smooth vector field v on M, define its *weighted divergence* $\operatorname{div}_{\mathbf{g},\mu} v$ by

$$\boxed{\operatorname{div}_{\mathbf{g},\mu} v = \frac{1}{D} \operatorname{div}_{\mathbf{g}} (Dv)}.$$

It follows immediately from this definition and (1.40) that the following extension of Theorem 1.9 takes place: for all smooth vector fields v and functions u,

$$\int_M \operatorname{div}_{\mathbf{g},\mu} v \, u \, d\mu = -\int_M \langle v, \nabla u \rangle_{\mathbf{g}} d\mu, \tag{1.50}$$

provided v or u has a compact support.

Define the *weighted Laplace operator* Δ_μ by

$$\boxed{\Delta_{\mathbf{g},\mu} = \operatorname{div}_{\mathbf{g},\mu} \circ \nabla_{\mathbf{g}}},$$

that is,

$$\Delta_{\mathbf{g},\mu} u = \frac{1}{D} \operatorname{div}_{\mathbf{g}} (D \nabla_{\mathbf{g}} u).$$

The Green formulas remain true, that is, if u and v are smooth functions on M and one of them has a compact support then

$$\int_M u \Delta_{\mathbf{g},\mu} v \, d\mu = -\int_M \langle \nabla u, \nabla v \rangle_{\mathbf{g}} d\mu = \int_M v \Delta_{\mathbf{g},\mu} u \, d\mu. \tag{1.51}$$

In the local coordinates $x^1, ..., x^n$, we have

$$\operatorname{div}_{\mathbf{g},\mu} v = \frac{1}{\rho} \frac{\partial}{\partial x^i} \left(\rho v^i \right) \tag{1.52}$$

and

$$\Delta_{\mathbf{g},\mu} = \frac{1}{\rho} \frac{\partial}{\partial x^i} \left(\rho g^{ij} \frac{\partial}{\partial x^j} \right). \tag{1.53}$$

where $\rho = D\sqrt{\det g}$. Note also that $d\mu = \rho d\lambda$, where λ is the Lebesgue measure in U.

Sometimes is it useful to know that the right hand side of (1.53) can be expanded as follows:

$$\Delta_{\mathbf{g},\mu} = g^{ij} \frac{\partial^2}{\partial x^i \partial x^j} + \left(\frac{1}{\rho} \frac{\partial \rho}{\partial x^i} g^{ij} + \frac{\partial g^{ij}}{\partial x^i} \right) \frac{\partial}{\partial x^j}. \tag{1.54}$$

Example. Consider the weighted manifold $(\mathbb{R}, \mathbf{g}, \mu)$ where $\mathbf{g}$ is the canonical Euclidean metric and $d\mu = D d\lambda$. Then by (1.53) or (1.54)

$$\Delta_{\mathbf{g},\mu} f = \frac{1}{D}\frac{d}{dx}\left(D\frac{df}{dx}\right) = f'' + \frac{D'}{D} f'.$$

Let $(M, \mathbf{g}, \mu)$ be a weighted manifold and D be the density function of measure μ. Define the *induced measure* μ_S on a submanifold S by the condition that μ_S has the density function $D|_S$ with respect to the Riemannian measure of $\mathbf{g}_S$. Hence, we obtain the weighted manifold $(S, \mathbf{g}_S, \mu_S)$.

1.10 Product of manifolds

Let X, Y be smooth manifolds of dimensions n and m, respectively, and let $M = X \times Y$ be the direct product of X and Y as topological spaces. The space M consists of the couples (x, y) where $x \in X$ and $y \in Y$, and it can be naturally endowed with a structure of a smooth manifold. Indeed, if U and V are charts on X and Y respectively, with the coordinates $x^1, ..., x^n$ and $y^1, ..., y^m$ then $U \times V$ is a chart on M with the coordinates $x^1, ..., x^n, y^1, ..., y^m$. The atlas of all such charts makes M into a smooth manifold.

We claim that, for any point $(x_0, y_0) \in M$, the tangent space $T_{(x_0,y_0)}M$ is naturally identified as the direct sum $T_{x_0}X \oplus T_{y_0}Y$ of the linear spaces. Indeed, any $\mathbb{R}$-differentiation $\xi \in T_{x_0}X$ can be considered as an $\mathbb{R}$-differentiation on functions $f(x, y)$ on M by freezing the variable $y = y_0$, that is

$$\xi(f) = \xi(f(\cdot, y_0)).$$

This identifies $T_{x_0}X$ as a subspace of $T_{(x_0,y_0)}M$, and the same applied to $T_{y_0}Y$.

Any $\xi \in T_{(x_0,y_0)}M$ has an expansion in the basis $\{\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}, \frac{\partial}{\partial y^1}, ..., \frac{\partial}{\partial y^m}\}$ as follows:

$$\xi = \sum_{i=1}^{n} \xi^i \frac{\partial}{\partial x^i} + \sum_{j=1}^{m} \xi^{j+n} \frac{\partial}{\partial y^j}.$$

Since

$$\xi_X := \xi^i \frac{\partial}{\partial x^i} \in T_{x_0}X \quad \text{and} \quad \xi_Y := \xi^{j+n}\frac{\partial}{\partial y^j} \in T_{y_0}Y,$$

we see that any $\xi \in T_{(x_0,y_0)}M$ splits in a unique way into the sum

$$\xi = \xi_X + \xi_Y \ \text{ where } \xi_X \in T_{x_0}X \text{ and } \xi_Y \in T_{y_0}Y,$$

which exactly means that

$$T_{(x_0,y_0)}M = T_{x_0}X \oplus T_{y_0}Y. \tag{1.55}$$

If $\mathbf{g}_X$ and $\mathbf{g}_Y$ are Riemannian metric tensors on X and Y, respectively, then define the metric tensor $\mathbf{g}$ on M as the *direct sum*

$$\mathbf{g} = \mathbf{g}_X + \mathbf{g}_Y, \tag{1.56}$$

as follows: for any two tangent vectors $\xi, \eta \in T_{(x_0,y_0)}M$, set

$$\langle \xi, \eta \rangle_{\mathbf{g}(x_0,y_0)} = \langle \xi_X, \eta_X \rangle_{\mathbf{g}_X(x_0)} + \langle \xi_Y, \eta_Y \rangle_{\mathbf{g}_Y(y_0)}$$

In the local coordinates $x^1, ..., x^n, y^1, ..., y^m$, we have then

$$\mathbf{g}_X + \mathbf{g}_Y = (g_X)_{ij}\, dx^i dx^j + (g_Y)_{kl}\, dy^k dy^l.$$

The Riemannian manifold $(M, \mathbf{g})$ is called the *Riemannian* (or direct) *product* of $(X, \mathbf{g}_X)$ and $(Y, \mathbf{g}_Y)$.

Before we discuss the Riemannian measure on $(M, \mathbf{g})$, let us first briefly recall the notion of the product of measures. Given two measure spaces $(X, \mathcal{A}_1, \mu_1)$ and $(Y, \mathcal{A}_2, \mu_2)$ where μ_i is a σ-finite measure defined on the σ-algebra $\mathcal{A}_i$, let us define a *product measure* $\mu = \mu_1 \times \mu_2$ on the product set $M = X \times Y$ as follows. First we define μ on the subsets of M of the form $A \times B$ where $A \in \mathcal{A}_1$ and $B \in \mathcal{A}_2$ by

$$\mu (A \times B) = \mu_1 (A)\, \mu_2 (B).$$

Observing that the sets of the type $A \times B$ form a semi-ring, one can extend then μ to a σ-algebra on M by using the Carathéodory extension theorem. One of the most important properties of the product measure μ is the Fubini theorem: if $f(x, y)$ is a non-negative μ-measurable function on M then

$$\int_M f d\mu = \int_Y \left(\int_X f(x, y)\, d\mu_1 (x) \right) d\mu_2 (y) = \int_X \left(\int_Y f(x, y)\, d\mu_2 (y) \right) d\mu_1 (x).$$

Note that the matrix g of the metric tensor $\mathbf{g}$ has the block diagonal form

$$g = \begin{pmatrix} \boxed{g_X} & 0 \\ 0 & \boxed{g_Y} \end{pmatrix},$$

which implies a similar form for g^{-1}. It follows that

$$\det g = \det g_X \det g_Y. \tag{1.57}$$

Let ν_X and ν_Y be the Riemannian measures on X and Y. In the local coordinates we have

$$d\nu_X = \sqrt{\det g_X} dx \quad \text{and} \quad d\nu_Y = \sqrt{\det g_Y} dy$$

where dx and dy are Lebesgue measures in the corresponding charts $U \subset X$ and $V \subset Y$, respectively. Let λ be the Lebesgue measure in the chart $U \times V$, so that $d\lambda = dxdy$. Then the Riemannian measure ν of M is given by

$$d\nu = \sqrt{\det g} d\lambda = \sqrt{\det g_X} \sqrt{\det g_Y} dxdy = d\nu_X d\nu_Y.$$

Hence, the measure ν is the product of measures ν_X and ν_Y. This fact can also be written also as follows:

$$\boxed{\nu = \nu_X \times \nu_Y}.$$

Consequently, we obtain by Fubini' theorem that, for any non-negative measurable function $f = f(x, y)$ on M,

$$\int_M f d\nu = \int_Y \left(\int_X f(x, y)\, d\nu_X (x) \right) d\nu_Y (y) = \int_X \left(\int_Y f(x, y)\, d\nu_Y (y) \right) d\nu_X (x).$$

Denoting by $z^1, ..., z^{n+m}$ the coordinates $x^1, ..., x^n, y^1, ..., y^m$, we obtain the following expression of the Laplace operator $\Delta_{\mathbf{g}}$ on $(M, \mathbf{g})$:

$$\begin{aligned}
\Delta_{\mathbf{g}} &= \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial z^i}\left(\sqrt{\det g}\, g^{ij} \frac{\partial}{\partial z^j}\right) \\
&= \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^i}\left(\sqrt{\det g}\, g_X^{ij} \frac{\partial}{\partial x^j}\right) + \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial y^i}\left(\sqrt{\det g}\, g_Y^{ij} \frac{\partial}{\partial y^j}\right) \\
&= \frac{1}{\sqrt{\det g_X}} \frac{\partial}{\partial x^i}\left(\sqrt{\det g_X}\, g_X^{ij} \frac{\partial}{\partial x^j}\right) + \frac{1}{\sqrt{\det g_Y}} \frac{\partial}{\partial y^i}\left(\sqrt{\det g_Y}\, g_Y^{ij} \frac{\partial}{\partial y^j}\right),
\end{aligned}$$

where we have used (1.57) and the fact that $\det g_X$ depends only in x while $\det g_Y$ depends only on y. It follows that

$$\boxed{\Delta_{\mathbf{g}} = \Delta_{\mathbf{g}_X} + \Delta_{\mathbf{g}_Y}.} \tag{1.58}$$

Example. The Riemannian manifold $(\mathbb{R}^{n+m}, \mathbf{g}_{\mathbb{R}^{n+m}})$ is the Riemannian product of $(\mathbb{R}^n, \mathbf{g}_{\mathbb{R}^n})$ and $(\mathbb{R}^m, \mathbf{g}_{\mathbb{R}^m})$ because

$$\mathbf{g}_{\mathbb{R}^{n+m}} = \left(dx^1\right)^2 + ... \left(dx^n\right)^2 + \left(dx^{n+1}\right)^2 + ... + \left(dx^{n+m}\right)^2 = \mathbf{g}_{\mathbb{R}^n} + \mathbf{g}_{\mathbb{R}^m}.$$

Also, we see directly that

$$\Delta_{\mathbb{R}^{n+m}} = \frac{\partial^2}{\left(\partial x^1\right)^2} + ... + \frac{\partial^2}{\left(\partial x^n\right)^2} + \frac{\partial^2}{\left(\partial x^{n+1}\right)^2} + ... + \frac{\partial^2}{\left(\partial x^{n+m}\right)^2} = \Delta_{\mathbb{R}^n} + \Delta_{\mathbb{R}^m}.$$

There are other possibilities to define a Riemannian tensor $\mathbf{g}$ on the product manifold $M = X \times Y$. For example, if $\psi(x)$ is a smooth positive function on X then consider the metric tensor

$$\mathbf{g} = \mathbf{g}_X + \psi^2(x)\, \mathbf{g}_Y. \tag{1.59}$$

The Riemannian manifold $(M, \mathbf{g})$ with this metric is called a *warped product*[3] of $(X, \mathbf{g}_X)$ and $(Y, \mathbf{g}_Y)$. In the local coordinates, we have

$$\mathbf{g} = (g_X)_{ij}\, dx^i dx^j + \psi^2(x)\, (g_Y)_{kl}\, dy^k dy^l.$$

Let $(X, \mathbf{g}_X, \mu_X)$ and $(Y, \mathbf{g}_Y, \mu_Y)$ be weighted manifold. Setting $M = X \times Y$, $\mathbf{g} = \mathbf{g}_X + \mathbf{g}_Y$ and $\mu = \mu_X \times \mu_Y$, we obtain a weighted manifold $(M, \mathbf{g}, \mu)$ that is the direct product of the weighted manifolds $(X, \mathbf{g}_X, \mu_X)$ and $(Y, \mathbf{g}_Y, \mu_Y)$. If $D_X(x)$ and $D_Y(y)$ are the density functions on X and Y, respectively, then the density function of M is $D(x, y) = D_X(x)\, D_Y(y)$.

A computation similar to the above shows that

$$\Delta_{\mathbf{g},\mu} = \Delta_{\mathbf{g}_X,\mu_X} + \Delta_{\mathbf{g}_Y,\mu_Y}.$$

[3] verzerrtes Produkt

1.11 Polar coordinates in $\mathbb{R}^n, \mathbb{S}^n, \mathbb{H}^n$

Euclidean space. Set $\mathbb{R}^n_* = \mathbb{R}^n \setminus \{o\}$ where o is the origin. Every point $x \in \mathbb{R}^n_*$ can be represented in the *polar coordinates* as a couple (r, θ) where $r := |x| > 0$ is the *polar radius* and $\theta := \frac{x}{|x|} \in \mathbb{S}^{n-1}$ is the *polar angle*. Conversely, a couple (r, θ) determines $x \in \mathbb{R}^n_*$ uniquely by $x = r\theta$, which establishes a homeomorphism between $\mathbb{R}_+ \times \mathbb{S}^{n-1}$ and $\mathbb{R}^n_*$.

The polar coordinates can be considered as local coordinates in $\mathbb{R}^n_*$. Indeed, let Ω be any chart on $\mathbb{S}^{n-1}$ with coordinates $\theta^1, ..., \theta^{n-1}$. Then $U := \mathbb{R}_+ \times \Omega$ is a chart in $\mathbb{R}^n_*$ with coordinates $\left(r, \theta^1, ..., \theta^{n-1}\right)$.

Proposition 1.12 *The canonical Euclidean metric $\mathbf{g}_{\mathbb{R}^n}$ has the following representation in the polar coordinates:*

$$\mathbf{g}_{\mathbb{R}^n} = dr^2 + r^2 \mathbf{g}_{\mathbb{S}^{n-1}}, \tag{1.60}$$

where $\mathbf{g}_{\mathbb{S}^{n-1}}$ is the canonical spherical metric and dr^2 is the canonical metric in $\mathbb{R}_+$.

Proof. Let $\theta^1, ..., \theta^{n-1}$ be local coordinates on $\mathbb{S}^{n-1}$ and let

$$\mathbf{g}_{\mathbb{S}^{n-1}} = \gamma_{ij} d\theta^i d\theta^j. \tag{1.61}$$

Then $r, \theta^1, ..., \theta^{n-1}$ are local coordinates on $\mathbb{R}^n$, and (1.60) means that

$$\mathbf{g}_{\mathbb{R}^n} = dr^2 + r^2 \gamma_{ij} d\theta^i d\theta^j. \tag{1.62}$$

We will first prove (1.62) with some functions $\gamma_{ij} = \gamma_{ij}(\theta)$ and then verify that (1.61) holds with the same functions γ_{ij}. 15.11.16

We start with the identity $x = r\theta$, which implies that the Cartesian coordinates $x^1, ..., x^n$ can be expressed via the polar coordinates $r, \theta^1, ..., \theta^{n-1}$ as follows:

$$x^i = r f^i\left(\theta^1, ..., \theta^{n-1}\right), \tag{1.63}$$

where f^i is the x^i-coordinate in $\mathbb{R}^n$ of the point $\theta \in \mathbb{S}^{n-1}$. Clearly, $f^1, ..., f^n$ are smooth functions of $\theta^1, ..., \theta^{n-1}$ and

$$\left(f^1\right)^2 + ... + \left(f^n\right)^2 \equiv 1. \tag{1.64}$$

Considering x^i, r and f^i as functions in the chart in question and using the product rule for d, we obtain

$$dx^i = d\left(r f^i\right) = f^i dr + r df^i.$$

It follows that

$$\left(dx^i\right)^2 = \left(f^i\right)^2 dr^2 + (r dr)\left(f^i df^i\right) + \left(f^i df^i\right)(r dr) + r^2\left(df^i\right)^2. \tag{1.65}$$

Applying d to the identity (1.64), we obtain

$$\sum_i f^i df^i = 0. \tag{1.66}$$

Adding up the identities (1.65) for all $i = 1, ..., n$ and using (1.64) and (1.66), we obtain

$$\mathbf{g}_{\mathbb{R}^n} = \sum_i \left(dx^i\right)^2 = dr^2 + r^2 \sum_i \left(df^i\right)^2.$$

Next, we have

$$df^i = \frac{\partial f^i}{\partial \theta^j} d\theta^j,$$

$$\left(df^i\right)^2 = \frac{\partial f^i}{\partial \theta^j} \frac{\partial f^i}{\partial \theta^k} d\theta^j d\theta^k,$$

which implies

$$\sum_i \left(df^i\right)^2 = \gamma_{jk} d\theta^j d\theta^k, \tag{1.67}$$

where $\gamma_{jk} = \sum_i \frac{\partial f^i}{\partial \theta^j} \frac{\partial f^i}{\partial \theta^k}$ are smooth functions of $\theta^1, ..., \theta^{n-1}$. Hence, we have proved the identity (1.62).

We are left to verify that $\gamma_{ij} d\theta^i d\theta^j$ is the canonical spherical metric. Indeed, the metric $\mathbf{g}_{\mathbb{S}^{n-1}}$ is obtained restricting of the metric $\mathbf{g}_{\mathbb{R}^n}$ to $\mathbb{S}^{n-1}$. On $\mathbb{S}^{n-1}$ we have $r \equiv 1$ and, hence, $dr = 0$. Therefore, substituting in (1.62) $r = 1$ and $dr = 0$, we obtain (1.61). ■

Sphere. Consider now the polar coordinates on the n-dimensional sphere

$$\mathbb{S}^n := \left\{x \in \mathbb{R}^{n+1} : |x| = 1\right\}.$$

For any $x = (x^1, ..., x^{n+1}) \in \mathbb{R}^{n+1}$ set

$$x' = \left(x^1, ..., x^n\right),$$

that is, x' is the projection of x onto $\mathbb{R}^n = \{x \in \mathbb{R}^{n+1} : x^{n+1} = 0\}$.

Let $p = (0, ...0, 1)$ be the north pole of $\mathbb{S}^n$ and $q = -p$ be the south pole of $\mathbb{S}^n$. For any point $x \in \mathbb{S}^n \setminus \{p, q\}$, define the polar coordinates of x on $\mathbb{S}^n$ as a pair (r, θ) where $r \in (0, \pi)$ and $\theta \in \mathbb{S}^{n-1}$ are given by

$$\cos r = x^{n+1} \quad \text{and} \quad \theta = \frac{x'}{|x'|}. \tag{1.68}$$

Clearly, the polar radius r is the angle between the vectors x and p. In fact, r can be regarded as the latitude of the point x measured from the pole. The polar angle θ gives direction in the hyperplane $\mathbb{R}^n$ and can be regarded as the longitude of the point x (see Fig. 1.7 and 1.8).

As in the case of the polar coordinates in the Euclidean space, the polar coordinates (r, θ) on $\mathbb{S}^n$ can be regarded as local coordinates $r, \theta^1, ..., \theta^{n-1}$ in a chart $U = (0, \pi) \times \Omega$ where Ω is any chart on $\mathbb{S}^{n-1}$ with the local coordinates $\theta^1, ..., \theta^{n-1}$.

Proposition 1.13 *The canonical spherical metric $\mathbf{g}_{\mathbb{S}^n}$ has the following representation in the polar coordinates:*

$$\mathbf{g}_{\mathbb{S}^n} = dr^2 + \sin^2 r \mathbf{g}_{\mathbb{S}^{n-1}}. \tag{1.69}$$

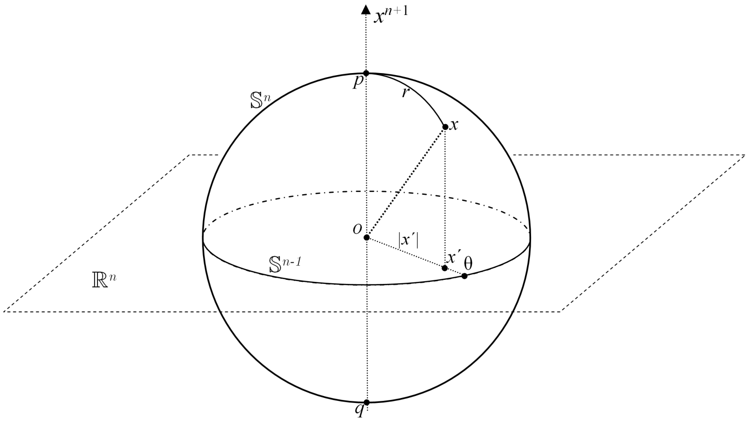

Figure 1.7: Polar coordinates on $\mathbb{S}^n$

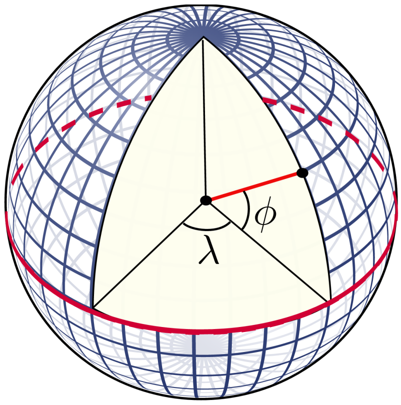

Figure 1.8: A picture from Wikipedia: the geographical latitude ϕ and longitude λ on the Earth considered as $\mathbb{S}^2$. In our notation $r = \frac{\pi}{2} - \phi$ and $\theta = \lambda + \text{const}$.

Proof. Let $\theta^1, ..., \theta^{n-1}$ are local coordinates on $\mathbb{S}^{n-1}$ and let us write down the metric $\mathbf{g}_{\mathbb{S}^n}$ in the local coordinates $r, \theta^1, ..., \theta^{n-1}$. Obviously, for any point $x \in \mathbb{S}^n \setminus \{p, q\}$, we have

$$|x'| = \sqrt{1 - (x^{n+1})^2} = \sqrt{1 - \cos^2 r} = \sin r$$

whence

$$x' = (\sin r)\, \theta.$$

Hence, the Cartesian coordinates $x^1, ..., x^{n+1}$ of the point x can be expressed as follows:

$$\begin{aligned} x^i &= \sin r\, f^i\left(\theta^1, ..., \theta^{n-1}\right),\ i = 1, ..., n, \\ x^{n+1} &= \cos r, \end{aligned}$$

where f^i are the same functions as in (1.63). Therefore, we obtain using (1.64), (1.66), and (1.67),

$$\begin{aligned}
\mathbf{g}_{\mathbb{S}^n} &= \left(dx^1\right)^2 + ... + (dx^n)^2 + \left(dx^{n+1}\right)^2 \\
&= \sum_{i=1}^n \left(f^i \cos r dr + \sin r df^i\right)^2 + \sin^2 r dr^2 \\
&= \sum_{i=1}^n \left[\left(f^i\right)^2 \cos^2 r dr^2 + \sin^2 r \left(df^i\right)^2 + \sin r \cos r dr f^i df^i + f^i df^i \sin r \cos r dr\right] \\
&\quad + \sin^2 r dr^2 \\
&= \left(\cos^2 r + \sin^2 r\right) dr^2 + \sin^2 r \sum_{i=1}^n \left(df^i\right)^2 \\
&= dr^2 + \sin^2 r\, \gamma_{ij} d\theta^i d\theta^j .
\end{aligned}$$

Since we already know that $\gamma_{ij} d\theta^i d\theta^j$ is the canonical metric on $\mathbb{S}^{n-1}$, we obtain (1.69).
∎

Hyperbolic space. The *hyperbolic space* $\mathbb{H}^n$, $n \geq 2$, is defined as follows. Consider in $\mathbb{R}^{n+1}$ a *hyperboloid* H given by the equation[4]

$$\left(x^{n+1}\right)^2 - (x')^2 = 1, \quad x^{n+1} > 0, \tag{1.70}$$

where as above $x' = (x^1, ..., x^n) \in \mathbb{R}^n$. By Lemma 1.6, H is a submanifold of $\mathbb{R}^{n+1}$ of dimension n.

Consider in $\mathbb{R}^{n+1}$ the *Minkowski metric*

$$\mathbf{g}_{Mink} = \left(dx^1\right)^2 + ... + (dx^n)^2 - \left(dx^{n+1}\right)^2, \tag{1.71}$$

which is a bilinear symmetric form in any tangent space $T_x\mathbb{R}^{n+1}$ but not positive definite. Hence, $\mathbf{g}_{Mink}$ is not a Riemannian metric, but is a called a *pseudo*-Riemannian metric. Nevertheless we can restrict $\mathbf{g}_{Mink}$ to H, so set

$$\mathbf{g}_H = \mathbf{g}_{Mink}|_H .$$

We will prove below that $\mathbf{g}_H$ is positive definite so that $(H, \mathbf{g}_H)$ is a Riemannian manifold. By definition, this manifold is called the *hyperbolic space* of dimension n and is denoted by $\mathbb{H}^n$. The metric $\mathbf{g}_H$ is called the *canonical hyperbolic metric* and is denoted also by $\mathbf{g}_{\mathbb{H}^n}$.

Our main purpose here is to introduce the polar coordinates in $\mathbb{H}^n$ and to represent $\mathbf{g}_{\mathbb{H}^n}$ in the polar coordinates. As a by-product, we will see that $\mathbf{g}_{\mathbb{H}^n}$ is positive definite.

Consider the point $p = (0, ..., 0, 1)$ that is called the pole of $\mathbb{H}^n$. For any point $x \in \mathbb{H}^n \setminus \{p\}$, define its polar coordinates as a pair (r, θ) where $r > 0$ and $\theta \in \mathbb{S}^{n-1}$ are given by

$$\cosh r = x^{n+1} \quad \text{and} \quad \theta = \frac{x'}{|x'|} \tag{1.72}$$

(see Fig. 1.9). The value of r is called the *hyperbolic angle* between the vectors x and p. It is possible to prove that the area of the sector bounded by the arc of the hyperbola between p and x and by the segments $[o, p]$, $[o, x]$ is equal to $r/2$.

[4] For comparison, the equation of $\mathbb{S}^n$ can be written in the form $\left(x^{n+1}\right)^2 + (x')^2 = 1$.

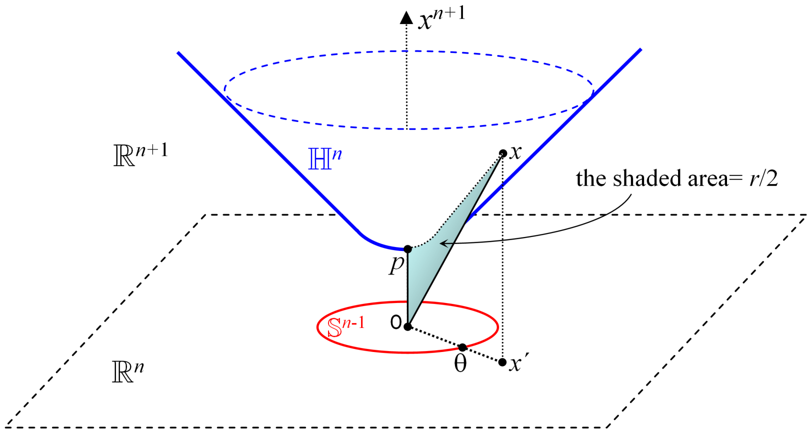

Figure 1.9: Polar coordinates on $\mathbb{H}^n$

Proposition 1.14 *The canonical hyperbolic metric $\mathbf{g}_{\mathbb{H}^n}$ has the following representation in the polar coordinates:*

$$\mathbf{g}_{\mathbb{H}^n} = dr^2 + \sinh^2 r \mathbf{g}_{\mathbb{S}^{n-1}}. \tag{1.73}$$

Consequently, $\mathbf{g}_{\mathbb{H}^n}$ is a Riemannian metric.

Proof. Let $\theta^1, ..., \theta^{n-1}$ be local coordinates on $\mathbb{S}^{n-1}$ and let us write down the metric $\mathbf{g}_{\mathbb{H}^n}$ in the local coordinates $r, \theta^1, ..., \theta^{n-1}$. For any point $x \in \mathbb{H}^n \setminus \{p\}$, we have

$$|x'| = \sqrt{|x^{n+1}|^2 - 1} = \sqrt{\cosh^2 r - 1} = \sinh r,$$

whence

$$x' = (\sinh r)\, \theta.$$

Hence, the Cartesian coordinates $x^1, ..., x^{n+1}$ of the point x can be expressed as follows:

$$\begin{array}{rcl} x^i & = & \sinh r f^i\left(\theta^1, ..., \theta^{n-1}\right),\ i = 1, ..., n, \\ x^{n+1} & = & \cosh r, \end{array}$$

where f^i are the same functions as in (1.63). It follows that

$$\begin{array}{rcl} \mathbf{g}_{\mathbb{H}^n} & = & \left(dx^1\right)^2 + ... + \left(dx^n\right)^2 - \left(dx^{n+1}\right)^2 \\ & = & \displaystyle\sum_{i=1}^n \left(f^i \cosh r dr + \sinh r df^i\right)^2 - \sinh^2 r dr^2 \\ & = & \displaystyle\sum_{i=1}^n \left[(f^i)^2 \cosh^2 r dr^2 + \sinh^2 r (df^i)^2 + \sinh r \cosh r dr f^i df^i + f^i df^i \sinh r \cosh r dr\right] \\ & & - \sinh^2 r dr^2 \\ & = & \left(\cosh^2 r - \sinh^2 r\right) dr^2 + \sinh^2 r \displaystyle\sum_{i=1}^n \left(df^i\right)^2 \\ & = & dr^2 + \sinh^2 r\, \gamma_{ij} d\theta^i d\theta^j. \end{array}$$

Since $\gamma_{ij}d\theta^i d\theta^j$ is the canonical metric on $\mathbb{S}^{n-1}$, we obtain (1.73).

Let us verify that $\mathbf{g}_{\mathbb{H}^n}$ is a Riemannian metric. We see from (1.73) that the tensor $\mathbf{g}_{\mathbb{H}^n}(x)$ is positive definite on $T_x\mathbb{H}^n$ for any $x \in \mathbb{H}^n \setminus \{p\}$. For the case $x = p$, let us use the local coordinates $x^1, ..., x^n$ on $\mathbb{H}^n$. Since x^{n+1} as a function on $\mathbb{H}^n$ attains its minimum at p, we see that $dx^{n+1}(p) = 0$ and the restriction of $\mathbf{g}_{Mink}$ onto $T_p\mathbb{H}^n$ becomes $(dx^1)^2 + ... + (dx^n)^2$ that is positive definite. $\blacksquare$

17.11.16

1.12 Model manifolds

Definition. An n-dimensional Riemannian manifold $(M, \mathbf{g})$ is called a *Riemannian model* if the following two conditions are satisfied:

1. There is a chart on M that covers all M, and the image of this chart in $\mathbb{R}^n$ is a ball
$$B_{r_0} := \{x \in \mathbb{R}^n : |x| < r_0\}$$
of radius $r_0 \in (0, +\infty]$ (in particular, if $r_0 = \infty$ then $B_{r_0} = \mathbb{R}^n$).

2. The metric $\mathbf{g}$ in the polar coordinates (r, θ) in the above chart has the form
$$\mathbf{g} = dr^2 + \psi^2(r)\,\mathbf{g}_{\mathbb{S}^{n-1}}, \tag{1.74}$$
where $\psi(r)$ is a smooth positive function on $(0, r_0)$.

The number r_0 is called the radius of the model M.

It follows that M is homeomorphic to B_{r_0}. To simplify the terminology and notation, we usually identify a model M with the ball B_{r_0}. Then the polar coordinates (r, θ) are defined in $M \setminus \{o\}$ where o is the origin of $\mathbb{R}^n$. If $\theta^1, ..., \theta^{n-1}$ are the local coordinates on $\mathbb{S}^{n-1}$ and
$$\mathbf{g}_{\mathbb{S}^{n-1}} = \gamma_{ij}d\theta^i d\theta^j,$$
then $r, \theta^1, ..., \theta^{n-1}$ are local coordinates on $M \setminus \{o\}$, and (1.74) is equivalent to
$$\mathbf{g} = dr^2 + \psi^2(r)\,\gamma_{ij}d\theta^i d\theta^j. \tag{1.75}$$

Observe also that away from a neighborhood of o, $\psi(r)$ may be any smooth positive function. However, $\psi(r)$ should satisfy certain conditions near o to ensure that the metric (1.74) extends smoothly to o.

For example, the results of Section 1.11 imply the following:

- $\mathbb{R}^n$ is a model with the radius $r_0 = \infty$ and $\psi(r) = r$;
- $\mathbb{S}^n \setminus \{q\}$ is a model with the radius $r_0 = \pi$ and $\psi(r) = \sin r$;
- $\mathbb{H}^n$ is a model with the radius $r_0 = \infty$ and $\psi(r) = \sinh r$.

Lemma 1.15 *On a model manifold* $(M, \mathbf{g})$ *with metric* (1.74), *the Riemannian measure* ν *is given in the polar coordinates in* $B_{r_0} \setminus \{o\}$ *by*

$$d\nu = \psi(r)^{n-1} \, dr d\sigma, \tag{1.76}$$

where dr *denotes the Lebesgue measure on* $(0, r_0)$ *and* $d\sigma$ *denotes the Riemannian measure on* $\mathbb{S}^{n-1}$.

The Laplace operator $\Delta_{\mathbf{g}}$ *has in the polar coordinates the form*

$$\Delta_{\mathbf{g}} = \frac{\partial^2}{\partial r^2} + \frac{d \ln \psi^{n-1}}{dr} \frac{\partial}{\partial r} + \frac{1}{\psi^2(r)} \Delta_{\mathbb{S}^{n-1}} \, . \tag{1.77}$$

Remark. The formula (1.76) can be used to integrate functions over M using the polar coordinates. Indeed, if f is any non-negative measurable function on M then we have

$$\begin{aligned} \int_M f d\nu &= \int_{M \setminus \{0\}} f d\nu = \int_0^{r_0} \int_{\mathbb{S}^{n-1}} f(r, \theta) \, \psi(r)^{n-1} \, dr d\sigma \\ &= \int_0^{r_0} \left(\int_{\mathbb{S}^{n-1}} f(r, \theta) \, d\sigma \right) \psi(r)^{n-1} \, dr. \end{aligned}$$

Proof. Let Ω be a chart on $\mathbb{S}^{n-1}$ with coordinates $\theta^1, ..., \theta^{n-1}$. Then $U = (0, r_0) \times \Omega$ is a chart on M with coordinates $r, \theta^1, ..., \theta^{n-1}$. For simplicity of notation, set $\theta^0 = r$. Let $g = (g_{ij})$ be the matrix of the tensor $\mathbf{g}$ in the chart U, where the indices i, j vary from 0 to $n-1$. It follows from (1.75)

$$g = \begin{pmatrix} 1 & 0 \quad \cdots \quad 0 \\ 0 & \\ \vdots & \boxed{\quad \psi^2(r) \, \gamma_{ij} \quad} \\ 0 & \end{pmatrix}. \tag{1.78}$$

In particular, we have

$$\det g = \psi^{2(n-1)} \det \gamma, \tag{1.79}$$

where $\gamma = (\gamma_{ij})$. By (1.33), the Riemannian measure ν on M is given by

$$d\nu = \sqrt{\det g} d\lambda,$$

where λ is the Lebesgue measure in the chart U. Denoting by dr the Lebesgue measure in $(0, r_0)$ and by $d\theta$ the Lebesgue measure in the chart Ω on $\mathbb{S}^{n-1}$, we have

$$d\lambda = dr d\theta.$$

Using also that

$$d\sigma = \sqrt{\det \gamma} d\theta.$$

we obtain

$$d\nu = \psi^{n-1} \sqrt{\det \gamma} dr d\theta = \psi^{n-1} dr d\sigma,$$

which proves (1.76).

It follows from (1.78) that

$$\left(g^{ij}\right)=g^{-1}=\left(\begin{array}{c|c} 1 & 0\ \cdots\ 0 \\ \hline 0 & \\ \vdots & \psi^{-2}\left(r\right)\gamma^{ij} \\ 0 & \end{array}\right) \tag{1.80}$$

where $(\gamma^{ij})=\left(\gamma_{ij}\right)^{-1}$. By (1.48) of $\Delta_{\mathbf{g}}$, we have in the local coordinates $\theta^0,...,\theta^{n-1}$

$$\Delta_{\mathbf{g}}=\frac{1}{\sqrt{\det g}}\sum_{i,j=0}^{n-1}\frac{\partial}{\partial\theta^i}\left(\sqrt{\det g}g^{ij}\frac{\partial}{\partial\theta^j}\right). \tag{1.81}$$

Since $g^{00}=1$, $g^{0i}=0$ for $i\geq 1$, it follows that

$$\Delta_{\mathbf{g}}=\frac{1}{\sqrt{\det g}}\frac{\partial}{\partial r}\left(\sqrt{\det g}\frac{\partial}{\partial r}\right)+\sum_{i,j=1}^{n-1}\frac{1}{\sqrt{\det g}}\frac{\partial}{\partial\theta^i}\left(\sqrt{\det g}g^{ij}\frac{\partial}{\partial\theta^j}\right). \tag{1.82}$$

Applying (1.80) and (1.79) and noticing that ψ depends only on r and γ_{ij} depend only on $\theta^1,...,\theta^{n-1}$, we obtain

$$\begin{aligned}\frac{1}{\sqrt{\det g}}\frac{\partial}{\partial r}\left(\sqrt{\det g}\frac{\partial}{\partial r}\right) &= \frac{\partial^2}{\partial r^2}+\left(\frac{\partial}{\partial r}\ln\sqrt{\det g}\right)\frac{\partial}{\partial r}\\ &= \frac{\partial^2}{\partial r^2}+\left(\frac{d}{dr}\ln\psi^{n-1}\right)\frac{\partial}{\partial r}\end{aligned}$$

and

$$\begin{aligned}\sum_{i,j=1}^{n-1}\frac{1}{\sqrt{\det g}}\frac{\partial}{\partial\theta^i}\left(\sqrt{\det g}g^{ij}\frac{\partial}{\partial\theta^j}\right) &= \sum_{i,j=1}^{n-1}\frac{\psi^{-2}\left(r\right)}{\sqrt{\det\gamma}}\frac{\partial}{\partial\theta^i}\left(\sqrt{\det\gamma}\gamma^{ij}\frac{\partial}{\partial\theta^j}\right)\\ &= \frac{1}{\psi^2\left(r\right)}\Delta_{\mathbb{S}^{n-1}}.\end{aligned}$$

Substituting into (1.82), we obtain (1.77). ■

Example. In $\mathbb{R}^n$, we have $\psi(r)=r$. In this case the Riemannian measure ν coincides with the Lebesgue measure λ_n, and we obtain

$$d\lambda=r^{n-1}drd\sigma. \tag{1.83}$$

It follows from (1.77) that

$$\Delta_{\mathbb{R}^n}=\frac{\partial^2}{\partial r^2}+\frac{n-1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\Delta_{\mathbb{S}^{n-1}}. \tag{1.84}$$

Consider the case $n=2$. If θ denotes the angle on $\mathbb{S}^1$ then we have $\mathbf{g}_{\mathbb{S}^1}=d\theta^2$ (Exercise 11) and, hence, $\Delta_{\mathbb{S}^{n-1}}=\frac{\partial^2}{\partial\theta^2}$. It follows that in the planar polar coordinates

$$\Delta_{\mathbb{R}^2}=\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}.$$

In $\mathbb{S}^n$, we have $\psi(r) = \sin r$ and, hence,

$$d\nu = \sin^{n-1} r \, dr d\sigma \tag{1.85}$$

and

$$\Delta_{\mathbb{S}^n} = \frac{\partial^2}{\partial r^2} + (n-1)\cot r \frac{\partial}{\partial r} + \frac{1}{\sin^2 r}\Delta_{\mathbb{S}^{n-1}}\,. \tag{1.86}$$

In $\mathbb{H}^n$, we have $\psi(r) = \sinh r$ and, hence,

$$d\nu = \sinh^{n-1} r \ dr d\sigma$$

and

$$\Delta_{\mathbb{H}^n} = \frac{\partial^2}{\partial r^2} + (n-1)\coth r \frac{\partial}{\partial r} + \frac{1}{\sinh^2 r}\Delta_{\mathbb{S}^{n-1}}\,. \tag{1.87}$$

The formula (1.86) can be iterated in dimension to obtain a full expansion of $\Delta_{\mathbb{S}^n}$ in the polar coordinates.

Example. Denote by σ_n the Riemannian measure on $\mathbb{S}^n$. The measure σ_n is frequently referred to as an *area*. Then (1.85) can be rewritten as follows:

$$d\sigma_n = \sin^{n-1} r \, dr d\sigma_{n-1}.$$

In particular, using Fubini's theorem, we obtain

$$\begin{aligned} \sigma_n(\mathbb{S}^n) &= \int_{\mathbb{S}^n} d\sigma_n = \int_0^\pi \int_{\mathbb{S}^{n-1}} \sin^{n-1} r \, dr d\sigma_{n-1} \\ &= \int_0^\pi \sin^{n-1} r \, dr \int_{\mathbb{S}^{n-1}} d\sigma_{n-1} \\ &= \sigma_{n-1}\left(\mathbb{S}^{n-1}\right) \int_0^\pi \sin^{n-1} r \, dr. \end{aligned}$$

Set

$$\omega_n := \sigma_{n-1}\left(\mathbb{S}^{n-1}\right) = \int_{\mathbb{S}^{n-1}} d\sigma_{n-1}, \tag{1.88}$$

that is, ω_n is the total area of the unit sphere on $\mathbb{R}^n$. Hence, we obtain the inductive formula

$$\omega_{n+1} = \omega_n \int_0^\pi \sin^{n-1} r \, dr. \tag{1.89}$$

For $n = 2$ we now that $\mathbf{g}_{\mathbb{S}^1} = d\theta^2$ and, hence, $d\sigma_1 = d\theta$, which implies that

$$\omega_2 = \int_0^{2\pi} d\theta = 2\pi.$$

By (1.89) we obtain

$$\omega_3 = 2\pi \int_0^\pi \sin r dr = 4\pi,$$

$$\omega_4 = 4\pi \int_0^\pi \sin^2 r = 2\pi^2.$$

Remark. Let us explain the meaning of the term $\psi^{n-1}(r)$ in (1.76). Consider for any $R \in (0, r_0)$ the sphere

$$S_R = \{x \in \mathbb{R}^n : |x| = R\}$$

as a submanifold of M of dimension $n-1$. It follows from (1.74) that the induced metric on S_R in the coordinates $\theta^1, ..., \theta^{n-1}$ is given by

$$(g_{S_R})_{ij} = \psi(R)^2 \gamma_{ij}(\theta)\, d\theta^i d\theta^j.$$

Denoting by σ_R the corresponding Riemannian measure on S_R (that is also called an area), we obtain

$$d\sigma_R = \sqrt{\det\left(\psi(R)^2 \gamma_{ij}(\theta)\right)} d\theta = \psi(R)^{n-1} \sqrt{\det \gamma} d\theta = \psi(R)^{n-1} d\sigma.$$

Using a chart Ω on $\mathbb{S}^{n-1}$ that covers almost all $\mathbb{S}^{n-1}$, we obtain that the total measure of S_R is

$$\sigma_R(S_R) = \int_{S_R} d\sigma_R = \int_\Omega \psi(R)^{n-1} \sqrt{\det \gamma} d\theta = \psi(R)^{n-1} \int_{\mathbb{S}^{n-1}} d\sigma = \omega_n \psi(R)^{n-1}. \tag{1.90}$$

Hence, $\omega_n \psi(R)^{n-1}$ is the total area of S_R.

Definition. A weighted manifold $(M, \mu, \mathbf{g})$ is called a *weighted model* if $(M, \mathbf{g})$ is a Riemannian model as above, and the density function D of the measure μ depends only on the polar angle r.

Lemma 1.16 *On a weighted model manifold $(M, \mathbf{g}, \mu)$ with metric (1.74) and the density function $D(r)$, the measure μ is given in the polar coordinates by*

$$d\mu = D(r)\, \psi^{n-1}(r)\, dr d\sigma, \tag{1.91}$$

where dr denotes the Lebesgue measure on $(0, r_0)$ and $d\sigma$ denotes the Riemannian measure on $\mathbb{S}^{n-1}$. The weighted Laplace operator $\Delta_{\mathbf{g},\mu}$ has in the polar coordinates the form

$$\Delta_{\mathbf{g},\mu} = \frac{\partial^2}{\partial r^2} + \frac{d}{dr} \ln\left(D\psi^{n-1}\right) \frac{\partial}{\partial r} + \frac{1}{\psi^2(r)} \Delta_{\mathbb{S}^{n-1}}. \tag{1.92}$$

Proof. Since

$$d\mu = D d\nu,$$

the identity (1.91) follows immediately from the identity (1.76) of Lemma 1.15.

By definition of the weighted Laplacian, we have

$$\begin{aligned}
\Delta_{\mathbf{g},\mu} f &= \operatorname{div}_{\mathbf{g},\mu}(\nabla f) = \frac{1}{D} \operatorname{div}_{\mathbf{g}}(D \nabla f) = \Delta_{\mathbf{g}} f + \frac{1}{D} \langle \nabla D, \nabla f \rangle \\
&= \Delta_{\mathbf{g}} f + \langle \nabla \ln D, \nabla f \rangle.
\end{aligned}$$

Using the notation $\theta^0 = r$ and the matrix (g^{ij}) given by (1.80), we obtain

$$\begin{aligned}
\langle \nabla \ln D, \nabla f \rangle &= \sum_{i,j=0}^{n-1} g^{ij} \frac{\partial \ln D}{\partial \theta^i} \frac{\partial f}{\partial \theta^j} = \frac{\partial \ln D}{\partial r} \frac{\partial f}{\partial r} + \sum_{i,j=1}^{n-1} g^{ij} \frac{\partial \ln D}{\partial \theta^i} \frac{\partial f}{\partial \theta^j} \\
&= \frac{d \ln D}{dr} \frac{\partial f}{\partial r},
\end{aligned}$$

because $\frac{\partial \ln D}{\partial \theta^i} = 0$ for all $i \geq 1$. Using the representation of $\Delta_{\mathbf{g}}$ from Lemma 1.15, we obtain

$$\Delta_{\mathbf{g},\mu} f = \frac{\partial^2 f}{\partial r^2} + \frac{d \ln \psi^{n-1}}{dr}\frac{\partial f}{\partial r} + \frac{1}{\psi^2(r)}\Delta_{\mathbb{S}^{n-1}} f + \frac{d \ln D}{dr}\frac{\partial f}{\partial r}.$$

Observing that

$$\ln \psi^{n-1} + \ln D = \ln\left(D\psi^{n-1}\right),$$

we obtain (1.92). ∎

Let $(M, \mathbf{g}, \mu)$ be any weighted manifold with the density function D. For any submanifold S of M, we have defined the induced Riemannian metric $\mathbf{g}_S$ on S. Let us define the induced measure μ_S as the measure on S with the density function $D|_S$ with respect to the Riemannian measure ν_S of S. Then $(S, \mathbf{g}_S, \mu_S)$ is a weighted manifold.

If $(M, \mathbf{g}, \mu)$ is a weighted model as above then the sphere

$$S_R = \{x \in \mathbb{R}^n : |x| = R\}$$

where $R \in (0, r_0)$, is a submanifold, so we obtain the induced metric $\mathbf{g}_{S_R}$ and the corresponding Riemannian measure σ_R as above, as well as the induced measure μ_{S_R} that we denote simply by μ_R and refer to as a *weighted area*. Since on S_R we have $D \equiv D(R)$, it follows from the definition of μ_R that

$$d\mu_R = D(R)\, d\sigma_R.$$

In particular, the total weighted area of S_R is given by

$$\mu_R(S_R) = D(R)\,\sigma_R(S_R) = \omega_n D(R)\,\psi(R)^{n-1},$$

which gives a geometric meaning to the term $D\psi^{n-1}$ that appears in (1.91) and (1.92). 22.11.16

The function

$$\boxed{S(r) := \omega_n D(r)\,\psi^{n-1}(r)}, \tag{1.93}$$

that coincides with $\mu_R(S_R)$, is called the *area function* of the weighted model $(M, \mathbf{g}, \mu)$. We can rewrite (1.92) in terms of this function as follows:

$$\boxed{\Delta_{\mathbf{g},\mu} = \frac{\partial^2}{\partial r^2} + \frac{S'(r)}{S(r)}\frac{\partial}{\partial r} + \frac{1}{\psi^2(r)}\Delta_{\mathbb{S}^{n-1}}}. \tag{1.94}$$

We see that the Laplace operator is directly related to such a geometric property of the manifold as the area function.

For any $R \in (0, r_0)$ consider the Euclidean ball

$$B_R = \{x \in \mathbb{R}^n : |x| < R\}.$$

as a subset of M. In the polar coordinates, we have

$$B_R \setminus \{o\} = \left\{(r, \theta) \in M : r \in (0, R), \theta \in \mathbb{S}^{n-1}\right\}.$$

It follows from (1.91) that

$$\mu(B_R) = \int_0^R \int_{\mathbb{S}^{n-1}} D\psi^{n-1} dr d\sigma = \omega_n \int_0^R D\psi^{n-1} dr = \int_0^R S(r)\, dr. \tag{1.95}$$

The function $V(R) := \mu(B_R)$ is called the *volume function* of the model manifold. Hence, we obtain

$$\boxed{V(R) = \int_0^R S(r)\,dr}.$$

For example, in $\mathbb{R}^n$ we have $\psi(r) = r^{n-1}$ and $D = 1$, which implies

$$S(r) = \omega_n r^{n-1}$$

and

$$V(R) = \frac{\omega_n}{n} R^n. \tag{1.96}$$

1.13 Length of paths and the geodesic distance

Let M be a smooth manifold.

Definition. A *path* (or *parametric curve*) on M is any continuous mapping $\gamma : I \to M$ where I is any interval in $\mathbb{R}$.

In the local coordinates $x^1, ..., x^n$, the path is given by its components $x^i = \gamma^i(t)$. If $\gamma^i(t)$ are C^k functions of t then the path γ is also called C^k.

Definition. For any C^1 path $\gamma : I \to M$ and for any $t \in I$, define the *velocity* $\dot\gamma(t)$ as the following $\mathbb{R}$-differentiation at $x = \gamma(t)$:

$$\begin{aligned}\dot\gamma(t) &: C^\infty(M) \to \mathbb{R} \\ \dot\gamma(t)(f) &= \frac{d}{dt} f(\gamma(t)) \quad \text{for any } f \in C^\infty(M).\end{aligned} \tag{1.97}$$

Indeed, it is easy to see that the mapping $\dot\gamma(t)$ defined by (1.97) satisfies the definition of an $\mathbb{R}$-differentiation at the point $x = \gamma(t)$: it is linear and satisfies the product rule, because so does the ordinary derivative $\frac{d}{dt}$. Hence, $\dot\gamma(t) \in T_{\gamma(t)}M$.

Let us express the tangent vector $\dot\gamma(t)$ in the local coordinates $x^1, ..., x^n$. Applying the chain rule, we obtain

$$\dot\gamma(t)(f) = \frac{d}{dt} f\left(\gamma^1(t), ..., \gamma^n(t)\right) = \frac{\partial f}{\partial x^i}\frac{d\gamma^i}{dt} = \dot\gamma^i \frac{\partial f}{\partial x^i}, \tag{1.98}$$

where using the notation

$$\dot\gamma^i \equiv \frac{d\gamma^i}{dt}.$$

Rewriting (1.98) in the operator form as follows

$$\dot\gamma = \dot\gamma^i \frac{\partial}{\partial x^i}, \tag{1.99}$$

we see that $\dot\gamma(t)$ has in the basis $\left\{\frac{\partial}{\partial x^i}\right\}$ the components $\dot\gamma^i(t)$.

As one of the consequences of (1.99), we obtain that any tangent vector $\xi \in T_xM$ can be represented as the velocity of a path; for example, one can take the path $\gamma^i(t) = x^i + t\xi^i$.

Let now $(M, \mathbf{g})$ be a Riemannian manifold. Recall that length of a tangent vector $\xi \in T_x M$ is defined by $|\xi|_{\mathbf{g}} = \sqrt{\langle \xi, \xi \rangle_{\mathbf{g}}}$.

Definition. For any C^1 path $\gamma : I \to M$, define its *length* $\ell_{\mathbf{g}}(\gamma)$ by

$$\ell_{\mathbf{g}}(\gamma) = \int_I |\dot{\gamma}(t)|_{\mathbf{g}}\, dt. \tag{1.100}$$

If the interval I is bounded and closed then clearly $\ell(\gamma) < \infty$. If the image of γ is contained in a chart U with coordinates $x^1, ..., x^n$ then

$$|\dot{\gamma}(t)|_{\mathbf{g}} = \sqrt{g_{ij}(\gamma(t))\, \dot{\gamma}^i(t)\, \dot{\gamma}^j(t)}$$

and hence

$$\ell_{\mathbf{g}}(\gamma) = \int_I \sqrt{g_{ij}\dot{\gamma}^i \dot{\gamma}^j}\, dt.$$

For example, if $(g_{ij}) \equiv \mathrm{id}$ then

$$\ell_{\mathbf{g}}(\gamma) = \int_I \sqrt{\left(\dot{\gamma}^1\right)^2 + ... + \left(\dot{\gamma}^n\right)^2}\, dt.$$

Assume in what follows that the interval I is bounded and closed, say, $I = [a, b]$, and extend the definition of $\ell_{\mathbf{g}}(\gamma)$ to piecewise C^1 paths γ. A path $\gamma : [a, b] \to M$ is called *piecewise* C^1 if it is continuous on $[a, b]$ and there is a finite partition $a = t_0 < t_1 < ... < t_N = b$ of the interval $[a, b]$ so that γ is C^1 on each of the intervals $[t_k, t_{k+1}]$. Then the velocity $\dot{\gamma}(t)$ is defined for all $t \neq t_k$ and the integral (1.100) still makes sense, so that the length $\ell_{\mathbf{g}}(\gamma)$ is well defined for piecewise C^1 paths and, moreover, is finite.

Let us use the paths to define a distance function on the manifold $(M, \mathbf{g})$. We say that a path $\gamma : [a, b] \to M$ connects points x and y if $\gamma(a) = x$ and $\gamma(b) = y$.

Definition. The *geodesic distance* $d(x, y)$ between any two points $x, y \in M$ is defined by

$$d(x, y) = \inf\left\{\ell_{\mathbf{g}}(\gamma) : \gamma \text{ is a piecewise } C^1\text{-path connecting } x \text{ and } y\right\}. \tag{1.101}$$

If the infimum in (1.101) is attained on a path γ then γ is called *a shortest* (or *a minimizing*) *geodesics* between x and y. If there is no path connecting x and y then, by definition, $d(x, y) = +\infty$.

For example, consider $\mathbb{R}^n$ with the canonical metric $\mathbf{g}_{\mathbb{R}^n}$. Then the geodesic distance of $(\mathbb{R}^n, \mathbf{g}_{\mathbb{R}^n})$ coincides with the Euclidean distance $|x - y|$, and the straight line segment $[x, y]$ between $x, y \in \mathbb{R}^n$ is the shortest geodesic (see Exercise 33).

Our purpose is to show that the geodesic distance is a metric[5] on M, and the topology of the metric space (M, d) coincides with the original topology of the smooth manifold M (see Theorem 1.20 below). We start with the following observation.

[5]We allow a metric $d(x, y)$ to take value $+\infty$. It can always be replaced by a *finite* metric

$$\widetilde{d}(x, y) := \frac{d(x, y)}{1 + d(x, y)},$$

which determines the same topology as $d(x, y)$.

Lemma 1.17 *The geodesic distance satisfies the following properties.*

(i) $d(x,y) \in [0,+\infty]$ *and* $d(x,x) = 0$.

(ii) Symmetry: $d(x,y) = d(y,x)$.

(iii) The triangle inequality: $d(x,y) \le d(x,z) + d(y,z)$.

Proof. (i) That $d(x,y) \in [0,\infty]$ is obvious from (1.101). Given $x \in M$, consider a constant path $\gamma : [0,1] \to M$ defined by $\gamma(t) \equiv x$. Clearly, $\dot{\gamma}(t) \equiv 0$ and $\ell_{\mathbf{g}}(\gamma) = 0$ whence $d(x,x) = 0$ follows.

(ii) If $\gamma : [a,b] \to M$ connects x and y, that is, $\gamma(a) = x$ and $\gamma(b) = y$ then consider a path

$$\widetilde{\gamma}(t) = \gamma(a+b-t)$$

that is also defined on $[a,b]$. Clearly, $\widetilde{\gamma}(a) = y$ and $\widetilde{\gamma}(b) = x$ so that $\widetilde{\gamma}$ connects y and x. It is obvious from the definition that $\ell(\widetilde{\gamma}) = \ell(\gamma)$, which implies by (1.101) that $d(x,y) = d(y,x)$.

(iii) Consider any piecewise C^1 path $\gamma_1 : [a_1,b_1] \to M$ connecting x and z, and a piecewise C^1 path $\gamma_2 : [a_2,b_2]$ connecting z and y. By a shift of the variable t, we can always assume that $b_1 = a_2$. Define the path $\gamma : [a_1,b_2]$ connecting x and y, as follows:

$$\gamma(t) = \begin{cases} \gamma_1(t), & t \in [a_1,b_1], \\ \gamma_2(t), & t \in [a_2,b_2]. \end{cases}$$

This path is continuous because $b_1 = a_2$ and $\gamma_1(b_1) = z = \gamma_2(a_2)$, and piecewise C^1 because so are γ_1 and γ_2. It follows from (1.101) that

$$d(x,y) \le \ell(\gamma) = \ell(\gamma_1) + \ell(\gamma_2).$$

Taking infimum with respect to γ_1 and γ_2, we obtain

$$d(x,y) \le d(x,z) + d(z,y),$$

which finishes the proof. ∎

We still need to verify that $d(x,y) > 0$ for all distinct points x,y. A crucial step towards that is contained in the following lemma.

Lemma 1.18 *For any point $p \in M$, there is a chart $U \ni p$ and $C \ge 1$ such that, for all $x,y \in \overline{U}$,*

$$C^{-1}|x-y| \le d(x,y) \le C|x-y|, \tag{1.102}$$

where $|x-y|$ is the Euclidean distance in U.

Proof. Fix a point $p \in M$ and a chart W around p with local coordinates $x^1,...,x^n$. Let V be the Euclidean ball $B_r(p)$ of radius r centered at p where $r > 0$ is so small that $\overline{V} \subset W$.

For any $x \in \overline{V}$ and any tangent vector $\xi \in T_xM$, its length $|\xi|_{\mathbf{g}}$ in the metric $\mathbf{g}$ is given by

$$|\xi|_{\mathbf{g}}^2 = g_{ij}(x)\,\xi^i\xi^j.$$

Denoting for simplicity the Euclidean metric $\mathbf{g}_{\mathbb{R}^n}$ in W by $\mathbf{e}$, we have

$$|\xi|_{\mathbf{e}}^2 = \sum_{i=1}^n \left(\xi^i\right)^2.$$

Since the matrix $(g_{ij}(x))$ is positive definite and continuously depends on x, there is a constant $C \geq 1$ such that

$$C^{-2} \sum_{i=1}^n \left(\xi^i\right)^2 \leq g_{ij}(x)\,\xi^i\xi^j \leq C^2 \sum_{i=1}^n \left(\xi^i\right)^2,$$

for all $x \in \overline{V}$ and $\xi \in T_xM$. Hence, we obtain

$$C^{-1}|\xi|_{\mathbf{e}} \leq |\xi|_{\mathbf{g}} \leq C\,|\xi|_{\mathbf{e}}.$$

It follows that, for any piecewise C^1 path γ in $\overline{V}$,

$$C^{-1}\ell_{\mathbf{e}}(\gamma) \leq \ell_{\mathbf{g}}(\gamma) \leq C\ell_{\mathbf{e}}(\gamma). \tag{1.103}$$

24.11.16

Connecting two points $x, y \in V$ by a straight line segment γ and noticing that the image of γ is contained in V and $\ell_{\mathbf{e}}(\gamma) = |x-y|$ we obtain

$$d(x,y) \leq \ell_{\mathbf{g}}(\gamma) \leq C\ell_{\mathbf{e}}(\gamma) = C\,|x-y|,$$

which proves the upper bound in (1.102).

Define now the set U by $U = B_{\frac{1}{3}r}(p)$. Let γ be any piecewise C^1 path on M connecting points $x, y \in \overline{U}$. If γ stays in V then we have $\ell_{\mathbf{e}}(\gamma) \geq |x-y|$. Combining with (1.103), we obtain

$$\ell_{\mathbf{g}}(\gamma) \geq C^{-1}|x-y|. \tag{1.104}$$

If γ does not stay in V then γ intersects the sphere ∂V, say at a point z (see Fig. 1.10). Denoting by $\widetilde{\gamma}$ be the part of γ that connects in $\overline{V}$ the point x to the point $z \in \partial V$, we obtain

$$\ell_{\mathbf{g}}(\gamma) \geq \ell_{\mathbf{g}}(\widetilde{\gamma}) \geq C^{-1}|x-z| \geq C^{-1}\frac{2}{3}r \geq C^{-1}|x-y|,$$

where we have used (1.104) for the path $\widetilde{\gamma}$ and $|x-y| \leq \frac{2}{3}r$. Hence, (1.104) holds for all paths γ connecting x and y, which implies

$$d(x,y) \geq C^{-1}|x-y|.$$

■

Corollary 1.19 *We have $d(x,y) > 0$ for all distinct points $x, y \in M$. Consequently, the geodesic distance $d(x,y)$ satisfies the axioms of a metric and, hence, (M,d) is a metric space.*

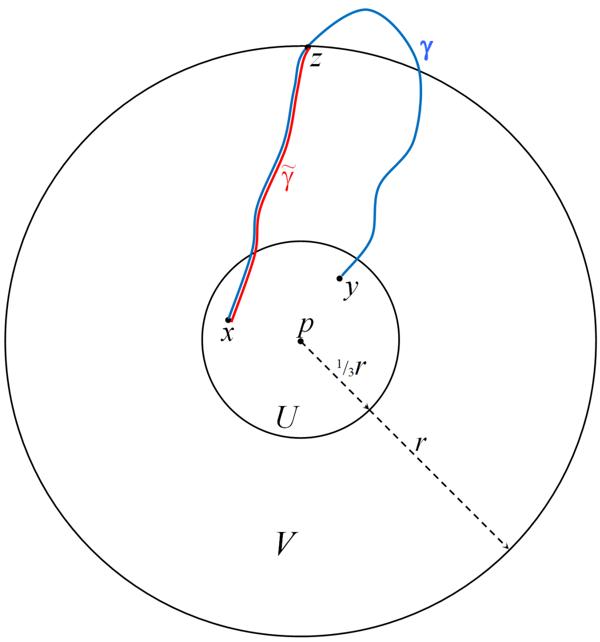

Figure 1.10: Path γ connecting the points x, y intersects ∂V at a point z.

In fact, if the manifold M is connected then the metric d is finite, that is, $d(x,y) < \infty$ for all $x, y \in M$ (see Exercise 32).

Proof. Fix a point $p \in M$ and let us prove that $d(p,x) > 0$ for any $x \neq p$. Let U be a chart around p as in Lemma 1.18. We can assume that U is a Euclidean ball $B_\varepsilon(p)$ of some radius $\varepsilon > 0$. If $x \in U$ then by (1.102)

$$d(p,x) \geq C^{-1}|p-x| > 0.$$

Assume that $x \notin U$. Then any path γ connecting p and x must intersect the boundary ∂U, say at a point z, which implies by (1.102) that

$$\ell_{\mathbf{g}}(\gamma) \geq d(p,z) \geq C^{-1}|p-z| = C^{-1}\varepsilon$$

(see Fig. 1.11)

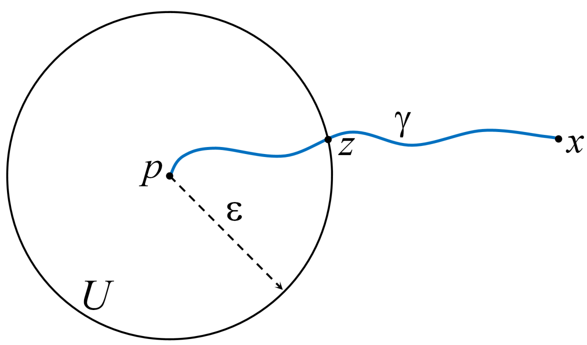

Figure 1.11: If $x \notin U$ then any path γ connecting p and x contains a point $z \in \partial U$

Taking inf in all such γ, we obtain that $d(p,x) \geq C^{-1}\varepsilon > 0$, which finishes the proof. ■

Definition. For any $x \in M$ and $r > 0$, denote by $B(x,r)$ the *geodesic ball* of radius r centered at $x \in M$, that is

$$B(x,r) = \{y \in M : d(x,y) < r\}.$$

In other words, $B(x,r)$ are the metric balls in the metric space (M,d). By definition, the topology of any metric space is generated by metric balls, which form a base of this topology. Note that the metric balls are open sets in this topology.

Theorem 1.20 *The topology of the metric space (M,d) coincides with the original topology of the smooth manifold M.*

Proof. Recall that the topology of M inside any chart U coincides with the Euclidean topology of U that is determined by the Euclidean distance function. Denote by T_M the original topology of M and by T_d – the topology of the metric space (M,d). To prove the identity of the two topologies, it suffices to prove that their local bases at any point are equivalent. A local base of T_d at a point $p \in M$ is given by the geodesic balls $B(p,r)$ with small radii $r > 0$, and a local base of T_M at p is given by the Euclidean balls $B_r(p)$ in any chart U containing p, also with small enough $r > 0$.

Hence, in order to obtained the identity of the two topologies, it suffices to prove the following: for any $p \in M$ there is a chart U containing p and $C > 1$ such that, for any small enough $r > 0$,

$$B_{C^{-1}r}(p) \subset B(p,r) \subset B_{Cr}(p). \tag{1.105}$$

Fix a point $p \in M$ and let U be a chart constructed in Lemma 1.18, where (1.102) holds. We can assume that U coincides with the Euclidean ball $B_\varepsilon(p)$ of some radius $\varepsilon > 0$. Then we will prove the inclusions (1.105) for any $r < C^{-1}\varepsilon$, where C is the constant from (1.102).

Indeed, if $x \in B_{C^{-1}r}(p)$ then $x \in U$ and

$$d(x,p) \leq C|x-p| < r,$$

whence $x \in B(p,r)$.

To prove the second inclusion in (1.105), let us first verify that $B(p,r) \subset U$. Indeed, if $x \notin U$ then any path γ connecting p and x contains a point $z \in \partial U$ (see Fig. 1.11). By (1.102), we obtain

$$\ell_{\mathbf{g}}(\gamma) \geq d(z,p) \geq C^{-1}|y-p| = C^{-1}\varepsilon \geq r,$$

whence $d(x,p) \geq r$ and $x \notin B(p,r)$. Therefore, if $x \in B(p,r)$ then $x \in U$ and, hence,

$$|x-p| \leq Cd(x,p) < Cr,$$

which implies $x \in B_{Cr}(p)$. ■

1.14 Smooth mappings and isometries

Let X and Y be two smooth manifolds of dimension n and m, respectively. A continuous mapping $\Phi : Y \to X$ is called smooth if it is represented in any charts of X and Y by smooth functions. More precisely, this means the following. Let $x^1, ..., x^n$ be the local coordinates in a chart $U \subset X$, and $y^1, ..., y^m$ be the local coordinates in a chart $V \subset Y$, and let $\Phi(V) \subset U$. Then the mapping Φ in V is given by n equations $x^i = \Phi^i(y^1, ..., y^m)$, where all functions Φ^i are smooth[6].

The mapping $\Phi : Y \to X$ allows to transfer various objects and structures either from Y to X, or back from X to Y. The corresponding operators in the case "from Y to X" are called "*push-forward*" operators and are also denoted by Φ, and in the case "from X to Y" they are called "*pullback*" operators and are denoted by Φ_*.

Definition. For any function $f : X \to \mathbb{R}$ define the pullback function $\Phi_* f : Y \to \mathbb{R}$ by

$$\Phi_* f = f \circ \Phi,$$

that is

$$(\Phi_* f)(y) = f(\Phi(y)) \text{ for any } y \in Y.$$

Clearly, if f is smooth then $\Phi_* f$ is also smooth. For example, pulling back the coordinate function x^i in chart of X, we obtain

$$\Phi_* x^i = x^i \circ \Phi = \Phi^i.$$

Now fix a point $b \in Y$ and set $a = \Phi(b) \in X$.

Definition. For any tangent vector $\xi \in T_b Y$, define its push-forward $\Phi\xi \in T_a V$ by

$$(\Phi\xi)(f) = \xi(\Phi_* f) \quad \text{for any } f \in C^\infty(X). \tag{1.106}$$

Clearly, $\Phi\xi$ is a linear mapping from $C^\infty(X)$ to $\mathbb{R}$. The fact that it satisfies the product rule and, hence, is an $\mathbb{R}$-differentiation, can be verified easily from (1.106).
29.11.16 Alternatively, we will see that by computing $\Phi\xi$ in the local coordinates as follows.

Considering f as a function of $x^1, ..., x^n$ and $\Phi_* f$ as a function of $y^1, ..., y^m$, we obtain, for any tangent vector $\xi = \xi^j \frac{\partial}{\partial y^j}$ at b, that

$$(\Phi\xi)(f) = \xi^j \frac{\partial}{\partial y^j}(\Phi_* f)\Big|_{y=b} = \xi^j \frac{\partial}{\partial y^j} f(\Phi(y))\Big|_{y=b} = \xi^j \frac{\partial \Phi^i}{\partial y^j}(b) \frac{\partial f}{\partial x^i}(a).$$

that is,

$$\Phi\xi = \xi^j \frac{\partial \Phi^i}{\partial y^j} \frac{\partial}{\partial x^i},$$

[6] By the continuity of Φ, for any $y \in Y$ and for any chart U in X containing $x := \Phi(y)$, there is a chart V in Y containing y such that $\Phi(V) \subset U$. Hence, the mapping Φ can be written in the coordinate form in a neighborhood of any point $y \in Y$.

where $\frac{\partial \Phi^i}{\partial y^j}$ is taken at b and $\frac{\partial}{\partial x^i}$ is taken at a. In particular, we see that $\Phi\xi \in T_aX$. It follows also that the components of $\Phi\xi$ in the basis $\left\{\frac{\partial}{\partial x^i}\right\}$ in T_aX are

$$\boxed{(\Phi\xi)^i = \frac{\partial \Phi^i}{\partial y^j}\xi^j}. \tag{1.107}$$

Denote by J the Jacobi matrix $\left(\frac{\partial \Phi^i}{\partial y^j}\right)$ where $i = 1, ..., n$ is the row index, $j = 1, ..., m$ is the column index. Denoting by ξ_{col} the column vector with components $\xi^1, ..., \xi^m$ and understanding $(\Phi\xi)_{\text{col}}$ similarly, (1.107) can be written in terms of matrix multiplication as follows:

$$(\Phi\xi)_{\text{col}} = J\xi_{\text{col}}. \tag{1.108}$$

Definition. The push-forward mapping

$$T_bY \ni \xi \mapsto \Phi\xi \in T_aX \tag{1.109}$$

is called the *tangent map* of Φ at b.

For convenience we use for the tangent map the same notation Φ. There are other commonly used notation for the tangent map such as $T_b\Phi$ or $d\Phi$. The tangent map (1.109) is also called the *differential* of Φ at b. The reason for the latter is the identity (1.107), where the Jacobi matrix $\left(\frac{\partial \Phi^i}{\partial y^j}\right)$ is used.

Definition. For any tangent covector $v \in T_a^*X$, define its pullback $\Phi_*v \in T_b^*Y$ by the following duality relation:

$$\langle \Phi_*v, \xi\rangle = \langle v, \Phi\xi\rangle \quad \forall \xi \in T_bY.$$

The pull-back mapping

$$T_a^*X \ni v \mapsto \Phi_*v \in T_b^*Y$$

is called the *cotangent map* of Φ at a.

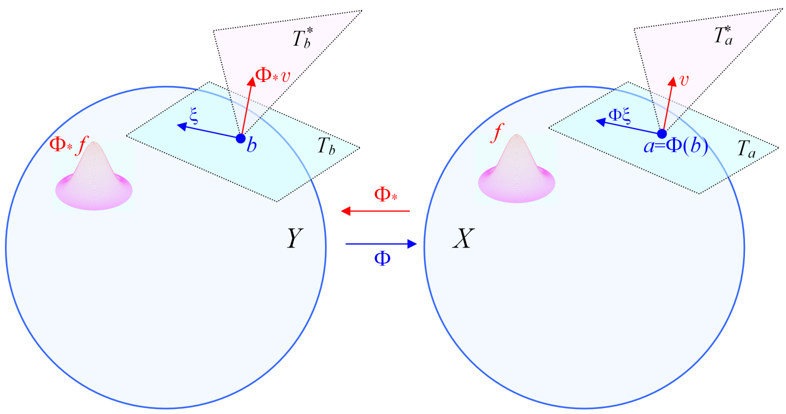

Figure 1.12: The pullback objects are red, the push-forward objects are blue.

Observe that, for any $f \in C^\infty(X)$,

$$\boxed{\Phi_*df = d(\Phi_*f)}$$

because

$$\langle \Phi_* df, \xi \rangle = \langle df, \Phi \xi \rangle = (\Phi \xi) f = \xi (\Phi_* f) = \langle d (\Phi_* f), \xi \rangle .$$

In the local coordinates, we obtain

$$\Phi_* dx^i = d \left(\Phi_* x^i \right) = d\Phi^i = \frac{\partial \Phi^i}{\partial y^j} dy^j .$$

It follows that, for any covector $v \in v_i dx^i \in T_a^* X$,

$$\Phi_* v = v_i \Phi_* dx^i = v_i \frac{\partial \Phi^i}{\partial y^j} dy^j ,$$

which implies that the pullback $\Phi_* v$ has in the basis $\{dy^j\}$ in $T_b^* Y$ the following components:

$$\boxed{(\Phi_* v)_j = v_i \frac{\partial \Phi^i}{\partial y^j}}.$$

Denoting by v_{row} the row vector with components $(v_1, ..., v_n)$ and understanding $(\Phi_* v)_{\text{row}}$ similarly, we obtain the matrix identity

$$(\Phi_* v)_{\text{row}} = v_{\text{row}} J.$$

Suppose that we have three manifolds X, Y, Z and two smooth maps

$$Z \xrightarrow{\Psi} Y \xrightarrow{\Phi} X.$$

Then pullback of functions satisfies the following identity

$$\boxed{(\Phi \circ \Psi)_* f = \Psi_* (\Phi_* f)} \quad \forall f \in C^\infty (M) ,$$

because

$$(\Phi \circ \Psi)_* f = f \circ (\Phi \circ \Psi) = (f \circ \Phi) \circ \Psi = \Psi_* (\Phi_* f) .$$

Similarly, the push-forward operation for tangent vectors satisfies the identity

$$\boxed{(\Phi \circ \Psi) \xi = \Phi (\Psi \xi)} \quad \forall \xi \in T_c Z,$$

because for any $f \in C^\infty (X)$, we have

$$(\Phi \circ \Psi) \xi (f) = \xi ((\Phi \circ \Psi)_* f) = \xi (\Psi_* (\Phi_* f)) = (\Psi \xi) (\Phi_* f) = \Phi (\Psi \xi) (f) .$$

In the same way, we obtain

$$\boxed{(\Phi \circ \Psi)_* v = \Psi_* (\Phi_* v)} \quad \forall v \in T_a^* X.$$

We notice that the push-forward of a composition is the composition of push-forwards, while the pullback of composition is the composition of pullbacks in the opposite order.

Returning to the case of one smooth mapping $\Phi : Y \to X$, assume that we are given a Riemannian metric tensor $\mathbf{g}$ on X. Then define its pullback $\Phi_* \mathbf{g}$ as a bilinear form on $T_b Y$ by

$$\Phi_* \mathbf{g} (\xi, \eta) = \mathbf{g} (\Phi \xi, \Phi \eta) \quad \text{for all } \xi, \eta \in T_b Y, \tag{1.110}$$

where $\mathbf{g}$ is taken at the point $a = \Phi(b)$ and $\Phi_*\mathbf{g}$ – at b. Obviously, $\Phi_*\mathbf{g}$ is a symmetric, non-negative definite, bilinear form on T_bY. Clearly, $\Phi_*\mathbf{g}$ at b is positive definite if and only if the tangent map $\Phi : T_bY \to T_aX$ is injective (that is, $\xi \neq 0 \Rightarrow \Phi\xi \neq 0$).

By the choice of the basis $\left\{\frac{\partial}{\partial x^i}\right\}$ in T_aX and $\left\{\frac{\partial}{\partial y^j}\right\}$ in T_bY, these spaces are identified with $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. By (1.108), the tangent map then is given by the Jacobi matrix

$$J = \left(\frac{\partial\Phi^i}{\partial y^j}\right),$$

where $i = 1, ..., n$ is the row index and $j = 1, ..., m$ is the column index. Recall that the linear mapping $J : \mathbb{R}^m \to \mathbb{R}^n$ is injective if and only if $\operatorname{rank} J = m$. Recall also that the rank of a matrix does not exceed the number of columns as well as the number of rows. If the rank of a matrix is equal to one of these numbers then the matrix is called *non-singular*. The condition $\operatorname{rank} J = m$ means that J is non-singular and $m \leq n$. Hence, we obtain the following statement.

Claim. *If $m \leq n$ and the Jacobi matrix J of $\Phi : Y \to X$ is non-singular at all points of Y then, for any Riemannian metric $\mathbf{g}$ on X, its pullback $\Phi_*\mathbf{g}$ is a Riemannian metric on Y.*

In the local coordinates, using $\mathbf{g} = g_{ij}dx^i dx^j$, we obtain

$$\Phi_*\mathbf{g} = g_{ij}\Phi_*\left(dx^i\right)\Phi_*\left(dx^j\right) = g_{ij}\frac{\partial\Phi^i}{\partial y^k}\frac{\partial\Phi^j}{\partial y^l}dy^k dy^l,$$

whence

$$(\Phi_*g)_{kl} = g_{ij}\frac{\partial\Phi^i}{\partial y^k}\frac{\partial\Phi^j}{\partial y^l}. \tag{1.111}$$

This identity can be rewritten in the form of matrix multiplication as follows:

$$\boxed{\Phi_*g = J^T g J}$$

where $g = (g_{ij})$ is an $n \times n$ matrix and $\Phi_*g = \left((\Phi_*g)_{kl}\right)$ is an $m \times m$ matrix.

Assume from now on that Y and X have the same dimension n. A mapping $\Phi : Y \to X$ is called a *diffeomorphism* if it is smooth and the inverse mapping $\Phi^{-1} : X \to Y$ exists and is also smooth. In this case, the tangent maps

$$\Phi : T_bY \to T_aX \quad \text{and} \quad \Phi^{-1} : T_aX \to T_bY$$

are also mutually inverse, which implies that the tangent map Φ is injective. Consequently, the pullback $\Phi_*\mathbf{g}$ of a Riemannian metric $\mathbf{g}$ on X is a Riemannian metric on Y.

Definition. Two Riemannian manifolds $(X, \mathbf{g}_X)$ and $(Y, \mathbf{g}_Y)$ and are called *isometric* if there is a diffeomorphism $\Phi : Y \to X$ such that

$$\Phi_*\mathbf{g}_X = \mathbf{g}_Y.$$

Such a mapping Φ is called a *Riemannian isometry.*

The relation "isometric" is denoted by the symbol $\cong$. It is easy to see that the relation $\cong$ between Riemannian manifolds is reflexive (the identity map is isometry), symmetric (because if Φ is an isometry then also Φ^{-1} is an isometry) and transitive (since the composition of two isometries is isometry). Hence, the relation $\cong$ is an equivalence relation between Riemannian manifolds. Two manifolds that are isometric have exactly the same properties as Riemannian manifolds and frequently can be regarded as the same manifold.

Definition. Two weighted manifolds $(Y, \mathbf{g}_Y, \mu_Y)$ and $(X, \mathbf{g}_X, \mu_X)$ are called *isometric* if there is a Riemannian isometry $\Phi : Y \to X$ such that

$$\Phi_* D_X = D_Y,$$

where D_X and D_Y are the density functions of μ_X and μ_Y, respectively.

Lemma 1.21 *Let Φ be an isometry of two weighted manifolds as above. Then the following is true:*

(a) *For any non-negative measurable function f on X,*

$$\int_Y (\Phi_* f)\, d\mu_Y = \int_X f\, d\mu_X. \tag{1.112}$$

(b) *For any $f \in C^\infty(X)$,*

$$\Phi_* (\Delta_X f) = \Delta_Y (\Phi_* f), \tag{1.113}$$

where Δ_Y and Δ_X are the weighted Laplace operators on Y and X, respectively.

Remark. The identity (1.112) can be rewritten as follows:

$$\int_X f(x)\, d\mu_X(x) = \int_Y f(\Phi(y))\, d\mu_Y(y),$$

and in this form it can be regarded as change of variables $x = \Phi(y)$ in integration. Note that this identity does not contain the determinant of the Jacobi matrix like in the classical formula (1.36) because the determinant is hidden in the definitions of the measures μ_X and μ_Y.

01.12.16

Proof. Because of a partition of unity, it suffices to prove the both identities (1.112) and (1.113) when f is supported in a chart U on X. Let $\varphi : U \to \mathbb{R}^n$ be a homeomorphism from U onto an open set $W \subset \mathbb{R}^n$ that exists by the definition of a chart. Denoting by $x^1, ..., x^n$ the Cartesian coordinates in W, we obtain the local coordinates $x^1, ..., x^n$ in U.

Consider the set $V = \Phi^{-1}(U) \subset Y$. Since both mappings

$$V \xrightarrow{\Phi} U \xrightarrow{\varphi} W$$

are homeomorphisms, we obtain a homeomorphism $V \xrightarrow{\psi} W$ where $\psi = \varphi \circ \Phi$, so that the Cartesian coordinates $x^1, ..., x^n$ serve also as local local coordinates in V (see Fig. 1.13).

Using in the both charts the coordinates $x^1, ..., x^n$, we obtain that the mapping $\Phi : V \to U$ in these coordinates is identical. Indeed, if a point $p \in V$ has coordinates

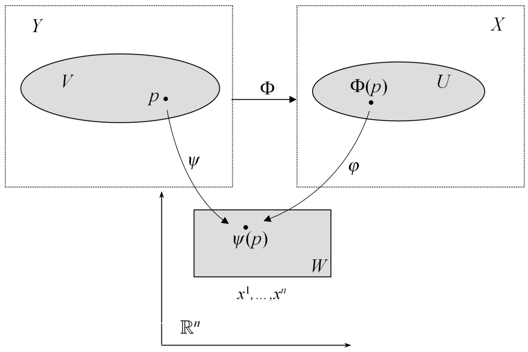

Figure 1.13: Mappings Φ, φ, ψ

$x^1, ..., x^n$ then $\psi(p)$ has in W the same coordinates, which implies that the point $\varphi^{-1}(\psi(p)) = \Phi(p)$ has in U the same coordinates.

Hence, the Riemannian metrics $\mathbf{g}_X$ and $\mathbf{g}_Y$ in the local coordinates $x^1, ..., x^n$ are identical, and so are the density functions. Then both equalities (1.112) and (1.113) are trivially satisfied. ■

Let $\Phi : M \to M$ be a diffeomorphism of a smooth manifold M. Then Φ is an isometry of a weighed manifold $(M, \mathbf{g}, \mu)$ provided

$$\Phi_* \mathbf{g} = \mathbf{g} \quad \text{and} \quad \Phi_* D = D.$$

The first of these conditions can be rewritten in the local coordinates in terms of matrices as follows:

$$J^T g J = g. \tag{1.114}$$

If Φ is an isometry of $(M, \mathbf{g}, \boldsymbol{\mu})$ then we obtain by (1.113) that $\Delta_{\mathbf{g},\mu}$ commutes with Φ_*. Alternatively, (1.113) can be written in the form

$$(\Delta_{\mathbf{g},\mu} f) \circ \Phi = \Delta_{\mathbf{g},\mu}(f \circ \Phi).$$

The set of all isometries of $(M, \mathbf{g}, \mu)$ is called the *group of isometries* of $(M, \mathbf{g}, \mu)$, because this set forms obviously a group with respect to operation of composition.

Example. Any translation $\Phi(x) = x + a$ in $\mathbb{R}^n$ is an isometry of the Riemannian manifold $(\mathbb{R}^n, \mathbf{g}_{\mathbb{R}^n})$, because the Jacobi matrix of the translation is id. Consider the orthogonal group $O(n)$, that is, the set of all $n \times n$ matrices A such that $A^T A = \mathrm{id}$ (in particular, this includes all the rotations in $\mathbb{R}^n$). If $A \in O(n)$ then the orthogonal transformation $\Phi(x) = Ax$ of $\mathbb{R}^n$ has the Jacobi matrix $J = A$. Since $g_{\mathbb{R}^n} = \mathrm{id}$, we see that (1.114) is satisfied, so that the orthogonal transformation is an isometry of $(\mathbb{R}^n, \mathbf{g}_{\mathbb{R}^n})$. Consequently, the Laplace operator in $\mathbb{R}^n$ commutes with orthogonal transformations.

Since A is invariant on $\mathbb{S}^{n-1}$, we see that the orthogonal transformation is also an isometry of $(\mathbb{S}^{n-1}, \mathbf{g}_{\mathbb{S}^{n-1}})$.

Example. Let $(M, \mathbf{g}, \mu)$ be a weighted model with polar coordinates (r, θ) (see Section 1.12) and let Φ be an isometry of $\mathbb{S}^{n-1}$. Then Φ induces an isometry of $(M, \mathbf{g}, \mu)$ by

$$\Phi(r, \theta) = (r, \Phi(\theta)).$$

In particular, $\Delta_{\mathbf{g},\mu}$ commutes with the rotations of the polar angle θ.

Chapter 2

Weak Laplace operator and spectrum

2.1 Regularity theory in $\mathbb{R}^n$

Consider in a domain $\Omega \subset \mathbb{R}^n$ an operator

$$Lu = \sum_{i,j=1}^{n} \partial_i \left(a_{ij}\partial_j u\right) + \sum_{i=1}^{n} b_j \partial_j u + cu, \tag{2.1}$$

where the coefficients a_{ij}, b_j, c are C^∞ smooth functions in Ω. Assume that (a_{ij}) is uniformly elliptic with the ellipticity constant λ and that the coefficients b_j, c are bounded in Ω, say, also by λ.

Definition. For any $u \in W^{1,2}_{loc}(\Omega)$ and $f \in L^2_{loc}(\Omega)$, we say that the equation $Lu = f$ holds weakly in Ω if, for any $\varphi \in \mathcal{D}(\Omega) := C_0^\infty(\Omega)$,

$$-\int_\Omega \sum_{i,j=1}^{n} a_{ij}\partial_j u\, \partial_i \varphi\, dx + \int_\Omega \sum_{i=1}^{n} b_j \partial_j u\, \varphi dx + \int_\Omega cudx = \int_\Omega f\varphi dx.$$

The following theorem was proved in *EDE*, Theorems 2.1, 2.8, 2.10.

Theorem 2.1 *Let L be the operator (2.1). If $u \in W^{1,2}_{loc}(\Omega)$ and $Lu \in W^{k,2}_{loc}(\Omega)$ then $u \in W^{k+2,2}_{loc}(\Omega)$. Moreover, for any open set $U \Subset \Omega$,*

$$\|u\|_{W^{k+2,2}(U)} \leq C\left(\|u\|_{W^{1,2}(\Omega)} + \|Lu\|_{W^{k,2}(\Omega)}\right), \tag{2.2}$$

where $C = C(U, \Omega, n, \lambda)$.

Consider a more general operator

$$Lu = \frac{1}{\rho(x)} \sum_{i,j=1}^{n} \partial_i \left(a_{ij}\partial_j u\right) + \sum_{i=1}^{n} b_j \partial_j u + cu, \tag{2.3}$$

where a_{ij}, b_j and c are as above and ρ is a smooth positive function in Ω that is bounded between two positive constants. We say that the equation $Lu = f$ holds weakly in Ω if, for any $\varphi \in \mathcal{D}(\Omega)$,

$$-\int_\Omega \sum_{i,j=1}^n a_{ij}\partial_j u\, \partial_i\left(\frac{1}{\rho}\varphi\right)dx + \int_\Omega \sum_{i=1}^n b_j\partial_j u\,\varphi dx + \int_\Omega cudx = \int_\Omega f\varphi dx. \tag{2.4}$$

Consider also an auxiliary operator

$$\widetilde{L}u = \sum_{i,j=1}^n \partial_i\left(a_{ij}\partial_j u\right) + \sum_{i=1}^n \rho b_j\partial_j u + \rho cu,$$

where all terms are obtain by multiplying those of Lu by ρ. Then $Lu = f$ is equivalent to $\widetilde{L}u = \rho f$, which can be seen by replacing the test function φ in (2.4) by $\psi = \varphi/\rho$.

Corollary 2.2 *Let L be the operator (2.3). If $u \in W^{1,2}_{loc}(\Omega)$ and $Lu \in W^{k,2}_{loc}(\Omega)$ then $u \in W^{k+2,2}_{loc}(\Omega)$. Moreover, for any open set $U \Subset \Omega$,*

$$\|u\|_{W^{k+2,2}(U)} \leq C\left(\|u\|_{W^{1,2}(\Omega)} + \|Lu\|_{W^{k,2}(\Omega)}\right), \tag{2.5}$$

where $C = C(U, \Omega, n, k, \lambda, \rho)$.

Proof. If $Lu = f$ where $f \in W^{k,2}_{loc}$ then also $\widetilde{L}u = \rho f \in W^{k,2}_{loc}$. Since Theorem 2.1 applies to the operator $\widetilde{L}$, we conclude that $u \in W^{k+2,2}_{loc}(\Omega)$. Choose an open set V such that $U \Subset V \Subset \Omega$. Applying (2.2) to the operator $\widetilde{L}$ in V, we obtain

$$\|u\|_{W^{k+2,2}(U)} \leq C\left(\|u\|_{W^{1,2}(V)} + \|\rho f\|_{W^{k,2}(V)}\right). \tag{2.6}$$

Since the function ρ and all its derivatives are bounded in V, it follows that

$$\|\rho f\|_{W^{k,2}(V)} \leq C'\|f\|_{W^{k,2}(V)},$$

where C' depends on $\|\rho\|_{C^k(V)} < \infty$. Substituting into (2.6), we obtain

$$\|u\|_{W^{k+2,2}(U)} \leq C''\left(\|u\|_{W^{1,2}(V)} + \|f\|_{W^{k,2}(V)}\right),$$

whence (2.5) follows. ■

In what follows we use the notation

$$L^k u := \underbrace{L\left(L\left(...Lu\right)\right)}_{k \text{ times } L}.$$

Assuming $u \in W^{1,2}_{loc}(\Omega)$ and $f \in L^2_{loc}(\Omega)$, let us define by induction in $k \in \mathbb{N}$ what it means that $L^k u = f$ weakly in Ω. If $k = 1$ then $L^k u = f$ is the same as $Lu = f$ that was defined above. If $k > 1$ then $L^k u = f$ means that $L^{k-1}u \in W^{1,2}_{loc}(\Omega)$ (which is defined by the inductive hypotheses), and $L^k u = f$ means that $L\left(L^{k-1}u\right) = f$.

Consequently, the equality $L^k u = f$ assumes that all the functions $u, Lu, ..., L^{k-1}u$ belong to $W^{1,2}_{loc}(\Omega)$, and $L\left(L^{k-1}u\right) = f$.

Corollary 2.3 *Let L be the operator (2.3). If*

$$u, Lu, ..., L^k u \in W^{1,2}_{loc}(\Omega)$$

then

$$u \in W^{2k+1,2}_{loc}(\Omega).$$

Moreover, for any open $U \Subset \Omega$,

$$\|u\|_{W^{2k+1,2}(U)} \le C \sum_{j=0}^{k} \left\| L^j u \right\|_{W^{1,2}(\Omega)}, \tag{2.7}$$

where $C = C(U, \Omega, n, k, \lambda, \rho)$.

Proof. Induction in k. If $k = 0$ then the statement is trivial. Induction step from $k-1$ to k, where $k \ge 1$. Set $v = Lu$. Then we have $L^{k-1} v \in W^{1,2}_{loc}(\Omega)$, and by the inductive hypothesis we conclude that $v \in W^{2k-1,2}_{loc}(\Omega)$. Moreover, choosing an open set V such that $U \Subset V \Subset \Omega$, we obtain

$$\|v\|_{W^{2k-1,2}(V)} \le C \sum_{j=0}^{k-1} \left\| L^j v \right\|_{W^{1,2}(\Omega)} = C \sum_{j=1}^{k} \left\| L^j u \right\|_{W^{1,2}(\Omega)}.$$

Therefore, $Lu = v \in W^{2k-1,2}_{loc}(\Omega)$, and by Corollary 2.2 we conclude that $u \in W^{2k+1,2}_{loc}(\Omega)$ and

$$\|u\|_{W^{2k+1,2}(U)} \le C \left(\|u\|_{W^{1,2}(V)} + \|v\|_{W^{2k-1,2}(V)} \right) \le C' \sum_{j=0}^{k} \left\| L^j u \right\|_{W^{1,2}(\Omega)},$$

which was to be proved. ■

Corollary 2.4 *Let L be the operator (2.3). If*

$$u, Lu, ..., L^k u \in W^{1,2}_{loc}(\Omega)$$

and, for some non-negative integer m,

$$2k + 1 > \frac{n}{2} + m,$$

then

$$u \in C^m(\Omega).$$

Moreover, for any compact set $K \subset \Omega$,

$$\|u\|_{C^m(K)} \le C \sum_{j=0}^{k} \left\| L^j u \right\|_{W^{1,2}(\Omega)}, \tag{2.8}$$

where $C = C(K, \Omega, n, k, m, \lambda, \rho)$.

Proof. Recall the Sobolev Embedding Theorem (PDE, Theorem 4.15): if U is an open subset of $\mathbb{R}^n$ and $l > m + \frac{n}{2}$ then we have an embedding

$$W^{l,2}_{loc}(U) \hookrightarrow C^m(U).$$

Moreover, if $u \in W^{l,2}(\Omega)$ then, for any compact $K \subset U$,

$$\|u\|_{C^m(K)} \le C\,\|u\|_{W^{l,2}(U)}. \tag{2.9}$$

Since $u \in W^{2k+1,2}_{loc}(\Omega)$, applying the first statement with $l = 2k+1$ and $U = \Omega$, we obtain $u \in C^m(\Omega)$. Now let U be any open neighborhood of K such that $U \Subset \Omega$. By (2.7) we have

$$\|u\|_{W^{2k+1,2}(U)} \le C \sum_{j=0}^{k} \left\|L^j u\right\|_{W^{1,2}(\Omega)},$$

which together with (2.9) implies (2.8). ■

06.12.16

2.2 Weak gradient and Sobolev spaces

Let $(M, \mathbf{g}, \mu)$ be a weighted manifold. Denote for simplicity by $\mathcal{D}(M)$ the space $C_0^\infty(M)$ and by $\vec{\mathcal{D}}(M)$ the space of smooth vector fields on M with compact supports. Denote by $\vec{L}^2(M,\mu)$ the space of all vector fields $v(x)$ on M such that all the components of v are measurable functions in all charts, and $|\nabla v|_{\mathbf{g}} \in L^2(M,\mu)$. Similarly we define $\vec{L}^2_{loc}$.

In what follows we write for simplicity $\nabla_{\mathbf{g}} = \nabla$ and $\operatorname{div}_{\mathbf{g},\mu} = \operatorname{div}$.

The space $\vec{L}^2(M,\mu)$ admits an inner product

$$(v,w)_{\vec{L}^2} := \int_M \langle v, w\rangle_{\mathbf{g}}\, d\mu,$$

and the corresponding norm is

$$\|v\|^2_{\vec{L}^2} = \int_M |\nabla v|^2_{\mathbf{g}}\, d\mu.$$

It is easy to prove that $\vec{L}^2$ is complete with respect to this norm and, hence, is a Hilbert space.

Definition. Fix a function $u \in L^2_{loc}$. A *weak gradient* of u is a vector field $v \in \vec{L}^2_{loc}$ (denoted also ∇u) such that, for any $\psi \in \vec{\mathcal{D}}$,

$$\int_M u \operatorname{div}\psi\, d\mu = -\int_M \langle v, \psi\rangle_{\mathbf{g}}\, d\mu. \tag{2.10}$$

Or, equivalently, ∇u is defined by the identity

$$(u, \operatorname{div}\psi)_{L^2} = -(\nabla u, \psi)_{\vec{L}^2}.$$

It follows from this definition that the weak gradient is uniquely defined. Note u is a smooth function then the classical gradient $v = \nabla u$ satisfies (2.10) by the divergence theorem, so that in this case the weak gradient exists and coincides with the classical gradient.

Definition. Define the Sobolev space:

$$W^1(M) = W^1(M, \mathbf{g}, \mu) = \left\{ u \in L^2(M, \mu) : \nabla u \in \vec{L}^2(M, \mu) \right\}$$

and the inner product in W^1:

$$(u, v)_{W^1} := (u, v)_{L^2} + (\nabla u, \nabla v)_{\vec{L}^2} = \int_M uv \, d\mu + \int_M \langle \nabla u, \nabla v \rangle_{\mathbf{g}} \, d\mu. \tag{2.11}$$

The associated norm given by

$$\|u\|^2_{W^1} = \|u\|^2_{L^2} + \|\nabla u\|^2_{L^2} = \int_M u^2 d\mu + \int_M |\nabla u|^2_{\mathbf{g}} \, d\mu. \tag{2.12}$$

Lemma 2.5 *$W^1(M)$ is a Hilbert space.*

Proof. It follows from (2.12) that the convergence $u_k \xrightarrow{W^1} u$ in $W^1(M)$ is equivalent to

$$u_k \xrightarrow{L^2} u \quad \text{and} \quad \nabla u_k \xrightarrow{L^2} \nabla u. \tag{2.13}$$

Let $\{u_k\}$ be a Cauchy sequence in $W^1(M)$. Then the sequence $\{u_k\}$ is Cauchy also in $L^2(M)$ and, hence, converges in L^2-norm to a function $u \in L^2(M)$. Similarly, the sequence $\{\nabla u_k\}$ is Cauchy in $\vec{L}^2(M)$ and, hence, converges in $\vec{L}^2$-norm to a vector field $v \in \vec{L}^2(M)$. It follows from the definition of the weak gradient that $\nabla u = v$ and, hence, (2.13) is satisfied. ∎

Since any open set $U \subset M$ is itself a manifold, we can define all the above spaces L^2 and W^1 also in U.

If U is in addition a chart, then we can define the spaces L^2 and W^1 considering U as a subset of $\mathbb{R}^n$, that is, using the Euclidean metric $\mathbf{e} = \mathbf{g}_{\mathbb{R}^n}$ and the Lebesgue measure λ. Denote these spaces by $L^2_{\mathbf{e}}$ and $W^1_{\mathbf{e}}$, respectively.

We say that a chart U on M is precompact if U as a set is precompact and $\overline{U}$ is contained in a larger chart.

Lemma 2.6 *If U is a precompact chart then*

$$\begin{aligned} L^2(U) &= L^2_{\mathbf{e}}(U) \\ W^1(U) &= W^1_{\mathbf{e}}(U) \end{aligned}$$

and

$$\begin{aligned} \|u\|_{L^2(U)} &\simeq \|u\|_{L^2_{\mathbf{e}}(U)} \\ \|u\|_{W^1(U)} &\simeq \|u\|_{W^1_{\mathbf{e}}(U)}. \end{aligned}$$

The sign $\simeq$ between two expressions means that the two expressions are comparable, that is, their ratio is bounded from above and below by positive constants.

Proof. In the chart U we have

$$d\mu = D\sqrt{\det g}d\lambda = \rho d\lambda$$

where λ is the Lebesgue measure in U and D is the density function. Since the function $\rho := D\sqrt{\det g}$ is bounded between two positive constants in U, we see that

$$\|u\|_{L^2(U)} \simeq \|u\|_{L^2_{\mathbf{e}}(U)}$$

and, hence,

$$L^2(U) = L^2_{\mathbf{e}}(U).$$

Let v be a measurable vector field on U. Denote by v^i be the coordinates of v in the basis $\left\{\frac{\partial}{\partial x^i}\right\}$. Let also $v_i = g_{ij}v^j$ be the covector coordinates of v. Note that

$$\|v\|^2_{\vec{L}^2} = \int_U |v|^2_{\mathbf{g}}\, d\mu = \int_U g^{ij}v_iv_j\rho d\lambda.$$

Since on U the matrix (ρg^{ij}) is uniformly elliptic, we obtain that in U

$$\rho g^{ij}v_iv_j \simeq v_1^2 + ... + v_n^2,$$

whence

$$\|v\|^2_{\vec{L}^2} \simeq \int_U \left(v_1^2 + ... + v_n^2\right) d\lambda = \|v\|^2_{\vec{L}^2_{\mathbf{e}}}.$$

Hence, identifying each vector field v in U with the Euclidean vector field $\{v_1, ..., v_n\}$, we obtain the identity

$$\vec{L}^2(U) = \vec{L}^2_{\mathbf{e}}(U).$$

Let $u \in W^1(U)$ and let $v = \nabla u$ be its weak gradient. For any $\psi = \psi^i \frac{\partial}{\partial x^i} \in \vec{\mathcal{D}}(U)$, we have

$$\int_M u \operatorname{div}\psi\, d\mu = \int_U u\frac{1}{\rho}\frac{\partial}{\partial x^i}\left(\rho\psi^i\right)\rho d\lambda = \int_U u\frac{\partial}{\partial x^i}\left(\rho\psi^i\right)d\lambda$$

and

$$\int_M \langle v, \psi\rangle_{\mathbf{g}}\, d\mu = \int_U g_{ij}v^j\psi^i\rho d\lambda = \int_U v_i\psi^i d\lambda.$$

By (2.10) we obtain

$$\int_U u\frac{\partial}{\partial x^i}\left(\rho\psi^i\right)d\lambda = -\int_U v_i\left(\rho\psi^i\right)d\lambda. \tag{2.14}$$

Fix $\varphi \in \mathcal{D}(U)$ and an index i. Choose then the vector field ψ as follows: $\psi_i = \varphi/\rho$ and $\psi_j = 0$ for all $j \neq i$. It follows from (2.14) that

$$\int_U u\frac{\partial\varphi}{\partial x^i}d\lambda = -\int_U v_i\varphi d\lambda,$$

that is, the function v_i satisfies the definition of the the weak derivative $\frac{\partial u}{\partial x^i}$ (in $U \subset \mathbb{R}^n$), so that

$$\frac{\partial u}{\partial x^i} = v_i.$$

Since $u \in L^2_{\mathbf{e}}(U)$ and $\{v_i\} \in \vec{L}^2_{\mathbf{e}}$, we obtain that $u \in W^1_{\mathbf{e}}(U)$ and

$$\|u\|^2_{W^1_{\mathbf{e}}} = \|u\|^2_{L^2_{\mathbf{e}}} + \|\{v_i\}\|^2_{\vec{L}^2_{\mathbf{e}}} \simeq \|u\|^2_{L^2} + \|v\|^2_{\vec{L}^2} = \|u\|^2_{W^1}.$$

Conversely, if $u \in W^1_{\mathbf{e}}(U)$ then $\left\{\frac{\partial u}{\partial x^i}\right\} \in \vec{L}^2_{\mathbf{e}}(U)$ and

$$\int_U u \frac{\partial}{\partial x^i}\left(\rho\psi^i\right) d\lambda = -\int_U \frac{\partial u}{\partial x^i}\left(\rho\psi^j\right) d\lambda.$$

Hence, the vector field v with covector components

$$v_i = \frac{\partial u}{\partial x^i}$$

belongs to $\vec{L}^2(U)$ and yields the weak gradient ∇u, which implies $u \in W^1(U)$. We conclude that

$$W^1_{\mathbf{e}}(U) = W^1(U).$$

■

2.3 Weak Laplacian

Here we write for simplicity $\Delta_{\mathbf{g},\mu} = \Delta$.

Definition. Let $u \in W^1_{loc}(\Omega)$ and $f \in L^2_{loc}(\Omega)$. We say that the equation $\Delta u = f$ is satisfied weakly in Ω, if, for any $\varphi \in \mathcal{D}(\Omega)$,

$$\int_\Omega \langle \nabla u, \nabla \varphi \rangle_{\mathbf{g}}\, d\mu = -\int_\Omega f\, \varphi d\mu, \tag{2.15}$$

that is,

$$(\nabla u, \nabla \varphi)_{\vec{L}^2} = -(f, \varphi)_{L^2}.$$

Of course, if u is a smooth function and $\Delta u = f$ is satisfied in the classical sense, then it is also satisfied in the weak sense, as it follows from the Green formula.

Theorem 2.7 *Let k, m be non-negative integers such that*

$$2k + 1 > m + \frac{n}{2}.$$

Let Ω be an open subset of a weighted manifold M and assume that

$$u, \Delta u, ..., \Delta^k u \in W^1_{loc}(\Omega). \tag{2.16}$$

Then $u \in C^m(\Omega)$. Moreover, for any compact set $K \subset \Omega$, that is contained in a chart,

$$\|u\|_{C^m(K)} \leq C \sum_{j=0}^k \left\|\Delta^j u\right\|_{W^1(\Omega)}, \tag{2.17}$$

where $C = C(K, \Omega, n, k, m, \mathbf{g}, D)$.

Proof. Let us choose a precompact chart U such that $K \subset U \Subset \Omega$. Let $x^1, ..., x^n$ be the coordinates in U. Consider in U the following differential operator

$$L = \frac{1}{\rho} \frac{\partial}{\partial x^i} \left(\rho g^{ij} \frac{\partial}{\partial x^j} \right),$$

where $\rho = D\sqrt{\det g}$. We know that, for a smooth function u, $Lu = \Delta u$ in U. Let us show that $Lu = \Delta u$ holds also if L and Δ are understood weakly.

As it follows from Lemma 2.6,

$$L^2_{loc}(U) = L^2_{\mathbf{e},loc}(U) \quad \text{and} \quad W^1_{loc}(U) = W^1_{\mathbf{e},loc}(U).$$

Assume $u \in W^1_{loc}(U)$ and $f \in L^2_{loc}(U)$. By (2.4), the equation $Lu = f$ weakly in U means that, for all $\psi \in \mathcal{D}(U)$

$$\int_U \sum_{i,j=1}^n \rho g^{ij} \partial_j u \, \partial_i \left(\frac{1}{\rho} \psi \right) d\lambda = - \int_U f \psi d\lambda. \tag{2.18}$$

By (2.15) the equation $\Delta u = f$ weakly in Umeans that, for all $\varphi \in \mathcal{D}(U)$,

$$\int_U \langle \nabla u, \nabla \varphi \rangle_{\mathbf{g}} \, d\mu = - \int_U f \, \varphi d\mu. \tag{2.19}$$

Since

$$\langle \nabla u, \nabla \varphi \rangle_{\mathbf{g}} = \langle du, \nabla \varphi \rangle = g^{ij} \frac{\partial u}{\partial x^j} \frac{\partial \varphi}{\partial x^i}$$

and

$$d\mu = \rho d\lambda,$$

we obtain that (2.19) is equivalent to

$$\int_U \rho g^{ij} \frac{\partial u}{\partial x^j} \frac{\partial \varphi}{\partial x^i} d\lambda = - \int_U \rho f \varphi d\lambda.$$

The change $\psi = \rho \varphi$ shows that the latter identity is equivalent to (2.18), which proves that the weak operators Δu and Lu are the same.

If (2.16) is satisfied then

$$u, \Delta u, ..., \Delta^k u \in W^1(U),$$

whence also

$$u, Lu, ..., L^k u \in W^1_{\mathbf{e}}(U).$$

Since $2k + 1 > \frac{n}{2} + m$, we obtain by Corollary 2.4 that $u \in C^m(U)$ and

$$\|u\|_{C^m(K)} \leq C \sum_{j=0}^k \left\| L^j u \right\|_{W^1_{\mathbf{e}}(U)}. \tag{2.20}$$

Since Ω can be covered by charts like U, we conclude that $u \in C^m(\Omega)$. The estimate (2.20) and $\|\cdot\|_{W^1_{\mathbf{e}}(U)} \simeq \|\cdot\|_{W^1(U)}$ (cf. Lemma 2.6) imply (2.17). ■

Remark. In the case $m = 0$ the norm

$$\|u\|_{C(K)} := \sup_K |u|$$

makes sense for any compact set K, not necessarily contained in a chart. In this case the estimate (2.17) holds also for any compact set K, because the latter can be covered by a finite number of precompact charts, and in each of them we can apply Theorem 2.7.

08.12.16

As an example of application of Theorem 2.7, let us prove the following statement.

Corollary 2.8 *Let a function $u \in W^1(\Omega)$ satisfy in Ω the equation $\Delta u = \alpha u$ in a weak sense, where α is a real number. Then $u \in C^\infty(\Omega)$. Moreover, for any compact set $K \subset \Omega$ that is contained in a chart, and for any non-negative integer m, we have*

$$\|u\|_{C^m(K)} \le C(1 + |\alpha|)^{\frac{m}{2}+\frac{n}{4}+\frac{1}{2}} \|u\|_{W^1(\Omega)}, \tag{2.21}$$

where $C = C(K, \Omega, n, m, \mathbf{g}, D)$.

Proof. We have $u \in W^1(\Omega)$ and $\Delta u = \alpha u \in W^1(\Omega)$. It follows that also $\Delta^2 u = \alpha^2 u \in W^1(\Omega)$ and, by induction, for any positive integer j, we obtain

$$\Delta^j u = \alpha^j u \in W^1(\Omega).$$

By Theorem 2.7 we conclude that $u \in C^\infty(\Omega)$.

By the estimate (2.17) of that theorem, we have, for any non-negative integers m, k such that $2k + 1 > \frac{n}{2} + m$, that

$$\|u\|_{C^m(K)} \le C \sum_{j=0}^{k} \left\|\Delta^j u\right\|_{W^1(\Omega)},$$

Since

$$\left\|\Delta^j u\right\|_{W^1} = |\alpha|^j \|u\|_{W^1},$$

it follows that

$$\|u\|_{C^m(K)} \le C \sum_{j=0}^{k} |\alpha|^j \|u\|_{W^1} \le C(1 + |\alpha|)^k \|u\|_{W^1}.$$

Choose k to be the smallest integer such that $2k + 1 > m + \frac{n}{2}$. Then

$$2k - 1 \le m + \frac{n}{2},$$

and, hence,

$$k \le \frac{m}{2} + \frac{n}{4} + \frac{1}{2},$$

whence (2.21) follows. ∎

Example. Consider the equation $\Delta u = \alpha u$ in $\Omega = (0, 2\pi)$. It becomes $u'' = \alpha u$ and if $\alpha < 0$ then one of the solution is $u(x) = \sin \beta x$ where $\beta = \sqrt{-\alpha}$. In this case

$$\|u\|_{L^2}^2 = \int_0^{2\pi} \sin^2 \beta x dx = \pi, \quad \|u'\|_{L^2}^2 = \beta^2 \int_0^{2\pi} \cos^2 \beta x dx = \beta^2 \pi = |\alpha| \pi$$

and

$$\|u\|_{W^1} = (1 + |\alpha|)^{1/2} \pi^{1/2}.$$

Assume that $|\alpha| \geq 1$ and, hence, $\beta \geq 1$. Then the functions $|\sin \beta x|$ and $|\cos \beta x|$ attain their maximum value 1 on $(0, 2\pi)$. Since

$$u^{(j)}(x) = \pm \beta^j \sin \beta x \quad \text{or} \quad \pm \beta^j \cos \beta x,$$

it follows that

$$\|u\|_{C^m(0,2\pi)} = \sup_{0 \leq j \leq m} \sup_{(0,2\pi)} \left|u^{(j)}\right| = \sup_{0 \leq j \leq m} \beta^j = \beta^m = |\alpha|^{m/2}.$$

It follows that

$$\|u\|_{C^m(0,2\pi)} \simeq (1 + |\alpha|)^{m/2 - 1/2} \|u\|_{W^1},$$

which shows that the exponent $m/2$ in (2.21) is correct.

2.4 Compact embedding theorem

Let Ω be an open subset of M. Clearly, $\mathcal{D}(\Omega) \subset W^1(\Omega)$. Define

$$W_0^1(M) = \text{the closure of } \mathcal{D}(\Omega) \text{ in } W^1(\Omega).$$

Theorem 2.9 (Compact embedding theorem) *If Ω is a precompact open subset of M then the identical embedding*

$$W_0^1(\Omega) \hookrightarrow L^2(\Omega)$$

is a compact operator.

Proof. In the case when $M = \mathbb{R}^n$ this theorem is known (Theorem 4.6 from PDE), and we will use it in order to prove that on an arbitrary manifold.

We need to prove that, for any bounded sequence $\{f_k\}$ in $W_0^1(\Omega)$, there is a subsequence $\{f_{k_i}\}$ that converges in $L^2(\Omega)$. Since $\mathcal{D}(\Omega)$ is dense in $W_0^1(\Omega)$, we can assume without loss of generality that all the functions f_k are in $\mathcal{D}(\Omega)$. Since $\Omega \subset M$ is relatively compact, there is a *finite* family $\{U_j\}_{j=1}^N$ of precompact charts such that

$$\overline{\Omega} \subset \bigcup_{j=1}^N U_j.$$

By Theorem 1.3, there exists a partition of unity at $\overline{\Omega}$ subordinate to $\{U_j\}$, that is, non-negative functions $\varphi_j \in \mathcal{D}(U_j)$ such that $\sum_{j=1}^N \varphi_j \equiv 1$ in $\overline{\Omega}$.

Let us prove that, for any j, the sequence $\left\{f_k \varphi_j\right\}_{k=1}^\infty$ is bounded in $W^1(\Omega)$. Indeed, suppressing indices k, j, we have

$$\|f\varphi\|_{L^2} \leq \sup|\varphi| \|f\|_{L^2} \leq \|f\|_{L^2} \leq \|f\|_{W^1},$$

$$\begin{aligned}\|\nabla (f\varphi)\|_{\vec{L}^2} &= \|\varphi \nabla f + f\nabla\varphi\|_{\vec{L}^2} \\ &\leq \sup\varphi \, \|\nabla f\|_{\vec{L}^2} + \sup|\nabla\varphi| \, \|f\|_{L^2} \\ &\leq C\,\|f\|_{W^1},\end{aligned}$$

where $C = 1 + \sup|\nabla\varphi| < \infty$. It follows that

$$\left\|f_k\varphi_j\right\|_{W^1(\Omega)} \leq C'\,\|f_k\|_{W^1(\Omega)},$$

which implies that, for any j, the sequence $\left\{f_k\varphi_j\right\}_{k=1}^{\infty}$ is bounded in $W^1(\Omega)$. It follows that this sequence is bounded also in $W^1(U_j) = W^1_{\mathbf{e}}(U_j)$. Since $f_k\varphi_j \in \mathcal{D}(U_j) \subset W_0^1(U_j)$, we can use the compact embedding theorem in $\mathbb{R}^n$ and conclude that there is a subsequence $\left\{f_{k_i}\varphi_j\right\}_{i=1}^{\infty}$ that converges in $L^2_{\mathbf{e}}(U_j) = L^2(U_j)$. By extending the limit function by 0 outside U_j, we obtain that $\left\{f_{k_i}\varphi_j\right\}_{i=1}^{\infty}$ converges in $L^2(\Omega)$.

Applying this procedure successively for each $j = 1, ..., N$, we obtain a subsequence $\{f_{k_i}\}$ such that $\left\{f_{k_i}\varphi_j\right\}_{i=1}^{\infty}$ converges in $L^2(\Omega)$ for *any* j. Since $\sum_{j=1}^{N}\varphi_j \equiv 1$ in Ω, we conclude that $\{f_{k_i}\}$ converges in $L^2(\Omega)$, which finishes the proof. ■

2.5 Resolvent operator

Fix an open set $\Omega \subset M$ and consider the following Dirichlet problem

$$\begin{cases} \Delta u - \alpha u = -f & \text{in } \Omega \\ u \in W_0^1(\Omega) \end{cases} \tag{2.22}$$

where α is a real parameter and f is a given function from $L^2(\Omega)$. A function $u \in W_0^1(\Omega)$ is called a *weak solution* of (2.22) if $\Delta u = \alpha u + f$ weakly in Ω; equivalently, this means that, for any $\varphi \in \mathcal{D}(\Omega)$,

$$(\nabla u, \nabla\varphi)_{\vec{L}^2} + \alpha\,(u,\varphi)_{L^2} = (f,\varphi)_{L^2}. \tag{2.23}$$

Theorem 2.10 (a) *For any $\alpha > 0$ and for any $f \in L^2(\Omega)$, the problem (2.22) has a unique solution u.*

(b) *Define the resolvent operator R_α by*

$$\begin{aligned} R_\alpha \;:\; & L^2(\Omega) \to L^2(\Omega) \\ & R_\alpha f = u, \end{aligned}$$

where u is the solution of (2.22). Then the operator R_α is linear, bounded with $\|R_\alpha\| \leq \alpha^{-1}$, injective, positive definite, self-adjoint operator in $L^2(\Omega)$.

(c) *If Ω is precompact then the operator R_α is compact.*

Proof. (a) All terms in (2.23) are bounded linear functionals of $\varphi \in W^1(\Omega)$, because

$$|(f,\varphi)| \leq \|f\|_{L^2}\|\varphi\|_{L^2} \leq \|f\|_{L^2}\|\varphi\|_{W^1}$$

and similarly

$$|(\nabla u, \nabla\varphi)_{\vec{L}^2}| \leq \|\nabla u\|_{\vec{L}^2}\,\|\nabla\varphi\|_{\vec{L}^2} \leq \|u\|_{W^1}\,\|\varphi\|_{W^1}.$$

Hence, all terms in (2.23) are continuous in $\varphi \in W^1(\Omega)$. If (2.23) holds for all $\varphi \in \mathcal{D}(\Omega)$, then it holds also for all $\varphi \in W_0^1(\Omega)$ because $\mathcal{D}(\Omega)$ is dense in $W_0^1(\Omega)$.

Denote the left hand side of (2.23) by $[u,\varphi]_\alpha$, that is,

$$[u,\varphi]_\alpha := (\nabla u, \nabla \varphi)_{\vec{L}^2} + \alpha (u,\varphi)_{L^2},$$

and observe that $[\cdot,\cdot]_\alpha$ is an inner product in W_0^1. If $\alpha = 1$ then $[\cdot,\cdot]_\alpha$ coincides with the standard inner product in W_0^1. For any $\alpha > 0$ and $u \in W_0^1$, we have

$$\min(\alpha, 1) \|u\|_{W^1}^2 \le [u,u]_\alpha \le \max(\alpha,1) \|u\|_{W^1}^2,$$

or shortly

$$[u,u]_\alpha \simeq \|u\|_{W^1}^2.$$

Therefore, the space W_0^1 with the inner product $[\cdot,\cdot]_\alpha$ is complete.

Rewrite the equation (2.23) in the form

$$[u,\varphi]_\alpha = (f,\varphi)_{L^2} \quad \forall \varphi \in W_0^1(\Omega). \tag{2.24}$$

Since the right hand side $\varphi \mapsto (f,\varphi)_{L^2}$ is a bounded functional of $\varphi \in W_0^1$, the equation (2.24) has a unique solution $u \in W_0^1(\Omega)$ by the Riesz representation theorem[1].

(b) Substituting $\varphi = u$ in (2.23) we obtain

$$\|\nabla u\|_{L^2}^2 + \alpha \|u\|_{L^2}^2 = (f,u)_{L^2}. \tag{2.25}$$

It follows that

$$\alpha \|u\|_{L^2}^2 \le \|f\|_{L^2} \|u\|_{L^2},$$

which implies $\|u\|_{L^2} \le \alpha^{-1} \|f\|_{L^2}$ and, hence,

$$\|R_\alpha\| := \sup_{f \in L^2 \setminus \{0\}} \frac{\|R_\alpha f\|_{L^2}}{\|f\|_{L^2}} \le \alpha^{-1} < \infty.$$

Hence, R_α is bounded.

If $u = R_\alpha f = 0$ then we obtain from (2.23) that $(f,\varphi)_{L^2} = 0$ for all $\varphi \in \mathcal{D}(\Omega)$. Since $\mathcal{D}(\Omega)$ is dense in $L^2(\Omega)$, it follows $f = 0$. Hence, R_α is injective.

It follows from (2.25) that if $f \neq 0$ then

$$(R_\alpha f, f)_{L^2} = (u,f)_{L^2} = \|\nabla u\|_{L^2}^2 + \alpha \|u\|_{L^2}^2 > 0,$$

because $u \neq 0$ by the injectivity. Hence, R_α is positive definite.

Since R_α is a bounded operator, in order to prove that it is self-adjoint it suffices to prove that it is symmetric, that is

$$(R_\alpha f, g)_{L^2} = (f, R_\alpha g)_{L^2} \quad \text{for all } f, g \in L^2(\Omega).$$

Setting $R_\alpha f = u$, $R_\alpha g = v$, and choosing $\varphi = v$ in (2.23), we obtain

$$(\nabla u, \nabla v)_{\vec{L}^2} + \alpha (u,v)_{L^2} = (f, R_\alpha g)_{L^2}.$$

Since the left hand side is symmetric in u, v, we conclude that the right hand side is
13.12.16 symmetric in f, g, which implies that R_α is symmetric.

(c) Consider an operator $\widetilde{R}_\alpha$ defined by

$$\begin{aligned} \widetilde{R}_\alpha &: L^2(\Omega) \to W_0^1(\Omega) \\ &\widetilde{R}_\alpha f = u \end{aligned}$$

It follows from (2.25) that

$$\|u\|_{W^1}^2 \leq C \|f\|_{L^2} \|u\|_{L^2}$$

where $C = \max(1, \alpha^{-1})$. Since $\|u\|_{L^2} \leq \alpha^{-1} \|f\|_{L^2}$, we obtain that

$$\|u\|_{W^1} \leq C' \|f\|_{L^2},$$

where $C' = (C\alpha^{-1})^{1/2}$, and

$$\left\|\widetilde{R}_\alpha\right\| := \sup_{f \in L^2(\Omega) \setminus \{0\}} \frac{\left\|\widetilde{R}_\alpha f\right\|_{W^1}}{\|f\|_{L^2}} \leq C' < \infty.$$

Therefore, $\widetilde{R}_\alpha$ is a bounded operator. The operator R_α can be represented as the following composition

$$L^2(\Omega) \overset{\widetilde{R}_\alpha}{\to} W_0^1(\Omega) \overset{I}{\hookrightarrow} L^2(\Omega)$$

where I is the identical embedding. Since $\widetilde{R}_\alpha$ is a bounded operator and I is compact by Theorem 2.9, we conclude that $R_\alpha = I \circ \widetilde{R}_\alpha$ is compact. ■

2.6 Eigenvalue problem

Consider in an open set $\Omega \subset M$ the following weak eigenvalue problem:

$$\begin{cases} \Delta v + \lambda v = 0 & \text{weakly in } \Omega \\ v \in W_0^1(\Omega) \setminus \{0\} \end{cases} \tag{2.26}$$

where $\lambda \in \mathbb{R}$ is a spectral parameter. Any solution v to (2.26) is called an eigenfunction of Δ in Ω, and the corresponding value of λ – a (Dirichlet) eigenvalue of Δ in Ω.

If λ is an eigenvalue of Δ in Ω, then consider the *eigenspace*

$$E_\lambda = \left\{v \in W_0^1(\Omega) : \Delta v + \lambda v = 0\right\}.$$

Clearly, E_λ is a subspace of $W_0^1(\Omega)$. The equation $\Delta v + \lambda v = 0$ means that, for any $\varphi \in W_0^1(\Omega)$

$$(\nabla v, \nabla \varphi)_{\vec{L}^2} = \lambda (v, \varphi)_{L^2}. \tag{2.27}$$

The both sides of this equation are continuous functionals of $v \in W_0^1(\Omega)$, which implies that E_λ is a *closed* subspace of $W_0^1(\Omega)$. The *multiplicity* of λ is defined as $\dim E_\lambda$ (finite or ∞).

[1]The Riesz representation theorem says the following: if l is a bounded linear functional on a Hilbert space H, then the equation

$$(u, \varphi)_H = l(\varphi) \quad \forall \varphi \in H$$

has a unique solution $u \in H$. The proof of this theorem amounts to construction of a vector orthogonal to the null space of l.

Theorem 2.11 *Assume that Ω is precompact. There exists an orthonormal basis $\{v_k\}_{k=1}^{\infty}$ in $L^2(\Omega)$ that consists of eigenfunctions of Δ in Ω; the corresponding eigenvalues λ_k are non-negative reals, and the sequence $\{\lambda_k\}$ is monotone increasing and diverges to $+\infty$ as $k \to \infty$.*

Besides, $v_k \in C^{\infty}(\Omega)$ for all k, the sequence $\{v_k\}_{k=1}^{\infty}$ is an orthogonal basis also in $W_0^1(\Omega)$, and the sequence $\{\lambda_k\}_{k=1}^{\infty}$ contains any eigenvalues λ of Δ in Ω exactly m times where m is the multiplicity of λ. In particular, any eigenvalue has a finite multiplicity.

Proof. Any eigenfunction of the Laplace operator is C^{∞} by Corollary 2.8, in particular, $v_k \in C^{\infty}(\Omega)$.

Let v be an eigenfunction of Δ in Ω with the eigenvalue λ. Rewrite the equation $\Delta v + \lambda v = 0$ in the form

$$\Delta v - v = -(1+\lambda)v.$$

By Theorem 2.10, this equation for $v \in W_0^1(\Omega)$ is equivalent to

$$v = R((1+\lambda)v),$$

where $R = R_1$. If $1+\lambda = 0$ then it follows $v = 0$ which contradicts to the definition of an eigenfunction. Therefore, $1+\lambda \neq 0$, which implies

$$Rv = \frac{1}{1+\lambda}v.$$

Hence, if v is an eigenfunction of Δ in Ω with an eigenvalue λ then v is an eigenfunction of the operator R in $L^2(\Omega)$ with the eigenvalue $\frac{1}{1+\lambda}$.

Conversely, if $v \in L^2(\Omega)$ is an eigenfunction of R with an eigenvalue α, that is,

$$Rv = \alpha v,$$

then $\alpha \neq 0$ by the injectivity of R, which implies $v = \frac{1}{\alpha}Rv \in W_0^1(\Omega)$. Hence, v is an eigenfunction of Δ in Ω with the eigenvalue λ that is determined by $\frac{1}{1+\lambda} = \alpha$, that is, $\lambda = \frac{1}{\alpha} - 1$.

Recall the Hilbert-Schmidt theorem: if H is a separable ∞-dimensional Hilbert space and A is a compact self-adjoint operator in H, then there exists an orthonormal basis $\{v_k\}_{k=1}^{\infty}$ in H that consists of the eigenvectors of A, the corresponding eigenvalues α_k are real, and the sequence $\{\alpha_k\}$ goes to 0 as $k \to \infty$.

Since R is a self-adjoint, compact operator in $L^2(\Omega)$, by the Hilbert-Schmidt theorem there is an orthonormal basis $\{v_k\}_{k=1}^{\infty}$ in $L^2(\Omega)$ that consists of the eigenfunctions of R, and if α_k denotes the eigenvalue of v_k then $\alpha_k \to 0$ as $k \to \infty$.

Since R is positive definite, we obtain that $\alpha_k > 0$, because

$$0 < (Rv_k, v_k)_{L^2} = \alpha_k \|v_k\|_{L^2}^2.$$

Any sequence of positive reals that goes to 0 can be rearranged to become monotone decreasing. Hence, by rearranging the sequences $\{v_k\}$ and $\{\alpha_k\}$, we achieve that $\{\alpha_k\}$ is monotone decreasing.

It follows that v_k is an eigenfunction of Δ in Ω with the eigenvalue $\lambda_k = \frac{1}{\alpha_k} - 1$. Clearly, $\{\lambda_k\}$ is monotone increasing and $\lambda_k \to +\infty$ as $k \to \infty$.

Let us show that $\lambda_k \geq 0$. Indeed, if v is an eigenfunction of Δ in Ω with an eigenvalue λ, then it follows from $\Delta v = -\lambda v$ that, for any $\varphi \in W_0^1(\Omega)$,

$$\boxed{(\nabla v, \nabla \varphi)_{\vec{L}^2} = \lambda (v, \varphi)_{L^2}}. \tag{2.28}$$

Substituting $\varphi = v$, we obtain

$$\boxed{\lambda = \frac{\|\nabla v\|_{\vec{L}^2}^2}{\|v\|_{L^2}^2}} \geq 0. \tag{2.29}$$

Let mention for the future the following consequence of (2.29):

$$\boxed{\|v\|_{W^1}^2 = (\lambda + 1) \|v\|_{L^2}^2}. \tag{2.30}$$

Let us verify that the sequence $\{v_k\}$ is orthogonal in $W_0^1(\Omega)$. Setting in (2.28) $v = v_k$ and $\varphi = v_l$, we obtain, for all $k \neq l$,

$$(\nabla v_k, \nabla v_l)_{\vec{L}^2} = \lambda_k (v_k, v_l)_{L^2} = 0,$$

which implies

$$(v_k, v_l)_{W^1} = 0.$$

In order to show that $\{v_k\}_{k=1}^{\infty}$ is a basis in $W_0^1(\Omega)$, it suffices to show that, for any $\varphi \in W_0^1(\Omega)$, if $(v_k, \varphi)_{W^1} = 0$ for any $k \geq 1$, then $\varphi = 0$. Indeed, by (2.28) we have

$$(\nabla v_k, \nabla \varphi)_{\vec{L}^2} = \lambda_k (v_k, \varphi)_{L^2}$$

whence

$$(v_k, \varphi)_{W^1} = (\nabla v_k, \nabla \varphi)_{\vec{L}^2} + (v_k, \varphi)_{L^2} = (\lambda_k + 1)(v_k, \varphi)_{L^2}.$$

Since $(v_k, \varphi)_{W^1} = 0$, it follows that also $(v_k, \varphi)_{L^2} = 0$. By the completeness of $\{v_k\}$ in $L^2(\Omega)$ we conclude that $\varphi = 0$.

Before we prove the remaining claim about the multiplicity of eigenvalues, let us verify that if v and w are two eigenfunctions with distinct eigenvalues λ and μ, then u and w are orthogonal in $L^2(\Omega)$ and $W^1(\Omega)$. Indeed, setting $\varphi = w$ in (2.28), we obtain

$$(\nabla v, \nabla w)_{\vec{L}^2} = \lambda (v, w)_{L^2} \tag{2.31}$$

and in the same way

$$(\nabla v, \nabla w)_{\vec{L}^2} = \mu (v, w)_{L^2}$$

whence

$$(\lambda - \mu)(v, w)_{L^2} = 0.$$

Since $\lambda \neq \mu$, we conclude that $(v, w)_{L^2} = 0$. It follows from (2.31) that also $(u, w)_{W^1} = 0$.

Assume that λ is an eigenvalue of Δ with multiplicity m, that is, $\dim E_\lambda = m$. In the next argument we regard E_λ as a subspace of $W_0^1(\Omega)$ and use only W^1 inner product. Assume that λ occurs l times in the sequence $\{\lambda_k\}_{k=1}^{\infty}$, say, at (necessarily consecutive) positions $i+1, ..., i+l$. Since $\lambda_k \to \infty$, we have $l < \infty$. The functions $v_{i+1}, ..., v_{i+l}$ belong to E_λ and are linearly independent, which implies $l \leq m$. Let us show that $l = m$. Assume from the contrary that $l < m$. Then there is a non-zero element $w \in E_\lambda$ that is orthogonal to span$\{v_{i+1}, ..., v_{i+l}\}$. We claim that w is orthogonal to all v_k. Indeed, w is orthogonal to $v_{i+1}, ..., v_{i+l}$ by construction, and w is orthogonal to all other v_k because their eigenvalues are different from λ. However, a non-zero element of $W_0^1(\Omega)$ cannot be orthogonal to all v_k because $\{v_k\}_{k=1}^\infty$ is a basis in $W_0^1(\Omega)$. ■

In what follows we denote by $\{\lambda_k(\Omega)\}_{k=1}^\infty$ the sequence of the eigenvalues of Δ in Ω in the (non-strictly) increasing order, that is, $\lambda_k(\Omega) \le \lambda_{k+1}(\Omega)$.

The sequence $\{\lambda_k(\Omega)\}_{k=1}^\infty$ contains certain information about the domain Ω (and about the metric tensor $\mathbf{g}$ in Ω). There was a famous question of Mark Kac stated in 1966 as follows:

"*Can one hear the shape of a drum*?"

The point is that if we consider Ω as a drum then the frequencies of vibration of the drum when hit are exactly $\sqrt{\lambda_k(\Omega)}$ (provided the metric tensor $\mathbf{g}$ is properly chosen depending on the material of the drum). Therefore, hearing the overtones of the drum allows (at least theoretically) to recover the sequence $\{\lambda_k(\Omega)\}$, and the main question is whether this sequence contains enough information to restore Ω and $\mathbf{g}$, up to isometry. In general the answer is negative, but constructing counterexamples is quite difficult.

2.7 The bottom eigenvalue

As before, let Ω be a precompact open subset of M. The value $\lambda_1(\Omega)$ is called the *bottom eigenvalue* of Ω.

Theorem 2.12 *Let $(M, \mathbf{g}, \mu)$ be a connected weighted manifold. If $\Omega \subset M$ is a non-empty relatively compact open set such that $M \setminus \overline{\Omega}$ is non-empty then $\lambda_1(\Omega) > 0$.*

In general $\lambda_1(\Omega) = 0$ is possible, for example, if M is a compact manifold (say, $\mathbb{S}^n$) and $\Omega = M$. Indeed, in this case $v \equiv 1 \in \mathcal{D}(\Omega)$ is an eigenfunction of Ω with the eigenvalue $\lambda = 0$ so that $\lambda_1(\Omega) = 0$. This example shows also that the condition that $M \setminus \overline{\Omega}$ is non-empty is essential for the positivity of $\lambda_1(\Omega)$.

The assumption about the connectedness of M is also essential. Indeed, let M consist of two disjoint copies of $\mathbb{S}^n$, so that M is disconnected Let Ω be one of the copies of $\mathbb{S}^n$. Then $M \setminus \overline{\Omega}$ is non-empty but still $\lambda_1(\Omega) = 0$ because again $\varphi \equiv 1$ is an eigenfunction of Ω with the eigenvalue $\lambda = 0$.

Recall that if Ω is a bounded domain in $\mathbb{R}^n$ then $\lambda_1(\Omega) > 0$ can be proved by using Friedrich's inequality. On a general manifold this tool is not available, so we have to
15.12.16 use a different argument.

Proof. Assume that $\lambda_1(\Omega) = 0$ so that there is an eigenfunction v of Δ in Ω with the eigenvalue 0, that is, $v \in W_0^1(\Omega)$ and $\Delta v = 0$ weakly in Ω. By Corollary 2.8 we have $v \in C^\infty(\Omega)$. We will prove that $v = 0$ in Ω which will contradict to the fact that v is an eigenfunction. It suffices to prove that $v = 0$ in any connected component. Hence, we can assume without loss of generality, that Ω is connected.

By (2.29) we have $\|\nabla v\|_{\vec{L}^2} = 0$ that is, $\nabla v = 0$ in Ω. Since Ω is connected, we conclude that $v \equiv \text{const}$ in Ω. If $v \neq 0$ in Ω then we can assume without loss of generality, that $v \equiv 1$ in Ω.

The set $\overline{\Omega}$ is closed and its complement is non-empty by hypothesis. The sets $\overline{\Omega}$ and $\overline{M \setminus \overline{\Omega}}$ are closed and their union is M. Since M is connected, these sets cannot be disjoint. Hence, there is a point x_0 that belongs to both $\overline{\Omega}$ and $\overline{M \setminus \overline{\Omega}}$.

Let U be any connected open neighborhood of x_0. Note that, by the choice of x_0, the set U intersects both Ω and $M \setminus \overline{\Omega}$. Consider the set $\Omega' = \Omega \cup U$ that is a connected open set. Note that, by construction, $\Omega' \setminus \overline{\Omega}$ is non-empty.

Since $v \in W_0^1(\Omega)$, extending v to Ω' by setting $v = 0$ in $\Omega' \setminus \Omega$, we obtain that $v \in W_0^1(\Omega')$. Since $v = 0$ on $\Omega' \setminus \Omega$, we have also $\nabla v = 0$ in $\Omega' \setminus \Omega$ a.e. (EDE, Lemma 1.5). Since also $\nabla v = 0$ in Ω, we conclude that $\nabla v = 0$ in Ω'. This implies that

$$\langle \nabla v, \nabla \varphi \rangle = 0 \ \forall \varphi \in \mathcal{D}(\Omega),$$

that is, $\Delta v = 0$ weakly in Ω'. It follows that $v \in C^\infty(\Omega')$. Since $\nabla v = 0$ in Ω', we conclude that $v \equiv \text{const}$ in Ω', which contradicts to the facts that $v = 1$ in Ω and $v = 0$ in $\Omega' \setminus \overline{\Omega}$. $\blacksquare$

Chapter 3

The heat semigroup in compact domains

As before, $(M, \mathbf{g}, \mu)$ is a weighted manifold and Δ is the weighted Laplace operator on M.

3.1 The heat equation and caloric functions

Let I be an interval in $\mathbb{R}$. Consider in $I \times M$ the heat *equation*

$$\frac{\partial u}{\partial t} = \Delta u,$$

where $u = u(t, x)$ is a function of $t \in I$ and $x \in M$. This equation can be understood in the classical sense: the function $u(t, x)$ is differentiable in t, is C^2 in x, and, for all $(t, x) \in I \times M$, $\frac{\partial u}{\partial t}(t, x) = \Delta u(t, x)$.

However, we will understand the heat equation in a weak sense, and the solution u will be regarded as a path in $L^2(M)$.

Definition. For a function $u : I \to L^2(M)$, define its L^2-derivative $u'(t) \in L^2(M)$ at $t \in I$ by

$$u'(t) = \lim_{s \to 0} \frac{u(t+s) - u(t)}{s},$$

where the limit is understood in the norm of $L^2(M)$, that is,

$$\left\| \frac{u(t+s) - u(t)}{s} - u'(t) \right\|_{L^2} \to 0 \text{ as } s \to 0.$$

Notation for the L^2-derivative: $u'(t)$ or $\frac{du}{dt}$.

Notation for function u: for any $t \in I$, $u(t)$ is an element of $L^2(M)$, so that $u(t)(x)$ makes sense. For simplicity, we use instead the notation $u(t, x)$. Then $u(t, \cdot)$ has the same meaning as $u(t)$.

Definition. A function $u : I \to L^2(M)$ is called *caloric* in $I \times M$ if

1. u is L^2-differentiable at any $t \in I$;

2. for any $t \in I$, we have $u(t) \in W^1(M)$ and $\Delta u(t) \in L^2(M)$, where Δ is understood in the weak sense;
3. for any $t \in I$, we have $\frac{du}{dt} = \Delta u(t)$.

In this case we also say that the heat equation $\frac{du}{dt} = \Delta u$ is satisfied weakly in $I \times M$.

Example. Assume that $v \in W^1(M)$ satisfies weakly in M the equation

$$\Delta v + \lambda v = 0.$$

Then the function $u(t,x) = e^{-\lambda t} v(x)$ is caloric in $\mathbb{R} \times M$. Indeed, u can be regarded as a mapping

$$\begin{aligned} u \;\; &: \;\; \mathbb{R} \to L^2(M) \\ & \quad\; u(t) = e^{-\lambda t} v \end{aligned}$$

Since v does not depend in t, we obtain

$$u'(t) = -\lambda e^{-\lambda t} v.$$

On the other hand,

$$\Delta u = e^{-\lambda t} \Delta v = -\lambda e^{-\lambda t} v,$$

whence $u' = \Delta u$ follows.

3.2 The mixed problem

Let Ω be an open subset of a weighted manifold M. Since Ω can be regarded as a manifold, the above notion of a caloric function is defined in $I \times \Omega$ for any interval $I \subset \mathbb{R}$.

Consider the following initial-boundary problem (shortly, *mixed problem*) in $\mathbb{R}_+ \times \Omega$:

$$\begin{cases} \frac{du}{dt} = \Delta u & \text{weakly in } \mathbb{R}_+ \times \Omega, \\ u(t,\cdot) \in W_0^1(\Omega) & \text{for any } t > 0, \\ u(t,\cdot) \xrightarrow{L^2} f & \text{as } t \to 0+, \end{cases} \tag{3.1}$$

where $f \in L^2(\Omega)$ is a given function. In other words, we look for a caloric function in $\mathbb{R}_+ \times \Omega$ that satisfies the appropriately understood boundary condition $u = 0$ on $\partial\Omega$ and the initial condition $u|_{t=0} = f$.

Theorem 3.1 *The mixed problem (3.1) has at most one solution.*

Proof. Assuming that u solves the mixed problem, consider the function

$$J(t) := \|u(t,\cdot)\|_{L^2}^2 = (u(t), u(t))$$

and prove that it is monotone decreasing in $t \in (0, +\infty)$. For that, we use the following product rule for L^2-derivatives: if $u(t)$ and $v(t)$ are L^2-differentiable functions then the numerical function $t \mapsto (u(t), v(t))$ is differentiable and

$$\frac{d}{dt}(u, v) = (\frac{d}{dt}u, v) + (u, \frac{d}{dt}v),$$

which is proved in the same way, as the usual product rule for scalar functions (see Exercise 65). In particular, we obtain that the function $J(t)$ is differentiable on $(0, +\infty)$ and

$$J'(t) = \frac{d}{dt}(u, u) = 2(\frac{du}{dt}, u) = 2(\Delta u, u).$$

By the definition of Δu, we have, for any $\varphi \in W_0^1(\Omega)$,

$$(\Delta u, \varphi) = -\langle \nabla u, \nabla \varphi \rangle.$$

Since $u \in W_0^1(\Omega)$, setting here $\varphi = u$ we obtain

$$(\Delta u, u) = -\langle \nabla u, \nabla u \rangle \leq 0,$$

whence $J'(t) \leq 0$ follows. Hence, $J(t)$ is a monotone decreasing function.

To prove the uniqueness of the solution is suffices to show that $f = 0$ implies $u = 0$. Indeed, if $u(t) \xrightarrow{L^2} 0$ as $t \to 0+$ then also $J(t) \to 0$. Since $J(t)$ is non-negative and decreasing, we conclude $J(t) \equiv 0$ for $t > 0$ and $u(t) = 0$, which was to be proved. ∎

Now we prove the existence of solution of (3.1) in precompact domains using the method of separation of variables.

Theorem 3.2 *Let Ω be precompact. Let $\{v_k\}_{k=1}^{\infty}$ be an orthonormal basis in $L^2(\Omega)$ that consists of eigenfunctions of Δ in Ω, and let $\{\lambda_k\}_{k=1}^{\infty}$ be the sequence of the corresponding eigenvalues, in the increasing order. For any $f \in L^2(\Omega)$, consider the eigenfunction expansion*

$$f = \sum_{k=1}^{\infty} a_k v_k$$

and, for any $t \geq 0$, set

$$u(t) = \sum_{k=1}^{\infty} e^{-\lambda_k t} a_k v_k.$$

Then $u(t)$ solves the mixed problem (3.1).

Let us first prove two lemmas.

Lemma 3.3 *Let $\{a_k\}_{k=1}^{\infty}$ be a sequence of reals.*

(a) If

$$\sum_{k=1}^{\infty} \lambda_k a_k^2 < \infty, \tag{3.2}$$

then the series $\sum a_k v_k$ converges in $W^1(\Omega)$ and, hence,

$$f := \sum_{k=1}^{\infty} a_k v_k \in W_0^1(\Omega)$$

and

$$\|f\|_{W^1}^2 = \sum_{k=1}^{\infty} (\lambda_k + 1)\, a_k^2. \tag{3.3}$$

(*b*) *If*

$$\sum_{k=1}^{\infty} \lambda_k^2 a_k^2 < \infty \tag{3.4}$$

then

$$\Delta f = -\sum_{k=1}^{\infty} \lambda_k a_k v_k \in L^2(\Omega) \tag{3.5}$$

By Exercise 42, the condition (3.2) is also necessary for $f \in W_0^1(\Omega)$, and (3.4) is also necessary for $\Delta f \in L^2(\Omega)$.

Proof. (*a*) Since the sequence $\{v_k\}$ is orthonormal in $L^2(\Omega)$, the series

$$\sum_{k=1}^{\infty} a_k v_k$$

converges in L^2 if and only if

$$\sum_{k=1}^{\infty} \|a_k v_k\|_{L^2}^2 = \sum_{k=1}^{\infty} a_k^2 < \infty.$$

The sequence $\{v_k\}$ is orthogonal also in $W_0^1(\Omega)$ and, by (2.30),

$$\|v_k\|_{W^1} = \sqrt{\lambda_k + 1}.$$

Hence, by (3.2)

$$\sum_{k=1}^{\infty} \|a_k v_k\|_{W^1}^2 = \sum_{k=1}^{\infty} a_k^2 (\lambda_k + 1) < \infty,$$

which implies that the series $\sum a_k v_k$ converges in W^1 and, hence, $f \in W_0^1(\Omega)$. Then (3.3) holds by the Parseval identity.

(*b*) Since (3.4) implies (3.2), we have $f \in W_0^1(\Omega)$. Consider the partial sums

$$f_N = \sum_{k=1}^{N} a_k v_k \quad \text{and} \quad g_N = \sum_{k=1}^{N} \lambda_k a_k v_k.$$

We have

$$\Delta f_N = \sum_{k=1}^{N} a_k \Delta v_k = -\sum_{k=1}^{N} \lambda_k a_k v_k = -g_N.$$

where we have used $\Delta v_k = -\lambda_k v_k$. Hence, for any $\varphi \in W_0^1(\Omega)$, we have

$$(\nabla f_N, \nabla \varphi)_{\vec{L}^2} = (g_N, \varphi)_{L^2}. \tag{3.6}$$

Letting $N \to \infty$ and using that $f_N \overset{W^1}{\to} f$ and

$$g_N \overset{L^2}{\to} g := \sum_{k=1}^{\infty} \lambda_k a_k v_k,$$

we obtain

$$(\nabla f, \nabla \varphi)_{\vec{L}^2} = (g, \varphi)_{L^2},$$

that is, $\Delta f = -g$, which was to be proved. ■

20.12.16

Corollary 3.4 *Let $f = \sum a_k v_k \in L^2(\Omega)$. If, for some non-negative integer j,*

$$\sum_{k=1}^{\infty} \lambda_k^{2j+1} a_k^2 < \infty \tag{3.7}$$

then

$$\Delta^j f = (-1)^j \sum_{k=1}^{\infty} \lambda_k^j a_k v_k \in W_0^1(\Omega),$$

where the series converges in $W^1(\Omega)$.

Proof. The case $j = 0$ is equivalent to Lemma 3.3(a).

Inductive step from j to $j+1$. By the inductive hypothesis, we have

$$g := \Delta^j f = (-1)^j \sum_{k=1}^{\infty} \lambda_k^j a_k v_k = \sum_{k=1}^{\infty} b_k v_k \in W_0^1(\Omega),$$

where

$$b_k = (-1)^j \lambda_k^j a_k.$$

Since

$$\sum_{k=1}^{\infty} \lambda_k^2 b_k^2 = \sum \lambda_k^{2j+2} a_k^2 < \infty,$$

we obtain by Lemma 3.3(b) that

$$\Delta g = -\sum_{k=1}^{\infty} \lambda_k b_k v_k \in L^2(\Omega).$$

Moreover, since

$$\sum_{k=1}^{\infty} \lambda_k (\lambda_k b_k)^2 = \sum_{k=1}^{\infty} \lambda_k^{2j+3} a_k^2 < \infty,$$

we obtain by Lemma 3.3(a) that $\Delta g \in W_0^1(\Omega)$. It remains to observe that

$$\Delta^{j+1} f = \Delta g = (-1)^{j+1} \sum_{k=1}^{\infty} \lambda_k^{j+1} a_k v_k,$$

which finishes the inductive step. ■

Lemma 3.5 (Dominated convergence theorem) *Consider a sequence of functions* $\{\gamma_k(t)\}_{k=1}^{\infty}$ *defined on some interval* I *containing* 0. *Assume that all* $\gamma_k(t)$ *are continuous at* $t=0$ *and that the sequence* $\{\gamma_k\}$ *is uniformly bounded on* I, *that is,*

$$C := \sup_{k\in\mathbb{N}} \sup_{t\in I} \gamma_k(t) < \infty.$$

Let $\sum_{k=1}^{\infty} h_k$ *be a convergent orthogonal series in a Hilbert space* H. *Then*

$$\sum_{k=1}^{\infty} \gamma_k(t)\, h_k \to \sum_{k=1}^{\infty} \gamma_k(0)\, h_k \quad \text{as } t\to 0,$$

where the converges is in the norm of H.

We will apply this lemma for $H=L^2$ and for $H=W_0^1$.

Proof. The convergence of $\sum h_k$ is equivalent to

$$\sum_{k=1}^{\infty} \|h_k\|^2 < \infty.$$

Since all functions $\gamma_k(t)$ are uniformly bounded, we obtain that

$$\sum_{k=1}^{\infty} \gamma_k(t)^2 \|h_k\|^2 < \infty,$$

which implies that the series

$$w(t) := \sum_{k=1}^{\infty} \gamma_k(t)\, h_k$$

converges for any $t\in I$. We need to prove that $w(t)\to w(0)$ as $t\to 0$. We have

$$w(t)-w(0) = \sum_{k=1}^{\infty} (\gamma_k(t)-\gamma_k(0))\, h_k,$$

whence by the Parseval identity

$$\|w(t)-w(0)\|_{L^2}^2 = \sum_{k=1}^{\infty} (\gamma_k(t)-\gamma_k(0))^2 \|h_k\|^2.$$

To prove that this goes to 0 as $t\to 0$, let us fix some $\varepsilon>0$ and choose N so big that

$$\sum_{k=N}^{\infty} \|h_k\|^2 < \varepsilon.$$

Then

$$\begin{aligned}\|w(t)-w(0)\|_{L^2}^2 &= \sum_{k=1}^{N} (\gamma_k(t)-\gamma_k(0))^2 \|h_k\|^2 \\ &\quad + \sum_{k=N}^{\infty} (\gamma_k(t)-\gamma_k(0))^2 \|h_k\|^2.\end{aligned}$$

The first (finite) sum goes to 0 as $t \to 0$ by the continuity of all γ_k at 0. The second sum is bounded by

$$\sum_{k=N}^{\infty} (2C)^2 \|h_k\|^2 = 4C^2 \sum_{k=N}^{\infty} \|h_k\|^2 \leq 4C^2 \varepsilon.$$

It follows that

$$\limsup_{t\to 0} \|w(t) - w(0)\|^2 \leq 4C^2 \varepsilon.$$

Since ε is arbitrary, we obtain $w(t) \xrightarrow{H} w(0)$. ■

Proof of Theorem 3.2. Fix $t > 0$. By Lemma 3.3(a), in order to prove that $u(t) \in W_0^1(\Omega)$, it suffices to verify that

$$\sum_{k=1}^{\infty} \lambda_k e^{-2\lambda_k t} a_k^2 < \infty,$$

and the latter is true because

$$\sup_{\lambda \geq 0} \lambda e^{-2\lambda t} = \frac{1}{t} \sup_{\lambda \geq 0} (\lambda t) e^{-2\lambda t} = \frac{1}{t} \sup_{\xi \geq 0} \xi e^{-2\xi} < \infty \tag{3.8}$$

and

$$\sum_{k=1}^{\infty} a_k^2 < \infty.$$

Let us show that $\Delta u(t) \in L^2(\Omega)$ for all $t > 0$. By Lemma 3.3(b), it suffices to verify that

$$\sum_{k=1}^{\infty} \lambda_k^2 e^{-2\lambda_k t} a_k^2 < \infty,$$

and this is true because similarly to (3.8)

$$\sup_{\lambda \geq 0} \lambda^2 e^{-2\lambda t} < \infty.$$

Besides, we obtain by (3.5) that

$$\Delta u(t) = -\sum_{k=1}^{\infty} \lambda_k e^{-\lambda_k t} a_k v_k.$$

Let us show that $u(t) \xrightarrow{L^2} f$ as $t \to 0$. Indeed, since

$$e^{-\lambda_k t} \to 1 \text{ as } t \to 0$$

and all functions $e^{-\lambda_k t}$ are bounded by 1 for all k and $t \geq 0$, we conclude by Lemma 3.5

$$u(t) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k t} v_k \xrightarrow{L^2} \sum_{k=1}^{\infty} a_k v_k = f.$$

Let us compute $\frac{du}{dt}$ at any $t > 0$. Observe that

$$\begin{aligned} \frac{u(t+s) - u(t)}{s} &= \sum_{k=1}^{\infty} \frac{e^{-\lambda_k(t+s)} - e^{-\lambda_k t}}{s} a_k v_k \\ &= \sum_{k=1}^{\infty} \frac{e^{-s\lambda_k} - 1}{s} e^{-\lambda_k t} a_k v_k. \end{aligned} \tag{3.9}$$

Fix $t > 0$ and consider the functions

$$\gamma_k(s) = \frac{e^{-s\lambda_k} - 1}{s} e^{-\lambda_k t}.$$

Clearly, we have as $s \to 0$

$$\gamma_k(s) \to -\lambda_k e^{-\lambda_k t} =: \gamma_k(0).$$

In order to be able to apply Lemma 3.5, we need to verify that the functions $\gamma_k(s)$ are uniformly bounded for all k and for all s near 0. This is equivalent to the following: there is $\varepsilon > 0$ such that

$$\sup_{\lambda \geq 0} \sup_{s \in [-\varepsilon, \varepsilon]} \frac{e^{-s\lambda} - 1}{s} e^{-\lambda t} < \infty.$$

In fact, we will take $\varepsilon = t/2$. Let us apply the inequality

$$\left| e^{\theta} - 1 \right| \leq |\theta| \, e^{|\theta|},$$

for any $\theta \in \mathbb{R}$, which follows from

$$\left| e^{\theta} - 1 \right| = \left| \int_0^{\theta} e^{\xi} d\xi \right| \leq |\theta| \, e^{|\theta|}.$$

Setting here $\theta = -\lambda s$, we obtain

$$\left| e^{-s\lambda} - 1 \right| \leq \lambda \, |s| \, e^{\lambda |s|},$$

whence, for all $s \in [-t/2, t/2]$,

$$\left| \frac{e^{-\lambda s} - 1}{s} e^{-t\lambda} \right| \leq \lambda e^{-\lambda t} e^{\lambda |s|} \leq \lambda e^{-\lambda t/2}.$$

Therefore, we have

$$\sup_{\lambda \geq 0} \sup_{s \in [-t/2, t/2]} \left| \frac{e^{-\lambda s} - 1}{s} e^{-t\lambda} \right| \leq \sup_{\lambda \geq 0} \lambda e^{-\lambda t/2} < \infty. \tag{3.10}$$

Returning to (3.9), we obtain

$$\frac{u(t+s) - u(t)}{s} \xrightarrow{L^2} -\sum_{k=1}^{\infty} \lambda_k e^{-\lambda_k t} a_k v_k,$$

whence

$$\frac{du}{dt} = -\sum_{k=1}^{\infty} \lambda_k e^{-\lambda_k t} a_k v_k = \Delta u(t),$$

which finishes the proof. ∎

Define for any $t \geq 0$ the operator $P_t^\Omega : L^2(\Omega) \to L^2(\Omega)$ as follows: if $f = \sum_{k=1}^{\infty} a_k v_k \in L^2(\Omega)$, then

$$P_t^\Omega f = \sum_{k=1}^{\infty} e^{-\lambda_k t} a_k v_k. \tag{3.11}$$

Theorem 3.6 *The operators P_t^Ω have the following properties:*
(*a*) $\left\| P_t^\Omega \right\| \leq 1$;
(*b*) $P_t^\Omega \to \mathrm{id}$ *as $t \to 0+$ in the strong operator topology;*
(*c*) $P_t^\Omega P_s^\Omega = P_{t+s}^\Omega$ *(the semigroup identity);*
(*d*) P_t^Ω *is self-adjoint.*

One says that the family $\left\{ P_t^\Omega \right\}$ is a *continuous contraction semigroup* in $L^2(\Omega)$. 10.01.17

Proof. (*a*) For any $f = \sum_{k=1}^{\infty} a_k v_k \in L^2(\Omega)$, we have

$$\|f\|_{L^2}^2 = \sum_{k=1}^{\infty} a_k^2$$

and

$$\left\| P_t^\Omega f \right\|_{L^2}^2 = \sum_{k=1}^{\infty} e^{-2\lambda_k t} a_k^2 \leq \sum_{k=1}^{\infty} a_k^2 = \|f\|_{L^2}^2,$$

whence $\left\| P_t^\Omega \right\| \leq 1$.

(*b*) We already know that, for any $f \in L^2(\Omega)$, $P_t^\Omega f \xrightarrow{L^2} f$ which exactly means that $P_t^\Omega \to \mathrm{id}$ in the strong operator topology (but not in the operator norm).

(*c*) We have, for any $f \in L^2(\Omega)$ as above,

$$\begin{aligned} P_t^\Omega P_s^\Omega f &= P_t^\Omega \left(\sum_{k=1}^{\infty} e^{-\lambda_k s} a_k v_k \right) \\ &= \sum_{k=1}^{\infty} e^{-\lambda_k t} e^{-\lambda_k s} a_k v_k \\ &= \sum_{k=1}^{\infty} e^{-\lambda_k (t+s)} a_k v_k \\ &= P_{t+s}^\Omega f, \end{aligned}$$

which proves the claim.

(*d*) For $f = \sum_{k=1}^{\infty} a_k v_k \in L^2(\Omega)$ and $g = \sum_{k=1}^{\infty} b_k v_k \in L^2(\Omega)$, we have

$$(P_t f, g)_{L^2} = \sum_{k=1}^{\infty} \left(e^{-\lambda_k t} a_k \right) b_k = \sum_{k=1}^{\infty} a_k \left(e^{-\lambda_k t} b_k \right) = \left(f, P_t^\Omega g \right)_{L^2},$$

which means that P_t^Ω is self-adjoint. ∎

3.3 Smoothness properties

Theorem 3.7 *Let Ω be a precompact open subset of M. For any $f \in L^2(\Omega)$ and $t > 0$, the function $P_t^\Omega f$ belongs to $C^\infty(\Omega)$. Moreover, for any compact set $K \subset \Omega$ and any $t > 0$,*

$$\left\| P_t^\Omega f \right\|_{C(K)} \le C \left(1 + t^{-1}\right)^{\frac{n}{4}+1} \|f\|_{L^2}, \tag{3.12}$$

where $C = C(\Omega, K, \mathbf{g}, D, n)$.

Proof. Let $f = \sum_{k=1}^\infty a_k v_k$, so that

$$u(t) := P_t^\Omega f = \sum_{k=1}^\infty e^{-\lambda_k t} a_k v_k.$$

By Corollary 3.4 we obtain that, for any non-negative j,

$$\Delta^j u(t) = (-1)^j \sum_{k=1}^\infty \lambda_k^j e^{-\lambda_k t} a_k v_k \in W_0^1(\Omega), \tag{3.13}$$

because

$$\sum_{k=1}^\infty \lambda_k^{2j+1} \left(e^{-\lambda_k t} a_k\right)^2 < \infty, \tag{3.14}$$

and the latter is true because $\sum a_k^2 < \infty$ and, for any $q \ge 0$,

$$\sup_{\lambda \ge 0} \lambda^q e^{-2\lambda t} = \sup_{\lambda \ge 0} t^{-q} (\lambda t)^q e^{-2\lambda t} = t^{-q} \sup_{\xi \ge 0} \xi^q e^{-2\xi} = \frac{C_q}{t^q} < \infty. \tag{3.15}$$

By Theorem 2.7, we conclude that $u(t) \in C^\infty(\Omega)$ for any $t > 0$.

Let us prove the estimate (3.12). By Lemma 3.3 we have

$$\left\| \Delta^j u \right\|_{W^1}^2 = \sum_{k=1}^\infty (\lambda_k + 1) \left(\lambda_k^j e^{-\lambda_k t} a_k\right)^2.$$

Since $\sum a_k^2 = \|f\|_{L^2}^2$ and by (3.15)

$$\sup_{\lambda \ge 0} (\lambda + 1) \left(\lambda^j e^{-\lambda t}\right)^2 \le \sup_{\lambda \ge 0} \lambda^{2j+1} e^{-2\lambda t} + \sup_{\lambda \ge 0} \lambda^{2j} e^{-2\lambda t} \le \frac{C_{2j+1}}{t^{2j+1}} + \frac{C_{2j}}{t^{2j}} \le \frac{C_j'\left(1 + t^{-1}\right)}{t^{2j}},$$

we obtain

$$\left\| \Delta^j u \right\|_{W^1}^2 \le \frac{C_j'\left(1 + t^{-1}\right)}{t^{2j}} \|f\|_{L^2}^2.$$

By Theorem 2.7, we have the following estimate

$$\|u\|_{C(K)} \le C \sum_{j=0}^k \left\| \Delta^j u \right\|_{W^1(\Omega)},$$

provided $2k+1>\frac{n}{2}$. Since

$$\begin{aligned}\sum_{j=0}^{k}\left\|\Delta^j u\right\|_{W^1(\Omega)} &\leq C\sum_{j=0}^{k}\frac{\left(1+t^{-1}\right)^{1/2}}{t^j}\|f\|_{L^2}\\ &\leq C\left(1+t^{-1}\right)^{1/2}\left(1+t^{-1}\right)^{k}\|f\|_{L^2}\\ &= C\left(1+t^{-1}\right)^{k+1/2}\|f\|_{L^2},\end{aligned}$$

it follows

$$\|u\|_{C(K)}\leq C\left(1+t^{-1}\right)^{k+1/2}\|f\|_{L^2}.$$

Choosing minimal k with $2k+1>n/2$, that is, $2k-1\leq n/2$ and, hence,

$$k+\frac{1}{2}\leq\frac{n}{4}+1,$$

whence (3.12) follows. ■

Theorem 3.8 *Under the conditions of Theorem 3.7, the function $u(t,x)=P_t^{\Omega}f(x)$ belongs to $C^\infty(\mathbb{R}_+\times\Omega)$, that is, $u(t,x)$ is smooth jointly in (t,x). Moreover, u satisfies in $\mathbb{R}_+\times\Omega$ the heat equation $\partial_t u=\Delta u$ in the classical sense.*

Proof. By Theorem 3.7, the function $P_t^{\Omega}f$ is smooth for any $t>0$, so that $u(t,x)=P_tf(x)$ is defined pointwise for all $t>0$ and $x\in\Omega$. Let us prove first that u is jointly continuous in (t,x). For that it suffices to show that $u(t,x)$ is continuous in t locally uniformly in x. In fact, we will prove that, for any $t>0$, for any positive integer m and for any compact set $K\subset\Omega$ that is covered by a chart U,

$$u(t+s,\cdot)\overset{C^m(K)}{\longrightarrow}u(t,\cdot)\quad\text{as } s\to 0, \tag{3.16}$$

which will settle the joint continuity of u.

By Theorem 2.7, we have

$$\|v\|_{C^m(K)}\leq C\sum_{j=0}^{k}\left\|\Delta^j v\right\|_{W^1(\Omega)}$$

provided $2k+1>m+\frac{n}{2}$ and $\Delta^j v\in W^1(\Omega)$ for all $j\leq k$. Hence, in order to prove (3.16) it suffices to show that, for any non-negative integer j and for any $t>0$,

$$\Delta^j u(t+s,\cdot)\overset{W^1(\Omega)}{\longrightarrow}\Delta^j u(t,\cdot)\ \text{ as } s\to 0. \tag{3.17}$$

As in the proof of Theorem 3.7, we have, for any non-negative integer j,

$$\Delta^j u(t,\cdot)=(-1)^j\sum_{k=1}^{\infty}\lambda_k^j e^{-\lambda_k t}a_k v_k\in W_0^1(\Omega),$$

whence

$$\begin{aligned}\Delta^j u(t+s,\cdot) &= (-1)^j\sum_{k=1}^{\infty}\lambda_k^j e^{-\lambda_k(t+s)}a_k v_k\\ &= (-1)^j\sum_{k=1}^{\infty}e^{-\lambda_k(s+t/2)}\lambda_k^j e^{-\lambda_k t/2}a_k v_k. \end{aligned}\tag{3.18}$$

By Lemma 3.3(a) and (3.14), the series

$$\sum_{k=1}^{\infty}\lambda_k^j e^{-\lambda_k t/2} a_k v_k$$

converges in $W_0^1(\Omega)$. Since the functions

$$\gamma_k(s) = e^{-\lambda_k(s+t/2)}$$

are uniformly bounded for all k and $s\in[-t/2,t/2]$, by Lemma 3.5 we can pass to the limit under the summation sign in (3.18) as $s\to 0$, which yields (3.17) and, hence, (3.16).

By (3.16), for any partial derivative in x-variables,

$$\partial^\alpha = \frac{\partial^{|\alpha|}}{(\partial x^1)^{\alpha_1}(\partial x^2)^{\alpha_2}\dots(\partial x^n)^{\alpha_n}},$$

we obtain

$$\partial^\alpha u(t+s,\cdot)\overset{C(K)}{\longrightarrow}\partial^\alpha u(t,\cdot)\quad\text{as } s\to 0,$$

which implies that $\partial^\alpha u$ is jointly continuous in (t,x).

Next, let us prove the existence of the partial derivative $\partial_t u$. For that let us verify that, for any $t>0$,

$$\frac{u(t+s,\cdot)-u(t,\cdot)}{s}\overset{C^m(K)}{\longrightarrow}\Delta u(t,\cdot)\quad\text{as } s\to 0 \tag{3.19}$$

(note for comparison, that in the proof of Theorem 3.2, we proved a similar convergence in the sense of $L^2(\Omega)$). By Theorem 2.7 it suffices to prove that, for any non-negative integer j,

$$\Delta^j\frac{u(t+s,\cdot)-u(t,\cdot)}{s}\overset{W^1(\Omega)}{\longrightarrow}\Delta^{j+1}u(t,\cdot)\quad\text{as } s\to 0. \tag{3.20}$$

By (3.18), we have

$$\begin{aligned}\Delta^j\frac{u(t+s,\cdot)-u(t,\cdot)}{s} &= (-1)^j\sum_{k=1}^{\infty}\lambda_k^j\frac{e^{-\lambda_k(t+s)}-e^{-\lambda_k t}}{s}a_k v_k\\ &= (-1)^{j+1}\sum_{k=1}^{\infty}\frac{1-e^{-\lambda_k s}}{s}e^{-\lambda_k t/2}\lambda_k^j e^{-\lambda_k t/2}a_k v_k.\end{aligned}$$

Observe that the series

$$\sum_{k=1}^{\infty}\lambda_k^j e^{-\lambda_k t/2} a_k v_k$$

converges in $W_0^1(\Omega)$ and the functions

$$\gamma_k(s) = \frac{1-e^{-\lambda_k s}}{s}e^{-\lambda_k t/2}$$

are uniformly bounded in k and $s\in[-t/4,t/4]$ (cf. the estimate (3.10) from the proof of Theorem 3.2). Since

$$\gamma_k(s)\to\lambda_k e^{-\lambda_k t/2}\quad\text{as } s\to 0,$$

we conclude by Lemma 3.5 that

$$\Delta^j \frac{u(t+s,\cdot)-u(t,\cdot)}{s} \overset{W^1}{\longrightarrow} (-1)^{j+1}\sum_{k=1}^{\infty}\lambda_k e^{-\lambda_k t/2}\lambda_k^j e^{-\lambda_k t/2} a_k v_k$$
$$= (-1)^{j+1}\sum_{k=1}^{\infty}\lambda_k^{j+1} e^{-\lambda_k t} a_k v_k = \Delta^{j+1} u(t,\cdot),$$

which proves (3.20) and, hence, (3.19). 12.01.17

It follows from (3.19) that $\partial_t u$ exists in the classical sense for all $t>0$ and $x \in M$, and

$$\partial_t u = \Delta u. \tag{3.21}$$

In particular, $\partial_t u$ is continuous in (t,x). Moreover, (3.19) also yields that, for any partial derivative ∂^α in x,

$$\frac{\partial^\alpha u(t+s,\cdot)-\partial^\alpha u(t,\cdot)}{s} \overset{C(K)}{\longrightarrow} \partial^\alpha \Delta u(t,\cdot),$$

which implies that $\partial_t(\partial^\alpha u)$ exists and

$$\partial_t(\partial^\alpha u) = \partial^\alpha \Delta u. \tag{3.22}$$

Consequently, $\partial_t(\partial^\alpha u)$ is continuous in (t,x).

Finally, we will prove that an arbitrary partial derivative of u is continuous in $\mathbb{R}_+ \times \Omega$. It suffices to prove the latter in the domain $(s,\infty)\times\Omega$, for any $s>0$. Fix $s>0$, a positive integer j, and consider the function

$$w(t) := \Delta^j u(t+s) = (-1)^j \sum_{k=1}^{\infty}\lambda_k^j e^{-\lambda_k(t+s)} a_k v_k = \sum_{k=1}^{\infty} e^{-\lambda_k t} b_k v_k$$

where

$$b_k = (-\lambda_k)^j e^{-\lambda_k s} a_k.$$

Since $\sum b_k^2 < \infty$, all the above arguments work for function w instead of u. Applying (3.22) to the function w instead of u, we obtain that, in $(s,\infty)\times\Omega$,

$$\partial_t\left(\partial^\alpha \Delta^j u\right) = \partial^\alpha\left(\Delta^{j+1} u\right). \tag{3.23}$$

It follows that

$$\partial_t^j(\partial^\alpha u) = \partial_t^{j-1}(\partial^\alpha \Delta u) = \partial_t^{j-2}\left(\partial^\alpha \Delta^2 u\right) = ... = \partial^\alpha \Delta^j u.$$

Applying this identity successively to an arbitrary partial derivative of the form

$$\partial_t^{j_1}\partial^{\alpha_1}\partial_t^{j_2}\partial^{\alpha_2}....\partial_t^{j_{p-1}}\partial^{\alpha_{p-1}}\partial_t^{j_p}\partial^{\alpha_p} u,$$

we bring this derivative to the form $\partial^\alpha \Delta^j u$ where $\alpha = \alpha_1 + ... + \alpha_p$ and $j = j_1 + ... + j_p$. It follows that any partial derivative of u exists and is continuous in (t,x), which finishes the proof. ■

3.4 Weak maximum principle

So far we have studied the following properties of the weighted Laplace operator:

- spectral properties, that is, eigenvalues and eigenfunctions;
- smoothness properties (for example, smoothness of solutions of mixed problems)

In this section, we consider properties of different kind, related to the *maximum principle.*

The spectral properties of more general differential and integral operators are studied in the spectral theory. The smoothness properties are characteristic to a larger class of *hypoelliptic operators.* Finally, the properties based on the maximum principle, are typical for *Markov operators* that are generators of Markov processes.

The Laplace operator is especially important as it belongs to the intersection of these areas of Mathematics.

Definition. A function $u : I \to L^2(\Omega)$ is called *subcaloric* in $I \times \Omega$ if

1. u is L^2-differentiable at any $t \in I$;
2. for any $t \in I$, we have $u(t) \in W^1(\Omega)$ and $\Delta u(t) \in L^2(\Omega)$, where Δ is understood in the weak sense;
3. for any $t \in I$, we have
$$\frac{du}{dt} \leq \Delta u(t). \tag{3.24}$$

In the same way, u is called *supercaloric* if

$$\frac{du}{dt} \geq \Delta u(t).$$

Definition. For functions $u, v \in W^1(\Omega)$ we write

$$u \leq v \operatorname{mod} W_0^1(\Omega) \tag{3.25}$$

if there is a function $w \in W_0^1(\Omega)$ such that $u \leq v + w$ in Ω.

The condition (3.25) can be regarded as a weak version of "$u \leq v$ on $\partial\Omega$".

Theorem 3.9 (Weak parabolic maximum principle) *Let u be a subcaloric function in $(0,T) \times \Omega$ such that*

(i) for any $t \in (0,T)$,
$$u(t) \leq 0 \operatorname{mod} W_0^1(\Omega). \tag{3.26}$$

(ii) $u(t)_+ \stackrel{L^2(\Omega)}{\longrightarrow} 0$ as $t \to 0$.

Then $u(t) \leq 0$ for all $t \in (0,T)$.

The condition (i) can be regarded as a weak version of "$u(t) \leq 0$ on $\partial\Omega$" and (ii) is a weak version of "$u(0) \leq 0$".

Theorem 3.9 can be reformulated as the *minimum principle* for supercaloric functions as follows. Assume that $u(t)$ is supercaloric in $(0,T) \times \Omega$ such that

(i) $u(t) \geq 0 \operatorname{mod} W_0^1(\Omega)$ for any $t \in (0,T)$;

(ii) $u(t)_- \xrightarrow{L^2} 0$ as $t \to 0$.

Then $u(t) \geq 0$ for all $t \in (0,T)$.

Example. Assuming that Ω is precompact, let $u(t)$ be a solution of the weak mixed problem (3.1) in Ω with the initial function $f \in L^2(\Omega)$. The function u is caloric in $\mathbb{R}_+ \times \Omega$ and, hence, supercaloric. Moreover, we have $u(t) \in W_0^1(\Omega)$ for all $t > 0$, that is, $u(t) = 0 \operatorname{mod} W_0^1(\Omega)$. We also know that $u(t) \xrightarrow{L^2} f$ as $t \to 0$. In particular, if $f \geq 0$ then $u(t)_- \xrightarrow{L^2} 0$ as $t \to 0$. Hence, by the minimum principle we conclude that $u(t) \geq 0$ for all $t > 0$. Similarly, if $f \leq 0$ then $u(t) \leq 0$. Consequently, if $f = 0$ then $u = 0$, which recovers the uniqueness result of Theorem 3.1.

For the proof of Theorem 3.9, we need the following lemma.

Lemma 3.10 *If $u \in W^1(\Omega)$ then the relation*

$$u \leq 0 \mod W_0^1(\Omega) \tag{3.27}$$

holds if and only if $u_+ \in W_0^1(\Omega)$.

Proof. In the proof we use the following facts:

- if $v \in W_0^1(\Omega)$ then also $v_+ \in W_0^1(\Omega)$ (*EDE*, Lemma 1.7).
- if $v_k \in W_0^1(\Omega)$ and $v_k \xrightarrow{W^1} v \in W_0^1(\Omega)$ then also $(v_k)_+ \xrightarrow{W^1} v_+$ (see Exercises).

If $u_+ \in W_0^1(\Omega)$ then (3.27) is satisfied because $u \leq u_+$. Conversely, we need to prove that if $u \leq w$ for some $w \in W_0^1(\Omega)$ then $u_+ \in W_0^1(\Omega)$.

Assume first that $w \in \mathcal{D}(\Omega)$, and let φ be a cutoff function of $\operatorname{supp} w$ in Ω (see Fig. 3.1). Then we have the following identity:

$$u_+ = ((1-\varphi)w + \varphi u)_+ . \tag{3.28}$$

Indeed, if $\varphi = 1$ then (3.28) is obviously satisfied. If $\varphi < 1$ then $w = 0$ and, hence, $u \leq 0$, so that the both sides of (3.28) vanish. Since $\varphi u \in W_0^1(\Omega)$ and $(1-\varphi)w \in \mathcal{D}(\Omega)$, it follows that

$$(1-\varphi)w + \varphi u \in W_0^1(\Omega).$$

By (3.28) we conclude that $u_+ \in W_0^1(\Omega)$.

For a general $w \in W_0^1(\Omega)$, let $\{w_k\}$ be a sequence of functions from $\mathcal{D}(\Omega)$ such that $w_k \xrightarrow{W^1} w$. Then we have

$$u_k := u + (w_k - w) \leq w_k,$$

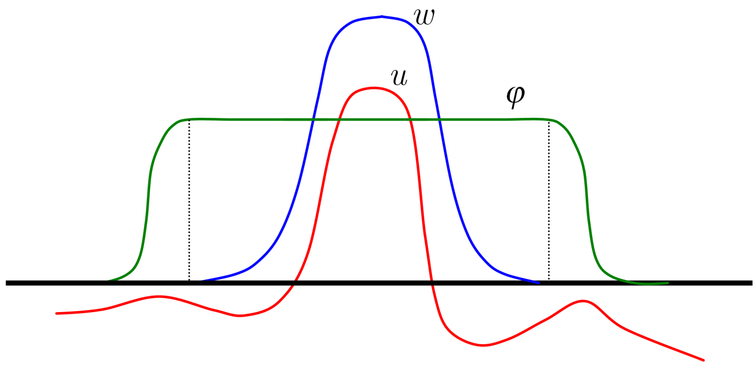

Figure 3.1: Functions u, w, φ

which implies by the first part of the proof that $(u_k)_+ \in W_0^1$. Since $u_k \overset{W^1}{\longrightarrow} u$, it follows that $(u_k)_+ \overset{W^1}{\longrightarrow} u_+$ (Exercise 67), whence we conclude that $u_+ \in W_0^1$. ■

17.01.17

Proof of Theorem 3.9. The inequality (3.24) means that, for any fixed $t \in (0, T)$ and any non-negative function $v \in \mathcal{D}(\Omega)$,

$$(u', v)_{L^2} \leq -(\nabla u, \nabla v)_{\vec{L}^2}, \tag{3.29}$$

where $u' \equiv \frac{du}{dt}$. Clearly, (3.29) extends to all non-negative functions $v \in W_0^1(\Omega)$.

Let a function $\varphi \in C^\infty(\mathbb{R})$ be such that, for some positive constant C,

$$\begin{cases} \varphi(s) = 0, & s \leq 0, \\ \varphi(s) > 0, & s > 0, \\ 0 \leq \varphi'(s) \leq C, & s \in \mathbb{R}. \end{cases} \tag{3.30}$$

By (3.26) and Lemma 3.10, we have $u(t)_+ \in W_0^1(\Omega)$, for any $t \in (0, T)$. Since φ is Lipschitz and vanishes at 0, we obtain that the function $\varphi(u(t)) = \varphi\left(u(t)_+\right)$ is also in $W_0^1(\Omega)$ and

$$\nabla\varphi(u) = \varphi'(u_+)\nabla u_+ = \varphi'(u)\nabla u,$$

where we drop the argument t for simplicity (EDE, Lemma 1.6). Setting $v = \varphi(u(t))$ in (3.29), we obtain

$$(u', \varphi(u))_{L^2} \leq -(\nabla u, \varphi'(u)\nabla u)_{\vec{L}^2} = -\int_\Omega \varphi'(u)|\nabla u|^2 d\mu \leq 0. \tag{3.31}$$

Let $\psi \in C^\infty(\mathbb{R})$ be another function satisfying (3.30). Using the product rule and the chain rule for L^2 derivatives (Exercises 65,66), we obtain

$$\begin{aligned} \frac{d}{dt}(u, \psi(u))_{L^2} &= (u', \psi(u))_{L^2} + (u, \psi'(u)u')_{L^2} \\ &= (u', \psi(u))_{L^2} + (u', \psi'(u)u)_{L^2} \\ &= (u', \psi(u) + \psi'(u)u)_{L^2}. \end{aligned} \tag{3.32}$$

Now choose ψ from the condition that

$$\psi(s) + \psi'(s)s = \varphi(s) \quad \forall s \in \mathbb{R},$$

that is, $(\psi(s)\,s)' = \varphi(s)$, which gives

$$\psi(s) = \frac{1}{s}\int_0^s \varphi(t)\,dt. \tag{3.33}$$

Note that the function ψ defined by (3.33) is C^∞ smooth on $\mathbb{R}$ because

$$\psi(s) = \frac{1}{s}\int_0^1 \varphi(s\xi)\,d(s\xi) = \int_0^1 \varphi(s\xi)\,d\xi.$$

It is easy to see from (3.33) that ψ satisfies (3.30). By (3.32) and (3.31) we obtain

$$\frac{d}{dt}(u,\psi(u))_{L^2} = (u',\varphi(u))_{L^2} \le 0.$$

Hence, $(u,\psi(u))_{L^2}$ as a function of t is decreasing in $(0,T)$. Since $\psi(s) \le Cs$ for any $s \ge 0$, we obtain that

$$(u,\psi(u))_{L^2} = (u_+,\psi(u_+))_{L^2} \le C(u_+,u_+)_{L^2} = C\|u_+\|^2_{L^2}.$$

By hypothesis, $\|u_+\|_{L^2} \to 0$ as $t \to 0$. Hence, the function $t \mapsto (u_+,\psi(u_+))_{L^2}$ is non-negative, decreasing on $(0,T)$ and goes to 0 as $t \to 0$. It follows that $(u_+,\psi(u_+))_{L^2} = 0$ for all $t \in (0,T)$, which implies that $u_+(t) = 0$ for all $t \in (0,T)$. Therefore, $u(t) \le 0$ for all $t \in (0,T)$, which was to be proved. ∎

Using the maximum/minimum principle, we prove further properties of the heat semigroup $P_t^\Omega f$, for any precompact open set $\Omega \subset M$.

Corollary 3.11 (Positivity-preserving property) *If $f \ge 0$ then $P_t^\Omega f \ge 0$.*

Proof. Consider the function $u(t) = P_t^\Omega f$ that is caloric, satisfies $u(t) = 0 \bmod W_0^1(\Omega)$ because $u(t) \in W_0^1(\Omega)$, and $u(t)_- \xrightarrow{L^2} 0$ as $t \to 0$ because $u(t) \xrightarrow{L^2} f \ge 0$. By the minimum principle we conclude that $u(t) \ge 0$, that is, $P_t^\Omega f \ge 0$. ∎

Corollary 3.12 (Minimality property of P_t^Ω) *Let u be a supercaloric function on $(0,T)\times\Omega$ such that*

(i) $u(t) \ge 0 \bmod W_0^1(\Omega)$ *for any* $t \in (0,T)$*;*

(ii) $L^2\text{-}\lim_{t\to0} u(t) \ge f$ *for some* $f \in L^2(\Omega)$.

Then, for all $t \in (0,T)$,

$$u(t) \ge P_t^\Omega f. \tag{3.34}$$

Proof. The function $v(t) = P_t^\Omega f - u(t)$ is obviously subcaloric in $(0,T)\times\Omega$ and satisfies the conditions:

(i) $v(t) \le 0 \bmod W_0^1(\Omega)$, because $P_t^\Omega f = 0 \bmod W_0^1(\Omega)$ and $u \ge 0 \bmod W_0^1(\Omega)$ for $t \in (0,T)$;

(ii) $v(t)_+ \xrightarrow{L^2} 0$ as $t \to 0$, because $L^2\text{-}\lim v(t) = L^2\text{-}\lim P_t^\Omega f - L^2\text{-}\lim u(t) \le f - f = 0$.

By Theorem 3.9, we conclude that $v(t) \leq 0$ whence (3.34) follows. ∎

Corollary 3.12 implies the following minimality property of $P_t^\Omega f$: if $f \geq 0$ then the function $u(t) = P_t^\Omega f$ is the *minimal* non-negative caloric function that satisfies the initial condition $u(t) \overset{L^2}{\to} f$. Indeed, this function is non-negative, caloric and satisfies the initial condition by Corollary 3.11 and Theorem 3.2. If $u(t)$ is any other function with these properties then by Corollary 3.12 we have (3.34), which means the minimality of $P_t^\Omega f$.

Corollary 3.13 (Submarkovian property) *If $f \leq 1$ then $P_t^\Omega f \leq 1$. Consequently, for any $f \in L^\infty(\Omega)$, we have $P_t^\Omega f \in L^\infty(\Omega)$ and*

$$\left\|P_t^\Omega f\right\|_{L^\infty} \leq \|f\|_{L^\infty}. \tag{3.35}$$

Proof. If $f \leq 1$ then consider the function $u(t) \equiv 1$ that is caloric and satisfies all the conditions of Corollary 3.12. It follows that $1 \geq P_t^\Omega f$, which was to be proved. For the proof of (3.35) it suffices to assume that $\|f\|_{L^\infty} = 1$. Then $f \leq 1$ implies $P_t^\Omega f \leq 1$, and $f \geq -1$ implies $P_t^\Omega f \geq -1$. Consequently, $\left\|P_t^\Omega f\right\|_{L^\infty} \leq 1$. ∎

In the next statement we compare $P_t^\Omega f$ in different domains. Any function $f \in L^2(\Omega)$ can be considered as an element of $L^2(M)$ by setting $f = 0$ outside Ω. In the same way, extend the function $P_t^\Omega f$ to the whole M by setting $P_t^\Omega f = 0$ in $M \setminus \Omega$.

Corollary 3.14 (Monotonicity property) *If $\Omega_1 \subset \Omega_2$ then $P_t^{\Omega_1} f \leq P_t^{\Omega_2} f$ for all non-negative $f \in L^2(\Omega_1)$.*

Proof. Consider the function $u(t) = P_t^{\Omega_2} f$ that is non-negative and caloric in $\mathbb{R}_+ \times \Omega_2$. Then it is also non-negative and caloric in $\mathbb{R}_+ \times \Omega_1$. Since $u(t) \overset{L^2(\Omega_2)}{\to} f$, it follows that also $u(t) \overset{L^2(\Omega_1)}{\to} f$. We conclude by Corollary 3.12 that $u(t) \geq P_t^{\Omega_1} f$, which was to be proved. ∎

3.5 The heat kernel in precompact domains

In this section we will prove that the operator P_t^Ω in $L^2(\Omega)$ has an *integral kernel*, that is, a function $p_t^\Omega(x,y)$ defined for $t > 0$ and $x, y \in \Omega$ such that

$$P_t^\Omega f(x) = \int_\Omega p_t^\Omega(x,y) f(y)\, d\mu(y).$$

If so then the function $p_t^\Omega(x,y)$ is called the *heat kernel* of Δ in Ω. We start with the following improvement of Theorem 3.7.

Theorem 3.15 *Let Ω be a precompact open subset of M. For any $f \in L^2(\Omega)$ and $t > 0$,*

$$\left\|P_t^\Omega f\right\|_{C(\Omega)} \leq C\left(1 + t^{-1}\right)^{\frac{n}{4}+1} \|f\|_{L^2(\Omega)}, \tag{3.36}$$

where $C = C(\Omega, \mathbf{g}, D, n)$.

19.01.17

As we will see from the proof, the constant C depends on a small open neighborhood of $\overline{\Omega}$.

Proof. The estimate (3.36) is an improvement of the estimate (3.12) of Theorem 3.7 where the norm in the left hand side was taken in $C(K)$ for a compact subset $K \subset \Omega$. In contrast, the estimate (3.36) provides the pointwise upper bound for $P_t^\Omega f$ uniformly in the entire domain Ω.

For the proof, let us first choose a precompact open subset V of M that covers $\overline{\Omega}$. Extend f to V by setting $f = 0$ in $V \setminus \Omega$. Applying the estimate (3.12) of Theorem 3.7 in the domain V and with $K = \overline{\Omega}$, we obtain that

$$\left\|P_t^V f\right\|_{C(\Omega)} \leq C\left(1+t^{-1}\right)^{\frac{n}{4}+1}\|f\|_{L^2(V)},$$

where $C = C(V, \Omega, \mathbf{g}, D, n)$. Assume further that f is non-negative. By Corollaries 3.11 and 3.14, we have

$$0 \leq P_t^\Omega f \leq P_t^V f,$$

whence it follows that

$$\left\|P_t^\Omega f\right\|_{C(\Omega)} \leq C\left(1+t^{-1}\right)^{\frac{n}{4}+1}\|f\|_{L^2(\Omega)}.$$

If f is signed then $f = f_+ - f_-$ and

$$\begin{aligned}
\left\|P_t^\Omega\right\|_{C(\Omega)} &= \left\|P_t^\Omega f_+ - P_t^\Omega f_-\right\|_{C(\Omega)} \\
&\leq \left\|P_t^\Omega f_+\right\|_{C(\Omega)} + \left\|P_t^\Omega f_-\right\|_{C(\Omega)} \\
&\leq C\left(1+t^{-1}\right)^{\frac{n}{4}+1}\left(\|f_+\|_{L^2(\Omega)} + \|f_-\|_{L^2(\Omega)}\right) \\
&\leq 2C\left(1+t^{-1}\right)^{\frac{n}{4}+1}\|f\|_{L^2(\Omega)},
\end{aligned}$$

which finishes the proof. ■

Corollary 3.16 *For any $t > 0$ and $x \in \Omega$, there exists a function $q_{t,x} \in L^2(\Omega)$ such that*

$$P_t^\Omega f(x) = \int_\Omega q_{t,x}(y) f(y)\, d\mu(y) \tag{3.37}$$

for all $f \in L^2(\Omega)$. Besides, we have

$$\|q_{t,x}\|_{L^2} \leq C\left(1+t^{-1}\right)^{\frac{n}{4}+1} =: \Phi(t), \tag{3.38}$$

where C is the same as in Theorem 3.15.

Proof. Fix $t > 0$ and $x \in \Omega$ and consider a linear mapping

$$\begin{array}{l} L^2(\Omega) \to \mathbb{R} \\ \quad f \ \mapsto \ P_t^\Omega f(x) \end{array} \tag{3.39}$$

By (3.36) we have

$$\left|P_t^\Omega f(x)\right| \leq \Phi(t)\|f\|_{L^2}. \tag{3.40}$$

Hence, the mapping (3.39) is bounded and, by the Riesz representation theorem, there is a function $q_{t,x} \in L^2(\Omega)$ such that

$$P_t^\Omega f(x) = (q_{t,x}, f)_{L^2},$$

which proves the first claim. Setting here $f = q_{t,x}$ and observing that

$$P_t^\Omega q_{t,x}(x) = \int_\Omega q_{t,x}^2 d\mu = \|q_{t,x}\|_{L^2}^2,$$

we obtain from (3.40) that

$$\|q_{t,x}\|_{L^2}^2 \le \Phi(t) \|q_{t,x}\|_{L^2},$$

whence (3.38) follows. ■

Definition. Define the *heat kernel* $p_t^\Omega(x,y)$ of Δ in Ω by the identity

$$p_t^\Omega(x,y) = q_{t,x}(y).$$

Clearly, $p_t^\Omega(x,y)$ (and $q_{t,x}(y)$) is the integral kernel of the operator P_t^Ω.

So far the heat kernel is defined as an L^2-function of y, for any $t>0$ and $x \in \Omega$. The next theorem shows that it is in fact a smooth function of t, x, y.

Theorem 3.17 *Let Ω be a non-empty precompact open subset of a weighted manifold M. Let $\{v_k\}_{k=1}^\infty$ be an orthonormal basis in $L^2(\Omega)$ that consists of the eigenfunctions of Δ in Ω. Then heat kernel $p_t^\Omega(x,y)$ admits the following eigenfunction expansion*

$$p_t^\Omega(x,y) = \sum_{k=1}^\infty e^{-\lambda_k t} v_k(x) v_k(y), \tag{3.41}$$

where the series converges absolutely and uniformly in $(t,x,y) \in [\varepsilon,\infty) \times \Omega \times \Omega$, for any $\varepsilon > 0$.

Besides, the series (3.41) converges in $C^m([\varepsilon,\infty) \times K \times K)$, for any positive integer m, for any $\varepsilon > 0$ and any compact subset $K \subset \Omega$ that is contained in a chart. Consequently, $p_t(x,y) \in C^\infty(\mathbb{R}_+ \times \Omega \times \Omega)$.

For the proof, we need the notion of *trace* of operators. Let A be a non-negative definite operator in a Hilbert space H, that is, $(Au,u) \ge 0$ for all $u \in H$. Let $\{h_k\}_{k=1}^\infty$ be an orthonormal basis in H. Define the trace of A by

$$\operatorname{trace} A = \sum_{k=1}^\infty (Ah_k, h_k).$$

The right hand side here is a series with non-negative terms, so its sum is always defined as an element of $[0,\infty]$. It is a general fact that the value of trace A does not depend on the choice of a basis. We do not prove this in general because in our specific case of $A = P_t^\Omega$ this will be clear otherwise.

Lemma 3.18 *In the setting of* Corollary 3.16, *we have, for any* $t > 0$,

$$\operatorname{trace} P_{2t}^{\Omega} = \int_{\Omega} \|q_{t,x}\|_{L^2}^2 d\mu(x) < \infty. \tag{3.42}$$

Besides, we have

$$\operatorname{trace} P_{2t}^{\Omega} = \sum_{k=1}^{\infty} e^{-2t\lambda_k(\Omega)}. \tag{3.43}$$

Consequently, the series (3.43) *converges.*

Proof. Observe first that operator P_{2t}^{Ω} is non-negative definite because for any $f \in L^2(\Omega)$, we have, by the semigroup property and symmetry of the heat semigroup,

$$\left(P_{2t}^{\Omega} f, f\right) = \left(P_t^{\Omega} P_t^{\Omega} f, f\right) = \left(P_t^{\Omega} f, P_t^{\Omega} f\right) \geq 0.$$

To prove (3.42), choose any orthonormal basis $\{h_k\}_{k=1}^{\infty}$ in L^2. Using (3.37), we obtain

$$\begin{aligned}
\operatorname{trace} P_{2t}^{\Omega} &= \sum_k \left(P_{2t}^{\Omega} h_k, h_k\right) = \sum_k \left(P_t^{\Omega} h_k, P_t^{\Omega} h_k\right) \\
&= \sum_k \int_{\Omega} \left(P_t^{\Omega} h_k(x)\right)^2 d\mu(x) \\
&= \sum_k \int_{\Omega} (q_{t,x}, h_k)^2 d\mu(x). \qquad (3.44)
\end{aligned}$$

Applying the Parseval identity in the basis $\{h_k\}$, we obtain

$$\sum_k (q_{t,x}, h_k)^2 = \|q_{t,x}\|_{L^2}^2. \tag{3.45}$$

Hence, (3.44) and (3.45) yield

$$\operatorname{trace} P_{2t}^{\Omega} = \int_{\Omega} \|q_{t,x}\|_{L^2}^2 d\mu(x),$$

which proves the first part of (3.42). The finiteness of the trace follows from (3.38):

$$\operatorname{trace} P_{2t}^{\Omega} \leq \int_{\Omega} \Phi(t)^2 d\mu(x) = \Phi(t)^2 \mu(\Omega) < \infty. \tag{3.46}$$

Finally, let us compute the trace in the orthonormal basis $\{v_k\}_{k=1}^{\infty}$ of the eigenfunctions of Δ in Ω, that is,

$$\operatorname{trace} P_{2t}^{\Omega} = \sum_{k=1}^{\infty} \left(P_{2t}^{\Omega} v_k, v_k\right)_{L^2}.$$

Observe that by the definition (3.11) of the operator P_{2t}^{Ω}, we have

$$P_{2t}^{\Omega} v_k = e^{-2t\lambda_k(\Omega)} v_k.$$

Hence, we obtain

$$\operatorname{trace} P_{2t}^{\Omega} = \sum_{k=1}^{\infty} \left(e^{-2t\lambda_k(\Omega)} v_k, v_k\right)_{L^2} = \sum_{k=1}^{\infty} e^{-2t\lambda_k(\Omega)},$$

which finishes the proof. ■

Proof of Theorem 3.17. Let us write $\lambda_k = \lambda_k(\Omega)$. By Lemma 3.18, we have, for any $t > 0$.

$$\sum_{k=1}^{\infty} e^{-t\lambda_k} < \infty. \tag{3.47}$$

As above, let $\{v_k\}$ be an orthonormal basis in $L^2(\Omega)$ that consists of eigenfunctions of Δ in Ω. Let us first prove that the series

$$\sum_{k=1}^{\infty} e^{-t\lambda_k} v_k(x)\, v_k(y) \tag{3.48}$$

converges absolutely and uniformly in the domain $t \geq \varepsilon, x \in \Omega, y \in \Omega$. By the Weierstrass M-test, it suffices to prove that

$$\sum_{k=1}^{\infty} \sup_{t\geq\varepsilon, x,y\in\Omega} \left|e^{-t\lambda_k} v_k(x)\, v_k(y)\right| < \infty. \tag{3.49}$$

Recall that, by Theorem 3.15, for any $f \in L^2(\Omega)$,

$$\sup_{x\in\Omega} \left|P_t^{\Omega} f(x)\right| \leq \Phi(t)\, \|f\|_{L^2},$$

where $\Phi(t)$ is defined in (3.38). Applying this to $f = v_k$ and using (3.55) and $\|v_k\|_{L^2} = 1$, we obtain

$$\sup_{x\in\Omega} \left|e^{-t\lambda_k} v_k(x)\right| \leq \Phi(t). \tag{3.50}$$

It follows that

$$\sup_{x,y\in\Omega} \left|e^{-2t\lambda_k} v_k(x)\, v_k(y)\right| \leq \Phi(t)^2.$$

Since function $\Phi(t)$ is decreasing in t, we obtain, for any $\varepsilon > 0$,

$$\sum_{k=1}^{\infty} \sup_{t\geq\varepsilon, x,y\in\Omega} \left|e^{-3t\lambda_k} v_k(x)\, v_k(y)\right| \leq \sum_{k=1}^{\infty} \sup_{x,y\in\Omega} \left|e^{-3\varepsilon\lambda_k} v_k(x)\, v_k(y)\right| \leq \Phi(\varepsilon)^2 \sum_{k=1}^{\infty} e^{-\varepsilon\lambda_k} < \infty, \tag{3.51}$$

where the right hand side is finite by (3.47). Renaming $3t$ to t and 3ε to ε, we obtain (3.49).

Now let us show that the series (3.48) converges in $C^m([\varepsilon, \infty) \times K \times K)$. Again, it suffices to prove that

$$\sum_{k=1}^{\infty} \left\|e^{-\lambda_k t} v_k(x)\, v_k(y)\right\|_{C^m([\varepsilon,\infty)\times K\times K)} < \infty. \tag{3.52}$$

By Corollary 2.8 and (2.30), we have

$$\|v_k\|_{C^m(K)} \leq C(\lambda_k + 1)^{\frac{m}{2}+\frac{n}{4}+\frac{1}{2}} \|v_k\|_{W^1(\Omega)} = C(\lambda_k + 1)^{\sigma}, \tag{3.53}$$

where $\sigma = \frac{m}{2} + \frac{n}{4} + 1$. For any partial derivative

$$\partial_x^{\alpha} \partial_y^{\beta} \partial_t^{\gamma},$$

where α, β are n-dimensional multiindices and γ is a non-negative integer such that

$$|\alpha| + |\beta| + \gamma \leq m,$$

we have

$$\partial_x^\alpha \partial_y^\beta \partial_t^\gamma \left(e^{-\lambda_k t} v_k(x) v_k(y)\right) = (-\lambda_k)^\gamma e^{-\lambda_k t} \partial_x^\alpha v_k(x) \partial_y^\beta v_k(y).$$

It follows from (3.53) that

$$\sup_{t \geq \varepsilon, x \in K, y \in K} \left|\partial_x^\alpha \partial_y^\beta \partial_t^\gamma \left(e^{-\lambda_k t} v_k(x) v_k(y)\right)\right| \leq C \lambda_k^\gamma e^{-\lambda_k \varepsilon} (\lambda_k + 1)^{2\sigma}.$$

Since $\gamma \leq m < 2\sigma$, we obtain

$$\sup_{t \geq \varepsilon, x \in K, y \in K} \left|\partial_x^\alpha \partial_y^\beta \partial_t^\gamma \left(e^{-\lambda_k t} v_k(x) v_k(y)\right)\right| \leq C e^{-\lambda_k \varepsilon} (\lambda_k + 1)^{4\sigma},$$

whence

$$\left\|e^{-\lambda_k t} v_k(x) v_k(y)\right\|_{C^m([\varepsilon, \infty) \times K \times K)} \leq C e^{-\lambda_k \varepsilon} (\lambda_k + 1)^{4\sigma}. \tag{3.54}$$

Finally, it follows from (3.47) that

$$\sum_{k=1}^\infty e^{-\lambda_k \varepsilon} (\lambda_k + 1)^{4\sigma} \leq \sup_{\lambda > 0} \left[(\lambda + 1)^{4\sigma} e^{-\lambda \varepsilon / 2}\right] \sum_{k=1}^\infty e^{-\lambda_k \varepsilon / 2} < \infty,$$

which proves (3.52).

We are let to prove that the sum of the series (3.48) is equal to $p_t^\Omega(x, y)$. Using the notation $q_{t,x}$ as above and noticing that

$$(q_{t,x}, v_k)_{L^2} = P_t^\Omega v_k(x) = e^{-t\lambda_k} v_k(x), \tag{3.55}$$

we obtain the following expansion of $q_{t,x}$ in the basis $\{v_k\}$:

$$q_{t,x} = \sum_{k=1}^\infty e^{-t\lambda_k} v_k(x) v_k, \tag{3.56}$$

that is,

$$p_t^\Omega(x, y) = \sum_{k=1}^\infty e^{-t\lambda_k} v_k(x) v_k(y),$$

where the series converges in $L^2(\Omega)$ in variable y, for any fixed $x \in \Omega$ and $t > 0$. Since this series converges also in $C(\Omega)$ in variable y, it determines a continuous function of y that is a continuous version of the L^2 function of y. Hence, we see that $p_t^\Omega(x, y)$ is defined for all $t > 0$ and $x, y \in \Omega$, and it is C^∞ jointly in t, x, y by the previous argument. ■

Remark. In Theorem 3.8, we have proved that the function $u(t, x) = P_t^\Omega(x)$ is jointly C^∞ in $t > 0$ and $x \in \Omega$, for any $f \in L^2(\Omega)$. The above approach to the proof of Theorem 3.17 can be used in order to obtain a simpler proof of Theorem 3.8. Indeed, we know that

$$P_t f = \sum_{k=1}^\infty e^{-\lambda_k t} a_k v_k, \tag{3.57}$$

where the series converges in $L^2(\Omega)$ for any $t>0$, and the coefficients a_k are determined from

$$f=\sum_{k=1}^{\infty} a_k v_k.$$

We claim that, for any compact set $K\subset\Omega$, any $\varepsilon>0$ and any positive integer m,

$$\sum_{k=1}^{\infty}\left\|e^{-\lambda_k t}a_k v_k(x)\right\|_{C^m([\varepsilon,\infty)\times K)}<\infty,$$

which will imply that $u\in C^\infty(\mathbb{R}_+\times\Omega)$. Indeed, we have by the Cauchy-Schwarz inequality

$$\begin{aligned}\sum_{k=1}^{\infty}\left\|e^{-\lambda_k t}a_k v_k(x)\right\|_{C^m([\varepsilon,\infty)\times K)} &\le \left(\sum_{k=1}^{\infty}a_k^2\right)^{1/2}\\ &\quad\times\left(\sum_{k=1}^{\infty}\left\|e^{-\lambda_k t}v_k(x)\right\|^2_{C^m([\varepsilon,\infty)\times K)}\right)^{1/2}.\end{aligned}$$

Clearly, the first series converges by the Parseval identity, while the second series converges by the same argument as above, because similarly to (3.54)

$$\left\|e^{-\lambda_k t}v_k(x)\right\|_{C^m([\varepsilon,\infty)\times K)}\le Ce^{-\lambda_k\varepsilon}(\lambda_k+1)^{3\sigma}.$$

24.01.17

Moreover, using (3.50), we obtain in the same way that

$$\sum_{k=1}^{\infty}\left\|e^{-\lambda_k t}a_k v_k(x)\right\|_{C([\varepsilon,\infty)\times\Omega)}<\infty,$$

which implies that the series (3.57) converges absolutely and uniformly in $[\varepsilon,\infty)\times\Omega$.

Remark. If the boundary $\partial\Omega$ is smooth, for example, a C^1-submanifold, then one can show that $v_k\in C\left(\overline{\Omega}\right)$ and $v_k|_{\partial\Omega}=0$ (similarly to the proof of Theorem 4.5 in *EDE*). The fact that the series in (3.57) and (3.41) converge absolutely and uniformly in $(t,x,y)\in[\varepsilon,\infty)\times\Omega\times\Omega$, implies that $P_t^\Omega f(x)=0$ when $x\in\partial\Omega$ and also $p_t(x,y)=0$ when one of the points x,y belongs to $\partial\Omega$.

3.6 Further properties of the heat kernel

As above, let Ω be a precompact open subset of M. In the previous section, we have constructed the heat kernel $p_t^\Omega(x,y)$ that is a C^∞-function of $(t,x,y)\in\mathbb{R}_+\times\Omega\times\Omega$ given by the series (3.41). This function is the integral kernel of P_t^Ω, that is, for all $f\in L^2(\Omega)$, $x\in\Omega$ and $t>0$ that

$$P_t^\Omega f(x)=\int_\Omega p_t^\Omega(x,y)f(y)\,d\mu(y). \tag{3.58}$$

Further properties of the heat kernel are stated in the following theorem.

Theorem 3.19 *In any precompact domain $\Omega \subset M$, the heat kernel has the following properties.*

(a) *Positivity: $p_t^\Omega(x,y) \geq 0$, for all $x,y \in \Omega$ and $t>0$.*

(b) *Submarkovian property: for all $x \in \Omega$ and $t>0$*

$$\int_\Omega p_t^\Omega(x,y)\,d\mu(y) \leq 1. \tag{3.59}$$

(c) *Symmetry: $p_t^\Omega(x,y) \equiv p_t^\Omega(y,x)$, for all $x,y \in \Omega$ and $t>0$.*

(d) *The heat equation: for any fixed $y \in \Omega$, the function $(t,x) \mapsto p_t(x,y)$ is caloric in $\mathbb{R}_+ \times \Omega$; moreover, it solves the heat equation $\partial_t u = \Delta u$ also in the classical sense.*

(e) *The boundary condition: $p_t^\Omega(\cdot,y) \in W_0^1(\Omega)$, for all $y \in \Omega$ and $t>0$.*

(f) *The semigroup identity: for all $x,y \in \Omega$ and $t,s>0$,*

$$p_{t+s}^\Omega(x,y) = \int_\Omega p_t^\Omega(x,z)\,p_s^\Omega(z,y)\,d\mu(z). \tag{3.60}$$

(g) *If $\Omega_1 \subset \Omega_2$ then $p_t^{\Omega_1}(x,y) \leq p_t^{\Omega_2}(x,y)$ for all $x,y \in \Omega_1$ and $t>0$.*

Proof. (a) Assume from the contrary that $p_{t_0}(x_0,y_0) < 0$ at some (t_0,x_0,y_0). By the continuity of the heat kernel, there is an open neighborhood U of y_0 such that $p_{t_0}(x_0,y) < 0$ for all $y \in U$. Choose a non-negative non-zero function $f \in \mathcal{D}(U)$. Then we have

$$P_{t_0}^\Omega f(x_0) = \int_U p_{t_0}^\Omega(x_0,y)\,f(y)\,d\mu(y) < 0,$$

while by Corollary 3.11 we must have $P_{t_0}^\Omega f(x_0) \geq 0$. This contradiction shows that $p_t(x,y) \geq 0$.

(b) By Corollary 3.13, $f \leq 1$ implies $P_t^\Omega f(x) \leq 1$ for all $x \in M$ and $t>0$. Taking $f = 1_\Omega$, we obtain

$$\int_\Omega p_t^\Omega(x,y)\,d\mu(y) \leq 1,$$

which was to be proved.

(c) The symmetry follows trivially from the eigenfunction expansion (3.41).

$(d)+(e)$ Fix $y \in \Omega$. As follows from the proof of Theorem 3.17, the series

$$u(t) := p_t^\Omega(\cdot,y) = \sum_{k=1}^\infty e^{-\lambda_k t} v_k(y)\,v_k, \tag{3.61}$$

converges in $L^2(\Omega)$ for any $t>0$. Indeed, (3.61) is obtained from (3.56) by switching the variables x and y and using the symmetry of the heat kernel. For any $t>s>0$, we obtain using (3.57) that

$$u(t) = \sum_{k=1}^\infty e^{-\lambda_k t} v_k(y)\,v_k = \sum_{k=1}^\infty e^{-\lambda_k(t-s)}\left(e^{-\lambda_k s} v_k(y)\right) v_k = P_{t-s}^\Omega u(s). \tag{3.62}$$

Since $u(s) \in L^2(\Omega)$, by the properties of the heat semigroup (Theorem 3.2) we obtain that $u(t)$ is caloric in the domain $t > s$ and $u(t) \in W_0^1(\Omega)$ for any $t > s$. Since s is arbitrary, we the same properties hold for $t > 0$.

Since the function $u(t,x) = p_t^\Omega(x,y)$ is C^∞-smooth, its L^2-derivative $\frac{d}{dt}u$ coincides with the classical derivative and the classical Laplacian Δu coincides with the weak Laplacian, whence it follows that u satisfies the classical heat equation $\partial_t u = \Delta u$. Alternatively, the latter can be seen by computing $\partial_t u$ and Δu by means of differentiating the series (3.61) term-by-term, which is possible because that series converges in any C^m.

(g) Rewriting the identity (3.62) by using (3.58) and the definition (3.61) of the function u, we obtain

$$p_t^\Omega(x,y) = \int_\Omega p_{t-s}^\Omega(x,z)\, u(s,z)\, d\mu(z) = \int_\Omega p_{t-s}^\Omega(x,z)\, p_s^\Omega(z,y)\, d\mu(z),$$

which is equivalent to (3.60).

(h) For all $t > 0$ and $x, y \in \Omega_1$, set

$$q_t(x,y) := p_t^{\Omega_2}(x,y) - p_t^{\Omega_1}(x,y).$$

By Corollary 3.14, for any non-negative $f \in L^2(\Omega_1)$ we have

$$\int_\Omega q_t(x,y)\, f(y)\, d\mu(y) = P_t^{\Omega_2} f(x) - P_t^{\Omega_1} f(x) \geq 0.$$

Arguing as in the proof of (a), we conclude that $q_t(x,y) \geq 0$, which finishes the proof.
∎

3.7 The initial condition

As we know, for any $f \in L^2(\Omega)$, we have

$$P_t^\Omega f \overset{L^2}{\to} f \quad \text{as } t \to 0.$$

Here we show that the convergence is "better" is the function f is "better".

Theorem 3.20 *Let Ω be a precompact open subset of M.*

(a) For any function $f \in \mathcal{D}(\Omega)$, we have

$$P_t^\Omega f \to f \text{ as } t \to 0, \tag{3.63}$$

where the convergence is in $C^m(K)$, for any positive integer m and any compact set $K \subset \Omega$ that is contained in a chart.

(b) For any open set $U \subset \Omega$ and for any $x \in U$, we have

$$\int_U p_t^\Omega(x,y)\, d\mu(y) \to 1 \text{ as } t \to 0, \tag{3.64}$$

where the convergence is local uniform in U.

(c) For any $f \in C_b(\Omega)$, the convergence (3.63) is locally uniform in Ω, that is, in $C(K)$ for any compact subset $K \subset \Omega$.

Proof. (a) If $f \in \mathcal{D}(\Omega)$ then also $\Delta^j f \in \mathcal{D}(\Omega) \subset W_0^1(\Omega)$ for any non-negative integer j. Hence, if $f = \sum_{k=1}^{\infty} a_k v_k$ then

$$\Delta^j f = (-1)^j \sum_{k=1}^{\infty} \lambda_k^j a_k v_k$$

(cf. Exercise 42), where the series converges in $W^1(\Omega)$. On the other hand, we have

$$\Delta^j P_t^{\Omega} f = (-1)^j \sum_{k=1}^{\infty} \lambda_k^j e^{-\lambda_k t} a_k v_k \in W_0^1(\Omega)$$

(see the identity (3.13) in the proof of Theorem 3.7). By Lemma 3.5 (with $H = W_0^1(\Omega)$ and $\gamma_k(t) = e^{-\lambda_k t}$), we obtain

$$\sum_{k=1}^{\infty} \lambda_k^j e^{-\lambda_k t} a_k v_k \overset{W^1(\Omega)}{\longrightarrow} \sum_{k=1}^{\infty} \lambda_k^j a_k v_k \quad \text{as } t \to 0,$$

that is

$$\Delta^j \left(P_t^{\Omega} f - f\right) \overset{W^1(\Omega)}{\longrightarrow} 0 \quad \text{as } t \to 0.$$

By Theorem 2.7 we conclude that

$$P_t^{\Omega} f - f \overset{C^m(K)}{\longrightarrow} 0 \quad \text{as } t \to 0, \tag{3.65}$$

which was to be proved.

(b) Let f be a cutoff function of $\{x\}$ in U, that is, $f \in \mathcal{D}(\Omega)$, $f = 1$ in a neighborhood of x and $0 \le f \le 1$. Then by (a)

$$\int_U p_t^{\Omega}(x,y)\, d\mu(y) \ge \int_{\Omega} p_t^{\Omega}(x,y) f(y)\, d\mu(y) = P_t^{\Omega} f(x) \to f(x) = 1$$

as $t \to 0$, where the convergence is local uniform in x. Since also

$$\int_U p_t^{\Omega}(x,y)\, d\mu(y) \le 1,$$

the convergence (3.64) follows.

(c) We have

$$\begin{aligned} P_t^{\Omega} f(x) - f(x) &= \int_{\Omega} p_t^{\Omega}(x,y)\left(f(y) - f(x)\right) d\mu(y) \\ &\quad + \left(\int_{\Omega} p_t^{\Omega}(x,y)\, d\mu(y) - 1\right) f(x). \end{aligned}$$

By (b) we obtain that

$$\left(\int_{\Omega} p_t^{\Omega}(x,y)\, d\mu(y) - 1\right) f(x) \to 0 \tag{3.66}$$

as $t \to 0$, where the convergence is local uniform in x. Choose an open set U containing x and such that $|f(y) - f(x)| \leq \varepsilon$ for any $y \in U$, where $\varepsilon > 0$ is prescribed. Then we have

$$\begin{aligned}
\left|\int_\Omega p_t^\Omega(x,y)\,(f(y)-f(x))\,d\mu(y)\right| &\leq \left|\int_U p_t^\Omega(x,y)\,(f(y)-f(x))\,d\mu(y)\right| \\
&\quad + \left|\int_{\Omega\setminus U} p_t^\Omega(x,y)\,(f(y)-f(x))\,d\mu(y)\right| \\
&\leq \varepsilon\int_U p_t^\Omega(x,y)\,d\mu(y) + 2\sup|f|\int_{\Omega\setminus U} p_t^\Omega(x,y)\,d\mu(y) \\
&\leq \varepsilon + 2\sup|f|\left(1-\int_U p_t^\Omega(x,y)\,d\mu(y)\right).
\end{aligned}$$

As $t \to 0$ we obtain using (b) that

$$\limsup_{t\to 0}\left|\int_\Omega p_t^\Omega(x,y)\,(f(y)-f(x))\,d\mu(y)\right| \leq \varepsilon.$$

Since ε is arbitrary, it follows that

$$\int_\Omega p_t^\Omega(x,y)\,(f(y)-f(x))\,d\mu(y) \to 0$$

as $t \to 0$. Combining with (3.66), we obtain $P_t^\Omega f(x) \to f(x)$ as $t \to 0$. Finally, it remains to observe that the above argument yields also the local uniform convergence in x. ■

Remark. The convergence (3.63) implies that, for any $y \in M$,

$$\int_\Omega p_t^\Omega(x,y)\,f(x)\,d\mu(x) \to f(y) \quad \text{as } t \to 0,$$

which means that $p_t^\Omega(\cdot,y) \to \delta_y$ where δ_y is the Dirac delta-function, and the convergence to δ_y is understood in the sense of distributions.

Remark. Recall that, for any $f \in L^2(\Omega)$, the function

$$u(t,x) = P_t^\Omega f(x)$$

solves the heat equation in $\mathbb{R}_+ \times \Omega$ in the classical sense and with the initial condition

$$u(t,\cdot) \overset{L^2(\Omega)}{\longrightarrow} f \quad \text{as } t \to 0.$$

If $f \in C_b(\Omega)$ then by Theorem 3.20 we have also that

$$u(t,\cdot) \overset{C(K)}{\longrightarrow} f \quad \text{as } t \to 0.$$

If in addition the boundary $\partial\Omega$ is a C^1-submanifold then $u(t,x)$ extends continuously to $\overline{\Omega}$ and vanishes on $\partial\Omega$, for any $t > 0$. Hence, we conclude that in this case the function u solves the *classical* mixed problem:

$$\begin{cases} \partial_t u = \Delta u, \\ u(t,\cdot)|_{\partial\Omega} = 0, \\ u(t,x) \to f(x) \ \text{as } t \to 0, \end{cases}$$

where the convergence to the initial function is locally uniform in Ω.

Chapter 4

Global heat semigroup

26.01.17

In this Chapter we construct the heat semigroup $\{P_t\}$ and the heat kernel $p_t(x,y)$ on the entire weighted manifold M.

4.1 Convergence issues

Let us first observe the following consequence of Theorem 2.7.

Proposition 4.1 *Let $\{u_k\}$ be a sequence of smooth functions on a weighted manifold M, each satisfying the same equation*

$$\Delta u_k = f,$$

where $f \in C^\infty(M)$. If, for some $u \in W^1_{loc}(M)$,

$$u_k \overset{W^1_{loc}(M)}{\longrightarrow} u \quad \textit{as } k \to \infty,$$

then the function u is C^∞-smooth in M and satisfies the equation $\Delta u = f$.

Proof. For any indices k,l we have $\Delta(u_k - u_l) = 0$ and, hence, $\Delta^j(u_k - u_l) = 0$, where j is any positive integer. This implies by Theorem 2.7 that

$$\|u_k - u_l\|_{C^m(K)} \leq C\|u_k - u_l\|_{W^1(\Omega)},$$

where Ω is any precompact open neighborhood of K. Since

$$\|u_k - u_l\|_{W^1(\Omega)} \to 0 \text{ as } k,l \to \infty,$$

it follows that also

$$\|u_k - u_l\|_{C^m(K)} \to 0 \text{ as } k,l \to \infty.$$

Hence, $\{u_k\}$ converges in $C^m(K)$, and the limit is necessarily u. Since m is arbitrary, this implies that $u \in C^\infty(M)$ and u satisfies $\Delta u = f$. ■

In this Chapter we accept without proof the following theorem that extends Proposition 4.1 to the heat equation and relaxes the W^1_{loc}-converges to that of L^1_{loc}.

Theorem 4.2 ([3], Theorem 7.4) *Let I be an open interval in $\mathbb{R}$ and M be a weighted manifold. Let $\{u_k\}$ be a sequence of smooth functions on the manifold $N := I \times M$, each satisfying the same equation*

$$\partial_t u_k - \Delta u_k = f,$$

where $f \in C^\infty(N)$. If, for some $u \in L^1_{loc}(N)$,

$$u_k \overset{L^1_{loc}(N)}{\longrightarrow} u \quad \text{as } k \to \infty$$

then the function u is C^∞-smooth in N and satisfies the equation

$$\partial_t u - \Delta u = f.$$

The proof of this theorem requires the regularity theory for the parabolic equations, that is similar to that of the elliptic equations.

4.2 The heat semigroup on M

Given a non-negative function $f \in L^2_{loc}(M)$, let us construct a function $P_t f$ for any $t > 0$ as follows. For any precompact open set $\Omega \subset M$ and $t > 0$, define $P_t^\Omega f$ as a function on M as follows:

$$P_t^\Omega f = \begin{cases} P_t^\Omega(f\mathbf{1}_\Omega) & \text{in } \Omega \\ 0, & \text{outside } \Omega. \end{cases}$$

Fix an *exhaustion sequence* $\{\Omega_k\}_{k=1}^\infty$ of M by precompact open subsets.

Lemma 4.3 *If $f \geq 0$ then the sequence of functions $\left\{P_t^{\Omega_k} f\right\}$ is monotone increasing in k. Moreover, the limit $\lim_{k\to\infty} P_t^{\Omega_k} f(x)$ does not depend on the choice of $\{\Omega_k\}$.*

Proof. Let us show that $P_t^{\Omega_k} f \geq P_t^{\Omega_{k-1}} f$. Outside Ω_{k-1} this is obvious because $P_t^{\Omega_k} f \geq 0 = P_t^{\Omega_{k-1}} f$. In Ω_{k-1} we have, using Corollaries 3.11 and 3.14, that

$$P_t^{\Omega_k} f = P_t^{\Omega_k}(f\mathbf{1}_{\Omega_k}) = P_t^{\Omega_k}\left(f\mathbf{1}_{\Omega_k\setminus\Omega_{k-1}}\right) + P_t^{\Omega_k}\left(f\mathbf{1}_{\Omega_{k-1}}\right) \geq P_t^{\Omega_{k-1}}\left(f\mathbf{1}_{\Omega_{k-1}}\right) = P_t^{\Omega_{k-1}} f.$$

If these is one more exhaustion sequence $\{\Omega'_k\}$ then for any Ω_k there is $\Omega'_m \supset \Omega_k$ which implies

$$P_t^{\Omega_k} f \leq P_t^{\Omega'_m} f$$

and, hence,

$$\lim_{k\to\infty} P_t^{\Omega_k} f \leq \lim_{k\to\infty} P_t^{\Omega'_k} f.$$

Since the opposite inequality is true by the same argument, we obtain the identity of the two limits. ■

For any non-negative function $f \in L^2_{loc}(M)$ and for all $t > 0$ and $x \in M$, set

$$P_t f(x) := \lim_{k\to\infty} P_t^{\Omega_k} f(x).$$

In general, $P_t f(x)$ may take values in $[0, \infty]$.

Lemma 4.4 *If $P_t f \in L^2_{loc}(\mathbb{R}_+ \times M)$ then the function $P_t f$ is C^∞ smooth and solves the heat equation in $\mathbb{R}_+ \times M$.*

Proof. Indeed, by the dominated convergence theorem, we obtain that

$$P_t^{\Omega_k} f \overset{L^2_{loc}(\mathbb{R}_+\times M)}{\longrightarrow} P_t f.$$

Since each of the functions $(t,x) \mapsto P_t^{\Omega_k} f$ solves the heat equation in $\mathbb{R}_+ \times \Omega_1$ and L^2_{loc}-convergence implies that in L^1_{loc}, it follows from Theorem 4.2 that $P_t f$ is C^∞-smooth in $\mathbb{R}_+ \times \Omega_1$ and solves in this domain the heat equation. Since Ω_1 can be chosen arbitrarily, we obtain that the same properties of $P_t f$ are true in $\mathbb{R}_+ \times M$. ■

Lemma 4.5 *Let $u(t,x)$ be a non-negative smooth solution to the heat equation in $\mathbb{R}_+ \times M$ such that*

$$u(t,\cdot) \overset{L^2_{loc}}{\longrightarrow} f \quad \text{as } t \to 0, \tag{4.1}$$

for some $f \in L^2_{loc}(M)$. Then $P_t f(x)$ is also a smooth solution to the heat equation in $\mathbb{R}_+ \times M$, satisfying the initial condition (4.1), and

$$u(t,x) \geq P_t f(x), \tag{4.2}$$

for all $t > 0$ and $x \in M$.

Proof. For any precompact open set $\Omega \subset M$, the function $u(t,x)$ is non-negative and caloric in $\mathbb{R}_+ \times \Omega$, and satisfies $u(t,\cdot) \overset{L^2(\Omega)}{\to} f$. By the minimality property of P_t^Ω (Corollary 3.12), we conclude that

$$u(t,x) \geq P_t^\Omega f(x),$$

whence (4.2) follows by letting $\Omega \to M$ (that is, by considering $\Omega = \Omega_k$ for an exhaustion sequence $\{\Omega_k\}$ and letting $k \to \infty$). Hence, the function $P_t f$ belongs to $L^2_{loc}(\mathbb{R}_+ \times M)$, and by Lemma 4.4 we conclude that $P_t f$ is smooth and satisfies the heat equation.

Finally, $P_t f \overset{L^2_{loc}}{\longrightarrow} f$ as $t \to 0$ follows from

$$f \overset{L^2(\Omega)}{\leftarrow} P_t^\Omega f \leq P_t f \leq u(t,\cdot) \overset{L^2(\Omega)}{\to} f$$

as $t \to 0$. ■

If f is a signed function from $L^2_{loc}(M)$, then consider $P_t f_+$ and $P_t f_-$. If they both are in $L^2_{loc}(\mathbb{R}_+ \times M)$ then we define

$$P_t f := P_t f_+ - P_t f_-.$$

In this case $P_t f$ also solves the heat equation in $\mathbb{R}_+ \times M$.

Theorem 4.6 *For any $f \in L^2(M)$, the function $P_t f$ belongs to $L^2_{loc}(\mathbb{R}_+ \times M)$ and, hence, is C^∞-smooth and solves the heat equation in $\mathbb{R}_+ \times M$. Besides, for any $t > 0$,*

$$\|P_t f\|_{L^2(M)} \leq \|f\|_{L^2(M)} \tag{4.3}$$

and

$$P_t f \overset{L^2(M)}{\to} f \quad \text{as } t \to 0. \tag{4.4}$$

Proof. If (4.3) is already proved then it implies $P_t f \in L^2_{loc}(\mathbb{R}_+ \times M)$. Hence, we need only to prove (4.3) and (4.4). Assume first that $f \geq 0$. Then we have, for any precompact domain $\Omega \subset M$, that

$$\left\| P_t^\Omega f \right\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)}.$$

Hence, letting $\Omega \to M$, we obtain (4.3).

In order to prove (4.4) for a non-negative f, observe that we have the following conditions:

$$\begin{array}{ccc} P_t^{\Omega_k} f & \leq & P_t f \\ {\scriptstyle t\to 0}\downarrow^{L^2(M)} & & \leq_{\|\cdot\|_{L^2}} \\ f\mathbf{1}_{\Omega_k} & \underset{k\to\infty}{\overset{L^2(M)}{\longrightarrow}} & f \end{array}$$

Using Lemma 4.7 to be stated and proved below, we conclude that $P_t f \overset{L^2(M)}{\to} f$ as $t \to 0$.

Let now f be signed. Then $f = f_+ - f_-$ where both f_+ and f_- belong to $L^2(M)$. Hence, we conclude that

$$P_t f = P_t f_+ - P_t f_-$$

is in $L^2_{loc}(\mathbb{R}_+ \times M)$. To prove (4.3), we have

$$\begin{aligned} \|P_t f\|_{L^2}^2 &= \|P_t f_+ - P_t f_-\|_{L^2}^2 \\ &= \|P_t f_+\|_{L^2}^2 + \|P_t f_-\|_{L^2}^2 - 2\left(P_t f_+, P_t f_-\right)_{L^2} \\ &\leq \|P_t f_+\|_{L^2}^2 + \|P_t f_-\|_{L^2}^2 \\ &\leq \|f_+\|_{L^2}^2 + \|f_-\|_{L^2}^2 = \|f\|_{L^2}^2. \end{aligned}$$

And for (4.4) we have, as $t \to 0$,

$$P_t f = P_t f_+ - P_t f_- \overset{L^2}{\to} f_+ - f_- = f.$$

■

Now we prove the lemma used in the above proof.

Lemma 4.7 *Let $\{u_{ik}\}$ be a double sequence of non-negative functions from $L^2(M)$ such that, for any k,*

$$u_{ik} \overset{L^2}{\to} f_k \in L^2(M) \quad \text{as } i \to \infty$$

and

$$f_k \overset{L^2}{\to} f \in L^2(M) \quad \text{as } k \to \infty.$$

Let $\{u_i\}$ be a sequence of functions from $L^2(M)$ such that, for all i, k,

$$u_{ik} \leq u_i \quad \text{and} \quad \|u_i\|_{L^2} \leq \|f\|_{L^2}.$$

Then $u_i \overset{L^2}{\to} f$ as $i \to \infty$.

Proof. All the hypotheses can be displayed in schematic form in the following diagram:

$$\begin{array}{ccc} u_{ik} & \leq & u_i \\ {\scriptstyle i\to\infty}\downarrow{\scriptstyle L^2} & & \leq_{\|\cdot\|_{L^2}} \\ f_k & \xrightarrow[k\to\infty]{L^2} & f \end{array}$$

where all notation are self-explanatory. We need to prove that also $u_i \xrightarrow{L^2} f$ as $i \to \infty$.

Given $\varepsilon > 0$, we have, for large enough k,

$$\|f - f_k\|_{L^2} \leq \varepsilon.$$

Fix one of such indices k. Then, for large enough i, we have

$$\|f_k - u_{ik}\|_{L^2} \leq \varepsilon$$

so that

$$\|f - u_{ik}\|_{L^2} \leq 2\varepsilon. \tag{4.5}$$

Let us show that, for such i,

$$\|f - u_i\|_{L^2}^2 \leq \Phi(\varepsilon), \tag{4.6}$$

with some function Φ such that $\Phi(\varepsilon) \to 0$ as $\varepsilon \to 0$, which will settle the claim.

Set

$$g = (f - u_i)_+ \text{ and } h = (f - u_i)_-,$$

and estimate the L^2-norms of g and h separately. By condition $u_{ik} \leq u_i$, we have

$$f - u_i \leq f - u_{ik}$$

whence

$$g = (f - u_i)_+ \leq (f - u_{ik})_+$$

and by (4.5)

$$\|g\|_{L^2} \leq 2\varepsilon. \tag{4.7}$$

In order to prove a similar estimate for $\|h\|_{L^2}$, let us first prove the following inequality, any any $x \in M$:

$$h^2 \leq u_i^2 + 2fg - f^2. \tag{4.8}$$

Indeed, in the domain $\{f \geq u_i\}$ we have $h = 0$, $g = f - u_i$, and (4.8) follows from

$$u_i^2 + 2fg - f^2 = u_i^2 + 2f(f - u_i) - f^2 = u_i^2 - 2fu_i + f^2 = (u_i - f)^2 \geq 0 = h^2.$$

In the domain $\{f < u_i\}$ we have $g = 0$, $h = u_i - f$ and (4.8) follows from

$$u_i^2 + 2fg - f^2 = u_i^2 - f^2 = (u_i + f)(u_i - f) \geq (u_i - f)^2 = h^2.$$

Integrating (4.8) over M and substituting $\|u_i\|_{L^2} \leq \|f\|_{L^2}$ and $\|g\|_{L^2} \leq 2\varepsilon$ (cf. (4.7)), we obtain

$$\|h\|_{L^2}^2 \leq \|u_i\|_{L^2}^2 + 2(f,g)_{L^2} - \|f\|_{L^2}^2 \leq 2(f,g)_{L^2} \leq 2\|f\|_{L^2}\|g\|_{L^2} \leq 4\varepsilon\|f\|_{L^2}.$$

It follows that

$$\|f - u_i\|_{L^2}^2 = \|g\|_{L^2}^2 + \|h\|_{L^2}^2 \leq 4\varepsilon^2 + 4\varepsilon\|f\|_{L^2},$$

which proves (4.6). ∎

31.01.17

4.3 The global heat kernel

For any precompact open set $\Omega \subset M$, extend the heat kernel $p_t^\Omega(x,y)$ from $x,y \in \Omega$ to $x,y \in M$ by setting $p_t^\Omega(x,y) = 0$ if one of the points x,y is outside Ω. For any exhaustion sequence $\{\Omega_k\}$, the sequence $\left\{p_t^{\Omega_k}(x,y)\right\}$ is monotone increasing by Theorem 3.19 and, hence, has the limit

$$p_t(x,y) = \lim_{k\to\infty} p_t^{\Omega_k}(x,y),$$

that is independent of the choice of $\{\Omega_k\}$ (the proof is similar to that of Lemma 4.3).

Definition. The function is called *the heat kernel* of Δ in M.

Theorem 4.8 *The heat kernel has the following properties.*

(a) *Finiteness and smoothness:* $p_t(x,y) \in C^\infty(\mathbb{R}_+ \times M \times M)$

(b) *Positivity:* $p_t(x,y) \geq 0$;

(c) *Submarkovian property:*

$$\int_M p_t(x,y)\,d\mu(y) \leq 1$$

(d) *Symmetry:* $p_t(x,y) = p_t(y,x)$.

(e) *The heat equation: for any fixed $y \in M$, the function $u(t,x) = p_t(x,y)$ solves the heat equation $\partial_t u = \Delta u$ in $\mathbb{R}_+ \times M$.*

(f) *Approximation of identity: for any open set $U \subset M$ and for any $x \in U$,*

$$\int_U p_t(x,y)\,d\mu(y) \to 1 \text{ as } t \to 0, \tag{4.9}$$

where the convergence is locally uniform in x. Moreover, for any $f \in C_b(M)$,

$$P_t f(x) \to f(x) \quad \text{as } t \to 0,$$

where the convergence is locally uniform in x.

(g) *The semigroup identity:*

$$p_{t+s}(x,y) = \int_\Omega p_t(x,z)\,p_s(x,y)\,d\mu(z).$$

(h) *The heat semigroup kernel: for all non-negative $f \in L^2_{loc}(M)$ (and for all $f \in L^2(M)$),*

$$P_t f(x) = \int_M p_t(x,y)\,f(y)\,d\mu(y). \tag{4.10}$$

Proof. $(b)+(c)+(d)+(g)$ follow immediately from the corresponding properties of $p_t^\Omega(x,y)$ by letting $\Omega\to M$. Note that at this moment we allow $p_t(x,y)$ to take the value ∞ which will be excluded in (a).

(a) The submarkovian property implies that $p_t(x,y)\in L^1_{loc}(\mathbb{R}_+\times M\times M)$. Consequently, the pointwise convergence

$$p_t^{\Omega_k}(x,y)\to p_t(x,y) \tag{4.11}$$

is also in $L^1_{loc}(\mathbb{R}_+\times M\times M)$. Consider the weighted product $M\times M$ and observe that the function

$$u(t,(x,y))=p_t^\Omega(x,y)$$

solves in $\mathbb{R}_+\times\Omega\times\Omega$ the following equation

$$2\partial_t u=\Delta_x u+\Delta_y u=\Delta_{(x,y)}u$$

where Δ_x and Δ_y denote the Laplace operators on M with respect to the variables x,y while $\Delta_{(x,y)}$ denotes the Laplace operator on $M\times M$ (see (1.58).

Hence, up to the time change $2t\to t$, the functions $p_t^{\Omega_k}(x,y)$ satisfy the heat equation in $\mathbb{R}_+\times\Omega_k\times\Omega_k$. By Theorem 4.2, we conclude that the limit $p_t(x,y)$ is C^∞-smooth on $\mathbb{R}_+\times M\times M$.

(e) Now apply the same argument with a fixed y. Since the function $(t,x)\mapsto p_t(x,y)$ is smooth, it is in $L^1_{loc}(\mathbb{R}_+\times M)$ and, hence, the convergence (4.11) is also in $L^1_{loc}(\mathbb{R}_+\times M)$, whence we obtain by Theorem 4.2 that $p_t(x,y)$ satisfies the heat equation in $\mathbb{R}_+\times M$.

(h) If $f\in L^2_{loc}(M)$ then, for any precompact open set $\Omega\subset M$, we have $f\in L^2(\Omega)$ and, hence,

$$P_t^\Omega f(x)=\int_\Omega p_t^\Omega(x,y)f(y)\,d\mu(y).$$

If f is non-negative then passing to the limit as $\Omega\to M$, we obtain (4.10) by the monotone convergence theorem.

If $f\in L^2(M)$ is signed then we have by the above argument the identity (4.10) for f_+ and f_-. Since by Theorem 4.6 the functions P_tf_+ and P_tf_- are finite (moreover, they are smooth), it follows that P_tf is well define and satisfies (4.10).

(f) Without loss of generality, we can assume that U is precompact. Let Ω be any precompact open set containing U. Then we have by Theorem 3.20

$$\int_U p_t(x,y)\,d\mu(y)\geq\int_U p_t^\Omega(x,y)\,d\mu(y)\to 1\ \text{ as } t\to 0,$$

while

$$\int_U p_t(x,y)\,d\mu(y)\leq 1,$$

whence (4.9) follows. The second claim is proved in the same way as that in Theorem 3.20. ■

4.4 Fundamental solutions

Definition. A C^∞-function $u(t,x)$ of $t>0$ and $x\in M$ is called a *fundamental solution* (of the heat equation) in M at $y\in M$ if

(i) $\partial_t u=\Delta u$ in $\mathbb{R}_+\times M$;

(ii) for any $f\in\mathcal{D}(M)$,

$$\int_M u(t,x)\,f(x)\,d\mu(x)\to f(y)\quad\text{as }t\to 0,$$

that will be shortly written as

$$u(t,\cdot)\to\delta_y\ \text{ as }t\to 0.$$

If in addition $u\geq 0$ and, for all $t>0$,

$$\int_M u(t,x)\,d\mu(x)\leq 1,\tag{4.12}$$

then u is called a *regular* fundamental solution.

Example. It is known that the following Gauss-Weierstrass function in $\mathbb{R}^n$ is a regular fundamental solution at 0:

$$u(t,x)=\frac{1}{(4\pi t)^{n/2}}\exp\left(-\frac{|x|^2}{4t}\right).$$

Lemma 4.9 *Let $u(t,x)$ be a smooth non-negative function on $\mathbb{R}_+\times M$ satisfying (4.12). Fix $y\in M$. Then the following conditions are equivalent:*

(a) $u(t,\cdot)\to\delta_y$ *as* $t\to 0$.

(b) *For any open set U containing y,*

$$\int_U u(t,\cdot)\,d\mu\to 1\ \text{ as }t\to 0.\tag{4.13}$$

(c) *For any $f\in C_b(M)$,*

$$\int_M u(t,\cdot)\,f\,d\mu\to f(y)\ \text{ as }t\to 0.\tag{4.14}$$

In particular, if u is a regular fundamental solution at y, then u satisfies (b) and (c).

Proof. The implication $(c)\Rightarrow(a)$ is trivial because $u(t,\cdot)\to\delta_y$ is equivalent to (4.14) for all $f\in\mathcal{D}(M)$.

The rest of the proof is practically identical to the proof of Theorem 3.20$(b),(c)$.

$(a) \Rightarrow (b)$. Let $f \in \mathcal{D}(U)$ be a cutoff function of the set $\{y\}$ in U. Then (4.14) holds for this f. Since $f(y) = 1$ and

$$\int_M u(t,\cdot) f\, d\mu \leq \int_U u(t,\cdot)\, d\mu \leq 1,$$

(4.13) follows from (4.14).

$(b) \Rightarrow (c)$. For any open set U containing y, we have

$$\begin{aligned}\int_M u(t,x) f(x)\, d\mu(x) &= \int_{M\setminus U} u(t,x) f(x)\, d\mu(x) \\ &\quad + \int_U u(t,x)\left(f(x) - f(y)\right)\, d\mu(x) \\ &\quad + f(y) \int_U u(t,x)\, d\mu(x).\end{aligned}$$

The last term here tends to $f(y)$ by (4.13). The other terms are estimated as follows:

$$\left|\int_{M\setminus U} u(t,x) f(x)\, d\mu\right| \leq \sup|f| \int_{M\setminus U} u(t,x)\, d\mu(x) \tag{4.15}$$

and

$$\begin{aligned}\left|\int_U u(t,x)\left(f(x) - f(y)\right) d\mu\right| &\leq \sup_{x\in U}|f(x) - f(y)| \int_U u(t,x)\, d\mu(x) \\ &\leq \sup_{x\in U}|f(x) - f(y)|. \end{aligned} \tag{4.16}$$

Obviously, the right hand side of (4.15) tends to 0 as $t \to 0$ due to (4.12) and (4.13). By the continuity of f at y, the right hand side of (4.16) can be made arbitrarily small uniformly in t by choosing U to be a small enough neighborhood of y. Combining the above three lines, we obtain (4.14). ∎

Remark. As we see from the last part of the proof, (4.14), in fact, holds for arbitrary $f \in L^\infty(M)$ provided f is continuous at the point y.

4.5 Heat kernel as a fundamental solution

Theorem 4.10 *For any $y \in M$, the heat kernel $p_t(x,y)$ is the minimal regular fundamental solution of the heat equation at y.*

Proof. The heat kernel is a regular fundamental solution by Theorem 4.8.

Let $u(t,x)$ be another regular fundamental solution at y. Fix $s > 0$. The function $t,x \mapsto u(t+s,x)$ satisfies the heat equation in $\mathbb{R}_+ \times M$ and, hence, $u(t+s,x)$ can be considered as a non-negative solution to the Cauchy problem in $\mathbb{R}_+ \times M$ with the initial function $f(x) = u(s,x)$. Since u is a smooth function, we have $f \in L^2_{loc}(M)$ and

$$u(t+s,\cdot) \xrightarrow{L^2_{loc}} f \text{ as } t \to 0.$$

By Lemma 4.5, we conclude that, for all $t > 0$ and $x \in M$,

$$u(t+s,x) \geq P_t f(x) = \int_M p_t(x,z)\, u(s,z)\, d\mu(z). \tag{4.17}$$

Fix now $t > 0$, $x \in M$ and choose an open set $\Omega \Subset M$ containing y. Then $p_t(x,\cdot) \in C_b(\Omega)$ and, by Lemma 4.9 in Ω,

$$\int_\Omega p_t(x,z)\, u(s,z)\, d\mu(z) \to p_t(x,y) \quad \text{as } s \to 0.$$

Hence, letting $s \to 0$ in (4.17), we obtain $u(t,x) \geq p_t(x,y)$, which was to be proved. ■

Theorem 4.11 *Let $u(t,x)$ be a regular fundamental solution to the heat equation at $y \in M$. If $u(t,x) \to 0$ as $x \to \infty$ where the convergence is uniform in $t \in (0,T)$ for any $T > 0$, then $u(t,x) \equiv p_t(x,y)$.*

Proof. By Theorem 4.10, we have $u(t,x) \geq p_t(x,y)$ so that we only need to prove the opposite inequality.

Fix some $\varepsilon > 0$. By the hypothesis $u(t,x) \to 0$ as $x \to \infty$, there is a compact set K such that $u(t,x) < \varepsilon$ for all $x \in M \setminus K$ and $t \in (0,T)$. Choose any precompact open set Ω containing K. Fix also some $s > 0$, set

$$f(x) = u(s,x),$$

and consider function

$$v(t,x) = u(t+s,x) - P_t^\Omega f(x) - \varepsilon.$$

The function $(t,x) \mapsto u(t+s,x)$ solves the heat equation in $(0,T-s) \times M$ which implies that it is caloric in $(0,T-s) \times \Omega$. Since the latter is true also for $P_t^\Omega f(x)$ and for the constant function ε, we see that $v(t,x)$ is caloric in $(0,T-s) \times \Omega$.

For each $t \in (0,T-s)$, we have

$$v(t,x) < 0 \quad \forall x \in \Omega \setminus K,$$

which implies that $\operatorname{supp} v(t,\cdot) \subset K$ and, hence, $v(t,\cdot)_+ \in W_0^1(\Omega)$, that is,

$$v(t,\cdot) \leq 0 \operatorname{mod} W_0^1(\Omega).$$

As $t \to 0$, we have

$$u(t+s,\cdot) \overset{\Omega}{\rightrightarrows} u(s,\cdot) = f,$$

which implies that also

$$u(t+s,\cdot) \overset{L^2(\Omega)}{\longrightarrow} f.$$

Since also

$$P_t^\Omega f \overset{L^2(\Omega)}{\longrightarrow} f,$$

it follows that

$$v(t,\cdot) \overset{L^2(\Omega)}{\longrightarrow} -\varepsilon$$

and, hence,

$$v(t,\cdot)_+ \overset{L^2(\Omega)}{\longrightarrow} 0 \quad \text{as } t \to 0.$$

By Theorem 3.9, we conclude that $v(t,x) \le 0$ for all $t \in (0, T-s)$ and $x \in \Omega$.

It follows that in Ω

$$u(t+s,\cdot) \le P_t^{\Omega} u(s,\cdot) + \varepsilon$$

whence, for any $x \in \Omega$,

$$u(t+s,x) \le \int_\Omega p_t(x,z)\, u(s,z)\, d\mu(z) + \varepsilon.$$

Letting here $s \to 0$ and applying Lemma 4.9 in Ω with function $f = p_t(x,\cdot) \in C_b(\Omega)$, we obtain that

$$\int_\Omega p_t(x,z)\, u(s,z)\, d\mu(z) \to p_t(x,y)$$

and, hence,

$$u(t,x) \le p_t(x,y) + \varepsilon,$$

for all $x \in \Omega$. Since Ω is arbitrary, this inequality holds for all $x \in M$. Finally, since $\varepsilon > 0$ is arbitrary, we conclude $u(t,x) \le p_t(x,y)$, which finishes the proof. ■

Example. As we know, the Gauss-Weierstrass function

$$p_t(x,y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{4t}\right) \tag{4.18}$$

is a regular fundamental solution of the heat equation in $\mathbb{R}^n$. By Theorem 4.11, we conclude that $p_t(x,y)$ is the heat kernel on $\mathbb{R}^n$ because $p_t(x,y) \to 0$ as $x \to \infty$ uniformly in t.

02.02.17

4.6 Heat kernel and isometries

Theorem 4.12 *Let $J : M \to M$ be an isometry of a weighted manifold $(M, \mathbf{g}, \mu)$. Then the heat kernel of M is J-invariant, that is, for all $t > 0$ and $x, y \in M$,*

$$p_t(Jx, Jy) = p_t(x,y) \tag{4.19}$$

Proof. Let us first show that the function $u(t,x) = p_t(Jx, Jy)$ is a regular fundamental solution at y. Indeed, by Lemma 1.21, for any smooth function f on M,

$$(\Delta f)(Jx) = \Delta(f(Jx)).$$

Applying this for $f = p_t(\cdot, Jy)$, we obtain

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial t} p_t(Jx, Jy) = (\Delta p_t)(Jx, Jy) = \Delta u,$$

so that u solves the heat equation.

By Lemma 1.21, we have the identity

$$\int_M f(Jx)\,d\mu(x) = \int_M f(z)\,d\mu(z), \tag{4.20}$$

for any non-negative function f. It follows that

$$\int_M u(t,x)\,d\mu(x) = \int_M p_t(z,Jy)\,d\mu(z) \le 1$$

and, similarly, for any open set U containing y,

$$\int_U u(t,x)\,d\mu(x) = \int_{JU} p_t(z,Jy)\,d\mu(z) \to 1 \text{ as } t\to 0.$$

Therefore, u is a regular fundamental solution. By Theorem 4.10, we conclude that

$$u(t,x) \ge p_t(x,y),$$

that is,

$$p_t(Jx,Jy) \ge p_t(x,y).$$

Applying the same argument to J^{-1} instead of J, we obtain the opposite inequality, which finishes the proof. ∎

Example. By Exercise 56, for any four points $x,y,x',y' \in \mathbb{H}^n$ such that

$$d(x',y') = d(x,y),$$

there exists a Riemannian isometry $J:\mathbb{H}^n\to\mathbb{H}^n$ such that $Jx'=x$ and $Jy'=y$. By Theorem 4.12, we conclude

$$p_t(x',y') = p_t(x,y).$$

Hence, $p_t(x,y)$, as a function of x,y, depends only on the distance $d(x,y)$.

The same applies to the heat kernel on $\mathbb{S}^n$.

4.7 Heat kernel on model manifolds

Let $(M,\mathbf{g},\mu)$ be a weighed model as in Section 1.12. That is, M is a ball $B_{r_0} = \{|x|<r_0\}$ in $\mathbb{R}^n$ (with $r_0\in(0,\infty]$) with a metric $\mathbf{g} = dr^2+\psi^2(r)\,\mathbf{g}_{\mathbb{S}^{n-1}}$ (where (r,θ) are the polar coordinates) and a density function $D=D(r)$. Let $S(r)$ be the area function of $(M,\mathbf{g},\mu)$ that is, $S(r) := \omega_n\psi^{n-1}(r)\,D(r)$, and let $p_t(x,y)$ be the heat kernel.

Let $(M,\widetilde{\mathbf{g}},\widetilde{\mu})$ be another weighted model based on the same smooth manifold M, and let $\widetilde{S}(r)$ and $\widetilde{p}_t(x,y)$ be its area function and heat kernel, respectively.

Theorem 4.13 *If $S(r)\equiv\widetilde{S}(r)$ then $p_t(x,o)=\widetilde{p}_t(y,o)$ for all $x,y\in M$ such that $|x|=|y|$.*

Note that the area function $S(r)$ does not fully identify the structure of the weighted model unless the latter is a Riemannian model. Nevertheless, $p_t(x,0)$ is completely determined by this function.

Proof. Let us first show that $p_t(x,o) = p_t(y,o)$ if $|x| = |y|$. Indeed, there is a rotation J of $\mathbb{R}^n$ such that $Jx = Jy$ and $Jo = o$. Since J is an isometry of $(M, \mathbf{g}, \mu)$, we obtain by Theorem 4.12 that p_t is J-invariant, which implies the claim.

By Lemma 4.9, the fact that a smooth non-negative function $u(t,x)$ on $\mathbb{R}_+ \times M$ is a regular fundamental solution at 0, is equivalent to the conditions

$$\begin{cases} \partial_t u = \Delta u, \\ \displaystyle\int_M u(t,x)\, d\mu(x) \le 1, \\ \displaystyle\int_{B_\varepsilon} u(t,x)\, d\mu(x) \to 1 \quad \text{as } t \to 0, \end{cases} \tag{4.21}$$

for all $0 < \varepsilon < r_0$. The heat kernel $p_t(x,o)$ is a regular fundamental solution on $(M, \mathbf{g}, \mu)$ at the point o, and it depends only on t and $r = |x|$ so that we can write $p_t(x,o) = u(t,r)$.

Using the fact that u does not depend on the polar angle, we obtain from (1.94)

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{S'(r)}{S(r)} \frac{\partial u}{\partial r}.$$

For $0 < \varepsilon < r_0$, we have by (1.91), (1.88), (1.93)

$$\int_{B_\varepsilon} u d\mu = \frac{1}{\omega_n} \int_0^\varepsilon \int_{\mathbb{S}^{n-1}} u\,(t,r)\, S(r)\, d\theta dr = \int_0^\varepsilon u(t,r)\, S(r)\, dr.$$

Hence, we obtain the following equivalent form of (4.21):

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial r^2} + \dfrac{S'(r)}{S(r)} \dfrac{\partial u}{\partial r}, \\ \displaystyle\int_0^{r_0} u(t,r)\, S(r)\, dr \le 1, \\ \displaystyle\int_0^\varepsilon u(t,r)\, S(r)\, dr \to 1 \quad \text{as } t \to 0. \end{cases} \tag{4.22}$$

It is important that all the conditions in (4.22) depend on the geometry of M only via the area function $S(r)$. Since by hypothesis $S(r) = \widetilde{S}(r)$, the conditions (4.22) are satisfied also with S replaced by $\widetilde{S}$, which means that $u(t,r)$ is a regular fundamental solution at 0 also on the manifold $(M, \widetilde{\mathbf{g}}, \widetilde{\mu})$. By Theorem 4.10, we conclude that $u(t,|x|) \ge \widetilde{p}_t(x,0)$, that is,

$$p_t(x,0) \ge \widetilde{p}_t(x,0).$$

The opposite inequality follows in the same way by switching p_t and $\widetilde{p}_t$, which finishes the proof. ∎

4.8 Heat kernel and change of measure

Let $(M, \mathbf{g}, h)$ be a weighted manifold. Any smooth positive function h on M determines a new measure $\widetilde{\mu}$ on M by

$$d\widetilde{\mu} = h^2 d\mu, \tag{4.23}$$

and, hence, a new weighted manifold $(M, \mathbf{g}, \widetilde{\mu})$. Denote by $\widetilde{\Delta}$ and $\widetilde{p}_t$ respectively the Laplace operator and the heat kernel on $(M, \mathbf{g}, \widetilde{\mu})$.

Theorem 4.14 *Let h be a smooth positive function on M that satisfies the equation*

$$\Delta h + \alpha h = 0, \tag{4.24}$$

where α is a real constant. Then the following identities holds

$$\widetilde{\Delta} = \frac{1}{h} \circ \Delta \circ h + \alpha \,\mathrm{id}, \tag{4.25}$$

$$\widetilde{p}_t(x, y) = e^{\alpha t} \frac{p_t(x, y)}{h(x)\, h(y)}, \tag{4.26}$$

for all $t > 0$ and $x, y \in M$.

The change of measure (4.23) satisfying (4.24) and the associated change of operator (4.25) are referred to as Doob's *h-transform.*

Proof. By the definition of the weighted Laplace operator, we obtain, for any smooth function f on M,

$$\begin{aligned} \widetilde{\Delta} f &= \frac{1}{h^2} \mathrm{div}_{\mathbf{g},\mu}(h^2 \nabla f) = \mathrm{div}_{\mathbf{g},\mu}(\nabla f) + \frac{1}{h^2} \langle \nabla h^2, \nabla f \rangle_{\mathbf{g}} \\ &= \Delta f + 2\langle \frac{\nabla h}{h}, \nabla f \rangle_{\mathbf{g}}. \end{aligned} \tag{4.27}$$

On the other hand, using the equation (4.24) and the product rule for Δ, we obtain

$$\begin{aligned} \frac{1}{h}\Delta(hf) &= \frac{1}{h}\left(h\Delta f + 2\langle \nabla h, \nabla f\rangle_{\mathbf{g}} + f\Delta h\right) \\ &= \Delta f + 2\langle \frac{\nabla h}{h}, \nabla f \rangle_{\mathbf{g}} + f\frac{\Delta h}{h} \\ &= \widetilde{\Delta} f - \alpha f. \end{aligned}$$

Hence, we have proved the identity

$$\widetilde{\Delta} f = \frac{1}{h}\Delta(hf) + \alpha f, \tag{4.28}$$

that is equivalent to (4.25). We have proved this identity for smooth f, but similarly it holds when Δ is understood in the weak sense.

In order to prove (4.26), it suffices to prove the same identity for the heat kernels $\widetilde{p}_t^{\Omega}$ and p_t^{Ω} for any precompact open set $\Omega \subset M$. If v is an eigenfunction of Δ in Ω with an eigenvalue λ then we have

$$\widetilde{\Delta}\left(\frac{v}{h}\right) = \frac{1}{h}(\Delta + \alpha) v = (-\lambda + \alpha)\frac{v}{h}$$

that is, $\frac{v}{h}$ is an eigenfunction of $\widetilde{\Delta}$ with the eigenvalue $\lambda-\alpha$ (of course, the same holds for the weak eigenfunctions). Observe that the mapping

$$u \mapsto \frac{u}{h}$$

is an isometry from $L^2(\Omega,\mu)$ to $L^2(\Omega,\widetilde{\mu})$ because for any $u \in L^2(\Omega,\mu)$,

$$\left\|\frac{u}{h}\right\|^2_{L^2(\Omega,\widetilde{\mu})} = \int_\Omega \left(\frac{u}{h}\right)^2 h^2 d\mu = \int_\Omega u^2 d\mu = \|u\|^2_{L^2(\Omega,\mu)}.$$

Therefore, if $\{v_k\}$ is an orthonormal basis in $L^2(\Omega,\mu)$ that consists of the eigenfunctions of Δ with eigenvalues $\{\lambda_k\}$, then the sequence $\left\{\frac{v_k}{h}\right\}$ is an orthonormal basis in $L^2(\Omega,\widetilde{\mu})$ that consists of the eigenfunctions of $\widetilde{\Delta}$ with eigenvalues $\{\lambda_k-\alpha\}$. Therefore, we obtain

$$\begin{aligned}\widetilde{p}^\Omega_t(x,y) &= \sum_k e^{-(\lambda_k-\alpha)t}\frac{v_k(x)}{h(x)}\frac{v_k(y)}{h(y)}\\ &= \frac{e^{\alpha t}}{h(x)h(y)}\sum_k e^{-\lambda_k t}v_k(x)v_k(y) = \frac{e^{\alpha t}p^\Omega_t(x,y)}{h(x)h(y)},\end{aligned}$$

which was to be proved. ∎

Example. The heat kernel in $(\mathbb{R}^1, \mathbf{g}_{\mathbb{R}^1}, \mu)$ with the Lebesgue measure μ is given by

$$p_t(x,y) = \frac{1}{(4\pi t)^{1/2}}\exp\left(-\frac{|x-y|^2}{4t}\right). \tag{4.29}$$

Let h be any positive smooth function on $\mathbb{R}^1$ that determines a new measure $\widetilde{\mu}$ on $\mathbb{R}^1$ by $d\widetilde{\mu} = h^2 d\mu$. Then we have $\Delta = \frac{d^2}{dx^2}$ and

$$\widetilde{\Delta} = \frac{1}{h^2}\frac{d}{dx}\left(h^2\frac{d}{dx}\right) = \frac{d^2}{dx^2} + 2\frac{h'}{h}\frac{d}{dx} \tag{4.30}$$

(cf. (4.27)). The equation (4.24) becomes

$$h'' + \alpha h = 0,$$

which is satisfied, for example, if $h(x) = \cosh\beta x$ and $\alpha = -\beta^2$. In this case, we have by (4.30)

$$\widetilde{\Delta} = \frac{d^2}{dx^2} + 2\beta\coth\beta x\frac{d}{dx}.$$

By Theorem 4.14, we obtain

$$\begin{aligned}\widetilde{p}_t(x,y) &= e^{\alpha t}\frac{p_t(x,y)}{h(x)h(y)}\\ &= \frac{1}{(4\pi t)^{1/2}}\frac{1}{\cosh\beta x\cosh\beta y}\exp\left(-\frac{|x-y|^2}{4t} - \beta^2 t\right).\end{aligned}$$

Example. Consider in $\mathbb{R}^1$ measure μ is given by

$$d\mu = e^{x^2} dx,$$

where dx is the Lebesgue measure. Then, by (4.30) with $h = e^{\frac{1}{2}x^2}$,

$$\Delta = \frac{d^2}{dx^2} + 2x\frac{d}{dx}. \tag{4.31}$$

We claim that the heat kernel $p_t(x, y)$ of $(\mathbb{R}, \mathbf{g}_{\mathbb{R}}, \mu)$ is given by the explicit formula:

$$p_t(x, y) = \frac{1}{(2\pi \sinh 2t)^{1/2}} \exp\left(\frac{2xye^{-2t} - x^2 - y^2}{1 - e^{-4t}} - t\right), \tag{4.32}$$

that is called the *Mehler kernel*. It is a matter of a routine (but hideous) computation to verify that the function (4.32) does solve the heat equation and satisfy the conditions of Lemma 4.9, which implies that is it a regular fundamental solution. It is easy to see that

$$p_t(x, y) \le \frac{1}{(4\pi t)^{1/2}} \exp\left(-\frac{|x - y|^2}{4t} - t\right),$$

which implies that $p_t(x, y) \to 0$ as $x \to \infty$ uniformly in t. Hence, we conclude by Theorem 4.11 that $p_t(x, y)$ is indeed the heat kernel.

Example. Continuing the previous example, it easily follows from (4.31) that function

$$h(x) = e^{-x^2}$$

satisfies the equation

$$\Delta h + 2h = 0.$$

Clearly, the change of measure $d\widetilde{\mu} = h^2 d\mu$ is equivalent to

$$d\widetilde{\mu} = e^{-x^2} dx.$$

By Theorem 4.14 and (4.32), we obtain that the heat kernel $\widetilde{p}_t$ of $(\mathbb{R}, \mathbf{g}_{\mathbb{R}}, \widetilde{\mu})$ is given by

$$\begin{aligned}\widetilde{p}_t(x, y) &= e^{2t}\frac{p_t(x, y)}{h(x)h(y)} = p_t(x, y)\exp\left(x^2 + y^2 + 2t\right) \\ &= \frac{1}{(2\pi \sinh 2t)^{1/2}} \exp\left(\frac{2xye^{-2t} - (x^2 + y^2)e^{-4t}}{1 - e^{-4t}} + t\right).\end{aligned}$$

4.9 Heat kernel on $\mathbb{H}^3$

As was shown in Example 4.6, the heat kernel $p_t(x, y)$ in the hyperbolic space $\mathbb{H}^n$ is a function of $r = d(x, y)$ and t.

Theorem 4.15 *The heat kernel of $\mathbb{H}^3$ is given by the following formula:*

$$p_t(x, y) = \frac{1}{(4\pi t)^{3/2}} \frac{r}{\sinh r} \exp\left(-\frac{r^2}{4t} - t\right). \tag{4.33}$$

The following formulas for $p_t(x,y)$ in $\mathbb{H}^n$ are known: if $n=2m+1$ then

$$p_t(x,y)=\frac{(-1)^m}{(2\pi)^m(4\pi t)^{1/2}}\left(\frac{1}{\sinh r}\frac{\partial}{\partial r}\right)^m e^{-m^2t-\frac{r^2}{4t}}, \tag{4.34}$$

which in the case $n=3$ gives (4.33), and if $n=2m+2$ then

$$p_t(x,y)=\frac{(-1)^m\sqrt{2}}{(2\pi)^m(4\pi t)^{3/2}}e^{-\frac{(2m+1)^2}{4}t}\left(\frac{1}{\sinh r}\frac{\partial}{\partial r}\right)^m\int_r^\infty\frac{se^{-\frac{s^2}{4t}}ds}{(\cosh s-\cosh r)^{\frac12}}. \tag{4.35}$$

In particular, the heat kernel in $\mathbb{H}^2$ is given by

$$p_t(x,y)=\frac{\sqrt{2}}{(4\pi t)^{3/2}}e^{-\frac14 t}\int_r^\infty\frac{se^{-\frac{s^2}{4t}}ds}{(\cosh s-\cosh r)^{\frac12}}. \tag{4.36}$$

Of course, once the formula is known, one can prove it by checking that it is a regular fundamental solution (which, however, is quite involved) and that $p_t(x,y)\to0$ as $x\to\infty$.

We will give here a non-computational proof of (4.33), which to some extend also explains why the heat kernel has this shape.

Proof. By Theorem 4.12, it suffices to prove (4.33) in the case $y=o$ where o is the origin in $\mathbb{H}^3$. Let (r,θ) be the polar coordinates in $\mathbb{H}^3\setminus\{o\}$. As we know, $\mathbb{H}^3$ can be considered as a model manifold bases on $\mathbb{R}^3$ (see Sections 1.12 and 4.7), and the area function of $\mathbb{H}^3$ is given by

$$S(r)=4\pi\sinh^2 r.$$

Recall also that the Laplacian in the polar coordinates has the following expression:

$$\Delta_{\mathbb{H}^3}=\frac{\partial^2}{\partial r^2}+2\coth r\frac{\partial}{\partial r}+\frac{1}{\sinh^2 r}\Delta_{\mathbb{S}^2}. \tag{4.37}$$

Denote by μ the Riemannian measure of $\mathbb{H}^3$.

For a smooth positive function h on $\mathbb{H}^3$, depending only on r, consider the weighted model $(\mathbb{H}^3,\widetilde{\mu})$ where $d\widetilde{\mu}=h^2d\mu$. The area function of $(\mathbb{H}^3,\widetilde{\mu})$ is given by

$$\widetilde{S}(r)=h^2(r)\,S(r).$$

Choose function h as follows:

$$h(r)=\frac{r}{\sinh r},$$

so that

$$\widetilde{S}(r)=4\pi r^2,$$

that is, $\widetilde{S}(r)$ is equal to the area function of $\mathbb{R}^3$. By a miraculous coincidence, the function h happens to satisfy in $\mathbb{H}^3\setminus\{o\}$ the equation

$$\Delta h+h=0, \tag{4.38}$$

which follows from (4.37) by a straightforward computation. The function h extends by continuity to the origin o by setting $h(o)=1$. In fact, the extended function is smooth in $\mathbb{H}^3$ and satisfies (4.38) in the entire $\mathbb{H}^3$ (Exercise 52).

Denoting by $\widetilde{p}_t$ the heat kernel of $(\mathbb{H}^3, \widetilde{\mu})$, we obtain by Theorem 4.14 that

$$\widetilde{p}_t(x,y) = \frac{e^t p_t(x,y)}{h(x)\,h(y)}. \tag{4.39}$$

Since the area functions of the weighted models $(\mathbb{H}^3, \widetilde{\mu})$ and $\mathbb{R}^3$ are the same, we conclude by Theorem 4.13 that their heat kernels at the origin are the same, that is

$$\widetilde{p}_t(x,o) = \frac{1}{(4\pi t)^{3/2}} \exp\left(-\frac{r^2}{4t}\right).$$

Combining with (4.39), we obtain

$$p_t(x,o) = e^{-t}\widetilde{p}_t(x,o)\,h(x)\,h(o) = \frac{1}{(4\pi t)^{3/2}} \frac{r}{\sinh r} \exp\left(-\frac{r^2}{4t} - t\right),$$

which was to be proved. ■

4.10 Heat kernel on $\mathbb{S}^1$

In this section $p_t(x,y)$ is the heat kernel of the Laplace operator on the circle $\mathbb{S}^1$. We identify $\mathbb{S}^1$ with the quotient $\mathbb{R}/2\pi\mathbb{Z}$, that is, consider elements of $\mathbb{S}^1$ as real numbers modulo $2\pi k$ with $k \in \mathbb{Z}$.

Proposition 4.16 *For all $t > 0$ and $x, y \in \mathbb{S}^1$,*

$$p_t(x,y) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{\infty} e^{-k^2 t} \cos k(x-y), \tag{4.40}$$

where the series converges absolutely and uniformly in $(t,x,y) \in [\varepsilon,\infty) \times \Omega \times \Omega$, for any $\varepsilon > 0$.

Proof. By Theorem 3.17, the heat kernel of a compact manifold M (or a precompact open subset of any manifold) is given by the eigenfunction expansion

$$p_t(x,y) = \sum_{k=1}^{\infty} e^{-\lambda_k t} v_k(x)\, v_k(y), \tag{4.41}$$

where $\{v_k\}$ is an orthonormal basis in $L^2(M)$ that consists of eigenfunctions of Δ, and $\{\lambda_k\}$ are their eigenvalues, and the convergence is absolute and uniform in $(t,x,y) \in [\varepsilon,\infty) \times \Omega \times \Omega$, for any $\varepsilon > 0$.

By Exercise 50, the eigenvalues of Δ on $\mathbb{S}^1$ are given by the sequence $\{m^2\}_{m=0}^{\infty}$ where the eigenvalue 0 has the eigenfunction const and the eigenvalue m^2 with $m \geq 1$ has two independent eigenfunctions $\cos m\theta$ and $\sin m\theta$. Since

$$\int_{\mathbb{S}^1} d\theta = \int_0^{2\pi} d\theta = 2\pi$$

and

$$\int_{\mathbb{S}^1} \cos^2 m\theta d\theta = \int_0^{2\pi} \cos^2 m\theta d\theta = \pi, \qquad \int_{\mathbb{S}^1} \sin^2 m\theta d\theta = \int_0^{2\pi} \sin^2 m\theta d\theta = \pi,$$

we obtain the following orthonormal basis in $L^2(\mathbb{S}^1)$ that consists of the eigenfunctions of Δ:

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \ldots, \frac{\cos mx}{\sqrt{\pi}}, \frac{\sin mx}{\sqrt{\pi}}, \ldots$$

By (4.41) we obtain

$$\begin{aligned} p_t(x,y) &= \frac{1}{2\pi} + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2 t} \cos mx \cos my + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2 t} \sin mx \sin my \\ &= \frac{1}{2\pi} + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2 t} \cos m(x-y), \end{aligned}$$

which was to be proved. ■

Proposition 4.17 *Let* $q_t(x,y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{4t}\right)$ *be the heat kernel in* $\mathbb{R}^1$. *Then the heat kernel* $p_t(x,y)$ *of* $\mathbb{S}^1$ *is given by*

$$p_t(x,y) = \sum_{n\in\mathbb{Z}} q_t(x+2\pi n, y). \tag{4.42}$$

Proof. Set

$$\widetilde{q}_t(x,y) = \sum_{n\in\mathbb{Z}} q_t(x+2\pi n, y)$$

and observe that the series converges in any reasonable sense because $q_t(x,y)$ decays quickly in $|x-y|$. Using the fact that $q_t(x,y)$ satisfies the heat equation in t, x for any fixed y, it is easy to show that so does $\widetilde{q}_t(x,y)$.

Next, we obtain

$$\begin{aligned} \int_{\mathbb{S}^1} \widetilde{q}_t(x,y)\,dx &= \sum_{n\in Z}\int_0^{2\pi} q_t(x+2\pi n, y)\,dx = \\ &= \sum_{n\in Z}\int_{2\pi n}^{2\pi(n+1)} q_t(z,y)\,dz = \int_{-\infty}^{\infty} q_t(z,y)\,dz = 1 \end{aligned}$$

that is,

$$\int_{\mathbb{S}^1} \widetilde{q}_t(x,y)\,dx = 1. \tag{4.43}$$

Also, we have

$$\int_{y-\varepsilon}^{y+\varepsilon} \widetilde{q}_t(x,y)\,dx \geq \int_{y-\varepsilon}^{y+\varepsilon} q_t(x,y)\,dx \to 1 \text{ as } \varepsilon \to 0.$$

Hence, $\widetilde{q}_t(x,y)$ is a regular fundamental solution to the heat equation on $\mathbb{S}^1$. By Theorem 4.10, we obtain

$$\widetilde{q}_t(x,y) \geq p_t(x,y).$$

It follows from (4.40) that

$$\int_{\mathbb{S}^1} p_t(x,y)\,dx = 1,$$

which together with (4.43) implies the identity $\widetilde{q}_t(x,y) = p_t(x,y)$. ■

Corollary 4.18 *(*The Poisson summation formula*) For all $t>0$, we have the following identity*

$$\sum_{k\in\mathbb{Z}} e^{-k^2t} = \sqrt{\frac{\pi}{t}} \sum_{n\in\mathbb{Z}} \exp\left(-\frac{\pi^2 n^2}{t}\right). \tag{4.44}$$

Proof. Rewrite (4.40) as follows

$$p_t(x,y) = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} e^{-k^2t} \cos k(x-y). \tag{4.45}$$

In particular, for $x=y=0$ we obtain

$$p_t(0,0) = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} e^{-k^2t}. \tag{4.46}$$

From (4.42) at $x=y=0$, we obtain

$$p_t(0,0) = \sum_{n\in\mathbb{Z}} \frac{1}{(4\pi t)^{1/2}} \exp\left(-\frac{\pi^2 n^2}{t}\right).$$

Comparing the above two lines, we obtain (4.44). ■

Chapter 5

* Stochastic completeness

Definition. A weighted manifold $(M, \mathbf{g}, \mu)$ is called *stochastically complete* if the heat kernel $p_t(x, y)$ satisfies the identity

$$\int_M p_t(x, y)\, d\mu(y) = 1, \tag{5.1}$$

for all $t > 0$ and $x \in M$.

The condition (5.1) can also be stated as $P_t 1 \equiv 1$. Recall that in general we have $0 \le P_t 1 \le 1$ as it follows from Corollaries 3.11 and 3.13.

If the condition (5.1) fails, that is, $P_t 1 \not\equiv 1$ then the manifold M is called *stochastically incomplete.*

Our purpose here is to provide conditions for the stochastic completeness (or incompleteness) in various terms.

5.1 Uniqueness for the bounded Cauchy problem

Fix $0 < T \le \infty$, set $I = (0, T)$ and consider the Cauchy problem in $I \times M$

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, & \text{in } I \times M, \\ u|_{t=0} = f, \end{cases} \tag{5.2}$$

where f is a given function from $C_b(M)$. The problem (5.2) is understood in the classical sense, that is, $u \in C^\infty(I \times M)$ and $u(t, x) \to f(x)$ locally uniformly in $x \in M$ as $t \to 0$. Here we consider the question of the *uniqueness* of a bounded solution of (5.2).

Theorem 5.1 *Fix $\alpha > 0$ and $T \in (0, \infty]$. For any weighted manifold M, the following conditions are equivalent.*

(a) M is stochastically complete.

(b) The equation $\Delta v = \alpha v$ in M has the only bounded non-negative solution $v = 0$.

(c) The Cauchy problem (5.2) in $(0, T) \times M$ has at most one bounded solution.

Remark. As we will see from the proof, in condition (b) the assumption that v is non-negative can be dropped without violating the statement.

Proof. We first assume $T<\infty$ and prove the following sequence of implications

$$\neg(a) \Longrightarrow \neg(b) \Longrightarrow \neg(c) \Longrightarrow \neg(a),$$

where $\neg$ means the negation of the statement.

Proof of $\neg(a) \Rightarrow \neg(b)$. So, we assume that M is stochastically incomplete and prove that there exists a non-zero bounded solution to the equation $-\Delta v+\alpha v=0$. Consider the function

$$P_t 1(x)=\int_M p_t(x,y)\, d\mu(y),$$

which by Lemma 4.4 is C^∞ smooth, $0\leq P_t 1\leq 1$ and, by the hypothesis of stochastic incompleteness, $P_t 1\not\equiv 1$. Consider also the function

$$w(x)=\int_0^\infty e^{-\alpha t} P_t 1(x)\, dt. \tag{5.3}$$

Let us verify that $w\in C^\infty(M)$, it satisfies the estimate

$$0\leq w\leq \alpha^{-1} \tag{5.4}$$

and the equation

$$-\Delta w+\alpha w=1. \tag{5.5}$$

The inequalities (5.4) follows from $0\leq P_t 1\leq 1$. To prove the other properties, consider an exhaustion $\{\Omega_i\}$ of M and define in Ω_k the function

$$w_i(x)=\int_0^\infty e^{-\alpha t} P_t^{\Omega_i} f(x)\, dt,$$

where $f=1_{\Omega_i}$. Expanding $f=\sum_{k=1}^\infty a_k v_k$ in the basis of eigenfunctions of Δ in Ω_i, we obtain

$$P_t^{\Omega_i} f=\sum_{k=1}^\infty e^{-\lambda_k t} a_k v_k$$

whence

$$w_i=\sum_{k=1}^\infty\left(\int_0^\infty e^{-(\lambda_k+\alpha)t} dt\right) a_k v_k=\sum_{k=1}^\infty \frac{a_k}{\lambda_k+\alpha} v_k.$$

It follows that $w_i\in W_0^1(\Omega)$ and

$$-\Delta w_i=\sum_{k=1}^\infty \frac{\lambda_k a_k}{\lambda_k+\alpha} v_k\in L^2(\Omega_i).$$

Hence,

$$-\Delta w_i+\alpha w_i=\sum_{k=1}^\infty \frac{\lambda_k a_k}{\lambda_k+\alpha} v_k+\sum_{k=1}^\infty \frac{\alpha a_k}{\lambda_k+\alpha} v_k=\sum_{k=1}^\infty a_k v_k=f=1.$$

Similarly to Corollary 2.8, we conclude that $w_i \in C^\infty(\Omega_i)$. Since $w_i \nearrow w$ as $i \to \infty$, we obtain by (an extension of) Proposition 4.1 that w is C^∞ smooth and satisfies (5.5).

It follows from $P_t 1(x) \not\equiv 1$ that there exist $x \in M$ and $t > 0$ such that $P_t 1(x) < 1$. Then (5.3) implies that, for this value of x, we have a *strict* inequality $w(x) < \alpha^{-1}$. Hence, $w \not\equiv \alpha^{-1}$.

Finally, consider the function $v = 1 - \alpha w$, which by (5.5) satisfies the equation $\Delta v = \alpha v$. It follows from (5.4) that $0 \le v \le 1$, and $w \not\equiv \alpha^{-1}$ implies $v \not\equiv 0$. Hence, we have constructed a non-zero non-negative bounded solution to $\Delta v = \alpha v$, which finishes the proof.

Proof of $\neg(b) \Rightarrow \neg(c)$. Let v be a bounded non-zero solution to equation $\Delta v = \alpha v$. By Corollary 2.8, $v \in C^\infty(M)$. Then the function

$$u(t,x) = e^{\alpha t} v(x) \tag{5.6}$$

satisfies the heat equation because

$$\Delta u = e^{\alpha t} \Delta v = \alpha e^{\alpha t} v = \partial_t u.$$

Hence, u solves the Cauchy problem in $\mathbb{R}_+ \times M$ with the initial condition $u(0,x) = v(x)$, and this solution u is bounded on $(0,T) \times M$ (note that T is *finite*). Let us compare $u(t,x)$ with the function $P_t v(x)$. Since $v \in C_b(M)$, the function $P_t v(x)$ solves the heat equation and satisfies the initial condition with the function v in the classical sense (cf. Lemma 4.9). It follows from Corollary 3.13 that

$$\sup |P_t v| \le \sup |v|,$$

whereas by (5.6)

$$\sup |u(t,\cdot)| = e^{\alpha t} \sup |v| > \sup |v|.$$

Therefore, $u \not\equiv P_t v$, and the bounded Cauchy problem in $(0,T) \times M$ has two different solutions with the same initial function v.

Proof of $\neg(c) \Rightarrow \neg(a)$. Assume that the problem (5.2) has two different bounded solutions with the same initial function. Subtracting these solutions, we obtain a non-zero bounded solution $u(t,x)$ to (5.2) with the initial function $f = 0$. Without loss of generality, we can assume that $0 < \sup u \le 1$. Consider the function $w = 1 - u$, for which we have $0 \le \inf w < 1$. The function w is a non-negative solution to the Cauchy problem (5.2) with the initial function $f = 1$. By Lemma 4.5, we conclude that $w(t,\cdot) \ge P_t 1$. Hence, $\inf P_t 1 < 1$ and M is stochastically incomplete.

Finally, let us prove the equivalence of (a), (b), (c) in the case $T = \infty$. Since the condition (c) with $T = \infty$ is weaker than that for $T < \infty$, it suffices to show that (c) with $T = \infty$ implies (a). Assume from the contrary that M is stochastically incomplete, that is, $P_t 1 \not\equiv 1$. Then the functions $u_1 \equiv 1$ and $u_2 = P_t 1$ are two different bounded solutions to the Cauchy problem (5.2) in $\mathbb{R}_+ \times M$ with the same initial function $f \equiv 1$, so that (a) fails, which was to be proved. ■

5.2 Geodesic completeness

Let $(M, \mathbf{g})$ be a Riemannian manifold and $d(x,y)$ be the geodesic distance on M (see Section 1.13 for the definition). The manifold $(M, \mathbf{g})$ is said to be *metrically complete* if the metric space (M,d) is complete, that is, any Cauchy sequence in (M,d) converges.

A smooth path $\gamma(t) : (a, b) \to M$ is called a *geodesics* if, for any $t \in (a, b)$ and for all s close enough to t, the path $\gamma|_{[t,s]}$ is a shortest path between the points $\gamma(t)$ and $\gamma(s)$. A Riemannian manifold $(M, \mathbf{g})$ is called *geodesically complete* if, for any $x \in M$ and $\xi \in T_x M \setminus \{0\}$, there is a geodesics $\gamma : [0, +\infty) \to M$ of infinite length such that $\gamma(0) = x$ and $\dot{\gamma}(0) = \xi$. It is known that, on a geodesically complete connected manifold, any two points can be connected by a shortest geodesics.

We state the following theorem without proof.

Hopf-Rinow Theorem. *For a Riemannian manifold $(M, \mathbf{g})$, the following conditions are equivalent:*

(a) $(M, \mathbf{g})$ *is metrically complete.*

(b) $(M, \mathbf{g})$ *is geodesically complete.*

(c) *All geodesic balls in M are relatively compact sets.*

This theorem will not be used, but it motivates us to give the following definition.

Definition. A Riemannian manifold $(M, \mathbf{g})$ is said to be *complete* if all the geodesic balls in M are relatively compact.

For example, any compact manifold is complete.

5.3 Stochastic completeness and the volume growth

Define the *volume function* $V(x, r)$ of a weighted manifold $(M, \mathbf{g}, \mu)$ by

$$V(x, r) := \mu(B(x, r)),$$

where $B(x, r)$ is the geodesic ball. Note that $V(x, r) < \infty$ for all $x \in M$ and $r > 0$ provided M is complete.

Recall that a manifold M is *stochastically* complete, if the heat kernel $p_t(x, y)$ satisfies the identity

$$\int_M p_t(x, y) \, d\mu(y) = 1,$$

for all $x \in M$ and $t > 0$ (see Section 5.1). The result of this section is the following volume test for the stochastic completeness.

Theorem 5.2 *Let $(M, \mathbf{g}, \mu)$ be a complete connected weighted manifold. If, for some point $x_0 \in M$,*

$$\int^{\infty} \frac{r dr}{\ln V(x_0, r)} = \infty, \tag{5.7}$$

then M is stochastically complete.

Condition (5.7) holds, in particular, if

$$V(x_0, r) \le \exp\left(C r^2\right). \tag{5.8}$$

As a consequence we see that both $\mathbb{R}^n$ and $\mathbb{H}^n$ are stochastically complete.

Fix $0 < T \leq \infty$, set $I = (0, T)$ and consider the following Cauchy problem in $I \times M$

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta_\mu u & \text{in } I \times M, \\ u|_{t=0} = 0. \end{cases} \tag{5.9}$$

A solution is sought in the class $u \in C^\infty(I \times M)$, and the initial condition means that $u(t,x) \to 0$ locally uniformly in $x \in M$ as $t \to 0$ (cf. Section 5.1). By Theorem 5.1, the stochastic completeness of M is equivalent to the uniqueness property of the Cauchy problem in the class of bounded solutions. In other words, in order to prove Theorem 5.2, it suffices to verify that the only bounded solution to (5.9) is $u \equiv 0$.

The assertion will follow from the following more general fact.

Theorem 5.3 *Let $(M, \mathbf{g}, \mu)$ be a complete connected weighted manifold, and let $u(x,t)$ be a solution to the Cauchy problem (5.9). Assume that, for some $x_0 \in M$ and for all $R > 0$,*

$$\int_0^T \int_{B(x_0,R)} u^2(x,t)\, d\mu(x) dt \leq \exp\left(f(R)\right), \tag{5.10}$$

where $f(r)$ is a positive increasing function on $(0, +\infty)$ such that

$$\int^\infty \frac{r dr}{f(r)} = \infty. \tag{5.11}$$

Then $u \equiv 0$ in $I \times M$.

Theorem 5.3 provides the uniqueness class (5.10) for the Cauchy problem. The condition (5.11) holds if, for example, $f(r) = Cr^2$, but fails for $f(r) = Cr^{2+\varepsilon}$ when $\varepsilon > 0$.

Before we embark on the proof, let us mention the following consequence.

Corollary 5.4 *If $M = \mathbb{R}^n$ and $u(t,x)$ be a solution to (5.9) satisfying the condition*

$$|u(t,x)| \leq C \exp\left(C|x|^2\right) \quad \textit{for all } t \in I,\ x \in \mathbb{R}^n, \tag{5.12}$$

then $u \equiv 0$. Moreover, the same is true if u satisfies instead of (5.12) the condition

$$|u(t,x)| \leq C \exp\left(f(|x|)\right) \quad \textit{for all } t \in I,\ x \in \mathbb{R}^n, \tag{5.13}$$

where $f(r)$ is a convex increasing function on $(0, +\infty)$ satisfying (5.11).

Proof. Since (5.12) is a particular case of (5.13) for the function $f(r) = Cr^2$, it suffices to treat the condition (5.13). In $\mathbb{R}^n$ we have $V(x,r) = cr^n$. Therefore, (5.13) implies that

$$\int_0^T \int_{B(0,R)} u^2(x,t)\, d\mu(x) dt \leq CR^n \exp\left(f(R)\right) = C\exp(\widetilde{f}(R)),$$

where $\widetilde{f}(r) := f(r) + n \ln r$. The convexity of f implies that $\ln r \leq Cf(r)$ for large enough r. Hence, $\widetilde{f}(r) \leq Cf(r)$ and function $\widetilde{f}$ also satisfies the condition (5.11). By Theorem 5.3, we conclude $u \equiv 0$. ■

The class of functions u satisfying (5.12) is called the *Tikhonov class*, and the conditions (5.13) and (5.11) define the *Täcklind class*. The uniqueness of the Cauchy problem in $\mathbb{R}^n$ in each of these classes are classical results.

Proof of Theorem 5.2. By Theorem 5.1, it suffices to verify that the only bounded solution to the Cauchy value problem (5.9) is $u \equiv 0$. Indeed, if u is a bounded solution of (5.9), then setting

$$S := \sup |u| < \infty$$

we obtain

$$\int_0^T \int_{B(x_0,R)} u^2(t,x)d\mu(x) \leq S^2 TV(x_0,R) = \exp\left(f\left(R\right)\right),$$

where

$$f(r) := \ln\left(S^2 TV(x_0,r)\right).$$

It follows from the hypothesis (5.7) that the function f satisfies (5.11). Hence, by Theorem 5.3, we obtain $u \equiv 0$. ∎

Proof of Theorem 5.3. Denote for simplicity $B_r = B(x_0,r)$. The main technical part of the proof is the following claim.

Claim. *Let $u\,(t,x)$ solve the heat equation in $(b,a)\times M$ where $b<a$ are reals, and assume that $u\,(t,x)$ extends to a continuous function in $[b,a]\times M$. Assume also that, for all $R>0$,*

$$\int_a^b \int_{B_R} u^2(x,t)\, d\mu(x)dt \leq \exp\left(f(R)\right),$$

where f is a function as in Theorem 5.2*. Then, for any $R>0$ satisfying the condition*

$$a-b \leq \frac{R^2}{8f(4R)}, \tag{5.14}$$

the following inequality holds:

$$\int_{B_R} u^2(a,\cdot)d\mu \leq \int_{B_{4R}} u^2(b,\cdot)d\mu + \frac{4}{R^2}. \tag{5.15}$$

Let us first show how this Claim allows to prove that any solution u to (5.9), satisfying (5.10), is identical 0. Extend $u\,(t,x)$ to $t=0$ by setting $u\,(0,x)=0$ so that u is continuous in $[0,T)\times M$. Fix $R>0$ and $t\in(0,T)$. For any non-negative integer k, set

$$R_k = 4^k R$$

and, for any $k \geq 1$, choose (so far arbitrarily) a number τ_k to satisfy the condition

$$0 < \tau_k \leq c\frac{R_k^2}{f(R_k)}, \tag{5.16}$$

where $c=\frac{1}{128}$. Then define a decreasing sequence of times $\{t_k\}$ inductively by $t_0=t$ and $t_k = t_{k-1}-\tau_k$ (see Fig. 5.1).

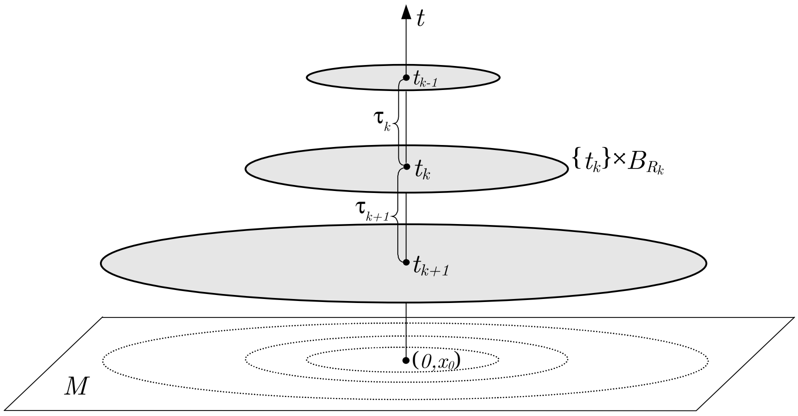

Figure 5.1: The sequence of the balls B_{R_k} and the time moments t_k.

If $t_k \geq 0$ then function u satisfies all the conditions of the Claim with $a = t_{k-1}$ and $b = t_k$, and we obtain from (5.15)

$$\int_{B_{R_{k-1}}} u^2(t_{k-1}, \cdot)d\mu \leq \int_{B_{R_k}} u^2(t_k, \cdot)d\mu + \frac{4}{R_{k-1}^2}, \tag{5.17}$$

which implies by induction that

$$\int_{B_R} u^2(t, \cdot)d\mu \leq \int_{B_{R_k}} u^2(t_k, \cdot)d\mu + \sum_{i=1}^{k} \frac{4}{R_{i-1}^2}. \tag{5.18}$$

If it happens that $t_k = 0$ for some k then, by the initial condition in (5.9),

$$\int_{B_{R_k}} u^2(t_k, \cdot)d\mu = 0.$$

In this case, it follows from (5.18) that

$$\int_{B_R} u^2(t, \cdot)d\mu \leq \sum_{i=1}^{\infty} \frac{4}{R_{i-1}^2} = \frac{C}{R^2},$$

which implies by letting $R \to \infty$ that $u(\cdot, t) \equiv 0$ (here we use the connectedness of M).

Hence, to finish the proof, it suffices to construct, for any $R > 0$ and $t \in (0, T)$, a sequence $\{t_k\}$ as above that vanishes at a finite k. The condition $t_k = 0$ is equivalent to

$$t = \tau_1 + \tau_2 + ... + \tau_k\,. \tag{5.19}$$

The only restriction on τ_k is the inequality (5.16). The hypothesis that $f(r)$ is an increasing function implies that

$$\int_R^\infty \frac{rdr}{f(r)} \leq \sum_{k=0}^{\infty} \int_{R_k}^{R_{k+1}} \frac{rdr}{f(r)} \leq \sum_{k=0}^{\infty} \frac{R_{k+1}^2,}{f(R_k)}$$

which together with (5.11) yields

$$\sum_{k=1}^{\infty}\frac{R_k^2}{f(R_k)}=\infty.$$

Therefore, the sequence $\{\tau_k\}_{k=1}^{\infty}$ can be chosen to satisfy simultaneously (5.16) and

$$\sum_{k=1}^{\infty}\tau_k=\infty.$$

By diminishing some of τ_k, we can achieve (5.19) for any finite t, which finishes the proof.

Now we prove the above Claim. Since the both integrals in (5.15) are continuous with respect to a and b, we can slightly reduce a and slightly increase b; hence, we can assume that $u\,(t,x)$ is not only continuous in $[b,a]\times M$ but also smooth.

Let $\rho(x)$ be a Lipschitz function on M (to be specified below) with the Lipschitz constant 1. Fix a real $s\notin[b,a]$ (also to be specified below) and consider the following the function

$$\xi(t,x):=\frac{\rho^2(x)}{4(t-s)},$$

which is defined on $\mathbb{R}\times M$ except for $t=s$, in particular, on $[b,a]\times M$. By the weak gradient $\nabla\rho$ is in $L^{\infty}(M)$ and satisfies the inequality $|\nabla\rho|\le 1$, which implies, for any $t\neq s$,

$$|\nabla\xi\,(t,x)|\le\frac{\rho\,(x)}{2\,(t-s)}.$$

Since

$$\frac{\partial\xi}{\partial t}=-\frac{\rho^2\,(x)}{4\,(t-s)^2},$$

we obtain

$$\frac{\partial\xi}{\partial t}+|\nabla\xi|^2\le 0. \tag{5.20}$$

For a given $R>0$, define a function $\varphi\,(x)$ by

$$\varphi\,(x)=\min\left(\left(3-\tfrac{d(x,x_0)}{R}\right)_+,1\right)$$

(see Fig. 5.2). Obviously, we have $0\le\varphi\le1$ on M, $\varphi\equiv1$ in B_{2R}, and $\varphi\equiv0$ outside B_{3R}. Since the function $d\,(\cdot,x_0)$ is Lipschitz with the Lipschitz constant 1, we obtain that φ is Lipschitz with the Lipschitz constant $1/R$. Then we have $|\nabla\varphi|\le1/R$. By the completeness of M, all the balls in M are relatively compact sets, which implies $\varphi\in Lip_0\,(M)$.

Consider the function $u\varphi^2e^{\xi}$ as a function of x for any fixed $t\in[b,a]$. Since it is obtained from locally Lipschitz functions by taking product and composition, this function is locally Lipschitz on M. Since this function has a compact support, it belongs to $Lip_0\,(M)$, whence

$$u\varphi^2e^{\xi}\in W_c^1\,(M)\,.$$

Multiplying the heat equation

$$\frac{\partial u}{\partial t}=\Delta_\mu u$$

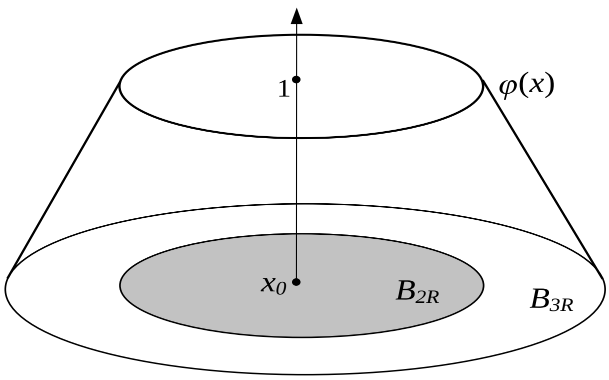

Figure 5.2: Function $\varphi(x)$

by $u\varphi^2 e^\xi$ and integrating it over $[b,a]\times M$, we obtain

$$\int_b^a\int_M \frac{\partial u}{\partial t} u\varphi^2 e^\xi d\mu dt = \int_b^a\int_M (\Delta_\mu u)\, u\varphi^2 e^\xi d\mu dt. \tag{5.21}$$

Since both functions u and ξ are smooth in $t\in[b,a]$, the time integral on the left hand side can be computed as follows:

$$\frac{1}{2}\int_b^a \frac{\partial(u^2)}{\partial t}\varphi^2 e^\xi dt = \frac{1}{2}\left[u^2\varphi^2 e^\xi\right]_b^a - \frac{1}{2}\int_b^a \frac{\partial \xi}{\partial t} u^2\varphi^2 e^\xi dt. \tag{5.22}$$

Using the Green formula to evaluate the spatial integral on the right hand side of (5.21), we obtain

$$\int_M (\Delta_\mu u)\, u\varphi^2 e^\xi d\mu = -\int_M \langle \nabla u, \nabla(u\varphi^2 e^\xi)\rangle d\mu.$$

Applying the product rule and the chain rule to compute $\nabla(u\varphi^2 e^\xi)$, we obtain

$$\begin{aligned}
-\langle \nabla u, \nabla(u\varphi^2 e^\xi)\rangle &= -|\nabla u|^2\varphi^2 e^\xi - \langle \nabla u, \nabla\xi\rangle u\varphi^2 e^\xi - 2\langle \nabla u, \nabla\varphi\rangle u\varphi e^\xi \\
&\le -|\nabla u|^2\varphi^2 e^\xi + |\nabla u|\,|\nabla \xi|\,|u|\,\varphi^2 e^\xi \\
&\quad + \left(\frac{1}{2}|\nabla u|^2\varphi^2 + 2|\nabla\varphi|^2 u^2\right) e^\xi \\
&= \left(-\frac{1}{2}|\nabla u|^2 + |\nabla u|\,|\nabla\xi|\,|u|\right)\varphi^2 e^\xi + 2|\nabla\varphi|^2 u^2 e^\xi.
\end{aligned}$$

Combining with (5.21), (5.22), and using (5.20), we obtain

$$\begin{aligned}\left[\int_M u^2\varphi^2 e^\xi d\mu\right]_b^a &= \int_b^a\int_M \frac{\partial\xi}{\partial t}u^2\varphi^2 e^\xi\, d\mu dt + 2\int_b^a\int_M (\Delta_\mu u)\, u\varphi^2 e^\xi d\mu dt \\ &\le \int_b^a\int_M \left(-|\nabla\xi|^2 u^2 - |\nabla u|^2 + 2|\nabla u||\nabla\xi||u|\right)\varphi^2 e^\xi d\mu dt \\ &\quad +4\int_b^a\int_M |\nabla\varphi|^2 u^2 e^\xi d\mu dt \\ &= -\int_b^a\int_M \left(|\nabla\xi||u| - |\nabla u|\right)^2\varphi^2 e^\xi d\mu dt \\ &\quad +4\int_b^a\int_M |\nabla\varphi|^2 u^2 e^\xi d\mu dt\end{aligned}$$

whence

$$\left[\int_M u^2\varphi^2 e^\xi d\mu\right]_b^a \le 4\int_b^a\int_M |\nabla\varphi|^2 u^2 e^\xi d\mu dt. \tag{5.23}$$

Using the properties of function $\varphi(x)$, in particular, $|\nabla\varphi| \le 1/R$, we obtain from (5.23)

$$\int_{B_R} u^2(a,\cdot)e^{\xi(a,\cdot)}d\mu \le \int_{B_{4R}} u^2(b,\cdot)e^{\xi(b,\cdot)}d\mu + \frac{4}{R^2}\int_b^a\int_{B_{4R}\setminus B_{2R}} u^2 e^\xi d\mu dt. \tag{5.24}$$

Let us now specify $\rho(x)$ and s. Set $\rho(x)$ to be the distance function from the ball B_R, that is,

$$\rho(x) = (d(x,x_0) - R)_+$$

(see Fig. 5.3).

Set $s = 2a - b$ so that, for all $t \in [b,a]$,

$$a - b \le s - t \le 2(a-b),$$

whence

$$\xi(t,x) = -\frac{\rho^2(x)}{4(s-t)} \le -\frac{\rho^2(x)}{8\,(a-b)} \le 0. \tag{5.25}$$

Consequently, we can drop the factor e^ξ on the left hand side of (5.24) because $\xi = 0$ in B_R, and drop the factor e^ξ in the first integral on the right hand side of (5.24) because $\xi \le 0$. Clearly, if $x \in B_{4R}\setminus B_{2R}$ then $\rho(x) \ge R$, which together with (5.25) implies that

$$\xi(t,x) \le -\frac{R^2}{8\,(a-b)} \quad \text{in} \quad [b,a]\times B_{4R}\setminus B_{2R}.$$

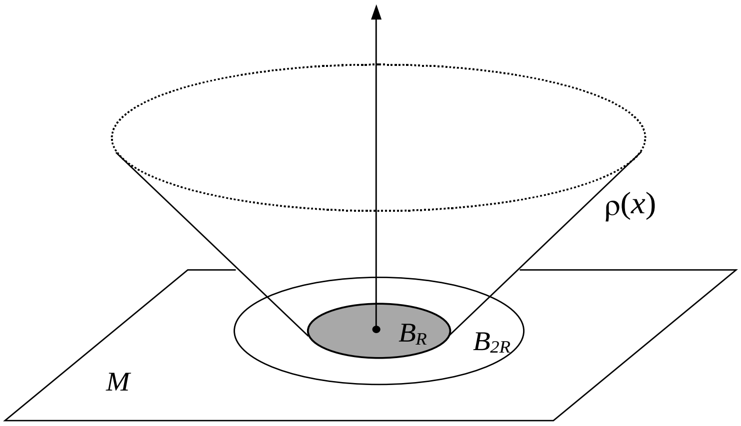

Figure 5.3: Function $\rho(x)$.

Hence, we obtain from (5.24)

$$\int_{B_R} u^2(a,\cdot)d\mu \le \int_{B_{4R}} u^2(b,\cdot)d\mu + \frac{4}{R^2}\exp\left(-\frac{R^2}{8\,(a-b)}\right)\int_b^a\int_{B_{4R}} u^2 d\mu dt.$$

By (5.10) we have

$$\int_b^a\int_{B_{4R}} u^2 d\mu dt \le \exp\left(f(4R)\right)$$

whence

$$\int_{B_R} u^2(a,\cdot)d\mu \le \int_{B_{4R}} u^2(b,\cdot)d\mu + \frac{4}{R^2}\exp\left(-\frac{R^2}{8\,(a-b)} + f(4R)\right).$$

Finally, applying the hypothesis (5.14), we obtain (5.15). ■

Chapter 6

* Gaussian estimates in the integrated form

As one can see from explicit examples of heat kernels (4.18), (4.32), (4.33), the dependence of the heat kernel $p_t(x,y)$ on the points x,y is frequently given by the term $\exp\left(-c\frac{d^2(x,y)}{t}\right)$ that is called the *Gaussian factor*. The Gaussian pointwise upper bounds of the heat kernel require certain additional assumptions about the manifold in question.

On the contrary, it is relatively straightforward to obtain the integrated upper bounds of the heat kernel, which is the main topic of this Chapter

6.1 The integrated maximum principle

Recall that any function $f \in Lip_{loc}(M)$ has the weak gradient $\nabla f \in \vec{L}^{\infty}_{loc}(M)$.

Theorem 6.1 (The integrated maximum principle) *Let $\xi(t,x)$ be a continuous function on $I \times M$, where $I \subset [0,+\infty)$ is an interval. Assume that, for any $t \in I$, $\xi(t,x)$ is locally Lipschitz in $x \in M$, the partial derivative $\frac{\partial \xi}{\partial t}$ exists and is continuous in $I \times M$, and the following inequality holds on $I \times M$:*

$$\frac{\partial \xi}{\partial t} + \frac{1}{2}|\nabla \xi|^2 \leq 0. \tag{6.1}$$

Then, for any function $f \in L^2(M)$, the function

$$J(t) := \int_M (P_t f)^2(x)\, e^{\xi(t,x)} d\mu(x) \tag{6.2}$$

is non-increasing in $t \in I$. Furthermore, for all $t, t_0 \in I$, if $t > t_0$ then

$$J(t) \leq J(t_0)\, e^{-2\lambda_1(M)(t-t_0)}. \tag{6.3}$$

Remark. Let $d(x)$ be a Lipschitz function on M with the Lipschitz constant 1. Then we have $|\nabla d| \leq 1$. It follows that the following functions satisfy (6.1):

$$\xi(t,x) = \frac{d^2(x)}{2t}$$

and

$$\xi(t,x)=ad(x)-\frac{a^2}{2}t,$$

where a is a real constant. In applications $d(x)$ is normally chosen to be the distance from x to some set.

Proof. Let us first reduce the problem to the case of non-negative f. Indeed, if f is signed then set $g=|P_{t_0}f|$ and notice that

$$|P_t f|=|P_{t-t_0}P_{t_0}f|\le P_{t-t_0}g.$$

Assuming that Theorem 6.1 has been already proved for function g, we obtain

$$\begin{aligned}\int_M (P_t f)^2 e^{\xi(t,\cdot)}d\mu &\le \int_M (P_{t-t_0}g)^2 e^{\xi(t,\cdot)}d\mu\\ &\le e^{-2\lambda_1(t-t_0)}\int_M g^2 e^{\xi(t_0,\cdot)}d\mu\\ &= e^{-2\lambda_1(t-t_0)}\int_M (P_{t_0}f)^2 e^{\xi(t_0,\cdot)}d\mu.\end{aligned}$$

Hence, we can assume in the sequel that $f\ge 0$. It suffices to prove that, for any relatively compact open set $\Omega\subset M$, the function

$$J_\Omega(t):=\int_\Omega \left(P_t^\Omega f\right)^2(x)\,e^{\xi(t,x)}d\mu(x)$$

is non-increasing in $t\in I$. Since $u(t,\cdot):=P_t^\Omega f\in L^2(\Omega)$ and $\xi(t,\cdot)$ is bounded in Ω, the function $J_\Omega(t)$ is finite (unlike $J(t)$ that a priori may be equal to ∞). Note also that $J_\Omega(t)$ is continuous in $t\in I$. Indeed, the path $t\mapsto u(t,\cdot)$ is continuous in $t\in[0,+\infty)$ in $L^2(\Omega)$ and the path $t\mapsto e^{\frac{1}{2}\xi(t,\cdot)}$ is obviously continuous in $t\in I$ in the sup-norm in $C_b(\Omega)$, which implies that the path $t\mapsto u(t,\cdot)\,e^{\frac{1}{2}\xi(t,\cdot)}$ is continuous in $t\in I$ in $L^2(\Omega)$.

To prove that $J_\Omega(t)$ is non-increasing in I it suffices to show that the derivative $\frac{dJ_\Omega}{dt}$ exists and is non-positive for all $t\in I_0:=I\setminus\{0\}$. Fix some $t\in I_0$. Since the functions $\xi(t,\cdot)$ and $\frac{\partial\xi}{\partial t}(t,\cdot)$ are continuous and bounded in $\overline{\Omega}$, they both belong to $C_b(\Omega)$. Therefore, the partial derivative $\frac{\partial\xi}{\partial t}$ is at the same time the derivative $\frac{d\xi}{dt}$ in $C_b(\Omega)$. In the same way, the function $e^{\xi(t,\cdot)}$ is differentiable in $C_b(\Omega)$ and

$$\frac{de^\xi}{dt}=\frac{\partial e^\xi}{\partial t}=e^\xi\frac{\partial\xi}{\partial t}. \tag{6.4}$$

The function $u(t,\cdot)$ is $L^2(\Omega)$-differentiable and its L^2 derivative $\frac{du}{dt}$ is given by

$$\frac{du}{dt}=\Delta u. \tag{6.5}$$

Using the product rules for L^2 derivatives, we conclude that ue^ξ is differentiable in $L^2(\Omega)$ and

$$\frac{d}{dt}\left(ue^\xi\right)=\frac{du}{dt}e^\xi+u\frac{de^\xi}{dt}. \tag{6.6}$$

It follows that the inner product $\left(u, ue^{\xi}\right) = J_{\Omega}(t)$ is differentiable as a real valued function of t and, by the product rule and by (6.4), (6.5), (6.6),

$$\begin{aligned}
\frac{dJ_{\Omega}}{dt} &= \left(\frac{du}{dt}, ue^{\xi}\right) + \left(u, \frac{d\left(ue^{\xi}\right)}{dt}\right) \\
&= 2\left(\frac{du}{dt}, ue^{\xi}\right) + \left(u^2, \frac{de^{\xi}}{dt}\right) \\
&= 2\left(\Delta u, ue^{\xi}\right) + \left(u^2, \frac{\partial \xi}{\partial t}e^{\xi}\right). \qquad (6.7)
\end{aligned}$$

By the chain rule for Lipschitz functions, we have $e^{\xi(t,\cdot)} \in Lip_{loc}(M)$. Since the function $e^{\xi(t,\cdot)}$ is bounded and Lipschitz in Ω and $u(t,\cdot) \in W_0^1(\Omega)$, we obtain that $ue^{\xi} \in W_0^1(\Omega)$. By the Green formula, we obtain

$$2\left(\Delta u, ue^{\xi}\right) = -2\int_{\Omega} \langle \nabla u, \nabla\left(ue^{\xi}\right)\rangle d\mu.$$

Since both functions u and $e^{\xi(t,\cdot)}$ are locally Lipschitz, the product rule and the chain rule apply for expanding $\nabla\left(ue^{\xi}\right)$. Substituting the result into (6.7) and using (6.1), we obtain

$$\begin{aligned}
\frac{dJ_{\Omega}}{dt} &\leq -2\int_{\Omega}\left(|\nabla u|^2 e^{\xi} + ue^{\xi}\langle \nabla u, \nabla \xi\rangle + \frac{1}{4}u^2 |\nabla \xi|^2 e^{\xi}\right) d\mu \\
&= -2\int_{\Omega}\left(\nabla u + \frac{1}{2}u\nabla\xi\right)^2 e^{\xi} d\mu, \qquad (6.8)
\end{aligned}$$

whence $\frac{dJ_{\Omega}}{dt} \leq 0$. To prove (6.3), observe that

$$\left(\nabla u + \frac{1}{2}u\nabla\xi\right)e^{\xi/2} = \nabla(ue^{\xi/2}).$$

Since $ue^{\xi/2} \in W_0^1(\Omega)$, we can apply the variational principle, which yields

$$\begin{aligned}
\int_{\Omega}\left(\nabla u + \frac{1}{2}u\nabla\xi\right)^2 e^{\xi} d\mu &= \int_{\Omega} |\nabla(ue^{\xi/2})|^2 d\mu \\
&\geq \lambda_1(\Omega)\int_{\Omega} |ue^{\xi/2}|^2 d\mu \\
&= \lambda_1(\Omega) J_{\Omega}(t). \qquad (6.9)
\end{aligned}$$

Hence, (6.8) yields

$$\frac{dJ_{\Omega}}{dt} \leq -2\lambda_1(\Omega) J_{\Omega}(t),$$

whence (6.3) follows. ∎

6.2 The Davies-Gaffney inequality

For any set A on a weighted manifold M and any $r > 0$, denote by A_r the r-neighborhood of A, that is,

$$A_r = \{x \in M : d(x, A) < r\}.$$

Write also $A_r^c = (A_r)^c = M \setminus A_r$.

Theorem 6.2 *Let A be a measurable subset of a weighted manifold M. Then, for any function $f \in L^2(M)$ and for all positive r, t,*

$$\int_{A_r^c} (P_t f)^2 \, d\mu \le \int_{A^c} f^2 d\mu + \exp\left(-\frac{r^2}{2t} - 2\lambda t\right) \int_A f^2 d\mu, \tag{6.10}$$

where $\lambda = \lambda_1(M)$. In particular, if $f \in L^2(A)$ then

$$\int_{A_r^c} (P_t f)^2 \, d\mu \le \|f\|_2^2 \exp\left(-\frac{r^2}{2t} - 2\lambda t\right) \tag{6.11}$$

(see Fig. 6.1).

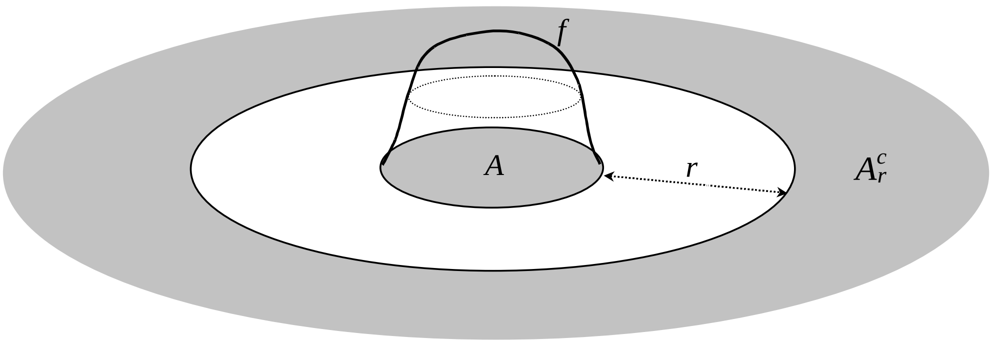

Figure 6.1: Sets A and A_r^c

Proof. Fix some $s > t$ and consider the function

$$\xi(\tau, x) = \frac{d^2(x, A_r^c)}{2(\tau - s)},$$

defined for $x \in M$ and $\tau \in [0, s)$. Set also

$$J(\tau) := \int_M (P_\tau f)^2 \, e^{\xi(\tau,\cdot)} d\mu.$$

Since the function ξ satisfies the condition

$$\frac{\partial \xi}{\partial \tau} + \frac{1}{2} |\nabla \xi|^2 \le 0,$$

we obtain by Theorem 6.1 that

$$J(t) \le J(0)\exp(-2\lambda t). \tag{6.12}$$

Since $\xi(\tau, x) = 0$ for $x \in A_r^c$, we have

$$J(t) \ge \int_{A_r^c} (P_t f)^2 \, d\mu. \tag{6.13}$$

On the other hand, using the fact that $\xi(0, x) \le 0$ for all x and

$$\xi(0, x) \le -\frac{r^2}{2s} \quad \text{for all } x \in A,$$

we obtain

$$J(0) \le \int_{A^c} f^2 d\mu + \exp\left(-\frac{r^2}{2s}\right)\int_A f^2 d\mu. \tag{6.14}$$

Combining together (6.12), (6.13), (6.14) and letting $s \to t+$, we obtain (6.10).

The inequality (6.11) trivially follows from (6.10) and the observation that $\int_{A^c} f^2 d\mu = 0$. ■

Corollary 6.3 (The Davies-Gaffney inequality). *If A and B are two disjoint measurable subsets of M and $f \in L^2(A)$, $g \in L^2(B)$, then, for all $t > 0$,*

$$|(P_t f, g)| \le \|f\|_2 \|g\|_2 \exp\left(-\frac{d^2(A,B)}{4t} - \lambda t\right) \tag{6.15}$$

(see Fig. 6.2*).*

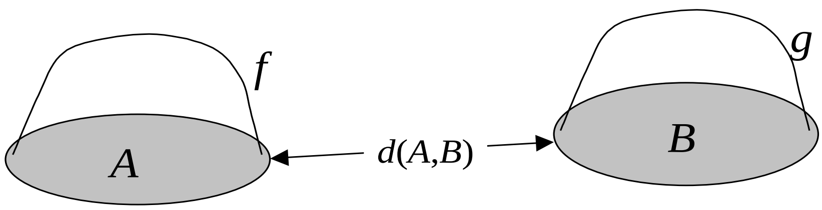

Figure 6.2: Sets A and B

Proof. Set $r = d(A, B)$. Then $B \subset A_r^c$ and by (6.11)

$$\int_B (P_t f)^2 \, d\mu \le \|f\|_2^2 \exp\left(-\frac{r^2}{2t} - 2\lambda t\right).$$

Applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} |(P_t f, g)| &\le \left(\int_B (P_t f)^2 \, d\mu\right)^{1/2} \|g\|_2 \\ &\le \|f\|_2 \|g\|_2 \exp\left(-\frac{r^2}{4t} - \lambda t\right), \end{aligned}$$

which was to be proved. ■

Note that (6.15) is in fact equivalent to (6.11) since the latter follows from (6.15) by dividing by $\|g\|_2$ and taking sup in all $g \in L^2(B)$ with $B = A_c^r$.

Assuming that the sets A and B in (6.15) have finite measures and setting $f = 1_A$ and $g = 1_B$, we obtain from (6.15)

$$(P_t 1_A, 1_B) \leq \sqrt{\mu(A)\mu(B)} \exp\left(-\frac{d^2(A,B)}{4t} - \lambda t\right),$$

or, in terms of the heat kernel,

$$\iint_{AB} p_t(x,y)d\mu(x)d\mu(y) \leq \sqrt{\mu(A)\mu(B)} \exp\left(-\frac{d^2(A,B)}{4t} - \lambda t\right). \tag{6.16}$$

This can be considered as an integrated form of the Gaussian upper bound of the heat kernel. Note that, unlike the pointwise bounds, the estimate (6.16) holds on an *arbitrary* manifold.

6.3 Upper bounds of higher eigenvalues

We give here an application of Corollary 6.3 to eigenvalue estimates on a compact weighted manifold M. As before, denote by $\lambda_k(M)$ be the k-th smallest eigenvalue of Δ counted with the multiplicity. Recall that $\lambda_k(M) \geq 0$ and $\lambda_1(M) = 0$.

Theorem 6.4 *Let M be a connected compact weighted manifold. Let $A_1, A_2, ..., A_k$ be $k \geq 2$ disjoint measurable sets on M, and set*

$$\delta := \min_{i \neq j} d(A_i, A_j).$$

Then

$$\lambda_k(M) \leq \frac{4}{\delta^2} \max_{i \neq j} \left(\ln \frac{2\mu(M)}{\sqrt{\mu(A_i)\mu(A_j)}}\right)^2. \tag{6.17}$$

In particular, if we have two sets $A_1 = A$ and $A_2 = B$ then (6.17) becomes

$$\lambda_2(M) \leq \frac{4}{\delta^2} \left(\ln \frac{2\mu(M)}{\sqrt{\mu(A)\mu(B)}}\right)^2, \tag{6.18}$$

where $\delta := d(A, B)$.

Proof. We first prove (6.18). Let $\{\varphi_k\}_{k=1}^{\infty}$ be an orthonormal basis in $L^2(M, \mu)$ that consists of the eigenfunctions of Δ, so that φ_k has the eigenvalue $\lambda_k = \lambda_k(M)$. By the eigenfunction expansion (3.41), we have for any $t > 0$

$$\begin{aligned} \iint_{AB} p_t(x,y)d\mu(x)d\mu(y) &= \sum_{i=1}^{\infty} e^{-t\lambda_i} \int_A \varphi_i(x)d\mu(x) \int_B \varphi_i(y)d\mu(y) \\ &= \sum_{i=1}^{\infty} e^{-t\lambda_i} a_i b_i, \end{aligned} \tag{6.19}$$

where

$$a_i = (1_A, \varphi_i) \quad \text{and} \quad b_i = (1_B, \varphi_i).$$

By the Parseval identity

$$\sum_{i=1}^{\infty} a_i^2 = \|1_A\|_2^2 = \mu(A) \quad \text{and} \quad \sum_{i=1}^{\infty} b_i^2 = \|1_B\|_2^2 = \mu(B).$$

Since $\lambda_1 = 0$, the first eigenfunction φ_1 is identical constant. By the normalization condition $\|\varphi_1\|_2 = 1$ we obtain $\varphi_1 \equiv 1/\sqrt{\mu(M)}$, which implies

$$a_1 = (1_A, \varphi_1) = \frac{\mu(A)}{\sqrt{\mu(M)}} \quad \text{and} \quad b_1 = (1_B, \varphi_1) = \frac{\mu(B)}{\sqrt{\mu(M)}}.$$

Therefore, (6.19) yields

$$\begin{aligned}
\iint_{AB} p_t(x,y)d\mu(x)d\mu(y) &= a_1 b_1 + \sum_{i=2}^{\infty} e^{-t\lambda_i} a_i b_i \\
&\geq a_1 b_1 - e^{-t\lambda_2} \left(\sum_{i=2}^{\infty} a_i^2 \right)^{1/2} \left(\sum_{i=2}^{\infty} b_i^2 \right)^{1/2} \\
&\geq \frac{\mu(A)\mu(B)}{\mu(M)} - e^{-t\lambda_2} \sqrt{\mu(A)\mu(B)}.
\end{aligned}$$

Comparing with (6.16), we obtain

$$\sqrt{\mu(A)\mu(B)} e^{-\frac{\delta^2}{4t}} \geq \frac{\mu(A)\mu(B)}{\mu(M)} - e^{-t\lambda_2} \sqrt{\mu(A)\mu(B)},$$

whence

$$e^{-t\lambda_2} \geq \frac{\sqrt{\mu(A)\mu(B)}}{\mu(M)} - e^{-\frac{\delta^2}{4t}}$$

Choosing t from the identity

$$e^{-\frac{\delta^2}{4t}} = \frac{1}{2} \frac{\sqrt{\mu(A)\mu(B)}}{\mu(M)},$$

we conclude

$$\lambda_2 \leq \frac{1}{t} \ln \frac{2\mu(M)}{\sqrt{\mu(A)\mu(B)}} = \frac{4}{\delta^2} \left(\ln \frac{2\mu(M)}{\sqrt{\mu(A)\mu(B)}} \right)^2,$$

which was to be proved.

Let us now turn to the general case $k > 2$. Consider the following integrals

$$J_{lm} := \int_{A_l} \int_{A_m} p(t,x,y) d\mu(x) d\mu(y)$$

and set

$$a_i^{(l)} := (1_{A_l}, \varphi_i).$$

Exactly as above, we have

$$
\begin{aligned}
J_{lm} &= \sum_{i=1}^{\infty} e^{-t\lambda_i} a_i^{(l)} a_i^{(m)} \\
&= \frac{\mu(A_l)\mu(A_m)}{\mu(M)} + \sum_{i=k}^{\infty} e^{-\lambda_i t} a_i^{(l)} a_i^{(m)} + \sum_{i=2}^{k-1} e^{-\lambda_i t} a_i^{(l)} a_i^{(m)} \\
&\geq \frac{\mu(A_l)\mu(A_m)}{\mu(M)} - e^{-\lambda_k t}\sqrt{\mu(A_l)\mu(A_m)} \\
&\quad + \sum_{i=2}^{k-1} e^{-\lambda_i t} a_i^{(l)} a_i^{(m)}. \qquad (6.20)
\end{aligned}
$$

On the other hand, by (6.16)

$$
J_{lm} \leq \sqrt{\mu(A_l)\mu(A_m)} e^{-\frac{\delta^2}{4t}}. \qquad (6.21)
$$

Therefore, we can further argue as in the case $k = 2$ provided the term in (6.20) can be discarded, which the case when

$$
\sum_{i=2}^{k-1} e^{-\lambda_i t} a_i^{(l)} a_i^{(m)} \geq 0. \qquad (6.22)
$$

Let us show that (6.22) can be achieved by *choosing* l, m. To that end, let us interpret the sequence

$$
a^{(j)} := (a_2^{(j)}, a_3^{(j)}, ..., a_{k-1}^{(j)})
$$

as a $(k-2)$-dimensional vector in $\mathbb{R}^{k-2}$. Here j ranges from 1 to k so that we have k vectors $a^{(j)}$ in $\mathbb{R}^{k-2}$. Let us introduce the inner product of two vectors $u = (u_2, ..., u_{k-1})$ and $v = (v_2, ..., v_{k-1})$ in $\mathbb{R}^{k-2}$ by

$$
\langle u, v \rangle_t := \sum_{i=2}^{k-1} e^{-\lambda_i t} u_i v_i \qquad (6.23)
$$

and apply the following elementary fact:

Lemma 6.5 *From any $n+2$ vectors in a n-dimensional Euclidean space, it is possible to choose two vectors with non-negative inner product.*

Note that $n + 2$ is the smallest number for which the statement of Lemma 6.5 is true. Indeed, choose an orthonormal basis $e_1, e_2, ..., e_n$ in the given Euclidean space and consider the vector

$$
v := -e_1 - e_2 - ... - e_n.
$$

Then any two of the following $n + 1$ vectors

$$
e_1 + \varepsilon v,\ e_2 + \varepsilon v,\ ....,\ e_n + \varepsilon v,\ v
$$

have a negative inner product, provided $\varepsilon > 0$ is small enough.

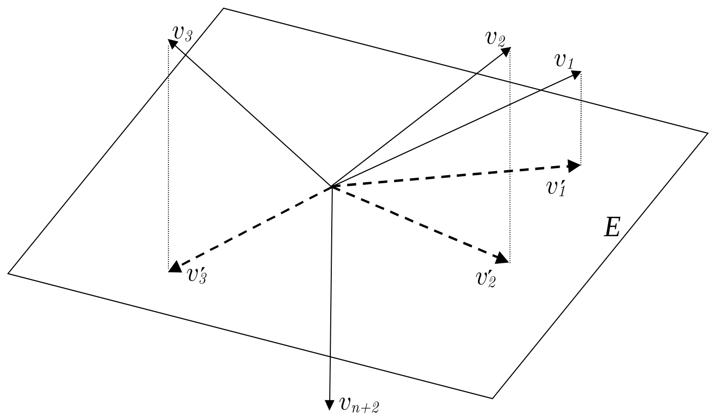

Figure 6.3: The vectors v_i' are the orthognal projections of v_i onto E.

Lemma 6.5 is easily proved by induction in n. The inductive basis for $n = 1$ is trivial. The inductive step is shown on Fig. 6.3. Indeed, assume that the $n+2$ vectors v_1, v_2, ..., v_{n+2} in $\mathbb{R}^n$ have pairwise obtuse angles. Denote by E the orthogonal complement of v_{n+2} in $\mathbb{R}^n$ and by v_i' the orthogonal projection of v_i onto E.

For any $i \leq n+1$, the vector v_i can be represented as

$$v_i = v_i' - \varepsilon_i v_{n+2} \, ,$$

where

$$\varepsilon_i = -\langle v_i, v_{n+2} \rangle > 0.$$

Therefore, we have

$$\langle v_i, v_j \rangle = \langle v_i', v_j' \rangle + \varepsilon_i \varepsilon_j \left| v_{n+2} \right|^2 .$$

By the inductive hypothesis, we have $\langle v_i', v_j' \rangle \geq 0$ for some i, j, which implies $\langle v_i, v_j \rangle \geq 0$, contradicting the assumption.

Now we can finish the proof of Theorem 6.4. Fix some $t > 0$. By Lemma 6.5, we can find l, m so that $\langle a^{(l)}, a^{(m)} \rangle_t \geq 0$; that is (6.22) holds. Then (6.20) and (6.21) yield

$$e^{-t\lambda_k} \geq \frac{\sqrt{\mu(A_l)\mu(A_m)}}{\mu(M)} - e^{-\frac{\delta^2}{4t}},$$

and we are left to choose t. However, t should not depend on l, m because we use t to define the inner product (6.23) *before* choosing l, m. So, we first write

$$e^{-t\lambda_k} \geq \min_{i,j} \frac{\sqrt{\mu(A_i)\mu(A_j)}}{\mu(M)} - e^{-\frac{\delta^2}{4t}}$$

and then define t by

$$e^{-\frac{\delta^2}{4t}} = \frac{1}{2} \min_{i,j} \frac{\sqrt{\mu(A_i)\mu(A_j)}}{\mu(M)},$$

whence (6.17) follows. ■