Analysis on Manifolds

Alexander Grigoryan

University of Bielefeld, WS 2025/26

Contents

Chapter 1

Riemannian manifolds

17-Oct-25 **Lecture 1**

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

1.1 Topological spaces and manifolds

Topological spaces. 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;
- the union of any family of open sets is open;
- the intersection of two open sets is open.

Closed sets. A subset F of M is called *closed* if its complement $F^c := M \setminus F$ is open. It follows that $\emptyset$ and M are closed sets, the intersection of any family of closed sets is closed, the union of two closed sets is closed.

Compact sets. A subset K of M is called *compact* if any open covering $\{\Omega_\alpha\}_{\alpha \in I}$ of K contains a finite subcover (where I is any index set). It is easy to prove that any closed subset of a compact set is also compact (Exercise 1).

Hausdorff spaces.

Definition. A topological space M is called *Hausdorff* if, for any two distinct points $x, y \in M$, there exist two disjoint open sets $U, V \subset M$ containing x and y, respectively. One says in this case that the sets U and V separate the points x, y.

In a Hausdorff space M, any compact subset K of M is closed (see Exercise 2).

Countable base.

Definition. 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.

Induced topology. Let M be a topological space and S be any subset of M. Then S itself is a topological space with the *induced* topology, that is, open sets in S are defined as intersections of open sets in M with S. If M has a countable base, then S also has countable base; if M is Hausdorff, the same is true also for S.

Continuous mappings. Let X and Y be two topological spaces. A mapping $F : X \to Y$ is called continuous if, for any open subset V of Y, the preimage $F^{-1}(Y)$ is an open subset of X. It is known that if F is continuous then, for any compact subset K of X, the image $F(K)$ is a compact subset of Y.

A mapping $F : X \to Y$ is called a homeomorphism if F is bijective, and both F and its inverse mapping are continuous.

Metric spaces. Any metric space (M, d) is a topological space with the following standard topology: a subset $\Omega \subset M$ is called open if for any $x \in \Omega$ there is a metric ball

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

with radius $r > 0$ that is a subset of Ω. It is easy to see that all metric balls are open sets.

The topology of a metric space is automatically Hausdorff because for any two distinct points $x, y \in M$, the balls $B(x, r/2)$ and $B(y, r/2)$ with $r = d(x, y)$ separate the points x, y.

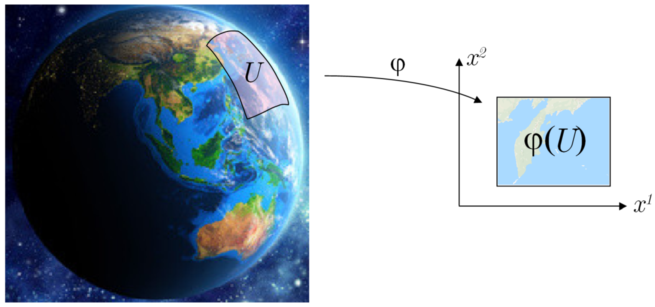
A metric space has a countable base if and only if it is *separable*, that is, if it contains a countable dense subset D. Indeed, if such a set D exists then all balls of rational radii centered at the points of D form a countable base. Conversely, if $\{B_j\}$ is a countable base then choosing one point in each B_j, we obtain a countable dense subset D of M.

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.

Local coordinates.

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 $\varphi(U)$ so that the coordinates $x^1, x^2, ..., x^n$ can be regarded as the Cartesian coordinates in a region in $\mathbb{R}^n$.



A chart on the surface of the earth

C-manifold.

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 an n-dimensional chart. The collection of all n-dimensional charts on M is called an *atlas*.

This terminology originates from geography and refers to a geographical atlas of the Earth, where each sheet can be regarded as (the image of) a 2-dimensional chart on the Earth's surface.

For example, $\mathbb{R}^n$ is a C-manifold and $U = \mathbb{R}^n$ is a single n-dimensional chart that covers $\mathbb{R}^n$.

Examples of C-manifolds. Let us consider some subsets of $\mathbb{R}^n$ that are C-manifolds. In all cases we use the induced topology of subsets.

Example. Let V be an open subset of $\mathbb{R}^n$ and $f : V \to \mathbb{R}^m$ be a continuous mapping. Then its *graph*

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

is a C-manifold because it is covered by a single n-dimensional chart (Γ, φ) where

$$\varphi : \Gamma \to V, \quad \varphi\,(x, f(x)) = x$$

is a homeomorphism.

Example. A subset M of $\mathbb{R}^{n+1}$ is called a *hypersurface* if it is locally a graph of a continuous real-valued function defined on an open subset of $\mathbb{R}^n$; that is, 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 with respect to one of the coordinates $x^1, ..., x^{n+1}$ of a continuous function $f : V \to \mathbb{R}$ defined on an open subset V of $\mathbb{R}^n$. Since $\Omega \cap M$ is a chart and M can be covered by such charts, we conclude that M is a C-manifold.

Example. Let $F : \mathbb{R}^{n+1} \to \mathbb{R}$ be a C^1-function. Consider the *null set* of F, that is, the set

$$M = \left\{ x \in \mathbb{R}^{n+1} : F(x) = 0 \right\},$$

and assume that $\nabla F(x) \neq 0$ for any point $x \in M$. Then M is a hypersurface and, hence, a C-manifold of dimension n (Exercise 7). Indeed, the condition $\nabla F(x) \neq 0$

means that one of the partial derivatives $\frac{\partial f}{\partial x^i}$ does not vanish at x. Then, by the implicit function theorem, the condition $F = 0$ in a small neighborhood of x is equivalent to $x^i = f(x^1, ..., \widehat{x^i}, ...x^n)$ for some C^1-function f depending on all $x^1, ..., x^n$ except for x^i. In other words, the set M in this neighborhood of x coincides with the graph of the function f. Since this is true for any $x \in M$, we conclude that M is a hypersurface.

For example, the unit sphere $\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$ is a hypersurface and, hence, a C-manifold of dimension n because it is the null set of the function $F(x) = \|x\|^2 - 1$, and $\nabla F = 2x \neq 0$ for all $x \in \mathbb{S}^n$.

Compact inclusion. Let M be a topological space. For any subset $A \subset M$, define the *closure* $\overline{A}$ of A as the intersection of all closed sets containing A. In other words, $\overline{A}$ is the smallest closed set containing A. We will use the relation $A \Subset B$ (*compact inclusion*) between two subsets A and B of M, which means the following: the closure $\overline{A}$ of A is a compact set and $\overline{A} \subset B$. A set A is called *precompact* (or relatively compact) if its closure $\overline{A}$ is compact.

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

Lemma 1.1 *For any C-manifold M, there is a countable family $\{U_i\}_{i=1}^{\infty}$ of charts covering all M and such that $U_i \Subset V_i$ for some chart V_i.*

Proof. 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. Hence, we obtain a covering of M by charts $\{U_x\}_{x \in M}$ such that each of then is compactly included in another chart. It remains to choose a countable subcover. By definition, manifold M has a countable base. Choose from this base only those elements that are contained in one of the sets U_x; let it be a sequence $\{B_j\}_{j=1}^{\infty}$. Since U_x is open, it is a union of some sets B_j. It follows that $\{B_j\}$ is a covering of M. Select for each B_j exactly *one* chart U_x containing B_j, say U_{x_j}. Thus, we obtain a countable family of charts $\{U_{x_j}\}$ covering M, which finishes the proof. ■

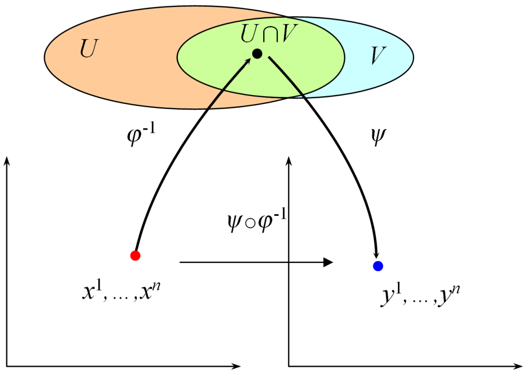
In particular, we see that a C-manifold M is a *locally compact* topological space.

22-Oct-25 Lecture 2

Change of coordinates. Let (U, φ) and (V, ψ) be two charts on a C-manifold M with the coordinates $x^1, ..., x^n$ and $y^1, ..., y^n$, respectively. Then in the intersection $U \cap V$ two coordinate systems are defined. The change of the coordinates from $x^1, ..., x^n$ to $y^1, ..., y^n$ is given then by *continuous* functions

$$y^i = y^i(x^1, ..., x^n) \tag{1.1}$$

because they are the components of the mapping $\psi \circ \varphi^{-1}$. Similarly, the change from $y^1, ..., y^n$ to $x^1, ..., x^n$ are given by continuous functions $x^i(y^1, ..., y^n)$ that are the components of the mapping $\varphi \circ \psi^{-1}$.



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

Smooth manifolds. Now we define the notion of a smooth manifold. Let k be a non-negative integer or $+\infty$.

Definition. A family $\mathcal{A}$ of charts on a C-manifold M is called a C^k-*atlas* if the charts from $\mathcal{A}$ cover all M and the change of coordinates in the intersection of any two charts from $\mathcal{A}$ is given by C^k-functions (that is, the functions (1.1) are C^k – all their partial derivatives of the order $\leq k$ are continuous).

Definition. Two C^k-atlases are said to be *compatible* if their union is again a C^k-atlas. A family 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 deal with smooth manifolds. By default, the term "manifold" will be used as a synonymous of "smooth manifold". By a chart on a smooth manifold we will always mean a chart from its C^∞-structure, that is, any chart compatible with the defining atlas $\mathcal{A}$.

Here are some examples of smooth manifolds.

1. $\mathbb{R}^n$ with the atlas consisting of a single chart $(\mathbb{R}^n, \mathrm{id})$.

2. The graph of any C^∞ mapping $f : V \to \mathbb{R}^m$ (where V is an open subset of $\mathbb{R}^n$) is a smooth manifold.

3. Any C^∞-hypersurface that is locally a graph of a C^∞-function, is a smooth manifold.

4. If $F : \mathbb{R}^{n+1} \to \mathbb{R}$ be a C^∞-function whose null set $M = \{F = 0\}$ is non-degenerate then M is a smooth manifold. For example, the unit sphere $\mathbb{S}^n$ is a smooth manifold.

5. 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.

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 $x^1, ..., x^n$. The set $C^k(M)$ is a linear space over $\mathbb{R}$ with respect to the usual addition of functions and multiplication by constant.

1.2 Cutoff functions and partition of unity

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

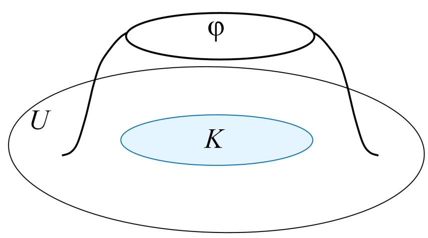
$$\operatorname{supp} f = \overline{\{x \in M : f(x) \neq 0\}},$$

where the bar stands for the closure. It follows from the definition of $\operatorname{supp} f$ that if $f \equiv 0$ outside a closed set $F \subset M$ then $\operatorname{supp} f \subset F$.

Denote by $C_0^k(M)$ the subspace of $C^k(M)$, which consists of functions with compact supports. If Ω is an open subset of M then $C_0^k(\Omega)$ denotes the set of all functions $f \in C_0^k(M)$ such that $\operatorname{supp} f \subset \Omega$.

Definition. Let M be a smooth manifold, U be an open subset of M and K be a compact subset of U. We say that a function function φ on M is a *cutoff function* of K in U if

- $\varphi \in C_0^\infty(U)$
- $\varphi \equiv 1$ in a neighborhood of K
- $0 \leq \varphi \leq 1$ on M.



A cutoff function φ of K in U

Lemma 1.2 *For any open subset U of $\mathbb{R}^n$ and any compact set $K \subset U$, there exists a cutoff function of K in U.*

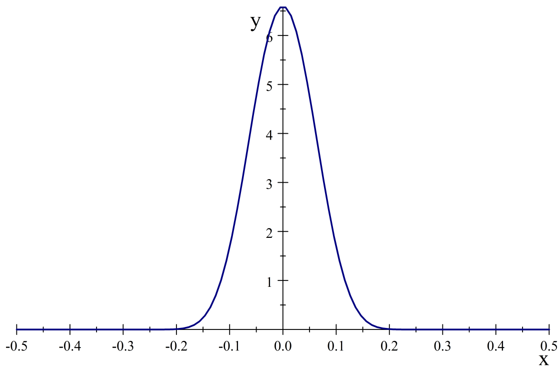
In the proof we use the notion of a mollifier. We say that a function $\psi \in C_0^\infty(\mathbb{R}^n)$ is a *mollifier* if $\operatorname{supp} \psi \subset B_1(0)$, $\psi \geq 0$, and

$$\int_{\mathbb{R}^n} \psi dx = 1. \tag{1.2}$$

For example, the following function

$$\psi(x) = \begin{cases} c\exp\left(-\frac{1}{\left(\frac{1}{4}-|x|^2\right)^2}\right), & |x| < 1/2 \\ 0, & |x| \geq 1/2 \end{cases} \tag{1.3}$$

is a mollifier, for a suitable normalizing constant $c > 0$.



The mollifier (1.3) in $\mathbb{R}$.

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

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

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

$$\int_{\mathbb{R}^n} \varphi_\varepsilon(x)dx = \int_{\mathbb{R}^n} \varepsilon^{-n}\psi\left(\frac{x}{\varepsilon}\right)dx = \int_{\mathbb{R}^n} \varepsilon^{-n}\psi(y)\left|\det\frac{dx}{dy}\right|dy = \int_{\mathbb{R}^n}\psi(y)\,dy = 1,$$

where we have used the change $x = \varepsilon y$ and the Jacobi matrix $\frac{dx}{dy} = \operatorname{diag}(\varepsilon, ..., \varepsilon)$ with the determinant ε^n.

Proof of Lemma 1.2. Let V be an open neighborhood of K such that $V \Subset U$, and set $f = 1_V$. Fix a mollifier ψ, some $\varepsilon \in (0,1)$ and consider the convolution

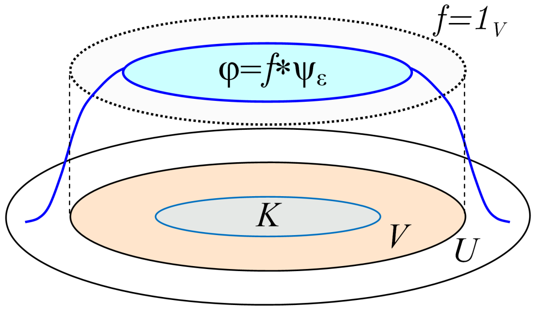
$$f * \psi_\varepsilon(x) = \int_{\mathbb{R}^n} f(x-y)\,\psi_\varepsilon(y)\,dy = \int_{B_\varepsilon(x)} f(z)\,\psi_\varepsilon(x-z)\,dz$$

and prove that $f * \psi_\varepsilon$ is a cutoff function of K in U provided ε is small enough.

Since f is Lebesgue integrable on $\mathbb{R}^n$, we have $f * \psi_\varepsilon \in C^\infty(\mathbb{R}^n)$. Clearly, we have

$$0 \leq f * \psi_\varepsilon(x) \leq \sup|f| \int_{\mathbb{R}^n} \psi_\varepsilon(y)\,dy = \sup|f| = 1.$$

Since f is supported in V and $\operatorname{supp}\psi_\varepsilon \subset B_\varepsilon(0)$, the convolution $f * \psi_\varepsilon$ is supported in a ε-neighborhood of V. Hence, if ε is small enough then $\operatorname{supp} f * \psi_\varepsilon \subset U$ so that $f * \psi_\varepsilon \in C_0^\infty(U)$.



Construction of a cutoff function

Besides, for small enough ε, we have $B_\varepsilon(x) \subset V$ for any $x \in K$, which implies $f|_{B_\varepsilon(x)} = 1$ and

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

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

The following statement provides a convenient tool 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 charts covering K. Then there exist functions $0 \leq \varphi_j \in C_0^\infty(U_j)$ 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.*

In fact, this statement is true for arbitrary open sets U_j not necessarily charts (see Exercises). A sequence of functions $\{\varphi_j\}$ 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 chart $U \supset K$, there exists a non-negative function $\varphi \in C_0^\infty(U)$ such that $\varphi \equiv 1$ in a neighborhood of K and $\varphi \leq 1$ on M; that is, φ is a cutoff function of K in U. Hence, Lemma 1.2 is a particular case of Theorem 1.3.

24-Oct-25 Lecture 3

Corollary 1.4 *Let $\{U_\alpha\}$ be an arbitrary family of charts covering M. 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 charts U_α and*

$$f = f_1 + ... + f_k \quad \text{on } M. \tag{1.4}$$

Proof. Let $K = \operatorname{supp} f$ and let $U_1, ..., U_k$ be a finite subfamily of $\{U_\alpha\}$ that covers K. By Theorem 1.3, there exists a partition of unity $\{\varphi_j\}_{j=1}^k$ at K subordinate to $\{U_j\}_{j=1}^k$. Set $f_j = f\varphi_j$ so that $f_j \in C_0^\infty(U_j)$. Then we have

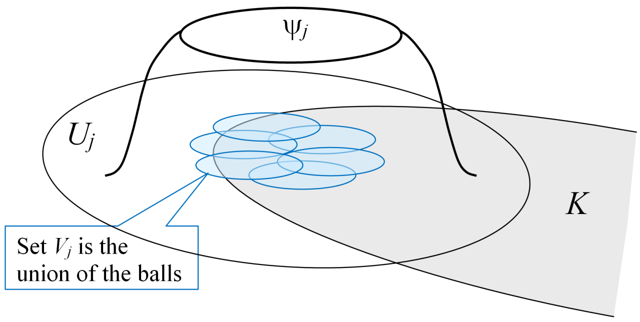
$$\sum_{j=1}^k f_j = f \quad \text{on } M,$$

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

Proof of Theorem 1.3. We claim that there exists a covering of K by a family $\{V_j\}_{j=1}^k$ of open sets such that $V_j \Subset U_j$. Indeed, since any point $x \in K$ belongs to a chart U_j, there is a ball B_x in this chart 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 $j = 1, ..., k$, define V_j as the union of all B_i's that are contained in U_j, that is

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

By construction, the set V_j is open, $V_j \Subset U_j$, and the union of all sets V_j covers K.



Function ψ_j is a cutoff function of V_j in U_j.

Considering the chart U_j as a subset of $\mathbb{R}^n$, we conclude by Lemma 1.2 that there exists a cutoff function $\psi_j \in C_0^\infty(U_j)$ of $\overline{V_j}$ in U_j. Now we consider U_j as a subset of M and extend ψ_j to M by setting $\psi_j = 0$ in $M \setminus U_j$, so that $\psi_j \in C_0^\infty(M)$.

Now we define functions φ_j, $j = 1, ..., k$, by

$$\varphi_j = \psi_j (1 - \psi_1) ... \left(1 - \psi_{j-1}\right), \tag{1.5}$$

that is,

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

Obviously, all functions φ_j are smooth,non-negative and $\operatorname{supp} \varphi_j \subset \operatorname{supp} \psi_j$. It follows that $\varphi_j \in C_0^\infty(U_j)$. Let us verify for any $m = 1, ..., k$ the following identity

$$1 - \sum_{j=1}^m \varphi_j = (1 - \psi_1) ... (1 - \psi_m). \tag{1.6}$$

Indeed, for $m = 1$ it is trivial. If it is true for some m, then, using the induction hypothesis and (1.5), we obtain

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

which proves the induction step.

It follows from (1.6) with $m = k$ that

$$\sum_{j=1}^{k} \varphi_j = 1 - (1 - \psi_1) \dots (1 - \psi_k) \leq 1. \tag{1.7}$$

Since $\psi_j = 1$ on V_j, (1.7) implies that $\sum_{j=1}^{k} \varphi_j \equiv 1$ on the union $\bigcup_{j=1}^{k} V_j$ that is an open neighborhood of K, which was to be proved. ■

1.3 Tangent space and tangent vectors

In $\mathbb{R}^n$ there is the notion of a direction and a directional derivative. If f is a function that is a defined in a neighborhood of a point $x_0 \in \mathbb{R}^n$ then, for any vector $v \in \mathbb{R}^n$, the directional derivative $\frac{\partial f}{\partial v}(x_0)$ is defined by

$$\frac{\partial f}{\partial v}(x_0) = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t}$$

provided the limit exists. If f is totally differentiable at x then $\frac{\partial f}{\partial v}$ is expressed via the partial derivatives by

$$\frac{\partial f}{\partial v}(x_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x_0) v^i$$

where v^i are the components of v. Hence, v determines the functional

$$\begin{aligned} C^\infty(\mathbb{R}^n) &\to \mathbb{R} \\ f &\mapsto \frac{\partial f}{\partial v}(x_0). \end{aligned}$$

Clearly, this functional determines back v because if we know $\sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x_0) v^i$ for all functions f then we can determine the components v^i, for example, by considering functions $f(x) = x^i$.

The purpose of this section is to define an analogous of the directional vector ξ on a smooth manifold, and we do it via the directional derivative.

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.

Example. 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 $\mathbb{R}$-differentiations $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$ that are clearly linearly independent.

Moreover, for any vector $v \in \mathbb{R}^n$ determines the $\mathbb{R}$-differentiation

$$\xi(f) = \frac{\partial f}{\partial v}(x_0)$$

so that $\frac{\partial}{\partial v} \in T_x\mathbb{R}^n$. Since

$$\frac{\partial f}{\partial v} = \sum_{i=1}^{n} v^i \frac{\partial f}{\partial x^i}$$

it follows that there is the following identity in $T_x\mathbb{R}^n$:

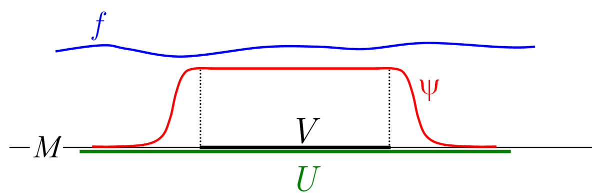
$$\frac{\partial}{\partial v} = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}.$$

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.*

In particular, $\dim T_{x_0}\mathbb{R}^n = n$, which implies that every $\mathbb{R}$-differentiation in $\mathbb{R}^n$ in a linear combination of the partial derivatives $\frac{\partial}{\partial x^i}$ and, hence, 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 a chart and $V \Subset U$ be an open subset of U. Then, for any function $f \in C^\infty(U)$, there exists a function $F \in C_0^\infty(M)$ such that $F = f$ on V.*

Proof. Indeed, let ψ be a cutoff function of $\overline{V}$ in U (see Lemma 1.2) that we extend to M by setting $\psi = 0$ outside U.

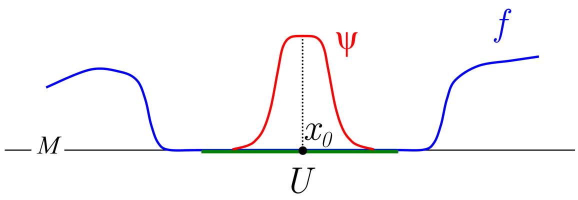


Functions f and ψ in Claim 1

Then the function $F = \psi f$ satisfies all the requirements as $F \in C_0^\infty(U)$ and $F = f$ on V. ∎

Claim 2. *Let $f \in C^\infty(M)$ and let $f \equiv 0$ in an open set U containing the point x_0. 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. By reducing U we can assume that U is a chart. Let V be an open neighborhood of x_0 that is compactly included in U. Let ψ be a cutoff function of $\overline{V}$ in U so that $\psi(x_0) = 1$. Then we have $f\psi \equiv 0$ on M, which implies $\xi(f\psi) = 0$.



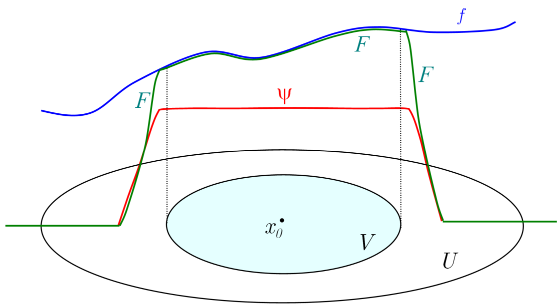
Functions f and ψ in Claim 2

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 open 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 small open neighborhood V of x_0.



Functions $f \in C^\infty(U)$ and $F \in C_0^\infty(M)$

Hence, define $\xi(f)$ by $\xi(f) := \xi(F)$. 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 $x \in B$,*

$$f(x) = f(o) + x^i h_i(x), \tag{1.8}$$

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

$$h_i(o) = \frac{\partial f}{\partial x^i}(o). \tag{1.9}$$

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

$$f(x) = f(o) + \int_0^1 \frac{d}{dt} f(tx)\, dt. \tag{1.10}$$

By the chain rule we have

$$\frac{d}{dt} f(tx) = \frac{\partial f}{\partial x^i}(tx) x^i$$

(also the summation over the repeated index i is assumed), whence it follows

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

Setting

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

we obtain (1.8). Clearly, $h_i \in C^\infty(B)$. The identity (1.9) follows from the line above by substitution $x = o$. ■

29-Oct-25 Lecture 4

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 $x \in B$,*

$$f(x) = f(o) + x^i \frac{\partial f}{\partial x^i}(o) + x^i x^j h_{ij}(x). \tag{1.11}$$

Proof. Applying (1.8) to the function h_i instead of f we obtain that there exist smooth functions h_{ij} in B, such that

$$h_i(x) = h_i(o) + x^j h_{ij}(x).$$

Substituting this into the representation (1.8) for f and using $h_i(o) = \frac{\partial f}{\partial x^i}(o)$ we obtain

$$f(x) = f(o) + x^i h_i(x) = f(o) + x^i \frac{\partial f}{\partial x^i}(o) + x^i x^j h_{ij}(x).$$

■

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 clearly linearly independent. We will prove that any tangent vector $\xi \in T_{x_0}M$ can be represented in the form

$$\xi = \xi^i \frac{\partial}{\partial x^i} \quad \text{where} \quad \xi^i = \xi\left(x^i\right). \tag{1.12}$$

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 (1.12) 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 (1.11) the following representation in a ball $B \subset U$ centred at o:

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

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

$$\xi(f) = \xi(1) f(o) + \xi\left(x^i\right) \frac{\partial f}{\partial x^i}(o) + \xi\left(x^i x^j h_{ij}\right). \tag{1.13}$$

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 $u_i = x^j h_{ij}$. By the linearity and the product rule, we have

$$\xi\left(x^i x^j h_{ij}\right) = \xi\left(x^i u_i\right) = \xi\left(x^i\right) u_i(o) + \xi(u_i) x^i(o) = 0,$$

because $x^i(o) = 0$ and $u_i(o) = x^j(o) h_{ij}(o) = 0$. Hence, in the right hand side of (1.13), the first and the third term vanish. Setting $\xi^i := \xi\left(x^i\right)$, we obtain

$$\xi(f) = \xi^i \frac{\partial f}{\partial x^i}(o), \tag{1.14}$$

which is equivalent to (1.12). ■

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 (1.14) takes the familiar form

$$\boxed{\frac{\partial f}{\partial \xi} = \xi^i \frac{\partial f}{\partial x^i}}, \tag{1.15}$$

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

Definition. A *vector field* on a smooth manifold M is a family $\{\xi(x)\}_{x \in M}$ of tangent vectors such that $\xi(x) \in T_x M$ 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.

1.4 Submanifolds

If M is a smooth manifold of dimension n then any open subset $\Omega \subset M$ trivially becomes a smooth manifold of the same dimension by restricting all charts to Ω.

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 a (embedded) *submanifold* of dimension $m < n$ if, for any point $x_0 \in S$, there is a chart $U \ni x_0$ in M with local coordinates $x^1, ..., x^n$ such that $S \cap U$ is given in the local coordinates by the equations

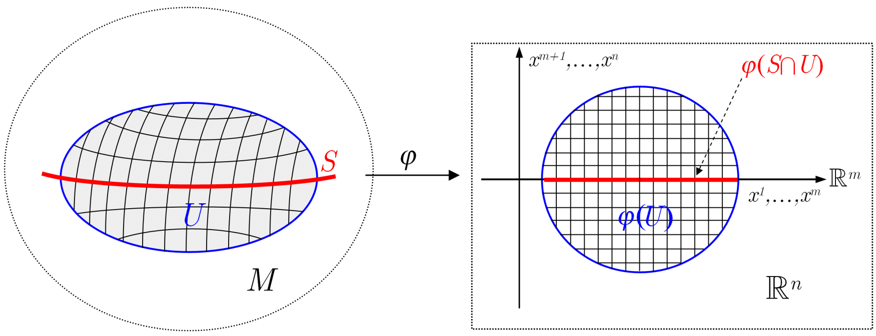
$$x^{m+1} = x^{m+2} = ... = x^n = 0. \tag{1.16}$$

Lemma 1.6 *For any chart U on M as above, the set $S \cap U$ is a chart on S of dimension m with coordinates $x^1, ..., x^m$. Besides, the collection of all such charts on S determines on S a C^∞-structure so that S is a smooth manifold of dimension m.*

Proof. Let $\varphi : U \to \mathbb{R}^n$ be a homeomorphism of U onto an open subset of $\mathbb{R}^n$ that determines the local coordinates $x^1, ..., x^n$ in U. Then the condition that $S \cap U$ is given by the equations (1.16) means that

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

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 of $S \cap U$ onto 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 $x^1, x^2, ..., x^m$. The family $\mathcal{A}$ of all such charts makes S into m-dimensional C-manifold.

We are left to verify that $\mathcal{A}$ is a C^∞-atlas. Let V be another chart on M with coordinates $y^1, ..., y^n$ such that $S \cap V$ is given in these coordinates by the equations

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

so that the coordinates in $S\cap V$ are $y^1, ..., y^m$. We need to verify that in $S\cap U\cap V$ the change of coordinates $y^i = y^i(x^1, ..., x^m)$ is given by C^∞-functions. Indeed, the change of coordinates in $U\cap V$ is given by C^∞ functions $y^i = y^i(x^1, ..., x^m, ..., x^n)$. It follows that on $S\cap U\cap V$ we have

$$y^i = y^i(x^1, ..., x^m, 0..., 0),$$

where the right hand side is clearly a C^∞ function of $x^1, ..., x^m$, which concludes the proof. ■

Example. 1. Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ be a smooth mapping. Then graph of f is a submanifold of $\mathbb{R}^{n+m}$ of dimension n (see Exercise 15 (a)).

2. Any smooth hypersurface in $\mathbb{R}^{n+1}$ is a submanifold of $\mathbb{R}^{n+1}$ of dimension n (see Exercise 15 (b)).

Lemma 1.7 *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 S of F, that is*

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

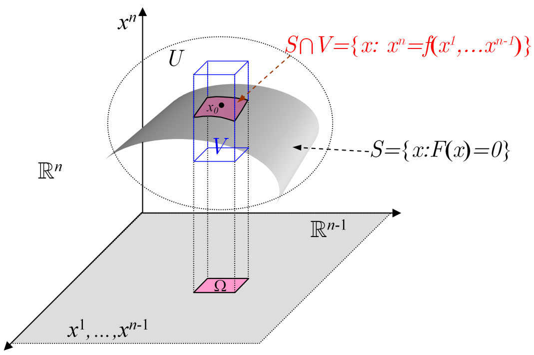
and assume that S is non-degenerate, that is,

$$\nabla F(x) \neq 0 \text{ for all } x \in S. \tag{1.17}$$

Then S is a submanifold of dimension $n-1$.

The condition (1.17) means that, for any $x \in S$ there exists a chart U containing x such that $\nabla F(x) \neq 0$ in the coordinates of this chart. It follows that the same condition will be satisfied in any other chart containing x.

Proof. Fix a point $x_0 \in S$ and a chart U containing x_0. The condition $\nabla F(x_0) \neq 0$ means that one of the partial derivatives $\frac{\partial F}{\partial x^i}$ does not vanish at x_0. Without loss of generality we can assume that $\frac{\partial F}{\partial x^n}(x_0) \neq 0$.



By the implicit function theorem, there exists an open subset Ω in $\mathbb{R}^{n-1}$ (=the subspace of $\mathbb{R}^n$ given by $x^n = 0$) and an open interval $I \in \mathbb{R}$ and a smooth function $f : \Omega \to I$ such that $V := \Omega \times I \subset U$ and the equation $F(x) = 0$ can be resolved in V with respect to the coordinate x^n as follows:

$$F(x) = 0 \Leftrightarrow \; x^n = f(x^1, ..., x^{n-1}) \; \text{ for all } x \in V.$$

Let us introduce new coordinates $y^1, ..., y^n$ in V as follows:

$$y^1 = x^1,\ y^2 = x^2,\ ...,\ y^{n-1} = x^{n-1},$$
$$y^n = x^n - f(x^1, ..., x^{n-1}).$$

Clearly, this change of coordinates is bijective, smooth in the both directions, and in the new coordinates the equation of S in V is $y^n = 0$. Hence, S is a $(n-1)$-dimensional submanifold of M. ■

05-Nov-25 Lecture 5

1.5 Cotangent space and differential

Let M be a smooth manifold of dimension n. For any point $x \in M$, we have defined the tangent space T_xM that consists of all $\mathbb{R}$-differentiations at the point x. Recall that $\mathbb{R}$-differentiation is a linear functional on $C^\infty(M)$ that satisfies the product rule at x. The elements of T_xM are also called *tangent vectors*. We have seen that T_xM is a linear space over $\mathbb{R}$ of dimension n.

Let V be any finitely dimensional linear space over $\mathbb{R}$ whose elements are called vectors. Then V has a *dual space* V^* that consists of all linear functionals on V; that is, the elements of V^* are linear functions $\omega : V \to \mathbb{R}$, that are called *covectors*. Clearly, V^* is itself a linear space over $\mathbb{R}$ with respect to addition of covectors and multiplication by constants: if $\omega_1, \omega_2 \in V^*$ then their sum $\omega_1 + \omega_2$ that is defined by

$$(\omega_1 + \omega_2)(\xi) = \omega_1(\xi) + \omega_2(\xi) \quad \text{for all } \xi \in V,$$

and is also a covector, and the product $\lambda\omega$ with a constant $\lambda \in \mathbb{R}$, that is defined by

$$(\lambda\omega)(\xi) = \lambda\omega(\xi) \quad \text{for all } \xi \in V$$

is also a covector.

For example, if $V = \mathbb{R}^n$ then every covector, that is, is a linear function $\omega : \mathbb{R}^n \to \mathbb{R}$, has the form

$$\omega(\xi) = \omega_1\xi^1 + \omega_2\xi^2 + ... + \omega_n\xi^n \quad \text{for any } \xi \in \mathbb{R}^n$$

with some coefficients $\omega_1, ..., \omega_n$. Hence, ω can be identified with the sequence $\{\omega_1, ..., \omega_n\}$ and $(\mathbb{R}^n)^*$ is isomorphic to $\mathbb{R}^n$. Consequently, for any n-dimensional linear space V, the dual space V^* has the same dimension n.

As any other n-dimensional linear space, T_xM possesses the *dual space* T_x^*M that consists of all linear functionals $\omega : T_xM \to \mathbb{R}$. The space T_x^*M is also an n-dimensional linear space over $\mathbb{R}$.

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* (or cotangent vectors).

For any $\omega \in T_x^*M$ and $\xi \in T_xM$, the value $\omega(\xi)$ will be also denoted by $\langle \omega, \xi \rangle$ and referred to as the *pairing* of ω and ξ. This notation reflects the fact that every vector $\xi \in T_xM$ can be regarded as a linear functional on T_x^*M given by $\xi(\omega) = \langle \omega, \xi \rangle$.

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

$$\begin{array}{l} T_xM \to \mathbb{R} \\ \xi \mapsto \xi(f) \end{array} \tag{1.18}$$

and observe that it is linear, by the definition of linear operations on T_xM:

$$(\xi_1 + \xi_2)(f) = \xi_1(f) + \xi_2(f), \quad (\lambda\xi)(f) = \lambda\xi(f).$$

Hence, the mapping (1.18) is a tangent covector.

Definition. The tangent covector $\xi \mapsto \xi(f)$ is called the *differential* of f at x and is denoted by df. In other words, df is a tangent covector defined by

$$\boxed{\langle df, \xi \rangle := \xi(f)} \text{ for any } \xi \in T_xM. \tag{1.19}$$

Let us fix a chart around the point x with the local coordinates $x^1, ..., x^n$. Then, for the tangent vector $\xi = \frac{\partial}{\partial x^i}$, we have

$$\langle df, \frac{\partial}{\partial x^i} \rangle = \frac{\partial f}{\partial x^i}.$$

Applying the above definition to the function $f = x^i$, we obtain a tangent covector dx^i.

Lemma 1.8 *$\{dx^i\}_{i=1}^n$ is a basis in the cotangent space T_x^*M.*

Proof. Indeed, let V be an n-dimensional linear space with a basis $\{e_1, ..., e_n\}$. For any $i = 1, ..., n$, define a covector $e^i \in V^*$ as follows: for any e_j set

$$\langle e^i, e_j \rangle = \delta_j^i := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (\delta_j^i \text{ is called the } \textit{Kronecker delta}),$$

and then extend $\langle e^i, \xi \rangle$ to all $\xi \in V$ by linearity. It is know from Linear Algebra that the sequence of covectors $\{e^1, ..., e^n\}$ is a basis in V^* that is called the *dual basis*.

Recall that $\left\{\frac{\partial}{\partial x^j}\right\}_{j=1}^n$ is a basis in T_xM. We claim that $\{dx^i\}_{i=1}^n$ is the dual basis in T_x^*M, which follows from

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

∎

Consequently, any tangent covector $\omega \in T_x^*M$ has an expansion in this basis:

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

where the coefficients $\omega_i \in \mathbb{R}$ are referred to as the *components* of ω in the basis $\{dx^i\}$.

Hence, for any tangent vector $\xi = \xi^j \frac{\partial}{\partial x^j}$, we obtain

$$\langle \omega, \xi \rangle = \langle \omega_i dx^i, \xi^j \frac{\partial}{\partial x^j} \rangle = \omega_i \xi^j \langle dx^i, \frac{\partial}{\partial x^j} \rangle = \omega_i \xi^j \delta^i_j = \omega_i \xi^i.$$

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

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

For example, for the covector df we obtain from (1.19) that

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

and, hence,

$$\boxed{df = \frac{\partial f}{\partial x^i} dx^i.} \tag{1.20}$$

1.6 Riemannian manifolds

Riemannian metric. Recall that, in a linear space V over $\mathbb{R}$, an *inner product* is any symmetric, positive definite, bilinear form g on V; that is, for any pair $\xi, \eta \in V$, $g(\xi, \eta)$ is a real number that satisfies the following conditions:

- the mappings $\xi \mapsto g(\xi, \eta)$ and $\eta \mapsto g(\xi, \eta)$ are linear (bilinearity);
- $g(\xi, \eta) = g(\eta, \xi)$ (symmetry)
- $g(\xi, \xi) \geq 0$ and $g(\xi, \xi) = 0 \Leftrightarrow \xi = 0$ (positive definiteness).

The value $g(\xi, \eta)$ is called the inner product of ξ and η and is frequently denoted simply by $\langle \xi, \eta \rangle$.

Let M be a smooth n-dimensional manifold.

Definition. 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 an inner product in the tangent space $T_x M$, *smoothly depending* on $x \in M$.

The metric tensor $\mathbf{g}$ determines an inner product $\langle \cdot, \cdot \rangle_{\mathbf{g}}$ in any tangent space $T_x M$ by

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

so that $T_x M$ becomes a Euclidean (=inner product) space.

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

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

where

$$\boxed{g_{ij}(x) = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle_{\mathbf{g}}}. \tag{1.21}$$

Clearly, $(g_{ij}(x))_{i,j=1}^{n}$ is a symmetric positive definite $n \times n$ matrix (where i is the row index and j is the column index). The functions $g_{ij}(x)$ are called the *components* of the metric tensor $\mathbf{g}$ in the coordinates $x^1, ..., x^n$.

Definition. The condition that $\mathbf{g}(x)$ *smoothly depends* on x means that all the components $g_{ij}(x)$ are C^∞-functions in any chart.

Definition. For any two covectors $u, v \in T_x^*M$, define their *tensor product* uv as a bilinear functional on T_xM by

$$uv\,(\xi, \eta) = \langle u, \xi \rangle \langle v, \eta \rangle .$$

The tensor product is normally denoted by $u \otimes v$, but we will suppress $\otimes$ in the notation.

Claim. *The metric tensor* $\mathbf{g}$ *can be represented in the local coordinates* $x^1, ..., x^n$ *as follows:*

$$\boxed{\mathbf{g} = g_{ij} dx^i dx^j}, \tag{1.22}$$

where $dx^i dx^j$ *is the tensor product of the covectors* dx^i *and* dx^j.

Proof. Indeed, since

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

it follows that, for all $\xi, \eta \in T_xM$,

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

which proves (1.22). ■

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

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

where $\left(dx^i\right)^2 = dx^i dx^i$. For this metric, we have $(g_{ij}) = \mathrm{id}$.

Lowering the index. Let $(M, \mathbf{g})$ be a Riemannian manifold. The metric tensor $\mathbf{g}$ can be regarded as a linear mapping

$$\mathbf{g}(x) : T_xM \to T_x^*M \tag{1.23}$$

as follows. For any vector $\xi \in T_xM$, define $\mathbf{g}(x)\xi \in T_x^*M$ by the identity

$$\langle \mathbf{g}(x)\xi, \eta \rangle = \langle \xi, \eta \rangle_{\mathbf{g}} \text{ for all } \eta \in T_xM, \tag{1.24}$$

Rewriting (1.24) in the local coordinates, we obtain

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

which implies

$$\boxed{(\mathbf{g}(x)\xi)_j = g_{ij}\xi^i}. \tag{1.25}$$

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.*

Raising the index.

Lemma 1.9 *The linear operator $\mathbf{g}(x) : T_xM \to T_x^*M$ is invertible. The inverse mapping*

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

has in the local coordinates the following form: for any $u \in T_x^*M$

$$\boxed{(\mathbf{g}^{-1}(x)u)^i = g^{ij}u_j}, \tag{1.26}$$

where the matrix (g^{ij}) is the inverse matrix of (g_{ij}), that is,

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

Proof. The operator $\mathbf{g}(x)$ is injective: indeed, if $\xi \neq 0$ then also $\mathbf{g}(x)\xi \neq 0$ because

$$\langle \mathbf{g}(x)\xi, \xi \rangle = \langle \xi, \xi \rangle_{\mathbf{g}} > 0.$$

Since the spaces T_xM and T_x^*M have the same dimensions, it follows that $\mathbf{g}(x)$ is bijective and, hence, invertible. Hence, its matrix (g_{ij}) is also invertible.

Fix $u \in T_x^*M$ and set $\xi = \mathbf{g}^{-1}(x)u$ so that $u = \mathbf{g}(x)\xi$. By (1.25) we have

$$u_j = g_{jk}\xi^k.$$

Using the fact that (g^{ij}) is the inverse matrix of (g_{ij}), we obtain

$$g^{ij}u_j = g^{ij}g_{jk}\xi^k = \delta^i_k\xi^k = \xi^i,$$

which is equivalent to (1.26). We have used the fact that $g^{ij}g_{jk}$ is the (i,k)-entry of the product of the matrices (g^{ij}) and (g_{jk}), and this product is the identity matrix $\mathrm{id} = \left(\delta_k^i\right)$. ■

Following the above convention and denoting the vector $\mathbf{g}^{-1}(x)u$ also by u, we obtain the following relation between the vector and covector components of u:

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

The operation of passing from u_j to u^i is called *raising the index*. Clearly, this is the inverse operation to lowering the index.

When we use the same notation u for a covector as well as for the corresponding vector then the covector components u_i are also called *covariant* components, while the vector components u^i are called *contravariant* components.

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

$$\langle u, v\rangle_{\mathbf{g}^{-1}} := \langle \mathbf{g}^{-1}(x)u, \mathbf{g}^{-1}(x)v\rangle_{\mathbf{g}}. \tag{1.27}$$

In the local coordinates we have

$$\langle u, v\rangle_{\mathbf{g}^{-1}} = g^{ij}u_iv_j \tag{1.28}$$

because by (1.27), (1.24) and (1.26)

$$\langle u, v\rangle_{\mathbf{g}^{-1}} = \langle \mathbf{g}^{-1}(x)u, \mathbf{g}^{-1}(x)v\rangle_{\mathbf{g}} = \langle u, \mathbf{g}^{-1}(x)v\rangle = u_i\left(\mathbf{g}^{-1}(x)v\right)^i = g^{ij}u_iv_j. \tag{1.29}$$

By elimination $\mathbf{g}$ from all the notations, we see from (1.29) that the expression $\langle u, v\rangle$ has the same value in the following four possible cases:

- u and v are covectors, and $\langle u, v\rangle$ is their inner product in T_x^*M;
- u and v are vectors, and $\langle u, v\rangle$ is their inner product in T_xM;
- u is a covector, v is a vector, and $\langle u, v\rangle$ is their pairing;
- u is a vector, v is a covector, and $\langle u, v\rangle$ is their pairing.

07-Nov-25 Lecture 6

Let $(M, \mathbf{g})$ be a Riemannian manifold. By definition, $\mathbf{g}(x)$ is for any $x \in M$ an inner product in the tangent space T_xM. We have observed that the inner product determines a linear mapping

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

from T_xM to the cotangent space T_x^*M (that is the dual space to T_xM). This linear mapping is given by the identity

$$\langle \mathbf{g}(x)\xi, \eta\rangle = \langle \xi, \eta\rangle_{\mathbf{g}} \text{ for all } \xi, \eta \in T_xM. \tag{1.30}$$

In the local coordinates, we have, for any $\xi \in T_xM$,

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

where (g_{ij}) is the matrix of the components of $\mathbf{g}$ (lowering the index).

We have proved that the mapping $\mathbf{g}(x)$ is bijective and, hence, it has the inverse mapping

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

In the local coordinates we have for any $u \in T_x^*M$

$$\left(\mathbf{g}^{-1}(x)u\right)^i = g^{ij}u_j$$

where (g^{ij}) is the inverse matrix of (g_{ij}) (raising the index). The operator $\mathbf{g}^{-1}(x)$ gives rise to the inner product $\langle \cdot, \cdot \rangle_{\mathbf{g}^{-1}}$ in the cotangent space T_x^*M.

Gradient.

Definition. For any $f \in C^\infty(M)$ define its *gradient* $\nabla f(x)$ at any point $x \in M$ by

$$\nabla f(x) = \mathbf{g}^{-1}(x)df(x), \tag{1.31}$$

that is, $\nabla f(x)$ is a vector that is obtained from the covector $df(x)$ by raising the index.

Applying (1.30) with $\xi = \nabla f(x)$, we obtain, for any $\eta \in T_xM$,

$$\langle \nabla f, \eta \rangle_{\mathbf{g}} = \langle \mathbf{g}(x)\nabla f, \eta \rangle = \langle df, \eta \rangle = \eta(f) = \frac{\partial f}{\partial \eta}$$

that is,

$$\boxed{\langle \nabla f, \eta \rangle_{\mathbf{g}} = \frac{\partial f}{\partial \eta}}, \tag{1.32}$$

which can be considered as an alternative definition of the gradient. In the local coordinates $x^1, ..., x^n$, the components $(\nabla f)^i$ of ∇f are obtained from the components $\frac{\partial f}{\partial x^j}$ of df by raising the index, that is,

$$\boxed{(\nabla f)^i = g^{ij}\frac{\partial f}{\partial x^j}} \tag{1.33}$$

If h is another smooth function on M then we obtain from (1.28)

$$\langle \nabla f, \nabla h \rangle_{\mathbf{g}} = \langle df, dh \rangle_{\mathbf{g}^{-1}} = g^{ij}\frac{\partial f}{\partial x^i}\frac{\partial h}{\partial x^j}. \tag{1.34}$$

Of course, if $(g^{ij}) = \mathrm{id}$ (as in $\mathbb{R}^n$) then (1.33) becomes

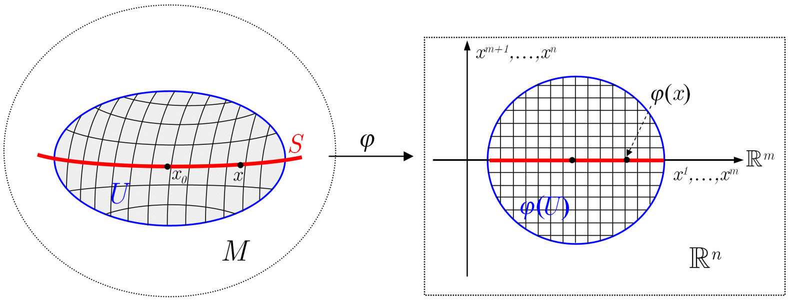
$$(\nabla f)^i = \frac{\partial f}{\partial x^i},$$

which matches the definition of the gradient in $\mathbb{R}^n$.

1.7 Riemannian metric on a submanifold

Let M be a smooth manifold of dimension n. Recall that a set $S \subset M$ is called a (embedded) *submanifold* of dimension m if, for any point $x_0 \in S$, there is a chart $U \ni x_0$ in M with local coordinates $x^1, ..., x^n$ such that a point $x \in U$ belongs to $S \cap U$ if and only if

$$x^{m+1} = x^{m+2} = ... = x^n = 0. \tag{1.35}$$



In this definition we have assumed $m < n$. However, it is valid also when $m = n$. In this case the number of equations in (1.35) is equal to $n - m = 0$, that is, the condition (1.35) is empty, which means that all points of U must lie in $S \cap U$, that is, $U \subset S$. Hence, S is a submanifold of M of dimension n if for any $x_0 \in S$ there is a chart $U \ni x_0$ such that $U \subset S$, which is equivalent to the fact that S is an open subset of M.

In what follows we assume that S is a submanifold of M of dimension $m \leq n$. Recall that any submanifold is also a smooth manifold with the induced smooth structure.

Tangent space on a submanifold. Fix a point $x_0 \in S$ and let ξ be an $\mathbb{R}$-differentiation on S at x_0. For any smooth function f on M, its restriction $f|_S$ is a smooth function on S. Hence, $\xi(f|_S)$ is defined, and by setting

$$\xi(f) := \xi(f|_S), \tag{1.36}$$

we extend ξ to an $\mathbb{R}$-differentiation on M at the same point x_0. In other words, (1.36) defines a linear mapping

$$T_{x_0}S \to T_{x_0}M. \tag{1.37}$$

Lemma 1.10 *The mapping (1.37) is injective and, hence, provides a natural identification of $T_{x_0}S$ as a subspace of $T_{x_0}M$.*

Proof. It suffices to prove that if $\xi \in T_{x_0}S$ is non-zero then its extension to $T_{x_0}M$ is also non-zero. Since $\xi \neq 0$ as an element of $T_{x_0}S$, there exists a smooth function $h \in C^\infty(S)$ such that $\xi(h) \neq 0$. Let $x^1, ..., x^n$ be a local coordinate system on M in a neighborhood of x_0 where S is given by the equations (1.35). Since $x^1, ..., x^m$ are the local coordinates on S, h is a smooth function of $x^1, ..., x^m$. Define a function f in the chart $x^1, ..., x^n$ on M by setting

$$f(x^1, ..., x^m, ..., x^n) = h(x^1, ..., x^m).$$

Then f is 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

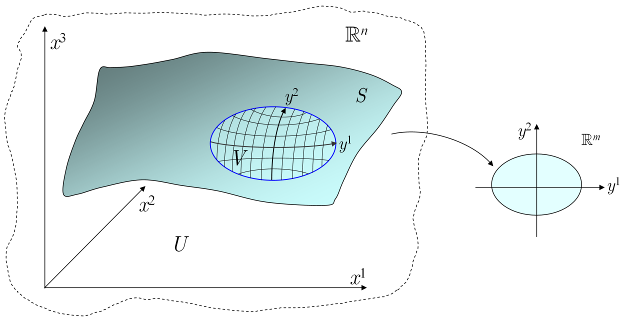
$$\xi(f) = \xi(f|_S) = \xi(h) \neq 0,$$

that is, ξ is non-zero as an element of $T_{x_0}M$. Hence, the mapping (1.37) is injective. ■

Let us describe the mapping (1.37) in the local coordinates. Let $x^1, ..., x^n$ be arbitrary local coordinates in a chart U in M and $y^1, ..., y^m$ be arbitrary local coordinates in a chart V on S. In the intersection $U \cap V$ we have the relations

$$x^i = x^i\left(y^1, ..., y^m\right), \quad i = 1, ..., n, \tag{1.38}$$

that express the x-coordinates of any point of $U \cap V$ via its y-coordinates. By renaming $U \cap V$ into V, we can assume in what follows that $V \subset U$.



Local coordinates $x^1, ..., x^n$ and $y^1, ..., y^m$

Any smooth function $f = f(x^1, ..., x^n)$ in U can be regarded in V as a smooth function of $y^1, ..., y^m$ using (1.38) that is, as the following function of $y^1, ..., y^m$:

$$f\left(x^1\left(y^1, ..., y^m\right), x^2\left(y^1, ..., y^m\right), ..., x^n\left(\left(y^1, ..., y^m\right)\right)\right).$$

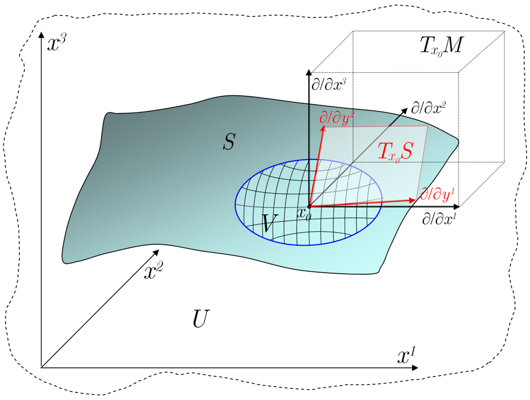
By the chain rule, we obtain

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

which can be rewritten in the operator form as follows:

$$\boxed{\frac{\partial}{\partial y^k} = \frac{\partial x^i}{\partial y^k}\frac{\partial}{\partial x^i}}. \tag{1.39}$$

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 (1.39) identifies explicitly $T_{x_0}S$ as a subspace of $T_{x_0}M$.



Tangent space $T_{x_0}S$ as a subspace of $T_{x_0}M$

Cotangent space on a submanifold 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 will also be denoted by ω. Hence, we obtain a linear mapping $T^*_{x_0}M \to T^*_{x_0}S$. This linear mapping is surjective as any linear functional on the subspace $T_{x_0}S$ can be extended to a linear functional on the entire space $T_{x_0}M$.

Assuming as above that $x^1, ..., x^n$ and $y^1, ..., y^m$ are local coordinate systems in M and S respectively, and considering all x^i's as functions of y^j's, let us represent $dx^i|_{T_{x_0}S}$ in the basis dy^j. Since by (1.39)

$$\langle dx^i, \frac{\partial}{\partial y^j}\rangle_{T_{x_0}S} = \langle dx^i, \frac{\partial}{\partial y^j}\rangle_{T_{x_0}M} = \langle dx^i, \frac{\partial x^l}{\partial y^j}\frac{\partial}{\partial x^l}\rangle = \frac{\partial x^l}{\partial y^j}\delta^i_l = \frac{\partial x^i}{\partial y^j},$$

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

$$\boxed{dx^i = \frac{\partial x^i}{\partial y^j}dy^j}. \tag{1.40}$$

Alternatively, (1.40) follows from (1.20) considering x^i as a function in the chart $y^1, ..., y^m$. Indeed, one can prove that, for any function $f \in C^\infty(M)$

$$(df)|_{T_{x_0}S} = d(f|_S)$$

(see Exercise 24). Applying this to the function $f = x^i$ and observing that, by the chain rule, for the restriction of x^i to S we have

$$dx^i = \frac{\partial x^i}{\partial y^j}dy^j$$

(see (1.20)), we obtain that the restriction of dx^i to $T_{x_0}S$ is given by the same formula.

Induced Riemannian metric Let $\mathbf{g}$ be a Riemannian metric on M. For any $x \in S$, 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* of S.

Lemma 1.11 *In the local coordinates $x^1, ..., x^n$ on M and $y^1, ..., y^m$ on S we have the identity*

$$\boxed{(g_S)_{ij} = g_{kl} \frac{\partial x^k}{\partial y^i} \frac{\partial x^l}{\partial y^j}}, \tag{1.41}$$

where g_{kl} are the components of $\mathbf{g}$ in the chart $x^1, ..., x^n$ and $(g_S)_{ij}$ are the components of $\mathbf{g}_S$ in the chart $y^1, ..., y^m$. In the matrix form, we have

$$\boxed{g_S^y = J^T g^x J} \tag{1.42}$$

where $g^x = (g_{kl})$, $g_S^y = ((g_S)_{ij})$ and J is the Jacobi matrix of the change $x = x(y)$, that is,

$$J = (J_{ki}) = \left(\frac{\partial x^k}{\partial y^i} \right). \tag{1.43}$$

Note that, in the matrix J in (1.43), $k = 1, ..., n$ is the row index and $i = 1, ..., m$ is the column index, so that J is an $n \times m$ matrix. Hence, the right hand side of (1.42) is the product of the three matrices of the following dimensions: $m \times n$, $n \times n$, $n \times m$ that results in an $m \times m$ matrix.

Proof. Restricting $\mathbf{g} = g_{kl} dx^k dx^l$ to $T_{x_0}S$, we obtain by (1.40)

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

Comparing with

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

we obtain (1.41). Next, we have by (1.41) and (1.43)

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

whence (1.42) follows. ■

In a particular case $m = n$, S is an open subset of M and the induced metric $\mathbf{g}_S$ coincides with the original metric $\mathbf{g}$, so that (1.42) provides the relation between the matrices g^x and g^y of $\mathbf{g}$ in two coordinate systems $x^1, ..., x^n$ and $y^1, ..., y^n$, respectively (cf. Exercise 22).

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 $\mathbf{g}_{\mathbb{S}^n}$.

1. Let us compute $\mathbf{g}_{\mathbb{S}^1}$ using the following chart on the unit circle $\mathbb{S}^1$. The upper semi-circle

$$U := \mathbb{S}^1 \cap \{x^2 > 0\}$$

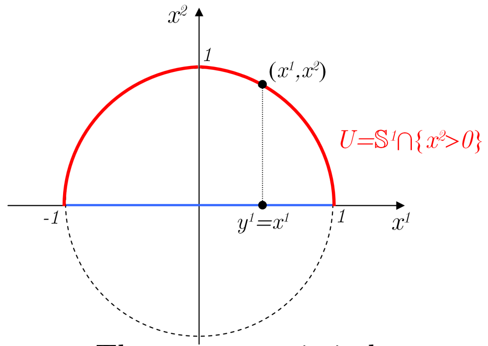
is the graph of a function $f(x^1) = \sqrt{1-(x^1)^2}$ on the interval $(-1,1)$ and, hence, is a chart on $\mathbb{S}^1$ with the local coordinate $y^1 = x^1$.

The relations between the coordinates x^1, x^2 in $\mathbb{R}^2$ and y^1 in $\mathbb{S}^1$ are:

$$x^1 = y^1 \quad \text{and} \quad x^2 = \sqrt{1-(y^1)^2}.$$

It follows that

$$dx^1 = dy^1 \quad \text{and} \quad dx^2 = \frac{-y^1}{\sqrt{1-(y^1)^2}} dy^1.$$



The upper semi-circle

Since

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

we obtain that

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

2. Let us compute $\mathbf{g}_{\mathbb{S}^2}$ using on the upper semi-sphere

$$U = \mathbb{S}^2 \cap \{x_3 > 0\}$$

the coordinates $y^1 = x^1$ and $y^2 = x^2$. We have then

$$x^3 = \sqrt{1-(x^1)^2-(x^2)^2} = \sqrt{1-(y^1)^2-(y^2)^2}$$

whence

$$dx^1 = dy^1, \quad dx^2 = dy^2$$

and

$$dx^3 = -\frac{y^1}{\sqrt{1-(y^1)^2-(y^2)^2}} dy^1 - \frac{y^2}{\sqrt{1-(y^1)^2-(y^2)^2}} dy^2.$$

It follows that

$$\begin{aligned}
\mathbf{g}_{\mathbb{S}^2} &= (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \\
&= (dy^1)^2 + (dy^2)^2 + \frac{(y^1)^2}{1-(y^1)^2-(y^2)^2}(dy^1)^2 + \frac{(y^2)^2}{1-(y^1)^2-(y^2)^2}(dy^2)^2 \\
&+ \frac{2y^1y^2}{1-(y^1)^2-(y^2)^2} dy^1 dy^2 \\
&= \frac{\left(1-(y^2)^2\right)(dy^1)^2 + \left(1-(y^1)^2\right)(dy^2)^2 + 2y^1y^2 dy^1 dy^2}{1-(y^1)^2-(y^2)^2}
\end{aligned}$$

See Exercise 25 for computation of the metric $\mathbf{g}_{\mathbb{S}^1}$ and $\mathbf{g}_{\mathbb{S}^2}$ in the polar coordinates.

12-Nov-25 **Lecture 7**

1.8 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 family of subsets of X such that $\mathcal{A}$ contains $\emptyset$, X and $\mathcal{A}$ is closed under operations of taking complement and countable unions (hence, also intersections). That is, if $\{A_i\}_{i=1}^{\infty}$ is a sequence of elements of $\mathcal{A}$ then also $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ and $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$.

A *measure* μ on a σ-algebra $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(\emptyset) = 0$ and μ is σ-additive, that is,

$$\mu\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) \text{ for all } A_i \in \mathcal{A}.$$

Given a measure μ on a σ-algebra $\mathcal{A}$ on M, one can define the notion of the integral

$$\int_X f d\mu$$

for a class of *measurable* functions. A function $f : X \to \mathbb{R}$ is called measurable if the set $\{x \in X : f(x) \leq c\}$ belongs to $\mathcal{A}$ for any $c \in \mathbb{R}$. In particular, for any set $A \in \mathcal{A}$ the function $f = \mathbf{1}_A$ is measurable and

$$\int_X \mathbf{1}_A d\mu = \mu(A).$$

The class of measurable functions is closed under arithmetic operations of functions and pointwise limit.

The most important example of a measure is the Lebesgue measure λ defined on the *Lebesgue σ-algebra* $\mathcal{L}(\mathbb{R}^n)$ that consists 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 $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{L}(\mathbb{R}^n)$ and that any Lebesgue measurable set is a union of a Borel set and a null set (=a set of measure zero). Note also that the Lebesgue measure generalizes and unifies the notions of length in $\mathbb{R}$, area in $\mathbb{R}^2$ and volume in $\mathbb{R}^3$.

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* and $\mathcal{B}(M)$ is called the Borel σ-algebra on M.

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 $\mathcal{L}(M)$ and referred to as the Lebesgue σ-algebra on M.

Since any open subset of M is measurable, that is, $\mathcal{L}(M)$ contains all open subsets of M, and $\mathcal{B}(M)$ is the smallest σ-algebra containing all open subsets of M, it follows that $\mathcal{B}(M) \subset \mathcal{L}(M)$.

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

For any chart U on M with the local coordinates $x^1, ..., x^n$, consider the matrix $g^x = (g_{ij})$ where g_{ij} are the components of the metric $\mathbf{g}$ in coordinates $x^1, ..., x^n$.

Theorem 1.12 *For any Riemannian manifold $(M, \mathbf{g})$, there exists a unique measure ν on $\mathcal{L}(M)$ such that, in any chart U on M with coordinates $x^1, ..., x^n$,*

$$\boxed{d\nu = \sqrt{\det g^x}\, dx}, \tag{1.44}$$

where dx denotes 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 $\det g^x > 0$ because the matrix g^x is positive definite; hence, the square root in (1.44) is well defined. The identity (1.44) means that, for any non-negative measurable function f on U,

$$\boxed{\int_U f d\nu = \int_U f \sqrt{\det g^x} dx}, \tag{1.45}$$

where U in the right hand side is regarded as a subset of $\mathbb{R}^n$. In particular, for any measurable set $A \subset U$, we have

$$\nu(A) = \int_A \sqrt{\det g^x}\, dx. \tag{1.46}$$

Proof. We need to construct measure ν with the domain $\mathcal{L}(M)$ that satisfies (1.46) in any chart U. Let us use (1.46) as definition of ν on the σ-algebra $\mathcal{L}(U)$ of Lebesgue measurable sets in U. We need to show that the measure ν defined by (1.46) in each chart, can be extended to $\mathcal{L}(M)$ and, moreover, this extension is unique.

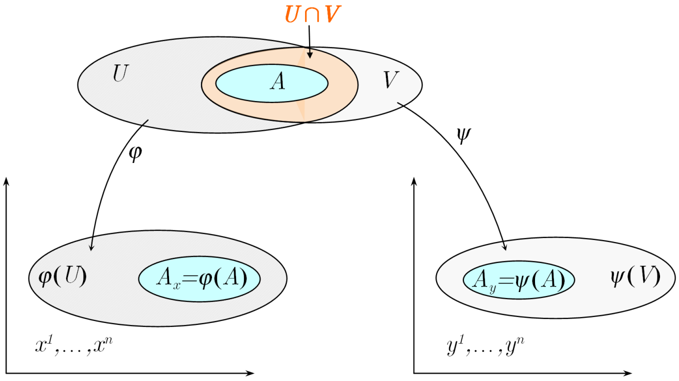
Step 1. Let us first prove that the measures that are defined by (1.46) in different charts, are compatible. That is, if U and V are two charts on M and A is a measurable set in $W := U \cap V$ then the integral in (1.46) has the same values in the both charts.

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 $A \subset W$,

$$\int_{A_x} \sqrt{\det g^x} dx = \int_{A_y} \sqrt{\det g^y} dy,$$

where dx and dy stand for the Lebesgue measures in U and V, respectively, A_x is the set A considered as a subset of U with coordinates $x^1, ..., x^n$, and A_y is the set A

considered as a subset of V with coordinates $y^1, ..., y^n$, respectively.



A set A in the intersection of two charts (U, φ) and (V, ψ).

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. (1.43)). By (1.42) we have

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

which implies

$$\det g^y = \det J^T \det g^x \det J = \det g^x (\det J)^2. \tag{1.47}$$

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

$$\int_{W_x} f(x)dx = \int_{W_y} f(x(y)) |\det J| \, dy. \tag{1.48}$$

Applying this for $f = 1_A \sqrt{\det g^x}$ and using (1.47), we obtain

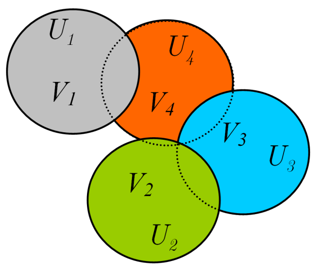
$$\int_{A_x} \sqrt{\det g^x} dx = \int_{A_y} \sqrt{\det g^x} |\det J| \, dy = \int_{A_y} \sqrt{\det g^x (\det J)^2} dy = \int_{A_y} \sqrt{\det g^y} dy,$$

which proves the claim.

Step 2. Let us prove that the measure ν on $\mathcal{L}(M)$ that satisfies (1.44) in all charts, is unique.

By Lemma 1.1, there is a countable family $\{U_i\}_{i=1}^{\infty}$ of relatively compact charts covering M and such that each $\overline{U}_i$ is contained in a chart. Consider the sets

$$V_1 = U_1$$
$$V_2 = U_2 \setminus U_1 = U_2 \cap U_1^c$$
$$V_3 = U_3 \setminus U_2 \setminus U_1 = U_3 \cap U_2^c \cap U_1^c$$
$$...$$
$$V_i = U_i \cap U_{i-1}^c \cap ... \cap U_1^c$$
$$...$$



Clearly, all the sets V_i are measurable and disjoint. Besides,

$$M = \bigsqcup_i V_i$$

because for any point $x \in M$ there is a minimal i such that $x \in U_i$ whence $x \in U_i \cap U_{i-1}^c ... \cap U_1^c = V_i$.

For any measurable set A on M, define the sets

$$A_i = A \cap V_i \tag{1.49}$$

Then we have $A_i \in \mathcal{L}(U_i)$ and $A = \bigsqcup_i A_i$.Therefore, for any measure ν, we should have

$$\nu(A) = \sum_i \nu(A_i). \tag{1.50}$$

However, the value $\nu(A_i)$ is uniquely determined by (1.44) because A_i is contained in the chart U_i. Hence, $\nu(A)$ is also uniquely defined, which was to be proved.

Step 3. Let us prove the existence of ν. For that fix some covering $\{U_i\}$ of M as above, and, for any measurable set A, define $\nu(A)$ by (1.50), 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 $\{B_k\}_{k=1}^{\infty}$ be a sequence of disjoint measurable sets in M such that

$$A = \bigsqcup_k B_k.$$

We need to prove that

$$\nu(A) = \sum_k \nu(B_k). \tag{1.51}$$

Defining the sets B_{ki} similarly to (1.49), that is, by

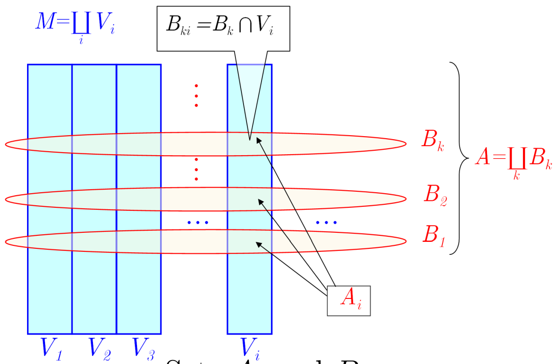
$$B_{ki} = B_k \cap V_i$$

we obtain that

$$B_k = \bigsqcup_i B_{ki}$$

as well as

$$A_i = A \cap V_i = \bigsqcup_k (B_k \cap V_i) = \bigsqcup_k B_{ki}.$$



Sets A_i and B_{ki}

Since ν is σ-additive in each chart U_i, we obtain

$$\nu(A_i) = \sum_k \nu(B_{ki})$$

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

$$\nu(A) \stackrel{\text{def}}{=} \sum_i \nu(A_i) = \sum_i \sum_k \nu(B_{ki}) = \sum_k \sum_i \nu(B_{ki}) \stackrel{\text{def}}{=} \sum_k \nu(B_k),$$

which proves (1.51). Here we have used the definition (1.50) twice: firstly, in the left hand side and, secondly, in the right hand side in the form

$$\nu(B_k) = \sum_i \nu(B_{ki}).$$

Step 4. Let $K \subset M$ be compact. By (1.50) we have

$$\nu(K) = \sum_i \nu(K_i), \tag{1.52}$$

where $K_i = K \cap V_i$. Since K is compact, it can be covered by a finite number of charts U_i so that the sum in (1.52) is finite. Since

$$\nu(K_i) = \int_{K_i} \sqrt{\det g^x}\, dx \le \int_{\overline{U}_i} \sqrt{\det g^x}\, dx,$$

$\overline{U}_i$ is a compact subset of a larger chart, and the function $\sqrt{\det g^x}$ is bounded on $\overline{U}_i$, we obtain $\nu(K_i) < \infty$ and, hence, $\nu(K) < \infty$.

Any non-empty open set $\Omega \subset M$ contains some non-empty chart U, whence it follows from (1.46) that

$$\nu(\Omega) \ge \nu(U) = \int_U \sqrt{\det g^x}\, dx > 0,$$

which finishes the proof. ∎

* **Alternative proof of existence of Riemannian measure.** 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 $\mathcal{L}(M)$. Since the union of sets U_i from Lemma 1.1 is M and $\nu(U_i) < \infty$, the measure ν on S is σ-finite and, hence, its extension to $\mathcal{L}(M)$ 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 1.13 *If $f \in C(M)$ and*

$$\int_M f\varphi\, d\nu = 0 \tag{1.53}$$

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

Proof. Assume that $f(x_0) \ne 0$ for some point $x_0 \in M$, say, $f(x_0) > 0$. Then, by the continuity of f, $f(x)$ is strictly positive in a open neighborhood Ω of x_0, that is, $f \ge c$ in Ω for some constant $c > 0$. Let φ be a cutoff function of $\{x_0\}$ in Ω. Then $\varphi \equiv 1$ in some open neighborhood U of x_0. Since $\nu(U) > 0$, it follows that

$$\int_M f\varphi\, d\nu = \int_\Omega f\varphi\, d\nu \ge \int_U f\, d\nu \ge c\nu(U) > 0,$$

which contradicts (1.53). ∎

1.9 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

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

Divergence satisfies the following identity any smooth vector field v in $\mathbb{R}^n$ and a smooth scalar function u with compact support in $\mathbb{R}^n$:

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

which can be deduced from the divergence theorem of Gauss. Alternatively, this identity is a consequence of Fubini's theorem and the integration by part formula: for all $w \in C^\infty(\mathbb{R}^n)$ and $u \in C_0^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} \frac{\partial w}{\partial x^i} u dx = -\int_{\mathbb{R}^n} w \frac{\partial u}{\partial x^i} dx \tag{1.54}$$

applied with $w = v^i$.

Recall that a vector field v on a manifold M is a collection $\{v(x)\}_{x\in M}$ of tangent vectors where $v(x) \in T_xM$ for any $x \in M$. 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 1.14 (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*

$$\boxed{\int_M (\operatorname{div} v)\, u\, d\nu = -\int_M \langle v, \nabla u\rangle d\nu}, \tag{1.55}$$

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

The identity (1.55) means a certain duality between the gradient and divergence.

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 (1.32), we have

$$\langle v, \nabla u\rangle_{\mathbf{g}} = \langle \nabla u, v\rangle_{\mathbf{g}} = \langle du, v\rangle = \frac{\partial u}{\partial x^i} v^i, \tag{1.56}$$

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

14-Nov-25 **Lecture 8**

Proof. The uniqueness of $\operatorname{div} v$ is simple: if there are two divergencies of v, say $(\operatorname{div} v)_1$ and $(\operatorname{div} v)_2$ then it follows from (1.55) that, for all $u \in C_0^\infty(M)$,

$$\int_M (\operatorname{div} v)_1 \, u \, d\nu = \int_M (\operatorname{div} v)_2 \, u \, d\nu.$$

Setting $f = (\operatorname{div} v)_1 - (\operatorname{div} v)_2$ and observing that $\int_M f u \, d\nu = 0$ for all $u \in C_0^\infty(M)$, we conclude by Lemma 1.13 that $f = 0$ on M and, hence, $(\operatorname{div} v)_1 = (\operatorname{div} v)_2$.

To prove the existence of $\operatorname{div} v$, let us first show that $\operatorname{div} v$ exists in any chart U, that is, there exists a smooth function $\operatorname{div} v$ in U such that (1.55) holds for all $u \in C_0^\infty(U)$. Let the coordinates in U be $x^1, ..., x^n$. Using (1.56), (1.45), and the integration-by-parts formula in U as a subset of $\mathbb{R}^n$, we obtain, for any $u \in C_0^\infty(U)$,

$$\begin{aligned}\int_U \langle v, \nabla u\rangle d\nu &= \int_U \frac{\partial u}{\partial x^i} v^i \sqrt{\det g^x} \, dx \\ &= -\int_U u \frac{\partial}{\partial x^i}\left(v^i \sqrt{\det g^x}\right) dx \\ &= -\int_U \frac{1}{\sqrt{\det g^x}} \frac{\partial}{\partial x^i}\left(v^i \sqrt{\det g^x}\right) u \, d\nu, \end{aligned} \tag{1.57}$$

where $g^x = (g_{ij})$ is the matrix of the metric $\mathbf{g}$ in U. Defining $\operatorname{div} v$ in U by

$$\boxed{\operatorname{div} v = \frac{1}{\sqrt{\det g^x}} \frac{\partial}{\partial x^i}\left(\sqrt{\det g^x} \, v^i\right)}, \tag{1.58}$$

we obtain the identity

$$\int_U \langle v, \nabla u\rangle d\nu = -\int_U (\operatorname{div} v) \, u \, d\nu \quad \text{for all } u \in C_0^\infty(U)$$

that coincides with (1.55). Hence, the formula (1.58) defines $\operatorname{div} v$ as a function in U.

Let us observe that if U and V are two charts then (1.58) defines two divergences: $(\operatorname{div} v)_1$ in U and $(\operatorname{div} v)_2$ in V, and these two divergencies coincide in the intersection $W = U \cap V$. Indeed, the both functions $(\operatorname{div} v)_1$ and $(\operatorname{div} v)_2$ satisfy the definition of the divergence in W for all $u \in C_0^\infty(W)$. Applying the statement about uniqueness of the divergence in W, we conclude that $(\operatorname{div} v)_1 = (\operatorname{div} v)_2$ in W. By the way, without using the uniqueness the proof of this fact is highly non-trivial.

Now let us define $\operatorname{div} v$ on the entire manifold M. Let us fix a family $\{U_\alpha\}$ of charts covering M. Defining by (1.58) the divergence $\operatorname{div} v$ in each chart U_α, we obtain that the function $\operatorname{div} v$ is well defined on the entire manifold M. Moreover, this function $\operatorname{div} v$ satisfies the identity (1.55), that is,

$$\int_M (\operatorname{div} v) \; u \, d\nu = -\int_M \langle v, \nabla u\rangle d\nu \tag{1.59}$$

for all functions $u \in C_0^\infty(M)$ that are *compactly supported in one of the charts* U_α.

We are left to extend the identity (1.59) to all functions $u \in C_0^\infty(M)$. By Corollary 1.4, any function $u \in C_0^\infty(M)$ can be represented as a finite sum $u_1 + ... + u_k$, where each u_j is smooth and compactly supported in one of the charts U_α. Hence, (1.59) holds for each of the functions u_j. By adding up all such identities for $j = 1, ..., k$, we obtain (1.59) for the function u. ■

It follows from (1.58) that

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

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

Corollary 1.15 *The identity (1.55) holds also if $u(x)$ is any smooth function on M and $v(x)$ is a compactly supported smooth vector field on M.*

Proof. Since $\operatorname{supp} v$ is compact, by Theorem 1.3, there exists a cutoff function of $\operatorname{supp} v$, that is, a function $\varphi \in C_0^\infty(M)$ such that $\varphi \equiv 1$ in a neighbourhood Ω of $\operatorname{supp} v$. Then $u\varphi \in C_0^\infty(M)$, and applying (1.55) with function $u\varphi$ instead of u, we obtain

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

■

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

$$\operatorname{div} v = \frac{1}{\sqrt{\det g^x}} \frac{\partial}{\partial x^i} \left(\sqrt{\det g^x} \, v^i \right), \tag{1.60}$$

and show by a direct computation that, in the intersection of any two charts, (1.60) 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 1.14 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

$$\frac{\partial}{\partial t} \ln \det a = b_k^l \frac{\partial a_l^k}{\partial t}. \tag{1.61}$$

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

$$I_i^k = \frac{\partial y^k}{\partial x^i} \text{ and } J_k^i = \frac{\partial x^i}{\partial y^k},$$

so that the matrices J and I are mutually inverse.

Let g^x be the matrix of the metric $\mathbf{g}$ in coordinates $x^1, ..., x^n$ and g^y be the matrix of $\mathbf{g}$ in coordinates $y^1, ..., y^n$. Let v_x^i be the components of the vector v in coordinates $x^1, ..., x^n$, and v_y^k be the components of the vector v in coordinates $y^1, ..., y^n$. Then we have

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

so that

$$v_y^k = v_x^i I_i^k.$$

Since by (1.47)

$$\sqrt{\det g^y} = \sqrt{\det g^x} \, |\det J| = \sqrt{\det g^x} \, |\det I|^{-1},$$

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

$$\begin{aligned}\operatorname{div} v &= \frac{1}{\sqrt{\det g^y}} \frac{\partial}{\partial y^k} \left(\sqrt{\det g^y} v_y^k \right) = \frac{\det I}{\sqrt{\det g^x}} J_k^j \frac{\partial}{\partial x^j} \left(\sqrt{\det g^x} v_x^i (\det I)^{-1} I_i^k \right) \\ &= \frac{1}{\sqrt{\det g^x}} \frac{\partial}{\partial x^j} \left(\sqrt{\det g^x} v_x^i \right) J_k^j I_i^k + v_x^i J_k^j I_i^k \det I \frac{\partial}{\partial x^j} (\det I)^{-1} + v_x^i J_k^j \frac{\partial I_i^k}{\partial x^j} \\ &= \frac{1}{\sqrt{\det g^x}} \frac{\partial}{\partial x^i} \left(\sqrt{\det g^x} v_x^i \right) - v_x^i \frac{\partial}{\partial x^i} \ln \det I + v_x^i J_k^j \frac{\partial I_i^k}{\partial x^j},\end{aligned}$$

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

$$-\frac{\partial}{\partial x^i} \ln \det I + J_k^j \frac{\partial I_i^k}{\partial x^j} = 0. \tag{1.62}$$

By (1.61), we have

$$\frac{\partial}{\partial x^i} \ln \det I = J_k^j \frac{\partial I_j^k}{\partial x^i}.$$

Noticing that

$$\frac{\partial I_j^k}{\partial x^i} = \frac{\partial^2 y^k}{\partial x^j \partial x^i} = \frac{\partial^2 y^k}{\partial x^i \partial x^j} = \frac{\partial I_i^k}{\partial x^j},$$

we obtain (1.62).

1.10 Laplace-Beltrami operator

Recall that the Laplace operator in $\mathbb{R}^n$ is defined on all functions $f \in C^\infty(\mathbb{R}^n)$ by

$$\Delta f = \sum_{i=1}^n \frac{\partial^2 f}{(\partial x^i)^2}. \tag{1.63}$$

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* Δ (frequently referred to simply as the Laplace operator) on any Riemannian manifold $(M, \mathbf{g})$ as follows: for any smooth function f on M, set

$$\boxed{\Delta f = \operatorname{div} (\nabla f)}, \tag{1.64}$$

that is, $\Delta = \operatorname{div} \circ \nabla$. 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 omitted when the metric $\mathbf{g}$ is fixed.

Clearly, 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 $(g^{ij}) = (g_{ij})^{-1}$. Combining with the formula (1.58) for divergence, we obtain

$$\boxed{\Delta f = \frac{1}{\sqrt{\det g^x}} \frac{\partial}{\partial x^i} \left(\sqrt{\det g^x} g^{ij} \frac{\partial f}{\partial x^j} \right)}, \tag{1.65}$$

For example, if $(g_{ij}) \equiv \mathrm{id}$ then also $(g^{ij}) \equiv \mathrm{id}$, and (1.65) takes the form (1.63). 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 (1.65) is a *second order elliptic differential operator*. One can rewrite (1.65) also as follows:

$$\Delta f = g^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j} + \frac{1}{\sqrt{\det g^x}} \frac{\partial}{\partial x^i} (\sqrt{\det g^x} g^{ij}) \frac{\partial f}{\partial x^j}.$$

Proposition 1.16 (The Green formula) *If u and v are smooth functions on a Riemannian manifold M and one of them has a compact support then*

$$\int_M u \, \Delta v \, d\nu = -\int_M \langle \nabla u, \nabla v \rangle d\nu = \int_M v \, \Delta u \, d\nu. \tag{1.66}$$

Proof. By symmetry it suffices to prove the left hand side identity in (1.66). By hypothesis either $\operatorname{supp} u$ or $\operatorname{supp} \nabla v$ is compact. By (1.64), Theorem 1.14 and Corollary 1.15, we obtain

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

which was to be proved. ■

1.11 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

$$d\mu = D d\nu \tag{1.67}$$

and defined on the σ-algebra $\mathcal{L}(M)$. The function D is called the *density function* of the measure μ.

The identity (1.67) means that, for any measurable set $A \subset M$,

$$\mu(A) = \int_A D \, d\nu$$

and, for any non-negative measurable function f on M,

$$\int_M f \, d\mu = \int_M f D \, d\nu.$$

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

$$\boxed{\operatorname{div}_{\mathbf{g},\mu} v = \frac{1}{D} \operatorname{div}_{\mathbf{g}} (Dv)}. \tag{1.68}$$

It follows immediately from this definition and (1.55) that the following extension of Theorem 1.14 takes place: for all smooth vector fields v and functions u,

$$\int_M (\operatorname{div}_{\mathbf{g},\mu} v) u \, d\mu = -\int_M \langle v, \nabla u \rangle_{\mathbf{g}} d\mu, \tag{1.69}$$

provided v or u has a compact support. Indeed, using (1.55) and (1.68), we obtain

$$\int_M \operatorname{div}_{\mathbf{g},\mu} v \, u \, d\mu = \int_M \frac{1}{D} \operatorname{div}_{\mathbf{g}} (Dv) \; u \, D d\nu = -\int_M \langle Dv, \nabla u \rangle_{\mathbf{g}} d\nu = -\int_M \langle v, \nabla u \rangle_{\mathbf{g}} d\mu.$$

Define the *weighted Laplace operator* $\Delta_{\mathbf{g},\mu}$ on all smooth functions u on M by

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

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

$$\int_M u \, \Delta_{\mathbf{g},\mu} v \, d\mu = -\int_M \langle \nabla u, \nabla v \rangle_{\mathbf{g}} \, d\mu = \int_M v \, \Delta_{\mathbf{g},\mu} u \, d\mu. \tag{1.70}$$

In the local coordinates $x^1, ..., x^n$ in a chart U, we have

$$d\mu = D d\nu = D \sqrt{\det g^x} dx = \rho dx,$$

where $\rho = D\sqrt{\det g^x}$ and dx is the Lebesgue measure in U. It follows from (1.58) and (1.68) that

$$\operatorname{div}_{\mathbf{g},\mu} v = \frac{1}{D\sqrt{\det g^x}} \frac{\partial}{\partial x^i} \left(\sqrt{\det g^x} \, D v^i \right) = \frac{1}{\rho} \frac{\partial}{\partial x^i} \left(\rho v^i \right) \tag{1.71}$$

and

$$\Delta_{\mathbf{g},\mu} f = \frac{1}{\rho} \frac{\partial}{\partial x^i} \left(\rho g^{ij} \frac{\partial f}{\partial x^j} \right). \tag{1.72}$$

Sometimes is it useful to know that the right hand side of (1.72) can be expanded as follows:

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

Example. Consider the weighted manifold $(\mathbb{R}^n, \mathbf{g}, \mu)$ where $\mathbf{g} = \mathbf{g}_{\mathbb{R}^n}$ is the canonical Euclidean metric and $d\mu = D d\nu = D d\lambda$. Then by (1.72)

$$\Delta_{\mathbf{g},\mu} f = \frac{1}{D} \frac{\partial}{\partial x^i} \left(D \frac{\partial f}{\partial x^i} \right) = \Delta f + \frac{1}{D} \langle \nabla D, \nabla f \rangle.$$

Let $(M, \mathbf{g}, \mu)$ be a weighted manifold and D be the density function of measure μ. On any submanifold S of M, define the *induced measure* μ_S by

$$d\mu_S = D|_S \, d\nu_S$$

where ν_S is the Riemannian measure on S of the induced metric $\mathbf{g}_S$. Hence, we obtain a weighted manifold $(S, \mathbf{g}_S, \mu_S)$.

1.12 Product of manifolds

Product of smooth 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 a base of topology in M is given by $U \times V$ where U is any open subset of X and V is any open subset of Y. It is easy to see that M is a Hausdorff topological space with countable base.

Besides, M 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 of dimension $n + m$.

Lemma 1.17 *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, that is,*

$$T_{(x_0,y_0)}M = T_{x_0}X \oplus T_{y_0}Y. \tag{1.73}$$

Proof. Any $\mathbb{R}$-differentiation $\xi \in T_{x_0}X$ can be considered as an $\mathbb{R}$-differentiation on smooth functions $f(x, y)$ on M at (x_0, y_0) 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 in the basis

$$\left\{ \frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}, \frac{\partial}{\partial y^1}, ..., \frac{\partial}{\partial y^m} \right\}$$

the components $\xi^1, ..., \xi^{n+m}$ so that:

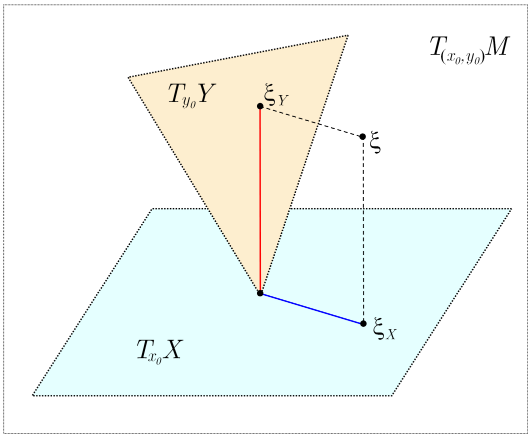
$$\xi = \sum_{i=1}^{n} \xi^i \frac{\partial}{\partial x^i} + \sum_{j=1}^{m} \xi^{n+j} \frac{\partial}{\partial y^j}.$$

Setting

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

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

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



Such a decomposition is obviously unique, whence (1.73) follows. ■

19-Nov-25 Lecture 9

Riemannian product. Let $\mathbf{g}_X$ and $\mathbf{g}_Y$ be Riemannian metric tensors on the manifolds X and Y, respectively. Define the Riemannian metric tensor $\mathbf{g}$ on $M = X \times Y$ as as follows: for any two tangent vectors $\xi, \eta \in T_{(x_0,y_0)}M$, set

$$\langle \xi, \eta \rangle_{\mathbf{g}(x_0,y_0)} = \langle \xi_X, \eta_X \rangle_{\mathbf{g}_X(x_0)} + \langle \xi_Y, \eta_Y \rangle_{\mathbf{g}_Y(y_0)} . \tag{1.74}$$

In particular, we have $|\xi|^2_{\mathbf{g}(x_0,y_0)} = |\xi_X|^2_{\mathbf{g}_X(x_0)} + |\xi_Y|^2_{\mathbf{g}_Y(y_0)}$ (a version of Pythagorean equation).

Definition. The metric $\mathbf{g}$ defined by (1.74) is called the *direct sum* of $\mathbf{g}_X$ and $\mathbf{g}_Y$ and is denoted by

$$\mathbf{g} = \mathbf{g}_X + \mathbf{g}_Y. \tag{1.75}$$

The Riemannian manifold $(M, \mathbf{g})$, where $M = X \times Y$ and $\mathbf{g}$ is given by (1.75), is called the *Riemannian* (or *direct*) *product* of $(X, \mathbf{g}_X)$ and $(Y, \mathbf{g}_Y)$.

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

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

so that the matrix of $\mathbf{g}$ has the form

$$g = \begin{pmatrix} \boxed{g_X} & \mathbf{0} \\ \mathbf{0} & \boxed{g_Y} \end{pmatrix}. \tag{1.76}$$

Measure on a Riemannian product. 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 the product set $M = X \times Y$ a *product measure*

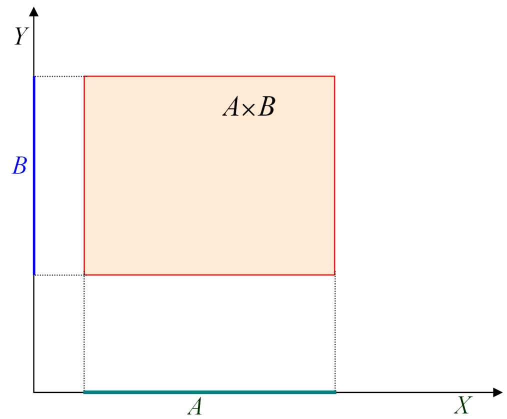
$$\mu = \mu_1 \times \mu_2$$

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 $\mathcal{A}$ on M by means of the Carathéodory extension theorem.



For example, this way one constructs the Lebesgue measure on $\mathbb{R}^n$: first construct it on $\mathbb{R}$ by extending the notion of length of intervals to $\mathcal{L}(\mathbb{R})$ and then use the notion of product of measures to define the Lebesgue measure on $\mathcal{L}(\mathbb{R}^n)$ by induction in n.

One of the most important properties of the product measure μ is the Fubini theorem: if $f(x, y)$ is a non-negative measurable function on M then

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

One writes these identities shortly as follows:

$$d\mu = d\mu_1(x) d\mu_2(y) = d\mu_2(y) d\mu_1(x).$$

Lemma 1.18 *Let $(X, \mathbf{g}_X)$ and $(Y, \mathbf{g}_Y)$ be Riemannian manifolds with the Riemannian measures ν_X and ν_Y, respectively, Let $(M, \mathbf{g})$ be the Riemannian product of $(X, \mathbf{g}_X)$ and $(Y, \mathbf{g}_Y)$, that is, $M = X \times Y$ and $\mathbf{g} = \mathbf{g}_X + \mathbf{g}_Y$. Then the Riemannian measure ν of $(M, \mathbf{g})$ is the product of ν_X and ν_Y, that is*

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

In other words, the operation of direct sum of Riemannian metrics is compatible with the operation of product of measures.

Proof. Let U be a chart on X with coordinates $x^1, ..., x^n$ and V be a chart in Y with coordinates $y^1, ..., y^m$. Then 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 U and V, respectively. Let λ be the Lebesgue measure in the chart $U \times V$, so that $d\lambda = dx\, dy$. Observe that (1.76) implies

$$\det g = \det g_X \det g_Y. \tag{1.77}$$

Then the Riemannian measure ν of M is given by

$$d\nu = \sqrt{\det g}\, d\lambda = \sqrt{\det g_X}\sqrt{\det g_Y}\, dx\, dy = d\nu_X d\nu_Y,$$

which was to be proved. ■

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

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

Laplace operator on the product. We continue using setup and notation of Lemma 1.18.

Lemma 1.19 *We have*

$$\boxed{\Delta_{\mathbf{g}} = \Delta_{\mathbf{g}_X} + \Delta_{\mathbf{g}_Y}}, \tag{1.78}$$

that is, for any $f \in C^\infty(M)$,

$$\Delta_{\mathbf{g}} f(x,y) = \Delta_{\mathbf{g}_X} f(x,y) + \Delta_{\mathbf{g}_Y} f(x,y),$$

where $\Delta_{\mathbf{g}_X}$ *acts on the variable* x *and* $\Delta_{\mathbf{g}_Y}$ *– on the variable* y.

Proof. It follows from (1.76) that a similar identity holds for the inverse matrices:

$$g^{-1} = \begin{pmatrix} \boxed{g_X^{-1}} & \mathbf{0} \\ \mathbf{0} & \boxed{g_Y^{-1}} \end{pmatrix}.$$

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 $(M, \mathbf{g})$:

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

where we have used (1.77) and the fact that $\det g_X$ depends only in x while $\det g_Y$ depends only on y. ■

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

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

Warped product of Riemannian manifolds. There are other possibilities to define a Riemannian metric $\mathbf{g}$ 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

$$\mathbf{g} = \mathbf{g}_X + \psi^2(x)\,\mathbf{g}_Y. \tag{1.79}$$

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

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

Product of weighted manifolds. Let $(X, \mathbf{g}_X, \mu_X)$ and $(Y, \mathbf{g}_Y, \mu_Y)$ be weighted manifold. Setting

$$M = X \times Y, \quad \mathbf{g} = \mathbf{g}_X + \mathbf{g}_Y, \quad \mu = \mu_X \times \mu_Y,$$

we obtain a weighted manifold $(M, \mathbf{g}, \mu)$ that is called 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

$$D(x, y) = D_X(x)\,D_Y(y).$$

Similarly to (1.78) we obtain that

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

1.13 Polar coordinates

1.13.1 Polar coordinates in $\mathbb{R}^n$

Every point $x \in \mathbb{R}^n \setminus \{0\}$ 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) with $r > 0$ and $\theta \in \mathbb{S}^{n-1}$ determines $x \in \mathbb{R}^n \setminus \{0\}$ uniquely by $x = r\theta$.

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

$$\theta^1, ..., \theta^{n-1}.$$

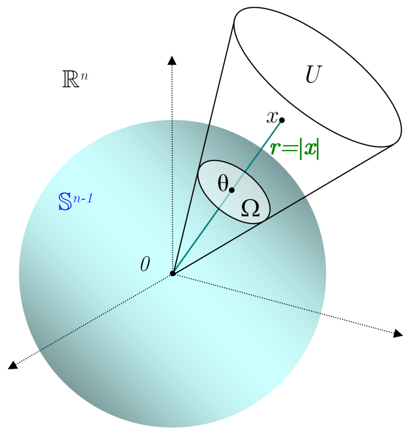
Then the set

$$U = \{x \in \mathbb{R}^n : r > 0, \ \theta \in \Omega\}$$

is a chart in $\mathbb{R}^n$ with coordinates

$$r, \theta^1, ..., \theta^{n-1}$$

that are called the *local polar coordinates*.



A chart U in $\mathbb{R}^n$

Proposition 1.20 *The canonical Euclidean metric* $\mathbf{g}_{\mathbb{R}^n}$ *in* $\mathbb{R}^n \{0\}$ *has the following representation in the local polar coordinates* $r, \theta^1, ..., \theta^{n-1}$*:*

$$\mathbf{g}_{\mathbb{R}^n} = dr^2 + r^2 \gamma_{ij} d\theta^i d\theta^j, \tag{1.80}$$

where γ_{ij} *are smooth functions of* $\theta^1, ..., \theta^{n-1}$*. Besides, the induced metric* $\mathbf{g}_{\mathbb{S}^{n-1}}$ *on* $\mathbb{S}^{n-1}$ *is given by*

$$\mathbf{g}_{\mathbb{S}^{n-1}} = \gamma_{ij} d\theta^i d\theta^j, \tag{1.81}$$

with the same functions γ_{ij}*. Consequently, we have*

$$\mathbf{g}_{\mathbb{R}^n} = dr^2 + r^2 \mathbf{g}_{\mathbb{S}^{n-1}}. \tag{1.82}$$

We see that $\mathbb{R}^n \setminus \{0\}$ can be regarded as a warped product of $(0, +\infty)$ and $\mathbb{S}^{n-1}$.

The identity (1.82) implies the following relation between the matrices $g_{\mathbb{R}^n}$ and $g_{\mathbb{S}^{n-1}}$:

$$g_{\mathbb{R}^n} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \boxed{r^2 g_{\mathbb{S}^{n-1}}} \end{pmatrix}.$$

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

$$x^i = r f^i \left(\theta^1, ..., \theta^{n-1}\right), \quad i = 1, ..., n, \tag{1.83}$$

where f^i is the i-th Cartesian 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

$$\left(f^1\right)^2 + ... + \left(f^n\right)^2 \equiv 1 \tag{1.84}$$

as $|\theta| = 1$. Considering x^i, r and f^i as functions in the chart U and using the product rule for d ($d(uv) = udv + vdu$, Exercise 21), we obtain from $x^i = rf^i$ that

$$dx^i = d\left(rf^i\right) = f^i dr + rdf^i.$$

It follows that

$$\begin{aligned}\left(dx^i\right)^2 &= \left(f^i dr\right)^2 + \left(f^i dr\right)\left(r df^i\right) + \left(r df^i\right)\left(f^i dr\right) + \left(r df^i\right)^2 \\ &= \left(f^i\right)^2 dr^2 + (r dr)\left(f^i df^i\right) + \left(f^i df^i\right)(r dr) + r^2\left(df^i\right)^2. \end{aligned} \tag{1.85}$$

Applying d to the identity (1.84), we obtain

$$\sum_i f^i df^i = 0. \tag{1.86}$$

Adding up the identities (1.85) for all $i = 1, ..., n$ and using (1.84) and (1.86), 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

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

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

which implies

$$\sum_i \left(df^i\right)^2 = \gamma_{jk} d\theta^j d\theta^k, \tag{1.87}$$

where

$$\gamma_{jk} = \sum_{i=1}^n \frac{\partial f^i}{\partial \theta^j}\frac{\partial f^i}{\partial \theta^k} \tag{1.88}$$

are smooth functions of $\theta^1, ..., \theta^{n-1}$. Hence, we have proved the identity (1.80).

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 (1.80) $r = 1$ and $dr = 0$, we obtain (1.81). Finally, (1.82) follows obviously from (1.80) and (1.81). ■

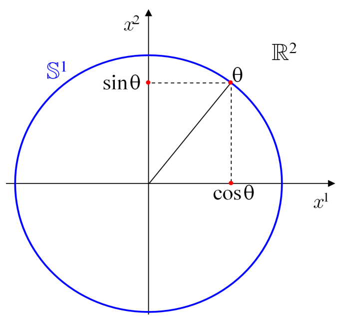
Example. Consider the case $n = 2$, that is, the polar coordinates in $\mathbb{R}^2$. Then the polar angle $\theta \in \mathbb{S}^1$ of a point $x \in \mathbb{R}^2 \setminus \{0\}$ can be identified with its angular measure $\theta \in (-\pi, \pi)$ provided

$$x \in \mathbb{R}^2 \setminus \left\{x : x^2 = 0, x^1 \le 0\right\}.$$

We have in this case

$$f^1(\theta) = \cos\theta \quad \text{and} \quad f^2(\theta) = \sin\theta,$$

which yields by (1.88)



$$\gamma_{11} = \left(\frac{\partial f^1}{\partial \theta}\right)^2 + \left(\frac{\partial f^2}{\partial \theta}\right)^2 = 1 \text{ and, hence, } \mathbf{g}_{\mathbb{S}^1} = d\theta^2 \text{ and } \mathbf{g}_{\mathbb{R}^2} = dr^2 + r^2 d\theta^2.$$

1.13.2 Polar coordinates in $\mathbb{S}^n$

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

$$x' := \left(x^1, ..., x^n \right) \in \mathbb{R}^n,$$

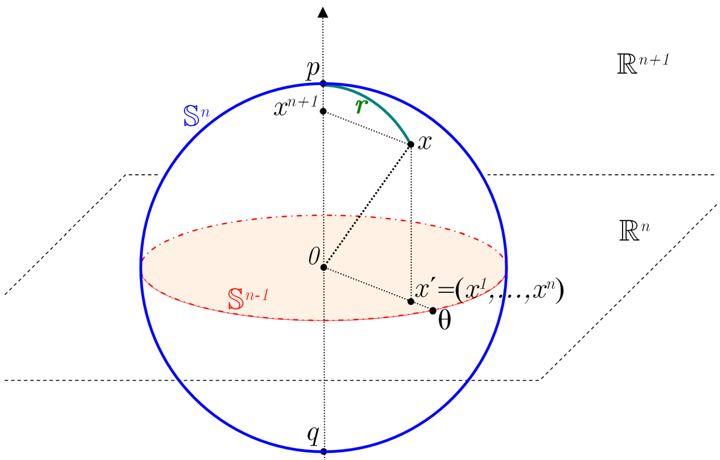
that is, x' is the projection of x onto

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

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 the formulas



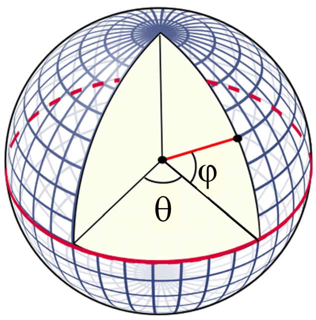
Polar coordinates on $\mathbb{S}^n$

$$\cos r = x^{n+1} \quad \text{and} \quad \theta = \frac{x'}{|x'|}. \tag{1.89}$$

Since $x^{n+1} = p \cdot x$, we have $\cos r = p \cdot x$ so that $r = \arccos(p \cdot x)$ is the angle between the unit vectors x and p. Hence, r is the length of the arc between x and p on the unit circle with the center at the origin that goes through 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.

Here are geographical latitude φ and longitude θ on the Earth regarded as $\mathbb{S}^2$. Using our notation, we have $r = \frac{\pi}{2} - \varphi$, where the latitude φ is measured from the equator $\mathbb{S}^1$. The longitude θ can be considered as a point on $\mathbb{S}^1$.

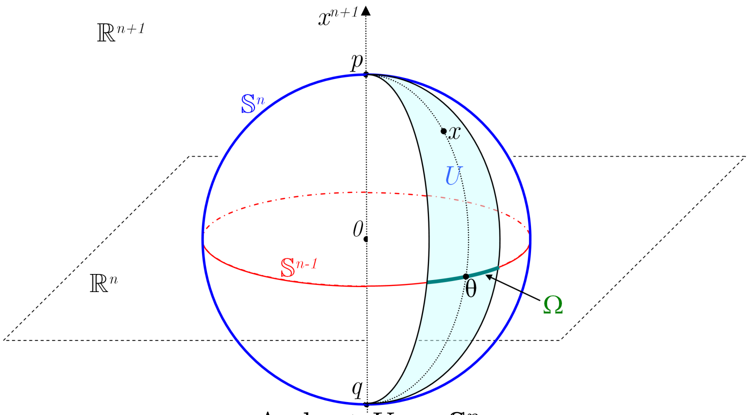


As in the case of the polar coordinates in the Euclidean space, the polar coordinates

(r,θ) on $\mathbb{S}^n$ can be regarded as local coordinates $r, \theta^1, ..., \theta^{n-1}$ in a chart

$$U = \{x \in \mathbb{S}^n : r \in (0,\pi), \ \theta \in \Omega\},$$

where Ω is any chart on $\mathbb{S}^{n-1}$ with the local coordinates $\theta^1, ..., \theta^{n-1}$.



A chart U on $\mathbb{S}^n$

Proposition 1.21 *The canonical spherical metric $\mathbf{g}_{\mathbb{S}^n}$ has in $\mathbb{S}^n \setminus \{p,q\}$ the following representation in the polar coordinates:*

$$\mathbf{g}_{\mathbb{S}^n} = dr^2 + \sin^2 r \, \mathbf{g}_{\mathbb{S}^{n-1}}. \tag{1.90}$$

We see that $\mathbb{S}^n \setminus \{p,q\}$ can be regarded as a warped product of $(0,\pi)$ and $\mathbb{S}^{n-1}$.

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 $x \in \mathbb{S}^n \setminus \{p,q\}$ 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 (1.83). It follows that

$$\begin{aligned} dx^i &= f^i \cos r dr + \sin r df^i, \quad i = 1, ..., n \\ dx^{n+1} &= -\sin r dr \end{aligned}$$

Hence, using (1.84), (1.86), and (1.87), we obtain

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

Since $\gamma_{ij} d\theta^i d\theta^j$ is the canonical metric on $\mathbb{S}^{n-1}$, we obtain (1.90). ■

21-Nov-25 Lecture 10

Recall that we have introduced in $\mathbb{R}^n$ the polar coordinates (r, θ) where $r \in (0, +\infty)$ and $\theta \in \mathbb{S}^{n-1}$, and proved that

$$\boxed{\mathbf{g}_{\mathbb{R}^n} = dr^2 + r^2\, \mathbf{g}_{\mathbb{S}^{n-1}}}. \tag{1.91}$$

Also, we have introduced in $\mathbb{S}^n$ the the polar coordinates (r, θ) where $r \in (0, \pi)$ and $\theta \in \mathbb{S}^{n-1}$, and proved that

$$\boxed{\mathbf{g}_{\mathbb{S}^n} = dr^2 + \sin^2 r\, \mathbf{g}_{\mathbb{S}^{n-1}}}. \tag{1.92}$$

These formulas allow to compute inductively the Riemannian metric on $\mathbb{S}^n$ and $\mathbb{R}^n$ in local coordinates.

It follows from (1.91) and (1.92) that the matrices of the corresponding Riemannian metrics are given by

$$g_{\mathbb{R}^n} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & r^2 \boxed{g_{\mathbb{S}^{n-1}}} \end{pmatrix}$$

and

$$g_{\mathbb{S}^n} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \sin^2 r \boxed{g_{\mathbb{S}^{n-1}}} \end{pmatrix}$$

where the boxed matrices are $(n-1) \times (n-1)$. Hence,

$$\sqrt{\det g_{\mathbb{R}^n}} = r^{n-1} \sqrt{\det g_{\mathbb{S}^{n-1}}} \quad \text{and} \quad \sqrt{\det g_{\mathbb{S}^n}} = \sin^{n-1} r \sqrt{\det g_{\mathbb{S}^{n-1}}}.$$

Denote by λ_n the Riemannian measure in $\mathbb{R}^n$ associated with the canonical Riemannian metric $\mathbf{g}_{\mathbb{R}^n}$ (that is, λ_n is the Lebesgue measure). Denote by σ_n the Riemannian measure on $\mathbb{S}^n$ associated with the canonical Riemannian metric $\mathbf{g}_{\mathbb{S}^n}$. Using the general formula for representation of a Riemannian measure $\nu_{\mathbf{g}}$ of a metric $\mathbf{g}$ in local coordinates $x^1, ..., x^n$

$$d\nu_{\mathbf{g}} = \sqrt{\det g}\, dx = \sqrt{\det g}\, dx^1 ... dx^n,$$

we obtain that λ_n and σ_n have the following representations in the polar coordinates:

$$\boxed{d\lambda_n = r^{n-1} dr\, d\sigma_{n-1}}$$

and

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

Example. Let (r, θ) be the polar coordinates in $\mathbb{R}^2$. We have already seen that

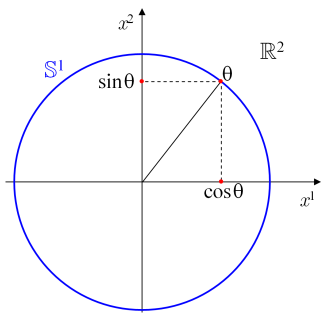
$$\mathbf{g}_{\mathbb{S}^1} = d\theta^2.$$

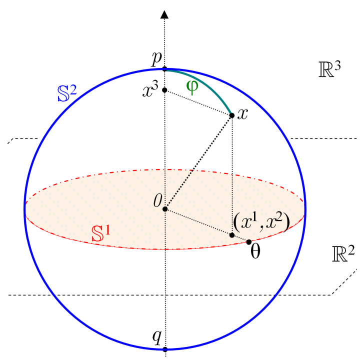
Hence, we obtain from (1.91) that

$$\mathbf{g}_{\mathbb{R}^2} = dr^2 + r^2 d\theta^2.$$

Let us compute $\mathbf{g}_{\mathbb{S}^2}$ in the polar coordinates (φ, θ) where φ is the polar radius on $\mathbb{S}^2$ and θ is the polar angle. By (1.92) we have

$$\mathbf{g}_{\mathbb{S}^2} = d\varphi^2 + \sin^2 \varphi\, d\theta^2.$$





Consequently, we have

$$d\sigma_1 = d\theta,$$
$$d\lambda_2 = r\,dr\,d\sigma_1 = r\,dr\,d\theta$$

and

$$d\sigma_2 = \sin\varphi\,d\varphi\,d\sigma_1 = \sin\varphi\,d\varphi\,d\theta.$$

For example, for a circle

$$B_R = \left\{x \in \mathbb{R}^2 : |x| < R\right\}$$

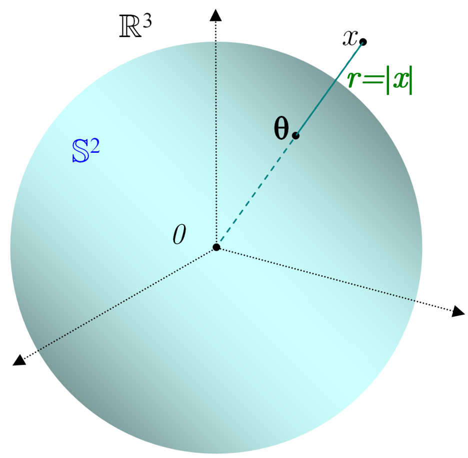
in $\mathbb{R}^2$ of radius R we have

$$\begin{aligned}\lambda_2(B_R) &= \int_{B_R} d\lambda_2 = \int_0^R \int_0^{2\pi} r\,dr\,d\theta \\ &= \left(\int_0^R r dr\right) \int_0^{2\pi} d\theta = \frac{R^2}{2} \cdot 2\pi = \pi R^2.\end{aligned}$$

Similarly we compute the Riemannian measure of $\mathbb{S}^2$:

$$\begin{aligned}\sigma_2(\mathbb{S}^2) &= \int_{\mathbb{S}^2} d\sigma_2 = \int_0^\pi \int_0^{2\pi} \sin\varphi\,d\varphi\,d\theta \\ &= \left(\int_0^\pi \sin\varphi\,d\varphi\right) \int_0^{2\pi} d\theta = 2 \cdot 2\pi = 4\pi.\end{aligned}$$

Example. Consider now the polar coordinates $(r, \boldsymbol{\theta})$ in $\mathbb{R}^3$ where r is the polar radius and $\boldsymbol{\theta} \in \mathbb{S}^2$. Expressing $\boldsymbol{\theta}$ in the spherical polar coordinates (φ, θ) of $\mathbb{S}^2$, we obtain the local polar coordinates (r, φ, θ) in $\mathbb{R}^3$ and



$$\begin{aligned}\mathbf{g}_{\mathbb{R}^3} &= dr^2 + r^2 \mathbf{g}_{\mathbb{S}^2} \\ &= dr^2 + r^2 d\varphi^2 + r^2 \sin^2\varphi\,d\theta^2.\end{aligned}$$

Consequently, we have

$$d\lambda_3 = r^2 dr\,d\sigma_2 = r^2 \sin\varphi\,dr\,d\varphi\,d\theta.$$

For example, let us compute the volume of a ball

$$B_R = \left\{x \in \mathbb{R}^3 : |x| < R\right\}$$

in $\mathbb{R}^3$ of radius R:

$$\begin{aligned}\lambda_3(B_R) &= \int_{B_R} d\lambda_3 = \int_0^R \int_0^\pi \int_0^{2\pi} r^2 \sin\varphi\,dr\,d\varphi\,d\theta \\ &= \left(\int_0^R r^2 dr\right) \left(\int_0^\pi \sin\varphi d\varphi\right) \int_0^{2\pi} d\theta \\ &= \frac{R^3}{3} \cdot 2 \cdot 2\pi = \frac{4\pi}{3} R^3.\end{aligned}$$

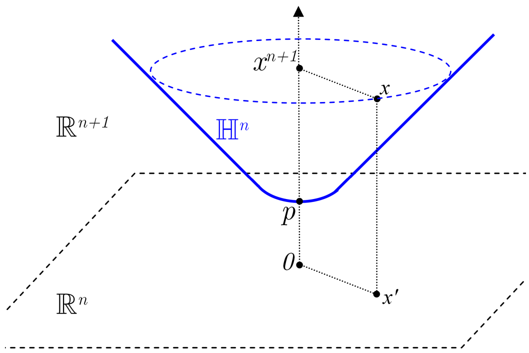
1.13.3 Polar coordinates in the hyperbolic space $\mathbb{H}^n$

Let us define *hyperbolic space* $\mathbb{H}^n$, $n \geq 2$, as follows. For that, consider in $\mathbb{R}^{n+1}$ a *semi-hyperboloid* H given by the equation

$$(x^{n+1})^2 - |x'|^2 = 1, \quad x^{n+1} > 0,$$

where as above $x' = (x^1, ..., x^n) \in \mathbb{R}^n$.

(For comparison, the equation of $\mathbb{S}^n$ is $(x^{n+1})^2 + |x'|^2 = 1$.)



By Lemma 1.7, H is a submanifold of $\mathbb{R}^{n+1}$ of dimension n.

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

$$\mathbf{g}_{Mink} = \left(dx^1\right)^2 + ... + (dx^n)^2 - \left(dx^{n+1}\right)^2, \tag{1.93}$$

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; it is 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.

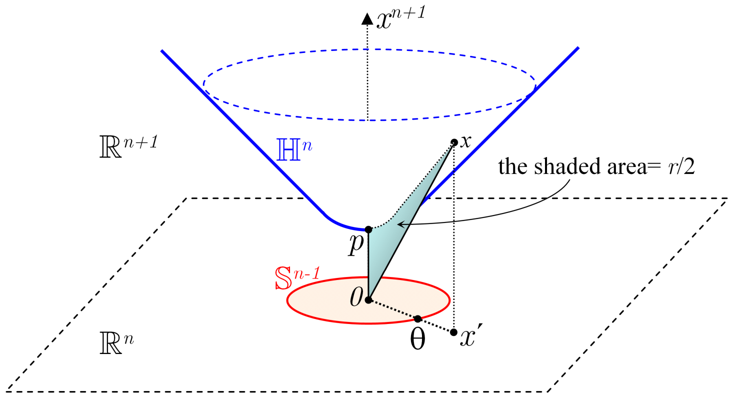
Definition. The manifold $(H, \mathbf{g}_H)$ is called the *hyperbolic space* of dimension n and is denoted by $\mathbb{H}^n$. The metric $\mathbf{g}_H$ is called the *canonical hyperbolic metric* and is denoted also by $\mathbf{g}_{\mathbb{H}^n}$.

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

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

Since $x^{n+1} > 1$, the equation $\cosh r = x^{n+1}$ has a unique



positive solution $r = \cosh^{-1} x^{n+1}$. 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 $r/2$.

Proposition 1.22 *The canonical hyperbolic metric* $\mathbf{g}_{\mathbb{H}^n}$ *has the following representation in the polar coordinates:*

$$\mathbf{g}_{\mathbb{H}^n} = dr^2 + \sinh^2 r\, \mathbf{g}_{\mathbb{S}^{n-1}}. \tag{1.94}$$

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

Proof. Let $\theta^1, ..., \theta^{n-1}$ be local coordinates on $\mathbb{S}^{n-1}$. Then $r, \theta^1, ..., \theta^{n-1}$ are the local coordinates on $\mathbb{H}^n$. Let us write down the metric $\mathbf{g}_{\mathbb{H}^n}$ in these coordinates. 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

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

Denoting by $f^i(\theta^1, ..., \theta^{n-1})$ the i-th Cartesian coordinate in $\mathbb{R}^n$ of $\theta \in \mathbb{S}^{n-1}$ (as in the proof of Proposition 1.20), we obtain that the Cartesian coordinates $x^1, ..., x^{n+1}$ of the point $x \in \mathbb{H}^n \setminus \{p\}$ admit the following representation:

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

Hence, we have, for $i = 1, ..., n$,

$$dx^i = f^i \cosh r\, dr + \sinh r\, df^i$$

and

$$dx^{n+1} = \sinh r\, dr.$$

It follows that

$$\begin{aligned} \mathbf{g}_{\mathbb{H}^n} &= \mathbf{g}_{Mink}|_{\mathbb{H}^n} = \left(dx^1\right)^2 + ... + (dx^n)^2 - \left(dx^{n+1}\right)^2 \\ &= \sum_{i=1}^n \left(f^i \cosh r\, dr + \sinh r\, df^i\right)^2 - \sinh^2 r\, dr^2 \\ &= \sum_{i=1}^n \left[(f^i)^2 \cosh^2 r\, dr^2 + \sinh r \cosh r\, dr\, f^i df^i + f^i df^i \sinh r \cosh r\, dr + \sinh^2 r (df^i)^2\right] \\ &\quad - \sinh^2 r\, dr^2 \\ &= \left(\cosh^2 r - \sinh^2 r\right) dr^2 + \sinh^2 r \sum_{i=1}^n \left(df^i\right)^2 \\ &= dr^2 + \sinh^2 r\, \gamma_{ij}\, d\theta^i d\theta^j, \end{aligned} \tag{1.95}$$

where $\gamma_{ij} d\theta^i d\theta^j$ is the canonical metric on $\mathbb{S}^{n-1}$. We have used here that

$$\sum_{i=1}^n \left(f^i\right)^2 = 1,\ \sum_{i=1}^n f^i df^i = 0 \text{ and } \sum_{i=1}^n \left(df^i\right)^2 = \gamma_{ij} d\theta^i d\theta^j,$$

see the proof of Proposition 1.20. Of course, (1.95) implies (1.94).

Let us verify that $\mathbf{g}_{\mathbb{H}^n}$ is a Riemannian metric. We see from (1.94) 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 $\mathbb{H}^n$ is a graph of a function $x^{n+1} = \sqrt{1 + |x'|^2}$ in $\mathbb{R}^n$. 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. ■

1.14 Model manifolds

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

1. M itself is a chart 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 some 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
$$\mathbf{g} = dr^2 + \psi^2(r)\,\mathbf{g}_{\mathbb{S}^{n-1}}, \tag{1.96}$$
where $\psi(r)$ is a smooth positive function on $(0, r_0)$.

The number r_0 is called the *radius* of the model M, and the function $\psi(r)$ is called the *profile* of M.

In particular, the model manifold M is homeomorphic to B_{r_0}. To simplify the terminology and notation, we usually identify M with the ball B_{r_0}. The origin o of $\mathbb{R}^n$ is called the *pole* of M. The Euclidean polar coordinates (r, θ) are clearly defined in $M \setminus \{o\}$. If $\theta^1, ..., \theta^{n-1}$ are the local coordinates on $\mathbb{S}^{n-1}$ then $r, \theta^1, ..., \theta^{n-1}$ are the local coordinates on $M \setminus \{o\}$. Since

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

we obtain by (1.96), that the metric $\mathbf{g}$ is given in these coordinates by

$$\mathbf{g} = dr^2 + \psi^2(r)\,\gamma_{ij} d\theta^i d\theta^j. \tag{1.97}$$

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 (1.96) extends smoothly to o.

For example, the results of Section 1.13.1 imply the following:

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

Measure and the Laplace-Beltrami operator on models.

Lemma 1.23 *On a model manifold $(M,\mathbf{g})$ with metric (1.96), the Riemannian measure ν is given in the polar coordinates (r,θ) in $B_{r_0}\setminus\{o\}$ by*

$$d\nu=\psi\left(r\right)^{n-1}dr\,d\sigma, \tag{1.98}$$

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

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

$$\Delta_{\mathbf{g}}f=\frac{\partial^2 f}{\partial r^2}+(n-1)\,\frac{\psi'(r)}{\psi(r)}\frac{\partial f}{\partial r}+\frac{1}{\psi^2\left(r\right)}\Delta_{\mathbb{S}^{n-1}}f\,, \tag{1.99}$$

where $\Delta_{\mathbb{S}^{n-1}}$ act on the variable $\theta\in\mathbb{S}^{n-1}$ in the function $f\left(r,\theta\right)$.

Remark. The formula (1.98) can be used to integrate functions over M using the polar coordinates. Indeed, if f is any non-negative measurable function on M then by (1.98)

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

Proof. Let Ω be a chart on $\mathbb{S}^{n-1}$ with coordinates $\theta^1,...,\theta^{n-1}$. Then

$$U=\left\{x\in M:r\in\left(0,r_0\right),\ \theta\in\Omega\right\}$$

is a chart on M with coordinates $r,\theta^1,...,\theta^{n-1}$. Let $g=(g_{ij})_{i,j=0}^{n-1}$ be the matrix of the tensor $\mathbf{g}$ in the chart U, where the index $i=0$ corresponds to the coordinate r and the index $i>0$ corresponds to θ^i. It follows from (1.97) that

$$g=\begin{pmatrix}1&0\ \cdots\ 0\\ 0&\\ \vdots&\boxed{\psi^2\left(r\right)\gamma_{ij}}\\ 0&\end{pmatrix}, \tag{1.100}$$

where $\gamma_{ij}(\theta)$ are the components of $\mathbf{g}_{\mathbb{S}^{n-1}}$ as in (1.81). In particular, we have

$$\det g=\psi^{2(n-1)}\left(r\right)\det\gamma, \tag{1.101}$$

where $\gamma=\left(\gamma_{ij}\right)$. By (1.44), the Riemannian measure σ on $\mathbb{S}^{n-1}$ is given in the chart Ω by

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

Similarly, the Riemannian measure ν on M is given in the chart U by

$$d\nu=\sqrt{\det g}\,dr\,d\theta^1...\,d\theta^{n-1}.$$

Using (1.101) we obtain that

$$\begin{aligned} d\nu &= \psi^{n-1}(r)\sqrt{\det\gamma}\,dr\,d\theta^1...\,d\theta^{n-1} \\ &= \psi^{n-1}(r)dr\,d\sigma, \end{aligned} \tag{1.102}$$

which proves (1.98). In fact, the identity (1.102) can be regarded as a detailed version of (1.98).

26-Nov-25 Lecture 11

It follows from (1.100) that

$$\left(g^{ij}\right) = g^{-1} = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & \boxed{\psi^{-2}(r)\,\gamma^{ij}} \\ 0 & \end{pmatrix}, \tag{1.103}$$

where $(\gamma^{ij}) = \left(\gamma_{ij}\right)^{-1}$. By (1.65), the Laplace-Beltrami operator $\Delta_{\mathbf{g}}$ has the following form in the local coordinates $\theta^0 = r, \theta^1, ..., \theta^{n-1}$:

$$\Delta_{\mathbf{g}} f = \frac{1}{\sqrt{\det g}} \sum_{i,j=0}^{n-1} \frac{\partial}{\partial\theta^i}\left(\sqrt{\det g}\, g^{ij}\frac{\partial f}{\partial\theta^j}\right). \tag{1.104}$$

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

$$\Delta_{\mathbf{g}} f = \frac{1}{\sqrt{\det g}}\frac{\partial}{\partial r}\left(\sqrt{\det g}\,\frac{\partial f}{\partial r}\right) + \sum_{i,j=1}^{n-1} \frac{1}{\sqrt{\det g}}\frac{\partial}{\partial\theta^i}\left(\sqrt{\det g}\, g^{ij}\frac{\partial f}{\partial\theta^j}\right). \tag{1.105}$$

Applying (1.103) and (1.101) and noticing that ψ depends only on r and γ_{ij} depend only on $\theta^1, ..., \theta^{n-1}$, we obtain

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

and

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

Substituting into (1.105), we obtain (1.99). ■

Example. The Riemannian measure in $\mathbb{R}^n$ coincides with the Lebesgue measure that will be denoted as above by λ_n. The induced Riemannian measure on $\mathbb{S}^n$ will be denoted as above by σ_n. The measure σ_n on $\mathbb{S}^n$ is frequently referred to as a (spherical) *area*.

Since in $\mathbb{R}^n$ we have $\psi(r) = r$, it follows from (1.98) that

$$d\lambda_n = r^{n-1} dr\, d\sigma_{n-1}. \tag{1.106}$$

In $\mathbb{S}^n$, we have $\psi(r) = \sin r$, and it follows from (1.98) that

$$d\sigma_n = \sin^{n-1} r\, dr\, d\sigma_{n-1}. \tag{1.107}$$

These two formulas we have seen already in Section 1.13.2.

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

$$d\nu_{\mathbb{H}^n} = \sinh^{n-1} r\ dr\, d\sigma_{n-1}.$$

Let us use (1.107) to compute inductively $\sigma_n(\mathbb{S}^n)$. Using integration in polar coordinates and Fubini's theorem, we obtain

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

It is customary to use the following notation:

$$\omega_n := \sigma_{n-1}\left(\mathbb{S}^{n-1}\right), \tag{1.108}$$

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

$$\omega_{n+1} = \omega_n \int_0^\pi \sin^{n-1} r\, dr. \tag{1.109}$$

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

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

Hence, using (1.109), we obtain

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

$$\omega_4 = 4\pi \int_0^\pi \sin^2 r\, dr = 2\pi^2,$$

$$\omega_5 = 2\pi^2 \int_0^\pi \sin^3 r\, dr = \frac{8}{3}\pi^2,$$

etc.

Example. Now let us give examples of computation of $\Delta_{\mathbf{g}}$ in polar coordinates. In $\mathbb{R}^n$ we have $\psi(r) = r$, and it follows from (1.99) that

$$\Delta_{\mathbb{R}^n} f = \frac{\partial^2 f}{\partial r^2} + \frac{n-1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\Delta_{\mathbb{S}^{n-1}} f\,. \tag{1.110}$$

In $\mathbb{S}^n$ we have $\psi(r) = \sin r$ and $\frac{\psi'}{\psi}(r) = \frac{\cos r}{\sin r} = \cot r$ so that

$$\Delta_{\mathbb{S}^n} f = \frac{\partial^2 f}{\partial r^2} + (n-1)\cot r\frac{\partial f}{\partial r} + \frac{1}{\sin^2 r}\Delta_{\mathbb{S}^{n-1}} f\,. \tag{1.111}$$

In $\mathbb{H}^n$ we have $\psi(r) = \sinh r$ and $\frac{\psi'}{\psi}(r) = \frac{\cosh r}{\sinh r} = \coth r$ so that

$$\Delta_{\mathbb{H}^n} f = \frac{\partial^2 f}{\partial r^2} + (n-1)\coth r\frac{\partial f}{\partial r} + \frac{1}{\sinh^2 r}\Delta_{\mathbb{S}^{n-1}} f\,. \tag{1.112}$$

Recall that

$$\Delta_{\mathbb{S}^1} f = \frac{\partial^2 f}{\partial \theta^2}$$

where θ is the angle variable on $\mathbb{S}^1$. Using (1.110) with $n = 2$, we obtain that, in the polar coordinates (r, θ) in $\mathbb{R}^2$,

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

and in $\mathbb{H}^2$

$$\Delta_{\mathbb{H}^2} f = \frac{\partial^2 f}{\partial r^2} + \coth r\frac{\partial f}{\partial r} + \frac{1}{\sinh^2 r}\frac{\partial^2 f}{\partial \theta^2}.$$

To state a similar formula in $\mathbb{S}^2$, let us denote the polar angle in $\mathbb{S}^2$ by φ instead of r having in mind that φ is the latitude measured from the pole. Then using the spherical coordinates (φ, θ), we obtain from (1.111) in the case $n = 2$ that

$$\Delta_{\mathbb{S}^2} f = \frac{\partial^2 f}{\partial \varphi^2} + \cot\varphi\frac{\partial f}{\partial \varphi} + \frac{1}{\sin^2\varphi}\frac{\partial^2 f}{\partial \theta^2}.$$

This formula allows us to reiterate (1.110), (1.111), (1.112) with $n = 3$ and obtain the following:

$$\begin{aligned}
\Delta_{\mathbb{R}^3} f &= \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\Delta_{\mathbb{S}^2} f \\
&= \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2 f}{\partial \varphi^2} + \cot\varphi\frac{\partial f}{\partial \varphi} + \frac{1}{\sin^2\varphi}\frac{\partial^2 f}{\partial \theta^2}\right) \\
&= \frac{\partial^2 f}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \varphi^2} + \frac{1}{r^2\sin^2\varphi}\frac{\partial^2 f}{\partial \theta^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{\cot\varphi}{r^2}\frac{\partial f}{\partial \varphi},
\end{aligned}$$

$$\begin{aligned}
\Delta_{\mathbb{S}^3} f &= \frac{\partial^2 f}{\partial r^2} + 2\cot r\frac{\partial f}{\partial r} + \frac{1}{\sin^2 r}\Delta_{\mathbb{S}^2} f \\
&= \frac{\partial^2 f}{\partial r^2} + 2\cot r\frac{\partial f}{\partial r} + \frac{1}{\sin^2 r}\left(\frac{\partial^2 f}{\partial \varphi^2} + \cot\varphi\frac{\partial f}{\partial \varphi} + \frac{1}{\sin^2\varphi}\frac{\partial^2 f}{\partial \theta^2}\right) \\
&= \frac{\partial^2 f}{\partial r^2} + \frac{1}{\sin^2 r}\frac{\partial^2 f}{\partial \varphi^2} + \frac{1}{\sin^2 r\sin^2\varphi}\frac{\partial^2 f}{\partial \theta^2} + 2\cot r\frac{\partial f}{\partial r} + \frac{\cot\varphi}{\sin^2 r}\frac{\partial f}{\partial \varphi}
\end{aligned}$$

and

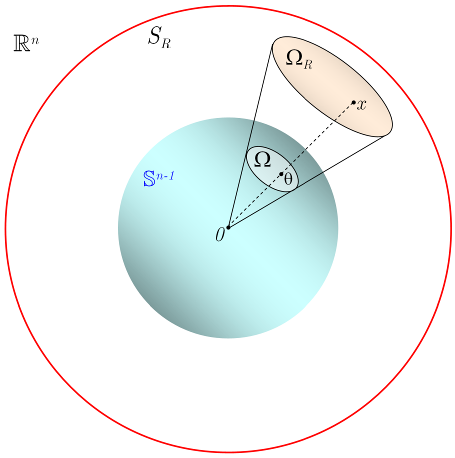
$$\begin{aligned}\Delta_{\mathbb{H}^3} f &= \frac{\partial^2 f}{\partial r^2} + 2\coth r \frac{\partial f}{\partial r} + \frac{1}{\sinh^2 r}\Delta_{\mathbb{S}^2} f \\ &= \frac{\partial^2 f}{\partial r^2} + 2\coth r \frac{\partial f}{\partial r} + \frac{1}{\sinh^2 r}\left(\frac{\partial^2 f}{\partial \varphi^2} + \cot\varphi \frac{\partial f}{\partial \varphi} + \frac{1}{\sin^2\varphi}\frac{\partial^2 f}{\partial \theta^2}\right) \\ &= \frac{\partial^2 f}{\partial r^2} + \frac{1}{\sinh^2 r}\frac{\partial^2 f}{\partial \varphi^2} + \frac{1}{\sinh^2 r \sin^2\varphi}\frac{\partial^2 f}{\partial \theta^2} + 2\coth r\frac{\partial f}{\partial r} + \frac{\cot\varphi}{\sinh^2 r}\frac{\partial f}{\partial \varphi}.\end{aligned}$$

The area function. Let us explain the meaning of the term $\psi^{n-1}(r)$ in (1.98). 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 $n-1$. Any chart Ω in $\mathbb{S}^{n-1}$ with coordinates $\theta^1, .., \theta^{n-1}$ gives rise to a chart Ω_R in S_R, also with coordinates $\theta^1, ..., \theta^{n-1}$:

$$\Omega_R = \{x = (r, \theta) \in \mathbb{R}^n : r = R,\ \theta \in \Omega\}.$$



Setting in (1.96) $r \equiv R$, we obtain 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 as before by σ the Riemannian measure on the unit sphere $\mathbb{S}^{n-1}$ and by σ_R the induced Riemannian measure on S_R (that is also called *area*), we obtain

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

It follows that

$$\sigma_R(S_R) = \psi(R)^{n-1}\sigma(\mathbb{S}^{n-1}) = \omega_n \psi(R)^{n-1}. \tag{1.113}$$

Definition. The function

$$\boxed{S(R) := \sigma_R(S_R) = \omega_n \psi(R)^{n-1}}$$

is called the *area function* of the model.

The are function determines the profile $\psi(R)$ and, hence, the model metric $\mathbf{g}$.

For example, in $\mathbb{R}^n$ we have

$$S(R) = \omega_n R^{n-1}, i$$

n $\mathbb{S}^n$:

$$S(R) = \omega_n (\sin R)^{n-1},$$

in $\mathbb{H}^n$:

$$S(R) = \omega_n (\sinh R)^{n-1}.$$

28-Nov-25 Lecture 12

Let restate the results of Lemma 1.23 in terms of the area function.

Corollary 1.24 *On a model manifold* $(M, \mathbf{g})$ *with metric* (1.96)*, we have*

$$d\nu = \frac{1}{\omega_n} S(r) dr\, d\sigma, \tag{1.114}$$

and

$$\Delta_{\mathbf{g}} = \frac{\partial^2}{\partial r^2} + \frac{S'(r)}{S(r)} \frac{\partial}{\partial r} + \frac{1}{\psi^2(r)} \Delta_{\mathbb{S}^{n-1}}. \tag{1.115}$$

The volume 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. It follows from (1.114) that

$$\nu(B_R) = \frac{1}{\omega_n} \int_0^R \int_{\mathbb{S}^{n-1}} S(r) dr d\sigma = \int_0^R S(r)\, dr.$$

Definition. The function

$$V(R) := \nu(B_R) = \int_0^R S(r)\, dr. \tag{1.116}$$

is called the *volume function* of the model manifold.

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

$$V(R) = \frac{\omega_n}{n} R^n. \tag{1.117}$$

In $\mathbb{S}^n$ we have $\psi(r) = \sin^{n-1} r$ whence

$$V(R) = \omega_n \int_0^R \sin^{n-1} r\, dr$$

and in $\mathbb{H}^n$

$$V(R) = \omega_n \int_0^R \sinh^{n-1} r\, dr.$$

For example, for $n = 2$, the volume function in $\mathbb{S}^2$ is

$$V(R) = 2\pi \int_0^R \sin r\, dr = 2\pi (1 - \cos R)$$

and in $\mathbb{H}^2$:

$$V(R) = 2\pi \int_0^R \sinh r\, dr = 2\pi (\cosh R - 1).$$

Weighted models. Finally, we discussed models in the class of weighted manifolds.

Definition. A weighted manifold $(M, \mathbf{g}, \mu)$ 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 1.25 *On a weighted model manifold $(M, \mathbf{g}, \mu)$ with metric (1.96) and the density function $D(r)$, the measure μ is given in the polar coordinates by*

$$d\mu = D(r)\, \psi^{n-1}(r)\, dr d\sigma \tag{1.118}$$

where σ is the Riemannian measure on $\mathbb{S}^{n-1}$. The weighted Laplace operator $\Delta_{\mathbf{g},\mu}$ has in the polar coordinates the form

$$\Delta_{\mathbf{g},\mu} = \frac{\partial^2}{\partial r^2} + \frac{d}{dr} \ln \left(D\psi^{n-1} \right) \frac{\partial}{\partial r} + \frac{1}{\psi^2(r)} \Delta_{\mathbb{S}^{n-1}}. \tag{1.119}$$

Proof. The identity (1.118) follows immediately from $d\mu = Dd\nu$ and $d\nu = \psi^{n-1}(r)\, dr d\sigma$ of Lemma 1.23 (cf. (1.98)).

By definition of the weighted Laplacian, we have

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

Using the notation $\theta^0 = r$ and the matrix (g^{ij}) given by (1.103), we obtain

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

because $\frac{\partial \ln D}{\partial \theta^i} = 0$ for all $i \geq 1$. Using the representation of $\Delta_{\mathbf{g}}$ from Lemma 1.23, we obtain

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

Finally, observing that

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

we obtain (1.119). ■

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, that is,

$$d\mu_S = D|_S \, d\nu_S.$$

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 (1.118) and (1.119).

The function

$$\boxed{S(r) := \mu_R(S_R) = \omega_n D(r) \, \psi^{n-1}(r)} \tag{1.120}$$

is called the *area function* of the weighted model $(M, \mathbf{g}, \mu)$. Using the area function, we can rewrite (1.118) and (1.119) as follows:

$$d\mu = \frac{1}{\omega_n} S(r) \, dr \, d\sigma$$

and

$$\boxed{\Delta_{\mathbf{g},\mu} = \frac{\partial^2}{\partial r^2} + \frac{S'(r)}{S(r)} \frac{\partial}{\partial r} + \frac{1}{\psi^2(r)} \Delta_{\mathbb{S}^{n-1}}} \tag{1.121}$$

as in Corollary 1.24.

1.15 Length of paths and the geodesic distance

A path and its velocity. 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 $x = \gamma(t)$:

$$\dot{\gamma}(t) : C^\infty(M) \to \mathbb{R}$$
$$\dot{\gamma}(t)(f) = \frac{d}{dt} f(\gamma(t)) \quad \text{for any } f \in C^\infty(M). \tag{1.122}$$

Indeed, it is easy to see that the mapping $\dot{\gamma}(t)$ defined by (1.122) 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 $\frac{d}{dt}$ (see Exercise 8). 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

$$\dot{\gamma}(t)(f) = \frac{d}{dt} f\left(\gamma^1(t), ..., \gamma^n(t)\right) = \frac{\partial f}{\partial x^i} \frac{d\gamma^i}{dt} = \dot{\gamma}^i \frac{\partial f}{\partial x^i}, \tag{1.123}$$

where using the notation

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

Rewriting (1.123) in the operator form as follows

$$\dot{\gamma} = \dot{\gamma}^i \frac{\partial}{\partial x^i}, \tag{1.124}$$

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 (1.124), 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 $\gamma^i(t) = x^i + t\xi^i$.

The length of a path. Let now $(M, \mathbf{g})$ be a Riemannian manifold. Recall that length of a tangent vector $\xi \in T_xM$ 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

$$\ell_{\mathbf{g}}(\gamma) = \int_I |\dot{\gamma}(t)|_{\mathbf{g}}\, dt. \tag{1.125}$$

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

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

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.

Definition. 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 $[t_k, t_{k+1}]$.

For a piecewise C^1 path γ, the velocity $\dot{\gamma}(t)$ is defined for all $t \neq t_k$ and the integral (1.125) still makes sense. Hence, the length $\ell_{\mathbf{g}}(\gamma)$ is well defined for piecewise C^1 paths and, moreover, is finite.

Geodesic distance. 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. Define the *geodesic distance* $d(x, y)$ between any two points $x, y \in M$ by

$$d(x, y) = \inf \left\{ \ell_{\mathbf{g}}(\gamma) : \gamma \text{ is a piecewise } C^1\text{-path connecting } x \text{ and } y \right\}. \tag{1.126}$$

If the infimum in (1.126) 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 between $x, y \in \mathbb{R}^n$ is the shortest geodesic (see Exercises).

Our purpose is to show that geodesic distance d is a metric on M, and the topology of the metric space (M, d) coincides with the original topology of the smooth manifold M (see Theorem 1.29 below).

03-Dec-25 Lecture 13

Before we proceed with the properties of geodesic distance, let us make some comments about the notions of curve and length. We have defined above the notion of a parametric curve as a C^1 mapping $\gamma : I \to M$ where I is an interval, and its length $\ell_{\mathbf{g}}(\gamma)$ by (1.125). One can define the notion of a curve differently: let us say that a (non-parametric) curve is an 1-dimensional submanifold of M. If S is such a curve then there is an induce Riemannian metric $\mathbf{g}_S$ and induced Riemannian measure ν_S. It is natural to define the length of such a curve S as $\nu_S(S)$.

A question arises what is the relation between the notions of parametric and non-parametric curve and the relation between the two notions of length.

The answer is given in Exercise 42. Suppose that a parametric curve $\gamma : (a, b) \to M$ is C^∞ smooth, injective, $\dot{\gamma}(t) \neq 0$ for all $t \in (a, b)$ and that γ is a homeomorphism of (a,b) onto the image $S = \gamma(a.b)$. Then (by Exercise 17) S is a submanifold of dimension 1 (that is, a non-parametric curve), and $\ell_{\mathbf{g}}(\gamma) = \nu_S(S)$. Hence, the two notions of length coincide when they both make sense.

Now let us return to the notion of geodesic distance and start with the following observations. Recall that

$$d(x,y) = \inf\left\{\ell_{\mathbf{g}}(\gamma) : \gamma \text{ is a piecewise } C^1\text{-path connecting } x \text{ and } y\right\}. \tag{1.127}$$

Lemma 1.26 *Geodesic distance satisfies the following properties.*

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

(*b*) *Symmetry:* $d(x,y) = d(y,x)$.

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

Remark. The remaining two issues that $d(x,y) < \infty$ and $d(x,y) > 0$ for $x \neq y$ will be addressed below.

Proof. (*a*) That $d(x,y) \in [0,\infty]$ is obvious from definition (1.127). 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.

(*b*) 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,

$$\widetilde{\gamma}(a) = \gamma(b) = y \quad \text{and} \quad \widetilde{\gamma}(b) = \gamma(a) = x$$

so that $\widetilde{\gamma}$ connects y and x. It is obvious from the definition that

$$\frac{d}{dt}\widetilde{\gamma}^i(t) = \frac{d}{dt}\left(\gamma^i(a+b-t)\right) = -\dot{\gamma}^i(a+b-t)$$

whence

$$\left|\frac{d}{dt}\widetilde{\gamma}(t)\right|_{\mathbf{g}} = |\dot{\gamma}(a+b-t)|_{\mathbf{g}}$$

and

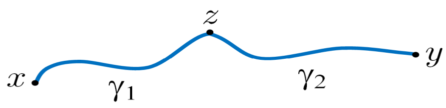
$$\begin{aligned}\ell_{\mathbf{g}}(\widetilde{\gamma}) &= \int_a^b \left|\frac{d}{dt}\widetilde{\gamma}(t)\right|_{\mathbf{g}} dt = \int_a^b |\dot{\gamma}(a+b-t)|_{\mathbf{g}}\, dt \quad (\text{change } s = a+b-t)\\ &= -\int_b^a |\dot{\gamma}(s)|_{\mathbf{g}}\, ds = \ell_{\mathbf{g}}(\gamma).\end{aligned}$$

Applying the definition (1.127), we obtain

$$\begin{aligned}d(x,y) &= \inf\{\ell_{\mathbf{g}}(\gamma) : \gamma \text{ connects } x \text{ and } y\}\\ &= \inf\{\ell_{\mathbf{g}}(\widetilde{\gamma}) : \widetilde{\gamma} \text{ connects } y \text{ and } x\}\\ &= d(y,x).\end{aligned}$$

(c) 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 parameter t, we can always assume that $a_2 = b_1$. Define a path $\gamma : [a_1, b_2]$ connecting x and y, as follows:

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



The path γ 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 (1.127) that

$$d(x, y) \le \ell_{\mathbf{g}}(\gamma) = \int_{a_1}^{b_2} |\dot\gamma(t)|_{\mathbf{g}}\, dt = \int_{a_1}^{b_1} |\dot\gamma_1(t)|_{\mathbf{g}}\, dt + \int_{a_2}^{b_2} |\dot\gamma_2(t)|_{\mathbf{g}}\, dt = \ell_{\mathbf{g}}(\gamma_1) + \ell_{\mathbf{g}}(\gamma_2).$$

Taking infimum in the right hand side 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 1.27 *For any point $p \in M$, there is a chart $U \ni p$ and a constant $C \ge 1$ such that, for all $x, y \in \overline{U}$,*

$$C^{-1}|x - y| \le d(x, y) \le C|x - y|, \tag{1.128}$$

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 $x^1, ..., x^n$. Without loss of generality, we assume that p is the origin of this coordinates system. Let V be the Euclidean ball $B_r(p)$ in the chart W 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 $\xi \in T_x M$, its length $|\xi|_{\mathbf{g}}$ in the metric $\mathbf{g}$ is given by

$$|\xi|_{\mathbf{g}}^2 = \sum_{i,j=1}^n g_{ij}(x)\, \xi^i \xi^j.$$

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

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

We claim that there is a constant $C \ge 1$ such that

$$C^{-2} \sum_{i=1}^n \left(\xi^i\right)^2 \le \sum_{i,j=1}^n g_{ij}(x)\, \xi^i \xi^j \le C^2 \sum_{i,j=1}^n \left(\xi^i\right)^2, \tag{1.129}$$

for all $x \in \overline{V}$ and $\xi \in \mathbb{R}^n$. Since the inequality (1.129) does not change if we multiply ξ by a positive constant, it suffices to prove (1.129) when $|\xi|_{\mathbf{e}} = 1$, that is, when $\xi \in \mathbb{S}^{n-1}$. Indeed, the function

$$\sum_{i,j=1}^n g_{ij}(x)\, \xi^i \xi^j, \quad x \in \overline{V}, \quad \xi \in \mathbb{S}^{n-1}$$

is continuous and positive, which implies that, in the compact domain $\overline{V} \times \mathbb{S}^{n-1}$, its maximum is finite and minimum is positive, whence (1.129) follows.

It follows from (1.129) that, for all $x \in \overline{V}$ and $\xi \in T_x M$,

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

Consequently, for any piecewise C^1 path γ in $\overline{V}$, we have

$$C^{-1} \ell_{\mathbf{e}} (\gamma) \le \ell_{\mathbf{g}} (\gamma) \le C \ell_{\mathbf{e}} (\gamma) . \tag{1.130}$$

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

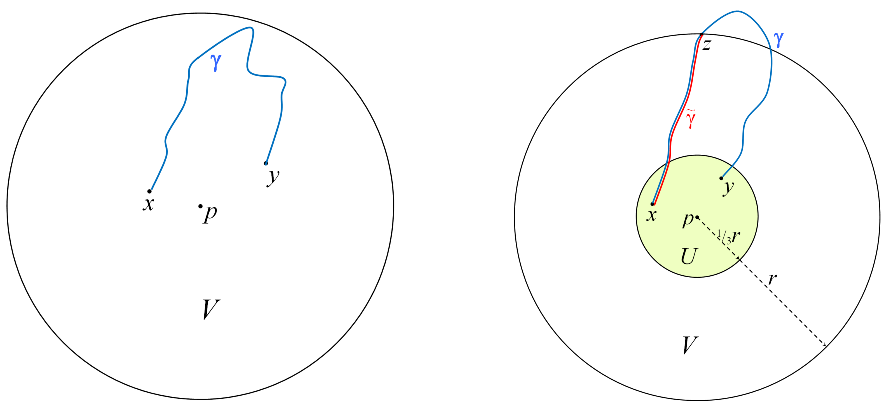
$$d (x, y) \le \ell_{\mathbf{g}} (\gamma) \le C \ell_{\mathbf{e}} (\gamma) = C |x - y| ,$$

which proves the upper bound in (1.128) (see Exercise 41 for the proof of the identity $\ell_{\mathbf{e}} (\gamma) = |x - y|$).

In order to prove the lower bound in (1.128), we have to reduce V further as follows. Define now the set U by $U = B_{\frac{1}{3} r} (p)$ and prove that, for any piecewise C^1 path γ on M connecting points $x, y \in \overline{U}$,

$$\ell_{\mathbf{g}} (\gamma) \ge C^{-1} |x - y| , \tag{1.131}$$

which will imply $d (x, y) \ge C^{-1} |x - y|$. For that, we consider two cases, when γ stays in V and when γ has some points outside V.



Case 1: γ stays in V Case 2: γ intersects ∂V at a point z.

If γ stays in V then (1.131) follows from (1.130) as follows:

$$\ell_{\mathbf{g}} (\gamma) \ge C^{-1} \ell_e (x, y) \ge C^{-1} |x - y| .$$

Consider now the second case when γ does not stay in V. Then γ intersects the boundary ∂V at some point z. Indeed, the image of γ is a connected set, and if it does not intersect ∂V then it is covered by two disjoint open sets V and $\overline{V}^c$, which is not possible as γ has non-empty intersection with each of them.

Denoting by $\widetilde{\gamma}$ the part of γ that connects in $\overline{V}$ the point x to the first point z on ∂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 (1.131) for the path $\widetilde{\gamma}$ and the inequalities

$$|x-z| \geq \frac{2}{3}r \quad \text{and} \quad |x-y| \leq |x|+|y| \leq \frac{2}{3}r,$$

and which finishes the proof. ■

Proposition 1.28 *We have $d(x,y) > 0$ for all distinct points $x, y \in M$. Hence, the geodesic distance $d(x,y)$ satisfies the axioms of a metric and (M,d) is a metric space.*

Remark. Here we allow a metric $d(x,y)$ to take value $+\infty$. It can always be replaced by a *finite* metric

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

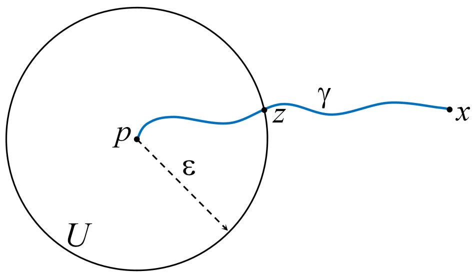
which determines the same topology as $d(x,y)$. However, if the manifold M is connected then the metric d is finite itself: $d(x,y) < \infty$ for all $x, y \in M$ (see Exercise 40).

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 1.27. We can always assume that U is a Euclidean ball $B_\varepsilon(p)$ of some radius $\varepsilon > 0$. If $x \in U$ then by (1.128)

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

Assume that $x \notin U$. Then any path γ connecting p and x must intersect the boundary ∂U, say at a point z. Denoting by $\widetilde{\gamma}$ the part of γ between p and z, we obtain by means of (1.128) that

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



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

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

The topology of the geodesic distance. As any metric, the geodesic distance induces a topology on M.

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, $B(x,r)$ is a metric ball 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 1.29 *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 τ_M the original topology of M and by τ_d – the topology of the metric space (M,d) (recall that a topology is a collection of open sets). We need to prove that $\tau_M = \tau_d$.

Fix a point $p \in M$ and denote by $B(p,r)$ the geodesic (metric) ball of radius $r > 0$. If U is a chart U containing p then denote by $B_r(p)$ the Euclidean ball in the chart U containing p, with small enough $r > 0$.

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

$$B_{C^{-1}r}(p) \subset B(p,r) \subset B_{Cr}(p). \tag{1.132}$$

Indeed, if $\Omega \in \tau_d$ (that is , Ω is open in the metric topology of d) then, for any $p \in \Omega$, there exists $r > 0$ such that $B(p,r) \subset \Omega$. By reducing r, we can assume that (1.132) is satisfied in the chart U as above. Hence, $B_{C^{-1}r}(p) \subset \Omega$ and, hence, Ω is open in the Euclidean topology of U, which implies that $\Omega \in \tau_M$. In the same way one proves that if $\Omega \in \tau_M$ then $\Omega \in \tau_d$.

Let us now prove the existence of a chart U as claimed above. Fix a point $p \in M$ and let U be a chart constructed in Lemma 1.27, where (1.128) holds, that is,

$$C^{-1}|x-y| \leq d(x,y) \leq C|x-y| \quad \text{for all } x,y \in \overline{U}. \tag{1.133}$$

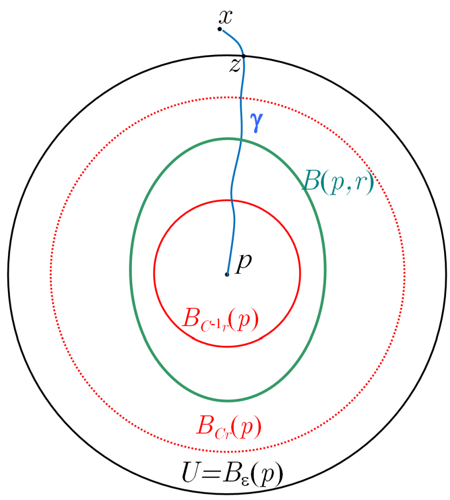
As above, we can assume that U coincides with the Euclidean ball $B_\varepsilon(p)$ of some radius $\varepsilon > 0$. Let prove the inclusions (1.132) for this chart U.

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

$$d(x,p) \leq C|x-p| < CC^{-1}r = r,$$

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

To prove the second inclusion in (1.132), let us first verify that $B(p,r) \subset U$ provided $r < C^{-1}\varepsilon$.



If $x \notin U$ then any path γ connecting x and p intersects ∂U

Indeed, if $x \notin U$ then any path γ connecting p and x contains a point $z \in \partial U$. By (1.133), we obtain

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

whence $d(x,p) \geq r$ and $x \notin B(p,r)$.

Therefore, if $x \in B(p,r)$ then $x \in U$ and, by (1.133),

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

which implies $x \in B_{Cr}(p)$. ■

1.16 Smooth mappings, push-forward and pullback

Smooth mapping and pullback of functions. Let X and Y be two smooth manifolds of dimension n and m, respectively.

Definition. A continuous mapping

$$\Phi : Y \to X$$

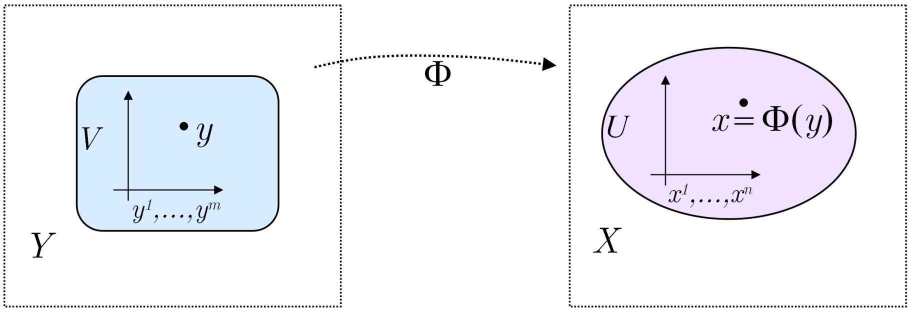
is called smooth if all its components in any charts of X and Y are smooth functions.

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 $V \subset Y$. Assume that $\Phi(V) \subset U$. Then the mapping Φ in V is given by n equations

$$x^i = \Phi^i\left(y^1, ..., y^m\right), \tag{1.134}$$

where all functions Φ^i are smooth. Note that, 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

(1.134) in a neighborhood of any point $y \in Y$. The mapping Φ can be regarded as a *parametric surface* of dimension m, similarly to the notion of a parametric curve.



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 in the case "from X to Y" they are called "*pullback*" operators.

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. The mapping

$$\begin{gathered}\Phi_* : C^\infty(Y) \to C^\infty(X) \\ \Phi_* f = f \circ \Phi\end{gathered}$$

is called a *pullback operator.*

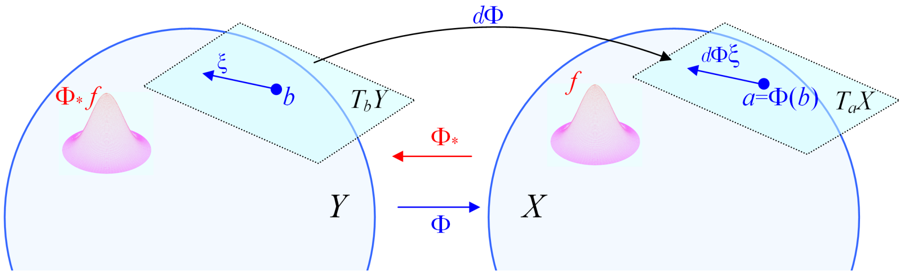
For example, for the coordinate function $f = x^i$ we obtain

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

Tangent map and push-forward of tangent vectors. Now fix a point $b \in Y$ and set $a = \Phi(b) \in X$.

Definition. Define the *push-forward* map $d\Phi : T_b Y \to T_a X$ as follows: for any tangent vector $\xi \in T_b Y$, define $d\Phi\xi$ as an $\mathbb{R}$-differentiation at a by

$$(d\Phi\xi)(f) = \xi(\Phi_* f) \quad \text{for any } f \in C^\infty(X). \tag{1.135}$$



Pullback and push-forward

Clearly, $d\Phi\xi$ is a linear mapping from $C^\infty(X)$ to $\mathbb{R}$. The fact that $d\Phi\xi$ is an $\mathbb{R}$-differentiation at a is stated in the next lemma.

Lemma 1.30 *For any $\xi \in T_bY$, its push-forward $d\Phi\xi$ is an $\mathbb{R}$-differentiation at $a \in X$, that is, $d\Phi\xi \in T_aX$. In the local coordinates $x^1, ..., x^n$ on X and $y^1, ..., y^m$ on Y, we have*

$$\boxed{(d\Phi\xi)^i = \xi^j \frac{\partial \Phi^i}{\partial y^j}(b)}. \tag{1.136}$$

Proof. For any $f \in C^\infty(X)$ and for any tangent vector $\xi = \xi^j \frac{\partial}{\partial y^j} \in T_bY$, we have

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

It follows that

$$d\Phi\xi = \xi^j \frac{\partial \Phi^i}{\partial y^j}(b) \frac{\partial}{\partial x^i},$$

where $\frac{\partial}{\partial x^i}$ is taken at a. It follows that $d\Phi\xi \in T_aX$ and that the components of $d\Phi\xi$ in the basis $\left\{\frac{\partial}{\partial x^i}\right\}$ in T_aX are given by (1.136). ■

Consider the Jacobi matrix

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

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 $(d\Phi\xi)_{\text{col}}$ similarly, (1.136) can be written in terms of matrix multiplication as follows:

$$(d\Phi\xi)_{\text{col}} = J(b)\xi_{\text{col}}. \tag{1.137}$$

Definition. The push-forward map

$$\begin{aligned} d\Phi : T_bY &\to T_aX \\ \xi &\mapsto d\Phi\xi \end{aligned} \tag{1.138}$$

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

Recall that, for a smooth mapping $\Phi : \mathbb{R}^m \to \mathbb{R}^n$, the differential $d\Phi(b)$ at a point $b \in \mathbb{R}^m$ is defined as a linear mapping $\mathbb{R}^m \to \mathbb{R}^n$ such that

$$\Phi(b+\xi) - \Phi(b) = d\Phi(b)\xi + o(|\xi|) \text{ as } \xi \to 0,$$

and it is expressed through the partial derivatives as follows:

$$(d\Phi(b)\xi)^i = \frac{\partial \Phi^i}{\partial y^j}(b)\xi^j,$$

which matches (1.136)

Cotangent map and pullback of tangent covectors. As above, let $\Phi : Y \to X$ be a smooth mapping of smooth manifolds X and Y, b be any point on Y and $a = \Phi(b) \in X$. Then $d\Phi$ is a linear mapping from T_bY to T_aX. As any other linear mapping, it has a dual mapping from T_a^*X to T_b^*Y as is given in the next definition.

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

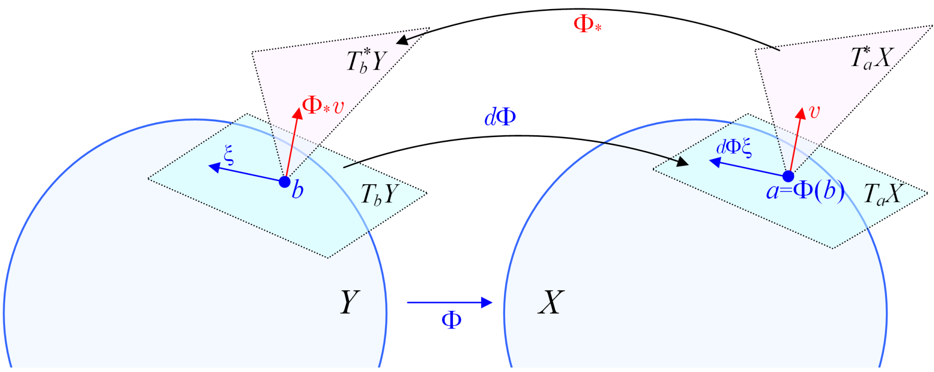
$$\langle \Phi_* v, \xi \rangle = \langle v, d\Phi \xi \rangle \quad \forall \xi \in T_bY. \tag{1.139}$$

The pull-back mapping

$$\begin{aligned} \Phi_* : T_a^*X &\to T_b^*Y \\ v &\mapsto \Phi_* v \end{aligned}$$

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

Remark. A natural notation for the dual map of $d\Phi$ would be $(d\Phi)^*$, which, however, is not commonly used, giving preference to Φ_*. One of the reasons for that is the identity (1.143) below.



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

Let $\Phi : Y \to X$ be smooth mapping of two smooth manifolds X, Y. For any real valued function f on X, we define its pullback $\Phi_* f$ as a function on Y by

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

Clearly, if $f \in C^\infty(X)$ then $\Phi_* f \in C^\infty(Y)$.

Fix a point $b \in Y$ and set $a = \Phi(b) \in X$. For any tangent vector $\xi \in T_b Y$, define its push-forward $d\Phi\xi \in T_a X$ as $\mathbb{R}$-differentiation as follows:

$$d\Phi\xi(f) = \xi(\Phi_* f) \quad \text{for any } f \in C^\infty(X). \tag{1.140}$$

If $x^1, ..., x^n$ are local coordinates in X and $y^1, ..., y^m$ are local coordinates in Y then we have by (1.136)

$$(d\Phi\xi)^i = \xi^j \frac{\partial \Phi^i}{\partial y^j}(b). \tag{1.141}$$

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

$$\langle \Phi_* v, \xi \rangle = \langle v, d\Phi\xi \rangle \quad \forall \xi \in T_b Y. \tag{1.142}$$

Lemma 1.31 *For any $f \in C^\infty(X)$,*

$$\Phi_* df = d(\Phi_* f). \tag{1.143}$$

In the local coordinates $x^1, ..., x^n$ on X and $y^1, ..., y^m$ on Y, we have for any $v \in T_a^* X$,

$$\boxed{(\Phi_* v)_j = v_i \frac{\partial \Phi^i}{\partial y^j}(b)}. \tag{1.144}$$

Proof. The both sides of (1.143) are tangent covectors in Y. For all $b \in Y$ and $\xi \in T_b Y$ we have

$$\langle \Phi_* df, \xi \rangle \overset{(1.142)}{=} \langle df, d\Phi\xi \rangle = (d\Phi\xi) f \overset{(1.140)}{=} \xi(\Phi_* f) = \langle d(\Phi_* f), \xi \rangle,$$

which proves (1.143) (in the second and forth equality we have used the definition of differential of a function).

To prove (1.144), observe that, for any $\xi \in T_b Y$,

$$\langle \Phi_* v, \xi \rangle \overset{(1.142)}{=} \langle v, d\Phi\xi \rangle = v_i (d\Phi\xi)^i \overset{(1.141)}{=} v_i \xi^j \frac{\partial \Phi^i}{\partial y^j}(b).$$

On the other hand, we have

$$\langle \Phi_* v, \xi \rangle = (\Phi_* v)_j \xi^j.$$

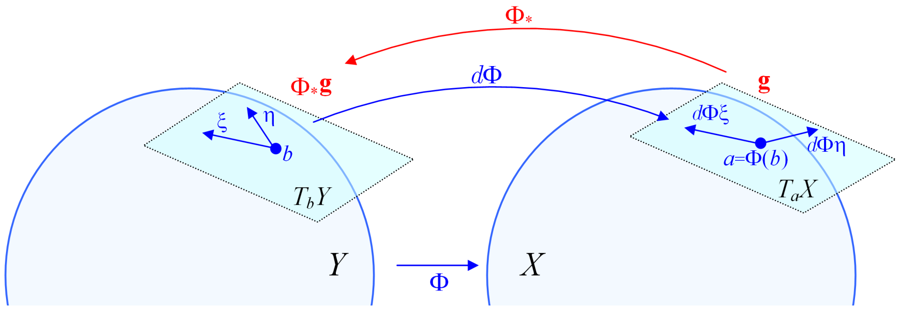
Comparison of these two identities yields (1.144). ■

Using as above the notation $J = \left(\frac{\partial \Phi^i}{\partial y^j} \right)$ for the Jacobi matrix of Φ, denoting by v_{row} the row vector with components $(v_1, ..., v_n)$ and understanding $(\Phi_* v)_{\text{row}}$ similarly, we obtain from (1.144) the matrix identity

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

Pullback of a Riemannian metric. Considering as above a smooth mapping $\Phi : Y \to X$, assume that we are given a bilinear form $\mathbf{g}$ on T_aX where $a = \Phi(b)$ for some $b \in Y$ (for example, $\mathbf{g}$ can be a Riemannian metric). Then define its pullback $\Phi_*\mathbf{g}$ as a bilinear form on T_bY by

$$\Phi_*\mathbf{g}(\xi, \eta) = \mathbf{g}(d\Phi\xi, d\Phi\eta) \quad \text{for all } \xi, \eta \in T_bY. \tag{1.145}$$



If $\mathbf{g}$ is symmetric then also $\Phi_*\mathbf{g}$ is symmetric. If $\mathbf{g}$ is positive definite then $\Phi_*\mathbf{g}$ is non-negative definite as

$$\Phi_*\mathbf{g}(\xi, \xi) = \mathbf{g}(d\Phi\xi, d\Phi\xi) \geq 0.$$

Clearly, $\Phi_*\mathbf{g}$ is positive definite if and only if the tangent map $d\Phi : T_bY \to T_aX$ is injective because in this case

$$\xi \neq 0 \Rightarrow d\Phi\xi \neq 0 \Rightarrow \mathbf{g}(d\Phi\xi, d\Phi\xi) > 0.$$

Proposition 1.32 *Let $m = \dim Y \leq n = \dim X$. Assume that the Jacobi matrix J of $\Phi : Y \to X$ has at all points of Y the maximal rank m. Then, for any Riemannian metric $\mathbf{g}$ on X, its pullback $\Phi_*\mathbf{g}$ is a Riemannian metric on Y.*

Besides, in the local coordinates $x^1, ..., x^n$ on X and $y^1, ..., y^m$ on Y, the matrices g^x of $\mathbf{g}$ and Φ_*g^y of $\Phi_*\mathbf{g}$ satisfy the following identity:

$$\boxed{\Phi_*g^y = J^T g^x J}. \tag{1.146}$$

Note that $g^x = (g_{ij})$ is an $n \times n$ matrix and $(\Phi_*g)^y = ((\Phi_*g)_{kl})$ is an $m \times m$ matrix.

Proof. Let us fix the bases $\left\{\frac{\partial}{\partial x^i}\right\}$ in T_aX and $\left\{\frac{\partial}{\partial y^j}\right\}$ in T_bY, respectively, where $a = \Phi(b)$. In these bases the tangent map $d\Phi : T_bY \to T_aX$ is given by (1.137) by multiplication 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. Since rank $J = m$, the image of $d\Phi$ is an m-dimensional subspace of T_aX, which implies that $d\Phi$ is injective. Consequently, $\Phi_*\mathbf{g}$ is positive definite and, hence, is a Riemannian metric on Y.

Using (1.145) and (1.141) we obtain, for all $\xi, \eta \in T_b Y$,

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

whence

$$\boxed{(\Phi_* g^y)_{kl} = g_{ij}^x \frac{\partial \Phi^i}{\partial y^k}\frac{\partial \Phi^j}{\partial y^l}}, \tag{1.147}$$

which is equivalent to (1.146). ■

Example. Let Y be a submanifold of X and let $\Phi : Y \to X$ be the identical embedding of Y into X. Let $\mathbf{g}$ be a Riemannian metric on X. We claim that

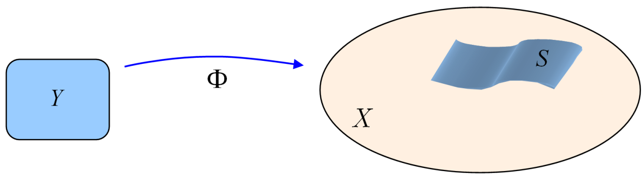
$$\Phi_* \mathbf{g} = \mathbf{g}_Y,$$

that is, the induced metric on Y coincides with the pullback metric. Indeed, in the local coordinates these two Riemannian metrics have the same matrices as by (1.146)

$$\Phi_* g^y = J^T g^x J,$$

and by Lemma 1.11 (see (1.42)) we have $g_Y^y = J^T g^x J$.

Remark. Any smooth mapping $\Phi : Y \to X$ can be regarded as a *parametric surface* in X of dimension m.



If Y is an interval then it becomes a parametric curve that is a parametric surface of dimension 1. By Exercise 17, if Φ satisfies the following three properties

- the mapping $\Phi : Y \to X$ is injective;
- the rank of the Jacobi matrix $J = \left(\frac{\partial \Phi^i}{\partial y^j}\right)$ of Φ is maximal at all points, that is, it is equal to m;
- Φ is a homeomorphism of Y onto its image $S := \Phi(Y) \subset X$;

then the image S is a submanifold of X of dimension m.

Push-forward and pullback of composition. Suppose that we have three manifolds X, Y, Z and two smooth mappings

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

so that their composition

$$\Phi \circ \Psi : Z \to X$$

is well defined and is a smooth mapping.

The pullback operation for functions satisfies the following identity

$$\boxed{(\Phi \circ \Psi)_* = \Psi_* \circ \Phi_*} \tag{1.148}$$

because for any function $f \in C^\infty(X)$

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

Lemma 1.33 (a) *Push-forward operation for tangent vectors on Z satisfies the identity*

$$\boxed{d(\Phi \circ \Psi) = d\Phi \circ d\Psi}.$$

(b) *Pullback operation for tangent covectors on X satisfies the identify*

$$\boxed{(\Phi \circ \Psi)_* = \Psi_* \circ \Phi_*}$$

Observe 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 *reverse* order.

Proof. (a) We need to prove that, for all $c \in Z$ and $\xi \in T_x Z$,

$$d(\Phi \circ \Psi)\xi = d\Phi(d\Psi\xi).$$

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

$$\begin{aligned}(d(\Phi \circ \Psi)\xi)(f) &\overset{(1.135)}{=} \xi((\Phi \circ \Psi)_* f) \overset{(1.148)}{=} \xi(\Psi_*(\Phi_* f)) \\ &\overset{(1.135)}{=} (d\Psi\xi)(\Phi_* f) \overset{(1.135)}{=} d\Phi(d\Psi\xi)(f),\end{aligned}$$

whence the claim follows.

(b) Fix $c \in Z$ and set $a = \Phi \circ \Psi(c) \in X$. Then, for any $v \in T_a^* X$ and $\xi \in T_c Z$

$$\langle (\Phi \circ \Psi)_* v, \xi \rangle = \langle v, d(\Phi \circ \Psi)\xi \rangle = \langle v, d\Phi(d\Psi\xi) \rangle = \langle \Phi_* v, d\Psi\xi \rangle = \langle \Psi_*(\Phi_* v), \xi \rangle,$$

whence the claim follows. ∎

12-Dec-25 Lecture 16

Diffeomorphism and isometry. Let now Y and X have the same dimension n.

Definition. 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

$$d\Phi : T_b Y \to T_a X \quad \text{and} \quad d\left(\Phi^{-1}\right) : T_a X \to T_b Y$$

(where $a = \Phi(b)$) are also mutually inverse because

$$d\Phi \circ d\Phi^{-1} = d(\Phi \circ \Phi^{-1}) = \mathrm{id},$$

which implies that the tangent map $d\Phi$ 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 exists 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 $\cong$ (\iso in Latex). It is easy to see that the relation $\cong$ between Riemannian manifolds is reflexive , symmetric and transitive so that $\cong$ is an equivalence relation between Riemannian manifolds. If two manifolds are isometric then they have exactly the same properties as Riemannian manifolds and normally can be identified with each other and regarded as the same manifold.

Example. Let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ be a diffeomorphism such that its Jacobi matrix J of Φ is at any point orthogonal, that is, $J^T J = \mathrm{id}$. Then Φ is an isometry of $(\mathbb{R}^n, \mathbf{g})$ where $\mathbf{g}$ is the canonical Euclidean metric, because in the Cartesian coordinates $x^1, ...x^n$ in the target $g^x = \mathrm{id}$, whence in the Cartesian coordinates $y^1, ..., y^n$ in the source

$$(\Phi_* g)^y = J^T g^x J = \mathrm{id} = g^y.$$

For example, let A be an orthogonal matrix (with constant coefficients) and $B \in \mathbb{R}^n$. Then $\Phi(y) = Ay + B$ satisfies the above conditions and, hence, is an isometry of $(\mathbb{R}^n, \mathbf{g})$.

Example. Let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ by given by $\Phi(y) = Ay$ where A is an orthogonal matrix. Then $|y| = 1$ implies $|\Phi(y)| = 1$ so that $\Psi = \Phi|_{\mathbb{S}^n}$ is a diffeomorphism of $\mathbb{S}^n$ to itself. It is easy to prove that Ψ is an isometry of $\mathbb{S}^n$. This family of isometries of $\mathbb{S}^n$ consists of rotations of the sphere and mirror symmetries.

Example. One can prove that the hyperbolic space $\mathbb{H}^n$ is isometric to each of the following Riemannian manifolds $(M, \mathbf{g})$:

1. (Exercise 39) Poincaré disk model: M is the unit ball $B_1 = \{x \in \mathbb{R}^n : |x| < 1\}$ and

$$\mathbf{g} = 4\frac{\mathbf{g}_{\mathbb{R}^n}}{\left(1 - |x|^2\right)^2} = 4\frac{\left(dx^1\right)^2 + ... + \left(dx^n\right)^2}{\left(1 - |x|^2\right)^2};$$

2. (see Exercises) Poincaré half-space model: M is the upper half-space $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x^n > 0\}$ and

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

Example. Let $(M, \mathbf{g})$ be a model manifold with polar coordinates (r, θ) (see Section 1.14) and let Ψ be an isometry of $\mathbb{S}^{n-1}$. Then Ψ induces an isometry of $(M, \mathbf{g})$ by $\Phi(r, \theta) = (r, \Psi(\theta))$.

* **Remark.** 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.

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:

$$J^T g J = g. \tag{1.149}$$

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.

Chapter 2

Weak Laplace operator and spectrum

In this Chapter we study the properties of the Laplace-Beltrami operator $\Delta_{\mathbf{g}}$ on a Riemannian manifold $(M, \mathbf{g})$, in particular, solvability of certain boundary value problems for this operator. We apply a standard approach of PDEs where one first proves the existence of so called *weak solutions* and then prove that they are smooth functions. Recall that the Laplace-Beltrami operator $\Delta_{\mathbf{g}}$ has in the local coordinates $x^1, ..., x^n$ the form

$$\Delta_{\mathbf{g}} u = \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^i} \left(\sqrt{\det g} g^{ij} \frac{\partial u}{\partial x^j} \right) = \frac{\partial}{\partial x^i} \left(g^{ij} \frac{\partial u}{\partial x^j} \right) + \frac{\partial}{\partial x^i} \log \sqrt{\det g} g^{ij} \frac{\partial u}{\partial x^j},$$

where $g = (g_{ij})$ and $g^{-1} = (g^{ij})$. Hence, we start with revision of local properties of such differential operators in $\mathbb{R}^n$.

2.1 Regularity theory in $\mathbb{R}^n$

Elliptic operators of second order. Consider in a domain $\Omega \subset \mathbb{R}^n$ an operator

$$Lu = \sum_{i,j=1}^{n} \partial_i (a_{ij} \partial_j u) + \sum_{i=1}^{n} b_j \partial_j u + cu, \tag{2.1}$$

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, which means that

$$\lambda^{-1} |\xi|^2 \leq \sum_{i,j=1}^{n} a_{ij}(x) \xi_i \xi_j \leq \lambda |\xi|^2 \quad \text{for all } x \in \Omega \text{ and } \xi \in \mathbb{R}^n.$$

Assume also that the coefficients b_j, c are bounded in Ω, say, also by λ.

A function $u \in C^2(\Omega)$ is called a strong (classical) solution of the equation $Lu = f$ if $Lu(x) = f(x)$ for all $x \in \Omega$. We need the notion of a weak solution that requires the use of another function space instead of $C^2(\Omega)$.

Weak derivatives and Sobolev space of 1^{st} order. If $u \in C^1(\Omega)$ and $\varphi \in C_0^1(\Omega)$ then the integration by parts yields the identity

$$\int_\Omega \varphi\, \partial_i u\, dx = -\int_\Omega u\, \partial_i \varphi\, dx.$$

This identity is used to define the notion of a *weak derivative* ∂_i.

Recall that the Lebesgue space $L^2(\Omega)$ consists of measurable functions u in Ω that are square integrable, that is,

$$\int_\Omega u^2 dx < \infty.$$

It is known that $L^2(\Omega)$ is a Hilbert space with the inner product

$$(u,v)_{L^2} = \int_\Omega uv\, dx.$$

Denote by $L^2_{loc}(\Omega)$ the space of functions u such that $u \in L^2(U)$ for any open set $U \Subset \Omega$. For example, if $u \in C(\Omega)$ then $u \in L^2_{loc}(\Omega)$.

Definition. Let $u, v \in L^2_{loc}(\Omega)$. The function v is called the weak derivative ∂_i of u and is denoted by $\partial_i u$ (or sometimes $\partial_i^{\text{weak}} u$) if the following identity is true

$$\int_\Omega \varphi\, v\, dx = -\int_\Omega u\, \partial_i \varphi\, dx \quad \text{for all } \varphi \in C_0^\infty(\Omega). \tag{2.2}$$

Functions φ in this definition are called *test functions*. Note that the both integrals in this definition converge because $\operatorname{supp} \varphi$ is compact. Of course, if $u \in C^1(\Omega)$ then its classical partial derivative $\partial_i u$ is also the weak derivative.

Example. Consider in $\mathbb{R}^n$ the function $u(x) = |x|^s$ where s can be also negative. It is easy to prove that if $s > -\frac{n}{2}$ then $u \in L^2_{loc}(\mathbb{R}^n)$, and if $s > 1 - \frac{n}{2}$ then its classical derivative $\partial_i u$ is also in $L^2_{loc}(\mathbb{R}^n)$ and is the weak derivative. However, $u \notin C^1(\mathbb{R}^n)$ if $s < 0$.

Definition. Define the *Sobolev space* $W^1(\Omega)$ as follows:

$$W^1(\Omega) = \left\{ u \in L^2(\Omega) : \partial_i u \in L^2(\Omega) \text{ for all } i = 1, ..., n \right\},$$

where $\partial_i u$ is the weak derivative.

In this case, we define the weak gradient of u as $\nabla u = (\partial_1 u, ..., \partial_n u)$. Observe that $|\nabla u| \in L^2(\Omega)$.

The notation $W^1(\Omega)$ is similar to $C^1(\Omega)$ where the letter "W" refers to *weak* derivative whereas "C" stands for *continuous* derivative.

The Sobolev space W^1 is a Hilbert space with respect to the inner product

$$(u,v)_{W^1} = \int_\Omega (uv + \nabla u \cdot \nabla v)\, dx = \int_\Omega \left(uv + \sum_{i=1}^n \partial_i u\, \partial_i v \right) dx. \tag{2.3}$$

Hence, the norm in $W^1(\Omega)$ is given by

$$\|u\|_{W^1}^2 = \int_\Omega \left(u^2 + |\nabla u|^2\right) dx.$$

17-Dec-25 Lecture 17

Denote by $W^1_{loc}(\Omega)$ the space of functions f on Ω such that $f \in W^1(U)$ for any open set $U \Subset \Omega$. Clearly, if $u \in C^1(\Omega)$ then $u \in W^1_{loc}(\Omega)$.

Weak derivatives and Sobolev spaces of higher order. Now let us define weak derivatives of higher order. We denoted by α a multiindex $\alpha = (\alpha_1, ..., \alpha_n)$ where α_i are non-negative integers. Set also $|\alpha| = \alpha_1 + ... + \alpha_n$. For any function $u \in C^k(\Omega)$ with $k \geq |\alpha|$, denote by $D^\alpha u$ the following partial derivative of u:

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} ... \partial x_n^{\alpha_n}}$$

that *is* understood in the classical sense. The integration by parts yields the following identity: for any $u \in C^k(\Omega)$ and $\varphi \in C_0^\infty(\Omega)$

$$\int_\Omega \varphi\, D^\alpha u\, dx = (-1)^{|\alpha|} \int_\Omega u\, D^\alpha \varphi\, dx$$

Motivated by this identity, we define $D^\alpha u$ also in the weak sense as follows.

Definition. Let $u, v \in L^2_{loc}(\Omega)$. The function v is called the weak derivative D^α of u and is denoted by $D^\alpha u$ (or sometimes $D_\alpha^{\text{weak}} u$) if the following identity is true

$$\int_\Omega \varphi\, v\, dx = (-1)^{|\alpha|} \int_\Omega u\, D^\alpha \varphi\, dx \quad \text{for all } \varphi \in C_0^\infty(\Omega). \tag{2.4}$$

If $u \in C^{|\alpha|}(\Omega)$ then the classical derivative $D^\alpha u$ gives also the weak derivative.

Observe that if the weak derivative $D^\alpha u$ exists then it is unique. Indeed, if v_1 and v_2 are two functions from $L^2_{loc}(\Omega)$ that satisfy (2.4) then, for all $\varphi \in C_0^\infty(\Omega)$,

$$\int_\Omega \varphi\, v_1\, dx = \int_\Omega \varphi\, v_2\, dx$$

and, hence,

$$\int_\Omega \varphi\,(v_1 - v_2)\, dx = 0,$$

which implies $v_1 = v_2$ a.e.. Therefore, v_1 and v_2 determine the same element in $L^2_{loc}(\Omega)$.

Example. Consider in $\mathbb{R}^n$ the function $u(x) = |x|^s$ where s can be also negative. One can prove that if $s > k - \frac{n}{2}$ then its classical derivative $D^\alpha u$ with $|\alpha| \leq k$ (that is defined in $\mathbb{R}^n \setminus 0$) is also in $L^2_{loc}(\mathbb{R}^n)$ and is the weak derivative of u.

Definition. For any $k \in \mathbb{N} \cup \{0\}$ define the Sobolev space $W^k(\Omega)$ by

$$W^k(\Omega) = \left\{u \in L^2(\Omega) : D^\alpha u \in L^2(\Omega) \text{ for all } \alpha \text{ with } |\alpha| \leq k\right\}.$$

One can prove that $W^k(\Omega)$ is the Hilbert space with the following inner product:

$$(u, v)_{W^k} = \int_\Omega \sum_{\{\alpha : |\alpha| \leq k\}} D^a u \, D^\alpha v \, dx.$$

The corresponding norm is

$$\|u\|^2_{W^k} = \int_\Omega \sum_{\{\alpha : |\alpha| \leq k\}} (D^a u)^2 \, dx.$$

Denote by $W^k_{loc}(\Omega)$ the space of functions f such that $f \in W^k(U)$ for any open set $U \Subset \Omega$. Clearly, if $u \in C^k(\Omega)$ then $u \in W^k_{loc}(\Omega)$. As follows from the above example, the function $u(x) = |x|^s$ belongs to $W^k_{loc}(\Omega)$ provided $s > k - \frac{n}{2}$.

Weak solutions of $Lu = f$. Let us return to the uniformly elliptic operator L given by (2.1), that is,

$$Lu = \sum_{i,j=1}^n \partial_i (a_{ij} \partial_j u) + \sum_{i=1}^n b_j \partial_j u + cu, \tag{2.5}$$

where a_{ij}, b_j, c are smooth functions in $\Omega \subset \mathbb{R}^n$. As before, denote by λ the ellipticity constant of L that also bounds $|b_j|$ and $|c|$ in Ω.

Assume that $u \in C^2(\Omega)$ satisfies the equation $Lu = f$ in Ω. Multiplying this equation by a test function $\varphi \in C_0^\infty(\Omega)$ and integrating over Ω, we obtain

$$\begin{aligned}
\int_\Omega f\varphi dx &= \int_\Omega \left(\sum_{i,j=1}^n \partial_i (a_{ij}\partial_j u) + \sum_{i=1}^n b_j \partial_j u + cu\right) \varphi \, dx \\
&= -\int_\Omega \sum_{i,j=1}^n a_{ij} \partial_j u \, \partial_i \varphi \, dx + \int_\Omega \sum_{j=1}^n b_j \partial_j u \, \varphi \, dx + \int_\Omega c \, u \, dx.
\end{aligned}$$

This identity motivates us for the following definition.

Definition. Let $u \in W^1_{loc}(\Omega)$ and $f \in L^2_{loc}(\Omega)$. We say that the equation $Lu = f$ is satisfied weakly in Ω if, for any $\varphi \in C_0^\infty(\Omega)$,

$$-\int_\Omega \sum_{i,j=1}^n a_{ij} \partial_j u \, \partial_i \varphi \, dx + \int_\Omega \sum_{j=1}^n b_j \partial_j u \, \varphi \, dx + \int_\Omega c \, u \, dx = \int_\Omega f\varphi dx.$$

Of course, if $u \in C^2(\Omega)$ and $f \in C(\Omega)$ then u satisfies $Lu = f$ weakly if and only if it satisfies the equation $Lu = f$ strongly.

Let us cite some results about weak solutions that were proved in the course Elliptic Differential Equations.

Theorem 2.1 *Let L be the operator (2.5). If $u \in W^1_{loc}(\Omega)$ and $f = Lu \in W^k_{loc}(\Omega)$ then $u \in W^{k+2}_{loc}(\Omega)$. Moreover, for any open set $U \Subset \Omega$,*

$$\|u\|_{W^{k+2}(U)} \leq C \left(\|u\|_{W^1(\Omega)} + \|Lu\|_{W^k(\Omega)} \right), \tag{2.6}$$

where $C = C(U, \Omega, n, \lambda)$.

Example. If $u \in W^1_{loc}(\Omega)$ and $f = Lu \in L^2_{loc}(\Omega)$ (case $k = 0$) then $u \in W^2_{loc}(\Omega)$. In fact, this means that the weak operator L can be understood as is stated in (2.5) where all partial derivatives ∂_i and ∂_j are understood in the weak sense.

For passage from weak derivative to the classical ones, we use the following theorem.

Theorem 2.2 *(*Sobolev Embedding Theorem*) If Ω is an open subset of $\mathbb{R}^n$ and l, m are non-negative integers such that*

$$l > m + \frac{n}{2},$$

then

$$u \in W^l_{loc}(\Omega) \Rightarrow u \in C^m(\Omega).$$

Moreover, for any open set $U \Subset \Omega$,

$$\|u\|_{C^m(U)} \leq C \|u\|_{W^l(\Omega)}.$$

Recall that

$$\|u\|_{C^m(U)} = \sup_{|\alpha| \leq m,\ x \in U} |D^\alpha u(x)|.$$

Example. Assume that $u \in W^1_{loc}(\Omega)$ and $f = Lu \in W^k_{loc}(\Omega)$ with $k > \frac{n}{2}$. By Theorem 2.1, we have $u \in W^l_{loc}(\Omega)$ where $l = k + 2$. Since

$$l = k + 2 > 2 + \frac{n}{2} = m + \frac{n}{2},$$

we conclude by Theorem 2.2 that $u \in C^2(\Omega)$. Consequently, the equation $Lu = f$ is satisfied in the classical sense.

Iterated operator L^k. Our purpose is to develop a similar theory of weak solutions on manifolds. The spaces $W^1(\Omega)$ can be easily defined if Ω is an open subset of a Riemannian manifold, by using the Riemannian gradient. However, the higher order Sobolev spaces $W^k(\Omega)$ require higher order derivatives that can be reasonably defined only in charts. Instead of second order derivatives, we will use the Laplace-Beltrami operator $\Delta_{\mathbf{g}}$, and instead of higher order derivatives we will use iterated Laplacian $\Delta^k_{\mathbf{g}}$.

Coming back to the operator L from (2.5), let us first define the notation

$$L^k u := \underbrace{L(L(...Lu))}_{k \text{ times } L}.$$

Of course, the operator L^k is well-defined on functions $u \in C^\infty(M)$.

Definition. Assuming $u \in W^1_{loc}(\Omega)$ 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

$$L^{k-1}u \in W^1_{loc}(\Omega) \quad \text{and} \quad L\left(L^{k-1}u\right) = f \text{ weakly,}$$

where $L^{k-1}u$ is defined by the inductive hypotheses.

For example, $L^2 u = f$ means that $Lu \in W^1_{loc}(\Omega)$ and $L(Lu) = f$.

Hence, for any $k \geq 1$, the equality $L^k u = f$ assumes that all the functions $u, Lu, ..., L^{k-1}u$ belong to $W^1_{loc}(\Omega)$, and $L\left(L^{k-1}u\right) = f$.

For $k = 0$, set $L^k = \mathrm{id}$.

Corollary 2.3 *Let L be the operator (2.5). If*

$$u, Lu, ..., L^k u \in W^1_{loc}(\Omega)$$

then

$$u \in W^{2k+1}_{loc}(\Omega).$$

Moreover, for any open set $U \Subset \Omega$,

$$\|u\|_{W^{2k+1}(U)} \leq C \sum_{j=0}^{k} \left\|L^j u\right\|_{W^1(\Omega)}, \tag{2.7}$$

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 $k \geq 1$. Set $v = Lu$. Then we have

$$v, Lv, ..., L^{k-1}v \in W^1_{loc}(\Omega),$$

and we conclude by the inductive hypothesis that $v \in W^{2k-1}_{loc}(\Omega)$. That is, $Lu = v \in W^{2k-1}_{loc}(\Omega)$, which yields by Theorem 2.1 that $u \in W^{2k+1}_{loc}(\Omega)$.

To prove (2.7), choose an open set V such that $U \Subset V \Subset \Omega$. By the inductive hypothesis, we have

$$\|Lu\|_{W^{2k-1}(V)} = \|v\|_{W^{2k-1}(V)} \leq C \sum_{j=0}^{k-1} \left\|L^j v\right\|_{W^1(\Omega)} = C \sum_{j=1}^{k} \left\|L^j u\right\|_{W^1(\Omega)}.$$

Combining this with Theorem 2.1, we obtain

$$\|u\|_{W^{2k+1}(U)} \leq C \left(\|u\|_{W^1(V)} + \|Lu\|_{W^{2k-1}(V)}\right) \leq C' \sum_{j=0}^{k} \left\|L^j u\right\|_{W^1(\Omega)},$$

which was to be proved. ■

Corollary 2.4 *Let L be the operator (2.5). If*

$$u, Lu, ..., L^k u \in W^1_{loc}(\Omega)$$

and, for some non-negative integer m,

$$2k + 1 > m + \frac{n}{2},$$

then

$$u \in C^m(\Omega).$$

Moreover, for any open set $U \Subset \Omega$,

$$\|u\|_{C^m(U)} \leq C \sum_{j=0}^{k} \left\|L^j u\right\|_{W^1(\Omega)}, \tag{2.8}$$

where $C = C(U, \Omega, n, k, m, \lambda, \rho)$.

Proof. By Corollary 2.3, we have $u \in W^{2k+1}_{loc}(\Omega)$. Since $2k + 1 > m + \frac{n}{2}$, we conclude by Theorem 2.2 with $l = 2k + 1$ that $u \in C^m(\Omega)$.

Now let V be any open set such that $U \Subset V \Subset \Omega$. By Theorem 2.2 we have

$$\|u\|_{C^m(U)} \leq C \|u\|_{W^{2k+1}(V)},$$

while by (2.7)

$$\|u\|_{W^{2k+1}(V)} \leq C \sum_{j=0}^{k} \left\|L^j u\right\|_{W^1(\Omega)}.$$

Combining these two inequalities, we obtain (2.8). ∎

19-Dec-25 **Lecture 18**

2.2 Weak gradient and Sobolev spaces

Let $(M, \mathbf{g}, \mu)$ be a weighted manifold. Denote by $\vec{L}^2(M, \mu)$ the space of all measurable vector fields[1] $v(x)$ on M such that $|v| \in L^2(M, \mu)$, where $|v| = |v|_{\mathbf{g}}$.

Similarly we define $\vec{L}^2_{loc}(M, \mu)$ as the space of all measurable vector field $v(x)$ on M such that $|v| \in L^2_{loc}(M, \mu)$.

The space $\vec{L}^2(M, \mu)$ admits an inner product

$$(v, w)_{\vec{L}^2} := \int_M \langle v, w \rangle \, d\mu,$$

[1] A vector field v on M is called measurable if all the components of v in any chart are measurable functions.

where $\langle v, w\rangle = \langle v, w\rangle_{\mathbf{g}}$, and the corresponding norm is

$$\|v\|^2_{\vec{L}^2} = \int_M |v|^2 \, d\mu.$$

It is easy to prove that $\vec{L}^2$ is complete with respect to this norm and, hence, is a Hilbert space.

Recall the divergence theorem: if u is a smooth function on M, ψ is a smooth vector field on M, and either u or ψ has a compact support then

$$\int_M u \operatorname{div} \psi \, d\mu = -\int_M \langle \nabla u, \psi\rangle d\mu,$$

where $\operatorname{div} = \operatorname{div}_{\mathbf{g},\mu}$ and $\nabla = \nabla_{\mathbf{g}}$. This identity was used to give definition of the divergence $\operatorname{div} \psi$. Now we use this identity to define a weak gradient.

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.

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 $\psi \in \vec{\mathcal{D}}$,

$$\int_M u \operatorname{div} \psi \, d\mu = -\int_M \langle v, \psi\rangle \, d\mu. \tag{2.9}$$

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 unique when it exists. Note u is a smooth function then the classical gradient $v = \nabla u$ satisfies (2.9) by the divergence theorem, so that in this case the weak gradient exists and coincides with the classical gradient.

Definition. Define the Sobolev space $W^1(M)$ by

$$W^1(M) = W^1(M, \mathbf{g}, \mu) := \left\{u \in L^2(M,\mu) : \nabla u \in \vec{L}^2(M,\mu)\right\}$$

and the inner product in W^1:

$$(u,v)_{W^1} := (u,v)_{L^2} + (\nabla u, \nabla v)_{\vec{L}^2} = \int_M uv \, d\mu + \int_M \langle \nabla u, \nabla v\rangle \, d\mu. \tag{2.10}$$

The associated norm given by

$$\|u\|^2_{W^1} = \|u\|^2_{L^2} + \|\nabla u\|^2_{L^2} = \int_M u^2 d\mu + \int_M |\nabla u|^2 \, d\mu. \tag{2.11}$$

Lemma 2.5 *$W^1(M)$ is a Hilbert space.*

Proof. It follows from (2.11) that the convergence $u_k \xrightarrow{W^1} u$ in $W^1(M)$ is equivalent to

$$u_k \xrightarrow{L^2} u \quad \text{and} \quad \nabla u_k \xrightarrow{L^2} \nabla u. \tag{2.12}$$

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 $v \in \vec{L}^2(M)$:

$$u_k \xrightarrow{L^2} u \quad \text{and} \quad \nabla u_k \xrightarrow{L^2} v. \tag{2.13}$$

It remains to verify that v is the weak gradient of u, which will imply (2.12). By the definition of the weak gradient of u_k, we have for any $\psi \in \vec{\mathcal{D}}$ that

$$(u_k, \operatorname{div}\psi)_{L^2} = -(\nabla u_k, \psi)_{\vec{L}^2}$$

Passing to the limit as $k \to \infty$, we obtain

$$(u, \operatorname{div}\psi)_{L^2} = -(v, \psi)_{\vec{L}^2}$$

that is, $v = \nabla u$, which finishes the proof. ■

Since any open set $U \subset M$ is itself a weighted manifold, we can define also the spaces $L^2(U)$ and $W^1(U)$.

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 $L^2_{\mathbf{e}}(U)$ and $W^1_{\mathbf{e}}(U)$, 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. In the next statement we use the relation $A \simeq B$ between two non-negative functions that means that $A \le CB$ and $B \le CA$ for some positive constant C that is the same for the specified range of variables.

Lemma 2.6 *Let U be a precompact chart in M. Then we have the following.*

(a) *For all measurable functions u on U,*

$$\|u\|_{L^2(U)} \simeq \|u\|_{L^2_{\mathbf{e}}(U)}.$$

Consequently, $L^2(U) = L^2_{\mathbf{e}}(U)$.

(b) *For all measurable vector fields v on U,*

$$\|v\|_{\vec{L}^2(U)} \simeq \|\widetilde{v}\|_{\vec{L}^2_{\mathbf{e}}(U)}$$

where $\widetilde{v} = (v_1, ..., v_n)$ and v_i are the covector components of v in the chart U. Hence, $\vec{L}^2(U) = \vec{L}^2_{\mathbf{e}}(U)$.

(c) *Let $u \in W^1(U)$ and let $v = \nabla_{\mathbf{g}} u$ be the weak gradient of u. Then $\frac{\partial u}{\partial x^i} = v_i$ weakly in U where v_i is the covector component of v.*

(d) $W^1(U) = W^1_{\mathbf{e}}(U)$ *and*

$$\|u\|_{W^1(U)} \simeq \|u\|_{W^1_{\mathbf{e}}(U)}.$$

Proof. (a) 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, $L^2(U) = L^2_{\mathbf{e}}(U)$.

(b) Let v be a measurable vector field on U. Denote by v^i be the components of v in the basis $\left\{\frac{\partial}{\partial x^i}\right\}$. Then $v_i = g_{ij}v^j$ be the covector components of v. We have

$$\|v\|^2_{\vec{L}^2} = \int_U |v|^2_{\mathbf{g}} \, d\mu = \int_U g^{ij} v_i v_j \rho \, d\lambda.$$

Since the matrix (ρg^{ij}) is uniformly elliptic[2] in U, we obtain that, for all $x \in U$,

$$\rho g^{ij}(x) v_i v_j \simeq v_1^2 + ... + v_n^2,$$

whence

$$\|v\|^2_{\vec{L}^2} \simeq \int_U \left(v_1^2 + ... + v_n^2\right) d\lambda = \|\widetilde{v}\|^2_{\vec{L}^2_{\mathbf{e}}}.$$

Hence, identifying each vector field v in U with the Euclidean vector field $\widetilde{v}$, we obtain the identity

$$\vec{L}^2(U) = \vec{L}^2_{\mathbf{e}}(U).$$

(c) Let $u \in W^1(U)$ and $v = \nabla_{\mathbf{g}} u$ be its week gradient. By (2.9) we have

$$\int_M u \operatorname{div} \psi \, d\mu = -\int_M \langle v, \psi \rangle_{\mathbf{g}} \, d\mu$$

for any vector field $\psi = \psi^i \frac{\partial}{\partial x^i} \in \vec{\mathcal{D}}(U)$. In the local coordinates, 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

$$\int_M \langle v, \psi \rangle \, d\mu = \int_U g_{ij} v^j \psi^i \rho d\lambda = \int_U v_i \left(\psi^i \rho\right) d\lambda.$$

Comparing these two identities and renaming $\rho\psi$ into ψ, we obtain

$$\int_U u \frac{\partial \psi^i}{\partial x^i} d\lambda = -\int_U v_i \psi^i d\lambda. \tag{2.14}$$

[2] If $(a_{ij}(x))$ is a positive definite matrix for any $x \in V$ (where V is an open subset of $\mathbb{R}^n$) and if $a_{ij} \in C(V)$ then (a_{ij}) is uniformly elliptic on any compact subset $K \subset V$. Indeed, as the trace $\sum_i a_{ii}$ is uniformly bounded on K, it follows that the maximal eigenvalue of (a_{ij}) is also uniformly bounded on K. Applying the same argument to the inverse matrix $(a_{ij})^{-1}$, we obtain that its maximal eigenvalue is also uniformly bounded from above, which implies that the minimal eigenvalue of (a_{ij}) is bounded from below by a positive constant, whence the uniform ellipticity follows.

Fix $\varphi \in \mathcal{D}(U)$, an index i, and choose the vector field ψ as follows: $\psi_i = \varphi$ 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 weak derivative $\frac{\partial u}{\partial x^i}$ (in $U \subset \mathbb{R}^n$), so that

$$\frac{\partial u}{\partial x^i} = v_i.$$

(d) Assume that $u \in W^1(U)$ and set $v = \nabla_{\mathbf{g}} u$. By (a) we have $u \in L^2(U) = L^2_{\mathbf{e}}(U)$. Since $v \in \vec{L}^2(U)$, by (b) we have $\widetilde{v} \in \vec{L}^2_{\mathbf{e}}(U)$. Since $\frac{\partial u}{\partial x^i} = v_i$ by (c), it follows that $u \in W^1_{\mathbf{e}}(U)$. Using again (a) and (b), we obtain

$$\|u\|^2_{W^1_{\mathbf{e}}} = \|u\|^2_{L^2_{\mathbf{e}}} + \|\widetilde{v}\|^2_{\vec{L}^2_{\mathbf{e}}} \simeq \|u\|^2_{L^2} + \|v\|^2_{\vec{L}^2} = \|u\|^2_{W^1}.$$

Conversely, assume that $u \in W^1_{\mathbf{e}}(U)$. Then $\frac{\partial u}{\partial x^i} \in L^2_{\mathbf{e}}(U)$ and, for any $\psi \in \vec{\mathcal{D}}(U)$,

$$\int_U u \frac{\partial \psi^i}{\partial x^i} d\lambda = -\int_U \frac{\partial u}{\partial x^i} \psi^i d\lambda.$$

Define in U a vector field v with covector components

$$v_i = \frac{\partial u}{\partial x^i}.$$

Clearly, $v \in \vec{L}^2(U)$ and v satisfies (2.14), which implies that v is the weak gradient $\nabla_{\mathbf{g}} u$. It follows that $u \in W^1(U)$, and we conclude that

$$W^1_{\mathbf{e}}(U) = W^1(U).$$

■

07-Jan-26 Lecture 19

2.3 Weak Laplacian

Let $(M, \mathbf{g}, \mu)$ be a weighted manifold. Here discuss the weighted Laplacian $\Delta = \Delta_{\mathbf{g},\mu}$ in the weak sense.

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 $\varphi \in \mathcal{D}(\Omega)$,

$$\int_\Omega \langle \nabla u, \nabla \varphi \rangle \, d\mu = -\int_\Omega f \, \varphi d\mu, \tag{2.15}$$

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

$$u,\Delta u,...,\Delta^k u\in W^1_{loc}(\Omega)\,, \tag{2.16}$$

where Δ is understood in the weak sense. Then $u\in C^m(\Omega)$. Moreover, for any precompact chart $U\Subset\Omega$,

$$\|u\|_{C^m(U)}\leq C\sum_{j=0}^{k}\left\|\Delta^j u\right\|_{W^1(\Omega)}, \tag{2.17}$$

where $C=C(U,\Omega,n,k,m,\mathbf{g},D)$.

Proof. Let us choose a precompact chart V such that $U\Subset V\Subset\Omega$. Let $x^1,...,x^n$ be the coordinates in V. Consider in V the following differential operator

$$L=\frac{1}{\rho}\frac{\partial}{\partial x^i}\left(\rho g^{ij}\frac{\partial}{\partial x^j}\right)=\frac{\partial}{\partial x^i}\left(g^{ij}\frac{\partial}{\partial x^j}\right)+\frac{1}{\rho}\frac{\partial}{\partial x^i}\left(\rho g^{ij}\right)\frac{\partial}{\partial x^j},$$

where $\rho=D\sqrt{\det g}$. We know that $Lu=\Delta u$ for all $u\in C^\infty(V)$. Let us show that $Lu=\Delta u$ holds also for $u\in W^1(V)$ when L and Δ are understood weakly.

By Lemma 2.6,

$$L^2(V)=L^2_{\mathbf{e}}(V)\quad\text{and}\quad W^1(V)=W^1_{\mathbf{e}}(V)\,.$$

Let $u\in W^1(V)$ and $f\in L^2(V)$. By definition, the equation $Lu=f$ weakly in V means that

$$\frac{\partial}{\partial x^i}\left(\rho g^{ij}\frac{\partial}{\partial x^j}\right)=\rho f\quad\text{weakly in }V,$$

that is, for all $\varphi\in\mathcal{D}(V)$,

$$\int_V\sum_{i,j=1}^{n}\rho g^{ij}\partial_j u\,\partial_i\varphi d\lambda=-\int_V\rho f\varphi d\lambda. \tag{2.18}$$

By (2.15) the equation $\Delta u=f$ weakly in V means that, for all $\varphi\in\mathcal{D}(V)$,

$$\int_V\langle\nabla u,\nabla\varphi\rangle\,d\mu=-\int_V f\,\varphi d\mu. \tag{2.19}$$

Since

$$\langle\nabla u,\nabla\varphi\rangle=\langle du,d\varphi\rangle=g^{ij}\frac{\partial u}{\partial x^j}\frac{\partial\varphi}{\partial x^i}$$

and

$$d\mu=\rho d\lambda,$$

we obtain that (2.19) is equivalent to

$$\int_V \rho g^{ij} \frac{\partial u}{\partial x^j} \frac{\partial \varphi}{\partial x^i} d\lambda = -\int_V \rho f \varphi d\lambda,$$

that is, to (2.18), which proves that the weak operators Δu and Lu coincide.

If (2.16) is satisfied then

$$u, \Delta u, ..., \Delta^k u \in W^1(V),$$

whence also

$$u, Lu, ..., L^k u \in W^1_{\mathbf{e}}(V).$$

Since $2k + 1 > \frac{n}{2} + m$, we obtain by Corollary 2.4 that $u \in C^m(V)$ and

$$\|u\|_{C^m(U)} \le C \sum_{j=0}^{k} \left\| L^j u \right\|_{W^1_{\mathbf{e}}(V)}. \tag{2.20}$$

Since Ω can be covered by charts like V, we conclude that $u \in C^m(\Omega)$. The estimate (2.20) and $\|\cdot\|_{W^1_{\mathbf{e}}(V)} \simeq \|\cdot\|_{W^1(V)}$ (cf. Lemma 2.6) imply (2.17). ■

Remark. In the case $m = 0$ the expression

$$\|u\|_{C(U)} := \sup_U |u|$$

makes sense for any precompact open set $U \Subset \Omega$, that is not necessarily a chart. In this case the estimate (2.17) holds also for any precompact open set U, because $\overline{U}$ can be covered by a finite number of precompact charts, and in each of them we can apply Theorem 2.7.

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 the weak sense, where α is a real number. Then $u \in C^\infty(\Omega)$. Moreover, for any precompact chart $U \Subset \Omega$ and for any non-negative integer m, we have

$$\|u\|_{C^m(U)} \le C (1 + |\alpha|)^{\frac{m}{2} + \frac{n}{4} + \frac{1}{2}} \|u\|_{W^1(\Omega)}, \tag{2.21}$$

where $C = C(U, \Omega, n, m, \mathbf{g}, D)$.

Proof. Since $u \in W^1(\Omega)$, we have also $\Delta u = \alpha u \in W^1(\Omega)$. It follows that also $\Delta^2 u = \Delta(\alpha u) = \alpha^2 u \in W^1(\Omega)$ and, by induction, for any positive integer j, we obtain

$$\Delta^j u = \alpha^j u \in W^1(\Omega).$$

By Theorem 2.7 with arbitrarily large m, 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

$$2k+1>m+\frac{n}{2}, \tag{2.22}$$

that

$$\|u\|_{C^m(U)} \leq C\sum_{j=0}^{k}\left\|\Delta^j u\right\|_{W^1(\Omega)},$$

Since

$$\left\|\Delta^j u\right\|_{W^1(\Omega)} = |\alpha|^j \|u\|_{W^1(\Omega)},$$

it follows that

$$\|u\|_{C^m(U)} \leq C\sum_{j=0}^{k}|\alpha|^j \|u\|_{W^1(\Omega)} \leq C\left(1+|\alpha|\right)^k \|u\|_{W^1(\Omega)}.$$

Choose k to be the smallest integer such that (2.22) holds. Then

$$2k-1\leq m+\frac{n}{2},$$

and, hence,

$$k\leq \frac{m}{2}+\frac{n}{4}+\frac{1}{2},$$

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

$$\begin{aligned}\|u\|_{L^2}^2 &= \int_0^{2\pi}\sin^2\beta x dx = \pi \\ \|u'\|_{L^2}^2 &= \beta^2\int_0^{2\pi}\cos^2\beta x dx = \beta^2\pi = |\alpha|\pi\end{aligned}$$

whence

$$\|u\|_{W^1}^2 = \|u\|_{L^2}^2 + \|u'\|_{L^2}^2 = \pi\left(1+|\alpha|\right)$$

and

$$\|u\|_{W^1} = \pi^{1/2}\left(1+|\alpha|\right)^{1/2}.$$

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 in $(0, 2\pi)$. Since

$$u^{(j)}(x) = \pm\beta^j\sin\beta x \quad \text{or} \quad \pm\beta^j\cos\beta x,$$

it follows that

$$\|u\|_{C^m(0,2\pi)} = \sup_{0\leq j\leq m}\sup_{(0,2\pi)}\left|u^{(j)}\right| = \sup_{0\leq j\leq m}\beta^j = \beta^m = |\alpha|^{m/2} \simeq \left(1+|\alpha|\right)^{m/2}.$$

It follows that

$$\|u\|_{C^m(0,2\pi)} \simeq \left(1+|\alpha|\right)^{\frac{m}{2}-\frac{1}{2}}\|u\|_{W^1(0,2\pi)},$$

which shows that the term $\left(1+|\alpha|\right)^{m/2}$ in (2.21) is an optimal one.

2.4 Resolvent operator

Fix an open set $\Omega \subset M$ and consider the following Dirichlet problem

$$\begin{cases} \Delta u - \alpha u = -f & \text{in } \Omega, \\ u \in W_0^1(\Omega), \end{cases} \tag{2.23}$$

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 (2.23) if $\Delta u = \alpha u + f$ weakly in Ω; equivalently, this means that, for any $\varphi \in \mathcal{D}(\Omega)$,

$$(\nabla u, \nabla \varphi)_{\vec{L}^2} + \alpha (u, \varphi)_{L^2} = (f, \varphi)_{L^2}. \tag{2.24}$$

Recall that $\mathcal{D}(\Omega) \subset W^1(\Omega)$. Define

$$W_0^1(\Omega) = \text{the closure of } \mathcal{D}(\Omega) \text{ in } W^1(\Omega).$$

Lemma 2.9 *If (2.24) holds for all $\varphi \in \mathcal{D}(\Omega)$ then it holds also for all $\varphi \in W_0^1(\Omega)$.*

Proof. All terms in (2.24) as functions of φ are bounded linear functionals of $\varphi \in W^1(\Omega)$, because

$$|(f, \varphi)| \le \|f\|_{L^2} \|\varphi\|_{L^2} \le \|f\|_{L^2} \|\varphi\|_{W^1}$$

and similarly

$$|(\nabla u, \nabla \varphi)_{\vec{L}^2}| \le \|\nabla u\|_{\vec{L}^2} \|\nabla \varphi\|_{\vec{L}^2} \le \|u\|_{W^1} \|\varphi\|_{W^1}.$$

Hence, all terms in (2.24) are continuous in $\varphi \in W^1(\Omega)$. If (2.24) 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 $W_0^1(\Omega)$. ∎

Theorem 2.10 (The resolvent operator)

(*a*) *For any $\alpha > 0$ and all $f \in L^2(\Omega)$, the problem (2.23) has a unique solution u.*

(*b*) *Define the resolvent operator R_α by*

$$R_\alpha : L^2(\Omega) \to L^2(\Omega)$$
$$R_\alpha f = u,$$

where u is the solution of (2.23). Then the operator R_α is

- *linear,*
- *bounded with the norm estimate $\|R_\alpha\| \le \alpha^{-1}$,*
- *injective,*
- *positive definite,*
- *self-adjoint in $L^2(\Omega)$.*

Proof. (a) Denote the left hand side of (2.24) 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 (2.24) in the form

$$[u,\varphi]_\alpha = (f,\varphi)_{L^2} \quad \forall\varphi \in W_0^1(\Omega)\,. \tag{2.25}$$

Since the right hand side $\varphi \mapsto (f,\varphi)_{L^2}$ is a bounded functional of $\varphi \in W_0^1$, the equation (2.25) has a unique solution $u \in W_0^1(\Omega)$ by the Riesz representation theorem.

Recall that 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 $u \in H$.

09-Jan-26 Lecture 20

(b) Substituting $\varphi = u$ in (2.24) we obtain

$$\|\nabla u\|_{L^2}^2 + \alpha\|u\|_{L^2}^2 = (f,u)_{L^2}\,. \tag{2.26}$$

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 (2.24) 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 (2.26) that if $f \ne 0$ then

$$(R_\alpha f,f)_{L^2} = (u,f)_{L^2} = \|\nabla u\|_{L^2}^2 + \alpha\|u\|_{L^2}^2 > 0,$$

because $u \ne 0$ 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 (2.24), 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 symmetric in f, g, which implies that R_α is symmetric. ∎

2.5 Compactness of resolvent

Recall that a linear operator $A : X \to Y$ in Banach spaces X, Y is called compact if, for any bounded set $S \subset X$, its image $f(S)$ is precompact in Y.

Equivalently, for any bounded sequence $\{x_k\}$ in X, the sequence $\{y_k\}$ with $y_k = f(x_k)$ is contains a convergent subsequence in Y.

Theorem 2.11 (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.

Remark. The embedding operator $I : W_0^1(\Omega) \to L^2(\Omega)$ is defined simply by $I(f) = f$ as W_0^1 is a subspace of L^2.

Proof. In the case when $M = \mathbb{R}^n$ this theorem is well known, and we will use it in order to prove that on an arbitrary manifold.

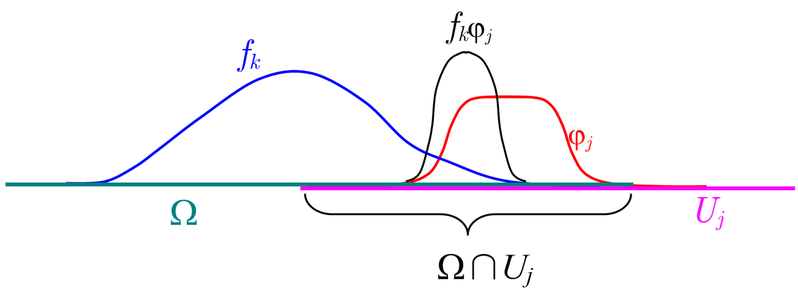
We need to prove that, for any bounded sequence $\{f_k\}$ in $W_0^1(\Omega)$, there is a subsequence $\{f_{k_i}\}$ that converges in $L^2(\Omega)$. Let us emphasize that the boundedness of $\{f_k\}$ is assumed in the W^1 norm while the convergence of $\{f_{k_i}\}$ must be proved in the weaker L^2 norm.

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, it follows from Lemma 1.1 that 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 $\overline{\Omega}$.

For any pair k, j, consider the function $f_k \varphi_j$ and observe that $\operatorname{supp}(f_k \varphi_j) \subset \Omega \cap U_j$ and, hence, $f_k \varphi_j \in \mathcal{D}(\Omega \cap U_j)$.



Let us prove that, for any fixed j, the sequence $\{f_k\varphi_j\}_{k=1}^{\infty}$ is bounded in $W^1(\Omega)$. Indeed, suppressing indices k, j, we have

$$\|f\varphi\|_{L^2} \le \sup|\varphi| \|f\|_{L^2} \le \|f\|_{L^2} \le \|f\|_{W^1}$$

and

$$\begin{aligned}\|\nabla(f\varphi)\|_{\vec{L}^2} &= \|\varphi\nabla f + f\nabla\varphi\|_{\vec{L}^2} \\ &\le \sup\varphi \|\nabla f\|_{\vec{L}^2} + \sup|\nabla\varphi| \|f\|_{L^2} \\ &\le C\|f\|_{W^1},\end{aligned}$$

where $C = 1 + \sup|\nabla\varphi| < \infty$. It follows that

$$\|f_k\varphi_j\|_{W^1(\Omega)} \le C' \|f_k\|_{W^1(\Omega)},$$

which implies that, for any j, the sequence $\{f_k\varphi_j\}_{k=1}^{\infty}$ is bounded in $W^1(\Omega)$.

Therefore, the sequence $\{f_k\varphi_j\}_{k=1}^{\infty}$ is bounded also in $W^1(\Omega \cap U_j)$ and, hence, in $W^1(U_j) = W^1_{\mathbf{e}}(U_j)$. Since U_j is a precompact open subset of $\mathbb{R}^n$, by the compact embedding theorem in $\mathbb{R}^n$ we conclude that there is a subsequence $\{f_{k_i}\varphi_j\}_{i=1}^{\infty}$ 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 $\{f_{k_i}\varphi_j\}_{i=1}^{\infty}$ converges in $L^2(\Omega)$.

Applying this procedure successively for each $j = 1, ..., N$, we obtain a subsequence $\{f_{k_i}\}$ such that $\{f_{k_i}\varphi_j\}_{i=1}^{\infty}$ converges in $L^2(\Omega)$ for *any* j. Since $\sum_{j=1}^{N}\varphi_j \equiv 1$ in Ω and, hence,

$$f_{k_i} = \sum_{j=1}^{N} f_{k_i}\varphi_j,$$

we conclude that $\{f_{k_i}\}$ converges in $L^2(\Omega)$, which finishes the proof. ∎

Theorem 2.12 *If $\Omega \subset M$ is precompact then the resolvent operator R_α in Ω is a compact operator in $L^2(\Omega)$ for any $\alpha > 0$.*

Proof. Consider an operator $\widetilde{R}_\alpha$ defined by

$$\widetilde{R}_\alpha : L^2(\Omega) \to W^1_0(\Omega)$$
$$\widetilde{R}_\alpha f = u$$

and first prove that this operator is bounded (which is a stronger statement than the boundedness of R_α). Recall that by (2.26)

$$\|\nabla u\|_{L^2}^2 + \alpha \|u\|_{L^2}^2 = (f, u)_{L^2},$$

which implies that

$$\|u\|_{W^1}^2 \leq C \|f\|_{L^2} \|u\|_{L^2},$$

where $C = \max(1, \alpha^{-1})$. Since by Theorem 2.10 $\|u\|_{L^2} \leq \alpha^{-1} \|f\|_{L^2}$, we obtain that

$$\|u\|_{W^1} \leq C' \|f\|_{L^2},$$

where $C' = (C\alpha^{-1})^{1/2}$. It follows that

$$\|\widetilde{R}_\alpha\| := \sup_{f \in L^2(\Omega)\setminus\{0\}} \frac{\|\widetilde{R}_\alpha f\|_{W^1}}{\|f\|_{L^2}} \leq C' < \infty,$$

that is, $\widetilde{R}_\alpha$ is a bounded operator.

The operator R_α can be represented as the following composition

$$L^2(\Omega) \stackrel{\widetilde{R}_\alpha}{\to} W_0^1(\Omega) \stackrel{I}{\hookrightarrow} L^2(\Omega),$$

where I is the identical embedding. Since $\widetilde{R}_\alpha$ is a bounded operator and I is compact by Theorem 2.11, we conclude that $R_\alpha = I \circ \widetilde{R}_\alpha$ is compact, because the composition of bounded and compact operators is always a compact operator. ■

14-Jan-26 Lecture 21

2.6 Eigenvalue problem

Consider in an open set $\Omega \subset M$ the following weak eigenvalue problem:

$$\begin{cases} \Delta v + \lambda v = 0 \ \text{ weakly in } \Omega, \\ v \in W_0^1(\Omega) \setminus \{0\}, \end{cases} \tag{2.27}$$

where $\lambda \in \mathbb{R}$ is a *spectral parameter*. Any solution v to (2.27) is called a (Dirichlet) eigenfunction of Δ in Ω, and the corresponding value of λ – a (Dirichlet) eigenvalue of Δ in Ω. Let us emphasize that a Dirichlet eigenfunction must be a non-zero element of $W_0^1(\Omega)$.

Recall that the equation in (2.27) means that for any $\varphi \in W_0^1(\Omega)$

$$\boxed{(\nabla v, \nabla \varphi)_{\vec{L}^2} = \lambda (v, \varphi)_{L^2}.} \tag{2.28}$$

If λ is an eigenvalue of Δ in Ω, then consider the *eigenspace*

$$E_\lambda = \left\{ v \in W_0^1(\Omega) : \Delta v + \lambda v = 0 \right\}.$$

Claim. *E_λ is a closed subspace of $W_0^1(\Omega)$.*

Proof. Clearly, E_λ is a subspace of $W_0^1(\Omega)$: if $v_1, v_2 \in E_\lambda$ then also $\alpha v_1 + \beta v_2 \in E_\lambda$ for any real α, β. Let us prove that E_λ is a closed set. For any fixed $\varphi \in W_0^1(\Omega)$, the both sides of (2.28) are continuous functionals of $v \in W_0^1(\Omega)$, which implies the set of v, satisfying (2.28) for a fixed φ, is closed. Since E_λ is the intersection of all these closed sets over all φ, we conclude that E_λ is also closed. ∎

The *multiplicity* of an eigenvalue λ is defined as $\dim E_\lambda$ (finite or ∞).

Theorem 2.13 *Assume that Ω is precompact. Then the following statements are true.*

(a) *There exists an orthonormal basis $\{v_k\}_{k=1}^\infty$ in $L^2(\Omega)$ that consists of the Dirichlet eigenfunctions of Δ in Ω.*

(b) *$v_k \in C^\infty(\Omega)$ for all k.*

(c) *The sequence $\{v_k\}_{k=1}^\infty$ is an orthogonal basis also in $W_0^1(\Omega)$.*

(d) *The eigenvalue λ_k of v_k is a non-negative real, and the sequence $\{\lambda_k\}$ is monotone increasing and diverges to $+\infty$ as $k \to \infty$.*

(e) *The sequence $\{\lambda_k\}_{k=1}^\infty$ contains any Dirichlet eigenvalues λ of Δ in Ω exactly m times where m is the multiplicity of λ. Consequently, any eigenvalue has a finite multiplicity.*

Proof. (b) Any eigenfunction of the Laplace operator is C^∞ by Corollary 2.8, in particular, $v_k \in C^\infty(\Omega)$.

(a) 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

$$v = R\left((1+\lambda)\, v\right), \tag{2.29}$$

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

$$v = \frac{1}{\alpha} Rv \in W_0^1(\Omega).$$

Comparing this with (2.29), we conclude that v is an eigenfunction of Δ in Ω with the eigenvalue λ that is determined by $1+\lambda = \frac{1}{\alpha}$, that is, $\lambda = \frac{1}{\alpha} - 1$.

Next we need 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 $k \to \infty$.

It follows that v_k is an eigenfunction of Δ in Ω with the eigenvalue

$$\lambda_k = \frac{1}{\alpha_k} - 1. \tag{2.30}$$

(c) Let us verify that the sequence $\{v_k\}$ is orthogonal in $W_0^1(\Omega)$. Applying (2.28) with $v = v_k$ 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

$$(v_k, v_l)_{W^1} = (v_k, v_l)_{L^2} + (\nabla v_k, \nabla v_l)_{\vec{L}^2} = 0.$$

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 $\varphi \in W_0^1(\Omega)$,

$$\text{if } (v_k, \varphi)_{W^1} = 0 \ \forall k \geq 1 \text{ then } \varphi = 0.$$

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 $\varphi = 0$.

(d) Since the resolvent 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.

By (2.30), $\{\lambda_k\}$ 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 by (2.28) we have, for any $\varphi \in W_0^1(\Omega)$,

$$(\nabla v, \nabla \varphi)_{\vec{L}^2} = \lambda (v, \varphi)_{L^2}. \tag{2.31}$$

Substituting $\varphi = v$, we obtain

$$\boxed{\lambda = \frac{\|\nabla v\|_{\vec{L}^2}^2}{\|v\|_{L^2}^2}} \geq 0. \tag{2.32}$$

Let mention for the future the following consequence of (2.32):

$$\boxed{\|v\|_{W^1}^2 = (\lambda + 1) \|v\|_{L^2}^2}. \tag{2.33}$$

(e) 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 (2.31), we obtain

$$(\nabla v, \nabla w)_{\vec{L}^2} = \lambda (v, w)_{L^2} \tag{2.34}$$

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 (2.34) that $(\nabla v, \nabla w)_{\vec{L}^2} = 0$ and, hence, $(u, w)_{W^1} = 0$.

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 $\text{span}\{v_{i+1}, ..., v_{i+l}\}$. 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 $W_0^1(\Omega)$. This contradiction shows that $l = m$, which concludes the proof. ■

16-Jan-26 Lecture 22

In what follows we denote by $\{\lambda_k(\Omega)\}_{k=1}^\infty$ the sequence of the Dirichlet eigenvalues of Δ in Ω in the (non-strictly) increasing order, that is, $\lambda_k(\Omega) \leq \lambda_{k+1}(\Omega)$.

Example. Let $M = \mathbb{R}$ and, hence $\Delta = \frac{d^2}{dx^2}$. Let us compute the eigenvalues of Δ in the interval $\Omega = (0, a) \subset \mathbb{R}$. The eigenvalue problem for Δ in Ω is

$$\begin{cases} v'' + \lambda v = 0 \text{ in } (0, a) \\ v(0) = v(a) = 0. \end{cases}$$

The ODE $v'' + \lambda v = 0$ has for positive λ the general solution

$$v(x) = C_1 \cos \sqrt{\lambda} x + C_2 \sin \sqrt{\lambda} x.$$

At $x = 0$ we obtain that $C_1 = 0$, and at $x = a$ we obtain that

$$\sin \sqrt{\lambda} a = 0,$$

which gives all solutions

$$\lambda_k = \left(\frac{\pi k}{a}\right)^2, \quad k \in \mathbb{N}.$$

The corresponding eigenfunctions $v_k(x) = \sin\frac{\pi k x}{a}$. It is possible to prove that $v_k \in W_0^1(0,a)$. Besides, the sequence $\left\{\sin\frac{\pi k x}{a}\right\}_{k=1}^{\infty}$ is known to be an orthogonal basis in $L^2(0,a)$, which implies that we have found all the eigenfunctions and eigenvalues of Δ in Ω.

Example. Let $M = \mathbb{R}^2$ and, hence, $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. Let us compute the eigenvalues of Δ in the rectangle $\Omega = (0,a)\times(0,b) \subset \mathbb{R}^2$. Observe first that if v is an eigenfunction of $\frac{d^2}{dx^2}$ in $U = (0,a)$ with an eigenvalue α and w is an eigenfunction in $W = (0,b)$ with an eigenvalue β then the function

$$v(x,y) = u(x)\,w(y)$$

vanishes on $\partial\Omega$ and satisfies

$$\begin{aligned}\Delta v &= \frac{\partial^2}{\partial x^2}u(x)\,w(y) + \frac{\partial^2}{\partial y^2}u(x)\,w(y)\\ &= \alpha v + \beta v = (\alpha+\beta)\,v,\end{aligned}$$

that is, v is an eigenfunction of Δ in Ω with the eigenvalue $\alpha+\beta$.

By the previous example, we have the following eigenvalues U and W

$$u_k(x) = \sin\frac{\pi k x}{a} \quad \text{and} \quad w_l(y) = \sin\frac{\pi l y}{b}$$

and the eigenvalues

$$\alpha_k = \left(\frac{\pi k}{a}\right)^2, \quad \beta_l = \left(\frac{\pi l}{b}\right)^2,$$

for arbitrary $k, l \in \mathbb{N}$. It follows that that Δ in Ω has the following eigenfunctions and eigenvalues:

$$\begin{aligned}v_{k,l}(x,y) &= \sin\frac{\pi k x}{a}\sin\frac{\pi l y}{b}\\ \lambda_{k,l} &= \pi^2\left(\left(\frac{k}{a}\right)^2 + \left(\frac{l}{b}\right)^2\right).\end{aligned}$$

Since $\{v_{k,l}\}$ forms an orthogonal basis in $L^2(\Omega)$, this list gives all eigenvalues of Δ in Ω.

For example, in the case $a = b = \pi$, the eigenvalues are

$$\lambda_{k,l} = k^2 + l^2,$$

that is, all sums of squares of two natural numbers. Setting $k, l = 1, 2, 3, 4, ...$ we obtain

$$\lambda_{1,1} = 2,\ \lambda_{1,2} = \lambda_{2,1} = 5,\ \lambda_{2,2} = 8,\ \lambda_{1,3} = \lambda_{3,1} = 10,\ \lambda_{2,3} = \lambda_{3,2} = 13,\ \lambda_{3,3} = 18,\ \lambda_{1,4} = \lambda_{4,1} = 17, ...$$

The sequence of the eigenvalues in the increasing order is

$$2, 5, 5, 8, 10, 10, 13, 13, 17, 17, 18, ...$$

In particular, the eigenvalues $5, 10, 13, 17$ have multiplicity 2.

Denote by $m(\lambda)$ the multiplicity of an arbitrary number λ in the sequence $\{\lambda_{k,l}\}$. Clearly, $m(\lambda)$ is equal to the number of ways in which λ can be represented as a sum of squares of two positive integers. For example, $m(50) = 3$ because

$$50 = 5^2 + 5^2 = 1^2 + 7^2 = 7^2 + 1^2.$$

An explicit formula for $m(\lambda)$ is obtained in Number Theory, using decomposition of λ into product of primes. In particular, it is known that

$$m(5^q) = q + 1$$

if q is an odd number. Consequently, $m(\lambda)$ can be arbitrarily large. For example, we have for $q = 3$

$$m(125) = 4,$$

and the corresponding representations of 125 in the form $k^2 + l^2$ are

$$125 = 2^2 + 11^2 = 11^2 + 2^2 = 5^2 + 10^2 = 10^2 + 5^2.$$

Example. Let $M = \mathbb{R}^n$ and, hence, $\Delta = \frac{\partial^2}{\partial x_1^2} + ... + \frac{\partial^2}{\partial x_n^2}$. Let us compute the eigenvalues of Δ in the box

$$\Omega = (0, a_1) \times (0, a_2) \times ... \times (0, a_n),$$

where $a_1, ..., a_n$ are positive reals. In the same way as above, we obtain the following eigenvalues and eigenfunctions of Δ in Ω:

$$\begin{aligned} v_{k_1,...,k_n}(x) &= \sin \frac{\pi k_1 x_1}{a_1} \dots \sin \frac{\pi k_n x_n}{a_n} \\ \lambda_{k_1,...,k_n} &= \pi^2 \left(\left(\frac{k_1}{a_1} \right)^2 + ... + \left(\frac{k_n}{a_n} \right)^2 \right), \end{aligned}$$

where $k_1, ..., k_n$ are arbitrary natural numbers.

Remark. For any bounded domain $\Omega \subset \mathbb{R}^n$, the following *Weyl's asymptotic* is known:

$$\lambda_k(\Omega) \sim c_n \left(\frac{k}{\operatorname{vol} \Omega} \right)^{2/n} \quad \text{as } k \to \infty,$$

where $c_n > 0$ depends on n only.

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 (Dirichlet) eigenvalue* of Ω.

Theorem 2.14 *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 $\lambda = 0$.

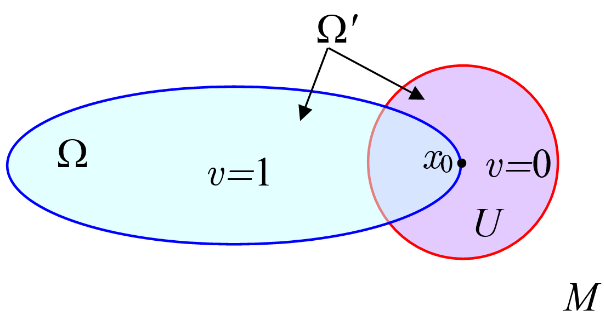
Recall that if Ω is a bounded domain in $\mathbb{R}^n$ then $\lambda_1(\Omega) > 0$ can be proved by using Friedrich's inequality. On a general manifold this tool is not available, so we have to use a different argument.

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 (2.32) 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 $v \not\equiv 0$ in Ω then we can assume without loss of generality, that $v \equiv 1$ in Ω.

The set $K := \overline{\Omega}$ is closed and non-empty, and its complement K^c is non-empty by hypothesis. The sets K and $\overline{K^c}$ 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 K and $\overline{K^c}$.

Let U be a precompact chart containing x_0. By shrinking U, we can assume that U is a ball in the local coordinates, in particular, U is connected. 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 $v \in W_0^1(\Omega)$, we can extend v to Ω' by setting $v = 0$ in $\Omega' \setminus \Omega$ and obtain that $v \in W_0^1(\Omega')$ (which follows from $\mathcal{D}(\Omega) \subset \mathcal{D}(\Omega')$). Since $v = 0$ on $\Omega' \setminus \Omega$, we have also $\nabla v = 0$ in $\Omega' \setminus \Omega$ a.e. (by the known result: if E is a measurable set in $\mathbb{R}^n$, $v \in W^1(\mathbb{R}^n)$ and $v = \text{const}$ a.e.in E then $\nabla v = 0$ a.e. in E). Since also $\nabla v = 0$ in Ω, we conclude that $\nabla v = 0$ in Ω' a.e.. This implies that

$$(\nabla v, \nabla \varphi)_{\vec{L}^2} = 0 \quad \forall \varphi \in \mathcal{D}(\Omega),$$

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 $\Omega' \setminus \overline{\Omega}$. ■

Proposition 2.15 (Variational characterization of λ_1) *For any precompact open set $\Omega \subset M$,*

$$\boxed{\lambda_1(\Omega) = \inf_{f \in \mathcal{D}(\Omega) \setminus \{0\}} \frac{\int_\Omega |\nabla f|^2 \, d\mu}{\int_\Omega f^2 d\mu}}. \tag{2.35}$$

See Exercise 71 for . The identity (2.35) can be used to *define* the value $\lambda_1(\Omega)$ for an *arbitrary* (not necessarily precompact) open set $\Omega \subset M$. In general, $\lambda_1(\Omega)$ does not have the meaning of the minimal eigenvalue of Δ in Ω but it is rather the minimal value of the *spectrum* of Δ as an unbounded self-adjoint operator in $L^2(\Omega)$.

Corollary 2.16 *Let $\Omega \subset M$ be a non-empty relatively compact open set such that $M \setminus \overline{\Omega}$ is non-empty. Then the Dirichlet problem*

$$\begin{cases} \Delta u = f & \text{weakly in } \Omega, \\ u \in W_0^1(\Omega), \end{cases}$$

has a unique solution for any $f \in L^2(\Omega)$.

Proof. By Theorem 2.14 we have $\lambda_1(\Omega) > 0$, and the unique solvability of the Dirichlet problem follows from Exercise 68. ■

Chapter 3

The heat equation

21-Jan-26 **Lecture 23**

As before, $(M, \mathbf{g}, \mu)$ is a weighted manifold and $\Delta = \Delta_{\mathbf{g},\mu}$ is the weighted Laplace operator on M.

3.1 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. Let us fix an open set $\Omega \subset M$.

Definition. For a function $u : I \to L^2(\Omega)$, define its L^2-derivative $u'(t) \in L^2(\Omega)$ at $t \in I$ by

$$u'(t) = \lim_{s \to 0} \frac{u(t+s) - u(t)}{s},$$

assuming the limit exists in the sense of the L^2-norm, that is,

$$\left\| \frac{u(t+s) - u(t)}{s} - u'(t) \right\|_{L^2(\Omega)} \to 0 \text{ as } s \to 0.$$

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 $L^2(\Omega)$, 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 $u : I \to L^2(\Omega)$ is called *caloric* in $I \times \Omega$ if

• function u is L^2-differentiable at any $t \in I$;

• 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;

• for any $t \in I$, we have $u'(t) = \Delta u(t)$.

In this case we also say that the heat equation $u' = \Delta u$ is satisfied weakly in $I \times \Omega$.

Example. Assume that $v \in W^1(\Omega)$ satisfies weakly in Ω the equation

$$\Delta v + \lambda v = 0$$

for some $\lambda \in \mathbb{R}$. Then the function $u(t, x) = e^{-\lambda t} v(x)$ is caloric in $\mathbb{R} \times \Omega$. Indeed, u can be regarded as a mapping

$$u : \mathbb{R} \to L^2(\Omega)$$
$$u(t) = e^{-\lambda t} v$$

Since v does not depend in t, we obtain

$$u'(t) = -\lambda e^{-\lambda t} v.$$

On the other hand, since $\Delta v = -\lambda v$, we have

$$\Delta u = e^{-\lambda t} \Delta v = -\lambda e^{-\lambda t} v,$$

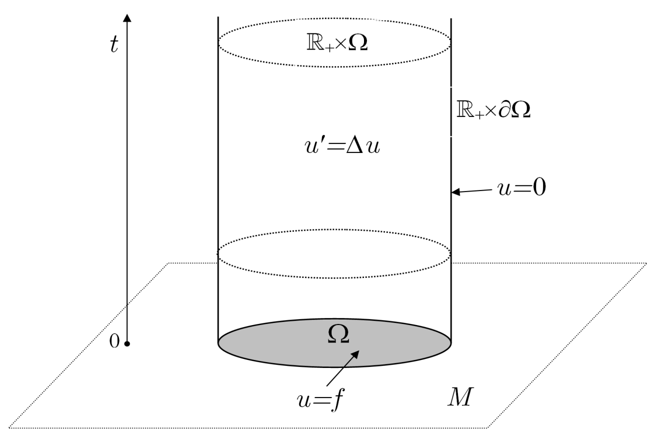
whence $u' = \Delta u$ follows.

3.2 The mixed problem

Let Ω be an open subset of a weighted manifold M. Consider the following initial-boundary problem (shortly, *mixed problem*) in $\mathbb{R}_+ \times \Omega$:

$$\begin{cases} u' = \Delta u & \text{weakly in } \mathbb{R}_+ \times \Omega, \\ u(t, \cdot) \in W_0^1(\Omega) & \text{for any } t > 0, \\ u(t, \cdot) \overset{L^2}{\to} f & \text{as } t \to 0+, \end{cases} \tag{3.1}$$

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 $\mathbb{R}_+ \times \partial\Omega$ 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

$$\frac{d}{dt}(u,v) = (u', v) + (u, v'),$$

which is proved in the same way, as the usual product rule for scalar functions (see Exercise 85). In particular, we obtain that the function $J(t)$ is differentiable on $(0,+\infty)$ and

$$J'(t) = \frac{d}{dt}(u,u) = 2(u', u) = 2(\Delta u, u).$$

By the definition of Δu, we have, for any $\varphi \in W_0^1(\Omega)$,

$$(\Delta u, \varphi) = -(\nabla u, \nabla \varphi).$$

Since $u \in W_0^1(\Omega)$, setting here $\varphi = u$ we obtain

$$(\Delta u, u) = -(\nabla u, \nabla u) \leq 0,$$

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 $u(t) \overset{L^2}{\to} 0$ 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.*

From now on, let Ω be a precompact open subset of M. Let $\{v_k\}_{k=1}^{\infty}$ be an orthonormal basis in $L^2(\Omega)$ that consists of eigenfunctions of Δ in Ω, and $\{\lambda_k\}_{k=1}^{\infty}$ be the sequence of the corresponding eigenvalues, in the increasing order.

Note that any $f \in L^2(\Omega)$ has an expansion in the basis $\{v_k\}$ that is called an *eigenfunction expansion*:

$$f = \sum_{k=1}^{\infty} a_k v_k$$

where a_k are real coefficients and the series converges in $L^2(\Omega)$. Moreover, since $\{v_k\}$ is orthonormal, we have

$$a_k = (f, v_k)_{L^2}.$$

Theorem 3.2 *Let a function $f \in L^2(\Omega)$ have the eigenfunction expansion*

$$f = \sum_{k=1}^{\infty} a_k v_k.$$

For any $t \geq 0$, consider the function

$$u(t) = \sum_{k=1}^{\infty} e^{-\lambda_k t} a_k v_k$$

as an element of $L^2(\Omega)$. Then $u(t)$ solves the mixed problem (3.1).

Hence, the problem (3.1) has a unique solution for any $f \in L^2(\Omega)$.

Remark. Since v_k satisfies $\Delta v_k + \lambda_k v_k = 0$ in Ω, the function $e^{-\lambda_k t} v_k(x)$ is caloric in $\mathbb{R}_+ \times \Omega$ and belongs to $W_0^1(\Omega)$ for any t. It follows that, for any finite N and any real a_k, the function

$$u(t) = \sum_{k=1}^{N} e^{-\lambda_k t} a_k v_k$$

is also caloric and belongs to $W_0^1(\Omega)$. As $t \to 0$ we have

$$u(t) \to f := \sum_{k=1}^{N} a_k v_k.$$

Hence, u solves (3.1) with this initial function f.

The idea of the proof of Theorem 3.2 is to justify the same approach when $N = \infty$ so that any initial function $f \in L^2(\Omega)$ can be handled in this way.

We prove first two lemmas.

Lemma 3.3 *Let $\{a_k\}_{k=1}^{\infty}$ be a sequence of reals. Assume that*

$$\sum_{k=1}^{\infty} a_k^2 < \infty \tag{3.2}$$

so that $f := \sum_{k=1}^{\infty} a_k v_k \in L^2(\Omega)$.

(*a*) *If*

$$\sum_{k=1}^{\infty} \lambda_k a_k^2 < \infty, \tag{3.3}$$

then the series $\sum_k a_k v_k$ converges in $W^1(\Omega)$ and, hence,

$$f \in W_0^1(\Omega).$$

Besides,

$$\|f\|_{W^1}^2 = \sum_{k=1}^{\infty} (\lambda_k + 1) a_k^2. \tag{3.4}$$

(*b*) *If*

$$\sum_{k=1}^{\infty} \lambda_k^2 a_k^2 < \infty \tag{3.5}$$

then

$$\Delta f = -\sum_{k=1}^{\infty} \lambda_k a_k v_k \in L^2(\Omega). \tag{3.6}$$

Remark. Since $\lambda_k \to \infty$, the condition (3.3) implies (3.2), while (3.5) implies (3.3).

Remark. By Exercise 73, the condition (3.3) is also necessary for $f \in W_0^1(\Omega)$, and (3.5) is also necessary for $\Delta f \in L^2(\Omega)$.

Proof. We use the fact that if $\{h_k\}_{k=1}^{\infty}$ is an orthogonal sequence in a Hilbert space H then the series $\sum_{k=1}^{\infty} h_k$ converges in H if and only if

$$\sum_{k=1}^{\infty} \|h_k\|^2 < \infty.$$

Besides, if $h = \sum_{k=1}^{\infty} h_k$ then, by the Parseval identity,

$$\|h\|^2 = \sum_{k=1}^{\infty} \|h_k\|^2.$$

(*a*) The sequence $\{v_k\}$ is orthogonal in $L^2(\Omega)$ as well as in $W_0^1(\Omega)$ and, by (2.33),

$$\|v_k\|_{W^1}^2 = (\lambda_k + 1)\|v_k\|_{L^2}^2 = \lambda_k + 1.$$

It follows that

$$\sum_{k=1}^{\infty} \|a_k v_k\|_{W^1}^2 = \sum_{k=1}^{\infty} a_k^2(\lambda_k + 1) = \sum_{k=1}^{\infty} a_k^2 \lambda_k + \sum_{k=1}^{\infty} a_k^2 < \infty. \tag{3.7}$$

Consequently, the series $\sum_{k=1}^{\infty} a_k v_k$ converges in $W^1(\Omega)$, and its sum f belongs to $W_0^1(\Omega)$. Finally, we obtain by the Parseval identity that

$$\|f\|_{W^1}^2 = \sum_{k=1}^{\infty} \|a_k v_k\|_{W^1}^2 = \sum_{k=1}^{\infty} a_k^2(\lambda_k + 1).$$

which proves (3.4).

(*b*) By (*a*) we have $f \in W_0^1(\Omega)$. By (3.5) we have

$$g := \sum_{k=1}^{\infty} \lambda_k a_k v_k \in L^2(\Omega).$$

We need to prove that $\Delta f = -g$ weakly. For that, consider first the partial sums

$$f_N = \sum_{k=1}^{N} a_k v_k \quad \text{and} \quad g_N = \sum_{k=1}^{N} \lambda_k a_k v_k.$$

Using that $\Delta v_k = -\lambda_k v_k$, we obtain

$$\Delta f_N = \sum_{k=1}^{N} a_k \Delta v_k = -\sum_{k=1}^{N} \lambda_k a_k v_k = -g_N.$$

Hence, for any $\varphi \in W_0^1(\Omega)$, we have

$$(\nabla f_N, \nabla \varphi)_{\vec{L}^2} = (g_N, \varphi)_{L^2}. \tag{3.8}$$

Letting $N \to \infty$ and using that $f_N \overset{W^1}{\to} f$ and $g_N \overset{L^2}{\to} g$, we obtain

$$(\nabla f, \nabla \varphi)_{\vec{L}^2} = (g, \varphi)_{L^2},$$

that is, $\Delta f = -g$, which was to be proved. ■

23-Jan-26 Lecture 24

Example. As an example of application of Lemma 3.3 consider the Dirichlet problem in Ω:

$$\begin{cases} \Delta u = -f \text{ weakly in } \Omega \\ u \in W_0^1(\Omega) \end{cases} \tag{3.9}$$

where f is a given function. Let us prove that if $\lambda_1 > 0$ then this problem has a solution for any $f \in L^2(\Omega)$. For that, let us reformulate Lemma 3.3 in terms of function u: if $u \in L^2(\Omega)$, the eigenfunction expansion of u is $u = \sum_{k=1}^\infty b_k v_k$, and

$$\sum_{k=1}^\infty \lambda_k^2 b_k^2 < \infty, \tag{3.10}$$

then $u \in W_0^1(\Omega)$ and

$$\Delta u = -\sum_{k=1}^\infty \lambda_k b_k v_k \in L^2(\Omega).$$

Hence, if

$$f = \sum_{k=1}^\infty a_k v_k$$

then the equation $\Delta u = -f$ can be reformulated in terms of the components of u and f as follows:

$$\lambda_k b_k = a_k.$$

Hence, if $\lambda_1 > 0$ then all $\lambda_k > 0$, and we can solve this equation is obtain

$$b_k = \frac{a_k}{\lambda_k}$$

and, hence,

$$u = \sum_{k=1}^\infty \frac{a_k}{\lambda_k} v_k. \tag{3.11}$$

The condition (3.10) is satisfies since

$$\sum_{k=1}^\infty \lambda_k^2 \left(\frac{a_k}{\lambda_k}\right)^2 = \sum_{k=1}^\infty a_k^2 < \infty$$

so that the function (3.11) is indeed a solution of (3.9).

Lemma 3.4 (A version of the 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 convergence 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 $\sum_k h_k$ 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

$$F(t) := \sum_{k=1}^{\infty} \gamma_k(t)\, h_k$$

converges for any $t \in I$. We need to prove that $F(t) \to F(0)$ as $t \to 0$. We have

$$F(t) - F(0) = \sum_{k=1}^{\infty} (\gamma_k(t) - \gamma_k(0))\, h_k,$$

whence by the Parseval identity

$$\|F(t) - F(0)\|^2 = \sum_{k=1}^{\infty} (\gamma_k(t) - \gamma_k(0))^2 \|h_k\|^2.$$

To prove that this quantity goes to 0 as $t \to 0$, let us fix some $\varepsilon > 0$ and choose N so big that

$$\sum_{k=N}^{\infty} \|h_k\|^2 < \varepsilon.$$

We have

$$\begin{aligned} \|F(t) - F(0)\|^2 &= \sum_{k=1}^{N} (\gamma_k(t) - \gamma_k(0))^2 \|h_k\|^2 \\ &\quad + \sum_{k=N}^{\infty} (\gamma_k(t) - \gamma_k(0))^2 \|h_k\|^2. \end{aligned}$$

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

$$\limsup_{t\to 0} \|F(t) - F(0)\|^2 \leq 4C^2\varepsilon.$$

Since ε is arbitrary, we obtain that $\|F(t) - F(0)\| \to 0$ as $t \to 0$, which was to be proved. ■

Proof of Theorem 3.2. Let

$$f = \sum_{k=1}^{\infty} a_k v_k \in L^2(\Omega).$$

We need to prove that the function

$$u(t) = \sum_{k=1}^{\infty} e^{-\lambda_k t} a_k v_k$$

solves the mixed problem (3.1), that is, it satisfies the following conditions:

1. $u(t) \in W_0^1(\Omega)$ for any $t > 0$ (the boundary condition);
2. $\Delta u(t) \in L^2(\Omega)$ for any $t > 0$ (computation of Δu);
3. $u(t) \overset{L^2}{\to} f$ as $t \to 0$ (the initial condition);
4. $u'(t)$ exists in $L^2(\Omega)$ and $u'(t) = \Delta u(t)$ for any $t > 0$ (computation of the time derivative).

Boundary condition. Fix $t > 0$. By Lemma 3.3(a), in order to prove that

$$u(t) = \sum_{k=1}^{\infty} \underline{e^{-\lambda_k t} a_k} v_k \in W_0^1(\Omega),$$

it suffices to verify that

$$\sum_{k=1}^{\infty} \lambda_k \underline{\left(e^{-\lambda_k t} a_k\right)^2} = \sum_{k=1}^{\infty} \lambda_k e^{-2\lambda_k t} a_k^2 < \infty.$$

Indeed, the latter is true because

$$\sum_{k=1}^{\infty} a_k^2 < \infty$$

and

$$\sup_k \lambda_k e^{-2\lambda_k t} \le \sup_{\lambda \ge 0} \lambda e^{-2\lambda t} = \frac{1}{t} \sup_{\lambda \ge 0} (\lambda t) e^{-2\lambda t} = \frac{1}{t} \sup_{\xi \ge 0} \xi e^{-2\xi} < \infty. \tag{3.12}$$

Computation of Δu. Let us show that $\Delta u(t) \in L^2(\Omega)$ for any $t > 0$. By Lemma 3.3(b), it suffices to verify that

$$\sum_{k=1}^{\infty} \lambda_k^2 \underline{e^{-2\lambda_k t} a_k^2} < \infty,$$

and the latter is true because similarly to (3.12)

$$\sup_{\lambda \ge 0} \lambda^2 e^{-2\lambda t} < \infty.$$

Besides, we obtain by (3.6) that

$$\Delta u(t) = -\sum_{k=1}^{\infty} \lambda_k \underline{e^{-\lambda_k t} a_k} v_k.$$

Initial condition. 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 uniformly bounded by 1 for all k and $t \ge 0$, we conclude by Lemma 3.4

$$u(t) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k t} v_k \xrightarrow{L^2} \sum_{k=1}^{\infty} a_k v_k = f.$$

Time derivative. Let us compute $u'(t)$ at any $t > 0$ and verify that $u'(t) = \Delta u$. Observe that

$$\begin{aligned} \frac{u(t+s) - u(t)}{s} &= \sum_{k=1}^{\infty} \frac{e^{-\lambda_k(t+s)} - e^{-\lambda_k t}}{s} a_k v_k \\ &= \sum_{k=1}^{\infty} \frac{e^{-s\lambda_k} - 1}{s} e^{-\lambda_k t} a_k v_k. \end{aligned} \tag{3.13}$$

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 3.4, 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

$$\sup_{\lambda \ge 0} \sup_{s \in [-\varepsilon, \varepsilon]} \left| \frac{e^{-s\lambda} - 1}{s} e^{-\lambda t} \right| < \infty. \tag{3.14}$$

In fact, we will take $\varepsilon = t/2$. Let us apply the inequality

$$\left|e^{\theta}-1\right| \leq |\theta|\, e^{|\theta|},$$

that is valid for any $\theta \in \mathbb{R}$ and that 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 $s \in [-t/2, t/2]$,

$$\left|\frac{e^{-\lambda s}-1}{s} e^{-t\lambda}\right| \leq \lambda e^{-\lambda t} e^{\lambda|s|} = \lambda e^{-\lambda(t-|s|)} \leq \lambda e^{-\lambda t/2}.$$

Therefore, we have

$$\sup_{\lambda \geq 0} \sup_{s \in [-t/2, t/2]} \left|\frac{e^{-\lambda s}-1}{s} e^{-t\lambda}\right| \leq \sup_{\lambda \geq 0} \lambda e^{-\lambda t/2} < \infty, \tag{3.15}$$

which proves (3.14). Returning to (3.13), we obtain

$$u'(t) = \lim_{s \to 0} \frac{u\,(t+s) - u\,(t)}{s} = -\sum_{k=1}^{\infty} \lambda_k e^{-\lambda_k t} a_k v_k = \Delta u(t),$$

which finishes the proof. ■

3.3 The heat semigroup

As before, let Ω be a precompact open subset of M, $\{v_k\}$ be the orthonormal basis in $L^2(\Omega)$ of eigenfunctions of Δ in Ω and $\{\lambda_k\}$ be the sequence of the corresponding eigenvalues.

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

$$P_t^{\Omega} f = \sum_{k=1}^{\infty} e^{-\lambda_k t} a_k v_k. \tag{3.16}$$

In particular, $P_0^{\Omega} f = f$ and, hence, $P_0^{\Omega} = \mathrm{id}$.

By Theorem 3.2, the function $u(t) = P_t^\Omega f$ is the unique solution of the mixed problem

$$\begin{cases} u' = \Delta u & \text{weakly in } \mathbb{R}_+ \times \Omega, \\ u(t) \in W_0^1(\Omega) & \text{for any } t > 0, \\ u(t) \overset{L^2}{\to} f & \text{as } t \to 0+ . \end{cases}$$

It is clear from the definition that the operator P_t^Ω is linear. We prove now the following simple properties of the operator P_t^Ω.

Theorem 3.5 *The operators P_t^Ω have the following properties:*

(a) $\|P_t^\Omega\| \le 1$ *(*contraction property*);*

(b) $P_t^\Omega \to \mathrm{id}$ *as $t \to 0+$ in the strong operator topology (*strong continuity*);*

(c) $P_t^\Omega P_s^\Omega = P_{t+s}^\Omega$ *(*the semigroup identity*);*

(d) P_t^Ω *is self-adjoint and non-negative definite.*

This family $\{P_t^\Omega\}_{t\ge 0}$ is called the *heat semigroup* in Ω. In the terminology of the theory of semigroups, $\{P_t^\Omega\}_{t\ge 0}$ is an *one-parameter strongly continuous contraction semigroup.*

Proof. (a) For any

$$f = \sum_{k=1}^{\infty} a_k v_k \in L^2(\Omega), \tag{3.17}$$

we have

$$\|f\|_{L^2}^2 = \sum_{k=1}^{\infty} a_k^2$$

and

$$\left\|P_t^\Omega f\right\|_{L^2}^2 = \sum_{k=1}^{\infty} e^{-2\lambda_k t} a_k^2 \le \sum_{k=1}^{\infty} a_k^2 = \|f\|_{L^2}^2 , \tag{3.18}$$

whence $\|P_t^\Omega\| \le 1$.

Remark. Since $\lambda_k \ge \lambda_1$ then (3.18) can be improved at follows:

$$\left\|P_t^\Omega f\right\|_{L^2}^2 \le e^{-2\lambda_1 t} \sum_{k=1}^{\infty} a_k^2 = e^{-2\lambda_t t} \|f\|_2^2$$

and, hence,

$$\|P_t^\Omega\| \le e^{-\lambda_1 t}.$$

If $\lambda_1 > 0$ then this inequality implies that $\|P_t^\Omega\| < 1$ for all $t > 0$ and, moreover, $\|P_t^\Omega\| \to 0$ as $t \to \infty$.

(b) We already know that $P_t^\Omega f \overset{L^2}{\to} f$ for any $f \in L^2(\Omega)$, which exactly means that $P_t^\Omega \to \mathrm{id}$ in the strong operator topology (but not in the operator norm).

28-Jan-26 **Lecture 25**

(c) Applying the identity (3.16), that is,

$$P_t^\Omega f = \sum_{k=1}^\infty e^{-\lambda_k t} a_k v_k$$

to $f = v_k$ we obtain

$$P_t^\Omega v_k = e^{-t\lambda_k} v_k.$$

Hence, for any $f = \sum_{k=1}^\infty a_k v_k$, we obtain

$$\begin{aligned} P_t^\Omega P_s^\Omega f &= P_t^\Omega \left(\sum_{k=1}^\infty e^{-\lambda_k s} a_k v_k \right) \\ &= \sum_{k=1}^\infty e^{-\lambda_k s} a_k P_t^\Omega v_k \\ &= \sum_{k=1}^\infty e^{-\lambda_k s} a_k e^{-\lambda_k t} v_k \\ &= \sum_{k=1}^\infty e^{-\lambda_k (t+s)} a_k v_k = P_{t+s}^\Omega f, \end{aligned}$$

which proves the claim. We have used here that P_t^Ω commutes with $\sum_{k=1}^\infty$ because this sum converges in L^2-norm and P_t^Ω is bounded and, hence, is continuous operator.

(d) For functions

$$f = \sum_{k=1}^\infty a_k v_k \in L^2(\Omega) \quad \text{and} \quad g = \sum_{k=1}^\infty b_k v_k \in L^2(\Omega)$$

we have

$$\left(P_t^\Omega f, g\right)_{L^2} = \sum_{k=1}^\infty \left(e^{-\lambda_k t} a_k\right) b_k = \sum_{k=1}^\infty a_k \left(e^{-\lambda_k t} b_k\right) = \left(f, P_t^\Omega g\right)_{L^2},$$

which means that P_t^Ω is symmetric and, hence, self-adjoint.

Finally, P_t^Ω is non-negative definite because by the previous identity we obtain for $f = g$

$$\left(P_t^\Omega f, f\right)_{L^2} = \sum_{k=1}^\infty a_k^2 e^{-\lambda_k t} \geq 0.$$

■

3.4 Smoothness of the heat semigroup

By definition, the function $P_t^\Omega f$ belongs to $W_0^1(\Omega)$ for any $t > 0$. Our next purpose is to prove that, in fact, $P_t^\Omega f \in C^\infty(\Omega)$ any $t > 0$.

We start with a lemma that is an extension of Lemma 3.3.

Lemma 3.6 *Let*

$$f=\sum_{k=1}^{\infty}a_kv_k\in L^2\left(\Omega\right).$$

If, for some non-negative integer j,

$$\sum_{k=1}^{\infty}\lambda_k^{2j+1}a_k^2<\infty \tag{3.19}$$

then

$$\Delta^j f=(-1)^j\sum_{k=1}^{\infty}\lambda_k^j a_k v_k\in W_0^1\left(\Omega\right),$$

where the series converges in $W^1\left(\Omega\right)$.

Proof. Induction in j. For $j=0$ the statement becomes as follows: if

$$\sum_{k=1}^{\infty}\lambda_k a_k^2<\infty$$

then $f\in W_0^1\left(\Omega\right)$, which is equivalent to Lemma 3.3(a).

In what follows we use also Lemma 3.3(b) that states the following: if

$$g=\sum_{k=1}^{\infty}b_kv_k\in L^2\left(\Omega\right)$$

and

$$\sum_{k=1}^{\infty}\lambda_k^2b_k^2<\infty$$

then

$$\Delta g=-\sum_{k=1}^{\infty}\lambda_kb_kv_k\in L^2\left(\Omega\right).$$

Inductive step from j to $j+1$. For that, we assume that

$$\sum_{k=1}^{\infty}\lambda_k^{2j+3}a_k^2<\infty$$

and prove that

$$\Delta^{j+1}f=(-1)^{j+1}\sum_{k=1}^{\infty}\lambda_k^{j+1}a_kv_k\in W_0^1(\Omega).$$

By the inductive hypothesis, we have

$$g:=\Delta^j f=(-1)^j\sum_{k=1}^{\infty}\lambda_k^j a_kv_k\in W_0^1\left(\Omega\right).$$

Set

$$b_k=(-1)^j\,\lambda_k^j a_k$$

so that $g=\sum_{k=1}^{\infty} b_k v_k$. Since

$$\Delta^{j+1} f=\Delta\left(\Delta^j f\right)=\Delta g,$$

we need to verify that $\Delta g \in W_0^1(\Omega)$. Since

$$\sum_{k=1}^{\infty} \lambda_k^2 b_k^2=\sum_{k=1}^{\infty} \lambda_k^{2j+2} a_k^2<\infty,$$

we obtain by Lemma 3.3(b) that

$$\Delta g=-\sum_{k=1}^{\infty} \lambda_k b_k v_k \in L^2(\Omega) .$$

Let us verify that, in fact, $\Delta g \in W_0^1(\Omega)$. By Lemma 3.3($a$) it suffices to show that

$$\sum_{k=1}^{\infty} \lambda_k\left(\lambda_k b_k\right)^2<\infty,$$

and this is true because

$$\sum_{k=1}^{\infty} \lambda_k\left(\lambda_k b_k\right)^2=\sum_{k=1}^{\infty} \lambda_k\left(\lambda_k(-1)^j \lambda_k^j a_k\right)^2=\sum_{k=1}^{\infty} \lambda_k^{2j+3} a_k^2<\infty,$$

Hence, $\Delta g \in W_0^1(\Omega)$. Since $\Delta^{j+1} f=\Delta g$,we conclude that $\Delta^{j+1} f \in W_0^1(\Omega)$ and

$$\Delta^{j+1} f=-\sum_{k=1}^{\infty} \lambda_k b_k v_k=(-1)^{j+1} \sum_{k=1}^{\infty} \lambda_k^{j+1} a_k v_k,$$

which finishes the prove. ■

Theorem 3.7 *Let Ω be a precompact open subset of M. For any $f \in L^2(\Omega)$ and $t>0$, we have*

$$P_t^{\Omega} f \in C^{\infty}(\Omega) .$$

Moreover, for any compact set $K \subset \Omega$ and any $t>0$,

$$\left\|P_t^{\Omega} f\right\|_{C(K)} \leq C\left(1+t^{-1}\right)^{\frac{n}{4}+1}\|f\|_{L^2}, \tag{3.20}$$

where $C=C(\Omega, K, \mathbf{g}, D, n)$.

Proof. Let $f=\sum_{k=1}^{\infty} a_k v_k$, so that

$$u(t):=P_t^{\Omega} f=\sum_{k=1}^{\infty} e^{-\lambda_k t} a_k v_k.$$

Let us prove that, for any $t>0$ and any non-negative j,

$$\Delta^j u(t) \in W_0^1(\Omega) .$$

By Lemma 3.6, we need to verify that

$$\sum_{k=1}^{\infty} \lambda_k^{2j+1} \left(e^{-\lambda_k t} a_k\right)^2 < \infty. \tag{3.21}$$

Indeed, we have

$$\sum_{k=1}^{\infty} \lambda_k^{2j+1} \left(e^{-\lambda_k t} a_k\right)^2 \leq \sup_k \lambda_k^{2j+1} e^{-2\lambda_k t} \sum_{k=1}^{\infty} a_k^2 < \infty$$

because $\sum a_k^2 < \infty$ and, for any $q \geq 0$,

$$\sup_{\lambda \geq 0} \lambda^q e^{-2\lambda t} = \sup_{\lambda \geq 0} t^{-q} (\lambda t)^q e^{-2\lambda t} = t^{-q} \sup_{\xi \geq 0} \xi^q e^{-2\xi} = \frac{C_q}{t^q} < \infty. \tag{3.22}$$

Hence, by Lemma 3.6, we obtain

$$\Delta^j u(t) = (-1)^j \sum_{k=1}^{\infty} \lambda_k^j e^{-\lambda_k t} a_k v_k \in W_0^1(\Omega). \tag{3.23}$$

By (3.23) and Theorem 2.7, we conclude that $u(t) \in C^\infty(\Omega)$ for any $t > 0$.

Let us prove the estimate (3.20). By Lemma 3.3 we have

$$\left\|\Delta^j u\right\|_{W^1}^2 = \sum_{k=1}^{\infty} (\lambda_k + 1) \left(\lambda_k^j e^{-\lambda_k t} a_k\right)^2 \leq \sup_k (\lambda_k + 1) \left(\lambda_k^j e^{-\lambda_k t}\right)^2 \sum_{k=1}^{\infty} a_k^2.$$

Since $\sum a_k^2 = \|f\|_{L^2}^2$ and by (3.22)

$$\sup_{\lambda \geq 0} (\lambda + 1) \left(\lambda^j e^{-\lambda t}\right)^2 \leq \sup_{\lambda \geq 0} \lambda^{2j+1} e^{-2\lambda t} + \sup_{\lambda \geq 0} \lambda^{2j} e^{-2\lambda t} \leq \frac{C_{2j+1}}{t^{2j+1}} + \frac{C_{2j}}{t^{2j}} \leq \frac{C_j' \left(1 + t^{-1}\right)}{t^{2j}},$$

we obtain

$$\left\|\Delta^j u\right\|_{W^1}^2 \leq \frac{C_j' \left(1 + t^{-1}\right)}{t^{2j}} \|f\|_{L^2}^2. \tag{3.24}$$

By Theorem 2.7, we have, for any precompact open chart $U \Subset \Omega$,

$$\|u\|_{C(U)} \leq C \sum_{j=0}^{k} \left\|\Delta^j u\right\|_{W^1(\Omega)}.$$

provided $2k + 1 > \frac{n}{2}$. Since K can be covered by a finite number of such charts U, we obtain that also

$$\|u\|_{C(K)} \leq C \sum_{j=0}^{k} \left\|\Delta^j u\right\|_{W^1(\Omega)}.$$

By (3.24) we have

$$\begin{aligned}\sum_{j=0}^{k}\left\|\Delta^j u\right\|_{W^1(\Omega)} &\le C\sum_{j=0}^{k}\frac{\left(1+t^{-1}\right)^{1/2}}{t^j}\|f\|_{L^2}\\ &= C\left(1+t^{-1}\right)^{1/2}\|f\|_{L^2}\sum_{j=0}^{k}\frac{1}{t^j}\\ &\le C\left(1+t^{-1}\right)^{1/2}\|f\|_{L^2}\left(1+\frac{1}{t}\right)^k\\ &= C\left(1+t^{-1}\right)^{k+1/2}\|f\|_{L^2},\end{aligned}$$

whence it follows

$$\|u\|_{C(K)} \le C\left(1+t^{-1}\right)^{k+1/2}\|f\|_{L^2}.$$

Finally, let k be the minimal integer such that $2k+1>\frac{n}{2}$. Then $2k-1\le\frac{n}{2}$, which implies

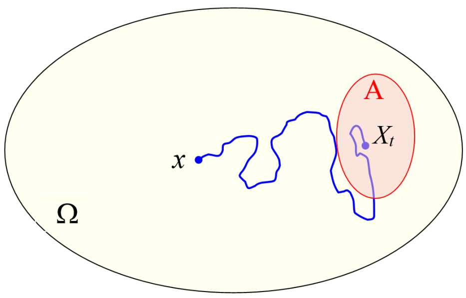
$$\begin{aligned}k-\frac{1}{2} &\le \frac{n}{4},\\ k+\frac{1}{2} &\le \frac{n}{4}+1,\end{aligned}$$

whence (3.20) follows. ■

Remark. It is possible to construct a diffusion process X_t in Ω generated by Δ and with vanishing condition on $\partial\Omega$. Analytically this process is described by the transition function

$$\mathbb{P}_x\left(X_t\in A\right)$$

that is the probability that the process X_t started at $x\in\Omega$ will be in a set $A\subset\Omega$ at time t.



One can represent the transition function in the form

$$\mathbb{P}_x\left(X_t\in A\right)=\mathbb{E}_x\mathbf{1}_A\left(X_t\right).$$

Consider a more general function

$$\mathbb{E}_x f\left(X_t\right)$$

where f is a bounded Borel function in Ω. It turns out that, for all $x\in\Omega$ and $t>0$,

$$\mathbb{E}_x f\left(X_t\right)=P_t^{\Omega}f\left(x\right).$$

In particular, the transition function can be expressed through the heat semigroup as follows: $\mathbb{P}_x\left(X_t\in A\right)=P_t^{\Omega}\mathbf{1}_A\left(x\right).$

3.5 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.

Let us first mention the two versions of the maximum principle for classical solutions of the boundary value problems.

1. Let Ω be a precompact open subset of M and assume that $u \in C^2\left(\overline{\Omega}\right)$ is *subharmonic* in Ω, that is, $\Delta u \geq 0$ in Ω. Then

$$\sup_{\Omega} u = \sup_{\partial\Omega} u.$$

In particular, if $u \leq 0$ on $\partial\Omega$ then $u \leq 0$ in Ω.

2. Fix some $T > 0$ and consider the cylinder $Q = (0, T) \times \Omega$. Let $u \in C^2\left(\overline{Q}\right)$ be *subcaloric* in Q in , that is,

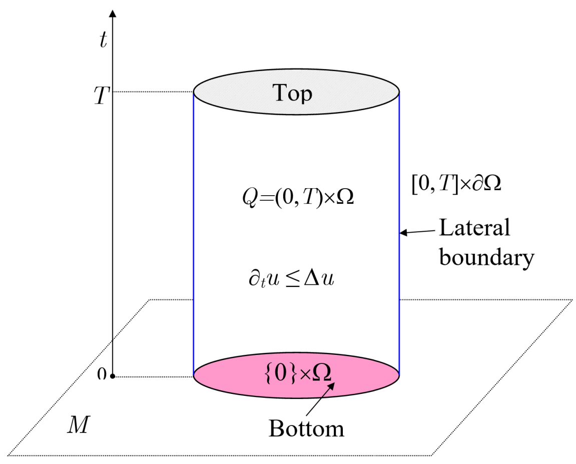
$$\partial_t u \leq \Delta u \text{ in } Q.$$

Then

$$\sup_{Q} u = \sup_{\partial_p Q} u,$$

where $\partial_p Q$ is the *parabolic boundary* of Q that is defined as the union of its bottom and the lateral boundary, that is

$$\partial_p Q := (\{0\} \times \Omega) \cup ([0, T] \times \partial\Omega).$$



In particular, if $u \leq 0$ on $\partial_p Q$ then $u \leq 0$ in Q.

We will prove here the version of such a maximum principle for weak solutions.

Definition. A function $u : I \to L^2(\Omega)$ in an interval I is called *subcaloric* 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 $u'(t) \le \Delta u(t)$.

In the same way, u is called *supercaloric* if $u'(t) \ge \Delta u(t)$.

Definition. For functions $u, v \in W^1(\Omega)$ we write

$$u \le v \bmod W_0^1(\Omega) \tag{3.25}$$

if $u \le v + w$ in Ω for some function $w \in W_0^1(\Omega)$.

Clearly, the relation (3.25) is transitive. It can be regarded as a weak version of "$u \le v$ on $\partial\Omega$".

Theorem 3.8 (Weak parabolic maximum principle) *Let a function* $u : (0,T) \to L^2(\Omega)$ *satisfy the following conditions:*

(i) *u is subcaloric in* $(0,T) \times \Omega$;

(ii) (the boundary condition) $u(t) \le 0 \bmod W_0^1(\Omega)$ *for any* $t \in (0,T)$*;*

(iii) (the initial condition) $u(t)_+ \overset{L^2(\Omega)}{\longrightarrow} 0$ *as* $t \to 0$.

Then $u(t) \le 0$ *for all* $t \in (0,T)$.

That is, if u satisfy the following conditions in the weak sense

$$\begin{cases} \partial_t u \le \Delta u & \text{in } (0,T) \times \Omega \\ u|_{\partial\Omega} \le 0 & \text{for all } t \in (0,T) \\ u|_{t=0} \le 0 & \end{cases}$$

then $u \le 0$ in $(0,T) \times \Omega$.

If u is supercaloric then $-u$ is subcaloric. Hence, Theorem 3.8 can be reformulated as the *minimum principle* for supercaloric functions as follows: if

(i) u is supercaloric in $(0,T) \times \Omega$;

(ii) $u(t) \ge 0 \bmod W_0^1(\Omega)$ for any $t \in (0,T)$;

(iii) $u(t)_- \overset{L^2}{\to} 0$ as $t \to 0$;

then $u(t) \ge 0$ 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, $u(t) = 0 \bmod W_0^1(\Omega)$. We also know that $u(t) \overset{L^2}{\to} f$ as $t \to 0$. In particular, if $f \ge 0$ then $u(t)_- \overset{L^2}{\to} 0$ as $t \to 0$ and, by the minimum principle, we conclude that $u(t) \ge 0$ for all $t > 0$. Similarly, if $f \le 0$ then $u(t) \le 0$.

Hence, if $f = 0$ then $u = 0$, which recovers the uniqueness result of Theorem 3.1.

For the proof of Theorem 3.8, we need the following lemma.

Lemma 3.9 *If $u \in W^1(\Omega)$ then the relation*

$$u \leq 0 \mod W_0^1(\Omega) \tag{3.26}$$

holds if and only if $u_+ \in W_0^1(\Omega)$ (that is, $u_+ = 0 \mod W_0^1(\Omega)$).

Proof. In the proof we use the following facts without proofs:

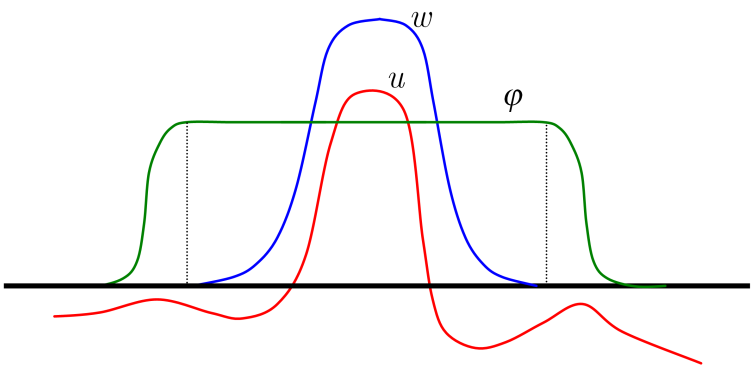
- If $u \in W^1(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$ then $u\varphi \in W_0^1(\Omega)$.
- If $v \in W_0^1(\Omega)$ then also $v_+ \in W_0^1(\Omega)$ and $\nabla v_+ = \mathbf{1}_{\{v>0\}} \nabla v$.
- If $v_k \in W_0^1(\Omega)$ and $v_k \overset{W^1}{\to} v \in W_0^1(\Omega)$ then also $(v_k)_+ \overset{W^1}{\to} v_+$.

These facts are known to be true in the case $M = \mathbb{R}^n$, but the general case can be derived from the case of $\mathbb{R}^n$ using partition of unity.

If $u_+ \in W_0^1(\Omega)$ then (3.26) 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. Then we have the following identity:

$$u_+ = ((1-\varphi)w + \varphi u)_+ . \tag{3.27}$$



Functions u, w, φ

Indeed, if $\varphi = 1$ then (3.27) is obviously satisfied. If $\varphi < 1$ then $w = 0$ and, hence, $u \leq 0$, so that the both sides of (3.27) 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 (3.27) we conclude that $u_+ \in W_0^1(\Omega)$.

30-Jan-26 **Lecture 26**

In the general case when $w \in W_0^1(\Omega)$, there exists a sequence $\{w_k\}$ of functions from $\mathcal{D}(\Omega)$ such that $w_k \xrightarrow{W^1} w$. Then we have

$$u_k := u + w_k - w \le w_k,$$

which implies by the first part of the proof that $(u_k)_+ \in W_0^1(\Omega)$. Since $u_k \stackrel{W^1}{\to} u$, it follows that $(u_k)_+ \stackrel{W^1}{\to} u_+$, whence we conclude that $u_+ \in W_0^1(\Omega)$. ∎

Proof of Theorem 3.8. Let a function $u : (0, T) \to L^2(\Omega)$ satisfy the following conditions:

(i) $u(t)$ is subcaloric function for any $t \in (0, T)$;

(ii) $u(t) \le 0 \bmod W_0^1(\Omega)$ for any $t \in (0, T)$;

(iii) $u(t)_+ \stackrel{L^2}{\to} 0$ as $t \to 0$.

Then we need to prove that $u(t) \le 0$ in Ω for all $t \in (0, T)$, that is, $u(t)_+ \equiv 0$. Recall that if $u(t)$ is caloric, $u(t) = 0 \bmod W_0^1(\Omega)$ and $u(t) \stackrel{L^2}{\to} 0$ as $t \to 0$ then $u(t) \equiv 0$ by Theorem 3.1. The present proof is a modification of the proof of Theorem 3.1.

By the condition (1) and the definition of subcaloric functions, Δu exists and belongs to $L^2(\Omega)$ for any $t \in (0, T)$, that is, for any test function $v \in \mathcal{D}(\Omega)$

$$(\Delta u, v)_{L^2} = -(\nabla u, \nabla v)_{\vec{L}^2}.$$

Clearly, this identity extends to all $v \in W_0^1(\Omega)$. The inequality $u'(t) \le \Delta u$ in the definition of subcaloric function implies that, for any *non-negative* function $v \in W_0^1(\Omega)$,

$$(u', v)_{L^2} \le -(\nabla u, \nabla v)_{\vec{L}^2}. \tag{3.28}$$

Let us explain first **the idea of the proof** that is similar to that of Theorem 3.1. By Lemma 3.9 we have $u_+(t) \in W_0^1(\Omega)$ for any $t \in (0, T)$. Substituting $v = u_+$ into (3.28), we obtain

$$(u', u_+)_{L^2} \le -(\nabla u, \nabla u_+)_{\vec{L}^2} = -\|\nabla u_+\|^2 \le 0,$$

where we have used that $\nabla u_+ = \mathbf{1}_{\{u>0\}} \nabla u$. In the other hand, we have $u_+ = \psi(u)$ where $\psi(s) = s_+$. Using the product and chain rules (Exercises 85, 86) that

$$\begin{aligned}
\frac{d}{dt}\|u_+\|^2 = \frac{d}{dt}(u, u_+)_{L^2} = \frac{d}{dt}(u, \psi(u))&\\
= (u', \psi(u)) + (u, \psi'(u)\, u')&\\
= (u', u_+) + \left(u, \mathbf{1}_{\{u>0\}} u'\right)&\\
= (u', u_+) + \left(u', \mathbf{1}_{\{u>0\}} u\right)&\\
= (u', u_+) + (u', u_+) \le 0&
\end{aligned}$$

that is,

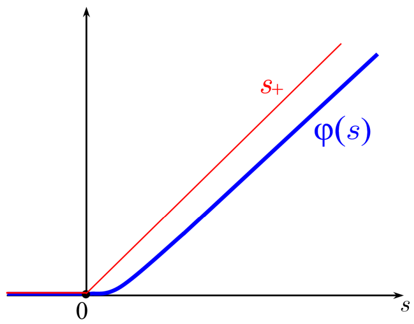
$$\frac{d}{dt}\|u_+\|^2 \le 0.$$

Since $\|u(t)_+\| \to 0$ as $t \to 0$, it follows that $\|u_+(t)\| \equiv 0$ whence $u(t) \le 0$.

However, there is a problem in this argument: the function $\psi(s)$ is not differentiable at $s = 0$ and, hence, the above application of the chain rule to $\frac{d}{dt}\psi(u(t))$ is invalid. To correct this, we need to replace the function $\psi(s) = s_+$ by a smooth approximation.

Let a function $\varphi \in C^\infty(\mathbb{R})$ be such that

$$\begin{cases} \varphi(s) = 0, & s \le 0, \\ \varphi(s) > 0, & s > 0, \\ 0 \le \varphi'(s) \le 1, & s \in \mathbb{R}. \end{cases} \tag{3.29}$$



In the next argument we use the following version of the chain rule: if $v \in W_0^1(\Omega)$ and φ is a smooth Lipschitz function on $\mathbb{R}$ that vanishes at 0, then $\varphi(v) \in W_0^1(\Omega)$ and

$$\nabla\varphi(v) = \varphi'(v)\,\nabla v.$$

Since $u(t)_+ \in W_0^1(\Omega)$ and the function φ from (3.29) is Lipschitz, we obtain that, for any $t \in (0, T)$,

$$\varphi(u(t)) = \varphi\left(u(t)_+\right) \in W_0^1(\Omega)$$

and

$$\nabla\varphi(u) = \nabla\varphi(u_+) = \varphi'(u_+)\,\nabla u_+ = \varphi'(u)\,\nabla u,$$

where we drop the argument t for simplicity.

Setting in (3.28)

$$v = \varphi(u(t))$$

we obtain

$$\begin{aligned} (u', \underline{\varphi(u)})_{L^2} &\le -\left(\nabla u, \nabla\varphi(u)\right)_{L^2} \\ &= -\left(\nabla u, \varphi'(u)\,\nabla u\right)_{\vec{L}^2} \\ &= -\int_\Omega \varphi'(u)\,|\nabla u|^2\,d\mu \le 0. \end{aligned} \tag{3.30}$$

Let $\psi \in C^\infty(\mathbb{R})$ be another function satisfying (3.29). Using the product rule and the chain rule for L^2 derivatives (Exercises 85, 86), we obtain

$$\begin{aligned}\frac{d}{dt}(u,\psi(u))_{L^2} &= (u',\psi(u))_{L^2} + (u,\psi'(u)u')_{L^2}\\ &= (u',\psi(u))_{L^2} + (u',\psi'(u)u)_{L^2}\\ &= \Big(u',\underline{\psi(u)+\psi'(u)u}\Big)_{L^2}. \end{aligned} \tag{3.31}$$

In order to be able to combine this with (3.30), we must have the identity

$$\psi(u)+\psi'(u)u=\varphi(u).$$

Hence, for a given φ, let find ψ to satisfy the equation

$$\psi(s)+\psi'(s)s=\varphi(s)\quad \forall s\in\mathbb{R}.$$

This equation it equivalent to $(\psi(s)s)' = \varphi(s)$, whence we obtain a solution ψ as follows:

$$\psi(s)=\frac{1}{s}\int_0^s\varphi(t)\,dt=\frac{1}{s}\int_0^1\varphi(s\xi)\,d(s\xi)=\int_0^1\varphi(s\xi)\,d\xi. \tag{3.32}$$

It is easy to see from (3.32) that $\psi\in C^\infty(\mathbb{R})$ and that ψ satisfies (3.29). Indeed, we have

$$\psi'(s)=\int_0^1\frac{d}{ds}\varphi(s\xi)\,d\xi=\int_0^1\xi\varphi'(s\xi)\,d\xi,$$

whence $\psi'(s)\geq 0$ and

$$\psi'(s)\leq\int_0^1\xi d\xi=\frac{1}{2}<1.$$

By (3.31) and (3.30) we obtain

$$\frac{d}{dt}(u,\psi(u))_{L^2}=(u',\varphi(u))_{L^2}\leq 0.$$

Hence, the function $F(t):=(u(t),\psi(u(t)))_{L^2}$ is decreasing in $t\in(0,T)$. Observe that

$$F(t)=(u,\psi(u))_{L^2}=\int_\Omega u\psi(u)\,d\mu=\int_\Omega u_+\psi(u_+)\,d\mu=(u_+,\psi(u_+))_{L^2}.$$

Since by (3.29) $\psi(s)\leq s$ for any $s\geq 0$, it follows that

$$F(t)=(u_+,\psi(u_+))_{L^2}\leq(u_+,u_+)_{L^2}=\|u_+\|_{L^2}^2.$$

By hypothesis, $\|u_+\|_{L^2}\to 0$ as $t\to 0$. Hence, the function $F(t)$ is non-negative, decreasing on $(0,T)$ and $F(t)\to 0$ as $t\to 0$. It follows that $F(t)=0$ for all $t\in(0,T)$, that is,

$$\int_\Omega u_+\psi(u_+)\,d\mu=0,$$

which implies that $u_+(t)=0$ for all $t\in(0,T)$. Therefore, $u(t)\leq 0$ 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 3.10 (Positivity-preserving property) *If* $f \geq 0$ *then* $P_t^\Omega f \geq 0$.

Proof. The function $u(t) = P_t^\Omega f$ is caloric. It 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 $u(t) \xrightarrow{L^2} f \geq 0$. By the minimum principle we conclude that $u(t) \geq 0$, that is, $P_t^\Omega f \geq 0$. ∎

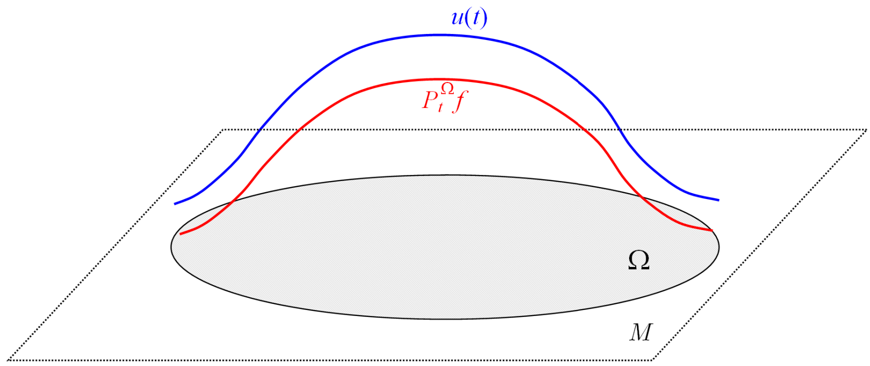
Corollary 3.11 (Minimality property of P_t^Ω) *Let* $u : (0,T) \to L^2(\Omega)$ *satisfy the following properties:*

(*i*) $u(t)$ *is supercaloric in* $(0,T) \times \Omega$

(*ii*) $u(t) \geq 0 \bmod W_0^1(\Omega)$ *for any* $t \in (0,T)$;

(*iii*) $L^2\text{-}\lim_{t\to 0} u(t) \geq f$ *for some* $f \in L^2(\Omega)$.

Then $u(t) \geq P_t^\Omega f$ *for all* $t \in (0,T)$.



Proof. The function $v(t) = P_t^\Omega f - u(t)$ satisfies the following conditions:

(*i*) $v(t)$ is subcaloric in $(0,T) \times \Omega$;

(*ii*) $v(t) \leq 0 \bmod W_0^1(\Omega)$ as $P_t^\Omega f = 0 \bmod W_0^1(\Omega)$ and $u(t) \geq 0 \bmod W_0^1(\Omega)$;

(*iii*) $v(t)_+ \xrightarrow{L^2} 0$ as $t \to 0$, as $L^2\text{-}\lim v(t) = L^2\text{-}\lim P_t^\Omega f - L^2\text{-}\lim u(t) \leq f - f = 0$.

By Theorem 3.8, we conclude that $v(t) \leq 0$ whence $P_t^\Omega f \leq u(t)$ follows. ∎

Corollary 3.11 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 $u(t) \xrightarrow{L^2} f$. Indeed, the function $P_t^\Omega f$ is non-negative, caloric and satisfies the initial condition by Corollary 3.10 and Theorem 3.2. If $u(t)$ is any other function with these properties then by Corollary 3.11 we have $u(t) \geq P_t^\Omega f$, which means the minimality of $P_t^\Omega f$.

Corollary 3.12 (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*

$$\left\| P_t^\Omega f \right\|_{L^\infty} \leq \|f\|_{L^\infty}. \tag{3.33}$$

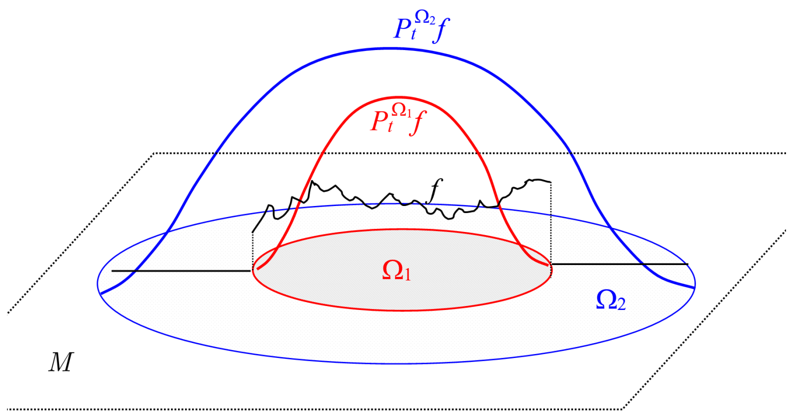
In other words, P_t^Ω is a contraction not only in $L^2(\Omega)$ but also in $L^\infty(\Omega)$.

Proof. If $f \leq 1$ then consider the function $u(t) \equiv 1$ that is caloric and satisfies all the conditions of Corollary 3.11. It follows that $1 \geq P_t^\Omega f$, which was to be proved.

For the proof of (3.33) 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 similarly $P_t^\Omega f \geq -1$. Consequently, $\left\|P_t^\Omega f\right\|_{L^\infty} \leq 1$. ■

In the next statement we compare functions $P_t^\Omega f$ 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 $M \setminus \Omega$. Since $P_t^\Omega f \in W_0^1(\Omega)$ for $t > 0$, it follows that also $P_t^\Omega f \in W_0^1(M)$.

Corollary 3.13 (Monotonicity property) *If $\Omega_1 \subset \Omega_2$ then $P_t^{\Omega_1} f \leq P_t^{\Omega_2} f$ in M for any non-negative $f \in L^2(\Omega_1)$ and all $t > 0$.*



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 $u(t) \overset{L^2(\Omega_2)}{\longrightarrow} f$, it follows that also $u(t) \overset{L^2(\Omega_1)}{\longrightarrow} f$. We conclude by Corollary 3.11 that $u(t) \geq P_t^{\Omega_1} f$ in Ω_1.

Since outside Ω_1 we have $P_t^{\Omega_1} f = 0 \leq P_t^{\Omega_2} f$, if follows that $P_t^{\Omega_1} f \leq P_t^{\Omega_2} f$ in M, which was to be proved. ■

3.6 The heat semigroup as integral operator

We start with the following improvement of Theorem 3.7.

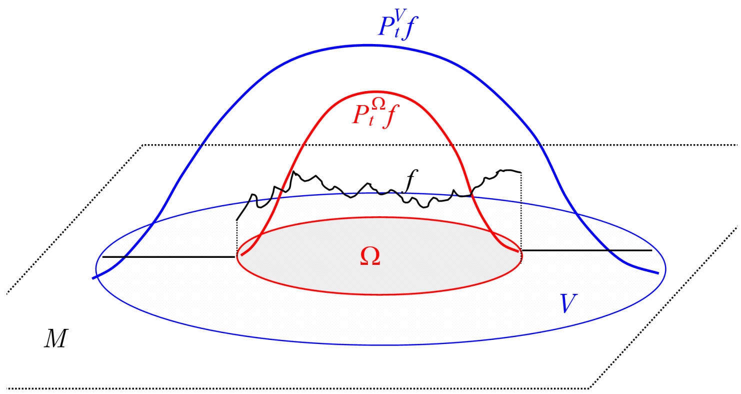
Theorem 3.14 *Let Ω be a precompact open subset of M. For any $f \in L^2(\Omega)$ and $t > 0$,*

$$\left\|P_t^\Omega f\right\|_{C(\Omega)} \leq C\left(1 + t^{-1}\right)^{\frac{n}{4}+1} \|f\|_{L^2(\Omega)}, \tag{3.34}$$

where $C = C(\Omega, \mathbf{g}, D, n)$.

The estimate (3.34) is an improvement of the estimate (3.20) 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 to that, the estimate (3.34) provides the pointwise upper bound for $P_t^\Omega f$ uniformly in the entire domain Ω. As we will see from the proof, the constant C depends on a small open neighborhood of $\overline{\Omega}$.

Proof. For the proof, let us first choose a precompact open subset V of M that covers $\overline{\Omega}$ (for example, one can take for any $x \in \overline{\Omega}$ a precompact open neighborhood V_x, then choose from $\{V_x\}_{x\in\overline{\Omega}}$ a finite family $V_{x_1}, ..., V_{x_N}$ that covers $\overline{\Omega}$ and set $V = \cup V_{x_i}$). Extend f to V by setting $f = 0$ in $V \setminus \Omega$.



Assume first that $f \geq 0$. By Corollaries 3.10 and 3.13, we have

$$0 \leq P_t^\Omega f \leq P_t^V f \text{ in } \Omega. \tag{3.35}$$

Applying the estimate (3.20) 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 $C = C(V, \Omega, \mathbf{g}, D, n)$. Combining with (3.35) we obtain

$$\left\|P_t^\Omega f\right\|_{C(\Omega)} \leq C\left(1 + t^{-1}\right)^{\frac{n}{4}+1} \|f\|_{L^2(\Omega)},$$

which proves (3.34) in the case when f is non-negative.

In the general case, when f is signed, we obtain, using $f = f_+ - f_-$ that

$$\begin{aligned} \left\|P_t^\Omega f\right\|_{C(\Omega)} &= \left\|P_t^\Omega f_+ - P_t^\Omega f_-\right\|_{C(\Omega)} \\ &\leq \left\|P_t^\Omega f_+\right\|_{C(\Omega)} + \left\|P_t^\Omega f_-\right\|_{C(\Omega)} \\ &\leq C\left(1 + t^{-1}\right)^{\frac{n}{4}+1} \left(\|f_+\|_{L^2(\Omega)} + \|f_-\|_{L^2(\Omega)}\right) \\ &\leq 2C\left(1 + t^{-1}\right)^{\frac{n}{4}+1} \|f\|_{L^2(\Omega)}, \end{aligned}$$

which finishes the proof. ∎

The estimate (3.34) will allow us to prove that the operator P_t^Ω is an *integral operator*.

04-Feb-26 Lecture 27

Theorem 3.15 (Existence of the integral kernel of P_t^Ω). *For any $t>0$ and $x\in\Omega$, there exists a function $q_{t,x}\in L^2(\Omega)$ such that*

$$P_t^\Omega f(x)=\int_\Omega q_{t,x}(y)\,f(y)\,d\mu(y) \tag{3.36}$$

for all $f\in L^2(\Omega)$. Besides, we have

$$\|q_{t,x}\|_{L^2}\le C\left(1+t^{-1}\right)^{\frac{n}{4}+1}=:\Phi(t), \tag{3.37}$$

where C is the same constant as in Theorem 3.14.

The function $q_{t,x}(y)$ is called the *integral kernel* of the operator P_t^Ω.

Proof. Fix $t>0$ and $x\in\Omega$ and consider the following linear functional on $L^2(\Omega)$:

$$\begin{array}{rcl}L^2(\Omega)&\to&\mathbb{R}\\ f&\mapsto&P_t^\Omega f(x)\end{array} \tag{3.38}$$

By (3.34) we have

$$\left|P_t^\Omega f(x)\right|\le\Phi(t)\,\|f\|_{L^2}. \tag{3.39}$$

Hence, the functional (3.38) 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 (3.39) that

$$\|q_{t,x}\|_{L^2}^2\le\Phi(t)\,\|q_{t,x}\|_{L^2},$$

whence (3.37) follows. ■

Our purpose in what follows is twofold:

1. Using the integral kernel $q_{t,x}$, we will prove that the function $P_t^\Omega f(x)$ is smooth jointly in t,x.

2. We will show that the function $q_{t,x}(y)$ has a version that is smooth jointly in t,x,y – this will be called the *heat kernel*.

3.7 The trace of the heat semigroup

Next, we need the notion of *trace* of operators. Let A be a non-negative definite operator in a Hilbert space H, that is, $(Au, u) \geq 0$ 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 but in our case of $A = P_t^{\Omega}$ this will be done by a specific argument.

Lemma 3.16 *For any precompact open set $\Omega \subset M$ and any $t > 0$,*

$$\operatorname{trace} P_{2t}^{\Omega} = \int_{\Omega} \|q_{t,x}\|_{L^2}^2 d\mu(x) < \infty. \tag{3.40}$$

Besides, we have

$$\operatorname{trace} P_t^{\Omega} = \sum_{k=1}^{\infty} e^{-t\lambda_k(\Omega)}. \tag{3.41}$$

Consequently, the series (3.41) converges for any $t > 0$, that is,

$$\sum_{k=1}^{\infty} e^{-t\lambda_k(\Omega)} < \infty.$$

Proof. Recall that the operator P_{2t}^{Ω} is non-negative definite by Theorem 3.5. To prove the identity in (3.40), choose any orthonormal basis $\{h_k\}_{k=1}^{\infty}$ in L^2. Using (3.36) with $f = h_k$, we obtain

$$\begin{aligned}\operatorname{trace} P_{2t}^{\Omega} &= \sum_{k=1}^{\infty} \left(P_{2t}^{\Omega} h_k, h_k\right)_{L^2} = \sum_{k=1}^{\infty} \left(P_t^{\Omega} h_k, P_t^{\Omega} h_k\right)_{L^2} \\ &= \sum_{k=1}^{\infty} \int_{\Omega} \left(P_t^{\Omega} h_k(x)\right)^2 d\mu(x) \\ &= \int_{\Omega} \sum_{k=1}^{\infty} (q_{t,x}, h_k)_{L^2}^2 \, d\mu(x).\end{aligned} \tag{3.42}$$

Applying the Parseval identity in the basis $\{h_k\}$, we obtain

$$\sum_{k=1}^{\infty} (q_{t,x}, h_k)_{L^2}^2 = \|q_{t,x}\|_{L^2}^2. \tag{3.43}$$

Hence, (3.42) and (3.43) yield

$$\operatorname{trace} P_{2t}^{\Omega} = \int_{\Omega} \|q_{t,x}\|_{L^2}^2 d\mu(x),$$

which proves the first part of (3.40). In particular, we see that trace P_{2t}^{Ω} does not depend on the choice of the basis $\{h_k\}$.

The second part of (3.40), that is, the finiteness of the trace, follows from Theorem 3.15 since for all $x \in \Omega$ and $t > 0$,

$$\|q_{t,x}\|_{L^2} \le C\left(1+t^{-1}\right)^{\frac{n}{4}+1} =: \Phi(t)$$

and, hence,

$$\operatorname{trace} P_{2t}^{\Omega} \le \int_{\Omega} \Phi(t)^2 \, d\mu(x) = \Phi(t)^2 \mu(\Omega) = C\left(1+t^{-1}\right)^{\frac{n}{2}+2} \mu(\Omega) < \infty. \tag{3.44}$$

Finally, in order to prove (3.41), let us compute the trace of P_t^{Ω} in the orthonormal basis $\{v_k\}_{k=1}^{\infty}$ of the eigenfunctions of Δ in Ω, that is,

$$\operatorname{trace} P_t^{\Omega} = \sum_{k=1}^{\infty} \left(P_t^{\Omega} v_k, v_k\right)_{L^2}.$$

Recall that by the definition (3.16) of the operator P_t^{Ω}, we have

$$P_t^{\Omega} v_k = e^{-t\lambda_k(\Omega)} v_k.$$

Hence, we obtain

$$\operatorname{trace} P_t^{\Omega} = \sum_{k=1}^{\infty} \left(e^{-t\lambda_k(\Omega)} v_k, v_k\right)_{L^2} = \sum_{k=1}^{\infty} e^{-t\lambda_k(\Omega)} \|v_k\|_{L^2}^2 = \sum_{k=1}^{\infty} e^{-t\lambda_k(\Omega)},$$

which finishes the proof. ■

As a by-product of the proof, we obtain from (3.44) (by changing $2t$ to t) the following useful inequality:

$$\boxed{\operatorname{trace} P_t^{\Omega} \le C\left(1+t^{-1}\right)^{\frac{n}{2}+2} \mu(\Omega).}$$

3.8 Smoothness of the heat semigroup in t, x

We know that, for any $t > 0$, the function $P_t^{\Omega} f(x)$ is smooth in $x \in \Omega$. Here we prove that $P_t^{\Omega} f(x)$ is smooth jointly in t and x.

Theorem 3.17 *Let Ω be a precompact open subset of M. For any $f \in L^2(\Omega)$, 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). Consequently, u satisfies in $\mathbb{R}_+ \times \Omega$ the heat equation $\partial_t u = \Delta u$ in the classical sense.*

Proof. We know that if

$$f = \sum_{k=1}^{\infty} a_k v_k.$$

then

$$u(t,x)=\sum_{k=1}^{\infty}e^{-\lambda_k t}a_k v_k(x), \tag{3.45}$$

where the series converges in $L^2(\Omega)$ for any $t>0$.

Let $U\Subset\Omega$ be a precompact chart. We claim that, for any $\varepsilon>0$ and any positive integer m,

$$\sum_{k=1}^{\infty}\left\|e^{-\lambda_k t}a_k v_k(x)\right\|_{C^m((\varepsilon,\infty)\times U)}<\infty,$$

which will imply that the series in (3.45) converges also in $C^m((\varepsilon,\infty)\times U)$, whence it follows that $u\in C^m((\varepsilon,\infty)\times U)$. Since $\varepsilon>0$ and m are arbitrary, this will imply that $u\in C^\infty(\mathbb{R}_+\times\Omega)$.

Recall that, by Corollary 2.8, if $\Delta v=\alpha v$ in the weak sense then $u\in C^\infty(\Omega)$ and

$$\|v\|_{C^m(U)}\leq C(1+|\alpha|)^{\frac{m}{2}+\frac{n}{4}+\frac{1}{2}}\|v\|_{W^1(\Omega)}.$$

Since $\Delta v_k=-\lambda_k v_k$, we obtain

$$\|v_k\|_{C^m(U)}\leq C(\lambda_k+1)^{\frac{m}{2}+\frac{n}{4}+\frac{1}{2}}\|v_k\|_{W^1(\Omega)}.$$

Since by (2.33)

$$\|v_k\|_{W^1(\Omega)}=(\lambda_k+1)^{1/2}\|v_k\|_{L^2(\Omega)}=(\lambda_k+1)^{1/2},$$

it follows that

$$\|v_k\|_{C^m(U)}\leq C(\lambda_k+1)^{\sigma}, \tag{3.46}$$

where $\sigma=\frac{m}{2}+\frac{n}{4}+1$.

For any partial derivative $\partial_x^\alpha\partial_t^\gamma$, where α is an n-dimensional multiindex and γ is a non-negative integer such that

$$|\alpha|+\gamma\leq m,$$

we have

$$\partial_x^\alpha\partial_t^\gamma\left(e^{-\lambda_k t}v_k(x)\right)=(-\lambda_k)^\gamma e^{-\lambda_k t}\partial_x^\alpha v_k(x).$$

Since $|\alpha|\leq m$ and by (3.46)

$$\sup_U|\partial_x^\alpha v_k(x)|\leq C(\lambda_k+1)^\sigma,$$

it follows that

$$\sup_{(\varepsilon,\infty)\times U}\left|\partial_x^\alpha\partial_t^\gamma\left(e^{-\lambda_k t}v_k(x)\right)\right|\leq C\lambda_k^\gamma e^{-\lambda_k\varepsilon}(\lambda_k+1)^\sigma\leq Ce^{-\lambda_k\varepsilon}(\lambda_k+1)^{\sigma+m}.$$

Since $(\lambda+1)^{\sigma+m}\leq C_\varepsilon e^{\lambda\varepsilon/2}$ for all $\lambda\geq0$, it follows that

$$\sup_{(\varepsilon,\infty)\times U}\left|\partial_x^\alpha\partial_t^\gamma\left(e^{-\lambda_k t}v_k(x)\right)\right|\leq C_\varepsilon e^{-\lambda_k\varepsilon/2},$$

which implies

$$\left\|e^{-\lambda_k t}v_k(x)\right\|_{C^m((\varepsilon,\infty)\times U)}\leq C_\varepsilon e^{-\lambda_k\varepsilon/2}. \tag{3.47}$$

By the Cauchy-Schwarz inequality and (3.47), we obtain

$$\sum_{k=1}^{\infty} \left\| e^{-\lambda_k t} a_k v_k(x) \right\|_{C^m((\varepsilon,\infty)\times U)} \leq \left(\sum_{k=1}^{\infty} a_k^2 \right)^{1/2} \left(\sum_{k=1}^{\infty} \left\| e^{-\lambda_k t} v_k(x) \right\|^2_{C^m((\varepsilon,\infty)\times U)} \right)^{1/2}$$

$$\leq C_\varepsilon \|f\|_{L^2} \left(\sum_{k=1}^{\infty} e^{-\lambda_k \varepsilon} \right)^{1/2} < \infty,$$

where the last series converges by Lemma 3.16. ■

06-Feb-26 Lecture 28

3.9 The heat kernel in precompact domains

By Theorem 3.15, the operator P_t^Ω in $L^2(\Omega)$ has an integral kernel $q_{t,x} \in L^2(\Omega)$, that is, for any $f \in L^2$ and all $t > 0$, $x \in \Omega$,

$$P_t^\Omega f(x) = \int_\Omega q_{t,x,}(y) f(y) d\mu.$$

Definition. Define the *heat kernel* $p_t^\Omega(x,y)$ of Δ in Ω by the identity

$$p_t^\Omega(x,y) = q_{t,x}(y).$$

So far $p_t^\Omega(x,y)$ is defined as an L^2-function of y, for any $t > 0$ and $x \in \Omega$. The next theorem shows that it has a smooth version of t, x, y.

Theorem 3.18 (Eigenfunction expansion of the heat kernel) *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, and $\{\lambda_k\}_{k=1}^\infty$ be corresponding eigenvalues. Then heat kernel $p_t^\Omega(x,y)$ admits the following eigenfunction expansion*

$$p_t^\Omega(x,y) = \sum_{k=1}^{\infty} e^{-\lambda_k t} v_k(x) v_k(y), \tag{3.48}$$

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 (3.48) converges in $C^m((\varepsilon,\infty) \times U \times U)$, for any positive integer m, for any $\varepsilon > 0$ and any precompact chart $U \Subset \Omega$. Consequently, $p_t^\Omega(x,y) \in C^\infty(\mathbb{R}_+ \times \Omega \times \Omega)$.

Proof. Let us first prove that the series

$$\sum_{k=1}^{\infty} e^{-t\lambda_k} v_k(x) v_k(y) \tag{3.49}$$

converges absolutely and uniformly in the domain $t > \varepsilon$, $x \in \Omega$, $y \in \Omega$. By the Weierstrass M-test, it suffices to prove that

$$\sum_{k=1}^{\infty} \sup_{t>\varepsilon, x,y\in\Omega} \left| e^{-t\lambda_k} v_k(x) v_k(y) \right| < \infty. \tag{3.50}$$

Recall that, by Theorem 3.14, for any $f \in L^2(\Omega)$ and $t > 0$,

$$\sup_{x\in\Omega} \left| P_t^\Omega f(x) \right| \le C \left(1 + t^{-1}\right)^{\frac{n}{4}+1} \|f\|_{L^2},$$

Applying this to $f = v_k$ and using that by (3.52)

$$P_t^\Omega v_k = e^{-t\lambda_k} v_k$$

and $\|v_k\|_{L^2} = 1$, we obtain

$$\sup_{x\in\Omega} \left| e^{-t\lambda_k} v_k(x) \right| \le C \left(1 + t^{-1}\right)^{\frac{n}{4}+1}, \tag{3.51}$$

whence

$$\sup_{x\in\Omega} |v_k(x)| \le C \left(1 + t^{-1}\right)^{\frac{n}{4}+1} e^{\lambda_k t}.$$

In particular, setting here $t = \varepsilon/4$, we obtain

$$\sup_{x\in\Omega} |v_k(x)| \le C_\varepsilon e^{\lambda_k \varepsilon/4}.$$

It follows that

$$\sup_{t>\varepsilon, x,y\in\Omega} \left| e^{-t\lambda_k} v_k(x) v_k(y) \right| \le C_\varepsilon e^{-\lambda_k \varepsilon} (e^{\lambda_k \varepsilon/4})^2 = C_\varepsilon e^{-\lambda_k \varepsilon/2}.$$

and (3.50) follows from $\sum_k e^{-\lambda_k \varepsilon/2} < \infty$ (Lemma 3.16).

Now let us prove that the sum of the series (3.49) is equal to $p_t^\Omega(x,y)$. Using the notation $q_{t,x}$ as above and noticing that

$$(q_{t,x}, v_k)_{L^2} = P_t^\Omega v_k(x) = e^{-t\lambda_k} v_k(x), \tag{3.52}$$

we obtain the following expansion of $q_{t,x}$ in the basis $\{v_k\}$:

$$q_{t,x} = \sum_{k=1}^{\infty} e^{-t\lambda_k} v_k(x) v_k, \tag{3.53}$$

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 jointly continuous in t,x,y by the previous argument.

Finally, let us show that the series (3.49) converges in $C^m\left((\varepsilon,\infty)\times U\times U\right)$, which will imply that the heat kernel $p_t^\Omega(x,y)$ is C^∞ smooth jointly in $t>0$ and $x,y\in\Omega$. Again, it suffices to prove that

$$\sum_{k=1}^{\infty}\left\|e^{-\lambda_k t}v_k\left(x\right)v_k\left(y\right)\right\|_{C^m((\varepsilon,\infty)\times U\times U)}<\infty. \tag{3.54}$$

By (3.46) we have

$$\|v_k\|_{C^m(U)}\leq C\left(\lambda_k+1\right)^{\sigma}, \tag{3.55}$$

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

$$\partial_x^\alpha\partial_y^\beta\partial_t^\gamma\left(e^{-\lambda_k t}v_k\left(x\right)v_k\left(y\right)\right)=\left(-\lambda_k\right)^\gamma e^{-\lambda_k t}\partial_x^\alpha v_k\left(x\right)\partial_y^\beta v_k\left(y\right).$$

It follows from (3.55) that

$$\sup_{t>\varepsilon,x\in U,y\in U}\left|\partial_x^\alpha\partial_y^\beta\partial_t^\gamma\left(e^{-\lambda_k t}v_k\left(x\right)v_k\left(y\right)\right)\right|\leq C\lambda_k^\gamma e^{-\lambda_k\varepsilon}\left(\lambda_k+1\right)^{2\sigma}.$$

Since $\gamma\leq m<2\sigma$, we obtain

$$\sup_{t>\varepsilon,x\in U,y\in U}\left|\partial_x^\alpha\partial_y^\beta\partial_t^\gamma\left(e^{-\lambda_k t}v_k\left(x\right)v_k\left(y\right)\right)\right|\leq Ce^{-\lambda_k\varepsilon}\left(\lambda_k+1\right)^{4\sigma}\leq C_\varepsilon e^{-\lambda_k\varepsilon/2},$$

whence

$$\left\|e^{-\lambda_k t}v_k\left(x\right)v_k\left(y\right)\right\|_{C^m((\varepsilon,\infty)\times U\times U)}\leq C_\varepsilon e^{-\lambda_k\varepsilon/2}. \tag{3.56}$$

Hence, (3.54) follows from $\sum_k e^{-\lambda_k\varepsilon/2}<\infty$. ■

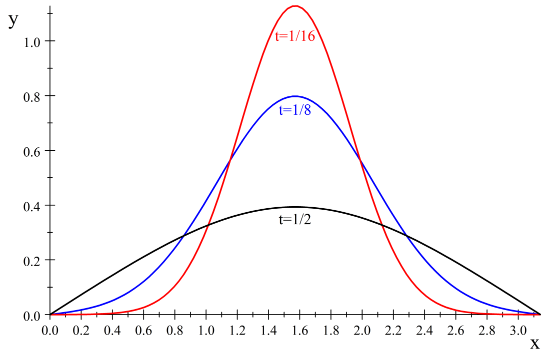
Remark. If the boundary $\partial\Omega$ is smooth, for example, a C^1-submanifold, then one can show that $v_k\in C\left(\overline{\Omega}\right)$ and $v_k|_{\partial\Omega}=0$. The fact that the series in (3.45) and 5(3.48) converge absolutely and uniformly in $(t,x,y)\in(\varepsilon,\infty)\times\Omega\times\Omega$, implies that $P_t^\Omega f\left(x\right)=0$ when $x\in\partial\Omega$ and also $p_t\left(x,y\right)=0$ when one of the points x,y belongs to $\partial\Omega$.

Example. Let $M=\mathbb{R}$ and $\Omega=(0,\pi)$. As we know, the eigenfunctions of Δ in $(0,\pi)$ are $\{\sin kx\}_{k=1}^\infty$ with the eigenvalues $\lambda_k=k^2$, where $k=1,2,....$ Since

$$\int_0^\pi\sin^2 kx dx=\frac{1}{2}\int_0^\pi\left(1-\cos 2kx\right)dx=\frac{\pi}{2},$$

the normalized eigenfunctions are $v_k(x)=\sqrt{\frac{2}{\pi}}\sin kx$. Hence, we obtain

$$p_t^\Omega(x,y)=\frac{2}{\pi}\sum_{k=1}^{\infty}e^{-k^2t}\sin kx\sin ky.$$



The graphs of the functions $x \mapsto p_t(x,y)$ for $y = \pi/2$ and for $t = \frac{1}{16}$ (red), $t = \frac{1}{8}$ (blue) and $t = \frac{1}{2}$ (black)

Clearly, $p_t(x,y) \to 0$ as $t \to \infty$. Consequently, also $P_t^{\Omega} f(x) \to 0$ as $t \to \infty$.

Example. If M is compact then we can set $\Omega = M$ and, hence, obtain the heat kernel $p_t(x,y)$ on the entire manifold M that is given by

$$p_t(x,y) = \sum_{k=1}^{\infty} e^{-\lambda_k t} v_k(x)\, v_k(y), \tag{3.57}$$

where $\{v_k\}$ is an orthonormal basis in $L^2(M)$ that consists of eigenfunctions of Δ in M and $\{\lambda_k\}$ are their eigenvalues.

Let us compute the heat kernel on $M = \mathbb{S}^1$. 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:

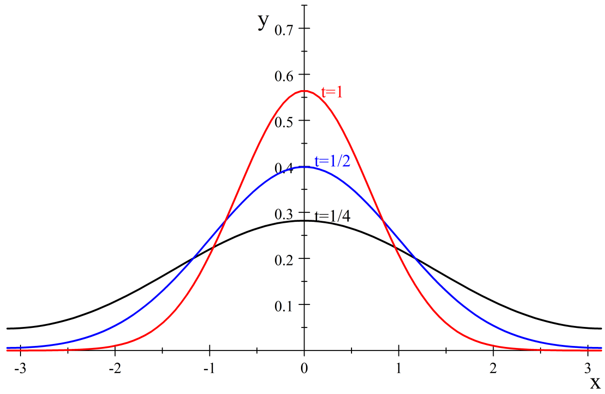
$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, ..., \frac{\cos mx}{\sqrt{\pi}}, \frac{\sin mx}{\sqrt{\pi}}, ...,$$

with the eigenvalues

$$0, 1, 1, ..., m^2, m^2, ... \ .$$

By (3.57) we obtain

$$\begin{aligned} p_t(x,y) &= \frac{1}{2\pi} + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2 t} \cos mx \cos my + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2 t} \sin mx \sin my \\ &= \frac{1}{2\pi} + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2 t} \cos m(x-y). \end{aligned} \tag{3.58}$$



The graphs of the heat kernel $p_t(x,0)$ on $\mathbb{S}^1$ for $t=1$ (red), $t=1/2$ (blue) and $t=1/4$ (black)

It follows from (3.58) that $p_t(x,y) \to \frac{1}{2\pi}$ as $t \to \infty$. One can derive that as $t \to \infty$

$$P_t f(x) \to \frac{1}{2\pi} \int_{\mathbb{S}^1} f(\theta)\, d\theta.$$

3.10 * 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 (3.48). 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

$$P_t^\Omega f(x) = \int_\Omega p_t^\Omega(x,y) f(y)\, d\mu(y). \tag{3.59}$$

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 $t > 0$*

$$\int_\Omega p_t^\Omega(x,y)\, d\mu(y) \leq 1. \tag{3.60}$$

(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* $t, s > 0$*,*

$$p_{t+s}^\Omega(x,y) = \int_\Omega p_t^\Omega(x,z)\, p_s^\Omega(z,y)\, d\mu(z). \tag{3.61}$$

(*g*) *Monotonicity: if* $\Omega_1 \subset \Omega_2$ *then* $p_t^{\Omega_1}(x,y) \le p_t^{\Omega_2}(x,y)$ *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 3.10 we must have $P_{t_0}^\Omega f(x_0) \ge 0$. This contradiction shows that $p_t(x,y) \ge 0$.

(b) By Corollary 3.12, $f \le 1$ implies $P_t^\Omega f(x) \le 1$ for all $x \in M$ and $t > 0$. Taking $f = 1_\Omega$, we obtain

$$\int_\Omega p_t^\Omega(x,y)\, d\mu(y) \le 1,$$

which was to be proved.

(c) The symmetry follows trivially from the eigenfunction expansion (3.48).

$(d) + (e)$ Fix $y \in \Omega$. As follows from the proof of Theorem 3.18, the series

$$u(t) := p_t^\Omega(\cdot, y) = \sum_{k=1}^\infty e^{-\lambda_k t} v_k(y)\, v_k, \tag{3.62}$$

converges in $L^2(\Omega)$ for any $t > 0$. Indeed, (3.62) is obtained from (3.53) by switching the variables x and y and using the symmetry of the heat kernel. For any $t > s > 0$, we obtain using (3.45) that

$$u(t) = \sum_{k=1}^\infty e^{-\lambda_k t} v_k(y)\, v_k = \sum_{k=1}^\infty e^{-\lambda_k (t-s)} \left(e^{-\lambda_k s} v_k(y)\right) v_k = P_{t-s}^\Omega u(s). \tag{3.63}$$

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 (3.62) term-by-term, which is possible because that series converges in any C^m.

(f) Rewriting the identity (3.63) by using (3.59) and the definition (3.62) 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 (3.61).

(g) 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 3.13, 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. ■

3.11 * The initial condition

As we know, for any $f\in L^2(\Omega)$, we have

$$P_t^\Omega f\overset{L^2}{\longrightarrow}f\ \text{ as } t\to 0.$$

Here we show that the convergence is "better" if 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

$$P_t^\Omega f\to f \text{ as } t\to 0, \tag{3.64}$$

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

$$\int_U p_t^\Omega(x,y)\,d\mu(y)\to 1 \text{ as } t\to 0, \tag{3.65}$$

where the convergence is local uniform in U.

(c) For any $f\in C_b(\Omega)$, the convergence (3.64) 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 $f=\sum_{k=1}^\infty a_kv_k$ then

$$\Delta^j f=(-1)^j\sum_{k=1}^\infty \lambda_k^j a_k v_k$$

(cf. Exercise 58), 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 By Lemma 3.6 and the identity (3.23) in the proof of Theorem 3.7). By Lemma 3.4 (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

$$P_t^\Omega f - f \overset{C^m(K)}{\longrightarrow} 0 \quad \text{as } t \to 0, \tag{3.66}$$

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 (3.65) follows.

(c) We have

$$\begin{aligned} P_t^\Omega f(x) - f(x) &= \int_\Omega p_t^\Omega(x,y)\,(f(y) - f(x))\, d\mu(y) \\ &\quad + \left(\int_\Omega p_t^\Omega(x,y)\, d\mu(y) - 1\right) f(x). \end{aligned}$$

By (b) we obtain that

$$\left(\int_\Omega p_t^\Omega(x,y)\, d\mu(y) - 1\right) f(x) \to 0 \tag{3.67}$$

as $t \to 0$, where the convergence is local uniform in x. Choose an open set U containing x and such that $|f(y) - f(x)| \le \varepsilon$ for any $y \in U$, where $\varepsilon > 0$ is prescribed. Then we have

$$\begin{aligned} \left|\int_\Omega p_t^\Omega(x,y)\,(f(y) - f(x))\, d\mu(y)\right| &\le \left|\int_U p_t^\Omega(x,y)\,(f(y) - f(x))\, d\mu(y)\right| \\ &\quad + \left|\int_{\Omega\setminus U} p_t^\Omega(x,y)\,(f(y) - f(x))\, d\mu(y)\right| \\ &\le \varepsilon \int_U p_t^\Omega(x,y)\, d\mu(y) + 2\sup|f| \int_{\Omega\setminus U} p_t^\Omega(x,y)\, d\mu(y) \\ &\le \varepsilon + 2\sup|f| \left(1 - \int_U p_t^\Omega(x,y)\, d\mu(y)\right). \end{aligned}$$

As $t \to 0$ we obtain using (b) that

$$\limsup_{t\to 0}\left|\int_\Omega p_t^\Omega(x,y)\,(f(y)-f(x))\,d\mu(y)\right| \le \varepsilon.$$

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 (3.67), 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 (3.64) 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

$$u(t,\cdot) \overset{L^2(\Omega)}{\to} f \quad \text{as } t \to 0.$$

If $f \in C_b(\Omega)$ then by Theorem 3.20 we have also that

$$u(t,\cdot) \overset{C(K)}{\to} f \quad \text{as } t \to 0.$$

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

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 $u \in W^1_{loc}(M)$,

$$u_k \overset{W^1_{loc}(M)}{\longrightarrow} u \quad \text{as } k \to \infty,$$

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

$$\|u_k - u_l\|_{C^m(K)} \le C\,\|u_k - u_l\|_{W^1(\Omega)},$$

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 3.10 and 3.13, 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 3.11), 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 $t > 0$,*

$$\|P_t f\|_{L^2(M)} \le \|f\|_{L^2(M)} \tag{4.3}$$

and

$$P_t f \overset{L^2(M)}{\longrightarrow} f \quad \text{as } t \to 0. \tag{4.4}$$

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 $f \ge 0$. Then we have, for any precompact domain $\Omega \subset M$, that

$$\left\|P_t^\Omega f\right\|_{L^2(\Omega)} \le \|f\|_{L^2(\Omega)}.$$

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:

$$\begin{array}{ccc} P_t^{\Omega_k} f & \le & P_t f \\ {\scriptstyle t\to 0}\downarrow^{L^2(M)} & & \le_{\|\cdot\|_{L^2}} \\ f\mathbf{1}_{\Omega_k} & \underset{k\to\infty}{\overset{L^2(M)}{\longrightarrow}} & f \end{array}$$

Using Lemma 4.7 to be stated and proved below, we conclude that $P_t f \overset{L^2(M)}{\longrightarrow} f$ 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

$$\begin{aligned} \|P_t f\|_{L^2}^2 &= \|P_t f_+ - P_t f_-\|_{L^2}^2 \\ &= \|P_t f_+\|_{L^2}^2 + \|P_t f_-\|_{L^2}^2 - 2\left(P_t f_+, P_t f_-\right)_{L^2} \\ &\le \|P_t f_+\|_{L^2}^2 + \|P_t f_-\|_{L^2}^2 \\ &\le \|f_+\|_{L^2}^2 + \|f_-\|_{L^2}^2 = \|f\|_{L^2}^2. \end{aligned}$$

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,

$$u_{ik} \le u_i \quad \text{and} \quad \|u_i\|_{L^2} \le \|f\|_{L^2}.$$

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:

$$\begin{array}{ccc} u_{ik} & \leq & u_i \\ {}_{i\to\infty}\downarrow_{L^2} & & \leq_{\|\cdot\|_{L^2}} \\ f_k & \overset{L^2}{\underset{k\to\infty}{\longrightarrow}} & f \end{array}$$

where all notation are self-explanatory. We need to prove that also $u_i \overset{L^2}{\to} f$ 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). ■

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 (1.78).

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 $P_t f_+$ and $P_t f_-$ are finite (moreover, they are smooth), it follows that $P_t f$ 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

$$u(t,\cdot) \to \delta_y \quad \text{as } t \to 0.$$

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,*

$$\int_U u(t,\cdot)\, d\mu \to 1 \quad \text{as } t \to 0. \tag{4.13}$$

(c) *For any $f \in C_b(M)$,*

$$\int_M u(t,\cdot) f\, d\mu \to f(y) \quad \text{as } t \to 0. \tag{4.14}$$

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

$$\int_M u(t,\cdot) f \, d\mu \le \int_U u(t,\cdot)\, d\mu \le 1,$$

(4.13) follows from (4.14).

$(b) \Rightarrow (c)$. For any open set U containing y, we have

$$\begin{aligned}\int_M u(t,x) f(x)\, d\mu(x) &= \int_{M\setminus U} u(t,x) f(x)\, d\mu(x) \\ &+ \int_U u(t,x)\left(f(x) - f(y)\right) d\mu(x) \\ &+ f(y)\int_U u(t,x)\, d\mu(x).\end{aligned}$$

The last term here tends to $f(y)$ by (4.13). The other terms are estimated as follows:

$$\left|\int_{M\setminus U} u(t,x) f(x)\, d\mu\right| \le \sup|f| \int_{M\setminus U} u(t,x)\, d\mu(x) \tag{4.15}$$

and

$$\begin{aligned}\left|\int_U u(t,x)\left(f(x) - f(y)\right) d\mu\right| &\le \sup_{x\in U}|f(x) - f(y)| \int_U u(t,x)\, d\mu(x) \\ &\le \sup_{x\in U}|f(x) - f(y)|.\end{aligned} \tag{4.16}$$

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

$$u(t+s,\cdot) \xrightarrow{L^2_{loc}} f \quad \text{as } t \to 0.$$

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,

$$v(t,\cdot) \leq 0 \bmod W_0^1(\Omega).$$

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 3.8, we conclude that $v(t,x) \leq 0$ for all $t \in (0, T-s)$ and $x \in \Omega$.

It follows that in Ω

$$u(t+s,\cdot) \leq P_t^\Omega u(s,\cdot) + \varepsilon$$

whence, for any $x \in \Omega$,

$$u(t+s,x) \leq \int_\Omega p_t(x,z)\, u(s,z)\, d\mu(z) + \varepsilon.$$

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,

$$u(t,x) \leq p_t(x,y) + \varepsilon,$$

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 $u(t,x) \leq p_t(x,y)$, 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.

4.6 Heat kernel and isometries

Lemma 4.12 *Let $\Phi : Y \to X$ be an isometry of two weighted manifolds $(X, \mathbf{g}_X, \mu_X)$ and $(Y, \mathbf{g}_Y, \mu_Y)$. Then the following is true:*

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

$$\int_Y (\Phi_* f)\, d\mu_Y = \int_X f\, d\mu_X. \tag{4.19}$$

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

$$\Phi_*(\Delta_X f) = \Delta_Y(\Phi_* f), \tag{4.20}$$

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

Remark. The identity (4.19) can be rewritten as follows:

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

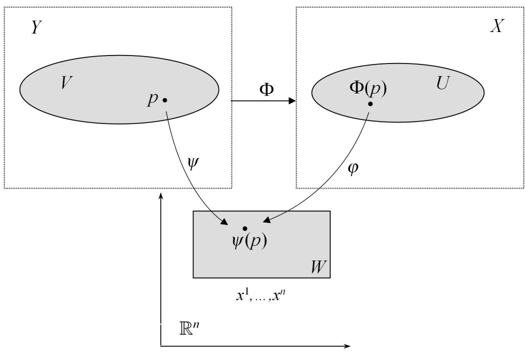
and in this form it can be regarded as change of variables $x = \Phi(y)$ in integration. Note that this identity does not contain the determinant of the Jacobi matrix like in the classical formula (1.48) because the determinant is hidden in the definitions of the measures μ_X and μ_Y.

Proof. Because of a partition of unity, it suffices to prove the both identities (4.19) and (4.20) 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

$$V \overset{\Phi}{\to} U \overset{\varphi}{\to} W$$

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



Mappings Φ, φ, ψ

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 $x^1, ..., x^n$ 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 (4.19) and (4.20) are trivially satisfied. ■

Theorem 4.13 *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* $x, y \in M$,

$$p_t(Jx, Jy) = p_t(x, y) \tag{4.21}$$

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 4.12, 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 4.12, we have the identity

$$\int_M f(Jx)\, d\mu(x) = \int_M f(z)\, d\mu(z), \tag{4.22}$$

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 4.13, 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 1.14. 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 4.14 *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 4.13 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

$$\begin{cases} \partial_t u = \Delta u, \\ \displaystyle\int_M u(t, x)\,d\mu(x) \le 1, \\ \displaystyle\int_{B_\varepsilon} u(t, x)\,d\mu(x) \to 1 \quad \text{as } t \to 0, \end{cases} \tag{4.23}$$

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 (1.121)

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{S'(r)}{S(r)}\frac{\partial u}{\partial r}.$$

For $0 < \varepsilon < r_0$, we have by (1.118), (1.108), (1.120)

$$\int_{B_\varepsilon} u\,d\mu = \frac{1}{\omega_n}\int_0^\varepsilon \int_{\mathbb{S}^{n-1}} u\,(t, r)\,S(r)\,d\theta dr = \int_0^\varepsilon u(t, r)\,S(r)\,dr.$$

Hence, we obtain the following equivalent form of (4.23):

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial r^2} + \dfrac{S'(r)}{S(r)}\dfrac{\partial u}{\partial r}, \\ \displaystyle\int_0^{r_0} u(t, r)\,S(r)\,dr \le 1, \\ \displaystyle\int_0^\varepsilon u(t, r)\,S(r)\,dr \to 1 \quad \text{as } t \to 0. \end{cases} \tag{4.24}$$

It is important that all the conditions in (4.24) depend on the geometry of M only via the area function $S(r)$. Since by hypothesis $S(r)=\widetilde{S}(r)$, the conditions (4.24) 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 $u(t,|x|)\geq \widetilde{p}_t(x,0)$, that is,

$$p_t(x,0)\geq \widetilde{p}_t(x,0).$$

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

$$d\widetilde{\mu}=h^2d\mu, \tag{4.25}$$

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 4.15 *Let h be a smooth positive function on M that satisfies the equation*

$$\Delta h+\alpha h=0, \tag{4.26}$$

where α is a real constant. Then the following identities holds

$$\widetilde{\Delta}=\frac{1}{h}\circ\Delta\circ h+\alpha\,\mathrm{id}, \tag{4.27}$$

$$\widetilde{p}_t(x,y)=e^{\alpha t}\frac{p_t(x,y)}{h(x)\,h(y)}, \tag{4.28}$$

for all $t>0$ and $x,y\in M$.

The change of measure (4.25) satisfying (4.26) and the associated change of operator (4.27) 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,

$$\begin{aligned}\widetilde{\Delta}f &= \frac{1}{h^2}\operatorname{div}_{\mathbf{g},\mu}(h^2\nabla f)=\operatorname{div}_{\mathbf{g},\mu}(\nabla f)+\frac{1}{h^2}\langle\nabla h^2,\nabla f\rangle_{\mathbf{g}}\\ &= \Delta f+2\langle\frac{\nabla h}{h},\nabla f\rangle_{\mathbf{g}}.\end{aligned} \tag{4.29}$$

On the other hand, using the equation (4.26) 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

$$\widetilde{\Delta} f = \frac{1}{h}\Delta(hf) + \alpha f, \tag{4.30}$$

that is equivalent to (4.27). We have proved this identity for smooth f, but similarly it holds when Δ is understood in the weak sense.

In order to prove (4.28), 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

$$\begin{aligned}
\widetilde{p}_t^\Omega(x,y) &= \sum_k e^{-(\lambda_k-\alpha)t}\frac{v_k(x)}{h(x)}\frac{v_k(y)}{h(y)} \\
&= \frac{e^{\alpha t}}{h(x)h(y)}\sum_k e^{-\lambda_k t} v_k(x) v_k(y) = \frac{e^{\alpha t} p_t^\Omega(x,y)}{h(x)h(y)},
\end{aligned}$$

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

$$p_t(x,y) = \frac{1}{(4\pi t)^{1/2}}\exp\left(-\frac{|x-y|^2}{4t}\right). \tag{4.31}$$

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

$$\widetilde{\Delta} = \frac{1}{h^2}\frac{d}{dx}\left(h^2\frac{d}{dx}\right) = \frac{d^2}{dx^2} + 2\frac{h'}{h}\frac{d}{dx} \tag{4.32}$$

(cf. (4.29)). The equation (4.26) 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 (4.32)

$$\widetilde{\Delta} = \frac{d^2}{dx^2} + 2\beta \coth \beta x \frac{d}{dx}.$$

By Theorem 4.15, 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 (4.32) with $h = e^{\frac{1}{2}x^2}$,

$$\Delta = \frac{d^2}{dx^2} + 2x\frac{d}{dx}. \tag{4.33}$$

We claim that the heat kernel $p_t(x,y)$ of $(\mathbb{R}, \mathbf{g}_{\mathbb{R}}, \mu)$ is given by the explicit formula:

$$p_t(x,y) = \frac{1}{(2\pi \sinh 2t)^{1/2}}\exp\left(\frac{2xye^{-2t} - x^2 - y^2}{1-e^{-4t}} - t\right), \tag{4.34}$$

that is called the *Mehler kernel*. It is a matter of a routine (but hideous) computation to verify that the function (4.34) 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 (4.33) 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 4.15 and (4.34), 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 4.16 *The heat kernel of $\mathbb{H}^3$ is given by the following formula:*

$$p_t(x,y) = \frac{1}{(4\pi t)^{3/2}} \frac{r}{\sinh r} \exp\left(-\frac{r^2}{4t} - t\right). \tag{4.35}$$

The following formulas for $p_t(x,y)$ in $\mathbb{H}^n$ are known: if $n = 2m+1$ then

$$p_t(x,y) = \frac{(-1)^m}{(2\pi)^m (4\pi t)^{1/2}} \left(\frac{1}{\sinh r}\frac{\partial}{\partial r}\right)^m e^{-m^2 t - \frac{r^2}{4t}}, \tag{4.36}$$

which in the case $n = 3$ gives (4.35), and if $n = 2m+2$ then

$$p_t(x,y) = \frac{(-1)^m \sqrt{2}}{(2\pi)^m (4\pi t)^{3/2}} e^{-\frac{(2m+1)^2}{4}t} \left(\frac{1}{\sinh r}\frac{\partial}{\partial r}\right)^m \int_r^\infty \frac{s e^{-\frac{s^2}{4t}} ds}{(\cosh s - \cosh r)^{\frac{1}{2}}}. \tag{4.37}$$

In particular, the heat kernel in $\mathbb{H}^2$ is given by

$$p_t(x,y) = \frac{\sqrt{2}}{(4\pi t)^{3/2}} e^{-\frac{1}{4}t} \int_r^\infty \frac{s e^{-\frac{s^2}{4t}} ds}{(\cosh s - \cosh r)^{\frac{1}{2}}}. \tag{4.38}$$

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 $p_t(x,y) \to 0$ as $x \to \infty$.

We will give here a non-computational proof of (4.35), which to some extend also explains why the heat kernel has this shape.

Proof. By Theorem 4.13, it suffices to prove (4.35) 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 1.14 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:

$$\Delta_{\mathbb{H}^3} = \frac{\partial^2}{\partial r^2} + 2\coth r \frac{\partial}{\partial r} + \frac{1}{\sinh^2 r}\Delta_{\mathbb{S}^2}. \tag{4.39}$$

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 $d\widetilde{\mu} = h^2 d\mu$. 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

$$\Delta h + h = 0, \tag{4.40}$$

which follows from (4.39) 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 (4.40) 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 4.15 that

$$\widetilde{p}_t(x,y) = \frac{e^t p_t(x,y)}{h(x)\,h(y)}. \tag{4.41}$$

Since the area functions of the weighted models $(\mathbb{H}^3, \widetilde{\mu})$ and $\mathbb{R}^3$ are the same, we conclude by Theorem 4.14 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 (4.41), 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 $\mathbb{S}^1$ and Poisson summation formula

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 4.17 *For all $t > 0$ and $x, y \in \mathbb{S}^1$,*

$$p_t(x,y) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{\infty} e^{-k^2 t} \cos k(x-y), \tag{4.42}$$

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 3.18, the heat kernel of a compact manifold M (or a precompact open subset of any manifold) is given by the eigenfunction expansion

$$p_t(x,y) = \sum_{k=1}^{\infty} e^{-\lambda_k t} v_k(x)\, v_k(y), \tag{4.43}$$

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:

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, ..., \frac{\cos mx}{\sqrt{\pi}}, \frac{\sin mx}{\sqrt{\pi}}, ...$$

By (4.43) we obtain

$$\begin{aligned} p_t(x,y) &= \frac{1}{2\pi} + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2t}\cos mx\cos my + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2t}\sin mx\sin my \\ &= \frac{1}{2\pi} + \frac{1}{\pi}\sum_{m=1}^{\infty} e^{-m^2t}\cos m(x-y), \end{aligned}$$

which was to be proved. ■

Proposition 4.18 *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*

$$p_t(x,y) = \sum_{n\in\mathbb{Z}} q_t(x+2\pi n, y). \tag{4.44}$$

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,

$$\int_{\mathbb{S}^1} \widetilde{q}_t(x,y)\,dx = 1. \tag{4.45}$$

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 (4.42) that

$$\int_{\mathbb{S}^1} p_t(x,y)\,dx = 1,$$

which together with (4.45) implies the identity $\widetilde{q}_t(x,y) = p_t(x,y)$. ∎

Corollary 4.19 *(*The Poisson summation formula*) For all $t > 0$, we have the following identity*

$$\sum_{k\in\mathbb{Z}} e^{-k^2t} = \sqrt{\frac{\pi}{t}} \sum_{n\in\mathbb{Z}} \exp\left(-\frac{\pi^2n^2}{t}\right). \tag{4.46}$$

Proof. Rewrite (4.42) as follows

$$p_t(x,y) = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} e^{-k^2t} \cos k(x-y). \tag{4.47}$$

In particular, for $x = y = 0$ we obtain

$$p_t(0,0) = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} e^{-k^2t}. \tag{4.48}$$

From (4.44) at $x = y = 0$, we obtain

$$p_t(0,0) = \sum_{n\in\mathbb{Z}} \frac{1}{(4\pi t)^{1/2}} \exp\left(-\frac{\pi^2n^2}{t}\right).$$

Comparing the above two lines, we obtain (4.46). ∎

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 3.10 and 3.12.

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 3.12 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 1.15 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

$$V(x_0,r) \leq \exp\left(Cr^2\right). \tag{5.8}$$

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

$$\int^\infty \frac{rdr}{f(r)} = \infty. \tag{5.11}$$

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*

$$|u(t,x)| \leq C \exp\left(C\,|x|^2\right) \quad \text{for all } t \in I,\ x \in \mathbb{R}^n, \tag{5.12}$$

then $u \equiv 0$. Moreover, the same is true if u satisfies instead of (5.12) the condition

$$|u(t,x)| \leq C \exp\left(f\left(|x|\right)\right) \quad \text{for all } t \in I,\ x \in \mathbb{R}^n, \tag{5.13}$$

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

$$\int_0^T \int_{B(x_0,R)} u^2(t,x) d\mu(x) \leq S^2 T V(x_0, R) = \exp(f(R)),$$

where

$$f(r) := \ln\left(S^2 T V(x_0, r)\right).$$

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 $R > 0$,*

$$\int_a^b \int_{B_R} u^2(x,t)\, d\mu(x) dt \leq \exp(f(R)),$$

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:

$$\int_{B_R} u^2(a,\cdot) d\mu \leq \int_{B_{4R}} u^2(b,\cdot) d\mu + \frac{4}{R^2}. \tag{5.15}$$

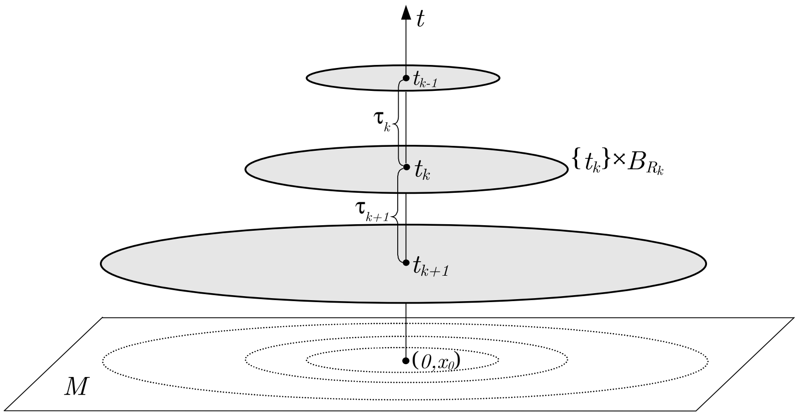
Let us first show how this Claim allows to prove that any solution u to (5.9), satisfying (5.10), is identical 0. Extend $u(t,x)$ to $t = 0$ by setting $u(0,x) = 0$ so that u is continuous in $[0,T) \times M$. Fix $R > 0$ and $t \in (0,T)$. For any non-negative integer k, set

$$R_k = 4^k R$$

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 $t_k = t_{k-1} - \tau_k$.



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 $|\nabla \rho| \leq 1$, which implies, for any $t \neq s$,

$$|\nabla \xi(t,x)| \leq \frac{\rho(x)}{2(t-s)}.$$

Since

$$\frac{\partial \xi}{\partial t} = -\frac{\rho^2(x)}{4(t-s)^2},$$

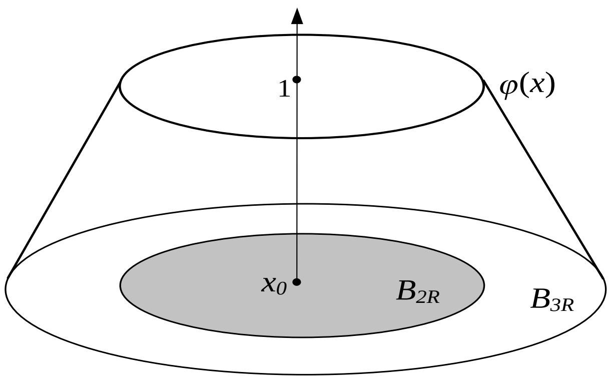
we obtain

$$\frac{\partial \xi}{\partial t} + |\nabla \xi|^2 \leq 0. \tag{5.20}$$

For a given $R > 0$, define a function $\varphi(x)$ by

$$\varphi(x) = \min\left(\left(3 - \tfrac{d(x,x_0)}{R}\right)_+, 1\right)$$

Obviously, we have $0 \leq \varphi \leq 1$ on M, $\varphi \equiv 1$ in B_{2R}, and $\varphi \equiv 0$ 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 $|\nabla \varphi| \leq 1/R$. By the completeness of M, all the balls in M are relatively compact sets, which implies $\varphi \in Lip_0(M)$.



Function $\varphi(x)$

Consider the function $u\varphi^2 e^\xi$ 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

$$u\varphi^2 e^\xi \in W_c^1(M).$$

Multiplying the heat equation

$$\frac{\partial u}{\partial t} = \Delta_\mu u$$

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

$$\begin{aligned}
-\langle \nabla u, \nabla(u\varphi^2 e^\xi)\rangle &= -|\nabla u|^2 \varphi^2 e^\xi - \langle \nabla u, \nabla \xi\rangle u\varphi^2 e^\xi - 2\langle \nabla u, \nabla\varphi\rangle u\varphi e^\xi \\
&\leq -|\nabla u|^2 \varphi^2 e^\xi + |\nabla u| |\nabla \xi| |u| \varphi^2 e^\xi \\
&\quad + \left(\frac{1}{2}|\nabla u|^2 \varphi^2 + 2|\nabla\varphi|^2 u^2\right) e^\xi \\
&= \left(-\frac{1}{2}|\nabla u|^2 + |\nabla u| |\nabla \xi| |u|\right)\varphi^2 e^\xi + 2|\nabla \varphi|^2 u^2 e^\xi.
\end{aligned}$$

Combining with (5.21), (5.22), and using (5.20), we obtain

$$\begin{aligned}\left[\int_M u^2\varphi^2 e^{\xi} d\mu\right]_b^a &= \int_b^a\int_M \frac{\partial \xi}{\partial t} u^2\varphi^2 e^{\xi}\, d\mu dt + 2\int_b^a\int_M (\Delta_\mu u)\, u\varphi^2 e^{\xi} d\mu dt \\ &\leq \int_b^a\int_M \left(-|\nabla\xi|^2 u^2 - |\nabla u|^2 + 2|\nabla u||\nabla\xi||u|\right)\varphi^2 e^{\xi} d\mu dt \\ &\quad +4\int_b^a\int_M |\nabla\varphi|^2 u^2 e^{\xi} d\mu dt \\ &= -\int_b^a\int_M (|\nabla\xi||u| - |\nabla u|)^2 \varphi^2 e^{\xi} d\mu dt \\ &\quad +4\int_b^a\int_M |\nabla\varphi|^2 u^2 e^{\xi} d\mu dt\end{aligned}$$

whence

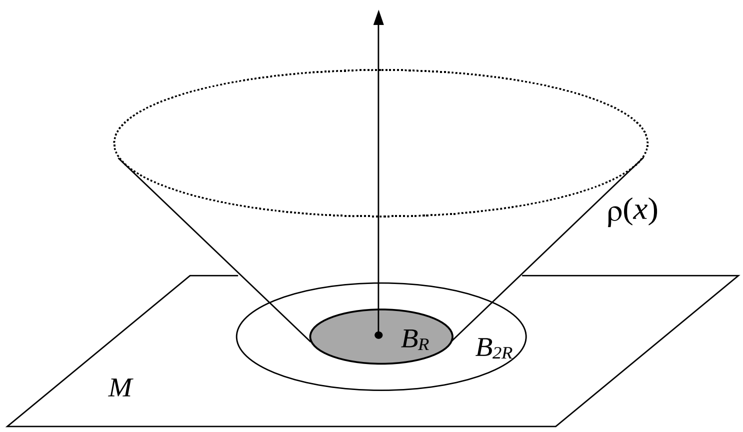
$$\left[\int_M u^2\varphi^2 e^{\xi} d\mu\right]_b^a \leq 4\int_b^a\int_M |\nabla\varphi|^2 u^2 e^{\xi} d\mu dt. \tag{5.23}$$

Using the properties of function $\varphi(x)$, in particular, $|\nabla\varphi| \leq 1/R$, we obtain from (5.23)

$$\int_{B_R} u^2(a,\cdot)e^{\xi(a,\cdot)}d\mu \leq \int_{B_{4R}} u^2(b,\cdot)e^{\xi(b,\cdot)}d\mu + \frac{4}{R^2}\int_b^a\int_{B_{4R}\setminus B_{2R}} u^2 e^{\xi} d\mu dt. \tag{5.24}$$

Let us now specify $\rho(x)$ and s. Set $\rho(x)$ to be the distance function from the ball B_R, that is,

$$\rho(x) = (d(x,x_0) - R)_+ .$$



Function $\rho(x)$.

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 $x \in B_{4R} \backslash B_{2R}$ then $\rho(x) \ge R$, which together with (5.25) implies that

$$\xi(t,x) \le -\frac{R^2}{8\,(a-b)} \quad \text{in} \quad [b,a] \times B_{4R} \backslash B_{2R}.$$

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), (4.34), (4.35), 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

$$|P_t f| = |P_{t-t_0} P_{t_0} f| \leq P_{t-t_0} g.$$

Assuming that Theorem 6.1 has been already proved for function g, we obtain

$$\begin{aligned}
\int_M (P_t f)^2 e^{\xi(t,\cdot)} d\mu &\leq \int_M (P_{t-t_0} g)^2 e^{\xi(t,\cdot)} d\mu \\
&\leq e^{-2\lambda_1(t-t_0)} \int_M g^2 e^{\xi(t_0,\cdot)} d\mu \\
&= e^{-2\lambda_1(t-t_0)} \int_M (P_{t_0} f)^2 e^{\xi(t_0,\cdot)} d\mu.
\end{aligned}$$

Hence, we can assume in the sequel that $f \geq 0$. 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),

$$\begin{aligned}\frac{dJ_\Omega}{dt} &= \left(\frac{du}{dt}, ue^{\xi}\right) + \left(u, \frac{d\left(ue^{\xi}\right)}{dt}\right) \\ &= 2\left(\frac{du}{dt}, ue^{\xi}\right) + \left(u^2, \frac{de^{\xi}}{dt}\right) \\ &= 2\left(\Delta u, ue^{\xi}\right) + \left(u^2, \frac{\partial \xi}{\partial t} e^{\xi}\right). \end{aligned} \tag{6.7}$$

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

$$\begin{aligned}\frac{dJ_\Omega}{dt} &\le -2\int_\Omega \left(|\nabla u|^2 e^{\xi} + ue^{\xi}\langle \nabla u, \nabla \xi\rangle + \frac{1}{4}u^2 |\nabla \xi|^2 e^{\xi}\right) d\mu \\ &= -2\int_\Omega \left(\nabla u + \frac{1}{2} u \nabla \xi\right)^2 e^{\xi} d\mu, \end{aligned} \tag{6.8}$$

whence $\frac{dJ_\Omega}{dt} \le 0$. 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

$$\begin{aligned}\int_\Omega \left(\nabla u + \frac{1}{2}u\nabla\xi\right)^2 e^{\xi} d\mu &= \int_\Omega |\nabla(ue^{\xi/2})|^2 d\mu \\ &\ge \lambda_1(\Omega) \int_\Omega |ue^{\xi/2}|^2 d\mu \\ &= \lambda_1(\Omega) J_\Omega(t). \end{aligned} \tag{6.9}$$

Hence, (6.8) yields

$$\frac{dJ_\Omega}{dt} \le -2\lambda_1(\Omega) J_\Omega(t),$$

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,

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

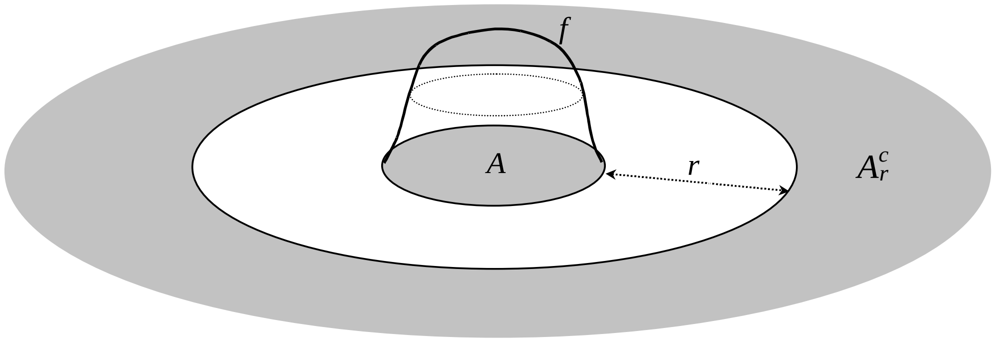
Write also $A_r^c = (A_r)^c = M \setminus A_r$.

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

$$\int_{A_r^c} (P_t f)^2 \, d\mu \le \|f\|_2^2 \exp\left(-\frac{r^2}{2t} - 2\lambda t\right). \tag{6.11}$$



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

$$J(t)\geq\int_{A_r^c}(P_tf)^2\,d\mu. \tag{6.13}$$

On the other hand, using the fact that $\xi(0,x)\leq 0$ for all x and

$$\xi(0,x)\leq-\frac{r^2}{2s}\ \text{ for all } x\in A,$$

we obtain

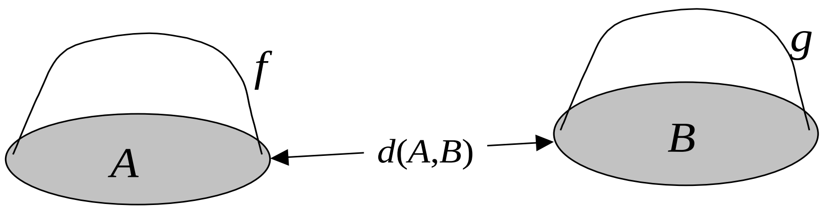
$$J(0)\leq\int_{A^c}f^2d\mu+\exp\left(-\frac{r^2}{2s}\right)\int_A f^2d\mu. \tag{6.14}$$

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 $\int_{A^c}f^2d\mu=0$. ■

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 $t>0$,*

$$|(P_tf,g)|\leq\|f\|_2\|g\|_2\exp\left(-\frac{d^2(A,B)}{4t}-\lambda t\right). \tag{6.15}$$



Sets A and B

Proof. Set $r=d(A,B)$. Then $B\subset A_r^c$ and by (6.11)

$$\int_B(P_tf)^2\,d\mu\leq\|f\|_2^2\exp\left(-\frac{r^2}{2t}-2\lambda t\right).$$

Applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}|(P_tf,g)| &\leq \left(\int_B(P_tf)^2\,d\mu\right)^{1/2}\|g\|_2\\ &\leq \|f\|_2\|g\|_2\exp\left(-\frac{r^2}{4t}-\lambda t\right),\end{aligned}$$

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 (3.48), 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

$$\begin{aligned} J_{lm} &= \sum_{i=1}^{\infty} e^{-t\lambda_i} a_i^{(l)} a_i^{(m)} \\ &= \frac{\mu(A_l)\mu(A_m)}{\mu(M)} + \sum_{i=k}^{\infty} e^{-\lambda_i t} a_i^{(l)} a_i^{(m)} + \sum_{i=2}^{k-1} e^{-\lambda_i t} a_i^{(l)} a_i^{(m)} \\ &\geq \frac{\mu(A_l)\mu(A_m)}{\mu(M)} - e^{-\lambda_k t}\sqrt{\mu(A_l)\mu(A_m)} \\ &\quad + \sum_{i=2}^{k-1} e^{-\lambda_i t} a_i^{(l)} a_i^{(m)}. \end{aligned} \tag{6.20}$$

On the other hand, by (6.16)

$$J_{lm} \leq \sqrt{\mu(A_l)\mu(A_m)} e^{-\frac{\delta^2}{4t}}. \tag{6.21}$$

Therefore, we can further argue as in the case $k = 2$ provided the term in (6.20) can be discarded, which the case when

$$\sum_{i=2}^{k-1} e^{-\lambda_i t} a_i^{(l)} a_i^{(m)} \geq 0. \tag{6.22}$$

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

$$\langle u, v \rangle_t := \sum_{i=2}^{k-1} e^{-\lambda_i t} u_i v_i \tag{6.23}$$

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

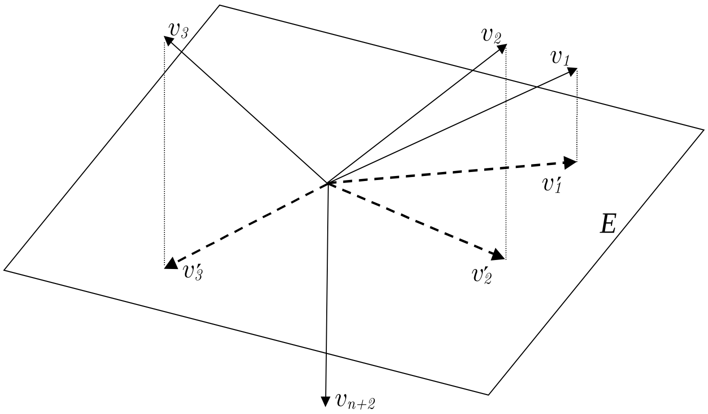
$$v := -e_1 - e_2 - ... - e_n.$$

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.

Lemma 6.5 is easily proved by induction in n. The inductive basis for $n = 1$ is trivial. The inductive step is shown on the diagram. 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.



The vectors v_i' are the orthognal projections 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

$$\langle v_i, v_j\rangle = \langle v_i', v_j'\rangle + \varepsilon_i\varepsilon_j\,|v_{n+2}|^2\,.$$

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. ■