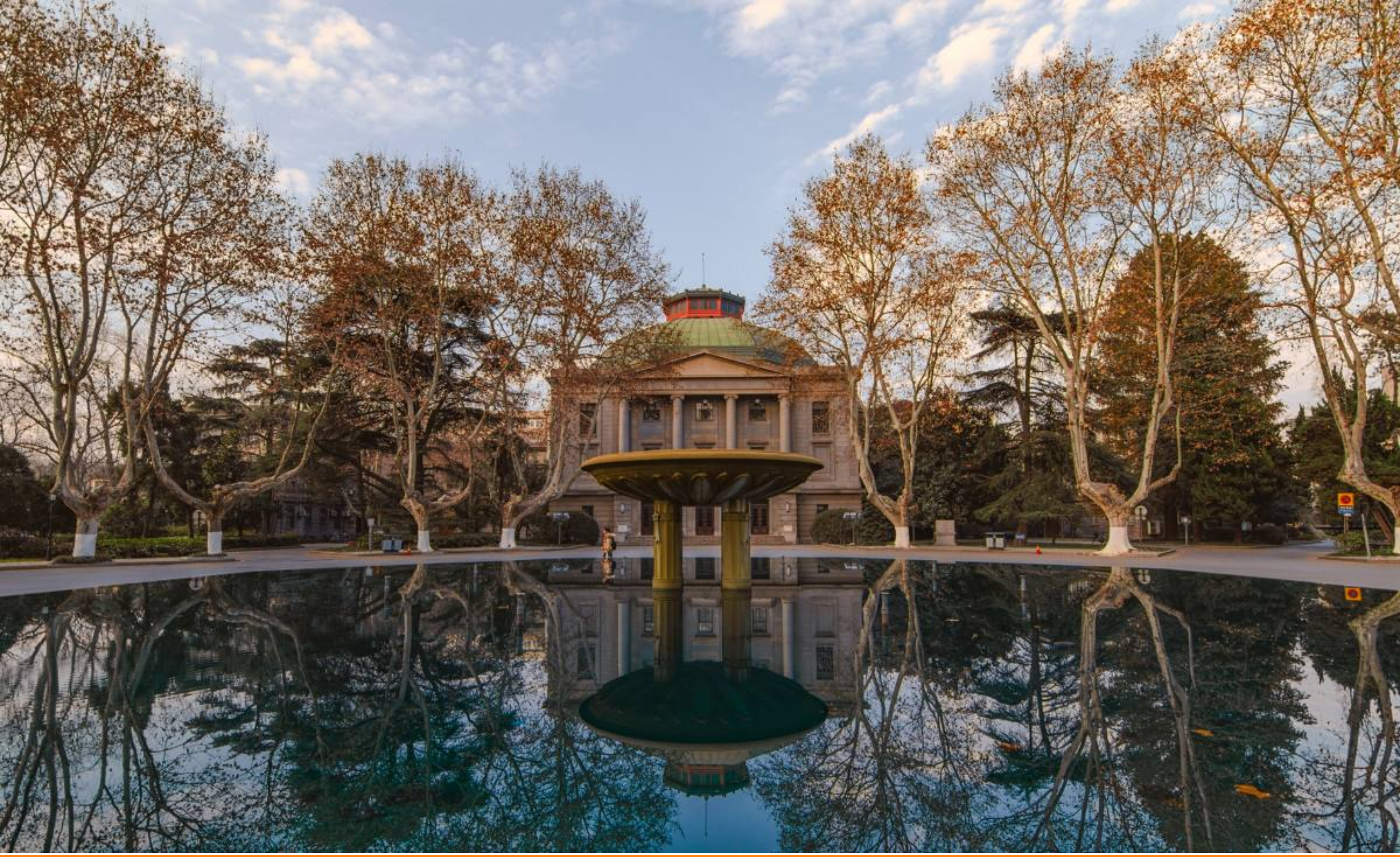

Differential manifolds

An introduction to Riemannian geometry and Ricci flow

Author: Yi Li

Institute: School of Mathematics and Shing-Tung Yau Center, Southeast University

Date: March 1, 2021

Version: 2.3;

Email: yilicms@seu.edu.cn; yilicms@gmail.com; yilicms@163.com

Rather less, but better. – Carl Friedrich Gauss

Main References

- Chow, Bennett; Lu, Peng; Ni, Lei. *Hamilton's Ricci flow*. GSM **77**, AMS/Science Press, 2006
- Lee, John M. *Riemannian manifolds: An introduction to curvature*. GTM **176**, Springer-Verlag, 1997.
- Petersen, Peter. *Riemannian geometry*. GTM **171**, Springer-Verlag, New York, 1998.
- Schoen, Richard; Yau, Shing-Tung. *Lectures on differential geometry*. Conference Proceedings and LEcture Notes in Geometry and Topology, **I**, International Press, Cambridge, MA, 1994.

Contents

Chapter 1 Differential manifolds

Introduction

- *Manifolds*
- *Tensors and forms*
- *Integration on manifolds*

1.1 Manifolds

Introduction

- *Differential manifolds*

The function $r^i : \mathbf{R}^m \to \mathbf{R}$ defined by

$$r^i(\boldsymbol{x}) := x^i,$$

where $\boldsymbol{x} = (x^1, \cdots, x^m) \in \mathbf{R}^m$, is called the ***i*-th canonical coordinate function** on $\mathbf{R}^m$. The canonical coordinate function on $\mathbf{R}$ will be denoted by r. If $f : X \to \mathbf{R}^m$ is a function on some set X, then we let

$$f^i := r^i \circ f,$$

where f^i is called the ***i*-th component function of *f***. If $f : \mathbf{R} \to \mathbf{R}$ and $x \in \mathbf{R}$, then we denote the **derivative of f at x** by

$$\left.\frac{d}{dr}\right|_x (f) = \left.\frac{df}{dr}\right|_x = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}.$$

If $f : \mathbf{R}^m \to \mathbf{R}$, $1 \le i \le m$, and $\boldsymbol{x} = (x^1, \cdots, x^m) \in \mathbf{R}^m$, then we denote the **partial derivative of f with respect to r^i at $\boldsymbol{x}$** by

$$\left.\frac{\partial}{\partial r^i}\right|_{\boldsymbol{x}} (f) = \left.\frac{\partial f}{\partial r^i}\right|_{\boldsymbol{x}} = \lim_{h\to 0} \frac{f(x^1, \cdots, x^{i-1}, x^i + h, x^{i+1}, \cdots, x^m) - f(\boldsymbol{x})}{h}.$$

If $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_m)$ is a m-tuple of nonnegative integers, then we set

$$|\boldsymbol{\alpha}| = \sum_{1\le i\le m} \alpha_i, \quad \boldsymbol{\alpha}! = \prod_{1\le i\le m} \alpha_i!, \quad \frac{\partial^{\boldsymbol{\alpha}}}{\partial \boldsymbol{r}^{\boldsymbol{\alpha}}} = \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial(r^1)^{\alpha_1} \cdots \partial(r^m)^{\alpha_m}}.$$

If $\boldsymbol{x} \in \mathbf{R}^m$, then $\mathbf{B}^m_r(\boldsymbol{x}) = \mathbf{B}^m(\boldsymbol{x}, r)$ will denote the **open ball** of radius r about $\boldsymbol{x}$. Write $\mathbf{B}^m_r := \mathbf{B}^m(\mathbf{0}, r)$. $\mathbf{C}_r$ will denote the **open cube** with sides of length $2r$ about the origin in $\mathbf{R}^m$. That is

$$\mathbf{C}_r := \{(x^1, \cdots, x^m) \in \mathbf{R}^m : |x^i| < r \text{ for all } 1 \le i \le m\}.$$

1.1.1 Differential manifolds

Let $U \subset \mathbf{R}^m$ be open and let $f : U \to \mathbf{R}$. We say that f is **differentiable of class C^k on U**, where $k \in \mathbf{N} \cup \{0, \infty\} \cup \{\omega\}$,

(i) ($k \in \mathbf{N} \cup \{0\}$) if the partial derivatives $\partial^{\boldsymbol{\alpha}} f / \partial \boldsymbol{r}^{\boldsymbol{\alpha}}$ exist and are continuous on U for $|\boldsymbol{\alpha}| \leq k$;

(ii) ($k = \infty$) if f is C^k for all $k \geq 0$;

(iii) ($k = \omega$) if f is locally given by convergent power series.

If $f : U \to \mathbf{R}^n$, then f is **differentiable of class C^k** if each of the component functions $f^i = r^i \circ f$ is C^k.

Definition 1.1. (Topological manifolds)

A **topological manifold $\mathcal{M}$ of dimension m** *is a Hausdorff space for which each point has a neighborhood homeomorphic to an open subset of $\mathbf{R}^m$. If φ is a homeomorphism of a connected open set $\mathcal{U} \subset \mathcal{M}$ onto an open subset $U := \varphi(\mathcal{U}) \subset \mathbf{R}^m$, φ is called a* **coordinate map**, *the functions $x^i := r^i \circ \varphi$ are called the* **coordinate functions**, *and the pair $(\mathcal{U}, \varphi)$ or $(\mathcal{U}, x^1, \cdots, x^m)$ is called a* **coordinate system**.

(a) A coordinate system $(\mathcal{U}, \varphi)$ is called a **cubic coordinate system** *if $\varphi(\mathcal{U})$ is an open cube about the origin in $\mathbf{R}^m$.*

(b) If $p \in \mathcal{U}$ and $\varphi(p) = 0$, then the coordinate system is said to be **centered at** *p.* ♣

Definition 1.2. (Differentiable structure)

A **differentiable structure $\mathscr{F}$ of class C^k**, *where $k \in \mathbf{N} \cup \{\infty\}$, on a topological manifold $\mathcal{M}$ of dimension m, is a collection of coordinate systems $(\mathcal{U}_\alpha, \varphi_\alpha)_{\alpha \in A}$ satisfying*

(a) $M = \cup_{\alpha \in A} \mathcal{U}_\alpha$;

(b) $\varphi_\alpha \circ \varphi_\beta^{-1}$ is C^k for all $\alpha, \beta \in A$;

(c) the collection $\mathscr{F}$ is maximal with respect to (b), that is, if $(\mathcal{U}, \varphi)$ is a coordinate system such that $\varphi \circ \varphi_\alpha^{-1}$ and $\varphi_\alpha \circ \varphi^{-1}$ are C^k for all $\alpha \in A$, then $(\mathcal{U}, \varphi) \in \mathscr{F}$. ♣

If $\mathscr{F}_0 := \{(U_\alpha, \varphi_\alpha) : \alpha \in A\}$ is any collection of coordinate systems satisfying properties (a) and (b), then there is a unique differentiable structure $\mathscr{F}$ containing $\mathscr{F}_0$. Namely, let

$$\mathscr{F} := \left\{ (\mathcal{U}, \varphi) : \varphi \circ \varphi_\alpha^{-1} \text{ and } \varphi_\alpha \circ \varphi^{-1} \text{ are } C^k \text{ for all } \varphi_\alpha \in \mathscr{F}_0 \right\}. \tag{1.1.1}$$

Hence, to find a differentiable structure on $\mathcal{M}^m$, we need only to find such a $\mathscr{F}_0$. Without loss of generality, we may say a differentiable structure is a collection of coordinate systems satisfying (a) and (b).

Replacing C^k by C^ω in Definition 1.2, we can define a **differentiable structure of class C^ω**. For a **complex analytic structure** on a $2n$-dimensional topological manifold, one requires that the coordinate systems have range in $\mathbf{C}^n$ and overlap holomorphically.

Definition 1.3. (Differentiable manifolds)

A m**-dimensional differentiable manifold of class** C^k *is a pair* $(\mathcal{M}, \mathscr{F})$ *consisting of an* m*-dimensional, second countable, topological manifold* $\mathcal{M}$ *together with a differentiable structure* $\mathscr{F}$ *of class* C^k*.* ♣

Unless we indicate otherwise, all manifolds are smooth manifolds or differentiable manifold of class C^∞. If $\mathcal{X}$ is a set, by a **manifold structure on** $\mathcal{X}$ we shall mean a choice of both a second countable topological manifold for $\mathcal{X}$ and a differentiable structure.

Example 1.1. (Some classical examples of differentiable manifolds)

(a) $\mathbf{R}^m$. *The standard differentiable structure on* $\mathbf{R}^m$ *is* $(\mathbf{R}^m, \mathbf{1})$*, where* $\mathbf{1}: \mathbf{R}^m \to \mathbf{R}^m$ *is the identity map.*

(b) *Finite dimensional real vector spaces. Let* V *be a finite dimensional real vector space. Then* V *has a nature manifold structure. If* $\mathbf{e} := (e_i)_{1 \le i \le m}$ *is a basis of* V*, the dual basis* $\mathbf{e}^* := (e_i^*)_{1 \le i \le m}$ *gives a global coordinate system on* V*:*

$$\varphi_{\mathbf{e}} : V \longrightarrow \mathbf{R}^m, \quad \varphi_{\mathbf{e}}(v) := (e_1^*(v), \cdots, e_m^*(v)), \ v \in V.$$

If $\tilde{\mathbf{e}} := (\tilde{e}_i)_{1 \le i \le m}$ *is another basis of* V *with*

$$\tilde{e}_i^* = \sum_{1 \le j \le m} a_{ij} e_j^*, \quad \boldsymbol{a} := (a_{ij})_{1 \le i,j \le m} \in \mathbf{GL}(m, \mathbf{R}),$$

we have $\varphi_{\tilde{\mathbf{e}}} = \boldsymbol{a}\varphi_{\mathbf{e}}$*. Consequently, this differentiable structure is independent of the choice of basis.*

(c) $\mathbf{C}^n$. *As a real* $2n$*-dimensional vector space,* $\mathbf{C}^n$ *has a natural manifold structure. If* $(e_i)_{1 \le i \le n}$ *is the canonical complex basis, then*

$$e_1, \cdots, e_n, \sqrt{-1}e_1, \cdots, \sqrt{-1}e_n$$

is a real basis for $\mathbf{C}^n$*, and its dual basis is the canonical global coordinate system on* $\mathbf{C}^n$*.*

(d) *The* $\boldsymbol{m}$**-sphere** *is the set*

$$\mathbf{S}^m := \left\{ \boldsymbol{a} = (a^1, \cdots, a^{m+1}) \in \mathbf{R}^{m+1} : \sum_{1 \le i \le m+1} (a^i)^2 = 1 \right\}.$$

Let $N := (0, \cdots, 0, 1)$ *and* $S := (0, \cdots, 0, -1)$*. Then the standard differentiable structure on* $\mathbf{S}^m$ *is*

$$(\mathbf{S}^m \setminus N, p_N), \quad (\mathbf{S}^m \setminus S, p_S),$$

where p_N *and* p_S *are stereographic projections from* N *and* S *respectively.*

(e) **(Open submanifolds)** *An open subset* $\mathcal{U}$ *of a differentiable manifold* $(\mathcal{M}, \mathscr{F})$ *is itself a differentiable manifold with differentiable structure*

$$\mathscr{F}_{\mathcal{U}} := \{(\mathcal{U}_\alpha \cap \mathcal{U}, \varphi_{\alpha|\mathcal{U}_\alpha \cap \mathcal{U}}) : (\mathcal{U}_\alpha, \varphi_\alpha) \in \mathscr{F}\}.$$

(f) **(Product manifolds)** *Let* $(\mathcal{M}_1, \mathscr{F}_1)$ *and* $(\mathcal{M}_2, \mathscr{F}_2)$ *be differentiable manifolds of* *dimensions m_1 and m_2 respectively. Define*

$$\varphi_\alpha \times \psi_\beta : \mathcal{U}_\alpha \times \mathcal{V}_\beta \longmapsto \mathbf{R}^{m_1} \times \mathbf{R}^{m_2}, \quad (x, y) \longmapsto (\varphi_\alpha(x), \psi_\beta(y)),$$

where $\mathscr{F}_1 = (\mathcal{U}_\alpha, \varphi_\alpha)_{\alpha \in A}$ and $\mathscr{F}_2 = (\mathcal{V}_\beta, \psi_\beta)_{\beta \in B}$. Then

$$\mathscr{F} = (\mathcal{U}_\alpha \times \mathcal{V}_\beta, \varphi_\alpha \times \psi_\beta)_{(\alpha,\beta) \in A \times B}$$

is a differentiable structure of $\mathcal{M}_1 \times \mathcal{M}_2$. amd $\dim(\mathcal{M}_1 \times \mathcal{M}_2) = m_1 + m_2$.

(g) *Let $\mathbf{T}^m = \mathbf{S}^1 \times \cdots \times \mathbf{S}^1$ (m times). It is called the* **m-dimensional torus**.

(h) *The* **general linear group** $\mathbf{GL}(m, \mathbf{R})$ *is the set of all $m \times m$ nonsingular real matrices. Define*

$$\mathbf{GL}(m, \mathbf{R}) \longrightarrow \mathbf{R}^{m^2}, \quad A = (a_{ij})_{1 \le i,j \le m} \longmapsto (a_{11}, \cdots, a_{1n}, \cdots, a_{m1}, \cdots, a_{mm}).$$

Then the determinant can be considered as a function of $\mathbf{R}^{m^2}$:

$$\det : \mathbf{R}^{m^2} \longrightarrow \mathbf{R}, \quad (a_{11}, \cdots, a_{1n}, \cdots, a_{m1}, \cdots, a_{mm}) \longmapsto \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}.$$

It is clear that $\det$ is a continuous function and $\mathbf{Ker}(\det)$ is a closed subset of $\mathbf{R}^{m^2}$. Consequently,

$$\mathbf{GL}(m, \mathbf{R}) = \mathbf{R}^{m^2} \setminus \mathbf{Ker}(\det)$$

is an open subset and then a differentiable manifold.

(i) *Let $T(m, n)$ be the space of all $m \times n$ real matrices. Then $T(m, n)$ can be regarded as $\mathbf{R}^{mn}$ and therefore is a real analytic (C^ω) manifold. Let $T(m, n; k)$ denote the space of all $m \times n$ real matrices of rank k (where $0 < k \le \min(m, n)$) with the induced topology of $T(m, n)$. Then $T(m, n; k)$ is a real analytic manifold of dimension $k(m + n - k)$. In fact, let $X_0 \in T(m, n)$. If $\operatorname{rank} X_0 \ge k$, there are permutation matrices P and Q such that*

$$PX_0Q = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix}$$

where A_0 is a nonsingular $k \times k$ matrix. There is an $\epsilon > 0$ (depending on A_0) such that if $||A - A_0||_{matrix} < \epsilon$, then A is nonsingular. ♠

Definition 1.4. (Smooth functions)

Let $\mathcal{U}$ be an open subset of a manifold $\mathcal{M}$ of dimension m. We say that $f : \mathcal{U} \to \mathbf{R}$ is a **C^∞-function on $\mathcal{U}$** *(denoted $f \in C^\infty(\mathcal{U})$) if $f \circ \varphi^{-1}$ is smooth for each coordinate map φ on $\mathcal{M}^m$.*

A continuous map $\psi : \mathcal{M} \to \mathcal{N}$ between two manifolds of dimensions m and n respectively, is said to be **of class C^∞**, *denoted $\psi \in C^\infty(\mathcal{M}, \mathcal{N})$, if $g \circ \psi$ is a smooth function on ψ^{-1}(domain of g) for all smooth functions g defined on open sets in $\mathcal{N}$.* ♣

Note that the continuous map $\psi : \mathcal{M} \to \mathcal{N}$ is smooth if and only if $\varphi \circ \psi \circ \tau^{-1}$ is smooth for each coordinate mapτ on $\mathcal{M}$ and φ on $\mathcal{N}$, $\dim \mathcal{M} = m$ and $\dim \mathcal{N} = n$.

$$\begin{tikzcd} \mathcal{M} \arrow[r, "\psi"] \arrow[d, "\tau"'] & \mathcal{N} \arrow[d, "\varphi"] \\ \mathbf{R}^m \arrow[r, "\varphi\circ\psi\circ\tau^{-1}"'] & \mathbf{R}^n \end{tikzcd}$$

Clearly that the composition of two smooth maps is again smooth. Observe that a mapping $\psi : \mathcal{M} \to \mathcal{N}$ is smooth if and only if for each $x \in \mathcal{M}$ there exists an open neighborhood $\mathcal{U}$ of x such that $\psi_{|\mathcal{U}}$ is smooth.

1.1.2 Partition of unity

A collection $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ of subsets of a topological space $\mathcal{X}$ is a **cover** of a set $\mathcal{W} \subset \mathcal{X}$ if $\mathcal{W} \subset \cup_{\alpha\in A}\mathcal{U}_\alpha$.

(i) It is an **open cover** if each $\mathcal{U}_\alpha$ is open.

(ii) A subcollection of the cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ which stills covers is called a **subcover**.

(iii) A **refinement** $\{\mathcal{V}_\beta\}_{\beta\in B}$ of the cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ is a cover such that for each β there is an $\alpha = \alpha(\beta)$ such that $\mathcal{V}_\beta \subset \mathcal{U}_\alpha$.

A collection $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ of subsets of $\mathcal{X}$ is **locally finite** if whenever $x \in \mathcal{X}$ there exists a neighborhood $\mathcal{U}_x$ of x such that $\mathcal{U}_x \cap \mathcal{U}_\alpha \neq \emptyset$ for only finitely many α.

A topological space is **paracompact** if every open cover has an open locally finite refinement.

Definition 1.5. (Partition of unity)

A **partition of unity on $\mathcal{X}$** *is a collection $\{\varphi_i\}_{i\in I}$ of smooth functions on $\mathcal{X}$ such that*

(a) The collection of supports $\{\mathrm{supp}(\varphi_i)\}_{i\in I}$ is locally finite,

(b) $0 \le \varphi_i \le 1$ on $\mathcal{X}$ for all $i \in I$, and

(c) $\sum_{i\in I}\varphi_i = 1$ on $\mathcal{X}$.

A partition of unity $\{\varphi_i\}_{i\in I}$ is **subordinate** *to the cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ if for each i there exists an α such that $\mathrm{supp}(\varphi_i) \subset \mathcal{U}_\alpha$. We say that it is subordinate to the cover $\{\mathcal{U}_i\}_{i\in I}$* **with the same index set as the partition of unity** *if $\mathrm{supp}(\varphi_i) \subset \mathcal{U}_i$ for each $i \in I$.* ♣

Proposition 1.1

Let $\mathcal{X}$ be a topological space which is locally compact (each point has at least one compact neighborhood), Hausdorff, and second countable. Then $\mathcal{X}$ is paracompact. In fact, each open cover of $\mathcal{X}$ has a countable, locally finite refinement consisting of open sets with compact closures. ♡

Proof. (1) There exists a sequence $\{\mathcal{G}_i\}_{i\in\mathbf{N}}$ of open sets such that

$$\overline{\mathcal{G}_i} \text{ is compact}, \quad \overline{\mathcal{G}_i} \subset \mathcal{G}_{i+1}, \quad X = \bigcup_{i\in\mathbf{N}} \mathcal{G}_i.$$

Lindelöf's theorem[1] says that any open cover of a second countable topological space has a countable subcover. Since open sets with compact closure consists of an open cover of $\mathcal{X}$, it follows that there is a countable basis $\{\mathcal{U}_i\}_{i\in\mathbf{N}}$ of the topology of $\mathcal{X}$, where $\mathcal{U}_i$ is an open set with compact closure. Let $\mathcal{G}_1 := \mathcal{U}_1$. Assume that $\mathcal{G}_k = \mathcal{U}_1 \cup \cdots \cup \mathcal{U}_{j_k}$. Let j_{k+1} be the smallest positive integer greater then j_k such that

$$\overline{\mathcal{G}_k} \subset \bigcup_{1\le i\le j_{k+1}} \mathcal{U}_i.$$

Define

$$\mathcal{G}_{k+1} := \bigcup_{1\le i\le j_{k+1}} \mathcal{U}_i.$$

Then we get a countable sequence $\{\mathcal{G}_i\}_{i\in\mathbf{N}}$ satisfying that $\overline{\mathcal{G}_i}$ is compact, $\overline{\mathcal{G}_i} \subset \mathcal{G}_{i+1}$, and

$$\mathcal{X} \subset \bigcup_{i\in\mathbf{N}} \mathcal{U}_i \subset \bigcup_{i\in\mathbf{N}} \mathcal{G}_i \subset \bigcup_{i\in\mathbf{N}} \overline{\mathcal{G}_i} \subset \mathcal{X}.$$

Therefore $X = \cup_{i\in\mathbf{N}}\mathcal{G}_i$.

(2) Let $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ be an arbitrary open cover. The set $\overline{\mathcal{G}_i} \setminus \mathcal{G}_{i-1}$ is compact and contained in the open set $\mathcal{G}_{i+1} \setminus \overline{\mathcal{G}_{i-2}}$. For each $i \ge 3$, we choose a finite subcover of the open cover $\{\mathcal{U}_\alpha \cap (\mathcal{G}_{i+1} \setminus \overline{\mathcal{G}_{i-2}})\}_{\alpha\in A}$ of $\overline{\mathcal{G}_i} \setminus \mathcal{G}_{i-1}$ and choose a finite subcover of the open cover $\{\mathcal{U}_\alpha \cap \mathcal{G}_3\}_{\alpha\in A}$ of the $\overline{\mathcal{G}_2}$. This collection of open sets is countable, locally finite refinement of the open cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$. In fact, let $\mathcal{W}_i := \overline{\mathcal{G}_i} \setminus \mathcal{G}_{i-1}$. Since $\mathcal{X}$ is locally compact, it follows that there exists a finite subcover $(\mathcal{U}_{ij})_{1\le j\le m_i}$ of the open cover $\{\mathcal{U}_\alpha \cap (\mathcal{G}_{i+1} \setminus \overline{\mathcal{G}_{i-2}})\}_{\alpha\in A}$ of W_i. Set

$$\mathcal{V}_{ij} := \mathcal{U}_{ij} \cap (\mathcal{G}_{i+1} \setminus \overline{\mathcal{G}_{i-2}}), \quad 1 \le j \le m_i.$$

Similarly, we can define $\mathcal{V}_{ij}$ for $i = 1, 2$. Hence $(\mathcal{V}_{ij})_{1\le j\le m_i}$ is an open subcover of $\mathcal{W}_i$ and

$$\bigcup_{i\in\mathbf{N}} \bigcup_{1\le j\le m_i} \mathcal{V}_{ij} = \bigcup_{i\in\mathbf{N}} \mathcal{W}_i = \mathcal{X}.$$

If $\Sigma := \{\mathcal{U}_\alpha\}_{\alpha\in A}$ and $\Sigma_0 := \{\mathcal{V}_{ij}\}_{i\in\mathbf{N},1\le j\le m_i}$, then Σ_0 is a refinement of Σ; moreover,

$$\mathcal{W}_i \cap \prod_{1\le j\le m_i} \mathcal{V}_{kj} = \emptyset$$

for $k \ne i-2, i-1, i, i+1, i+2$, which implies that Σ_0 is a locally finite. □

Consider the function

$$f(t) := \begin{cases} e^{-1/t}, & t > 0, \\ 0, & t \le 0 \end{cases} \tag{1.1.2}$$

which is nonnegative, smooth, and positive for $t > 0$. Then the function

$$g(t) := \frac{f(t)}{f(t) + f(1-t)} \tag{1.1.3}$$

is nonnegative, smooth and takes the value 1 for $t \ge 1$ and the value 0 for $t \le 0$. Set

$$h(t) := g(t+2)g(2-t) \tag{1.1.4}$$

[1] See: Li, Yi. Topology I, Theorem 1.17.

which is a nonnegative smooth function on $\mathbf{R}$ which is 1 on $[-1,1]$ and zero outside of $(-2,2)$. Finally, define

$$\varphi := (h\circ r^1)\cdots(h\circ r^m). \tag{1.1.5}$$

In general, we can show that there exists a nonnegative smooth function φ on $\mathbf{R}^m$ which equals 1 on the closed cube $\overline{\mathbf{C}_1}$ and zero on the complement of the open cube $\mathbf{C}_2$.

Theorem 1.1. (Existence of partitions of unity)

Let $\mathcal{M}$ be a manifold of dimension m and $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ an open cover of $\mathcal{M}$. Then there exists a countable partition of unity $\{\varphi_i\}_{i\in\mathbf{N}}$ subordinate to the cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ with $\mathrm{supp}(\varphi_i)$ compact for each i. If one does not require compact supports, then there is a partition of unity $\{\varphi_\alpha\}_{\alpha\in A}$ subordinate to the cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ (that is, $\mathrm{supp}(\varphi_\alpha)\subset\mathcal{U}_\alpha$ with at most countably many of the φ_α not identically zero). ♡

Proof. Let the sequence $\{\mathcal{G}_i\}_{i\in\mathbf{N}}$ cover $\mathcal{M}$ as in **Proposition 1.1** and set $\mathcal{G}_0=\emptyset$. For $x\in\mathcal{M}$ let i_x be the largest integer such that $x\in\mathcal{M}\setminus\overline{\mathcal{G}_{i_x}}$. Choose an $\alpha_x\in A$ such that $x\in\mathcal{U}_{\alpha_x}$ and let $(\mathcal{V},\tau)$ be a coordinate system centered at x such that

$$\mathcal{V}\subset\mathcal{U}_{\alpha_x}\cap(\mathcal{G}_{i_x+2}\setminus\overline{\mathcal{G}_{i_x}}),\quad \tau(\mathcal{V})\subset\overline{\mathbf{C}_2}.$$

Define

$$\psi_x := \begin{cases}\varphi\circ\tau, & \text{on } \mathcal{V},\\ 0, & \text{otherwise}\end{cases}$$

where φ is the function given by (1.1.5). Then ψ_x is a smooth function on $\mathcal{M}$ which has the value 1 on some open neighborhood $\mathcal{W}_x$ of x, and has compact support lying in $\mathcal{V}$. Since $\{\mathcal{G}_i\}_{i\in\mathbf{N}}$ is locally finite, for each $i\geq 1$, we can choose a finite set of points x in $\mathcal{M}$ whose corresponding $\mathcal{W}_x$-neighborhoods cover $\overline{\mathcal{G}_i}\setminus\mathcal{G}_{i-1}$. Order the corresponding ψ_x functions in a sequence $\{\psi_j\}_{j\in\mathbf{N}}$. The supports of the ψ_j form a locally finite family of subsets of $\mathcal{M}$. Therefore the function

$$\psi := \sum_{j\in\mathbf{N}}\psi_j$$

is a well-defined strictly positive smooth function on $\mathcal{M}$. For each $i\in\mathbf{N}$ define

$$\varphi_i := \frac{\psi_i}{\psi}.$$

Then the functions $\{\varphi_i\}_{i\in\mathbf{N}}$ form a partition of unity subordinate to the cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ with $\mathrm{supp}(\varphi_i)$ compact.

Let φ_α be identically zero if no φ_i has support in $\mathcal{U}_\alpha$ and be the sum of the φ_i with support in $\mathcal{U}_\alpha$. Then $\{\varphi_\alpha\}_{\alpha\in A}$ is a partition of unity subordinate to the cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ with at most countably many of the φ_α not identically zero. Note that the support of φ_α lies in $\mathcal{U}_\alpha$ and observe that the support of φ_α is not necessarily compact. □ □

Corollary 1.1

Let $\mathcal{U}$ be open in a manifold $\mathcal{M}$ of dimension m and let $\mathcal{E}$ be closed in $\mathcal{M}$ with $\mathcal{E} \subset \mathcal{U}$. Then there exists a smooth function $\varphi : \mathcal{M} \to \mathbf{R}$ such that

(a) $0 \le \varphi \le 1$ *on* $\mathcal{M}$,

(b) $\varphi \equiv 1$ *on* $\mathcal{E}$, *and*

(c) $\mathrm{supp}(\varphi) \subset \mathcal{U}$.

♡

Proof. Consider an open cover $(\mathcal{U}, \mathcal{M} \setminus \mathcal{E})$. By **Theorem 1.1**, we have a partition of unity (φ, ψ) subordinate to $(\mathcal{U}, \mathcal{M} \setminus \mathcal{E})$ with $\mathrm{supp}(\varphi) \subset \mathcal{U}$ and $\mathrm{supp}(\psi) \subset \mathcal{M} \setminus \mathcal{E}$. Since

$$\varphi + \psi \equiv 1 \quad \text{on } \mathcal{M},$$

it follows that $\varphi \equiv 1$ on $\mathcal{E}$. □

1.1.3 Tangent vectors

If $\boldsymbol{V} = (V^1, \cdots, V^m)$ is a vector at a point $\boldsymbol{p}$ and f is differential on a neighborhood of $\boldsymbol{p}$, we define

$$\boldsymbol{V}(f) := \sum_{1 \le i \le m} V^i \frac{\partial f}{\partial r^i}\bigg|_{\boldsymbol{p}} = \langle \boldsymbol{V}, \partial f \rangle |_{\boldsymbol{p}} \tag{1.1.6}$$

called the **directional derivative of f in the direction $\boldsymbol{V}$ at $\boldsymbol{p}$**.

Definition 1.6. (Germs)

Let x be a point of a manifold $\mathcal{M}$. Smooth functions f and g defined on open subsets containing x are said to have the same **germ at x** *if they agree on some neighborhood of x.*

♣

If $\mathcal{U}_f$ denotes an open subset of a smooth function f containing x, we set

$$\mathscr{F} := \{(f, \mathcal{U}_f) : x \in \mathcal{U}_f\}. \tag{1.1.7}$$

Definition 1.6 introduces an equivalence relation on $\mathscr{F}$:

$$(f, \mathcal{U}_f) \sim (g, \mathcal{U}_g) \quad \Longleftrightarrow \quad f = g \quad \text{on} \quad \mathcal{U} \subset \mathcal{U}_f \cap \mathcal{U}_g \tag{1.1.8}$$

for some open subset $\mathcal{U}$ containing x. The set of equivalence classes is denoted by

$$\widetilde{\mathscr{F}_x} := \mathscr{F} / \sim = \{[f] : (f, \mathcal{U}_f) \in \mathscr{F}\}. \tag{1.1.9}$$

The equivalence class $[f]$ is also written as $\boldsymbol{f}$. Define

$$\widetilde{\mathscr{F}_x} \longrightarrow \mathbf{R}, \quad \boldsymbol{f} \longmapsto \boldsymbol{f}(x) := f(x). \tag{1.1.10}$$

It is clear that the mapping (1.1.10) is well-defined.

Let

$$\mathscr{F}_x := \{\boldsymbol{f} \in \widetilde{\mathscr{F}_x} : \boldsymbol{f}(x) = 0\} \tag{1.1.11}$$

and $\mathscr{F}_x^k$ be the k-th power of $\mathscr{F}_x$.

Proposition 1.2

$\mathscr{F}_x$ is an ideal in $\widetilde{\mathscr{F}}_x$, and $\mathscr{F}_x^k$ is an ideal of $\widetilde{\mathscr{F}}_x$ consisting of all finite linear combinations of k-fold products of elements of $\mathscr{F}_x$. These form a descending sequence of ideals

$$\widetilde{\mathscr{F}}_x \supset \mathscr{F}_x \supset \mathscr{F}_x^2 \subset \mathscr{F}_x^3 \supset \cdots .$$

♡

Definition 1.7. (Tangent vectors)

A **tangent vector V at the point $x \in \mathcal{M}$** *is a linear derivation of the algebra $\widetilde{\mathscr{F}}_x$. That is, for $\boldsymbol{f}, \boldsymbol{g} \in \widetilde{\mathscr{F}}_x$ and $\lambda \in \mathbf{R}$,*

$$\begin{aligned} V(\boldsymbol{f} + \lambda \boldsymbol{g}) &= V(\boldsymbol{f}) + \lambda V(\boldsymbol{g}), \\ V(\boldsymbol{f} \cdot \boldsymbol{g}) &= \boldsymbol{f}(x) V(\boldsymbol{g}) + \boldsymbol{g}(x) V(\boldsymbol{f}). \end{aligned}$$

$T_x\mathcal{M}$ denotes the set of tangent vectors to $\mathcal{M}$ at x and is called the **tangent space to $\mathcal{M}$ at x**.

♣

Let $\mathcal{M}$ be an m-dimensional manifold. For $V, W \in T_x\mathcal{M}$ and $\lambda \in \mathbf{R}$, we define

$$(V + W)(\boldsymbol{f}) := V(\boldsymbol{f}) + \lambda W(\boldsymbol{f}), \quad \lambda V(\boldsymbol{f}) := \lambda V(\boldsymbol{f}). \tag{1.1.12}$$

In this way, $T_x\mathcal{M}$ becomes a real vector space. We will show in **Theorem 1.2** that

$$\dim T_x\mathcal{M} = \dim \mathcal{M} = m.$$

If $\boldsymbol{c}$ is the germ of a function with the constant value c on a neighborhood of x. For $V \in T_x\mathcal{M}$, we have

$$\begin{aligned} V(\boldsymbol{c}) &= V(c\mathbf{1}) = cV(\mathbf{1}), \\ V(\mathbf{1}) &= V(\mathbf{1} \cdot \mathbf{1}) = 1V(\mathbf{1}) + 1V(\mathbf{1}) = 2V(\mathbf{1}). \end{aligned}$$

Consequently,

$$V(\boldsymbol{c}) = 0. \tag{1.1.13}$$

Lemma 1.1

$T_x\mathcal{M}$ is naturally isomorphic to $(\mathscr{F}_x/\mathscr{F}_x^2)^*$.

♡

Proof. Define

$$\varphi : (\mathscr{F}_x/\mathscr{F}_x^2)^* \longrightarrow T_x\mathcal{M}, \quad \ell \longmapsto V_\ell,$$

where

$$V_\ell(\boldsymbol{f}) := \ell([\boldsymbol{f} - [f(x)]]), \quad \boldsymbol{f} \in \widetilde{\mathscr{F}}_x.$$

This is well-defined, since $\boldsymbol{f} - [f(x)] \in \mathscr{F}_x$.

For any $\boldsymbol{f}, \boldsymbol{g} \in \widetilde{\mathscr{F}}_x$ and $\lambda \in \mathbf{R}$, compute

$$\begin{aligned} V_\ell(\boldsymbol{f} + \lambda \boldsymbol{g}) &= V_\ell([f + \lambda g]) = \ell([[f + \lambda g] - [(f + \lambda g)(x)]]) \\ &= \ell([\boldsymbol{f} + \lambda \boldsymbol{g} - [f(x)] - \lambda[g(x)]]) = V_\ell(\boldsymbol{f}) + \lambda V_\ell(\boldsymbol{g}) \end{aligned}$$

and

$$\begin{aligned} V_\ell(\boldsymbol{f}\cdot\boldsymbol{g}) &= V_\ell([fg]) = \ell([[fg]-[(fg)(x)]]) \\ &= \ell([(\boldsymbol{f}-[f(x)])(\boldsymbol{g}-[g(x)]) \\ &\quad + [f(x)](\boldsymbol{g}-[g(x)]) + [g(x)](\boldsymbol{f}-[f(x)])]) \\ &= [f(x)]V_\ell(\boldsymbol{g}) + [g(x)]V_\ell(\boldsymbol{f}) + A \end{aligned}$$

where

$$A = \ell([(\boldsymbol{f}-[f(x)])(\boldsymbol{g}-[g(x)])]) = 0$$

because of $(\boldsymbol{f}-[f(x)])(\boldsymbol{g}-[g(x)]) \in \mathscr{F}_x^2$.

Conversely, define

$$\psi : T_x\mathcal{M} \longrightarrow (\mathscr{F}_x/\mathscr{F}_x^2)^*, \quad U \longmapsto \psi(U),$$

where

$$\psi(U)([\boldsymbol{f}]) := U(\boldsymbol{f}+[f(x)]), \quad \boldsymbol{f} \in \mathscr{F}_x.$$

Compute

$$\begin{aligned} (\varphi\circ\psi(U))(\boldsymbol{f}) &= \varphi(\psi(U))(\boldsymbol{f}) = \psi(U)([\boldsymbol{f}-[f(x)]]) \\ &= U(\boldsymbol{f}-[f(x)]+[f(x)]) = U(\boldsymbol{f}), \\ (\psi\circ\varphi(\ell))([\boldsymbol{f}]) &= \psi(\varphi(\ell))([\boldsymbol{f}]) = \varphi(\ell)(\boldsymbol{f}+[f(x)]) \\ &= \ell([\boldsymbol{f}+[f(x)]-[f(x)]]) = \ell([\boldsymbol{f}]). \end{aligned}$$

Thus ψ is the inverse of φ, and hence $T_x\mathcal{M}$ is isomorphic to $(\mathscr{F}_x/\mathscr{F}_x^2)^*$. □

Theorem 1.2

$\dim \mathscr{F}_x/\mathscr{F}_x^2 = \dim \mathcal{M}.$ ♡

Proof. The proof is based on the following

Lemma 1.2. (Taylor's expansion)

If g is of class C^k $(k \geq 2)$ on a convex open set U about $\boldsymbol{p}$ in $\mathbf{R}^m$, then for each $\boldsymbol{q} \in U$,

$$\begin{aligned} g(\boldsymbol{q}) &= g(\boldsymbol{p}) + \sum_{1\leq i\leq m} \left.\frac{\partial g}{\partial r^i}\right|_{\boldsymbol{p}} (r^i(\boldsymbol{q}) - r^i(\boldsymbol{p})) \qquad (1.1.14) \\ &\quad + \sum_{1\leq i,j\leq m} (r^i(\boldsymbol{q})-r^i(\boldsymbol{p}))(r^j(\boldsymbol{q})-r^j(\boldsymbol{p})) \int_0^1 (1-t) \left.\frac{\partial^2 g}{\partial r^i \partial r^j}\right|_{\boldsymbol{p}+t(\boldsymbol{q}-\boldsymbol{p})} dt. \end{aligned}$$

♡

Let $(\mathcal{U},\varphi)$ be a coordinate system about x with coordinate functions $x^1,\cdots,x^m$. For any $\boldsymbol{f} \in \mathscr{F}_x$, we have, for any $\boldsymbol{q} \in \varphi(\mathcal{U})$,

$$\begin{aligned} (f\circ\varphi^{-1})(\boldsymbol{q}) &= (f\circ\varphi^{-1})(\varphi(x)) + \sum_{1\leq i\leq m} \left.\frac{\partial(f\circ\varphi^{-1})}{\partial r^i}\right|_{\varphi(x)} (r^i(\boldsymbol{q}) - r^i(\varphi(x))) \\ &\quad + \sum_{1\leq i,j\leq m} (r^i(\boldsymbol{q}) - r^i(\varphi(x))(r^j(\boldsymbol{q}) - r^j(\varphi(x)))h(\boldsymbol{q}) \end{aligned}$$

by (1.1.14), where

$$h(\boldsymbol{q}) := \int_0^1 (1-t)\frac{\partial^2(f\circ\varphi^{-1})}{\partial r^i \partial r^j}\bigg|_{t\boldsymbol{q}+(1-t)\varphi(x)} dt$$

is a smooth function near $\varphi(x)$. Composing with φ yields

$$\begin{aligned} f(q) &= f(x) + \sum_{1\le i\le m} \frac{\partial(f\circ\varphi^{-1})}{\partial r^i}\bigg|_{\varphi(x)} ((r^i\circ\varphi)(q) - (r^i\circ\varphi)(x)) \\ &\quad + \sum_{1\le i,j\le m} ((r^i\circ\varphi)(q) - (r^i\circ\varphi)(x))((r^j\circ\varphi)(q) - (r^j\circ\varphi)(x))(h\circ\varphi)(q). \end{aligned}$$

Thus

$$f = \sum_{1\le i\le m} \frac{\partial(f\circ\varphi^{-1})}{\partial r^i}\bigg|_{\varphi(x)} (x^i - x^i(x)) + \sum_{1\le i,j\le m} (x^i - x^i(x))(x^j - x^j(x))(h\circ\varphi).$$

Consequently,

$$\boldsymbol{f} = \sum_{1\le i\le m} \frac{\partial(f\circ\varphi^{-1})}{\partial r^i}\bigg|_{\varphi(x)} (\boldsymbol{x}^i - [x^i(x)]) \bmod \mathscr{F}_x^2$$

and hence $([\boldsymbol{x}^i - [x^i(x)]])_{1\le i\le m}$ spans $\mathscr{F}_x/\mathscr{F}_x^2$. Suppose now that

$$\sum_{1\le i\le m} a_i(\boldsymbol{x}^i - [x^i(x)]) \in \mathscr{F}_x^2$$

for $a_i \in \mathbf{R}$. Since

$$\sum_{1\le i\le m} a_i(x^i - x^i(x))\circ\varphi^{-1} = \sum_{1\le i\le m} a_i(r^i - r^i(\varphi(x)))$$

it follows that

$$\sum_{1\le i\le m} a_i(\boldsymbol{r}^i - [r^i(\varphi(x))]) \in \mathscr{F}_{\varphi(x)}^2$$

which implies

$$0 = \frac{\partial}{\partial r^j}\bigg|_{\varphi(x)} \left(\sum_{1\le i\le m} a_i(r^i - r^i(\varphi(x)))\right) = \sum_{1\le i\le m} a_i\delta_{ij} = a_j, \quad 1\le j\le m.$$

Thus $\dim \mathscr{F}_x/\mathscr{F}_x^2 = m$. □

Corollary 1.2

For any $x\in\mathcal{M}$, we have $\dim T_x\mathcal{M} = \dim\mathcal{M}$. ♡

If f is smooth function defined on a neighborhood of $x\in\mathcal{M}$ and $V\in T_x\mathcal{M}$, we define

$$V(f) := V(\boldsymbol{f}). \tag{1.1.15}$$

Thus $V(f) = V(g)$ whenever f and g agree on a neighborhood of x, and

$$V(f+\lambda g) = V(f) + \lambda V(g), \quad V(fg) = f(x)V(g) + g(x)V(f).$$

This shows that we can treat tangent vectors as operating on functions rather than on their germs.

$$\begin{array}{ccc} \tilde{\mathscr{F}}_x & \xrightarrow{V} & \mathbf{R} \\ \uparrow & & \\ C^\infty(x) & \xrightarrow{V} & \mathbf{R} \end{array}$$

Definition 1.8. (Natural tangent vectors)

Let $(\mathcal{U}, \varphi)$ be a coordinate system with coordinate functions $x^1, \cdots, x^m$ and let $x \in \mathcal{U}$. For each $i \in \{1, \cdots, m\}$, we define a tangent vector $(\partial/\partial x^i)|_x \in T_x\mathcal{M}$ by

$$\left(\frac{\partial}{\partial x^i}\bigg|_x\right)(f) = \frac{\partial f}{\partial x^i}\bigg|_x := \frac{\partial(f \circ \varphi^{-1})}{\partial r^i}\bigg|_{\varphi(x)} \tag{1.1.16}$$

for each function f which is smooth near x. ♣

Note 1.1

$((\partial/\partial x^i)|_x)(f)$ depends only on the germ of f at x, and $(\partial/\partial x^i)_x$ is a tangent vector at x.

(a) $((\partial/\partial x_i)|_x)_{1 \le i \le m}$ is a basis of $T_x\mathcal{M}$ and dual to $[\boldsymbol{x}^i - [x^i(x)]]_{1 \le i \le m}$. Indeed,

$$\left(\frac{\partial}{\partial x^i}\bigg|_x\right)(x^j - x^j(x)) = \frac{\partial(r^j - r^j(\varphi(x)))}{\partial r^i}\bigg|_{\varphi(x)} = \delta_j.$$

(b) If $V \in T_x\mathcal{M}$, then

$$V = \sum_{1 \le i \le m} V(x^i)\frac{\partial}{\partial x^i}\bigg|_x.$$

Indeed, witting $V = \sum_{1 \le i \le m} a^i(\partial/\partial x^i)|_x$ we get

$$V(x^j) = V(x^j - x^j(x)) = \sum_{1 \le i \le m} a^i\left(\frac{\partial}{\partial x^i}\bigg|_x\right)(x^j - x^j(x)) = \sum_{1 \le i \le m} a^i\delta_{ij} = a_j.$$

(c) Suppose that $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ are coordinate systems about x with coordinate functions $x^1, \cdots, x^m$ and $y^1, \cdots, y^m$ respectively. Then

$$\frac{\partial}{\partial y^j}\bigg|_x = \sum_{1 \le i \le m}\left(\frac{\partial}{\partial y^j}\bigg|_x\right)(x^i)\frac{\partial}{\partial x^i}\bigg|_x = \sum_{1 \le i \le m}\frac{\partial x^i}{\partial y^j}\bigg|_x\frac{\partial}{\partial x^i}\bigg|_x.$$

In particular, if x^1 were equal to y^1, then

$$\frac{\partial}{\partial y^1}\bigg|_x = \frac{\partial}{\partial x^1}\bigg|_x + \sum_{2 \le i \le m}\frac{\partial x^i}{\partial y^1}\bigg|_x\frac{\partial}{\partial x^i}\bigg|_x \ne \frac{\partial}{\partial x^1}\bigg|_x.$$

(d) When $\mathcal{M} = \mathbf{R}^m$ with the canonical coordinate system $(\mathbf{R}^m, \mathbf{1}, r^1, \cdots, r^m)$, we obtain

$$\left(\frac{\partial}{\partial r^i}\bigg|_{\boldsymbol{x}}\right)(f) = \frac{\partial(f \circ \mathbf{1}^{-1})}{\partial r^i}\bigg|_{\mathbf{1}(\boldsymbol{x})} = \frac{\partial f}{\partial r^i}\bigg|_{\boldsymbol{x}}.$$

Thus the tangent vectors defined above are the ordinary partial derivative operators $(\partial/\partial r^i)$. In particular, $T_{\boldsymbol{x}}\mathbf{R}^m \cong \mathbf{R}^m$.

(e) We defined $\mathscr{F}_x$ and $\mathscr{F}_x^k$ in the C^∞ case and shows that it is finite. However, $\mathscr{F}_x/\mathscr{F}_x^2$ is always infinite dimensional in the C^k case for $1 \le k < \infty$. There are lots of ways to define tangent vectors in the C^k case so that $\dim T_x\mathcal{M} = \dim \mathcal{M}$ (all of which work in the C^∞ case too). ♣

Let $F : \mathcal{M} \to \mathcal{N}$ be a smooth map between two manifolds of dimensions m and n respectively, and $x \in \mathcal{M}$. The **differential of F at x** is the linear map

$$F_{*,x} : T_x\mathcal{M} \longrightarrow T_{F(x)}\mathcal{N}, \quad V \longmapsto F_{*,x}(V) \tag{1.1.17}$$

defined as follows: for any smooth function g defined on a neighborhood of $\psi(x)$, we define

$$F_{*,x}(V)(g) := V(g \circ F). \tag{1.1.18}$$

Clearly that $F_{*,x}$ is a linear map of $T_x\mathcal{M}$ into $T_{F(x)}\mathcal{N}$. The map F is called **nonsingular at x** if $F_{*,x}$ is nonsingular, that is, $\ker(F_{*,x}) = 0$. The dual map

$$F_x^* : T^*_{F(x)}\mathcal{N} := (T_{F(x)}\mathcal{N})^* \longrightarrow T_x^*\mathcal{M} := (T_x\mathcal{M})^*, \quad \omega \longmapsto \psi_x^*(\omega) \tag{1.1.19}$$

defined by

$$\psi_x^*(\omega)(V) := \omega(\psi_{*,x}(V)), \quad V \in T_x\mathcal{M}. \tag{1.1.20}$$

If f is a smooth function on $\mathcal{M}$, and if $V \in T_x\mathcal{M}$ and $f(x) = r_0$, then

$$\mathrm{grad}_x(f) := f_{*,x} : T_x\mathcal{M} \longrightarrow T_{f(x)}\mathbf{R} \cong \mathbf{R}, \quad V \longmapsto \mathrm{grad}_x(f)(V).$$

we have

$$\mathrm{grad}_x(f)(V) = (\mathrm{grad}_x(f)(V))(r)\frac{d}{dr}\bigg|_{r_0} = V(r \circ f)\frac{d}{dr}\bigg|_{r_0} = V(f)\frac{d}{dr}\bigg|_{r_0}. \tag{1.1.21}$$

Hence $\mathrm{grad}_x(f)$ can be viewed as an element of $T_x^*\mathcal{M}$. More precisely, define

$$df_x : T_x\mathcal{M} \longrightarrow \mathbf{R}, \quad V \longmapsto V(f). \tag{1.1.22}$$

The natural isomorphism $\partial_{r_0} : T_{f(x)}\mathbf{R} \to \mathbf{R}$ given by $\partial_{r_0}(a\frac{d}{dr}|_{r_0}) = a$ implies

$$(\partial_{r_0} \circ \mathrm{grad}_x(f))(V) = \partial_{r_0}\left(V(f)\frac{d}{dr}\bigg|_{r_0}\right) = V(f) = df_x.$$

If ω_{r_0} is the basis of the one-dimensional space $T_{r_0}^*\mathbf{R}$ dual to $\frac{d}{dr}|_{r_0}$, we arrive at

$$f_x^*(\omega_{r_0})(V) = \omega_{r_0}(f_{*,x}(V)) = \omega_{r_0}\left(V(f)\frac{d}{dr}\bigg|_{r_0}\right) = V(f)\omega_{r_0}\left(\frac{d}{dr}\bigg|_{r_0}\right) = V(f) = df_x(V).$$

Thus

$$df_x = f_x^*(\omega_{f(x)}). \tag{1.1.23}$$

Note 1.2

(a) Consider a smooth map $F : \mathcal{M} \to \mathcal{N}$ between two manifolds of dimensions m and n respectively, and $x \in \mathcal{M}$. Let $(\mathcal{U}, \varphi, x^1, \cdots, x^m)$ and $(\mathcal{V}, \psi, y^1, \cdots, y^n)$ be coordinate systems about x and $F(x)$ respectively. Then

$$\begin{aligned} F_{*,x}\left(\frac{\partial}{\partial x^j}\bigg|_x\right) &= \sum_{1\le i\le n} F_{*,x}\left(\frac{\partial}{\partial x^j}\bigg|_x\right)(y^i)\frac{\partial}{\partial y^i}\bigg|_{F(x)} \\ &= \sum_{1\le i\le n}\left(\frac{\partial}{\partial x^j}\bigg|_x\right)(y^i\circ F)\frac{\partial}{\partial y^i}\bigg|_{F(x)} = \sum_{1\le i\le n}\frac{\partial(y^i\circ F)}{\partial x^j}\bigg|_x\frac{\partial}{\partial y^i}\bigg|_{F(x)} \end{aligned}$$

by Note 1.1 (c). The matrix $(\partial(y^i \circ F)/\partial x^j)_{1\le i\le n, 1\le j\le m}$ is called the **Jacobian of F**.

(b) If $(\mathcal{U}, x^1, \cdots, x^m)$ is a coordinate system on $\mathcal{M}$ and $x \in \mathcal{U}$, then $\{dx^i|_x\}_{1\le i\le m}$ is the basis of $T_x^*\mathcal{M}$ dual to $\{\partial/\partial x^i|_x\}_{1\le i\le m}$ by (1.1.23). If f is a smooth function, then

$$df_x = \sum_{1\le i\le m}\frac{\partial f}{\partial x^i}\bigg|_x dx^i|_x. \tag{1.1.24}$$

In fact,

$$dx^i|_x\left(\frac{\partial}{\partial x^j}\bigg|_x\right) = \frac{\partial}{\partial x^j}\bigg|_x (x^i) = \delta_{ij}.$$

Since $df_x \in T_x^*\mathcal{M}$*, we can write* $df_x = \sum_{1\le i\le m} a_i dx^i|_x$*. Then*

$$\frac{\partial f}{\partial x^j}\bigg|_x = \frac{\partial}{\partial x^j}\bigg|_x (f) = df_x\left(\frac{\partial}{\partial x^j}\bigg|_x\right) = \sum_{1\le i\le m} a_i\delta_{ij} = a_j.$$

(c) **Chain rule.** *Let* $F:\mathcal{M}\to\mathcal{N}$ *and* $G:\mathcal{N}\to\mathcal{P}$ *be smooth maps. Then*

$$(F\circ G)_{*,x} = F_{*,G(x)}\circ G_{*,x}. \tag{1.1.25}$$

For any smooth function g *near* $F(G(x))$ *and* $V\in T_x\mathcal{M}$*, we have*

$$\begin{aligned}(F\circ G)_{*,x}(V)(g) &= V(g\circ\varphi\circ G) = V((g\circ F)\circ G) = G_{*,x}(V)(g\circ F)\\ &= F_{*,G(x)}(\psi_{G,x}(V))(g) = (F_{*,G(x)}\circ G_{*,x})(V)(g).\end{aligned}$$

(d) If $F:\mathcal{M}\to\mathcal{N}$ *and* $f:\mathcal{N}^n\to\mathbf{R}$ *are smooth, then*

$$d(f\circ F)_x = F_x^*(df_{\psi(x)}). \tag{1.1.26}$$

For any $V\in T_x\mathcal{M}$*, we obtain*

$$(F_x^*(df_{F(x)}))(V) = df_{F(x)}(F_{*,x}(V)) = F_{*,x}(V)(f) = V(f\circ F) = d(f\circ F)_x(V).$$

(e) A smooth mapping $\sigma:(a,b)\to\mathcal{M}$ *is called a* **smooth curve** *in* $\mathcal{M}$*. Let* $t\in(a,b)$*. Then the* **tangent vector to σ at t** *is the vector*

$$\dot\sigma(t) := \sigma_{*,t}\left(\frac{d}{dr}\bigg|_t\right)\in T_{\sigma(t)}\mathcal{M}. \tag{1.1.27}$$

If V *is any nonzero element of* $T_x\mathcal{M}$*, then*

$$V = \varphi^{-1}_{*,\mathbf{0}}\left(\frac{\partial}{\partial r^1}\bigg|_{\mathbf{0}}\right) \tag{1.1.28}$$

for some coordinate system $(\mathcal{U},\varphi)$ *centered at* x*. Hence* V *is the tangent vector at* $\mathbf{0}$ *to the curve* $\sigma(t) := \varphi^{-1}(t,0,\cdots,0)$*.*

Two smooth curves σ *and* τ *in* $\mathcal{M}^m$ *for which* $\sigma(t_0)=\tau(t_0)=x$ *have the same tangent vector at* t_0 *if and only if*

$$\frac{d(f\circ\sigma)}{dr}\bigg|_{t_0} = \frac{d(f\circ\tau)}{dr}\bigg|_{t_0}$$

for all functions f *which are smooth on a neighborhood of* x*.* ♣

If $\sigma:(a,b)\to\mathbf{R}^m$ is a curve in $\mathbf{R}^m$, then

$$\begin{aligned}\dot\sigma(t)(f) &= \sigma_{*,t}\left(\frac{d}{dr}\bigg|_t\right)(f) = \frac{d}{dr}\bigg|_t (f\circ\sigma) = \sum_{1\le i\le m}\frac{\partial f}{\partial r^j}\bigg|_{\sigma(t)}\frac{d\sigma^j}{dr}\bigg|_t\\ &= \left(\sum_{1\le j\le m}\frac{d\sigma^j}{dr}\bigg|_t\frac{\partial}{\partial r^j}\bigg|_{\sigma(t)}\right)(f);\end{aligned}$$

thus

$$\dot\sigma(t) = \sum_{1\le j\le m}\frac{d\sigma^j}{dr}\bigg|_t\frac{\partial}{\partial r^j}\bigg|_{\sigma(t)}.$$

If we identify this tangent vector with the element

$$\left(\frac{d\sigma^1}{dr}\bigg|_t, \cdots, \frac{d\sigma^m}{dr}\bigg|_t\right)$$

of $\mathbf{R}^m$, then we arrive at

$$\dot{\sigma}(t) = \lim_{h\to 0}\frac{\sigma(t+h)-\sigma(t)}{h}.$$

Consequently, with this identification out notion of tangent vector coincides with the geometric notion of a tangent to a curve in Euclidean space.

Theorem 1.3

Let $F : \mathcal{M} \to \mathcal{N}$ be a smooth mapping and $\mathcal{M}$ be connected. If $F_{*,x} \equiv 0$ for every $x \in \mathcal{M}$, then ψ is a constant map. ♡

Proof. Let $y \in F(\mathcal{M})$ and $x \in F^{-1}(y)$ (notice that $\psi^{-1}(y)$ is closed). Choose coordinate systems $(\mathcal{U}, x^1, \cdots, x^m)$ and $(\mathcal{V}, y^1, \cdots, y^n)$ about x and y respectively so that $F(\mathcal{U}) \subset \mathcal{V}$. For $x' \in \mathcal{U}$,

$$0 = F_{*,x'}\left(\frac{\partial}{\partial x^j}\bigg|_{x'}\right) = \sum_{1\le i\le n}\frac{\partial(y^i\circ F)}{\partial x^j}\bigg|_{x'}\frac{\partial}{\partial y^i}\bigg|_{F(x')}, \quad 1 \le j \le m$$

by **Note 1.1** (a), implying that

$$\frac{\partial(y^i\circ F)}{\partial x^j} \equiv 0, \quad 1 \le i \le n,\ 1 \le j \le m.$$

Thus $y^i \circ F$ are constant on $\mathcal{U}$ and hence $\psi(\mathcal{U})$ is constant. Since $F(x) = y$, it follows that $F(\mathcal{U}) = y$ and $F^{-1}(y) = \mathcal{U}$ that is open.

Because $F^{-1}(y)$ is open and closed in a connected manifold $\mathcal{M}$, we must have $\psi^{-1}(y) = \mathcal{M}$. Thus $F(\mathcal{M}) = y$. □

Let $\mathcal{M}$ be a smooth manifold with differentiable structure $\mathscr{F}$. Let

$$T\mathcal{M} := \bigcup_{x\in\mathcal{M}} T_x\mathcal{M}, \quad T^*\mathcal{M} := \bigcup_{x\in\mathcal{M}} T^*_x\mathcal{M}. \tag{1.1.29}$$

There are natural projections:

$$\begin{array}{ccc} T\mathcal{M} & \xrightarrow{\pi} & \mathcal{M} \\ & & \uparrow{\scriptstyle \pi^*} \\ & & T^*\mathcal{M} \end{array}$$

where

$$\pi(V) = x \quad \text{if } V \in T_x\mathcal{M}, \quad \pi^*(V) = x \quad \text{if } V \in T^*_x\mathcal{M}. \tag{1.1.30}$$

Let $(\mathcal{U}, \varphi) \in \mathscr{F}$ with coordinate functions $x^1, \cdots, x^m$. Define $\tilde{\varphi} : \pi^{-1}(\mathcal{U}) \to \mathbf{R}^{2m}$ and $\tilde{\varphi}^* : (\pi^*)^{-1}(\mathcal{U}) \to \mathbf{R}^{2m}$ by

$$\begin{aligned} \tilde{\varphi}(V) &= \left(x^1(\pi(V)), \cdots, x^m(\pi(V)), dx^1(V), \cdots, dx^m(V)\right), \\ \tilde{\varphi}^*(\tau) &= \left(x^1(\pi^*(\tau)), \cdots, x^m(\pi^*(\tau)), \tau\left(\frac{\partial}{\partial x^1}\right), \cdots, \tau\left(\frac{\partial}{\partial x^m}\right)\right) \end{aligned}$$

for all $V \in \pi^{-1}(\mathcal{U})$ and $\tau \in (\pi^*)^{-1}(\mathcal{U})$.

We now construct a topology and a differentiable structure on $T\mathcal{M}$.

(a) If $(\mathcal{U}, \varphi), (\mathcal{V}, \psi) \in \mathscr{F}$, then $\tilde{\psi} \circ \tilde{\varphi}^{-1}$ is smooth.

(b) The collection

$$\left\{\tilde{\varphi}^{-1}(\mathcal{W}) : \mathcal{W} \text{ open in } \mathbf{R}^{2m},\ (\mathcal{U}, \varphi) \in \mathscr{F}\right\}$$

forms a basis for a topology on $T\mathcal{M}$ which makes $T\mathcal{M}$ into a $2m$-dimensional, second countable topology manifold.

(c) Let $\widetilde{\mathscr{F}}$ be the maximal collection containing

$$\left\{(\pi^{-1}(\mathcal{U}), \tilde{\varphi}) : (\mathcal{U}, \varphi) \in \mathscr{F}\right\}.$$

Then $\tilde{\mathscr{F}}$ is a differentiable structure on $T\mathcal{M}^m$. Note that $\pi^{-1}(\mathcal{U}) \cong \mathcal{U} \times \mathbf{R}^m \subset \mathbf{R}^{2m}$.

The construction for $T^*\mathcal{M}$ goes similarly. $T\mathcal{M}$ and $T^*\mathcal{M}$ with these differentiable structures are called respectively the **tangent bundle** and the **cotangent bundle** of $\mathcal{M}$. Note that $\tilde{\varphi}$ and $\tilde{\varphi}^*$ are both one-to-one maps onto open subsets of $\mathbf{R}^{2m}$. We prove here only for $\tilde{\varphi}$.

$$\begin{array}{ccc} \pi^{-1}(\mathcal{U}) & \xrightarrow{\tilde{\varphi}} & \mathcal{U} \times \mathbf{R}^m \\ \downarrow & & \| \\ \varphi(\mathcal{U}) \times \mathbf{R}^m & = & \varphi(\mathcal{U}) \times \mathbf{R}^m \end{array}, \qquad \begin{array}{c} V \\ \downarrow \\ (x^1(\pi(V)), \cdots, x^m(\pi(V)), dx^1(V), \cdots, dx^m(V)) \end{array}$$

Define $\tilde{\psi} : \mathcal{U} \times \mathbf{R}^m \to \pi^{-1}(\mathcal{U})$ by

$$\tilde{\psi}(x, \boldsymbol{v}) := \sum_{1 \le j \le m} v^j \frac{\partial}{\partial x^j}\bigg|_x$$

and

$$\begin{array}{ccc} \mathcal{U} \times \mathbf{R}^m & \xrightarrow{\tilde{\psi}} & \pi^{-1}(\mathcal{U}) \\ \| & & \uparrow \\ \varphi(\mathcal{U}) \times \mathbf{R}^m & = & \varphi(\mathcal{U}) \times \mathbf{R}^m \end{array}, \qquad \begin{array}{c} (x, \boldsymbol{v}) \\ \downarrow \\ (x^1(x), \cdots, x^m(x), \boldsymbol{v}) \end{array}$$

Compute

$$\begin{aligned}
(\tilde{\varphi} \circ \tilde{\psi})(x, \boldsymbol{v}) &= \tilde{\varphi}\left(\sum_{1 \le j \le m} v^j \frac{\partial}{\partial x^j}\bigg|_x\right) = \left(\pi\left(\sum_{1 \le j \le m} v^j \frac{\partial}{\partial x^j}\bigg|_x\right),\right. \\
&\qquad \left. dx^1\left(\sum_{1 \le j \le m} v^j \frac{\partial}{\partial x^j}\bigg|_x\right), \cdots, dx^m\left(\sum_{1 \le j \le m} v^j \frac{\partial}{\partial x^j}\bigg|_x\right)\right) \\
&= (x, v^1, \cdots, v^m) = (x, \boldsymbol{v}), \\
(\tilde{\psi} \circ \tilde{\varphi})(V) &= \tilde{\psi}(\pi(V), dx^1(V), \cdots, dx^m(V)) = \sum_{1 \le j \le m} dx^j(V) \frac{\partial}{\partial x^j}\bigg|_x \\
&= \sum_{1 \le j \le m} V(x^j) \frac{\partial}{\partial x^j}\bigg|_x = V
\end{aligned}$$

since $dx^j(V) = V(x^j)$.

It will sometimes be convenient to write the points of $T\mathcal{M}$ as pairs (x, V) where $x \in \mathcal{M}$ and $V \in T_x\mathcal{M}$ (and similarly for $T^*\mathcal{M}$).

If $F:\mathcal{M}\to\mathcal{N}$ is a smooth map, then the differential of F defines a mapping of the tangent bundles

$$F_*:T\mathcal{M}\longrightarrow T\mathcal{N} \tag{1.1.31}$$

where

$$F_*(x,V):=(F(x),F_{*,x}(V)),\quad (x,V)\in T\mathcal{M}. \tag{1.1.32}$$

Note that F_* is a smooth map.

1.1.4 Submanifolds

Let $F:\mathcal{M}\to\mathcal{N}$ be smooth.

(a) F is an **immersion** if $F_{*,x}$ is nonsingular for each $x\in\mathcal{M}$ (that is, $F_{*,x}$ is one-to-one).

(b) The pair $(\mathcal{M},F)$ is a **submanifold** of $\mathcal{N}$ if F is a one-to-one immersion.

(c) F is an **imbedding** if F is a one-to-one immersion which is also a homeomorphism into; that is, F is open as a map into $F(\mathcal{M})$ with the relative topology.

(d) F is a **diffeomorphism** if F maps $\mathcal{M}$ one-to-one onto $\mathcal{N}$ and F^{-1} is smooth.

Clearly that

$$\text{imbedding}\subset\text{submanifold}\subset\text{immersion}.$$

Example 1.2

(a) Consider $F:\mathbf{R}\to\mathbf{R}^2$ given by

$$F(t):=\left(2\cos\left(t-\frac{\pi}{2}\right),\sin 2\left(t-\frac{\pi}{2}\right)\right).$$

Then F is an immersion. Note that

$$F_{*,t}\left(\left.\frac{d}{dr}\right|_t\right)=\frac{dF(t)}{dt}=\left(-2\sin\left(t-\frac{\pi}{2}\right),2\cos 2\left(t-\frac{\pi}{2}\right)\right).$$

(b) Consider $G:\mathbf{R}\to\mathbf{R}^2$ given by

$$G(t):=\left(2\cos\left(2\tan^{-1}t-\frac{\pi}{2}\right),\sin 2\left(2\tan^{-1}t-\frac{\pi}{2}\right)\right).$$

Then G is a submanifold but not an imbedding. Since $\lim_{|t|\to\infty}G(t)=(0,0)$ and $G(0)=(0,0)$, it follows that a neighborhood of $(0,0)$ on $G(\mathbf{R})$ is of the form $(-a,b)\cup(A,+\infty)\cup(-\infty,-B)$ for some $a,b,A,B>0$.

(c) Consider $H:\mathbf{R}\to\mathbf{R}^2$ given by

$$H(t):=(t,1).$$

Then H is an imbedding. ♠

The composition of diffeomorphisms is again a diffeomorphism. Let

$$\mathscr{M}:=\{\text{smooth manifolds}\}.$$

We write $\mathcal{M}\sim\mathcal{N}$ in $\mathscr{M}$ if there exists a diffeomorphism $F:\mathcal{M}\to\mathcal{N}$. It is clear that $\sim$ is an equivalence relation on $\mathscr{M}$.

Note 1.3

(1) Consider the real line $\mathbf{R}$ *and take* $\mathscr{F}$ *to be the maximal collection containing the identity map* $\mathbf{1}$*. Let* $\varphi : \mathbf{R} \to \mathbf{R}$ *given by* $t \mapsto t^3$*, and* $\mathscr{F}'$ *be the maximal collection containing* φ*. Hence* $(\mathbf{R}, \mathscr{F})$ *and* $(\mathbf{R}, \mathscr{F}')$ *both are smooth manifolds. Since*

$$\mathbf{1} \circ \varphi^{-1}(t) = \mathbf{1}(t^{1/3}) = t^{1/3},$$

it follows that $\mathbf{1} \circ \varphi^{-1}$ *is not smooth at* $t = 0$ *and then* $\mathscr{F} \neq \mathscr{F}'$*. Define*

$$f : (\mathbf{R}, \mathscr{F}) \longrightarrow (\mathbf{R}, \mathscr{F}'), \quad t \longmapsto t^{1/3}.$$

Since

$$\varphi \circ f \circ \mathbf{1}^{-1}(t) = t, \quad \mathbf{1} \circ f^{-1} \circ \varphi^{-1}(t) = t,$$

the map f *must be diffeomorphic.*

(2) Milnor (1956) showed that $\mathbf{S}^7$ *possesses non-diffeomorphic differentiable structures, and Kervaire (1961) found a topology manifold that possesses no differentiable structures.* ♣

If $F : \mathcal{M} \to \mathcal{N}$ is a diffeomorphism, then $F_{*,x}$ is an isomorphism. Conversely, whenever $F_{*,x}$ is an isomorphism, we can show that F is a diffeomorphism on a neighborhood of x.

Definition 1.9

A set $\{f_j\}_{1 \le j \le k}$ *of smooth functions defined on some neighborhood of* x *in a manifold* $\mathcal{M}$ *is called an* **independent set at** $\boldsymbol{x}$ *if the differentials* $df_{1,x}, \cdots, df_{k,x}$ *form an independent set in* $T_x^*\mathcal{M}$*.* ♣

Theorem 1.4. (Inverse function theorem)

Let $U \subset \mathbf{R}^m$ *be open and let* $f : U \to \mathbf{R}^m$ *be smooth. If the Jacobian matrix*

$$\left(\frac{\partial(r^i \circ f)}{\partial r^j} \right)_{1 \le i,j \le m}$$

is nonsingular at $\mathbf{r_0} \in U$*, then there exists an open set* V *with* $\mathbf{r_0} \in V \subset U$ *such that* $f_{|V}$ *maps* V *one-to-one onto the open set* $f(V)$*, and* $f_{|V}^{-1}$ *is smooth.* ♡

Corollary 1.3

Assume that $F : \mathcal{M} \to \mathcal{N}$ *is smooth, that* $x \in \mathcal{M}$*, and that* $F_{*,x} : T_x\mathcal{M} \to T_{F(x)}\mathcal{N}$ *is an isomorphism. Then there is a neighborhood* $\mathcal{U}$ *of* x *such that* $F : \mathcal{U} \to F(\mathcal{U})$ *is a diffeomorphism onto the open set* $F(\mathcal{U})$ *in* $\mathcal{N}$*.* ♡

Proof. Since $F_{*,x}$ is isomorphic, it follows that $m = \dim \mathcal{M} = \dim T_x\mathcal{M} = \dim T_x\mathcal{N} = \dim \mathcal{N} = n$. Choose coordinate systems $(\mathcal{V}, \varphi)$ about x and $(\mathcal{W}, \psi)$ about $F(x)$ with $F(\mathcal{V}) \subset \mathcal{W}$. Let $\varphi(x) = \boldsymbol{p}$ and $\psi(F(x)) = \boldsymbol{q}$. Consider the map $f := \psi \circ F \circ \varphi^{-1} : V := \varphi(\mathcal{V}) \to W := \psi(\mathcal{W})$. Observe that $V, W \subset \mathbf{R}^m$ and

$$d(f_{|V})_{\boldsymbol{p}} = d(\psi \circ F_{|\mathcal{V}})_{\varphi^{-1}(\boldsymbol{p})} \circ d(\varphi_{|V}^{-1})_{\boldsymbol{p}} = d(\psi_{|F(\mathcal{V})})_{F \circ \varphi^{-1}(\boldsymbol{p})} \circ d(F_{|\mathcal{V}})_{\varphi^{-1}(\boldsymbol{p})} \circ d(\varphi_{|V}^{-1})_{\boldsymbol{p}}.$$

Hence $f_{|V}$ is nonsingular at $\boldsymbol{p}$ and, by **Theorem 1.4**, there is a diffeomorphism $\alpha : \tilde{U} \to \alpha(\tilde{U})$ on a neighborhood $\tilde{U}$ of $\boldsymbol{p}$ with $\tilde{\mathcal{U}} \subset V = \varphi(\mathcal{V})$. Then $\alpha = \psi \circ F \circ \varphi^{-1}$ on $\tilde{U}$; in particular, $F = \psi^{-1} \circ \alpha \circ \varphi$ on $\mathcal{U} := \varphi^{-1}(\tilde{U})$. □

Corollary 1.4

If $f_1, \cdots, f_m$ is an independent set of functions at $x_0 \in \mathcal{M}$, then $\{f_i\}_{1 \le i \le m}$ forms a coordinate system on a neighborhood of x_0. ♡

Proof. Suppose that $f_i : \mathcal{U} \to \mathbf{R}$ for $1 \le i \le m$, and $x_0 \in \mathcal{U}$. Define a smooth map $\psi : \mathcal{U} \to \mathbf{R}^m$ by

$$\psi(x) := (f_1(x), \cdots, f_m(x)), \quad x \in \mathcal{U}.$$

Consider

$$\psi_{x_0}^* : T_{\psi(x_0)}^* \mathbf{R}^m \longrightarrow T_{x_0}^* \mathcal{U}$$

and observe that

$$\psi_{x_0}^*(dr_{\psi(x_0)}^i) = d(r^i \circ \psi)_{x_0} = df_{i,x_0}.$$

Because $df_{1,x_0}, \cdots, df_{m,x_0}$ is a basis of $T_{x_0}^* \mathcal{U}$, we conclude that $\psi_{x_0}^*$ is an isomorphism on $T_{\psi(x_0)}^* \mathbf{R}^m$ and then its dual ψ_{*,x_0} is also an isomorphism. By **Corollary 1.3**, ψ is a diffeomorphism on a neighborhood $\mathcal{V} \subset \mathcal{U}$ of x_0.

Now $(\mathcal{V}, \psi, y^1, \cdots, y^m)$ is a coordinate system of x_0, where $y^i := f_i$. □

Corollary 1.5

If $f_1, \cdots, f_k$, $k < m$, is an independent set of functions at $x \in \mathcal{M}$, then they form part of a coordinate system on a neighborhood of x. ♡

Proof. Choose a coordinate system $(\mathcal{U}, \varphi, x^1, \cdots, x^m)$ about x. Then $(dx_x^i)_{1 \le i \le m}$ is a basis of $T_x^* \mathcal{M}^m$. Since $f_1, \cdots, f_k$ is an independent set of functions at x, we can choose $m - k$ of the x^i so that $df_{1,x}, \cdots, df_{k,x}, dx_x^{i_1}, \cdots, dx_x^{i_{m-k}}$ is basis of $T_x^* \mathcal{M}^m$. Then apply **Corollary 1.4**. □ □

Corollary 1.6

Let $F : \mathcal{M} \to \mathcal{N}$ be smooth and assume that $F_{*,x} : T_x \mathcal{M} \to T_{F(x)} \mathcal{N}$ is surjective. If $x^1, \cdots, x^n$ form a coordinate system on some neighborhood of $F(x)$, then $x^1 \circ F, \cdots, x^n \circ F$ form part of a coordinate system on some neighborhood of x. ♡

Proof. Consider

$$F_{*,x} : T_x \mathcal{M} \longrightarrow T_{F(x)} \mathcal{N}, \quad F_x^* : T_{F(x)}^* \mathcal{N} \longrightarrow T_x^* \mathcal{M}.$$

If $F_x^*(\omega_1) = F_x^*(\omega_2)$ for $\omega_1, \omega_2 \in T_{F(x)}^* \mathcal{N}$, then, for any $V \in T_x \mathcal{M}$, we have

$$(\omega_1 - \omega_2) F_{*,x}(V) = 0.$$

The surjectivity of $F_{*,x}$ implies that $F_{*,x}(V_0) \neq 0$ for some $V_0 \in T_x\mathcal{M}$ and then $\omega_1 = \omega_2$. Thus F_x^* is injective.

We now can prove that the functions $\{x^i \circ F\}_{1\le i\le n}$ are independent of x. Indeed, let

$$\sum_{1\le i\le n} a_i (x^i \circ F)_{*,x} = 0, \quad a_i \in \mathbf{R}.$$

Since $(x^i \circ F)_{*,x} = F_x^*(dx^i_{F(x)})$, it follows that

$$F_x^*\left(\sum_{1\le i\le m} a_i dx^i_{F(x)}\right) = 0.$$

The injectivity of F_x^* implies $\sum_{1\le i\le m} a_i dx_x^i = 0$ and then $a_1 = \cdots = a_n = 0$. Finally, the result follows from **Corollary 1.5**. □

Corollary 1.7

If $f_1, \cdots, f_k$ is a set of smooth functions on a neighborhood of $x \in \mathcal{M}$ such that $T_x^*\mathcal{M}$ is spanned by $df_1, \cdots, df_k$, then a subset of the f_i forms a coordinate system on a neighborhood of x. ♡

Proof. Observe that $k \ge m$. Then there exist $f_{i_1}, \cdots, f_{i_m}$ so that there form a basis of $T_x^*\mathcal{M}$. Now the result follows from **Corollary 1.4**. □ □

Corollary 1.8

Let $F : \mathcal{M} \to \mathcal{N}$ be smooth and assume that $F_{*,x} : T_x\mathcal{M} \to T_{F(x)}\mathcal{N}$ is injective. If $x^1, \cdots, x^n$ form a coordinate system on a neighborhood of $F(x)$, then a subset of the functions $\{x^i \circ F\}_{1\le i\le n}$ forms a coordinate system on a neighborhood of x. ♡

Proof. Consider

$$F_{*,x} : T_x\mathcal{M} \longrightarrow T_{F(x)}\mathcal{N}, \quad F_x^* : T^*_{F(x)}\mathcal{N} \longrightarrow T_x^*\mathcal{M}.$$

If $U = F_{*,x}(V)$, we define $F^{-1}_{*,x}(U) := V$. Since $F_{*,x}$ is injective, it follows that $F^{-1}_{*,x}$ is well-defined. For $\omega \in T_x^*\mathcal{M}$, define

$$\tau := \omega \circ (F_{*,x})^{-1} \in T^*_{F(x)}\mathcal{N}.$$

Hence $F_x^*(\tau)(V) = \tau(F_{*,x}(V)) = \omega(v)$; thus, F_x^* is surjective.

Since $T^*_{F(x)}\mathcal{N}$ is spanned by $\{dx^i|_x\}_{1\le i\le n}$, the surjectivity of F_x^* implies that $T_x^*\mathcal{M}$ is spanned by $\{F_x^*(dx^i|_x)\}_{1\le i\le n} = \{d(x^i \circ F)_x\}_{1\le i\le n}$. Now the result follows from **Corollary 1.7**. □

Suppose one has a smooth mapping $F : \mathcal{M} \to \mathcal{N}$ factoring through a submanifold $(\mathcal{P}, G)$ of $\mathcal{M}$. That is, $F(\mathcal{N}) \subset G(\mathcal{P})$, whence there is a uniquely defined mapping F_0 of $\mathcal{N}$ into $\mathcal{P}$

such that $G \circ F_0 = F$.

$$\begin{tikzcd} \mathcal{N} \arrow[r, "F"] \arrow[dr, dotted, "\exists! F_0"'] & \mathcal{M} \\ & \mathcal{P} \arrow[u, "G"'] \end{tikzcd}$$

The problem is: When is F_0 smooth? This is certainly not always the case. Let $(\mathbf{R}, F)$ and $(\mathbf{R}, G)$ both be figure-8 submanifolds with precisely the same image sets, but with the difference that as $t \to \pm\infty$, $F(t)$ approaches the intersection along the horizontal direction, but $G(t)$ approaches along the vertical. Suppose also that $F(0) = G(0) = (0, 0)$. Then F_0 is not even continuous since

$$F_0^{-1}(-1, 1) = F^{-1}(G(-1, 1)) = (-\infty, -\alpha) \cup (\alpha, \infty) \cup \{0\}$$

for some $\alpha > 0$.

Theorem 1.5

Suppose that $F : \mathcal{N} \to \mathcal{M}$ is smooth, that $(\mathcal{P}, G)$ is a submanifold of $\mathcal{M}$, and that F factors through $(\mathcal{P}, G)$, that is, $F(\mathcal{N}) \subset G(\mathcal{P})$. Since G is injective, there is a unique mapping F_0 of $\mathcal{N}$ into $\mathcal{P}$ such that $G \circ F_0 = F$.

(a) F_0 is smooth if it is continuous.

(b) F_0 is continuous if G is an embedding.

♡

Proof. (b) is obvious, since $F_0 = G^{-1} \circ F$. So we may assume that F_0 is continuous. It suffices to show that $\mathcal{P}$ can be covered by coordinate systems $(\mathcal{U}, \varphi)$ such that the map $\varphi \circ F_0$ restricted to the open set $F_0^{-1}(\mathcal{U})$ is smooth.

Let $x \in \mathcal{P}$ and let $(\mathcal{V}, \psi)$ be a coordinate system on a neighborhood of $G(x)$ in $\mathcal{M}$. Since $(\mathcal{P}, G)$ is a submanifold, it follows that $G_{*,x}$ is injective. By **Corollary 1.8**, there is a projection $\pi : \mathbf{R}^m \to \mathbf{R}^p$ such that the map $\varphi := \pi \circ \psi \circ G$ is a coordinate system on a neighborhood $\mathcal{U}$ of x. Then

$$\varphi \circ F_{0|F_0^{-1}(\mathcal{U})} = \pi \circ \psi \circ G \circ F_{0|F_0^{-1}(\mathcal{U})} = \pi \circ \psi \circ F_{|F_0^{-1}(\mathcal{U})}$$

is smooth. □

Submanifolds $(\mathcal{N}_1, F_1)$ and $(\mathcal{N}_2, F_2)$ of $\mathcal{M}$ will be called **equivalent** if there exists a diffeomorphism $G : \mathcal{N} \to \mathcal{N}$ such that $F_1 = F_2 \circ G$.

$$\begin{tikzcd} \mathcal{N}_1 \arrow[r, "F_1"] \arrow[dr, dotted, "\exists G"'] & \mathcal{M} \\ & \mathcal{N}_2 \arrow[u, "F_2"'] \end{tikzcd}$$

Observe that $\dim \mathcal{N}_1 = \dim \mathcal{N}_2$. This is an equivalence relation on the collection of all submanifolds of $\mathcal{M}$.

(i) Each equivalence class ξ has a unique representative of the form $(\mathcal{A}, i)$ where $\mathcal{A}$ is a subset of $\mathcal{M}$ with a manifold structure such that the inclusion $i : \mathcal{A} \to \mathcal{M}$ is a smooth immersion.

Proof. For any $(\mathcal{N}, F) \in \xi$, let $(A, i) := (F(\mathcal{N}), i)$, where the manifold structure on $\mathcal{A}$ is induced from the diffeomorphism $F : \mathcal{N} \to F(\mathcal{N})$ and $i : F(\mathcal{N}) \hookrightarrow \mathcal{M}$ is the natural inclusion. Clearly that (A, i) is equivalent to $(\mathcal{N}, F)$.

If $(\mathcal{B}, j)$ is another representative of ξ, then $j = i \circ H$ for some diffeomorphism $H : \mathcal{B} \to \mathcal{A}$. Hence $\mathcal{B}$ admits a manifold structure such that the inclusion $j : \mathcal{B} \to \mathcal{M}$ is a smooth immersion. □ □

(ii) The conclusion of some theorems in the following sections state that there exist *unique* (uniqueness means up to equivalence. In particular, if the submanifolds of $\mathcal{M}$ are viewed as subsets $\mathcal{A} \subset \mathcal{M}^m$ with manifolds structures for which the inclusion maps are smooth immersions, then uniqueness means unique subset with unique second countable locally Euclidean topology and unique differentiable structure) submanifolds satisfying certain conditions.

In the case of a submanifold $(\mathcal{A}, i)$ of $\mathcal{M}$ where i is the inclusion map, we shall often drop the i and simply speak of the submanifold $\mathcal{A} \subset \mathcal{M}$.

(iii) Let $\mathcal{A}$ be a subset of $\mathcal{M}$. Then generally there is not a unique manifold structure on $\mathcal{A}$ such that $(\mathcal{A}, i)$ is a submanifold of $\mathcal{M}$, if there is one at all. However we have the following two uniqueness theorems which involve conditions on the topology on $\mathcal{A}$.

(a) Let $\mathcal{M}$ be a differentiable manifold and $\mathcal{A}$ a subset of $\mathcal{M}$. Fix a topology on $\mathcal{A}$. Then there is at most one differentiable structure on $\mathcal{A}$ such that $(\mathcal{A}, i)$ is a submanifold of $\mathcal{M}$, where i is the inclusion map.

(b) Let $\mathcal{M}$ be a differentiable manifold and $\mathcal{A}$ a subset of $\mathcal{M}$. If in the relative topology, $\mathcal{A}$ has a differentiable structure such that $(\mathcal{A}, i)$ is a submanifold of $\mathcal{M}$, then $\mathcal{A}$ has a unique manifold structure (that is, unique second countable locally Euclidean topology together with a unique differentiable structure) such that $(\mathcal{A}, i)$ is a submanifold of $\mathcal{M}$.

1.1.5 Implicit function theorem

Recall the following implicit function theorem in calculus.

Theorem 1.6. (Implicit function theorem)

Let $U \subset \mathbf{R}^{m-n} \times \mathbf{R}^n$ be open, and let $f : U \to \mathbf{R}^n$ be smooth. We denote the canonical coordinate system on $\mathbf{R}^{m-n} \times \mathbf{R}^n$ by $(r^1, \cdots, r^{m-n}, s^1, \cdots, s^n)$. Suppose that at the point $(\boldsymbol{r}_0, \boldsymbol{s}_0) \in U$, $f(\boldsymbol{r}_0, \boldsymbol{s}_0) = 0$, and that the matrix

$$\left(\left.\frac{\partial f^i}{\partial s^j}\right|_{(\boldsymbol{r}_0, \boldsymbol{s}_0)} \right)_{1 \le i,j \le n}$$

is nonsingular. Then there exists an open neighborhood V of $\boldsymbol{r}_0$ in $\mathbf{R}^{n-m}$ and an open neighborhood W of $\boldsymbol{s}_0$ in $\mathbf{R}^n$ such that $V \times W \subset U$, and there exists a smooth map

$g : V \to W$ *such that for each* $(\boldsymbol{p}, \boldsymbol{q}) \in V \times W$

$$f(\boldsymbol{p}, \boldsymbol{q}) = 0 \Longleftrightarrow \boldsymbol{q} = g(\boldsymbol{p}). \tag{1.1.33}$$

♡

Theorem 1.7

Assume that $F : \mathcal{M} \to \mathcal{N}$ *is smooth, that* y *is a point of* $\mathcal{N}$*, that* $\mathcal{P} := F^{-1}(y)$ *is nonempty, and that* $F_{*,x} : T_x\mathcal{M} \to T_{F(x)}\mathcal{N}$ *is surjective for all* $x \in \mathcal{P}$*. Then* $\mathcal{P}$ *has a unique manifold structure such that* $(\mathcal{P}, i)$ *is a submanifold of* $\mathcal{M}$*, where* i *is the inclusion map. Moreover,* $i : \mathcal{P} \to \mathcal{M}$ *is actually an imbedding and the dimension of* $\mathcal{P}$ *is* $\dim \mathcal{M} - \dim \mathcal{N}$*.* ♡

By **Corollary 1.6**, $m = \dim \mathcal{M} \geq n = \dim \mathcal{N}$.

Proof. By above remarks, it suffices to prove that in the relative topology, $\mathcal{P}$ has a differentiable structure such that $(\mathcal{P}, i)$ is a submanifold of $\mathcal{M}$ of dimension $p := m - n$. It is sufficient to prove that if $x \in \mathcal{P}$, then there exists a coordinate system on a neighborhood $\mathcal{U}$ of x in $\mathcal{M}$ for which $\mathcal{P} \cap \mathcal{U}$ is a coordinate system of the correct dimension. Let $y^1, \cdots, y^n$ be a coordinate system centered at y in $\mathcal{N}$. Since $F_{*,x} : T_x\mathcal{M} \to T_y\mathcal{N}$ is surjective, it follows from **Corollary 1.6** that the collection of functions $(x^i := y^i \circ F)_{1 \leq i \leq n}$ forms part of a coordinate system about $x \in \mathcal{M}$. Complete to a coordinate system $x^1, \cdots, x^n, x^{n+1}, \cdots, x^m$ on a neighborhood $\mathcal{U}$ of x. Then $\mathcal{P} \cap \mathcal{U} = \{x^1 = \cdots = x^n = 0\}$. By **Theorem 1.6**, $\mathcal{P} \cap \mathcal{U}$ is a submanifold of dimension $p = m - n$. □

Theorem 1.8

Assume that $F : \mathcal{M} \to \mathcal{N}$ *is smooth and that* $(\mathcal{O}, G)$ *is a submanifold of* $\mathcal{N}$*. Suppose that whenever* $x \in F^{-1}(G(\mathcal{O}))$*, then*

$$T_{F(x)}\mathcal{N} = F_{*,x}T_x\mathcal{M} + G_{*,G^{-1}(F(x))}T_{G^{-1}(F(x))}\mathcal{O} \tag{1.1.34}$$

(not necessarily a direct sum). Then if $\mathcal{P} := F^{-1}(G(\mathcal{O}))$ *is nonempty,* $\mathcal{P}$ *can be given a manifold structure so that* $(\mathcal{P}, i)$ *is a submanifold of* $\mathcal{M}$*, where* i *is the inclusion map with*

$$\dim \mathcal{P} = \dim \mathcal{M} - (\dim \mathcal{N} - \dim \mathcal{O}). \tag{1.1.35}$$

Moreover, if $(\mathcal{O}, G)$ *is an imbedded submanifold, then so is* $(\mathcal{P}, i)$*, and in this case there is a unique manifold structure on* $\mathcal{P}$ *such that* $(\mathcal{P}, i)$ *is a submanifold of* $\mathcal{M}$*.* ♡

Example 1.3

(a) Define the function $f : \mathbf{R}^{m+1} \to \mathbf{R}$ *by*

$$f(\boldsymbol{p}) := \sum_{1 \leq i \leq m+1} [r^i(\boldsymbol{p})]^2, \quad \boldsymbol{p} \in \mathbf{R}^{m+1}.$$

Then $f^{-1}(1)$ *has a unique manifold structure for which it is a submanifold of* $\mathbf{R}^{m+1}$ *under the inclusion map. This is nothing but the unit* m*-sphere.*

(b) We define a map F *from the general linear group* $\mathrm{GL}(m, \mathbf{R})$ *to the vector space of all*

real symmetric $m \times m$ matrices by

$$F : \mathbf{GL}(m, \mathbf{R}) \longrightarrow \mathbf{Sym}(m, \mathbf{R}), \quad A \longmapsto AA^T,$$

where A^T stands for the transpose of A. For the $m \times m$ identity matrix I_m, let $\mathbf{O}(m, \mathbf{R}) := F^{-1}(I_m)$. $\mathbf{O}(m, \mathbf{R})$ is a subgroup of $\mathbf{GL}(m, \mathbf{R})$ under matrix multiplication called the **orthogonal group**. *We shall check that $\mathbf{O}(m, \mathbf{R})$ has a unique manifold structure such that $(\mathbf{O}(m, \mathbf{R}), i)$ is a submanifold of $\mathbf{GL}(m, \mathbf{R})$ and that in this manifold structure i is an embedding and $\mathbf{O}(m, \mathbf{R})$ has dimension*

$$\dim \mathbf{O}(m, \mathbf{R}) = m^2 - \frac{m(m+1)}{2} = \frac{m(m-1)}{2}.$$

By **Theorem 1.7**, *we suffice to show that $F_{*,\sigma} : T_\sigma \mathbf{GL}(m, \mathbf{R}) \to T_{F(\sigma)} \mathbf{Sym}(m, \mathbf{R})$ is surjective for all $\sigma \in \mathbf{O}(m, \mathbf{R})$. Define the right translation $R_\sigma : \mathbf{GL}(m, \mathbf{R}) \to \mathbf{GL}(m, \mathbf{R})$ by $R_\sigma(\tau) := \tau\sigma^T$. Observe that R_σ is a diffeomorphism. For any $\sigma \in \mathbf{O}(m, \mathbf{R})$, we have*

$$F \circ R_\sigma(\tau) = F(\tau\sigma^T) = \tau\sigma^T(\tau\sigma^T)^T = \tau\sigma^T\sigma\tau^T = F(\tau).$$

Thus $F \circ R_\sigma = F$ and then

$$F_{*,\sigma} = (F \circ R_\sigma)_{*,\sigma} = F_{*,I} \circ (R_\sigma)_{*,\sigma}.$$

Hence we need only to check that $F_{*,I} : T_I \mathbf{GL}(m, \mathbf{R}) \to T_I \mathbf{Sym}(m, \mathbf{R})$ is surjective. Since $\mathbf{GL}(m, \mathbf{R})$ and $\mathbf{Sym}(m, \mathbf{R})$ can be viewed as submanifolds of a large Euclidean space, it suffices to check that F is surjective. For a symmetric $m \times m$ real matrix A with rank r, we have

$$A = B \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} B^T = B \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \cdot \left(B \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \right)^T = F\left(B \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \right)$$

for some matrix $B \in \mathbf{GL}(m, \mathbf{R})$. ♠

1.1.6 Vector fields

A mapping $\sigma : [a, b] \to \mathcal{M}$ is a **smooth curve in** $\mathcal{M}$ if σ can be extends to be a smooth mapping of $(a - \epsilon, b + \epsilon)$ into $\mathcal{M}$ for some $\epsilon > 0$. The curve $\sigma : [a, b] \to \mathcal{M}$ is said to be **piecewise smooth** if there exists a partition

$$a = \alpha_0 < \cdots < \alpha_k = b$$

such that $\sigma_{|[\alpha_i, \alpha_{i+1}]}$ is smooth for $i = 0, \cdots, k-1$.

(a) Piecewise smooth curves are necessarily continuous.

(b) If $\sigma : [a, b] \to \mathcal{M}$ is a smooth curve in $\mathcal{M}$, then its tangent vector

$$\dot{\sigma}(t) := \sigma_{*,t}\left(\left. \frac{d}{dr} \right|_t \right) \in T_{\sigma(t)}\mathcal{M}$$

is well-defined.

Definition 1.10. (Vector fields)

A **vector field $\boldsymbol{X}$ along a (smooth) curve $\boldsymbol{\sigma : [a,b] \to \mathcal{M}}$** *is a mapping $X : [a,b] \to T\mathcal{M}$ which lifts σ, that is,*

$$\pi \circ X = \sigma, \tag{1.1.36}$$

where $\pi : T\mathcal{M} \to \mathcal{M}$ is the natural projection. A vector field X is called a **smooth vector field along $\boldsymbol{\sigma}$** *if the mapping $X : [a,b] \to T\mathcal{M}$ is smooth. A* **vector field $\boldsymbol{X}$ on an open set $\boldsymbol{\mathcal{U}}$ in $\boldsymbol{\mathcal{M}}$** *is a lifting of $\mathcal{U}$ into $T\mathcal{M}$, that is, $X : \mathcal{U} \to T\mathcal{M}$ such that*

$$\pi \circ X = \mathbf{1}_{\mathcal{U}} = \textit{identity map on } \mathcal{U}. \tag{1.1.37}$$

A vector field X is called a **smooth vector field on an open set $\boldsymbol{\mathcal{U}}$** *if the mapping $X : \mathcal{U} \to T\mathcal{M}$ is smooth (written as $X \in C^\infty(\mathcal{U}, T\mathcal{M})$).* ♣

The set $C^\infty(\mathcal{U}, T\mathcal{M})$ of smooth vector fields over $\mathcal{U}$ forms a vector space over $\mathbf{R}$ and a module over the ring $C^\infty(\mathcal{U})$ of smooth functions on $\mathcal{U}$.

If X is a vector field on $\mathcal{U}$ and $x \in \mathcal{U}$, then $X_x := X(x) \in T_x\mathcal{M}$. If f is a smooth function on $\mathcal{U}$, we define

$$X(f) : \mathcal{U} \longrightarrow \mathbf{R}, \quad x \longmapsto X(f)(x) := X_x(f). \tag{1.1.38}$$

Thus $X(f)$ is a function on $\mathcal{U}$.

Proposition 1.3

Let X be a vector field on $\mathcal{M}$. Then the following are equivalent:

(a) X is smooth.

(b) If $(\mathcal{U}, x^1, \cdots, x^m)$ is a coordinate system on $\mathcal{M}$, and if $(a^i)_{1 \le i \le m}$ is the collection of functions on $\mathcal{U}$ defined by

$$X_{|\mathcal{U}} = \sum_{1 \le i \le m} a^i \frac{\partial}{\partial x^i},$$

then $a^i \in C^\infty(\mathcal{U})$.

(c) Whenever $\mathcal{V}$ is open in $\mathcal{M}$ and $f \in C^\infty(\mathcal{V})$, then $X(f) \in C^\infty(\mathcal{V})$. ♡

Proof. $(a) \Rightarrow (b)$: If X is smooth, then $X_{|\mathcal{U}}$ is also smooth. Since $dx^i : \pi^{-1}(\mathcal{U}) \to \mathbf{R}$ is smooth, it follows that

$$dx^i \circ X_{|\mathcal{U}} = \sum_{1 \le j \le m} a^j dx^i \circ \frac{\partial}{\partial x^j} = \sum_{1 \le j \le m} a^j \delta_{ij} = a_i$$

which is smooth on $\mathcal{U}$.

$(b) \Rightarrow (c)$: Let $(\mathcal{U}, x^1, \cdots, x^m)$ be a coordinate system on $\mathcal{M}$ with $\mathcal{U} \subset \mathcal{V}$. By (b), we arrive at

$$X(f)_{|\mathcal{U}} = \sum_{1 \le i \le m} a^i \frac{\partial f}{\partial x^i}$$

and a^i are all smooth on $\mathcal{U}$. Hence $X(f)$ is smooth on $\mathcal{U}$.

$(c) \Rightarrow (a)$: It suffices to prove that $X_{|\mathcal{U}}$ is smooth where $(\mathcal{U}, x^1, \cdots, x^m)$ is an arbitrary coordinate system on $\mathcal{M}$. To prove that $X_{|\mathcal{U}}$ is smooth, we need only to check that $X_{|\mathcal{U}}$ composed with the canonical coordinate functions on $\pi^{-1}(\mathcal{U})$ (for $T\mathcal{M}$) are smooth functions. Because

$$x^i \circ \pi \circ X_{|\mathcal{U}} = x^i, \quad dx^i \circ X_{|\mathcal{U}} = X(x^i),$$

we conclude that $X_{|\mathcal{U}}$ is smooth. □

If X, Y are smooth vector fields on $\mathcal{M}$, we define the **Lie bracket of X and Y** by

$$[X, Y]_x(f) := X_x(Yf) - Y_x(Xf) \tag{1.1.39}$$

where $x \in \mathcal{M}$ and f is any smooth function near x.

Proposition 1.4

Suppose X, Y, Z are smooth vector fields on $\mathcal{M}$.

(a) $[X, Y]$ *is a smooth vector field on* $\mathcal{M}$.

(b) If $f, g \in C^\infty(\mathcal{M})$*, then*

$$[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X. \tag{1.1.40}$$

(c) **(Jacobi identity)** $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0.$

(d) $[X, Y] = -[Y, X].$

♡

A vector space V with a bilinear operation $[\cdot, \cdot]$ satisfying (c) and (d) is called a **Lie algebra**. Consequently, $C^\infty(\mathcal{M}, T\mathcal{M}))$ together with the Lie bracket is a Lie algebra. In part (b), fX is a smooth vector field on $\mathcal{M}$ defined by

$$(fX)_x := f(x)X_x, \quad x \in \mathcal{M}^m. \tag{1.1.41}$$

Proof. (a) By **Proposition 1.3**, $[X, Y]$ is smooth.

(b) If $f, g \in C^\infty(\mathcal{M})$, then

$$[fX, gY]_x(h) = (fX)_x((gY)(h)) - (gY)_x((fX)(h)), \quad x \in \mathcal{M},$$

for any smooth function h near x. Using the fact that

$$(gY)(h)(p) = (gY)_p(h) = g(p)Y_p(h) = g(p)(Y(h))(p) = (g \cdot Y(h))(p),$$

we arrive at

$$\begin{aligned}
[fX, gY]_x(h) &= (f(x)X_x)(g \cdot Y(h)) - (g(x)Y_x)(f \cdot X(h)) \\
&= f(x)\,(g(x)X_x(Y(h)) + Y(h)(x)X_x(g)) \\
&\quad - g(x)\,(f(x)Y_x(X(h)) + X(h)(x)Y_x(f)) \\
&= f(x)g(x)\,[X_x(Y(h)) - Y_x(X(h))] \\
&\quad + f(x)(Y(h)(x))X_x(g) - g(x)(X(h)(x))Y_x(f) \\
&= (fg)(x)[X, Y]_x(h) + f(x)Y_x(h)X_x(g) - g(x)X_x(h)Y_x(f) \\
&= (fg[X, Y])_x(h) + (f(Xg)Y)_x(h) - (g(Yf)X)_x(h).
\end{aligned}$$

Thus $[fX, gY] = fg[X,Y] + f(Xg)Y - g(Yf)X$.

(c) It follows from part (b).

(d) For any $x \in \mathcal{M}$ and any function f smooth near x, we have

$$([X,Y]+[Y,X])_x(f) = (X_x(Yf) - Y_x(Xf)) + (Y_x(Xf) - X_x(Yf)) = 0.$$

Thus $[X,Y] = -[Y,X]$. □

Definition 1.11. (Integral curves)

Let X be a smooth vector field on $\mathcal{M}$. A smooth curve σ in $\mathcal{M}$ is an **integral curve of X** *if*

$$\dot{\sigma}(t) = X_{\sigma(t)} \tag{1.1.42}$$

for each t in the domain of σ. ♣

Let X be a smooth vector field on $\mathcal{M}$ and $x \in \mathcal{M}$. A smooth curve $\gamma : (a,b) \to \mathcal{M}^m$ is an integral curve of X if and only if

$$\gamma_{*,t}\left(\left.\frac{d}{dr}\right|_t\right) = X_{\gamma(t)}, \quad a < t < b. \tag{1.1.43}$$

Suppose that $0 \in (a,b)$ and $\gamma(0) = x$. Choose a coordinate system $(\mathcal{U}, \varphi)$ with coordinate functions $x^1, \cdots, x^m$ about x. By **Proposition 1.3**,

$$X_{|\mathcal{U}} = \sum_{1\le i\le m} f^i \frac{\partial}{\partial x^i} \tag{1.1.44}$$

where the f^i are smooth functions on $\mathcal{U}$. Moreover, for each t such that $\gamma(t) \in \mathcal{U}$, we get

$$\gamma_{*,t}\left(\left.\frac{d}{dr}\right|_t\right) = \sum_{1\le i\le m} \left.\frac{d(x^i\circ\gamma)}{dr}\right|_t \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)}. \tag{1.1.45}$$

Thus, from (1.1.45) and (1.1.44), the equation (1.1.43) becomes

$$\sum_{1\le i\le m} \left.\frac{d(x^i\circ\gamma)}{dr}\right|_t \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)} = \sum_{1\le i\le m} f^i(\gamma(t)) \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)}. \tag{1.1.46}$$

Hence, γ is an integral curve of X on $\gamma^{-1}(\mathcal{U})$ if and only if

$$\left.\frac{d(x^i\circ\gamma)}{dr}\right|_t = f^i(\gamma(t)), \quad 1 \le i \le m, \ t \in \gamma^{-1}(\mathcal{U}). \tag{1.1.47}$$

If we define $\gamma^i := x^i \circ \gamma = r^i \circ \varphi \circ \gamma$, then $\gamma(t) = \varphi^{-1}(\gamma^1(t), \cdots, \gamma^m(t))$ and the equation (1.1.47) becomes

$$\left.\frac{d\gamma^i}{dr}\right|_t = f^i \circ \varphi^{-1}(\gamma^1(t), \cdots, \gamma^m(t)), \quad 1 \le i \le m, \ t \in \gamma^{-1}(\mathcal{U}). \tag{1.1.48}$$

Observe that equation (1.1.48) is a system of first order ordinary differential equations.

Theorem 1.9

Let X be a smooth vector field on a manifold $\mathcal{M}$. For each $x \in \mathcal{M}$ there exists $a(x), b(x) \in$ $\mathbf{R} \cup \{\pm\infty\}$ and a smooth curve

$$\gamma_x : (a(x), b(x)) \longrightarrow \mathcal{M} \tag{1.1.49}$$

such that

(a) $0 \in (a(x), b(x))$ *and* $\gamma_x(0) = x$.

(b) γ_x *is an integral curve of* X.

(c) If $\mu : (c,d) \to \mathcal{M}$ *is a smooth curve satisfying conditions (a) and (b), then* $(c,d) \subset (a(x), b(x))$ *and* $\mu = \gamma_{|(c,d)}$. ♡

For each $t \in \mathbf{R}$, we define a transformation X_t with domain

$$\mathscr{D}_t := \{x \in \mathcal{M} : t \in (a(x), b(x))\} \tag{1.1.50}$$

by

$$X_t(x) := \gamma_x(t). \tag{1.1.51}$$

Theorem 1.10

Let X *be a smooth vector field on a manifold* $\mathcal{M}$ *and* γ_x *be obtained from* **Theorem 1.9**.

(d) For each $x \in \mathcal{M}$*, there exists an open neighborhood* $\mathcal{V}$ *of* x *and an* $\epsilon > 0$ *such that the map*

$$(t, p) \longmapsto X_t(p) \tag{1.1.52}$$

is defined and is smooth from $(-\epsilon, \epsilon) \times \mathcal{V}$ *into* $\mathcal{M}$.

(e) $\mathscr{D}_t$ *is open for each* t.

(f) $\cup_{t \geq 0} \mathscr{D}_t = \mathcal{M}$.

(g) $X_t : \mathscr{D}_t \to \mathscr{D}_{-t}$ *is a diffeomorphism with inverse* X_{-t}.

(h) Let s *and* t *be real numbers. Then the domain of* $X_s \circ X_t$ *is contained in but generally not equal to* $\mathscr{D}_{s+t}$*. However, the domain of* $X_s \circ X_t$ *is* $\mathscr{D}_{s+t}$ *in the case in which* s *and* t *both have the same sign. Moreover, on the domain of* $X_s \circ X_t$ *we have*

$$X_s \circ X_t = X_{s+t}. \tag{1.1.53}$$ ♡

Proof. Let $(a(x), b(x))$ be the union of all the open intervals which contain 0 and which are domains of integral curves of X satisfying the initial condition that the origin maps to x.

(f) Since $(a(x), b(x)) \neq \emptyset$, it follows that $\cup_{t \geq 0} \mathscr{D}_t = \mathcal{M}$.

(d) By the differentiability of the solutions of (1.1.48).

(h) Let $t \in (a(x), b(x))$. Then $s \mapsto \gamma_x(t+s)$ is an integral curve of X with the initial condition $0 \mapsto \gamma_x(t)$ and with maximal domain $(a(x) - t, b(x) - t)$. By the uniqueness, we obtain

$$(a(x) - t, b(x) - t) = (a(\gamma_x(t)), b(\gamma_x(t))), \tag{1.1.54}$$

and for s in the interval (1.1.54)

$$\gamma_{\gamma_x(t)}(s) = \gamma_x(t+s). \tag{1.1.55}$$

If x belong to the domain of $X_s \circ X_t$, then $t \in (a(x), b(x))$ and $s \in (a(\gamma_x(t)), b(\gamma_x(t)))$, so by (1.1.54), $s + t \in (a(x), b(x))$. Thus $x \in \mathscr{D}_{s+t}$.

The following example shows that the domain of $X_s \circ X_t$ is generally not equal to $\mathscr{D}_{s+t}$. Consider the vector field $\partial/\partial r^1$ on $\mathcal{M} := \mathbf{R}^2 \setminus \{\mathbf{0}\}$ with $s = -1$ and $t = 1$.

If s and t both have the same sign and if $x \in \mathscr{D}_{s+t}$, that is, $s + t \in (a(x), b(x))$, then $t \in (a(x), b(x))$ (in fact, if $s, t > 0$, then $t < s + t < b(x)$; if $t \leq a(x)$, we must have $t < 0$, a contradiction. Hence $t \in (a(x), b(x))$. Similarly, we can prove the same conclusion when $s, t < 0$) and, by (1.1.54), $s \in (a(\gamma_x(t)), b(\gamma_x(t)))$. Hence x is in the domain of $X_s \circ X_t$.

Part (e) and (g) are trivial if $t = 0$, so we may assume that $t > 0$ and that $x \in \mathscr{D}_t$. By part (d) and the compactness of $[0, t]$, there exists a neighborhood $\mathcal{W}$ of $\gamma_x([0, t])$ and an $\epsilon > 0$ such that the map (1.1.52) is defined and is smooth from $(-\epsilon, \epsilon) \times \mathcal{W}$ into $\mathcal{M}$. Choose a positive integer k so that $t/k \in (-\epsilon, \epsilon)$. Let $\alpha_1 := X_{t/k|\mathcal{W}}$ and let $\mathcal{W}_1 := \alpha_1^{-1}(\mathcal{W})$. Then for $i = 2, \cdots, k$ we define

$$\alpha_i := X_{t/k|\mathcal{W}_{i-1}}, \quad \mathcal{W}_i := \alpha_i^{-1}(\mathcal{W}_{i-1}).$$

Observe that α_i is a smooth map on the open set $\mathcal{W}_{i-1} \subset \mathcal{W}$ (where $\mathcal{W}_0 := \mathcal{W}$). It follows that $\mathcal{W}_k$ is an open subset of $\mathcal{W}$, that $\mathcal{W}_k$ contains x, and that by part (h),

$$\alpha_1 \circ \alpha_2 \circ \cdots \circ \alpha_{k|\mathcal{W}_k} = X_{t|\mathcal{W}_k}. \tag{1.1.56}$$

Consequently, $\mathcal{W}_k \subset \mathscr{D}_t$; hence $\mathscr{D}_t$ is open.

(g) Since $0 \in (a(x), b(x))$, it follows that

$$-t \in (a(x) - t, b(x) - t) = (a(\gamma_x(t)), b(\gamma_x(t)))$$

by (1.1.54). Thus X_t is a map of $\mathscr{D}_t$ to $\mathscr{D}_{-t}$. Using (1.1.53), we see that the inverse of X_t is X_{-t}. The smoothness of X_t follows from (1.1.56). Hence X_t is a diffeomorphism from $\mathscr{D}_t$ to $\mathscr{D}_{-t}$. □

Definition 1.12. (Complete vector fields)

A smooth vector field X on $\mathcal{M}$ is **complete** *if $\mathscr{D}_t = \mathcal{M}$ for all t (that is, the domain of γ_x is $\mathbf{R}$ for each $x \in \mathcal{M}$). In this case, the transformations X_t form a group of transformations of $\mathcal{M}$ parametrized by the real numbers called the* **1-parameter group of X**. ♣

In the case that $\mathcal{M}$ is compact, any smooth vector field on $\mathcal{M}$ is complete. If X is not complete, the transformations X_t do not form a group since their domains depend on t. In this case, we shall refer to the collection of transformations X_t as the local $\mathbf{1}$-parameter group of $\boldsymbol{X}$.

Example 1.4

An example of non-complete vector field is the vector field $\partial/\partial r^1$ on $\mathbf{R}^2 \setminus \{\mathbf{0}\}$. If $a > 0$, the domain of the maximal integral curve through $(a, 0)$ is $(-a, \infty)$. ♠

Definition 1.13. (Local smooth extensions)

Let $F : \mathcal{M} \to \mathcal{N}$ be smooth. A smooth vector field X along F (that is, $X \in C^\infty(\mathcal{M}, T\mathcal{N})$ and $\pi \circ X = F$) has **local smooth extension in $\mathcal{N}$** *if given $x \in \mathcal{M}$ there exist a* *neighborhood $\mathcal{U}$ of x and a neighborhood $\mathcal{V}$ of $F(x)$ such that $F(\mathcal{U}) \subset \mathcal{V}$, and there also exists a smooth vector field $\widetilde{X}$ on $\mathcal{V}$ such that*

$$\widetilde{X} \circ F_{|\mathcal{U}} = X_{|\mathcal{U}}. \tag{1.1.57}$$

Note 1.4

(1) If $F : \mathcal{M} \to \mathcal{N}$ is an immersion, then any smooth vector field along F has local smooth extension in $\mathcal{N}$.

(2) If F is not immersion, such extensions may not exist. Let

$$\alpha : \mathbf{R} \longrightarrow \mathbf{R}, \quad t \longmapsto t^3$$

and let

$$X(t) := \dot{\alpha}(t) = \alpha_{*,t}\left(\left.\frac{d}{dr}\right|_t\right).$$

Since α is a homeomorphism, there is a vector field $\widetilde{X}$ on $\mathbf{R}$ so that the following diagram commutes.

$$\begin{tikzcd} T\mathbf{R}^1 \arrow[r, "d\alpha"] & T\mathbf{R}^1 \\ \mathbf{R} \arrow[u, "\frac{d}{dr}"] \arrow[ur, "X"] \arrow[r, "\alpha"'] & \mathbf{R} \arrow[u, dotted, "\widetilde{X}"'] \end{tikzcd}$$

Now, X is a smooth vector field along α, but $\widetilde{X}$ is not a smooth vector field on $\mathbf{R}$. Indeed, letting $u := t^3$ yields

$$\widetilde{X}_u = \widetilde{X}_{\alpha(t)} = X(t) = \alpha_{*,t}\left(\left.\frac{d}{dr}\right|_t\right) = \left.\frac{d\alpha}{dr}\right|_t \left.\frac{d}{dr}\right|_{\alpha(t)} = 3t^2 \left.\frac{d}{dr}\right|_u = 3u^{2/3}\left.\frac{d}{dr}\right|_u.$$

Thus $\widetilde{X}(t) = 3t^{2/3}\frac{d}{dr}|_t$ and the function $t^{2/3}$ is not differentiable at the origin.

Proposition 1.5

Let $x \in \mathcal{M}$ and let X be a smooth vector field on $\mathcal{M}$ such that $X(x) \neq 0$. Then there exists a coordinate system $(\mathcal{U}, \varphi)$ with coordinate functions $x^1, \cdots, x^m$ on a neighborhood of x such that

$$X_{|\mathcal{U}} = \left.\frac{\partial}{\partial x^1}\right|_{\mathcal{U}}. \tag{1.1.58}$$

Proof. Since $X_x = X(x) \neq 0$, we can choose a coordinate system $(\mathcal{V}, \psi)$ centered at x with coordinate functions $y^1, \cdots, y^m$ such that

$$X_x = \left.\frac{\partial}{\partial y^1}\right|_x.$$

From **Theorem 1.10** (d), there exists an $\epsilon > 0$ and a neighborhood W of the origin in $\mathbf{R}^{m-1}$ such that the map

$$\sigma(t, a^2, \cdots, a^m) := X_t\left(\psi^{-1}(0, a^2, \cdots, a^m)\right)$$

is defined and is smooth for $(t, a^2, \cdots, a^m) \in (-\epsilon, \epsilon) \times W \subset \mathbf{R}^m$. Since

$$\sigma_{*,\mathbf{0}}\left(\frac{\partial}{\partial r^1}\bigg|_{\mathbf{0}}\right) = X_x = \frac{\partial}{\partial y^1}\bigg|_x, \quad \sigma_{*,\mathbf{0}}\left(\frac{\partial}{\partial r^i}\bigg|_{\mathbf{0}}\right) = \frac{\partial}{\partial y^i}\bigg|_x, \ 2 \le i \le m,$$

it follows that σ is nonsingular and that $\varphi := \sigma^{-1}$ is a coordinate map on some neighborhood $\mathcal{U}$ of x. Let $x^1, \cdots, x^m$ denote the coordinate functions of the coordinate system $(\mathcal{U}, \varphi)$. Then since

$$\sigma_{*,(t,a^2,\cdots,a^m)}\left(\frac{\partial}{\partial r^1}\bigg|_{(t,a^2,\cdots,a^m)}\right) = X_{\sigma(t,a^2,\cdots,a^m)},$$

we have $X_{|\mathcal{U}} = \frac{\partial}{\partial x^1}|_{\mathcal{U}}$. In fact, for any $p \in \mathcal{U}$, we write $\varphi(p) = \sigma^{-1}(p) = (t, a^2, \cdots, a^m)$ and hence

$$X_p = \sigma_{*,\varphi(p)}\left(\frac{\partial}{\partial r^1}\bigg|_{\varphi(p)}\right) = \sum_{1\le i\le m} \frac{\partial}{\partial r^1}\bigg|_{\varphi(p)} x^i \frac{\partial}{\partial x^i}\bigg|_p = \sum_{1\le i\le m} \delta_{1i}\frac{\partial}{\partial x^i}\bigg|_p = \frac{\partial}{\partial x^1}\bigg|_p.$$

Since p were arbitrary, we obtain the desired result. $\square$

Definition 1.14. (F-related vector fields))

Let $F : \mathcal{M} \to \mathcal{N}$ be smooth. Smooth vector fields X on $\mathcal{M}$ and Y on $\mathcal{N}$ are called ***F*-related** *if $F_* \circ X = Y \circ F$.*

$$\begin{tikzcd} \mathcal{M} \arrow[r, "F"] \arrow[d, "X"'] & \mathcal{N} \arrow[d, "Y"] \\ T\mathcal{M} \arrow[r, "F_*"'] & T\mathcal{N} \end{tikzcd}$$

If $x \in \mathcal{M}$, then

$$Y_{F(x)} = F_{*,x}(X_x).$$

♣

Proposition 1.6

Let $F : \mathcal{M} \to \mathcal{N}$ be smooth. Let X_1 and X_2 be smooth vector fields on $\mathcal{M}$ and let Y_1 and Y_2 be smooth vector fields on $\mathcal{N}$. If X_1 is F-related to Y_1 and if X_2 is F-related to Y_2, then $[X_1, X_2]$ is F-related to $[Y_1, Y_2]$. ♡

Proof. For any $x \in \mathcal{M}$ and $f \in C^\infty(\mathcal{N})$, we have

$$\begin{aligned}
F_{*,x}([X_1, X_2]_x)(f) &= [X_1, X_2]_x(f \circ F) \\
&= (X_1)_x(X_2(f \circ F)) - (X_2)_x(X_1(f \circ F)) \\
&= (X_1)_x((F_* \circ X_2)(f)) - (X_2)_x((F_* \circ X_1)(f)) \\
&= (X_1)_x((Y_2 \circ F)(f)) - (X_2)_x((Y_1 \circ F)(f)) \\
&= (X_1)_x(Y_2(f) \circ F) - (X_2)_x(Y_1(f) \circ F) \\
&= F_{*,x}((X_1)_x)(Y_2(f)) - F_{*,x}((X_2)_x)(Y_1(f)) \\
&= (Y_1)_{F(x)}(Y_2(f)) - (Y_2)_{F(x)}(Y_1(f)) \ = \ [Y_1, Y_2]_{F(x)}(f),
\end{aligned}$$

where we used $(Y_i \circ F)(f) = Y_i(f) \circ F$. $\square$

1.2 Tensors and forms

Introduction

- *Tensor and exterior algebras*
- *Tensor fields and differential forms*
- *Lie derivatives*

1.2.1 Tensor and exterior algebras

Throughout this subsection, U, V, W are finite dimensional real vector spaces, and V^* stands for the dual space of V.

Definition 1.15

Let $\mathscr{F}(V, W)$ be the free vector space over $\mathbf{R}$ *whose generators are the points of $V \times W$. That is, $\mathscr{F}(V, W)$ consists of all finite linear combinations of elements of $V \times W$. Let $\mathscr{R}(V, W)$ be the subspace of $\mathscr{F}(V, W)$ generated by the set of all elements of $\mathscr{F}(V, W)$ of the following forms:*

$$(v_1 + v_2, w) - (v_1, w) - (v_2, w), \quad (v, w_1 + w_2) - (v_1, w_1) - (v_1, w_2)$$

and

$$(av, w) - a(v, w), \quad (v, aw) - a(v, w)$$

whenever $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, and $a \in \mathbf{R}$. ♣

The quotient space

$$V \otimes W := \mathscr{F}(V, W)/\mathscr{R}(V, W) \tag{1.2.1}$$

is called the **tensor product of V and W**. The coset of $V \otimes W$ containing (v, w) is denoted by $v \otimes w$. Clearly that

$$\begin{aligned}
(v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w, \\
v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2, \\
a(v \otimes w) &= av \otimes w = v \otimes aw.
\end{aligned}$$

Proposition 1.7. (Universal mapping property)

Let $\varphi : V \times W \to V \otimes W$ denote the bilinear map $(v, w) \mapsto v \otimes w$. Then whenever U is a vector space and $\ell : V \times W \to U$ is a bilinear map, there exists a unique linear map $\tilde{\ell} : V \otimes W \to U$ such that the following diagram commutes:

$$\begin{tikzcd}
V \otimes W \arrow[dr, dotted, "\tilde{\ell}"] & \\
V \times W \arrow[u, "\varphi"] \arrow[r, "\ell"'] & U
\end{tikzcd}$$

The pair $(V \otimes W, \varphi)$ is said to solve the **universal mapping problem** *for bilinear maps* *with domain $V \times W$. Moreover, $(V \otimes W, \varphi)$ is unique with this property in the sense that if X is a vector space and $\tilde{\varphi} : V \times W \to X$ a bilinear map with the above universal mapping property, then there exists an isomorphism $\alpha : V \otimes W \to X$ such that $\alpha \circ \varphi = \tilde{\varphi}$.* ♡

Proof. Define

$$\tilde{\ell} : V \otimes W \longrightarrow U, \quad v \otimes w \longmapsto \ell(v, w).$$

Since ℓ is bilinear, the above mapping is well-defined and $\tilde{\ell} \circ \varphi = \ell$. On the other hand, there exists a unique bilinear map $\bar{\ell} : \mathscr{F}(V, W) \to U$, because $\mathscr{F}(V, W)$ is a free module. Hence $\mathscr{R}(V, W) \subset \ker(\bar{\ell})$ and then there exists a unique bilinear map $\tilde{\ell} : V \otimes W = \mathscr{F}(V, W)/\mathscr{R}(V, W) \to U$.

If X and $\tilde{\varphi}$ are above, consider

$$\begin{tikzcd} V \otimes W \arrow[dr, dotted, "\exists! \alpha"] & \\ V \otimes W \arrow[u, "\varphi"] \arrow[r, "\tilde{\varphi}"'] & X \end{tikzcd} \qquad \begin{tikzcd} X \arrow[dr, dotted, "\exists! \beta"] & \\ V \times W \arrow[u, "\tilde{\varphi}"] \arrow[r, "\varphi"'] & V \otimes W \end{tikzcd}$$

where

$$\alpha \circ \varphi = \tilde{\varphi}, \quad \beta \circ \tilde{\varphi} = \varphi.$$

Since

$$\alpha \circ \beta \circ \tilde{\varphi} = \alpha \circ \varphi = \tilde{\varphi}, \quad \beta \circ \alpha \circ \varphi = \beta \circ \tilde{\varphi} = \varphi,$$

it follows from the uniqueness that $\alpha \circ \beta = \mathbf{1}_X$, $\beta \circ \alpha = \mathbf{1}_{V \otimes W}$. So α is an isomorphism. □

Note 1.5

(a) *$V \otimes W$ is canonically isomorphic to $W \otimes V$.*

Proof. Consider

$$\begin{tikzcd} V \otimes W \arrow[dr, dotted, "\tilde{f}"] & \\ V \times W \arrow[u, "\varphi"] \arrow[r, "f"'] & W \otimes V \end{tikzcd}$$

the bilinear map $f(v, w) := w \otimes v$. By **Proposition 1.7**, there exists a unique linear map $\tilde{f} : V \otimes W \to W \otimes V$ such that $\tilde{f} \circ \varphi = f$ (and then $\tilde{f}(v \otimes w) = w \otimes v$. Similarly, there exists a linear map $\tilde{g} : W \otimes V \to V \otimes V$ with $\tilde{g}(w \otimes v) = v \otimes w$. Then

$$\begin{aligned} \tilde{g} \circ \tilde{f}(v \otimes w) &= \tilde{g}(w \otimes v) &= v \otimes w \\ \tilde{f} \circ \tilde{g}(w \otimes v) &= \tilde{f}(v \otimes w) &= w \otimes v. \end{aligned}$$

Thus $\tilde{g} \circ \tilde{f} = \mathbf{1}_{V \otimes W}$ and $\tilde{f} \circ \tilde{g} = \mathbf{1}_{W \otimes V}$. □

(b) *$V \otimes (W \otimes U)$ is canonically isomorphic to $(V \otimes W) \otimes U$.*

Proof. We first claim that there exists a bilinear map $\Phi : (V \otimes W) \otimes U \to V \otimes (W \otimes U)$.

For any $u \in U$, define

$$f_u : V \times W \longrightarrow V \otimes (W \otimes U), \quad (v, w) \longmapsto v \otimes (w \otimes u).$$

Then f_u is a bilinear map and, by **Proposition 1.7**, we have a unique bilinear map $\bar{f}_u : V \otimes W \to V \otimes (W \otimes U)$ such that

$$\bar{f}_u(v \otimes w) = v \otimes (w \otimes u).$$

On the other hand, define

$$g : (V \otimes W) \times U \longmapsto V \otimes (W \otimes U), \quad \left(\sum_{i \in I} v_i \otimes w_i, u\right) \longmapsto \bar{f}_u\left(\sum_{i \in I} v_i \otimes w_i\right).$$

Then g is a bilinear map and there exists a bilinear map $\Phi : (V \otimes W) \otimes U \to V \otimes (W \otimes U)$ such that

$$\Phi\left(\left(\sum_{i \in I} v_i \otimes w_i\right) \otimes u\right) = \sum_{i \in I} v_i \otimes (w_i \otimes u).$$

Similarly, we have a bilinear map

$$\Psi : V \otimes (W \otimes U) \longrightarrow (V \otimes W) \otimes U$$

satisfying

$$\Psi\left(v \otimes \left(\sum_{i \in I} w_i \otimes u_i\right)\right) = \sum_{i \in I} (v \otimes w_i) \otimes u_i.$$

Clearly that $\Phi \circ \Psi = \mathbf{1}_{V \otimes (W \otimes U)}$ and $\Psi \circ \Phi = \mathbf{1}_{(V \otimes W) \otimes U}$. □

(c) $V^* \otimes W \cong \mathrm{Hom}(V, W)$. Consider

$$\begin{tikzcd}
V^* \otimes W \arrow[dr, dotted, "\alpha"] & \\
V^* \times W \arrow[u, "\varphi"] \arrow[r, "\psi"] & \mathrm{Hom}(V, W)
\end{tikzcd}$$

where

$$\psi(f, w)(v) := f(v)w, \quad v \in V, \ w \in W, \ f \in V^*.$$

Observe that ψ is a bilinear map. From **Proposition 1.7**, *there exists a linear map $\alpha : V^* \otimes W \to \mathrm{Hom}(V, W)$ such that*

$$\alpha(f \otimes w)(v) = f(v)w.$$

Let $\alpha(f \otimes w) = 0$. Then $f(v)w = 0$ for all $v \in V$. If $w \neq 0$, then $f(v) = 0$. Consequently, $f \otimes w = 0$. Thus $\ker \alpha = 0$. To prove the surjectivity, choose any $F \in \mathrm{Hom}(V, W)$. If $e_1, \cdots, e_n$ is a basis of V, where $n = \dim V$, then

$$\alpha\left(\sum_{1 \le i \le n} r_i \otimes F(e_i)\right)(v) = \sum_{1 \le i \le n} r_i(v) F(e_i) = \sum_{1 \le i \le n} v^i F(e_i) = F(v)$$

where $r_i : V \to \mathbf{R}$ is the map given by $r_i(v) = v^i$ for $v = \sum_{1 \le i \le n} v^i e_i$. Thus α is *surjective. Finally,*

$$\dim(V^* \otimes W) = \dim \operatorname{Hom}(V, W)$$

and in particular

$$\dim(V \otimes W) = \dim V \cdot \dim W. \tag{1.2.2}$$

(d) Let $(e_i)_{1\le i\le \dim V}$ and $(f_j)_{1\le j\le \dim W}$ be bases for V and W respectively. Then $(e_i \otimes f_j)_{1\le i\le \dim V, 1\le j\le \dim W}$ is a basis of $V \otimes W$. ♣

Definition 1.16. (Tensors)

The **tensor space $\mathscr{T}^{r,s}V$ of type (r,s)** *associated with V is the vector space*

$$\underbrace{V \otimes \cdots \otimes V}_{r} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s}.$$

The direct sum

$$\mathscr{T}V := \bigoplus_{r,s\ge 0} \mathscr{T}^{r,s}V \tag{1.2.3}$$

where $\mathscr{T}^{0,0}V := \mathbf{R}$, is called the **tensor algebra of V**. *Elements of $\mathscr{T}V$ are finite linear combinations over $\mathbf{R}$ of elements of the various $\mathscr{T}^{r,s}V$ and are called* **tensors**. ♣

$\mathscr{T}V$ is a non-commutative, associative, graded algebra under $\otimes$ multiplication: if

$$\begin{aligned} u &= u_1 \otimes \cdots \otimes u_{r_1} \otimes u_1^* \otimes \cdots \otimes u_{s_1}^* \in \mathscr{T}^{r_1,s_1}V, \\ v &= v_1 \otimes \cdots \otimes v_{r_2} \otimes v_1^* \otimes \cdots \otimes v_{s_2}^* \in \mathscr{T}^{r_2,s_2}V \end{aligned}$$

their product $u \otimes v$ is defined by

$$u \otimes v := u_1 \otimes \cdots \otimes u_{r_1} \otimes v_1 \otimes \cdots \otimes v_{r_2} \otimes u_1^* \otimes \cdots \otimes u_{s_1}^* \otimes v_1^* \otimes \cdots \otimes v_{s_2}^* \in \mathscr{T}^{r_1+r_2, s_1+s_2}V.$$

Tensors in $\mathscr{T}^{r,s}$ are called **homogeneous of degree (r,s)**. A homogeneous tensor of degree (r,s) is called **decomposable** it if can be written as

$$v_1 \otimes \cdots \otimes v_r \otimes v_1^* \otimes \cdots \otimes v_s^*$$

where $v_i \in V$, $1 \le i \le r$, and $v_j^* \in V^*$, $1 \le j \le s$.

Definition 1.17. (Exterior algebras)

Denote

$$\mathscr{C}V := \bigoplus_{k\ge 0} \mathscr{T}^{k,0}V \tag{1.2.4}$$

the subalgebra of $\mathscr{T}V$. Let $\mathscr{I}V$ be the two-sided ideal in $\mathscr{C}V$ generated by the set of elements of the form $v \otimes v$ for $v \in V$ and set

$$\mathscr{I}^kV := \mathscr{I}(V) \cap \mathscr{T}^{k,0}V. \tag{1.2.5}$$

Observe that

$$\mathscr{I}V = \bigoplus_{k\geq 0} \mathscr{I}^k\mathcal{V} \tag{1.2.6}$$

and is a graded ideal in $\mathscr{C}V$. *The* **exterior algebra** $\wedge V$ *of* V *is the graded algebra*

$$\wedge V := \mathscr{C}V/\mathscr{I}V. \tag{1.2.7}$$

If we set

$$\wedge^k V := \mathscr{T}^{k,0}V/\mathscr{I}^k(V)\ (k \geq 2), \quad \wedge_1(V) := V, \quad \wedge_0(V) := \mathbf{R}, \tag{1.2.8}$$

then

$$\wedge V = \bigoplus_{k\geq 0} \wedge^k V. \tag{1.2.9}$$

♣

The multiplication in the algebra $\wedge V$ is denoted by $\wedge$ and is called the **wedge** or **exterior product**. In particular, the coset containing $v_1 \otimes \cdots \otimes v_k$ is $v_1 \wedge \cdots \wedge v_k$.

Definition 1.18. (Alternative maps)

A multi-linear map

$$h : \underbrace{V \times \cdots \times V}_{r} \longrightarrow W$$

is called **alternative** *if*

$$h(v_{\pi(1)}, \cdots, v_{\pi(r)}) = \operatorname{sgn}(\pi) h(v_1, \cdots, v_r), \quad v_1, \cdots, v_r \in V$$

for all permutations π *in the permutations group* S_r. *The vector space of all alternative multi-linear functions*

$$\underbrace{V \times \cdots \times V}_{r} \longrightarrow \mathbf{R}$$

will be denoted by $\mathscr{A}_r(V)$ *and for convenience we set* $\mathscr{A}_0(V) := \mathbf{R}$.

♣

Note 1.6

(a) If $u \in \wedge^k V$ *and* $v \in \wedge^\ell V$, *then* $u \wedge v \in \wedge^{k+\ell} V$ *and*

$$u \wedge v = (-1)^{k\ell} v \wedge u. \tag{1.2.10}$$

Proof. without loss of generality, we may assume that $u := u_1 \wedge \cdots \wedge u_k$ and $v = v_1 \wedge \cdots \wedge v_\ell$. Suppose that $1 \leq i < j \leq k$. Then

$$0 = u_1 \wedge \cdots \wedge (u_i + u_j) \wedge \cdots \wedge (u_i + u_j) \wedge \cdots \wedge u_k$$

where the first $u_i + u_j$ is in the i-th position while the second one in the j-th position. Direct computation shows that

$$u_1 \wedge \cdots \wedge u_i \wedge \cdots \wedge u_j \wedge \cdots \wedge u_k = -u_1 \wedge \cdots \wedge u_j \wedge \cdots \wedge u_i \wedge \cdots \wedge u_k.$$

Then

$$\begin{aligned} u \wedge v &= u_1 \wedge \cdots \wedge u_k \wedge v_1 \wedge \cdots \wedge v_\ell \\ &= -u_1 \wedge \cdots \wedge u_{k-1} \wedge v_1 \wedge u_k \wedge v_2 \wedge \cdots \wedge v_\ell \\ &= (-1)^k v_1 \wedge u_1 \wedge \cdots \wedge u_k \wedge v_2 \wedge \cdots \wedge v_\ell \\ &= (-1)^{k\ell} v_1 \wedge \cdots \wedge v_\ell \wedge u_1 \wedge \cdots \wedge u_k \ = \ (-1)^{k\ell} v \wedge u. \end{aligned}$$

□

(b) If $(e_i)_{1 \le i \le \dim V}$ is a basis of V, then

$$(e_I)_I, \quad I \subset \{1, \cdots, \dim V\}, \quad e_I := e_{i_1} \wedge \cdots \wedge e_{i_r} \text{ with } i_1 < \cdots < i_r,$$

is a basis of $\wedge V$ (If $I = \emptyset$, we require $e_I = 1$). In particular,

$$\wedge^{\dim V} V \cong \mathbf{R}, \quad \wedge^j V = \{0\} \text{ for } j > \dim V. \tag{1.2.11}$$

Moreover,

$$\dim(\wedge V) = 2^{\dim V}, \quad \dim(\wedge^k V) = \binom{\dim V}{k} \text{ for } 0 \le k \le \dim V. \tag{1.2.12}$$

♣

Proposition 1.8. (Universal mapping property)

Let

$$\varphi : \underbrace{V \times \cdots \times V}_{k} \longrightarrow \wedge^k V, \quad (v_1, \cdots, v_k) \longmapsto v_1 \wedge \cdots \wedge v_k$$

be the natural alternative multi-linear map. To each alternative multi-linear map

$$h : \underbrace{V \times \cdots \times V}_{k} \longrightarrow W$$

there corresponds uniquely a linear map $\tilde{h} : \wedge^k V \to W$ such that $\tilde{h} \circ \varphi = h$.

$$\begin{tikzcd} \wedge^k V \arrow[dr, dotted, "\exists ! \tilde{h}"] & \\ \underbrace{V \times \cdots \times V}_{k} \arrow[u, "\varphi"] \arrow[r, "h"'] & W \end{tikzcd}$$

The pair $(\wedge^k V, \varphi)$ is said to solve the **universal mapping problem** *for alternative multi-linear maps with domain $V \times \cdots \times V$ (k copies); and this is the unique solution in the sense that if X is a vector space and $\tilde{\varphi} : V \times \cdots \times V \to X$ an alternative multi-linear map also possessing the universal mapping property for alternative multi-linear maps with domains $V \times \cdots \times V$, then there is an isomorphism $\alpha : \wedge^k V \to X$ such that $\alpha \circ \varphi = \tilde{\varphi}$.* ♡

Proof. Define

$$\tilde{h}(v_1 \wedge \cdots \wedge v_k) := h(v_1, \cdots, v_k).$$

For $u_1 \otimes \cdots \otimes u_k \in \mathscr{I}^k V$, there exist $i, j \in \{1, \cdots, k\}$ with $i \neq j$ such that $u_i = u_j$. Then

$$h(u_{\pi(1)}, \cdots, u_{\pi(j)}, \cdots, u_{\pi(i)}, \cdots, u_k) = \mathrm{sgn}(\pi) \cdot h(u_1, \cdots, u_j, \cdots, u_i, \cdots, u_k)$$

where $\pi : (1, \cdots, i, \cdots, j \cdots, k) \mapsto (1, \cdots, j, \cdots, i, \cdots k)$ with $\mathrm{sgn}(\pi) = 2(j-i)-1$. Hence

$$h(u_1, \cdots, u_i, \cdots, u_j, \cdots, u_k) = -h(u_1, \cdots, u_j, \cdots, u_i, \cdots, u_k);$$

thus $h(u_1, \cdots, u_i, \cdots, u_j, \cdots, u_k) = 0$ and then $\tilde{h}$ is well-defined.

The reminder proof is similar to that in **Proposition 1.7**. □ □

In the special case $W := \mathbf{R}$, we can prove from **Proposition 1.8** that

$$(\wedge^k V)^* \cong \mathscr{A}_k(V). \tag{1.2.13}$$

Define

$$\Phi : (\wedge^k V)^* \longrightarrow \mathscr{A}_k(V), \quad f \longmapsto \Phi(f)$$

by

$$\Phi(f)(v_1, \cdots, v_k) := f(v_1 \wedge \cdots \wedge v_k),$$

and

$$\Psi : \mathscr{A}_k(V) \longrightarrow (\wedge^k V)^*, \quad h \longmapsto \Psi(h)$$

by

$$\Psi(h) := \tilde{h} : \wedge^k V \longrightarrow \mathbf{R}.$$

We first check that $\Phi(f) \in \mathscr{A}_k(V)$: for any permutation $\pi \in S_k$,

$$\begin{aligned} \Phi(f)(v_{\pi(1)}, \cdots, v_{\pi(k)}) &= f(v_{\pi(1)} \wedge \cdots \wedge v_{\pi(k)}) \\ &= f(\mathrm{sgn}(\pi) v_1 \wedge \cdots \wedge v_k) \;=\; \mathrm{sgn}(\pi)\Phi(f)(v_1 \wedge \cdots \wedge v_k). \end{aligned}$$

Compute

$$\begin{aligned} \Phi \circ \Psi(h)(v_1, \cdots, v_k) &= \Phi(\tilde{h})(v_1, \cdots, v_k) \\ &= \tilde{h}(v_1 \wedge \cdots \wedge v_k) \;=\; h(v_1, \cdots, v_k), \\ \Psi \circ \Phi(f)(v_1 \wedge \cdots \wedge v_k) &= \widetilde{\Phi(f)}(v_1 \wedge \cdots \wedge v_k) \\ &= \Phi(f)(v_1, \cdots, v_k) \;=\; f(v_1 \wedge \cdots \wedge v_k). \end{aligned}$$

Thus Φ is isomorphic.

We shall now consider various dualities between the spaces $\mathscr{T}^{r,s}V, \wedge^k V, \wedge V$ and the corresponding spaces $\mathscr{T}^{r,s}V^*, \wedge^k V^*, \wedge V^*$.

Definition 1.19. (Nonsingular pairings)

Let V and W be finite dimensional real vector spaces. A **pairing of V and W** *is a bilinear map*

$$(\,,\,) : V \times W \longrightarrow \mathbf{R}.$$

A pairing is called **nonsingular** *if whenever $w \neq 0$ in W, there exists an element $v \in V$ such that $(v, w) \neq 0$, and whenever $v \neq 0$ in V, there exists an element $w \in W$ such that $(v, w) \neq 0$.* ♣

Let V and W be nonsingularly paired by $(\ ,\ )$ and define

$$\varphi : V \longrightarrow W^*, \quad \varphi(v)(w) := (v, w), \quad v \in V, \ w \in W. \tag{1.2.14}$$

If $\varphi(v_1) = \varphi(v_2)$ for $v_1, v_2 \in V$, then

$$(v_1 - v_2, w) = 0, \quad \text{for all } w \in W.$$

Consequently, $v_1 - v_2 = 0$ by nonsingularity. Thus φ is injective. Similarly, defining

$$\psi : W \longrightarrow V^*, \quad \psi(w)(v) := (v, w), \quad v \in V, \ w \in W, \tag{1.2.15}$$

implies that ψ is injective and therefore

$$\dim V \le \dim W^* = \dim W \le \dim V^* = \dim V.$$

Thus φ and ψ are both isomorphisms.

A non-singular pairing of $\mathscr{T}^{r,s}V^*$ with $\mathscr{T}^{r,s}V$ is defined as follows:

$$\mathscr{T}^{r,s}V^* \times \mathscr{T}^{r,s}V \longrightarrow \mathbf{R}, \quad (v^*, u) \longmapsto (v^*, u)$$

where

$$v^* = v_1^* \otimes \cdots \otimes v_r^* \otimes u_{r+1} \otimes \cdots \otimes u_{r+s} \in \mathscr{T}^{r,s}V^*$$

and

$$u = u_1 \otimes \cdots \otimes u_r \otimes v_{r+1}^* \otimes \cdots \otimes v_{r+s}^* \in \mathscr{T}^{r,s}V$$

yields

$$(v^*, u) := v_1^*(u_1) \cdots v_{r+s}^*(u_{r+s}).$$

Clearly that this is a non-singular pairing. Indeed, if $u \neq 0$, then $u_1, \cdots, u_r$ and $v_{r+1}^*, \cdots, v_{r+s}^*$ are all nonzero. Choose elements $u_{r+1}, \cdots, u_{r+s} \in V$ such that $v_{r+1}^*(u_{r+1}), \cdots, v_{r+s}^*(u_{r+s})$ are nonzero. Setting $v_1^* = \cdots = v_r^*$ =nonzero constant yields $(v^*, u) \neq 0$. The above remark gives us an isomorphism

$$\mathscr{T}^{r,s}V^* \cong (\mathscr{T}^{r,s}V)^*. \tag{1.2.16}$$

Let $\mathscr{M}_{r,s}(V)$ be the vector space of all multi-linear functions

$$\underbrace{V \times \cdots \times V}_{r} \times \underbrace{V^* \times \cdots \times V^*}_{s} \longrightarrow \mathbf{R}.$$

By the universal mapping property, **Proposition 1.7**, we obtain

$$(\mathscr{T}^{r,s}V)^* \cong \mathscr{M}_{r,s}(V). \tag{1.2.17}$$

If $\tilde{h} \in (\mathscr{T}^{r,s}V)^*$, then the corresponding multi-linear function $h \in \mathscr{M}_{r,s}(V)$ satisfies

$$h(v_1, \cdots, v_r, v_1^*, \cdots, v_s^*) = \tilde{h}(v_1 \otimes \cdots \otimes v_r \otimes v_1^* \otimes \cdots \otimes v_s^*).$$

$$
\begin{tikzcd}
\mathscr{T}^{r,s}V \arrow[rr, dotted, "\tilde{h}"] & & \mathbf{R} \\
\underbrace{V\times\cdots\times V}_{r}\times\underbrace{V^*\times\cdots\times V^*}_{s} \arrow[u, "\varphi"] \arrow[urr, "h"'] & &
\end{tikzcd}
$$

Finally, from (1.2.16) and (1.2.17) we obtain an isomorphism

$$
\mathscr{T}^{r,s}V^* \cong (\mathscr{T}^{r,s}V)^* \cong \mathscr{M}_{r,s}(V). \tag{1.2.18}
$$

A **non-singular pairing of $\wedge^k V^*$ with $\wedge^k V$** is defined as follows:

$$
\wedge^k V^* \times \wedge^k V \longrightarrow \mathbf{R}, \quad (v^*, u) \longmapsto (v^*, u)
$$

where $v^* = v_1^* \wedge \cdots \wedge v_k^* \in \wedge^k V^*$ and $u = u_1 \wedge \cdots \wedge u_k \in \wedge^k V$ yields

$$
(v^*, u) := \det\left(v_i^*(u)\right).
$$

Clearly that this is a non-singular pairing. Suppose that $\det(v_i^*(u_j)) = 0$ for all $v^* = v_1^* \wedge \cdots \wedge v_k^* \in V^*$. Then the system

$$
\begin{aligned}
0 &= v_1^*(u_1)x^1 + \cdots + v_1^*(u_k)x^k \\
&\cdots \\
0 &= v_k^*(u_1)x^1 + \cdots + v_k^*(u_k)x^k
\end{aligned}
$$

has no zero solution $(x^1, \cdots, x^k) \in \mathbf{R}^k$. Consequently,

$$
0 = u_1 x^1 + \cdots + u_k x^k
$$

where $x^1, \cdots, x^k$ are not all equal to zero. Without loss of generality, we may assume that $x^1 \neq 0$; hence

$$
u_1 = -\frac{x^2}{x^1}u_2 - \cdots - \frac{x^k}{x^1}u_k
$$

and

$$
u = \left(-\frac{x^2}{x^1}u_2 - \cdots - \frac{x^k}{x^1}u_k\right) \wedge u_2 \wedge \cdots \wedge u_k = -\sum_{2\le i\le k} \frac{x^i}{x^1} u_i \wedge u_2 \wedge \cdots \wedge u_k = 0.
$$

Thus the pairing is non-singular and we have an isomorphism

$$
\wedge^k V^* \cong (\wedge^k V)^* \cong \mathscr{A}_k(V) \tag{1.2.19}
$$

by (1.2.13). Finally, we have

$$
\wedge V^* = \bigoplus_{k\ge 0} \wedge^k V^* \cong \bigoplus_{k\ge 0} (\wedge^k V)^* = (\wedge V)^*, \tag{1.2.20}
$$

$$
\wedge V^* = \bigoplus_{k\ge 0} \wedge^k V^* \cong \bigoplus_{k\ge 0} \mathscr{A}_k(V) =: \mathscr{A}(V). \tag{1.2.21}
$$

Note 1.7

(a) If $(e_i)_{1\le i\le \dim V}$ is a basis of V with dual basis $(e_i^*)_{1\le i\le \dim V}$ in V^*, then the bases $(e_I)_I$ and $(e_I^*)_I$ are dual bases of $\wedge V$ and $\wedge V^*$ under the isomorphism (1.2.20).

(b) Define $\alpha, \beta : \wedge^k V^* \times \wedge^k V \longrightarrow \mathbf{R}$, *where*

$$\alpha(v^*, u) := \det\left(v_i^*(u_j)\right), \tag{1.2.22}$$

$$\beta(v^*, u) := \frac{1}{k!}\det\left(v_i^*(u_j)\right) = \frac{1}{k!}\alpha(v^*, u) \tag{1.2.23}$$

for any $v^* = v_1^* \wedge \cdots \wedge v_k^* \in \wedge^k V^*$ *and* $u = u_1 \wedge \cdots \wedge u_k \in \wedge^k V$. *These two different isomorphisms induce different algebra structures* $\wedge_\alpha$ *and* $\wedge_\beta$ *on* $\mathscr{A}(V)$ *via (1.2.21). If* $f \in \mathscr{A}_p(V)$ *and* $g \in \mathscr{A}_q(V)$, *then*

$$f \wedge_\alpha g(v_1, \cdots, v_{p+q}) = \sum_{\pi \in \wedge_{p,q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \tag{1.2.24}$$

where $\pi \in \wedge_{p,q}$ *means that* $\pi(1) < \cdots < \pi(p)$ *and* $\pi(p+1) < \cdots < \pi(p+q)$, *and*

$$f \wedge_\beta g(v_1, \cdots, v_{p+q}) = \sum_{\pi \in S_{p+q}} \frac{\operatorname{sgn}(\pi)}{(p+q)!} f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}). \tag{1.2.25}$$

We now claim that

$$f \wedge_\alpha g = \frac{(p+q)!}{p!q!} f \wedge_\beta g. \tag{1.2.26}$$

We suffice to prove

$$\begin{aligned} & \sum_{\pi \in S_{p+q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \\ = \; & p!q! \sum_{\pi \in \wedge_{p,q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}). \end{aligned}$$

The left-hand side is equal to

$$\begin{aligned} & \sum_{\pi \in \wedge_{p,q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \\ + \; & \sum_{\pi \in S_{p+q} \setminus \wedge_{p,q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}); \end{aligned}$$

there are three cases of $\pi \in S_{p+q} \setminus \wedge_{p,q}$:

(A) $\pi(p+1) < \cdots < \pi(p+q)$,

(B) $\pi(1) < \cdots < \pi(p)$, *and*

(C) othercase.

Compute

$$\begin{aligned} & \sum_{\pi \in (A)} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \\ = \; & (p! - 1) \sum_{\sigma \in \wedge_{\pi;p}} \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma) f(v_{\sigma(\pi(1))}, \cdots, v_{\sigma(\pi(p))}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \\ = \; & (p! - 1) \sum_{\sigma \in \wedge_{\pi;p,0}} \operatorname{sgn}(\pi \circ \sigma) f(v_{\sigma(\pi(1))}, \cdots, v_{\sigma(\pi(p))}) g(v_{\sigma(\pi(p+1))}, \cdots, v_{\sigma(\pi(p+q))}) \\ = \; & (p! - 1) \sum_{\pi \in \wedge_{p,q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \end{aligned}$$

where $\sigma \in \wedge_{\pi;p}$ *means that* $\sigma(\pi(1)) < \cdots < \sigma(\pi(p))$ *and* $\sigma \in \wedge_{\pi;p,0}$ *means that* $\sigma(\pi(1)) < \cdots < \sigma(\pi(p))$ *and* $\sigma(\pi(i)) = \pi(i)$ *for* $p+1 \le i \le p+q$*. Similarly*

$$\begin{aligned}&\sum_{\pi\in(B)} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \\ = \;& (q!-1) \sum_{\sigma\in\wedge_{p,q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}), \\ &\sum_{\pi(C)} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \\ = \;& (p!-1)(q!-1) \sum_{\sigma\in\wedge_{p,q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}).\end{aligned}$$

Consequently the left-hand is equal to the term

$$\sum_{\pi\in\wedge_{p,q}} \operatorname{sgn}(\pi) f(v_{\pi(1)}, \cdots, v_{\pi(p)}) g(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)})$$

multiplied by the constant

$$(p!-1) + (q!-1) + (p!-1)(q!-1) + 1 = p!q!.$$

For example, for $p = q = 1$*, we have* $\mathscr{A}_1(V) \cong \wedge^1 V^* \cong (\wedge^1 V)^* = V^*$*. If* $\gamma, \delta \in V^* \cong \mathscr{A}_1(V)$ *and* $v, w \in V$*, then*

$$\gamma \wedge_\alpha \delta \in \mathscr{A}_2(V) \cong (\wedge^2 V)^*, \quad \gamma \wedge_\beta \delta \in \mathscr{A}_2(V) \cong (\wedge^2 V)^*.$$

Moreover,

$$\gamma \wedge_\alpha \delta(v, w) = \gamma(v)\delta(w) - \gamma(w)\delta(v), \quad \gamma \wedge_\beta \delta(v, w) = \frac{\gamma(v)\delta(w) - \gamma(w)\delta(v)}{2}.$$

♣

Let $\operatorname{End}(\wedge V)$ denote the vector space of all endomorphisms of $\wedge V$ (i.e., linear transformations from $\wedge V$ into $\wedge V$). Let $u \in \wedge V$.

(1) **Left multiplication by** u is the endomorphism $\epsilon_u = u\wedge \in \operatorname{End}(\wedge V)$ defined by

$$\epsilon_u v := u \wedge v, \quad v \in \wedge V. \tag{1.2.27}$$

(2) **Interior multiplication by** u is the endomorphism $\iota_u \in \operatorname{End}(\wedge V^*)$ defined by

$$(\iota_u v^*, w) := (v^*, \epsilon_u w), \quad v^* \in \wedge V^*, \; w \in \wedge V, \tag{1.2.28}$$

where $\iota_u v^* \in \wedge V^* \cong (\wedge V)^*$.

(3) If $u \in V = \wedge^1 V$, for each $k \in \mathbf{N}$, we have

$$\iota_u : \wedge^k V^* \longrightarrow \wedge^{k-1} V^*, \quad v_1^* \wedge \cdots \wedge v_k^* \longmapsto \iota_u(v_1^* \wedge \cdots \wedge v_k^*) \tag{1.2.29}$$

where

$$\iota_u(v_1^* \wedge \cdots \wedge v_k^*)(w_2 \wedge \cdots \wedge w_k) := v_1^* \wedge \cdots \wedge v_k^*(u \wedge w_2 \wedge \cdots w_k).$$

When $k = 1$, we get

$$\iota_u : V^* = \wedge^1 V^* \longrightarrow \wedge^0 V^* = \mathbf{R}, \quad v^* \longmapsto \iota_u v^*$$

and

$$\iota_u v^* = (\iota_u v^*, 1) = (v^*, \epsilon_u 1) = (v^*, u) = v^*(u).$$

An endomorphism l of $\wedge V$ is called

(a) a **derivation** if

$$l(u \wedge v) = l(u) \wedge v + u \wedge l(v), \quad u, v \in \wedge V. \tag{1.2.30}$$

(b) an **anti-derivation** if

$$l(u \wedge v) = l(u) \wedge v + (-1)^p u \wedge l(v), \quad u \in \wedge^p V, \ v \in \wedge V. \tag{1.2.31}$$

(c) **of degree k** if

$$l : \wedge^j V \longrightarrow \wedge^{j+k} V$$

for all j, where we assume that $\wedge^i V = \{0\}$ for $i < 0$.

Proposition 1.9

(a) $l \in \mathrm{End}(\wedge V)$ is an anti-derivation if and only if

$$l(v_1 \wedge \cdots \wedge v_j) = \sum_{1 \le i \le j} (-1)^{i+1} v_1 \wedge \cdots \wedge l(v_i) \wedge \cdots \wedge v_j, \quad \textit{for all } j.$$

(b) If $u \in V$, then ι_u is an anti-derivation of degree -1. ♡

Proof. (a) Suppose $l \in \mathrm{End}(\wedge V)$ is an anti-derivation and the result holds for $j-1$. We have

$$\begin{aligned} l(v_1 \wedge \cdots \wedge v_j) &= l((v_1 \wedge \cdots \wedge v_{j-1}) \wedge v_j) \\ &= l(v_1 \wedge \cdots \wedge v_{j-1}) \wedge v_j + (-1)^{j-1} v_1 \wedge \cdots \wedge v_{j-1} \wedge l(v_j) \\ &= \sum_{1 \le i \le j-1} (-1)^{i+1} v_1 \wedge \cdots \wedge l(v_i) \wedge \cdots \wedge v_{j-1} \wedge v_j \\ &\quad + (-1)^{j+1} v_1 \wedge \cdots \wedge v_{j-1} \wedge l(v_j) \\ &= \sum_{1 \le i \le j} (-1)^{i+1} v_1 \wedge \cdots \wedge l(v_i) \wedge \cdots \wedge v_j. \end{aligned}$$

Conversely, choose a basis $\{e_i\}_{1 \le i \le \dim V}$ of V. Any element $u \in \wedge^p V$ can be written as $u = \sum_{1 \le i_1 < \cdots < i_p \le \dim V} a^{i_1 \cdots i_p} e_{i_1} \wedge \cdots \wedge e_{i_p}$. If $v \in \wedge^q V$ with $v = \sum_{1 \le j_1 < \cdots < j_q \le \dim V} b^{j_1 \cdots j_q} e_{j_1} \wedge \cdots \wedge e_{i_q}$, then

$$\begin{aligned} l(u \wedge v) &= a^{i_1 \cdots i_p} b^{j_1 \cdots j_q} l\left(e_{i_1} \wedge \cdots \wedge e_{i_p} \wedge e_{j_1} \wedge \cdots \wedge e_{j_q}\right) \\ &= a^{i_1 \cdots i_p} b^{j_1 \cdots j_q} \left(\sum_{1 \le m \le p} (-1)^{m+1} e_{i_1} \wedge \cdots \wedge l(e_{i_m}) \wedge \cdots \wedge e_{i_p} \wedge \cdots \wedge e_{j_q} \right. \\ &\quad \left. + \sum_{1 \le n \le q} (-1)^{n+1+p} e_{i_1} \wedge \cdots \wedge e_{i_p} \wedge e_{j_1} \wedge \cdots \wedge l(e_{j_n}) \wedge \cdots \wedge e_{j_q} \right) \\ &= a^{i_1 \cdots i_p} \sum_{1 \le m \le p} (-1)^{m+1} e_{i_1} \wedge \cdots \wedge l(e_{i_m}) \wedge \cdots \wedge e_{i_p} \wedge v \\ &\quad + (-1)^p u \wedge b^{j_1 \cdots j_q} \sum_{1 \le n \le q} (-1)^{n+1} e_{j_1} \wedge \cdots \wedge l(e_{j_n}) \wedge \cdots \wedge e_{j_q} \\ &= l(u) \wedge v + (-1)^p u \wedge l(v). \end{aligned}$$

(b) By definition, ι_u is of degree -1. To prove the anti-derivation, we suffice to check part (a). For any $w_2 \wedge \cdots \wedge w_j \in \wedge^{j-1}V$ and $v_1^* \wedge \cdots \wedge v_j^* \in \wedge^j V^*$, we have

$$\begin{aligned}
\big(\iota_u(v_1^* \wedge \cdots \wedge v_j^*), w_2 \wedge \cdots \wedge w_j\big) &= (v_1^* \wedge \cdots \wedge v_j^*, u \wedge w_2 \wedge \cdots \wedge w_j) \\
= \sum_{1 \le i \le j} (-1)^{i+1} v_i^*(u) \det\left(v_k^*(w_\ell)\ (k = \{1, \cdots, j\} \setminus \{i\},\ \ell \in \{2, \cdots, j\})\right) \\
= \sum_{1 \le i \le j} (-1)^{i+1} v_i^*(u) \left(v_1^* \wedge \cdots \wedge \widehat{v_i^*} \wedge \cdots \wedge v_j^*, w_2 \wedge \cdots \wedge w_j\right) \\
= \left(\sum_{1 \le i \le j} (-1)^{i+1} v_1^* \wedge \cdots \wedge v_i^*(u) \wedge \cdots \wedge v_j^*, w_2 \wedge \cdots \wedge w_j\right) \\
= \left(\sum_{1 \le i \le j} (-1)^{i+1} v_1^* \wedge \cdots \wedge \iota_u v_i^* \wedge \cdots \wedge v_j^*, w_2 \wedge \cdots \wedge w_j\right).
\end{aligned}$$

Hence ι_u is an anti-derivation. □

Let $l : V \to W$ be a linear transformation. Then l defines an algebra homomorphism

$$l : \wedge V \longrightarrow \wedge W, \quad v_1 \wedge \cdots \wedge v_k \longmapsto l(v_1) \wedge \cdots \wedge l(v_k) \tag{1.2.32}$$

and $l(1) := 1$. The transpose $l^* : W^* \to V^*$ defines an algebra homomorphism

$$l^* : \wedge W^* \longrightarrow \wedge V^*, \quad w_1^* \wedge \cdots \wedge w_k^* \longmapsto l^*(w_1^*) \wedge \cdots \wedge l^*(w_k^*). \tag{1.2.33}$$

For any $w^* \in \wedge W^*$ and $v \in \wedge V$, we have

$$(l^*(w^*), v) = (w^*, l(v)). \tag{1.2.34}$$

For $w^* = w_1^* \wedge \cdots \wedge w_k^*$ and $v = v_1 \wedge \cdots \wedge v_k$, we have

$$\begin{aligned}
(l^*(w^*), v) &= (l^*(w_1^*) \wedge \cdots \wedge l^*(w_k^*), v_1 \wedge \cdots \wedge v_k) \\
&= \det\left(l^*(w_i^*)(v_j)\right) = \det\left(w_i^*(l(v_j))\right) \\
&= (w_1^* \wedge \cdots \wedge w_k^*, l(v_1) \wedge \cdots \wedge l(v_k)) = (w^*, l(v)).
\end{aligned}$$

1.2.2 Tensor fields and differential forms

Let $\mathcal{M}$ be a differentiable manifold. Define

$$\mathscr{T}^{r,s}T\mathcal{M} := \bigcup_{x \in \mathcal{M}} \mathscr{T}^{r,s}T_x\mathcal{M}, \tag{1.2.35}$$

$$\wedge^k T^*\mathcal{M} := \bigcup_{x \in \mathcal{M}} \wedge^k T_x^*\mathcal{M}, \tag{1.2.36}$$

$$\wedge T^*\mathcal{M} := \bigcup_{x \in \mathcal{M}} \wedge T_x^*\mathcal{M} \tag{1.2.37}$$

the **tensor bundle of type (r, s) over $\mathcal{M}$**, **exterior k bundle over $\mathcal{M}$**, and **exterior bundle over $\mathcal{M}$**, respectively. $\mathscr{T}^{r,s}T\mathcal{M}$, $\wedge^k T^*\mathcal{M}$, and $\wedge T^*\mathcal{M}$ have natural manifold structures such that the canonical projection maps to $\mathcal{M}^m$ are smooth. If $(\mathcal{U}, \varphi)$ is a coordinate system on

$\mathcal{M}$ with coordinate functions $x^1, \cdots, x^m$, then the basis[2] $(\partial/\partial x^i)_{1\le i\le m}$ of $T_x\mathcal{M}$ on $\mathcal{M}$ and $(dx^i)_{1\le i\le m}$ of $T_x^*\mathcal{M}$ on $\mathcal{M}$, for $x \in \mathcal{U}$, yield

(a) the basis of $\mathscr{T}^{r,s}T_x\mathcal{M}$:

$$\left(\frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}\right)_{1\le i_1,\cdots,i_r,j_1,\cdots,j_s\le m};$$

(b) the basis of $\wedge^k T_x^*\mathcal{M}$:

$$\left(dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right)_{1\le i_1<\cdots<i_k\le m};$$

(c) the basis of $\wedge T_x^*\mathcal{M}$:

$$(dx^i)_{1\le i\le m},\ (dx^{i_1}\wedge dx^{i_2})_{1\le i_1<i_2\le m},\ \cdots,\ \left(dx^1\wedge\cdots\wedge dx^m\right).$$

Definition 1.20. (Tensor fields and differential forms)

A smooth mapping of $\mathcal{M}$ into $\mathscr{T}^{r,s}T\mathcal{M}$, $\wedge^k T^*\mathcal{M}$, or $\wedge T^*\mathcal{M}$ whose composition with the canonical projection is the identity map is called a **(smooth) tensor field of type (r,s) on $\mathcal{M}$**, *a* **(differential) k-form on $\mathcal{M}$**, *or a* **(differential) form on $\mathcal{M}$**, *respectively.* ♣

A lifting $\alpha : \mathcal{M} \to \mathscr{T}^{r,s}T\mathcal{M}$ is a smooth tensor field of type (r,s) if and only if for each coordinate system $(\mathcal{U}, x^1, \cdots, x^m)$ on $\mathcal{M}$,

$$\alpha_{|\mathcal{U}} = a^{i_1\cdots i_r}_{j_1\cdots j_s}\frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s} \tag{1.2.38}$$

where $a^{i_1\cdots i_r}_{j_1\cdots j_s} \in C^\infty(\mathcal{U})$. A lifting $\beta : \mathcal{M} \to \wedge^k T^*\mathcal{M}$ is a differential k-form if and only if for any coordinate system $(\mathcal{U}, x^1, \cdots, x^m)$ on $\mathcal{M}$,

$$\beta_{|\mathcal{U}} = b_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k} \tag{1.2.39}$$

where $b_{i_1\cdots i_k} \in C^\infty(\mathcal{U})$.

Definition 1.21

Let

$$\mathcal{E}^k(\mathcal{M}) := C^\infty(\mathcal{M}, \wedge^k T^*\mathcal{M}) \tag{1.2.40}$$

denote the space of all smooth k-forms on $\mathcal{M}$, and

$$\mathcal{E}^*(\mathcal{M}) := C^\infty(\mathcal{M}, \wedge T^*\mathcal{M}) = \bigoplus_{0\le k\le m} \mathcal{E}^k(\mathcal{M}) \tag{1.2.41}$$

the space of all smooth forms on $\mathcal{M}$. ♣

Since $\wedge^0 T^*\mathcal{M} = \cup_{x\in\mathcal{M}} \wedge^0 T_x^*\mathcal{M} = \cup_{x\in\mathcal{M}}\mathbf{R} = \mathcal{M}\times\mathbf{R}$, it follows that

$$\mathcal{E}^0(\mathcal{M}) \cong C^\infty(\mathcal{M}). \tag{1.2.42}$$

We now consider operations on forms.

(1) For $\omega, \eta \in \mathcal{E}^*(\mathcal{M})$, define $\omega + \eta \in \mathcal{E}^*(\mathcal{M})$ by

$$(\omega+\eta)_x := \omega_x + \eta_x.$$

[2] We always omit the subscript x.

(2) For $\omega \in \mathcal{E}^*(\mathcal{M})$ and $c \in \mathbf{R}$, define $c\omega \in \mathcal{E}^*(\mathcal{M})$ by

$$(c\omega)_x := c\omega_x.$$

(3) For $\omega, \eta \in \mathcal{E}^*(\mathcal{M})$, define $\omega \wedge \eta \in \mathcal{E}^*(\mathcal{M})$ by

$$(\omega \wedge \eta)_x := \omega_x \wedge \eta_x.$$

Moreover, if $\omega \in \mathcal{E}^p(\mathcal{M})$ and $\eta \in \mathcal{E}^q(\mathcal{M})$, then $\omega \wedge \eta \in \mathcal{E}^{p+q}(\mathcal{M})$.

(4) For $\omega \in \mathcal{E}^*(\mathcal{M})$ and $f \in \mathcal{E}^0(\mathcal{M})$, define $d\omega \in \mathcal{E}^*(\mathcal{M})$ by

$$(f\omega)_x := f(x)\omega_x.$$

Clearly that $\mathcal{E}^*(\mathcal{M})$ is a module over the ring $C^\infty(\mathcal{M})$ and $(\mathcal{E}^*(\mathcal{M}), \wedge)$ is a graded algebra over $\mathbf{R}$.

Let $\mathfrak{X}(\mathcal{M})$ denote the $C^\infty(\mathcal{M})$-module of smooth vector fields on $\mathcal{M}$; that is

$$\mathfrak{X}(\mathcal{M}) := C^\infty(\mathcal{M}, T\mathcal{M}). \tag{1.2.43}$$

Consider $\mathscr{A}_k(\mathcal{M})$ the set of all alternative $C^\infty(\mathcal{M})$ multi-linear map

$$\underbrace{\mathfrak{X}(\mathcal{M}) \times \cdots \times \mathfrak{X}(\mathcal{M})}_{k} \longrightarrow C^\infty(\mathcal{M}).$$

Proposition 1.10

We have

$$\mathcal{E}^k(\mathcal{M}) \cong \mathscr{A}_k(\mathcal{M}). \tag{1.2.44}$$

Proof. Let $\omega \in \mathcal{E}^k(\mathcal{M})$ be a k-form on $\mathcal{M}$. For any $X_1, \cdots, X_k \in \mathfrak{X}(\mathcal{M})$, define

$$\omega(X_1, \cdots, X_k)(x) := \omega_x(X_1(x), \cdots, X_k(x)), \quad x \in \mathcal{M}. \tag{1.2.45}$$

Then ω can be viewed as an alternative multi-linear map of the module $\mathfrak{X}(\mathcal{M})$ into $C^\infty(\mathcal{M})$:

$$\begin{aligned} & \omega(X_1, \cdots, X_{i-1}, fX + gY, X_{i+1}, \cdots, X_k) \\ = \ & f\omega(X_1, \cdots, X_{i-1}, X, X_{i+1}, \cdots, X_k) + g\omega(X_1, \cdots, X_{i-1}, Y, X_{i+1}, \cdots, X_k) \end{aligned}$$

whenever $f, g \in C^\infty(\mathcal{M})$ and $X_1, \cdots, X_{i-1}, X, Y, X_{i+1}, \cdots, X_k \in \mathfrak{X}(\mathcal{M})$. Thus $\omega \in \mathscr{A}_k(\mathcal{M})$.

Conversely, choose an element $\omega \in \mathscr{A}_k(\mathcal{M})$. For any $v_1, \cdots, v_k \in T_x\mathcal{M}$, choose $V_1, \cdots, V_k \in \mathscr{X}(\mathcal{M})$ such that

$$V_i(x) = v_i, \quad 1 \le i \le k.$$

Define

$$\omega_x(v_1, \cdots, v_k) := \omega(V_1, \cdots, V_k)(x).$$

We shall check that $\omega_x(v_1, \cdots, v_k)$ is well-defined and independent of the choice of the extensions V_i to v_i. Without loss of generality, we may assume that $k = 1$. Then $X : \mathfrak{X}(\mathcal{M}) \to C^\infty(\mathcal{M})$. To any fixed $X \in \mathfrak{X}(\mathcal{M})$, we want to show that $\omega(X)(x)$ depends only on $X(x)$.

For this, it suffices to check $\omega(X)(x) = 0$ whenever $X(x) = 0$. Write

$$X = \sum_{1 \le i \le m} a^i \frac{\partial}{\partial x^i}, \quad a^i \in C^\infty(\mathcal{U})$$

in a coordinate system $(\mathcal{U}, x^1, \cdots, x^m)$ about x. Since $X(x) = 0$, it follows that $a^i(x) = 0$ for all i. By **Corollary 1.1**, there exist a smooth function φ on $\mathcal{M}$ and a neighborhood $\mathcal{V}$ of x such that

$$\mathcal{V} \subset \mathcal{U}, \quad \varphi_{|\mathcal{V}} = 1, \quad \varphi_{|\mathcal{M}^m \setminus \mathcal{U}} = 0.$$

Define the vector field X_i by

$$X_i := \begin{cases} \varphi \frac{\partial}{\partial x^i}, & \text{on } \mathcal{U}, \\ 0, & \text{on } \mathcal{M} \setminus \mathcal{U}. \end{cases}$$

Then X_i is a smooth vector field on $\mathcal{M}$. Define the smooth function $\tilde{a}^i$ on $\mathcal{M}$ by

$$\tilde{a}^i := \begin{cases} \varphi a_i, & \text{on } \mathcal{U}, \\ 0, & \text{on } \mathcal{M}^m \setminus \mathcal{U}. \end{cases}$$

Hence

$$\varphi^2 X = \sum_{1 \le i \le m} \varphi^2 a_i \frac{\partial}{\partial x^i} = \sum_{1 \le i \le m} \varphi a_i \left(\varphi \frac{\partial}{\partial x^i}\right);$$

on $\mathcal{U}$, we have

$$\varphi^2 X = \sum_{1 \le i \le m} \tilde{a}^i X_i \Longrightarrow X = \sum_{1 \le i \le m} \tilde{a}^i X_i + (1 - \varphi^2) X,$$

while, on $\mathcal{M} \setminus \mathcal{U}$,

$$\varphi^2 X = 0 = X_i = \tilde{a}^i.$$

Finally, we arrive at

$$X = \sum_{1 \le i \le m} \tilde{a}^i X_i + (1 - \varphi^2) X$$

on $\mathcal{M}$. Therefore

$$\begin{aligned} \omega(X)(x) &= \omega\left(\sum_{1 \le i \le m} \tilde{a}^i X_i + (1 - \varphi^2) X\right)(x) \\ &= \sum_{1 \le i \le m} \tilde{a}^i(x) \omega(X_i)(x) + ((1 - \varphi^2)(x)) \omega(X)(x) = 0. \end{aligned}$$

So $\omega(X)(x) = 0$ if $X(x) = 0$. □

If T is a tensor field of type (r, s), then we can consider T as a map

$$T : \underbrace{\mathcal{E}^1(\mathcal{M}) \times \cdots \times \mathcal{E}^1(\mathcal{M})}_{r} \times \underbrace{\mathfrak{X}(\mathcal{M}) \times \cdots \times \mathfrak{X}(\mathcal{M})}_{s} \longrightarrow C^\infty(\mathcal{M})$$

where

$$T(\omega_1, \cdots, \omega_r, X_1, \cdots, X_s)(x) := T_x\left(\omega_1(x), \cdots, \omega_r(x), X_1(x), \cdots, X_s(x)\right),$$

which is $C^\infty(\mathcal{M})$ multi-linear with respect to the $C^\infty(\mathcal{M})$-modules $\mathcal{E}^1(\mathcal{M})$ and $\mathfrak{X}(\mathcal{M})$.

If $\omega, \eta \in \mathcal{E}^1(\mathcal{M})$ and $X, Y \in \mathfrak{X}(\mathcal{M})$, we have

$$\omega \wedge \eta(X, Y) = \omega(X)\eta(Y) - \omega(Y)\eta(X). \tag{1.2.46}$$

Definition 1.22

If $f \in C^\infty(\mathcal{M})$, then df is a smooth mapping of $T\mathcal{M}$ into $\mathbf{R}$ which is linear on each tangent space. Thus df can be considered as a 1-form,

$$df : \mathcal{M} \longrightarrow \wedge^1 T^*\mathcal{M}, \quad x \longmapsto df(x) := df|_x \in T_x^*\mathcal{M} = \wedge^1 T_x^*\mathcal{M}. \tag{1.2.47}$$

The 1-form $df \in \mathcal{E}^1(\mathcal{M})$ is called the **exterior derivative** *of the 0-form f. Moreover, we obtain a map*

$$d : \mathcal{E}^0(\mathcal{M}) \longrightarrow \mathcal{E}^1(\mathcal{M}), \quad f \longmapsto df. \tag{1.2.48}$$

♣

Theorem 1.11. (Exterior differentiation)

There exists a unique anti-derivation $d : \mathcal{E}^*(\mathcal{M}) \to \mathcal{E}^*(\mathcal{M})$ of degree $+1$ such that

(1) $d^2 = 0$.

(2) Whenever $f \in C^\infty(\mathcal{M}) = \mathcal{E}^0(\mathcal{M})$, df is the differential of f.

d is called the **exterior differentiation operator** *of $\mathcal{E}^*(\mathcal{M})$.*

♡

Proof. (1) **Existence.** Let $x \in \mathcal{M}$ and define

$$\begin{aligned} \mathcal{E}_x^*(\mathcal{M}) &:= \{\text{smooth forms defined on open subsets of } \mathcal{M} \text{ containing } x\}, \\ \mathcal{E}_x^k(\mathcal{M}) &:= \{\text{smooth } k\text{-forms defined on open subsets of } \mathcal{M} \text{ containing } x\}. \end{aligned}$$

Observe that $\mathcal{E}_x^*(\mathcal{M}) = \oplus_{k \geq 0} \mathcal{E}_x^k(\mathcal{M})$. We fix a coordinate system $(\mathcal{U}, x^1, \cdots, x^m)$ about x. If $\omega \in \mathcal{E}_x^*(\mathcal{M})$, then

$$\omega_{|\text{domain}(\omega) \cap \mathcal{U}} = a_I dx^I$$

where $a_I \in C^\infty(\text{domain}(\omega) \cap \mathcal{U})$, I runs over all subsets of $\{1, \cdots, m\}$, and $dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_r}$ when $I = \{i_1 < \cdots < i_r\}$ or $dx^I = 1$ when $I = \emptyset$. Define $d\omega \in \mathcal{E}_x^*(\mathcal{M})$ by

$$(d\omega)_x := da_I|_x \wedge dx^I|_x \in \wedge T_x^*\mathcal{M}.$$

We first give the following properties:

(a) $\omega \in \mathcal{E}_x^r(\mathcal{M})$ implies $d\omega(x) \in \wedge^{r+1} T_x^*\mathcal{M}$.

(b) $d\omega(x)$ depends only on the germ of ω at x.

(c) $d(a_1\omega_1 + a_2\omega_2)(x) = a_1 d\omega_1(x) + a_2 d\omega_2(x)$, $a_i \in \mathbf{R}$ and $\omega_i \in \mathcal{E}_x^*(\mathcal{M})$, where the domain of $a_1\omega_1 + a_2\omega_2$ is domain$(\omega_1)\cap$domain(ω_2).

(d) $d(\omega_1 \wedge \omega_2)(x) = d\omega_1(x) \wedge \omega_2(x) + (-1)^r \omega_1(x) \wedge d\omega_2(x)$, $\omega_1 \in \mathcal{E}_x^r(\mathcal{M}^m)$ and $\omega_2 \in \mathcal{E}_x^*(\mathcal{M})$.

(e) If f is smooth on a neighborhood of x, then $d(df)(x) = 0$.

If $\omega = \sum_{i_1 < \cdots < i_r} a_{i_1 \cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r}$, then

$$d\omega(x) = \sum_{i_1 < \cdots < i_r} \sum_{1 \leq j \leq m} \left.\frac{\partial a_{i_1 \cdots i_r}}{\partial x^j}\right|_x dx^j|_x \wedge dx^{i_1}|_p \wedge \cdots \wedge dx^{i_r}|_x \in \wedge^{r+1} T_x^*\mathcal{M}.$$

Part (b) follows from the chain rule, while part (c) is obvious. From (b) and (c), it suffices to check (d) for $\omega_1 = f dx^{i_1} \wedge \cdots \wedge dx^{i_r}$ and $\omega_2 = g dx^{j_1} \wedge \cdots \wedge dx^{j_s}$ on some neighborhood of x. If $r = s = 0$, then

$$\begin{aligned} d(\omega_1 \wedge \omega_2)(x) &= d(fg)(x) = df(x) \cdot g(x) + f(x) \cdot dg(x) \\ &= d\omega_1(x) \wedge \omega_2(x) + (-1)^r \omega_1(x) \wedge d\omega_2(x). \end{aligned}$$

The second case is $r = 0$ or $s = 0$. Assume first that $r = 0$; then

$$\begin{aligned} d(\omega_1 \wedge \omega_2)(x) &= d(fg dx^{j_1} \wedge \cdots \wedge dx^{j_s})(x) \\ &= \sum_{1 \le j \le m} \left.\frac{\partial (fg)}{\partial x^j}\right|_x dx^j|_x \wedge dx^{j_1}|_x \wedge \cdots \wedge dx^{j_s}|_x \\ &= \sum_{1 \le j \le m} \left(f(x) \left.\frac{\partial g}{\partial x^j}\right|_x + g(x) \left.\frac{\partial f}{\partial x^j}\right|_x \right) dx^j|_x \wedge dx^{j_1}|_x \wedge \cdots \wedge dx^{j_s}|_x \\ &= f(x) d\omega_2(x) + df(x) \wedge \omega_2(x) \\ &= d\omega_1(x) \wedge \omega_2(x) + (-1)^r \omega_1(x) \wedge d\omega_2(x). \end{aligned}$$

For the case $s = 0$, we obtain

$$\begin{aligned} d(\omega_1 \wedge \omega_2)(x) &= d(fg dx^{i_1} \wedge \cdots \wedge dx^{i_r})(x) \\ &= \sum_{1 \le j \le m} \left(f(x) \left.\frac{\partial g}{\partial x^j}\right|_x + g(x) \left.\frac{\partial f}{\partial x^j}\right|_x \right) dx^j|_x \wedge dx^{i_1}|_x \wedge \cdots \wedge dx^{i_r}|_x \\ &= g(x) \wedge d\omega_1(x) + \sum_{1 \le j \le m} f(x) \left.\frac{\partial g}{\partial x^j}\right|_x (-1)^r dx^{i_1}|_x \wedge \cdots \wedge dx^{i_r}|_x \wedge dx^j|_x \\ &= g(x) \wedge d\omega_1(x) + (-1)^r \omega_1(x) \wedge dg(x) \\ &= d\omega_1(x) \wedge \omega_2 + (-1)^r \omega_1(x) \wedge d\omega_2(x). \end{aligned}$$

The third and last case is $r, s > 0$. If $\{i_1, \cdots, i_r\} \cap \{j_1, \cdots, j_s\} \neq \emptyset$, the result id obvious. Then we may assume that $\{i_1, \cdots, i_r\} \cap \{j_1, \cdots, j_s\} = \emptyset$. Write

$$(f dx^{i_1} \wedge \cdots \wedge dx^{i_r}) \wedge (g dx^{j_1} \wedge \cdots \wedge dx^{j_s}) = \varepsilon fg dx^{\ell_1} \wedge \cdots \wedge dx^{\ell_{r+s}}$$

where $\ell_1 < \cdots < \ell_{r+s}$ and $\varepsilon = \pm 1$. In particular,

$$dx^{i_1} \wedge \cdots \wedge dx^{i_r} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_s} = \varepsilon dx^{\ell_1} \wedge \cdots \wedge dx^{\ell_{r+s}}.$$

Compute

$$\begin{aligned} d(\omega_1 \wedge \omega_2)(x) &= d(\varepsilon fg dx^{\ell_1} \wedge \cdots \wedge dx^{\ell_{r+s}})(x) \\ &= \varepsilon \left(df(x) g(x) + f(x) dg(x) \right) \wedge dx^{\ell_1}|_x \wedge \cdots \wedge dx^{\ell_{r+s}}|_x \\ &= \varepsilon g(x) df(x) \wedge dx^{\ell_1}|_x \wedge \cdots \wedge dx^{\ell_{r+s}}|_x \\ &\quad + \varepsilon f(x) dg(x) \wedge dx^{\ell_1}|_x \wedge \cdots \wedge dx^{\ell_{r+s}}|_x \\ &= \left(df(x) \wedge dx^{i_1}|_x \wedge \cdots \wedge dx^{i_r}|_x \right) \wedge \left(g(x) dx^{j_1}|_x \wedge \cdots \wedge dx^{j_s}|_x \right) \\ &\quad + (-1)^r \left(f(x) dx^{i_1}|_x \wedge \cdots \wedge dx^{i_r}|_x \right) \wedge \left(dg(x) \wedge dx^{j_1}|_x \wedge \cdots \wedge dx^{j_s}|_x \right) \\ &= d\omega_1(x) \wedge \omega_2(x) + (-1)^r \omega_1(x) \wedge d\omega_2(x). \end{aligned}$$

For (e), on domain$(f) \cap \mathcal{U}$, we have

$$df = \sum_{1\le i\le m} \frac{\partial f}{\partial x^i} dx^i$$

so that

$$\begin{aligned} d(df)(x) &= \sum_{1\le i\le m} d\left(\frac{\partial f}{\partial x^i}\right)(x) \wedge dx^i(x) \\ &= \sum_{1\le i,j\le m} \left.\frac{\partial^2 f}{\partial x^i \partial x^j}\right|_x dx^j|_x \wedge dx^i|_x \;=\; -\sum_{1\le i,j\le m} \left.\frac{\partial^2 f}{\partial x^i \partial x^j}\right|_x dx^i|_x \wedge dx^j|_x \end{aligned}$$

which implies that $d(df)(x) = 0$.

We now claim that the definition of d at x is independent of the choice of coordinate systems. If $(\mathcal{V}, y^1, \cdots, y^m)$ is another coordinate system about x, we can define another operator d' on $\mathcal{E}^*_x(\mathcal{M})$ satisfying the above properties (a) – (e). For $\omega \in \mathcal{E}^*_x(\mathcal{M})$, we have

$$\begin{aligned} d'\omega(x) &= d'\left(a_{i_1\cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r}\right)(x) \\ &= d'(a_{i_1\cdots i_r})(x) \wedge dx^{i_1}|_x \wedge \cdots \wedge dx^{i_r}|_x \\ &\quad + \sum_{1\le k\le r} (-1)^{k-1} a_{i_1\cdots i_r}(x) dx^{i_1}|_x \wedge \cdots \wedge d'(dx^{i_k})(x) \wedge \cdots \wedge dx^{i_r}|_x \\ &= d(a_{i_1\cdots i_r})(x) \wedge dx^{i_1}|_x \wedge \cdots \wedge dx^{i_r}|_x \;=\; d\omega(x). \end{aligned}$$

If $\omega \in \mathcal{E}^*(\mathcal{M})$, we define $d\omega$ to be the form which as a lifting of $\mathcal{M}$ into $\wedge T^*\mathcal{M}$ sends x to $d\omega(x)$. We now check that $d^2 = 0$ and d is an anti-derivation of $\mathcal{E}^*(\mathcal{M})$ of degree $+1$. If $\omega = a_I dx^I$ near $x \in \mathcal{M}$, then

$$d(d\omega)(x) = d\left(da_I \wedge dx^I\right)(x) = d(da_I)(x) \wedge dx^I|_x + (-1)^{|I|+1} da_I \wedge d(dx^I)(x) = 0.$$

(2) **Uniqueness.** Let d' be an anti-derivation of $\mathcal{E}^*(\mathcal{M})$ of degree $+1$ satisfying (1) and (2).

(i) *If $\omega \in \mathcal{E}^*(\mathcal{M})$ with $\omega|_{\mathcal{W}} = 0$, where $\mathcal{W}$ is a neighborhood of x, then $d'\omega(x) = 0$.* Choose a neighborhood $\mathcal{U}$ of x such that $\overline{\mathcal{U}} \subset \mathcal{W}$ and $\overline{\mathcal{U}}$ is compact. By **Corollary 1.1**, there exists a smooth function φ on $\mathcal{M}$ such that

$$\varphi|_{\mathcal{U}} \equiv 0, \quad \varphi|_{\mathcal{M}\setminus\mathcal{W}} \equiv 1.$$

Then $\varphi\omega = \omega$ on $\mathcal{M}$ and $d'\varphi(x) = \varphi(x) = 0$. Hence

$$d'\omega(x) = d'(\varphi\omega)(x) = d'\varphi(x) \wedge \omega(x) + \varphi(x) d'\omega(x) = 0.$$

(ii) d' is defined only on elements of $\mathcal{E}^*(\mathcal{M})$, that is, on globally defined forms on $\mathcal{M}$. We wish to define d' on $\mathcal{E}^*_x(\mathcal{M})$ for each $x \in \mathcal{M}$. *If $\omega \in \mathcal{E}^*_x(\mathcal{M})$, we define $d'\omega$ as follows.* Let $\mathcal{V}_1$ be the domain of ω. By **Corollary 1.1**, there exist a smooth function ψ on $\mathcal{M}^m$ and a neighborhood $\mathcal{V}_2$ of x such that

$$\psi_{|\mathcal{V}_2} \equiv 1, \quad \psi_{|\mathcal{M}\setminus\mathcal{V}_1} \equiv 0, \quad \mathrm{supp}(\psi) \subset V_1$$

and $\mathcal{V}_2 \subset \overline{\mathcal{V}_2} \subset \mathcal{V}_1$ and $\overline{\mathcal{V}_2}$ is compact. Then $\psi\omega \in \mathcal{E}^*(\mathcal{M})$ and define

$$d'\omega(x) := d'(\psi\omega)(x).$$

We must check that the above definition is independent of the extension. Let $(\tilde{\psi}, \widetilde{\mathcal{V}_2})$ be another pair with above properties. Then

$$(\psi\omega - \tilde{\psi}\omega)_{|\mathcal{V}_2 \cap \widetilde{\mathcal{V}_2}} = 0;$$

thus $d'(\psi\omega - \tilde{\psi}\omega)(x) = 0$ and then $d'(\psi\omega)(x) = d'(\tilde{\psi}\omega)(x)$.

(iii) *$d'\omega(x)$ is defined for all $\omega \in \mathcal{E}^*_x(\mathcal{M})$ and satisfies properties (a) – (e).* The properties (a) – (c) are obvious. For $\omega_1, \omega_2 \in \mathcal{E}^*_x(\mathcal{M})$ we can find a suitable smooth function (see above) φ on $\mathcal{M}^m$ such that $\varphi\omega_1 \wedge \varphi\omega_2 \in \mathcal{E}^*(\mathcal{M})$. Then ($\omega_1 \in \mathcal{E}^r_x(\mathcal{M})$)

$$\begin{aligned} d'(\omega_1 \wedge \omega_2)(x) &= d'(\varphi\omega_1 \wedge \varphi\omega_2)(x) \\ &= d'(\varphi\omega_1)(x) \wedge (\varphi\omega_2)(x) + (-1)^r(\varphi\omega_1)(x) \wedge d'(\varphi\omega_2)(x) \\ &= d'\omega 1(x) \wedge \omega_2(x) + (-1)^r \omega_1(x) \wedge d'\omega_2(x). \end{aligned}$$

For any smooth function f near x, we have

$$d'(d'f)(x) = d'(d'(\varphi f))(x) = 0.$$

Whenever $\omega \in \mathcal{E}^*_x(\mathcal{M})$, we have proved

$$d'\omega(x) = d\omega(x).$$

This prove the uniqueness. □

From the above prove that

$$d\omega_{|\mathcal{U}} = d(\omega_{|\mathcal{U}}) \tag{1.2.49}$$

whenever $\mathcal{U}$ is an open subset in $\mathcal{M}$.

Definition 1.23. (Interior multiplication)

Let $X \in \mathfrak{X}(\mathcal{M})$ and $\omega \in \mathcal{E}^*(\mathcal{M})$. **Interior multiplication of ω by X** *is the form $\iota_X\omega$ defined by*

$$\iota_X\omega(x) := \iota_{X_x}(\omega_x), \quad x \in \mathcal{M}^m. \tag{1.2.50}$$

Then $\iota_X\omega$ is a smooth form and $\iota_X : \mathcal{E}^*(\mathcal{M}) \to \mathcal{E}^*(\mathcal{M})$ is an anti-derivation of degree -1. ♣

Let $F : \mathcal{M} \to \mathcal{N}$ be a smooth map and let $x \in \mathcal{M}$. Then we have the differential $F_{*,x} : T_x\mathcal{M} \to T_{F(x)}\mathcal{N}$, its transpose $F^*_x : T^*_{F(x)}\mathcal{N} \to T^*_x\mathcal{M}$, and the induced algebra homomorphism

$$F^*_x : \wedge T^*_{F(x)}\mathcal{N} \longrightarrow \wedge T^*_x\mathcal{M}$$

If $\omega \in \mathcal{E}^*(\mathcal{N})$, then we can pull ω back to a form on $\mathcal{M}$ by setting

$$F^* : \mathcal{E}^*(\mathcal{N}) \longrightarrow \mathcal{E}^*(\mathcal{M}), \quad \omega \longmapsto F^*\omega \tag{1.2.51}$$

where $(F^*\omega)_x := F^*_x(\omega_{F(x)})$.

Proposition 1.11

Let $F : \mathcal{M} \to \mathcal{N}$ be a smooth map. Then

(a) $F^* : \mathcal{E}^*(\mathcal{N}) \to \mathcal{E}^*(\mathcal{M})$ *and is an algebra homomorphism.*

(b) F^* *commutes with* d*, that is*

$$d(F^*\omega) = F^*(d\omega), \quad \omega \in \mathcal{E}^*(\mathcal{N}). \tag{1.2.52}$$

$$\begin{tikzcd} \mathcal{E}^*(\mathcal{N}) \arrow[r, "F^*"] \arrow[d, "d"'] & \mathcal{E}^*(\mathcal{M}) \arrow[d, "d"] \\ \mathcal{E}^*(\mathcal{N}) \arrow[r, "F^*"'] & \mathcal{E}^*(\mathcal{M}) \end{tikzcd}$$

(c) *For any* $\omega \in \mathcal{E}^k(\mathcal{N})$ *and* $X_1, \cdots, X_k \in \mathfrak{X}(\mathcal{M})$*, we have*

$$(F^*\omega)(X_1, \cdots, X_k)(x) = \omega_{F(x)}(F_{*,x}(X_1(x)), \cdots, F_{*,x}(X_k(x))). \tag{1.2.53}$$

♡

Proof. Part (a) is obvious and part (c) follows from (1.2.34). We now consider part (b). If $f \in C^\infty(\mathcal{N})$, then $F^*f = f \circ F \in C^\infty(\mathcal{M})$ and

$$(F^*(df))_x = F_x^*(df_{F(x)}) = d(f \circ F)_x = d(F^*f)_x$$

by (1.1.21).

Let $\omega \in \mathcal{E}^*(\mathcal{N})$ and $x \in \mathcal{M}$. Consider a coordinate system $(\mathcal{V}, y^1, \cdots, y^n)$ about $y := F(x)$ and choose a neighborhood $\mathcal{U}$ of x such that $F(\mathcal{U}) \subset \mathcal{V}$. Write

$$\omega_{|\mathcal{V}} = a_{i_1\cdots i_r} dy^{i_1} \wedge \cdots \wedge dy^{i_r}, \quad a_{i_1\cdots i_r} \in C^\infty(\mathcal{V}).$$

Then

$$F^*\omega|_{\mathcal{U}} = (a_{i_1\cdots i_r} \circ F) d(y^{i_1} \circ F) \wedge \cdots \wedge d(y^{i_r} \circ F)$$

which is a smooth form on $\mathcal{V}$. Hence $F^*\omega \in \mathcal{E}^*(\mathcal{M})$. Moreover

$$\begin{aligned} d(F^*\omega)_x &= d\Big((a_{i_1\cdots i_r} \circ F) d(y^{i_1} \circ F) \wedge \cdots \wedge d(y^{i_r} \circ F)\Big)_x \\ &= \Big(d(a_{i_1\cdots i_r} \circ F) \wedge d(y^{i_1} \circ F) \wedge \cdots \wedge d(y^{i_r} \circ F)\Big)_x \\ &= \Big(F^*\Big(da_{i_1\cdots i_r} \wedge dy^{i_1} \wedge \cdots \wedge dy^{i_r}\Big)\Big)_x = (F^*(d\omega))_x. \end{aligned}$$

Therefore, $d(F^*\omega) = F^*(d\omega)$. □

1.2.3 Lie derivatives

Fix a smooth vector field X on a differentiable manifold $\mathcal{M}$. Recall the local 1-parameter group $\{X_t\}_{t\geq 0}$ of transformations associated with X. If Y is another smooth vector field on $\mathcal{M}$, we define the derivative of Y with respect to X at the point $x \in \mathcal{M}$ as follows: Since $X_t : \mathscr{D}_t \to \mathscr{D}_{-t}$ is a diffeomorphism, $(X_{-t})_{*,X_t(x)}(Y_{X_t(x)}) \in T_x\mathcal{M}^m$. We define the **Lie**

derivative of Y with respect to X at x by

$$(\mathscr{L}_X Y)_x := \lim_{t\to 0}\frac{(X_{-t})_{*,X_t(x)}(Y_{X_t(x)}) - Y_x}{t} = \frac{d}{dt}\Big|_{t=0}\left((X_{-t})_{*,X_t(x)}(Y_{X_t(x)})\right) \tag{1.2.54}$$

because $(X_{-0})_{*,X_0(x)}(Y_{X_0(x)}) = Y_x$. Similarly, we can define the **Lie derivative of a differential form ω with respect to X at x** by

$$(\mathscr{L}_X\omega)_x := \lim_{t\to 0}\frac{(X_t)^*_x(\omega_{X_t(x)}) - \omega_x}{t} = \frac{d}{dt}\Big|_{t=0}\left((X_t)^*_x(\omega_{X_t(x)})\right). \tag{1.2.55}$$

The smoothness of $\mathscr{L}_X Y$ and $\mathscr{L}_X\omega$ is obvious. The Lie derivative $\mathscr{L}_X$ can be extended to arbitrary tensor fields in the obvious way. If T is a tensor field of type (r,s), then $(\mathscr{L}_X T)_x$ is given by

$$(\mathscr{L}_X T)_x := \frac{d}{dt}\Big|_{t=0}\Big((X_{-t})_{*,X_t(x)}(v_1\otimes\cdots\otimes v_r)\otimes(X_t)^*_x(v_1^*\otimes\cdots\otimes v_s^*)\Big) \tag{1.2.56}$$

if $T_{X_t(x)} = v_1\otimes\cdots\otimes v_t\otimes v_1^*\otimes\cdots\otimes v_s^*$.

Proposition 1.12

Let X be a smooth vector field on $\mathcal{M}$. Then

(a) $\mathscr{L}_X f = Xf$ *whenever* $f\in C^\infty(\mathcal{M})$.

(b) $\mathscr{L}_X Y = [X,Y]$ *for each smooth vector field* Y *on* $\mathcal{M}$.

(c) $\mathscr{L}_X : \mathcal{E}^*(\mathcal{M})\to\mathcal{E}^*(\mathcal{M})$ *is a derivation which commutes with* d.

(d) **(Cartan formula)** $\mathscr{L}_X = \iota_X\circ d + d\circ\iota_X$ *on* $\mathcal{E}^*(\mathcal{M})$.

(e) If $\omega\in\mathcal{E}^p(\mathcal{M})$ *and* $Y_0,\cdots,Y_p\in\mathfrak{X}(\mathcal{M})$, *then*

$$\begin{aligned}\mathscr{L}_{Y_0}(\omega(Y_1,\cdots,Y_p)) &= (\mathscr{L}_{Y_0}\omega)(Y_1,\cdots,Y_p)\\ &\quad+\sum_{1\le i\le p}\omega(Y_1,\cdots,Y_{i-1},\mathscr{L}_{Y_0}Y_i,Y_{i+1},\cdots,Y_p).\end{aligned}$$

(f) If $\omega\in\mathcal{E}^p(\mathcal{M})$ *and* $Y_0,\cdots,Y_p\in\mathfrak{X}(\mathcal{M})$, *then*

$$\begin{aligned}d\omega(Y_0,\cdots,Y_1) &= \sum_{0\le i\le p}(-1)^i Y_i\Big(\omega(Y_0,\cdots,\widehat{Y_i},\cdots,Y_p)\Big)\\ &\quad+\sum_{0\le i<j\le p}(-1)^{i+j}\omega([Y_i,Y_j],Y_0,\cdots,\widehat{Y_i},\cdots,\widehat{Y_j},\cdots,Y_p).\end{aligned}$$

♡

Proof. (a) Compute

$$\begin{aligned}(\mathscr{L}_X f)_x &= \lim_{t\to0}\frac{(X_t)^*_x(f(X_t(x))) - f(x)}{t} = \frac{d}{dt}\Big|_{t=0}((X_t)^*_x(f(X_t(x))))\\ &= \frac{d}{dt}\Big|_{t=0}((f\circ X_t)(x)) = \frac{d}{dt}\Big|_{t=0}(f(\gamma_x(t)))\\ &= \sum_{1\le i\le m}\frac{\partial f}{\partial x^i}\Big|_x\frac{\partial\gamma_x^i}{\partial t}\Big|_{t=0} = \sum_{1\le i\le m}\frac{\partial f}{\partial x^i}\Big|_x a^i = (Xf)_x\end{aligned}$$

if $X_x = \sum_{1\le i\le m} a^i\frac{\partial}{\partial x^i}$.

(b) We need only to show that $(\mathscr{L}_X Y)(f) = [X,Y]f$ for each $f\in C^\infty(\mathcal{M})$. Let $x\in\mathcal{M}$.

Then

$$(\mathscr{L}_X Y)_x(f) = \left(\lim_{t\to 0}\frac{(X_{-t})_{*,X_t(x)}(Y_{X_t(x)}) - Y_x}{t}\right)(f)$$
$$= \left.\frac{d}{dt}\right|_{t=0}\left[(X_{-t})_{*,X_t(x)}(Y_{X_t(x)})(f)\right] = \left.\frac{d}{dt}\right|_{t=0}\left[Y_{X_t(x)}(f\circ X_{-t})\right].$$

Define a real-valued function H on a neighborhood of $(0,0)\in\mathbf{R}^2$ by

$$H(t,u) := f\left(X_{-t}(Y_u(X_t(x)))\right).$$

Then by (a),

$$Y_{X_t(x)}(f\circ X_{-t}) = \left.\frac{\partial}{\partial r^2}\right|_{(t,0)} H(t,u)$$

and then

$$(\mathscr{L}_X Y)_x(f) = \left.\frac{\partial^2 H}{\partial r^1\partial r^2}\right|_{(0,0)}.$$

Set

$$K(t,u,s) := f\left(X_s(Y_u(X_t(x)))\right)$$

near a neighborhood of $(0,0,0)\in\mathbf{R}^3$. Then $H(t,u)=K(t,u,-t)$ and the chain rule implies

$$\left.\frac{\partial^2 H}{\partial r^1\partial r^2}\right|_{(0,0)} = \left.\frac{\partial^2 K}{\partial r^1\partial r^2}\right|_{(0,0,0)} - \left.\frac{\partial^2 K}{\partial r^3\partial r^2}\right|_{(0,0,0)}.$$

Since $K(t,u,0)=f(Y_u(X_t(x)))$, it follows that

$$\left.\frac{\partial K}{\partial r^2}\right|_{(t,0,0)} = Y_{X_t(x)}f = (Yf)(X_t(x)), \quad \left.\frac{\partial^2 K}{\partial r^1\partial r^2}\right|_{(0,0,0)} = X_x(Yf).$$

Similarly, using $K(0,u,s)=f(X_s(Y_u(x)))$, we arrive at

$$\left.\frac{\partial K}{\partial r^3}\right|_{(0,u,0)} = X_{Y_u(x)}f = (Xf)(Y_u(x)), \quad \left.\frac{\partial^2 K}{\partial r^2\partial r^3}\right|_{(0,0,0)} = Y_x(Xf).$$

Therefore, $(\mathscr{L}_x Y)_x(f) = X_x(Yf) - Y_x(Xf) = [X,Y]_x(f)$. By part (b), $\mathscr{L}_X Y$ is a smooth vector field.

(c) The derivation is clear. Next we check $\mathscr{L}$ commutes with d when applied to functions; that is

$$(\mathscr{L}_X(df))_x = d(\mathscr{L}_X f)_x, \quad f\in C^\infty(\mathcal{M}^m),\ x\in\mathcal{M}^m.$$

For $Y_x\in T_x\mathcal{M}$, the right-hand side gives

$$d(\mathscr{L}_X f)_x(Y_x) = Y_x(\mathscr{L}_X f) = Y_x\left(\left.\frac{d}{dt}\right|_{t=0}(f\circ X_t)\right)$$

where $f\circ X_t$ can be considered as a smooth function on $(-\varepsilon,\varepsilon)\times\mathcal{W}$ for some $\varepsilon>0$ and some neighborhood $\mathcal{W}$ of x in $\mathcal{M}$. The left-hand side gives

$$(\mathscr{L}_X(df))_x Y_x = \left(\left.\frac{d}{dt}\right|_{t=0}(X_t)^*_x(df_{X_t(x)})\right)Y_x = \left.\frac{d}{dt}\right|_{t=0}\left((X_t)^*_x(df_{X_t(x)})(Y_x)\right)$$
$$= \left.\frac{d}{dt}\right|_{t=0}\left(d(f\circ X_t)_x Y_x\right) = \left.\frac{d}{dt}\right|_{t=0}\left(Y_x(f\circ X_t)\right).$$

Let Y be an extension of Y_x to a smooth vector field on $\mathcal{W}$. Then d/dt and Y have canonical extensions to smooth vector fields $\widetilde{d/dt}$ and $\widetilde{Y}$ respectively on $(-\varepsilon,\varepsilon)\times\mathcal{W}$ where $[\widetilde{d/dt},\widetilde{Y}]=0$.

Hence

$$\begin{aligned} d(\mathscr{L}_X f)_x(Y_x) &= Y_x\left(\lim_{t\to 0}\frac{d}{dt}(f\circ X_t)\right) = Y_x\left(\lim_{t\to 0}\widetilde{\frac{d}{dt}}(f\circ X_t)\right) = \lim_{t\to 0}\widetilde{Y}\left(\widetilde{\frac{d}{dt}}(f\circ X_t)\right) \\ &= \lim_{t\to 0}\widetilde{\frac{d}{dt}}\left(\widetilde{Y}(f\circ X_t)\right) = \frac{d}{dt}\Big|_{t=0}\left(Y_x(f\circ X_t)\right) = (\mathscr{L}_X(df))_x Y_x. \end{aligned}$$

To see that the form $\mathscr{L}_X\omega$ is smooth and to check that $\mathscr{L}_X$ commutes with d on all of $\mathcal{E}^*(\mathcal{M})$, we simply express ω in local coordinates and compute using $(\mathscr{L}_X(df))_x Y_x = d(\mathscr{L}_X f)_x(Y_x)$, part (a), and the fact that $\mathscr{L}_X$ is a derivation.

The last three results can be verified by direct computations. □

1.2.4 Star transformation

Let $(V,\langle,\rangle)$ be an m-dimensional real inner product space. For $w = w_1\wedge\cdots\wedge w_p, v = v_1\wedge\cdots\wedge v_p \in \wedge^p V$, define

$$\langle w, v\rangle := \det(\langle w_i, v_j\rangle_{1\le i,j\le p}). \tag{1.2.57}$$

This defines an inner product on $\wedge^p V$ and then on $\wedge V$.

(a) If $(e_i)_{1\le i\le m}$ is an orthonormal basis of V, then $(e_{i_1}\wedge\cdots\wedge e_{i_r})_{1\le i_1<\cdots<i_r\le m}$ is an orthonormal basis of $\wedge V$.

(b) Since $\wedge^m V$ is one-dimensional, $\wedge^m V\setminus\{0\}$ has two components. An **orientation on V** is a choice of a component of $\wedge^m V\setminus\{0\}$.

Let $(V,\langle,\rangle)$ be an oriented inner product space and $\dim V = m$. The **star transformation**

$$* : \wedge V \longrightarrow \wedge V \tag{1.2.58}$$

is defined by requiring, for any orthonormal basis $\{e_i\}_{1\le i\le m}$ of V,

$$\begin{aligned} *(1) &= \pm e_1\wedge\cdots\wedge e_m, \quad *(e_1\wedge\cdots\wedge e_m) = \pm 1, \\ *(e_1\wedge\cdots\wedge e_p) &= \pm e_{p+1}\wedge\cdots\wedge e_m, \end{aligned}$$

where one takes "+" if $e_1\wedge\cdots\wedge e_m$ lies in the component of $\wedge^m V\setminus\{0\}$ determined by the orientation and "−" otherwise. Observe that

$$* : \wedge^p V \longrightarrow \wedge^{m-p} V. \tag{1.2.59}$$

Proposition 1.13

(1) On $\wedge^p V$, we have $*^2 = (-1)^{p(m-p)}$.

(2) For any $v, w\in\wedge^p V$, we have $\langle v, w\rangle = *(w\wedge *v) = *(v\wedge *w)$. ♡

Proof. (1) Compute

$$*^2(e_1\wedge\cdots\wedge e_p) = *(\pm e_{p+1}\wedge\cdots\wedge e_m) = \pm *(e_{p+1}\wedge\cdots\wedge e_m).$$

Since

$$\begin{aligned} e_1\wedge\cdots\wedge e_p\wedge e_{p+1}\wedge\cdots\wedge e_m &= (-1)^p e_{p+1}\wedge e_1\wedge\cdots\wedge e_p\wedge e_{p+2}\wedge\cdots\wedge e_m \\ &= [(-1)^p]^{m-p} e_{p+1}\wedge\cdots\wedge e_m\wedge e_1\wedge\cdots\wedge e_p \end{aligned}$$

it follows that $*(e_{p+1}\wedge\cdots\wedge e_m)=\pm(-1)^{p(m-p)}e_1\wedge\cdots\wedge e_p$ and then $*^2=(-1)^{p(m-p)}$ on $\wedge^p V$.

(2) For convenience, we may assume that the orientation is "+". Choose an orthonormal basis $\{e_i\}_{1\le i\le m}$ of V and write

$$
\begin{aligned}
v &= \sum_{1\le i_1<\cdots<i_p\le m} v^{i_1\cdots i_p}e_{i_1}\wedge\cdots\wedge e_{i_p} &=& \sum_{|I|=p} v^I e_I,\\
w &= \sum_{1\le j_1<\cdots<j_p\le m} w^{j_1\cdots j_p}e_{j_1}\wedge\cdots\wedge e_{j_p} &=& \sum_{|J|=p} w^J e_J
\end{aligned}
$$

where $I=\{1\le i_1<\cdots<i_p\le m\}$ and $e_I:=e_{i_1}\wedge\cdots\wedge e_{i_p}$. Then

$$
*(w\wedge *v)=*\left(\sum_{|I|=p} w\wedge v^I *(e_I)\right)=\sum_{|I|,|J|=p} v^I w^J *(e_J\wedge *(e_I)).
$$

For any $I=\{1\le i_1<\cdots<i_p\le m\}$, define $I^c:=\{1\le i_{p+1}<\cdots<i_m\le m\}$ where $i_{p+1},\cdots,i_m$ are obtained from the ordered set $\{1<\cdots<m\}$ by removing $i_1,\cdots,i_p$. Then

$$
*(e_I)=\mathrm{sgn}(I,I^c)e_{I^c}.
$$

Consequently,

$$
*(w\wedge *v)=\sum_{|I|,|J|=p} v^I w^J *(\mathrm{sgn}(I,I^c)e_J\wedge e_{I^c})=\sum_{I=J,|I|=p} v^I w^J\mathrm{sgn}(I,I^c)*(e_I\wedge e_{I^c}).
$$

Because $e_I\wedge e_{I^c}=\mathrm{sgn}(I,I^c)e_1\wedge\cdots\wedge e_m$, we arrive at

$$
*(w\wedge *v)=\sum_{|I|=p} v^I w^I *(1)=\sum_{|I|=p} v^I w^J=\langle v,w\rangle.
$$

Similarly, we can prove $*(v\wedge *w)=\langle v,w\rangle$. □

Let $\epsilon^*_\xi:\wedge^{p+1}V\to\wedge^p V$ denote the adjoint of left exterior multiplication by $\xi\in V$. That is

$$
\langle \epsilon^*_\xi v,w\rangle:=\langle v,\epsilon_\xi w\rangle=\langle v,\xi\wedge w\rangle,\quad v\in\wedge^{p+1}V,\ w\in\wedge^p V. \tag{1.2.60}
$$

We claim that

$$
\epsilon^*_\xi v=(-1)^{mp}*(\xi\wedge(*v)),\quad v\in\wedge^{p+1}V,\ \xi\in V. \tag{1.2.61}
$$

It suffices to prove that

$$
\langle(-1)^{mp}*(\xi\wedge(*v)),w\rangle=\langle v,\xi\wedge w\rangle
$$

for any $w\in\wedge^p V$. The left-hand side is equal to

$$
\begin{aligned}
(-1)^{mp}*(w\wedge **(\xi\wedge(*v))) &= (-1)^{mp}*\left(w\wedge(-1)^{mp-p^2}\xi\wedge(*v)\right)\\
&= (-1)^{-p^2}\langle v,w\wedge\xi\rangle \;=\; (-1)^{-p^2}\langle v,(-1)^p\xi\wedge w\rangle\\
&= (-1)^{p-p^2}\langle v,\xi\wedge w\rangle \;=\; \langle v,\xi\wedge w\rangle
\end{aligned}
$$

by Proposition 1.13.

1.2.5 Cartan's lemma

Let $\mathcal{M}$ be a differential manifold and $p \leq m$.

Theorem 1.12. (Cartan's lemma)

Let $\omega_1, \cdots, \omega_p$ be 1-forms on $\mathcal{M}$ which are linearly independent pointwise. If $\theta_1, \cdots, \theta_p$ are 1-forms on $\mathcal{M}$ with

$$\sum_{1 \leq i \leq p} \theta_i \wedge \omega_i = 0,$$

then there exist smooth functions A_{ij} on $\mathcal{M}$ with $A_{ij} = A_{ji}$ such that

$$\theta_i = \sum_{1 \leq j \leq p} A_{ij} \omega_j, \quad 1 \leq i \leq p.$$

♡

Proof. Choose $\omega_{p+1}, \cdots, \omega_m \in \mathcal{E}^1(\mathcal{M})$ so that $\omega_1, \cdots, \omega_p, \omega_{p+1}, \cdots, \omega_m$ form a basis of $T\mathcal{M}$ pointwise. Write

$$\theta_i = \sum_{1 \leq j \leq m} A_{ij} \omega_j, \quad 1 \leq i \leq p$$

for some smooth functions A_{ij} on $\mathcal{M}$. Hence

$$0 = \sum_{1 \leq i \leq p} \sum_{1 \leq j \leq m} A_{ij} \omega_j \wedge \omega_i = \sum_{1 \leq i,j \leq p} A_{ij} \omega_j \wedge \omega_i + \sum_{1 \leq i \leq p} \sum_{p+1 \leq j \leq m} A_{ij} \omega_j \wedge \omega_i.$$

Consequently,

$$0 = \sum_{1 \leq i,j \leq p} (A_{ij} - A_{ji}) \omega_j \wedge \omega_i - \sum_{1 \leq i \leq p} \sum_{p+1 \leq j \leq m} A_{ij} \omega_i \wedge \omega_j;$$

then $A_{ij} - A_{ji} = 0$ for all $1 \leq j < j \leq p$ and A_{ij} for $1 \leq i \leq p$ and $p + 1 \leq j \leq m$. Thus $\theta_i = \sum_{1 \leq j \leq p} A_{ij} \omega_j$ for each $i \in \{1, \cdots, p\}$. □

1.3 Integration on manifolds

Introduction

- *Orientation*
- *Integration on manifolds*
- *de Rham cohomology*

1.3.1 Orientation

Let V be a real space of dimension m. An **orientation on V** is a choice of a (connected) component of $\wedge^m V \setminus \{0\}$.

Definition 1.24. (Orientation)

Let $\mathcal{M}$ be a connected manifold of dimension m. Let O be the "0-section" of the exterior m-bundle $\wedge^m T^*\mathcal{M}$; that is

$$O := \bigcup_{x\in\mathcal{M}} \{0 \in \wedge^m T_x^*\mathcal{M}\}. \tag{1.3.1}$$

Since each $\wedge^m T_x^*\mathcal{M}\setminus\{0\}$ has exactly two components, it follows that $\wedge^m T^*\mathcal{M}\setminus O$ has at most two components. We say that $\mathcal{M}$ is **orientable** *if $\wedge^m T^*\mathcal{M}\setminus O$ has two components; and if $\mathcal{M}^m$ is orientable, an* **orientation on $\mathcal{M}$** *is a choice of one of the two components of $\wedge^m T^*\mathcal{M}\setminus O$.*

A non-connected manifold $\mathcal{M}$ is said to be orientable if each component of $\mathcal{M}$ is orientable, and an orientation is a choice of orientation on each component. ♣

Let $\mathcal{M}$ be oriented and let $e_1, \cdots, e_m$ be a basis of $T_x\mathcal{M}$ with dual basis $e_1^*, \cdots, e_m^*$. We say that the ordered basis $e_1, \cdots, e_m$ is **oriented** if $e_1^*\wedge\cdots\wedge e_m^*$ belongs to the orientation.

Let $\mathcal{M}$ and $\mathcal{N}$ be orientable m-dimensional manifolds, and let $F : \mathcal{M}\to\mathcal{N}$ be a smooth map. We say that F **preserves orientation** if the induced map

$$F^* : \wedge^m T^*\mathcal{N} \longrightarrow \wedge^m T^*\mathcal{M}$$

maps the component of $\wedge^m T^*\mathcal{N}\setminus O$ determining the orientation on $\mathcal{N}$ into the component of $\wedge^m T^*\mathcal{M}\setminus O$ determining the orientation on $\mathcal{M}$. Equivalently, F is orientation-preserving if $F_{*,x}$ sends oriented bases of $T_x\mathcal{M}$ into oriented bases of $T_{F(x)}\mathcal{N}$.

Proposition 1.14

Let $\mathcal{M}$ be a manifold of dimension m. Then the following are equivalent.

(a) $\mathcal{M}$ is orientable.

(b) There is a collection $\Psi := \{(\mathcal{V},\psi)\}$ of coordinate systems on $\mathcal{M}$ such that

$$\mathcal{M} = \bigcup_{(\mathcal{V},\psi)\in\Psi} \mathcal{V}, \quad \det\left(\frac{\partial x^i}{\partial y^j}\right) > 0 \text{ on } \mathcal{U}\cap\mathcal{V}$$

whenever $(\mathcal{U}, x^1, \cdots, x^m)$ and $(\mathcal{V}, y^1, \cdots, y^m)$ belong to Ψ.

(c) There is a nowhere-vanishing m-form on $\mathcal{M}$. ♡

Proof. Without loss of generality, we may assume that $\mathcal{M}$ is connected.

$(a) \Rightarrow (b)$. Given (a), choose an orientation on $\mathcal{M}$; that is, we choose one of the two components, called Λ, of $\wedge^m T^*\mathcal{M}\setminus O$. Let Ψ consist of all of those coordinate system $(\mathcal{V}, y^1, \cdots, y^m)$ on $\mathcal{M}$ such that the map of $\mathcal{V}$ into $\wedge^m T^*\mathcal{M}$ defined by

$$x \longmapsto (dy^1\wedge\cdots\wedge dy^m)(x)$$

has range in Λ. If $(\mathcal{U}, x^1, \cdots, x^m)$ and $(\mathcal{V}, y^1, \cdots, y^m)$ are any two coordinate systems on $\mathcal{M}$, then for $x \in \mathcal{U} \cap \mathcal{V}$,

$$\begin{aligned}(dx^1 \wedge \cdots \wedge dx^m)(x) &= \left(\frac{\partial x^1}{\partial y^{j_1}} dy^{j_1}\right) \wedge \cdots \wedge \left(\frac{\partial x^m}{\partial y^{j_m}} dy^{j_m}\right)(x) \\ &= \det\left(\left.\frac{\partial x^i}{\partial y^j}\right|_x\right)(dy^1 \wedge \cdots \wedge dy^m)(x).\end{aligned}$$

Then

$$\det\left(\left.\frac{\partial x^i}{\partial y^j}\right|_x\right) > 0$$

for each $x \in \mathcal{U} \cap \mathcal{V}$.

$(b) \Rightarrow (c)$. Let $\{\varphi_i\}_{i \in I}$ be a partition of unity subordinate to the cover of $\mathcal{M}$ given by the coordinate neighborhoods in the collection Ψ with φ_i subordinate to $(\mathcal{V}_i, x_i^1, \cdots, x_i^m)$. Let

$$\omega := \sum_{i \in I} \varphi_i dx_i^1 \wedge \cdots \wedge dx_i^m$$

be a global m-form on $\mathcal{M}$, where $\varphi_i dx_i^1 \wedge \cdots \wedge dx_i^m$ is defined to be the 0 outside of $\mathcal{V}_i$. On $\mathcal{V}_i \cap \mathcal{V}_j$, we have

$$\omega\left(\frac{\partial}{\partial x_j^1}, \cdots, \frac{\partial}{\partial x_j^m}\right) = \sum_{i \in I} \varphi_i \det\left(\frac{\partial x_i^\alpha}{\partial x_j^\beta}\right)$$

which is positive.

$(c) \Rightarrow (a)$. Let ω be a nowhere-vanishing m-form on $\mathcal{M}^m$ and let

$$\Lambda^+ := \bigcup_{x \in \mathcal{M}} \{a\omega_x : a \in \mathbf{R},\ a > 0\}, \quad \Lambda^- := \bigcup_{x \in \mathcal{M}^m} \{a\omega_x : a \in \mathbf{R},\ a > 0\}.$$

Then

$$\wedge^m T^*\mathcal{M} \setminus O = \Lambda^+ \coprod \Lambda^-$$

is the disjoint union of the above two open subsets Λ^+ and Λ^-, so $\wedge^m T^*\mathcal{M} \setminus O$ is disconnected and $\mathcal{M}$ is orientable. □

Example 1.5

(1) The **standard orientation** *on $\mathbf{R}^m$ is the one determined by the m-form $dr^1 \wedge \cdots \wedge dr^m$.*

(2) Let $\mathcal{M}$ be a manifold of dimension m and suppose that there is an immersion $f : \mathcal{M} \to \mathbf{R}^{m+1}$. A **normal vector field** *along $(\mathcal{M}, f)$ is a smooth map*

$$N : \mathcal{M} \longrightarrow T\mathbf{R}^{m+1}$$

such that $N(x) \in T_{f(x)}\mathbf{R}^{m+1}$ for any $x \in \mathcal{M}$ and is orthogonal to $f_{*,x}(T_x\mathcal{M}) \subset T_{f(x)}\mathbf{R}^{m+1}$. Such a manifold $\mathcal{M}$ is orientable if and only if there is a smooth nowhere-vanishing normal vector field along $(\mathcal{M}, f)$.

(3) $\mathbf{S}^m$ is orientable for each $m \geq 1$.

(4) $\mathbf{RP}^m$ is orientable if and only if m is odd. ♠

1.3.2 Integration on manifolds

Let F be a diffeomorphism of a bounded open set D in $\mathbf{R}^m$ with a bounded open set $F(D)$ and let

$$\mathcal{J}_F := \det\left(\frac{\partial F^i}{\partial r^j}\right), \quad F = (F^1, \cdots, F^m).$$

If f is a bounded continuous function on $F(D)$ and A is a nice subset of D (A will be polyhedral in most of application), then

$$\int_{F(A)} f = \int_A f \circ F |\mathcal{J}_F|. \tag{1.3.2}$$

Integration of m-forms in $\mathbf{R}^m$. The standard orientation on $\mathbf{R}^m$ is determined by the m-form $dr^1 \wedge \cdots \wedge dr^m$. If ω is an m-form on an open set $D \subset \mathbf{R}^m$, then

$$\omega = f dr^1 \wedge \cdots \wedge dx^m, \quad f \in C^\infty(D).$$

For any $A \subset D$, define

$$\int_A \omega := \int_A f \tag{1.3.3}$$

provided that the latter exists. Let F be a diffeomorphism of a bounded open set D in $\mathbf{R}^m$ with a bounded open set $F(D)$ and A be a nice subset of D. If ω is an m-form on $F(D)$, then

$$F^*\omega = f \circ F d(r^1 \circ F) \wedge \cdots \wedge d(r^m \circ F) = f \circ F dF^1 \wedge \cdots \wedge dF^m = (f \circ F)\mathcal{J}_F dr^1 \wedge \cdots \wedge dr^m.$$

Hence, by (1.3.2),

$$\int_{F(A)} \omega = \int_{F(A)} f = \int_A f \circ F |\mathcal{J}_F| = \pm \int_A f \circ F \mathcal{J}_F = \pm \int_A F^*\omega, \tag{1.3.4}$$

where one uses "$+$" if F is orientation-preserving and "$-$" if F is orientation-reversing.

Integration over chains. For each $p \geq 1$ let

$$\Delta^p := \left\{ (a^1, \cdots, a^p) \in \mathbf{R}^p : \sum_{1 \leq i \leq p} a_i \leq 0, \ a_i \geq 0 \right\}.$$

Δ^p is called the **standard p-simplex** in $\mathbf{R}^p$. For $p = 0$, we set $\Delta^0 := \{0\}$ called the **standard 0-simplex**.

Let $\mathcal{M}$ be a manifold of dimension m. A **singular p-simplex** σ in $\mathcal{M}$ is a map

$$\sigma : \Delta^p \longrightarrow \mathcal{M}$$

which extends to be a smooth map of a neighborhood of Δ^p in $\mathbf{R}^p$ into $\mathcal{M}$. A **p-chain** c in $\mathcal{M}$ (with real coefficient) is a finite linear combination

$$c = \sum_{i \in I} a^i \sigma_i, \quad |I| < \infty,$$

of p-simplices σ_i in $\mathcal{M}^m$ where $a_i \in \mathbf{R}$. For each $p \geq 0$ we define a collection of maps

$$k_i^p : \Delta^p \longrightarrow \Delta^{p+1}, \quad 0 \leq i \leq p+1$$

as follows:

(i) For $p=0$, define

$$k_0^0(0)=1, \quad k_1^0(0)=0.$$

(ii) For $p\geq 1$, define

$$\begin{aligned} k_0^p(a^1,\cdots,a^p) &:= \left(1-\sum_{1\leq i\leq p}a^i, a_1,\cdots,a^p\right), \\ k_i^p(a^1,\cdots,a^p) &:= (a_1,\cdots,a_{i-1},0,a_i,\cdots,a_p), \quad 1\leq i\leq p+1. \end{aligned}$$

If σ is a p-simplex in $\mathcal{M}^m$ with $p\geq 1$, we define its ***i*-th face** ($0\leq i\leq p$)

$$\sigma^i := \sigma\circ k_i^{p-1} \tag{1.3.5}$$

and define the **boundary** of σ

$$\partial\sigma := \sum_{0\leq i\leq p}(-1)^i\sigma^i. \tag{1.3.6}$$

If $c=\sum_{j\in J}a^j\sigma_j$ is a p-chain, then the boundary of c is given by

$$\partial c := \sum_{j\in J}a^j\partial\sigma_j = \sum_{j\in J}\sum_{0\leq i\leq p}(-1)^i a^j\sigma_j^i.$$

Clearly that

$$k_i^{p+1}\circ k_j^p = k_{j+1}^{p+1}\circ k_i^p, \quad p\geq 0, \ i\leq j. \tag{1.3.7}$$

$$\begin{tikzcd} \Delta^p \arrow[r, "k_j^p"] \arrow[d, "k_i^p"'] & \Delta^{p+1} \arrow[d, "k_i^{p+1}"] \\ \Delta^{p+1} \arrow[r, "k_{j+1}^{p+1}"'] & \Delta^{p+2} \end{tikzcd}$$

Moreover,

$$\partial^2 = 0. \tag{1.3.8}$$

For any p-simplex ($p\geq 2$) σ in $\mathcal{M}$, we have

$$\begin{aligned} \partial^2\sigma &= \partial\left(\sum_{0\leq i\leq p}(-1)^i\sigma^i\right) = \sum_{0\leq i\leq p}(-1)^i\partial\sigma^i = \sum_{0\leq i\leq p}(-1)^i\partial\left(\sigma\circ k_i^{p-1}\right) \\ &= \sum_{0\leq i\leq p}(-1)^i\sum_{0\leq j\leq p-1}(-1)^j\sigma\circ k_i^{p-1}\circ k_j^{p-2} \\ &= \left(\sum_{0\leq i\leq p, 0\leq j\leq p-1, i\leq j} + \sum_{0\leq i\leq p, 0\leq j\leq p-1, i>j}\right)(-1)^{i+j}\sigma\circ k_i^{p-1}\circ k_j^{p-2} \\ &= \sum_{0\leq i\leq p, 0\leq j\leq p-1, i\leq j}(-1)^{i+j}\sigma\circ k_{j+1}^{p-1}\circ k_i^{p-2} \\ &\quad + \sum_{0\leq i\leq p, 0\leq j\leq p-1, i>j}(-1)^{i+j}\sigma\circ k_i^{p-1}\circ k_j^{p-2} = 0. \end{aligned}$$

Let σ be a p-simplex in $\mathcal{M}$ and ω be a *continuous* p-form defined on a neighborhood of the image of σ.

(1) If $p = 0$, we define the integral of ω over σ by

$$\int_\sigma \omega := \omega(\sigma(0)). \tag{1.3.9}$$

(2) If $p \geq 1$, we define the integral of ω over σ by

$$\int_\sigma \omega := \int_{\Delta^p} \sigma^*\omega. \tag{1.3.10}$$

We now extend these integrals linearly to chains so that if $c = \sum_{i\in I} a^i \sigma_i$, then

$$\int_c \omega := \sum_{i\in I} a^i \int_{\sigma_i} \omega. \tag{1.3.11}$$

The **fundamental theorem** in calculus now can be stated as

$$\int_\sigma dF = \int_{\partial\sigma} F, \tag{1.3.12}$$

whenever $F \in C^\infty(\mathbf{R})$ and σ is a smooth 1-simplex in $\mathbf{R}$.

Theorem 1.13. (Stokes' theorem I)

Let c be a p-chain ($p \geq 1$) in a manifold $\mathcal{M}$ of dimension m and let ω be a smooth $(p-1)$-form defined on a neighborhood of the image of σ. Then

$$\int_c d\omega = \int_{\partial c} \omega. \tag{1.3.13}$$

♡

Proof. It suffices to consider the case

$$\int_\sigma d\omega = \int_{\partial\sigma} \omega$$

for any p-simplex σ. Since

$$\int_\sigma d\omega = \int_{\Delta^p} \sigma^*(d\omega) = \int_{\Delta^p} d(\sigma^*\omega)$$

by (1.3.10), and

$$\begin{aligned}
\int_{\partial\sigma} \omega &= \int_{\sum_{0\leq i\leq p}(-1)^i\sigma^i} \omega = \sum_{0\leq i\leq p} (-1)^i \int_{\sigma^i} \omega \\
&= \sum_{0\leq i\leq p} (-1)^i \int_{\Delta^{p-1}} (\sigma^i)^*\omega = \sum_{0\leq i\leq p} (-1)^i \int_{\Delta^{p-1}} (k_i^{p-1})^* (\sigma^*(\omega)).
\end{aligned}$$

Then we need only to prove

$$\int_{\Delta^p} d(\sigma^*\omega) = \sum_{0\leq i\leq p} (-1)^i \int_{\Delta^{p-1}} (k_i^{p-1})^* (\sigma^*\omega).$$

(1) $p = 1$. Then σ is a 1-complex and ω is a 0-form. Compute

$$\begin{aligned}
\int_{\Delta^1} d(\sigma^*\omega) &= \int_{\Delta^1} \frac{d}{dr}(\omega\circ\sigma)dr = \omega(\sigma(1)) - \omega(\sigma(0)), \\
\sum_{0\leq i\leq 1} (-1)^i \int_{\Delta^0} (k_i^0)^*(\sigma^*\omega) &= \int_{\Delta^0} (k_0^0)^*(\sigma^*\omega) - \int_{\Delta^0} (k_1^0)^*(\sigma^*\omega) \\
&= \int_{\Delta^0} (k_0^0)^*(\omega\circ\sigma) - \int_{\Delta^0} (k_1^0)^*(\omega\circ\sigma) \\
&= \omega(\sigma(k_0^0(0))) - \omega(\sigma(k_1^0(0))) = \omega(\sigma(1)) - \omega(\sigma(0)).
\end{aligned}$$

(2) $p \geq 2$. Write the $(p-1)$-form $\sigma^*\omega$ as

$$\sigma^*\omega := \sum_{1\leq j\leq p} a_j dr^1 \wedge \cdots \wedge \widehat{dr^j} \wedge \cdots \wedge dr^p,$$

where a_j are smooth functions on a neighborhood of Δ^p in $\mathbf{R}^p$. Compute

$$\begin{aligned}
\int_{\Delta^p} d(\sigma^*\omega) &= \sum_{1\leq j\leq p} \int_{\Delta^p} d\left(a_j dr^1 \wedge \cdots \wedge \widehat{dr^j} \wedge \cdots \wedge dr^p\right) \\
&= \sum_{1\leq j\leq p} \int_{\Delta^p} \sum_{1\leq k\leq p} \frac{\partial a_j}{\partial r^k} dr^k \wedge dr^1 \wedge \cdots \wedge \widehat{dr^j} \wedge \cdots \wedge dr^p \\
&= \sum_{1\leq j\leq p} \int_{\Delta^p} \frac{\partial a_j}{\partial r^j} dr^j \wedge dr^1 \wedge \cdots \wedge \widehat{dr^j} \wedge \cdots \wedge dr^p \\
&= \sum_{1\leq j\leq p} (-1)^{j-1} \int_{\Delta^p} \frac{\partial a_j}{\partial r^j} dr^1 \wedge \cdots \wedge dr^p.
\end{aligned}$$

We claim that

$$(k_i^{p-1})^* r^j = \begin{cases} r^j, & 1 \leq j \leq i-1, \\ 0, & j = i, \\ r^{j-1}, & i+1 \leq j \leq p, \end{cases}$$

for any $1 \leq i \leq p$, and

$$(k_0^{p-1})^* r^j = \begin{cases} 1 - \sum_{1\leq i\leq p-1} r^i, & j = 1, \\ r^{j-1}, & 1 < j \leq p. \end{cases}$$

In fact, from $k_i^{p-1} : \Delta^{p-1} \to \Delta^p$ and $(k_i^{p-1})^* : (\Delta^p)^* \to (\Delta^{p-1})^*$, we have, for any $\boldsymbol{f} = (f^1, \cdots, f^{p-1}) \in \Delta^{p-1}$,

$$\begin{aligned}
(k_i^{p-1})^* r^j(\boldsymbol{f}) &= r^j\left(k_i^{p-1}\boldsymbol{f}\right) = r^j(f^1, \cdots, f^{i-1}, 0, f^i, \cdots, f^{p-1}) \\
&= \begin{cases} f^j = r^j(\boldsymbol{f}), & 1 \leq j \leq i-1, \\ 0, & j = i, \\ f^{j-1} = r^{j-1}(\boldsymbol{f}), & i+1 \leq j \leq p \end{cases} \\
(k_0^{p-1})^* r^j(\boldsymbol{f}) &= r^j\left(k_0^{p-1}(f^1, \cdots, f^{p-1})\right) \\
&= r^j\left(1 - \sum_{0\leq i\leq p-1} f^i, f^1, \cdots, f^{p-1}\right) \\
&= \begin{cases} \left(1 - \sum_{0\leq i\leq p-1} r^i\right)(\boldsymbol{f}), & j = 1, \\ f^{j-1} = r^{j-1}(\boldsymbol{f}), & 1 < j \leq p. \end{cases}
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\sum_{0\leq i\leq p} (-1)^i \int_{\Delta^{p-1}} (k_i^{p-1})^* \left(a_j dr^1 \wedge \cdots \wedge \widehat{dr^j} \wedge \cdots \wedge dr^p\right) \\
= &\int_{\Delta^{p-1}} (k_0^{p-1})^* \left(a_j dr^1 \wedge \cdots \wedge \widehat{dr^j} \wedge \cdots \wedge dr^p\right) \\
&+ \sum_{1\leq i\leq p} (-1)^i \int_{\Delta^{p-1}} (k_i^{p-1})^* \left(a_j dr^1 \wedge \cdots \wedge \widehat{dr^j} \wedge \cdots \wedge dr^p\right);
\end{aligned}$$

the first term on the right-hand side is equal to

$$\int_{\Delta^{p-1}} a_j \left((k_0^{p-1})^*(r^1), \cdots, (k_0^{p-1})^*(r^p)\right) d(r^1 \circ k_0^{p-1}) \wedge \cdots \wedge \widehat{d(r^j \circ k_0^{p-1})} \wedge \cdots \wedge d(r^p \circ k_0^{p-1})$$

$$\int_{\Delta^{p-1}} a_j \left(1 - \sum_{1 \le i \le p-1} r^i, r^1, \cdots, r^{p-1}\right) d\left(- \sum_{1 \le i \le p-1} r^i\right) \wedge \cdots \wedge \widehat{dr^{j-1}} \wedge \cdots \wedge dr^{p-1}$$

$$\int_{\Delta^{p-1}} a_j \left(1 - \sum_{1 \le i \le p-1} r^i, r^1, \cdots, r^{p-1}\right) (-dr^{j-1}) \wedge dr^1 \wedge \cdots \wedge \widehat{dr^{j-1}} \wedge \cdots \wedge dr^{p-1}$$

$$= (-1)^{1+j-2} \int_{\Delta^{p-1}} a_j \left(1 - \sum_{1 \le i \le p-2} r^i, r^1, \cdots, r^{p-1}\right) dr^1 \wedge \cdots \wedge dr^{p-1}$$

while the second term to

$$\sum_{1 \le i \le p} (-1)^i \int_{\Delta^{p-1}} a_j(r^1, \cdots, r^{i-1}, 0, r^i, \cdots, r^{p-1})$$
$$d(r^1 \circ k_i^{p-1}) \wedge \cdots \wedge \widehat{d(r^j \circ k_i^{p-1})} \wedge \cdots \wedge d(r^p \circ k_i^{p-1})$$
$$= (-1)^j \int_{\Delta^{p-1}} a_j(r^1, \cdots, r^{j-1}, 0, r^j, \cdots, r^{p-1}) dr^1 \wedge \cdots \wedge dr^{p-1}.$$

Hence

$$\sum_{0 \le i \le p} (-1)^i \int_{\Delta^{p-1}} (k_i^{p-1})^* \left(a_j dr^1 \wedge \cdots \wedge \widehat{dr^j} \wedge \cdots \wedge dr^p\right)$$

$$(-1)^{j-1} \int_{\Delta^{p-1}} a_j \left(1 - \sum_{1 \le i \le p-1} r^i, r^1, \cdots, r^{p-1}\right) dr^1 \wedge \cdots \wedge dr^{p-1}$$

$$+ (-1)^j \int_{\Delta^{p-1}} a_j(r^1, \cdots, r^{j-1}, 0, r^j, \cdots, r^{p-1}) dr^1 \wedge \cdots \wedge dr^{p-1}.$$

Consider

$$\varphi_j(r^1, \cdots, r^{p-1}) := \begin{cases} (r^1, \cdots, r^{p-1}), & j = 1, \\ \left(1 - \sum_{1 \le i \le p-1} r^i, r^2, \cdots, r^{p-1}\right), & j = 2, \\ \left(r^2, \cdots, r^{j-1}, 1 - \sum_{1 \le i \le p-1} r^i, r^j, \cdots, r^{p-1}\right), & 3 \le j \le p. \end{cases}$$

Then $\varphi_j(\Delta^{p-1}) = \Delta^{p-1}$ and $|\mathcal{J}_{\varphi_j}| = 1$. By the change of variables formula, we obtain

$$(-1)^{j-1} \int_{\Delta^{p-1}} a_j \left(1 - \sum_{1 \le i \le p-1} r^i, r^1, \cdots, r^{p-1}\right) dr^1 \wedge \cdots \wedge dr^{p-1}$$

$$= (-1)^{j-1} \int_{\Delta^{p-1}} a_j \left(1 - \sum_{1 \le i \le p-1} r^i\varphi_j, r^1\varphi_j, \cdots, r^{p-1}\varphi_j\right) dr^1 \wedge \cdots \wedge dr^{p-1}$$

$$= (-1)^{j-1} \int_{\Delta^{p-1}} a_j \left(r^1, \cdots, r^{j-1}, 1 - \sum_{1 \le i \le p-1} r^i, r^j, \cdots, r^{p-1}\right) dr^1 \wedge \cdots \wedge dr^{p-1}.$$

Therefore we need only to show that

$$\int_{\Delta^p} \frac{\partial a_j}{\partial r^j} dr^1 \wedge \cdots \wedge dr^p$$

$$= \int_{\Delta^{p-1}} a_j \left(r^1, \cdots, r^{j-1}, 1 - \sum_{1 \le i \le p-1} r^i, r^j, \cdots, r^{p-1} \right) dr^1 \wedge \cdots \wedge dr^{p-1}$$

$$- \int_{\Delta^{p-1}} a_j(r^1, \cdots, r^{j-1}, 0, r^j, \cdots, r^{p-1}) dr^1 \wedge \cdots \wedge dr^{p-1}.$$

This can be seen as follows:

$$\begin{aligned}
\int_{\Delta^p} \frac{\partial a_j}{\partial r^j} dr^1 \wedge \cdots \wedge dr^p \;&=\; \int_{\sum_{1 \le i \le p} r^i \le 1, r_i \ge 0} \frac{\partial a_j}{\partial r^j} dr^1 \wedge \cdots \wedge dr^p \\
&\int_{\sum_{1 \le i \le p-1} t^i \le 1, t^i \ge 0} \left(\int_0^{1 - \sum_{1 \le i \le p-1} t^i} \frac{\partial a_j}{\partial r^j} dr^j \right) dt^1 \wedge \cdots \wedge dt^{p-1} \\
&\int_{\sum_{1 \le i \le p-1} t^i \le 1, t^i \ge 0} \Bigg(a_j \left(t^1, \cdots, t^{j-1}, 1 - \sum_{1 \le i \le p-1} t^i, t^j, \cdots, t^{p-1} \right) \\
&\qquad - a_j(t^1, \cdots, t^{j-1}, 0, t^j, \cdots, t^{p-1}) \Bigg) dt^1 \wedge \cdots \wedge dt^{p-1} \\
&\int_{\Delta^{p-1}} \Bigg(a_j \left(r^1, \cdots, r^{j-1}, 1 - \sum_{1 \le i \le p-1} r^i, r^j, \cdots, r^{p-1} \right) \\
&\qquad - a_j(t^1, \cdots, t^{j-1}, 0, t^j, \cdots, t^{p-1}) \Bigg) dr^1 \wedge \cdots \wedge dr^{p-1}
\end{aligned}$$

by the fundamental theorem of calculus. □

Integration on an oriented manifold. Let $\mathcal{M}$ be an m-dimensional oriented manifold. A subset $\mathcal{D}$ of $\mathcal{M}$ will be called a **regular domain** if for each point $x \in \mathcal{M}$ one of the following holds:

(a) there is an open neighborhood of x which is contained in $\mathcal{M} \setminus \mathcal{D}$;

(b) there is an open neighborhood of x which is contained in $\mathcal{D}$;

(c) there is a centered coordinate system $(\mathcal{U}, \varphi)$ about x such that $\varphi(\mathcal{U} \cap \mathcal{D}) = \varphi(\mathcal{U}) \cap \mathbf{H}^m$, where $\mathbf{H}^m$ is the half-space of $\mathbf{R}^m$ defined by $r^m \ge 0$.

Points of $\mathcal{D}$ of type (b) are called **interior points** ($\mathrm{Int}(\mathcal{D})$ or $\mathcal{D}^\circ$). Points of $\mathcal{D}$ of type (c) are called **boundary points** ($\partial\mathcal{D}$). $\partial\mathcal{D}$ is an embedded $(m-1)$-dimensional submanifold of $\mathcal{M}$.

Let $x \in \partial\mathcal{D}$ and $v \in T_x\mathcal{M}$. We call v an **outer vector to $\mathcal{D}$** if for each smooth curve $\alpha(t)$ in $\mathcal{M}^m$ with $\dot{\alpha}(0) = v$, there exists an $\varepsilon > 0$ such that $\alpha(t) \notin \mathcal{D}$ for all $t \in (0, \epsilon)$. Let v be an outer to $\partial\mathcal{D}$ at x, and let $v_1, \cdots, v_{m-1}$ be a basis of $T_x\partial\mathcal{D}$. Then we define $v_1, \cdots, v_{m-1}$ be an oriented basis of $T_x\partial\mathcal{D}$ if and only if $v, v_1, \cdots, v_{m-1}$ is an oriented basis of $T_x\mathcal{M}$. This definition is independent of the choice of the outer vector v and defines a smooth orientation on $\partial\mathcal{D}$.

An m-simplex σ in $\mathcal{M}$ is called **regular** if σ extends to a diffeomorphism on a neighborhood of Δ^m. An **oriented regular m-simplex** is one in which the map σ preserves orientation.

(i) Associated with a given regular domain $\mathcal{D}$, we consider only oriented regular m-simplices of the following two types:

  (α) $\sigma(\Delta^m) \subset \mathrm{Int}(\mathcal{D})$.

  (β) $\sigma(\Delta^m) \subset \mathcal{D}$ and $\sigma(\Delta^m) \cap \partial\mathcal{D} = \sigma(\Delta^{m-1})$; that is, precisely the m-th face of σ lies in $\partial\mathcal{D}$.

(ii) Cover $\mathcal{D}$ by open sets $\mathcal{U}$ of the following types:

  (α') $\mathcal{U}$ lies in the interior of an oriented regular m-simplex σ of type (α).

  (β') $\mathcal{U}$ is the image under a type (β) oriented regular m-simplex σ of an open set $\mathcal{V} \subset \mathbf{R}^m$ which is a neighborhood of a point in the m-th face of Δ^m, which intersects the boundary of Δ^m only in that m-th face, and whose image under σ is contained in $\sigma(\Delta^m) \cup (\mathcal{M} \setminus \mathcal{D})$.

Let ω be a *continuous* m-form with compact support and let $\mathcal{D}$ be a regular domain in $\mathcal{M}^m$. Since $\mathrm{supp}(\omega) \cap \mathcal{D}$ is compact, we can cover $\mathrm{supp}(\omega) \cap \mathcal{D}$ by a finitely many open sets $\mathcal{U}_1, \cdots, \mathcal{U}_k$ of type (α') or (β'). Let the associated oriented regular m-simplices be $\sigma_1, \cdots, \sigma_k$. Write $\mathcal{U} := \mathcal{M} \setminus (\mathrm{supp}(\omega) \cap \mathcal{D})$. Then $\mathcal{U}, \mathcal{U}_1, \cdots, \mathcal{U}_k$ is a cover of $\mathcal{M}$ and, therefore, there exists a partition of unity $\varphi, \varphi_1, \cdots, \varphi_k$ subordinate to this cover. Define

$$\int_{\mathcal{D}} \omega := \sum_{1 \le i \le k} \int_{\sigma_i} \varphi_i \omega. \tag{1.3.14}$$

Proposition 1.15

The definition (1.3.14) is independent of the cover and the partition of unity chosen. ♡

Proof. Let $\mathcal{V}, \mathcal{V}_1, \cdots, \mathcal{V}_\ell$ and $\psi, \psi_1, \cdots, \psi_\ell$ be another such cover and another such partition of unity respectively, with $\mathcal{V}_j$ associated with the oriented regular m-simplex τ_j. Since $\psi = 0$ on $\mathrm{supp}(\omega) \cap \mathcal{D}$, it follows that $\sum_{1 \le j \le \ell} \psi_j = 1$ on $\mathrm{supp}(\omega) \cap \mathcal{D}$ and

$$\sum_{1 \le i \le k} \int_{\sigma_i} \varphi_i \omega = \sum_{1 \le i \le k} \int_{\sigma_i} \sum_{1 \le j \le \ell} \psi_j \varphi_i \omega = \sum_{1 \le i \le k, 1 \le j \le \ell} \int_{\sigma_i} \psi_j \varphi_i \omega.$$

Similarly,

$$\sum_{1 \le j \le \ell} \int_{\tau_j} \psi_j \omega = \sum_{1 \le i \le k, 1 \le j \le \ell} \int_{\tau_j} \psi_j \varphi_i \omega.$$

Since $\sigma_i^{-1} \circ \tau_j$ is an orientation-preserving diffeomorphism on the open set where it is defined and $(\text{supp}(\psi_j\varphi_i\omega)) \cap \sigma_i(\Delta^m) = (\text{supp}(\psi_j\varphi_i\omega)) \cap \tau_j(\Delta^m)$, we arrive at

$$\begin{aligned}\int_{\sigma_i} \psi_i\varphi_i\omega &= \int_{\Delta^m} \sigma_i^*(\psi_j\varphi_i\omega) = \int_{(\sigma_i^{-1}\circ\tau_j)(\Delta^m)} \sigma_i^*(\psi_j\varphi_i\omega)\\ &= \int_{\Delta^m} (\sigma_i^{-1}\circ\tau_j)^*\left(\sigma_i^*(\psi_j\varphi_i\omega)\right)\\ &= \int_{\Delta^m} \tau_j^*(\sigma_i^{-1})^*\left(\sigma_i^*(\psi_j\varphi_i\omega)\right) = \int_{\Delta^m} \tau_j^*(\psi_j\varphi_i\omega) = \int_{\tau_j} \psi_j\varphi_j\omega.\end{aligned}$$

by (1.3.2). □

Using again (1.3.2), we can prove that

$$\int_{F(\mathcal{D})} \omega = \pm\int_{\mathcal{D}} F^*\omega \tag{1.3.15}$$

whenever F is a diffeomorphism on $\mathcal{M}$, where "+" if and only if F is orientation-preserving.

Theorem 1.14. (Stokes' theorem II)

Let $\mathcal{D}$ be a regular domain in an oriented m-dimensional manifold $\mathcal{M}$, and let ω be a smooth $(m-1)$-form of compact support. Then

$$\int_{\mathcal{D}} d\omega = \int_{\partial\mathcal{D}} \omega. \tag{1.3.16}$$

♡

Proof. Let $\varphi_1, \cdots, \varphi_k$ and $\sigma_1, \cdots, \sigma_k$ be chosen as in (1.3.14) relative to $(\text{supp}(\omega))\cap\mathcal{D}$. Since $\sum_{1\le i\le k}\varphi_i \equiv 1$ on a neighborhood of $(\text{supp}(\omega))\cap\mathcal{D}$ we have

$$\sum_{1\le i\le k} d(\varphi_i\omega) = \sum_{1\le i\le k}(d\varphi_i\wedge\omega + \varphi_i d\omega) = d\omega.$$

If σ_i is an m-simplex of type (α), then

$$\int_{\partial\sigma_i}\varphi_i\omega = 0 = \int_{\partial\mathcal{D}}\varphi_i\omega$$

since $\text{supp}(\varphi_i\omega) \subset \text{Int}(\sigma_i(\Delta^m)) \subset \text{Int}(\mathcal{D})$. If σ_i is an m-simplex of type (β), then $\varphi_i\omega$ is zero on the boundary of σ_i except possibly at points in the interior of σ_i^m. Hence

$$\int_{\partial\sigma_i}\varphi_i\omega = (-1)^m\int_{\sigma_i^m}\varphi_i\omega = (-1)^m(-1)^m\int_{\partial\mathcal{D}}\varphi_i\omega = \int_{\partial\mathcal{D}}\varphi_i\omega$$

because σ_i^m is an orientation-preserving regular $(m-1)$-simplex in $\partial\mathcal{D}$ if m is even, and is orientation-revising if m is odd. therefore

$$\begin{aligned}\int_{\mathcal{D}} d\omega &= \int_{\mathcal{D}}\sum_{1\le i\le k} d(\varphi_i\omega) = \sum_{1\le i\le k}\int_{\mathcal{D}} d(\varphi_i\omega) = \sum_{1\le i\le k}\int_{\sigma_i} d(\varphi_i\omega)\\ &= \sum_{1\le i\le k}\int_{\partial\sigma_i}\varphi_i\omega = \sum_{1\le i\le k}\int_{\partial\mathcal{D}}\varphi_i\omega = \int_{\partial\mathcal{D}}\omega\end{aligned}$$

by Theorem 1.13. □

A manifold is called **closed** if it is both compact and without boundary.

Corollary 1.9

Let ω be a smooth $(m-1)$-form on a closed oriented m-dimensional manifold $\mathcal{M}$. Then

$$\int_{\mathcal{M}} d\omega = 0.$$

♡

Integration on a Riemannian manifold. Let $\mathcal{M}$ be a **Riemannian manifold** of dimension m. That is, $\mathcal{M}$ is an m-dimensional manifold with a positive definite inner product $g_x := \langle \, , \rangle_x$ on each tangent space $T_x\mathcal{M}$ such that $x \mapsto \langle X, Y\rangle_x := \langle X_x, Y_x\rangle_x$ is a smooth function on $\mathcal{M}$ whenever X and Y are smooth vector fields on $\mathcal{M}$.

(1) The existence of Riemannian metrics on manifolds is obvious by using the partition of unity.

(2) Given a point $x \in \mathcal{M}$ we can always find[3] a neighborhood $\mathcal{U}$ of x and a collection $e_1, \cdots, e_m$ of smooth vector fields on $\mathcal{U}$ which are orthonormal in the sense that they form an orthonormal basis of the tangent space to $\mathcal{M}$ at each point of $\mathcal{U}$. such a collection $\{e_i\}_{1\le i\le m}$ is called a **local orthonormal frame field**.

(3) Since the inner product $\langle \, , \rangle_x$ is a non-singular pairing of $T_x\mathcal{M}$ with itself, it induces an isomorphism of $T_x\mathcal{M}$ with $T_x^*\mathcal{M}$,

$$\varphi : T_x\mathcal{M} \longrightarrow T_x^*\mathcal{M}, \quad v \longmapsto \varphi_v, \quad \varphi_v(w) := \langle v, w\rangle_x. \tag{1.3.17}$$

Consequently, $T^*\mathcal{M}$ is also a Riemannian manifold.

(4) Let $e_1, \cdots, e_m$ be a local orthonormal frame field on $\mathcal{U}$ and let $\omega_1, \cdots, \omega_m$ be the dual 1-forms. That is

$$\omega_i(e_j) = \delta_{ij} \quad \text{on} \quad \mathcal{U}. \tag{1.3.18}$$

We call $\{\omega_i\}_{1\le i\le m}$ a **local orthonormal coframe field** on $\mathcal{U}$.

(5) If $\{\omega_i\}_{1\le i\le m}$ and $\{\omega_i'\}_{1\le i\le m}$ are two local orthonormal coframe fields on $\mathcal{U}$ and $\mathcal{U}'$ respectively, then, on $\mathcal{U}\cap\mathcal{U}'$,

$$\omega_1 \wedge \cdots \wedge \omega_m = \pm\omega_1' \wedge \cdots \wedge \omega_m'.$$

Let $(\mathcal{M}, g)$ be an m-dimensional oriented Riemannian manifold. A local coframe field $\{\omega_i\}_{1\le i\le m}$ on $\mathcal{U}$ will be called **oriented** if $\omega_1 \wedge \cdots \wedge \omega_m$ belongs to the orientation at each point of $\mathcal{U}$. Choose a local oriented orthonormal coframe field $\{\omega_i\}_{1\le i\le m}$ at each point of $\mathcal{M}$. By (5) above, we have a globally defined m-form

$$\omega := \omega_1 \wedge \cdots \wedge \omega_m \tag{1.3.19}$$

nowhere-vanishing on $\mathcal{M}$. This form is called the **volume form** of the oriented Riemannian manifold $\mathcal{M}$. The **volume of $\mathcal{M}$** is

$$\mathrm{Vol}(g) := \int_{\mathcal{M}} \omega = \int_{\mathcal{M}} \omega_1 \wedge \cdots \wedge \omega_m. \tag{1.3.20}$$

[3]Start with a coordinate system $(\mathcal{U}, x^1, \cdots, x^m)$, apply the Gram-Schmidt procedure to orthonormalize $\partial/\partial x^1$, $\cdots$,$\partial/\partial x^m$, and do it simultaneously at all points of $\mathcal{U}$.

We define $*$ on $\wedge T_x^*\mathcal{M}$ so that get a linear operator

$$* : \mathcal{E}^p(\mathcal{M}) \longrightarrow \mathcal{E}^{m-p}(\mathcal{M}) \tag{1.3.21}$$

satisfying

$$*^2 = (-1)^{p(m-p)} \quad \text{on } \mathcal{E}^p(\mathcal{M}). \tag{1.3.22}$$

If f is a continuous function with compact support, we define

$$\int_{\mathcal{M}} f := \int_{\mathcal{M}} *f = \int_{\mathcal{M}} f\omega. \tag{1.3.23}$$

Note 1.8

We can define $\int_{\mathcal{M}} f$ on any Riemannian manifold $\mathcal{M}$. Choosing $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ a cover of $\mathcal{M}$ by interiors of regular m-simplices $\{\sigma_\alpha\}'_{\alpha\in A}$ and $\{\omega_i^\alpha\}_{1\le i\le m}$ a local orthonormal coframe field defined on a neighborhood of $\sigma_\alpha(\Delta^m)$, we can find smooth functions $\{h_\alpha\}_\alpha$ on neighborhoods of Δ^m such that

$$\sigma_\alpha^*(\omega_1^\alpha \wedge \cdots \wedge \sigma_m^\alpha) = h_\alpha dr^1 \wedge \cdots \wedge dr^m.$$

Let $\{\varphi_\alpha\}_{\alpha\in A}$ be a partition of unity subordinate to the cover $\{\mathcal{U}_\alpha\}_{\alpha\in A}$ and let f be a continuous function with compact support on $\mathcal{M}$. Define

$$\int_{\mathcal{M}} f := \sum_{\alpha\in A} \int_{\Delta^m} ((\varphi_\alpha f)\circ\sigma_\alpha)\,|h_\alpha| dr^1 \wedge \cdots \wedge dr^m. \tag{1.3.24}$$

This definition is independent of the cover and partition of unity chosen. In the case of an oriented Riemannian manifold, (1.3.23) agrees with (1.3.24). ♣

If f is a smooth function on $\mathbf{R}^m$, its gradient is defined by

$$\mathrm{grad}(f) := \sum_{1\le i\le m} \frac{\partial f}{\partial r^i}\frac{\partial}{\partial r^i}.$$

If $V = \sum_{1\le i\le m} v^i \frac{\partial}{\partial r^i}$ is a smooth vector field, its divergence is defined by

$$\mathrm{div}(V) := \sum_{1\le i\le m} \frac{\partial v^i}{\partial r^i}.$$

Let $(\mathcal{M}, g)$ be an oriented Riemannian manifold. If $v \in T_x\mathcal{M}$, we write

$$v_\flat = \varphi(v)$$

according to (1.3.16). If $\omega \in T_x^*\mathcal{M}$, we write

$$\omega^\sharp := \varphi^{-1}(\omega).$$

If f is a smooth function on $\mathcal{M}$, its **gradient** is the vector field

$$\mathrm{grad}(f) := (df)^\sharp. \tag{1.3.25}$$

If V is a smooth vector field, its **divergence** is the function

$$\mathrm{div}(V) := *d*V_\flat. \tag{1.3.26}$$

observe that $\mathrm{div}(V)$ is defined independent of orientability.

Theorem 1.15. (Divergence theorem)

If V is a smooth vector field with compact support on an oriented Riemannian manifold $\mathcal{M}$, if $\mathcal{D}$ is a regular domain in $\mathcal{M}$ and if $\mathbf{n}$ is the unit outer normal vector field on $\partial\mathcal{D}$, then

$$\int_{\mathcal{D}} \operatorname{div} V = \int_{\partial\mathcal{D}} \langle V, \mathbf{n} \rangle. \tag{1.3.27}$$

♡

Let $\mathcal{M}$ be an oriented Riemannian manifold, let f and g are smooth functions on $\mathcal{M}$, and let $\mathcal{D}$ be a regular domain in $\mathcal{M}$. The **Laplacian** of g, denoted Δg, is defined by

$$\Delta g := *d * dg. \tag{1.3.28}$$

If $\mathbf{n}$ is the unit outer normal vector field along $\partial\mathcal{D}$, we let $\frac{\partial g}{\partial \mathbf{n}}$ denote $\mathbf{n}(g)$. **Green's identities** say that

$$\begin{aligned} \int_{\mathcal{D}} f\Delta g &= -\int_{\mathcal{D}} \langle \operatorname{grad}(f), \operatorname{grad}(g) \rangle + \int_{\partial\mathcal{D}} f \frac{\partial g}{\partial \mathbf{n}}, & (1.3.29) \\ \int_{\mathcal{D}} (f\Delta g - g\Delta f) &= \int_{\partial\mathcal{D}} \left(f\frac{\partial g}{\partial \mathbf{n}} - g\frac{\partial f}{\partial \mathbf{n}} \right). & (1.3.30) \end{aligned}$$

1.3.3 de Rham cohomology

A p-form on a differentiable manifold $\mathcal{M}$ is called **closed** if $d\alpha = 0$. It is called **exact** if $\alpha = d\beta$ for some smooth $(p-1)$-form β. Write

$$\mathcal{Z}^p(\mathcal{M}) := \{\text{closed } p\text{-forms}\}, \quad \mathcal{B}^p(\mathcal{M}) := \{\text{exact } p\text{-forms}\}.$$

Since $\mathcal{B}^p(\mathcal{M}) \subset \mathcal{Z}^p(\mathcal{M})$, we define the ***p*-th de Rham cohomology group of $\mathcal{M}$**

$$H^p_{\mathrm{dR}}(\mathcal{M}) := \frac{\mathcal{Z}^p(\mathcal{M})}{\mathcal{B}^p(\mathcal{M})}. \tag{1.3.31}$$

Example 1.6

For the unit circle $\mathbf{S}^1$, we have

$$H^p_{\mathrm{dR}}(\mathbf{S}^1) = \begin{cases} \mathbf{R}, & p = 0, 1, \\ 0, & p \geq 2. \end{cases} \tag{1.3.32}$$

Since $\mathbf{S}^1$ is connected, it follows from Theorem 1.3 that f is constant. Hence $H^0_{\mathrm{dR}}(\mathbf{S}^1) \cong \mathbf{R}$. The polar coordinate function θ on $\mathbf{S}^1$ is not well-defined globally, but $d\theta$ is a globally well-defined nowhere-vanishing 1-form on $\mathbf{S}^1$ because $d\theta$ is the volume form of the induced Riemannian metric on $\mathbf{S}^1$ from $\mathbf{R}^2$.

(a) $d\theta$ is not exact. Otherwise, $d\theta = d\alpha$ for some smooth function f on $\mathbf{S}^1$. By Corollary 1.9, we get

$$0 = \int_{\mathbf{S}^1} d\alpha = \int_{\mathbf{S}^1} d\theta = 2\pi - 0 = 2\pi.$$

(b) All 1-forms on $\mathbf{S}^1$ are closed.

(c) If α is a 1-form, then $\alpha - c\,d\theta$ is exact for some constant c. Let $\alpha = f(\theta)d\theta$ and

define

$$c := \frac{1}{2\pi}\int_{\mathbf{S}^1} f(\theta)d\theta, \quad g(\theta) := \int_0^\theta [f(\tau)-c]d\tau.$$

Since

$$\begin{aligned}
g(\theta+2\pi n) &= \int_0^{\theta+2\pi n}[f(\tau)-c]d\tau \\
&= \int_0^\theta [f(\tau)-c]d\tau + \int_\theta^{\theta+2\pi n}[f(\tau)-c]d\tau \\
&= \int_0^\theta [f(\tau)-c]d\tau + \sum_{1\le k\le n}\int_{\theta+2\pi(k-1)}^{\theta+2\pi k}[f(\tau)-c]d\tau \\
&= \int_0^\theta [f(\tau)-c]d\tau + \sum_{1\le k\le n}\int_0^{2\pi} f(\tau)d\tau - 2\pi cn \\
&= \int_0^\theta [f(\tau)-c]d\tau + n2\pi c - 2\pi cn \ = \ g(\theta),
\end{aligned}$$

the function $g(\theta)$ is a well-defined smooth function on $\mathbf{S}^1$. Moreover

$$dg = [f(\theta)-c]d\theta = \alpha - cd\theta.$$

Thus $\alpha - cd\theta$ is exact.

Now the result follows from (a) – (c).

Let $F:\mathcal{M}\to\mathcal{N}$ be a smooth map. Since $d\circ F^* = F^*\circ d$

$$\begin{array}{ccc}
\mathcal{E}^*(\mathcal{N}) & \xrightarrow{F^*} & \mathcal{E}^*(\mathcal{M}) \\
{\scriptstyle d}\big\downarrow & & \big\downarrow{\scriptstyle d} \\
\mathcal{E}^*(\mathcal{N}) & \xrightarrow[F^*]{} & \mathcal{E}^*(\mathcal{M})
\end{array}$$

we have an induced homomorphism

$$F^*: H^p_{\rm dR}(\mathcal{N}) \longrightarrow H^p_{\rm dR}(\mathcal{M}) \tag{1.3.33}$$

for each integer $p\ge 0$. For $\alpha\in\mathcal{Z}^p(\mathcal{N})$, $d(F^*\alpha)=F^*(d\alpha)=0$ implies $F^*\alpha\in\mathcal{Z}^p(\mathcal{M})$. If $\alpha\in\mathcal{B}^p(\mathcal{N})$, then $d\beta=\alpha$ for some $\beta\in\mathcal{E}^{p-1}(\mathcal{N})$ and $F^*\alpha = F^*(d\beta) = d(F^*\beta)$; hence $F^*\alpha\in\mathcal{Z}^p(\mathcal{M})$.

(a) If $G:\mathcal{N}\to\mathcal{P}$ is another smooth map, then

$$(G\circ F)^* = F^*\circ G^*. \tag{1.3.34}$$

(b) The identity map $\mathbf{1}:\mathcal{M}\to\mathcal{M}$ induces

$$\mathbf{1}^* = \mathbf{1} \tag{1.3.35}$$

on the de Rham cohomology groups.

For each integer $p\ge 0$, we let $C_{{\rm sing},p}(\mathcal{M};\mathbf{R})$ denote the real vector space generated by the singular p-simplices in $\mathcal{M}$. That is, an element of $C_{{\rm sing},p}(\mathcal{M};\mathbf{R})$ is a singular p-chain in $\mathcal{M}$ with real coefficients. For $p<0$, define $C_{{\rm sing},p}(\mathcal{M};\mathbf{R})$ be the zero vector space. The boundary operator ∂ induces linear transformations

$$\partial_p : C_{\mathrm{sing},p}(\mathcal{M};\mathbf{R}) \longrightarrow C_{\mathrm{sing},p-1}(\mathcal{M};\mathbf{R}) \tag{1.3.36}$$

for each $p \in \mathbf{Z}$. Since $\partial_p \circ \partial_{p+1} = 0$, we define the **$p$-th differential singular homology group of $\mathcal{M}$ with real coefficients** by

$$H_{\mathrm{sing},p}(\mathcal{M};\mathbf{R}) := \frac{\ker(\partial_p)}{\mathrm{Im}(\partial_{p+1})}. \tag{1.3.37}$$

Elements of $\ker(\partial_p)$ are called **differentiable p-cycles** and elements of $\mathrm{Im}(\partial_{p+1})$ are called **differentiable p-boundaries**. Define a linear map

$$\varphi : H^p_{\mathrm{dR}}(\mathcal{M}) \longrightarrow H_{\mathrm{sing},p}(\mathcal{M};\mathbf{R})^*, \quad [\alpha] \longmapsto \varphi([\alpha])([z]) := \int_z \alpha. \tag{1.3.38}$$

We shall prove that (1.3.38) is well-defined. If $\alpha' = \alpha + d\beta$ for some $\beta \in \mathcal{E}^{p-1}(\mathcal{M})$ and if $z' = z + \partial_p w$ for some $w \in C_{\mathrm{sing},p+1}(\mathcal{M};\mathbf{R})$, we have

$$\begin{aligned}
\int_{z'} \alpha' &= \int_{z+\partial_p w} (\alpha + d\beta) = \int_z (\alpha + d\beta) + \int_{\partial_p w} (\alpha + d\beta) \\
&= \int_z \alpha + \int_z d\beta + \int_{\partial_p w} \alpha + \int_{\partial_p w} d\beta \\
&= \int_z \alpha + \int_{\partial z} \beta + \int_w d\alpha + \int_w d^2\beta = \int_z \alpha.
\end{aligned}$$

Theorem 1.16. (de Rham theorem)

If $\mathcal{M}$ is compact manifold, then (1.3.38) is an isomorphism. ♡

Chapter 2 Riemannian geometry

Introduction

2.1 Introduction

A fundamental question in Riemannian geometry is that given a restriction on the curvature of a Riemannian manifold, what topological conditions follow?

(1) **Myers' theorem:** If $\mathcal{M}$ is a complete m-dimensional manifold with Ricci curvature bounded below by a positive constant $(m-1)K$, then $\mathcal{M}$ has diameter at most $\pi/\sqrt{K}$. Topological consequence is that $\mathcal{M}$ is compact and has finite fundamental group.

(2) **Cartan-Hadamard theorem:** If $\mathcal{M}$ is a simply-connected, complete m-dimensional manifold with nonpositive sectional curvature, then $\mathcal{M}$ is diffeomorphic to $\mathbf{R}^m$ and each exponential map is a diffeomorphism.

2.2 Metrics, connections, curvatures and covariant differentiation

Introduction

Let $\mathcal{M}$ be an m-dimensional (smooth) manifold and $\odot^2 T^*\mathcal{M} = T^*\mathcal{M} \otimes_S T^*\mathcal{M}$ denote the subspace of $T^*\mathcal{M} \otimes T^*\mathcal{M}$ generated by elements of the form $X \otimes Y + Y \otimes X$. We also consider the subspace $\odot^2_+ T^*\mathcal{M}$, consisting of all positive-definite symmetric covariant 2-tensor fields, of $\odot^2 T^*\mathcal{M}$. Then a Riemannian metric g can be viewed as a section of the bundle $\odot^2_+ T^*\mathcal{M}$.

2.2.1 Metrics and connections

If V is a tangent field of $\mathcal{M}$, we denote by V_p or (p, V) the tangent vector at the point p. If we consider the tangent space $T_p\mathcal{M}$, its element is written as v.

Example 2.1

The Euclidean space $\mathbf{E}^m := (\mathbf{R}^m, g_{\text{can}})$ *is the simplest Riemannian manifold. The tangent bundle* $T\mathbf{R}^m$ *is naturally identified with the product manifold* $\mathbf{R}^m \times \mathbf{R}^m$ *via the map*

$$(p, V) \in \mathbf{R}^m \times \mathbf{R}^m \longmapsto \left(p, \frac{d}{dt}\Big|_{t=0}(p+tV)\right) \in T\mathbf{R}^m$$

so that the standard or canonical metric g_{can} *on* $\mathbf{R}^m$ *is defined by*

$$g_{\text{can}}\left((p, V), (p, W)\right) := V \cdot W.$$

♠

A **Riemannian isometry** between Riemannian manifolds $(\mathcal{M}, g_{\mathcal{M}})$ and $(\mathcal{N}, g_{\mathcal{N}})$ is a diffeomorphism $\varphi : \mathcal{M} \to \mathcal{N}$ such that $\varphi^* g_{\mathcal{N}} = g_{\mathcal{M}}$, i.e.,

$$g_{\mathcal{N}}(d\varphi(V), d\varphi(W)) = g_{\mathcal{M}}(V, W) \tag{2.2.1}$$

for all tangent fields V and W. In this case φ^{-1} is also a Riemannian isometry.

Suppose that we have an immersion (or embedding) $\varphi : \mathcal{M} \to \mathcal{N}$, and that $(\mathcal{N}, g_{\mathcal{N}})$ is a Riemannian manifold. We can then construct a Riemannian manifold on $\mathcal{M}$ by pulling back $g_{\mathcal{N}}$ to $g_{\mathcal{M}} := \varphi^* g_{\mathcal{N}}$ on $\mathcal{M}$, i.e.,

$$g_{\mathcal{M}}(V, W) = g_{\mathcal{N}}(d\varphi(V), d\varphi(W)).$$

Note that if $g_{\mathcal{M}}(V, V) = 0$, then since φ is an immersion, we have $V = 0$.

A **Riemannian immersion** (or **Riemannian embedding**) is thus an immersion (or embedding) $\varphi : (\mathcal{M}, g_{\mathcal{M}}) \to (\mathcal{N}, g_{\mathcal{N}})$ such that $g_{\mathcal{M}} = \varphi^* g_{\mathcal{N}}$. Riemannian immersions are also called **isometric immersions**, but as we shall see below they are *almost never* distance preserving.

Example 2.2

Define

$$\mathbf{S}^m(r) := \{\boldsymbol{x} \in \mathbf{R}^{m+1} : |\boldsymbol{x}| = r\}.$$

The metric induced from the embedding $\mathbf{S}^m(r) \hookrightarrow \mathbf{R}^{m+1}$ *is the canonical metric on* $\mathbf{S}^m(r)$. *The unit sphere, or standard sphere, is* $\mathbf{S}^m := \mathbf{S}^m(1) \subset \mathbf{R}^{m+1}$ *with the induced metric.*

♠

A **Riemannian submersion** $\varphi : (\mathcal{M}, g_{\mathcal{M}}) \to (\mathcal{N}, g_{\mathcal{N}})$ is a submersion $\varphi : \mathcal{M} \to \mathcal{N}$ such that for each $p \in \mathcal{M}$, $d\varphi_p : \mathrm{Ker}^{\perp}(d\varphi_p) \to T_{\varphi(p)}\mathcal{N}$ is a linear isometry. In other words, if $(p, V), (p, W) \in T_p\mathcal{M}$ are perpendicular to the kernel of $d\varphi_p : T_p\mathcal{M} \to T_{\varphi(p)}\mathcal{N}$, then

$$g_{\mathcal{M}}(V, W) = g_{\mathcal{N}}(d\varphi(V), d\varphi(W)).$$

This is also equivalent to saying that the adjoint $(d\varphi_p)^* : T_{\varphi(p)}\mathcal{N} \to T_p\mathcal{M}$ preserves inner products of vectors.

Example 2.3. (Hopf fibration)

The Hopf fibration $\mathbf{S}^3(1) \to \mathbf{S}^2(1/2)$ *can be written as*

$$(z, w) \longmapsto \left(\frac{|w|^2 - |z|^2}{2}, z\overline{w}\right) \tag{2.2.2}$$

if we think of $\mathbf{S}^3(1) \subset \mathbf{C}^2$ *and* $\mathbf{S}^2(1/2) \subset \mathbf{R} \oplus \mathbf{C}$. *The fiber containing* (z, w) *consists of the points* $(e^{\sqrt{-1}\theta}z, e^{\sqrt{-1}\theta}w)$, *and hence the tangent vectors that are perpendicular to those points are of the form* $\lambda(-\overline{w}, \overline{z})$, $\lambda \in \mathbf{C}$. *Calculate*

$$\left(\frac{|w + \lambda\overline{z}|^2 - |z - \lambda\overline{w}|^2}{2}, (z - \lambda\overline{w})\overline{(w + \lambda\overline{z})}\right) = \left(2\mathrm{Re}(\overline{\lambda}zw), -\lambda\overline{w}^2 + \overline{\lambda}z^2\right).$$

This term has length $|\lambda|$ *as well as the length of* $\lambda(-\overline{w}, \overline{z})$. *Hence the map is a Riemannian submersion.* ♠

For a Riemannian manifold $(\mathcal{M}, g)$, let $\mathrm{Iso}(\mathcal{M}, g)$ denote the group of Riemannian isometries $\varphi : (\mathcal{M}, g) \to (\mathcal{M}, g)$ and $\mathrm{Iso}_p(\mathcal{M}, g)$ the **isometry group at** p, i.e., those $\varphi \in \mathrm{Iso}(\mathcal{M}, g)$ with $\varphi(p) = p$. A Riemannian manifold is said to be **homogeneous** if its isometry group acts transitively, i.e., for any points $p, q \in \mathcal{M}$, there is an $\varphi \in \mathrm{Iso}(\mathcal{M}, g)$ such that $\varphi(p) = q$.

Example 2.4

(1) $\mathrm{Iso}(\mathbf{R}^m, g_{\mathrm{can}}) = \mathbf{R}^m \rtimes \mathbf{O}(m)$, *i.e.,*

$$\mathrm{Iso}(\mathbf{R}^m, g_{\mathrm{can}}) = \{\varphi|\mathbf{R}^m \to \mathbf{R}^m : F(x) = V + Ox,\ V \in \mathbf{R}^m \text{ and } O \in \mathbf{O}(m)\}.$$

(2) On the sphere

$$\mathrm{Iso}(\mathbf{S}^m(r), g_{\mathrm{can}}) = \mathbf{O}(m+1) = \mathrm{Iso}_{\mathbf{0}}(\mathbf{R}^{m+1}, g_{\mathrm{can}}).$$

Consequently, $\mathbf{O}(m+1)/\mathbf{O}(m) \cong \mathbf{S}^m$. ♠

If G is a Lie group, then the tangent space can be trivialized by

$$TG \cong G \times T_eG$$

by using the left (or right) translations on G. If H is a closed subgroup of G, then G/H is a manifold. If we endow G with a metric such that right translation by elements in H act by isometries, then there is unique Riemannian metric on G/H making the projection $G \to G/H$ into a Riemannian submersion. If in addition, the metric is also left invariant then G acts by isometries on G/H (on the left) thus making G/H into a homogeneous space.

For any tangent fields X and Y, we define

$$\langle X, Y\rangle_g := g(X, Y), \quad |X|_g := (\langle X, X\rangle_g)^2$$

and the angle of X and Y

$$\angle_g(X, Y) := \cos^{-1}\left(\frac{\langle X, Y\rangle_g}{|X|_g|Y|_g}\right).$$

Let $x^1, \cdots, x^m$ be a local coordinate system of $\mathcal{M}$. Then $\partial_1, \cdots, \partial_m$, where $\partial_i := \partial/\partial x^i$, form a local basis for $T\mathcal{M}$ and $dx^1, \cdots, dx^m$ form a dual basis for $T^*\mathcal{M}$. The metric may then be written in above local coordinates as

$$g = g_{ij} dx^i \otimes dx^j$$

where $g_{ij} := g(\partial/\partial x^i, \partial/\partial x^j)$. **Here and throughout the note, we follow the Einstein summation convention of summing over repeated indices. If the repeated indices both occur in upper line or lower line, we write down the summation symbol.**

Given a smooth immersion $\varphi : \mathcal{N} \to \mathcal{M}$ and a metric g on $\mathcal{M}$, the pull-back g to a metric on $\mathcal{N}$ is

$$(\varphi^* g)(V, W) = g(\varphi_* V, \varphi_* W),$$

where $\varphi_* : T\mathcal{N} \to T\mathcal{M}$ is the tangent map. If $(y^\alpha)_{\alpha=1}^n$ and $(x^i)_{i=1}^m$ are local coordinates on N and M, respectively, then

$$(\varphi^* g)_{\alpha\beta} = g_{ij} \frac{\partial \varphi^i}{\partial y^\alpha} \frac{\partial \varphi^j}{\partial y^\beta},$$

where $(\varphi^* g)_{\alpha\beta} := (\varphi^* g)(\partial/\partial y^\alpha, \partial/\partial y^\beta)$ and $\varphi^i := x^i \circ \varphi$. Given any covariant p-tensor field α on $\mathcal{M}$ and a smooth map $\varphi : \mathcal{N} \to \mathcal{M}$, we can define the pull-back of α to $\mathcal{N}$ by

$$(\varphi^* \alpha)(X_1, \cdots, X_p) := \alpha(\varphi_*(X_1), \cdots, \varphi_*(X_p)) \tag{2.2.3}$$

for all $X_1, \cdots, X_p \in C^\infty(T\mathcal{N})$.

Note 2.1

If φ is a diffeomorphism, then we do not distinguish between a metric g and its pull-back $\varphi^* g$. ♣

Example 2.5

(1) The canonical metric on $\mathbf{R}^m$ is

$$g_{\text{can}} = \delta_{ij} dx^i \otimes dx^j = \sum_{1 \le i \le m} dx^i \otimes dx^i.$$

(2) On $\mathbf{R}^2 \setminus \{$half line$\}$ we also have polar coordinates (r, θ). In these polar coordinates the canonical metric is

$$g_{\text{can}} = dr \otimes dr + r^2 d\theta \otimes d\theta.$$

Hence

$$g_{rr} = 1, \quad g_{\theta\theta} = r^2, \quad g_{r\theta} = g_{\theta r} = 0.$$

(3) A surface of revolution consist of a curve

$$\gamma(t) = (r(t), z(t)) : I \longrightarrow \mathbf{R}^2,$$

where $I \subset \mathbf{R}$ is open and $r(t) > 0$ for all t. By rotating this curve around the z-axis, we get a surface that can be represented as

$$(t, \theta) \longmapsto f(t, \theta) = (r(t)\cos\theta, r(t)\sin\theta, z(t)).$$

Then the metric is

$$g = \left(\dot{r}^2 + \dot{z}^2\right) dt \otimes dt + r^2 d\theta \otimes d\theta$$

so that

$$g_{tt} = \left(\dot{r}^2 + \dot{z}^2\right)^{1/2}, \quad g_{\theta\theta} = r, \quad g_{t\theta} = g_{\theta t} = 0.$$

(4) On $I \times \mathbf{S}^1$, we have rotationally symmetric metrics

$$g = \eta^2(t) dt \otimes dt + \varphi^2(t) d\theta \otimes d\theta.$$

Example (3) is a special case of (4).

(5) Let $\mathrm{sn}_k(t)$ denote the unique solution to

$$\ddot{x}(t) + k\dot{x}(t) = 0, \quad x(0) = 0, \quad \dot{x}(0) = 1.$$

Then we have a 1-parameter family

$$dt \otimes dt + \mathrm{sn}_k^2(t) d\theta \otimes d\theta \tag{2.2.4}$$

of rotationally symmetric metrics. When $k = 0$, this is $\mathbf{R}^2$; when $k > 0$, we get $\mathbf{S}^2(1/\sqrt{k})$; and when $k < 0$, we get the hyperbolic metrics

$$dt \otimes dt + r^2 \sinh^2\left(\frac{t}{r}\right) d\theta \otimes d\theta.$$

(6) Assume we have

$$dt \otimes dt + \varphi^2(t) d\theta \otimes d\theta$$

where $\varphi : [0, b) \to [0, \infty)$ is smooth, $\varphi(0) = 0$ and $\varphi(t) > 0$ for $t > 0$. Write

$$\varphi(t) = t\psi(t)$$

for some smooth function $\psi(t) > 0$ for $t > 0$. Introduce Cartesian coordinates

$$x = t\cos\theta, \quad y = t\sin\theta$$

near $t = 0$. Then

$$\begin{aligned} dt \otimes dt + \varphi^2(t) d\theta \otimes d\theta \quad &= \frac{x^2 + \psi^2(t) y^2}{x^2 + y^2} dx \otimes dx + \frac{xy - xy\psi^2(t)}{x^2 + y^2} dx \otimes dy \\ &\quad + \frac{xy - xy\psi^2(t)}{x^2 + y^2} dy \otimes dx + \frac{\psi^2(t) x^2 + y^2}{x^2 + y^2} dy \otimes dy \end{aligned}$$

so that

$$\begin{aligned} g_{xx} &= 1 + \frac{\psi^2(t) - 1}{t^2} \cdot y^2, \\ g_{xy} &= g_{yx} = \frac{1 - \psi^2(t)}{t^2} \cdot xy, \\ g_{yy} &= 1 + \frac{\psi^2(t) - 1}{t^2} \cdot x^2. \end{aligned}$$

To check for smoothness of the metric at $(x, y) = (0, 0)$ (or $t = 0$), it suffices to check that the function

$$\frac{\psi^2(t) - 1}{t^2}$$

is smooth at $t = 0$. First, it is clearly necessary that $\psi(0) = 1$; this is the vertical tangent condition. Second, if ψ is given by a power series we see that it must further satisfy: $\dot{\psi}(0) = \psi^{(3)} = \cdots = 0$. Those conditions are also sufficient when ψ is merely smooth. Translating back to φ, we get that the metric is smooth at $t = 0$ if and only if $\varphi^{(\text{even})}(0) = 0$ and $\dot{\varphi}(0) = 1$.

These conditions are all satisfied by the metrics $dt \otimes dt + \mathrm{sn}_k^2(t) d\theta \otimes d\theta$, where $t \in [0, \infty)$ when $k \leq 0$ and $t \in [0, \pi/\sqrt{k})$ for $k > 0$. Note that in this case $\mathrm{sn}_k(t)$ is real analytic. ♠

Example 2.6. (Doubly warped products)

(1) (Doubly warped products in general) *We can consider metrics on $I \times \mathbf{S}^{m-1}$ of the type*

$$dt \otimes dt + \varphi^2(t) ds_{m-1} \otimes ds_{m-1}$$

where $ds_{m-1} \otimes ds_{m-1}$ is the canonical metric on $\mathbf{S}^{m-1}(1) \subset \mathbf{R}^m$. Even more general are metrics of the type

$$dt \otimes dt + \varphi^2(t) ds_p \otimes ds_p + \psi^2(t) ds_q \otimes ds_q$$

on $I \times \mathbf{S}^p \times \mathbf{S}^q$.

(a) If $\varphi : (0, b) \to (0, \infty)$ is smooth and $\varphi(0) = 0$, then we get a smooth metric at $t = 0$ if and only if

$$\varphi^{(\text{even})}(0) = 0, \quad \dot{\varphi}(0) = 1,$$

and

$$\psi(0) > 0, \quad \psi^{(\text{odd})}(0) = 0.$$

The topology near $t = 0$ in this case is $\mathbf{R}^{p+1} \times \mathbf{S}^q$.

(b) If $\varphi : (0, b) \to (0, \infty)$ is smooth and $\varphi(b) = 0$, then we get a smooth metric at $t = b$ if and only if

$$\varphi^{(\text{even})}(b) = 0, \quad \dot{\varphi}(b) = -1,$$

and

$$\psi(b) > 0, \quad \psi^{(\text{odd})}(b) = 0.$$

The topology near $t = b$ in this case is again $\mathbf{R}^{p+1} \times \mathbf{S}^q$.

By adjusting and possibly changing the roles of these function we can get three different types of topologies:

- *$\varphi, \psi : [0, \infty) \to [0, \infty)$ are both positive on all of $(0, \infty)$. Then we have a smooth metric on $\mathbf{R}^{p+1} \times \mathbf{S}^q$ if φ, ψ satisfy (a).*
- *$\varphi, \psi : [0, b] \to [0, \infty)$ and both positive on $(0, b)$ and satisfy (a) and (b). Then we get a smooth metric on $\mathbf{S}^{p+1} \times \mathbf{S}^q$.*
- *$\varphi, \psi : [0, b] \to [0, \infty)$ as in the second type but the roles of ψ and φ are interchanged*

at $t = b$. Then we get a smooth metric on $\mathbf{S}^{p+q+1}$.

(2) (Spheres as warped products) *The sphere metric $dt \otimes dt + \sin^2(r) ds_{m-1} \otimes ds_{m-1}$ can be written as a rotationally symmetric metric. Consider the map*

$$F : (0, \pi) \times \mathbf{R}^n \longrightarrow \mathbf{R} \times \mathbf{R}^m, \quad (r, z) \longmapsto (t, x) = (\cos r, \sin r \cdot z),$$

which reduces to a map

$$G : (0, \pi) \times \mathbf{S}^{m-1} \longrightarrow \mathbf{R} \times \mathbf{R}^m, \quad (r, z) \longmapsto (\cos r, \sin r \cdot z).$$

Thus, G really maps into the unit sphere $\mathbf{S}^m$ in $\mathbf{R}^{m+1}$. Calculate

$$\begin{aligned} g_{\text{can}} &= dt \otimes dt + \sum_{i,j=1}^{m} \delta_{ij} dx^i \otimes dx^j \\ &= d\cos r \otimes d\cos r + \sum_{i,j=1}^{m} \delta_{ij} d(z^i \sin r) \otimes d(z^j \sin r) \\ &= dr \otimes dr + \sin^2 r \left(dz^1 \otimes dz^1 + \cdots + dz^m \otimes dz^m\right) \end{aligned}$$

which is the canonical metric $dt \otimes dt + \sin^2 r ds_{m-1} \otimes ds_{m-1}$.

The metric

$$dt \otimes dt + \sin^2 t\, ds_p \otimes ds_p + \cos^2 t\, ds_q \otimes ds_q, \quad t \in \left[0, \frac{\pi}{2}\right],$$

are also $\left(\mathbf{S}^{p+q+1}, g_{\text{can}}\right)$. Namely, we have $\mathbf{S}^p \subset \mathbf{R}^{p+1}$ and $\mathbf{S}^q \subset \mathbf{R}^{q+1}$, so we have map

$$\left(0, \frac{\pi}{2}\right) \times \mathbf{S}^p \times \mathbf{S}^q \longrightarrow \mathbf{R}^{p+1} \times \mathbf{R}^{q+1}, \quad (t, x, y) \longmapsto (x \cdot \sin t, y \cdot \cos t),$$

where $x \in \mathbf{R}^{p+1}$, $y \in \mathbf{R}^{q+1}$ have $|x| = |y| = 1$. The map is a Riemannian isometry.

(3) (The Hopf fibration) *On $\mathbf{S}^3(1)$, write the metric as*

$$dt \otimes dt + \sin^2 t\, d\theta_1 \otimes d\theta_1 + \cos^2 t\, d\theta_2 \otimes d\theta_2, \quad t \in \left[0, \frac{\pi}{2}\right],$$

and use complex coordinates

$$\left(t, e^{\sqrt{-1}\theta_1}, e^{\sqrt{-1}\theta_2}\right) \longmapsto \left(\sin t e^{\sqrt{-1}\theta_1}, \cos t e^{\sqrt{-1}\theta_2}\right)$$

to describe the isometric embedding

$$\left(0, \frac{\pi}{2}\right) \times \mathbf{S}^1 \times \mathbf{S}^1 \longrightarrow \mathbf{S}^3(1) \subset \mathbf{C}^2.$$

On $\mathbf{S}^2(1/2)$ use the metric

$$dr \otimes dr + \frac{\sin^2(2r)}{4} d\theta \otimes d\theta, \quad r \in \left[0, \frac{\pi}{2}\right],$$

with coordinates

$$\left(r, e^{\sqrt{-1}\theta}\right) \longrightarrow \left(\frac{1}{2}\cos(2r), \frac{1}{2}\sin(2r) e^{\sqrt{-1}\theta}\right).$$

The Hopf fibration in these coordinates looks like

$$\left(t, e^{\sqrt{-1}\theta_1}, e^{\sqrt{-1}\theta_2}\right) \longmapsto \left(t, e^{\sqrt{-1}(\theta_1 - \theta_2)}\right).$$

On $\mathbf{S}^3(1)$ *we have an orthogonal framing*

$$\left\{\partial_{\theta_1}+\partial_{\theta_2}, \partial_t, \frac{\cos^2 t\partial_{\theta_1}-\sin^2 t\partial_{\theta_1}}{\cos t\sin t}\right\},$$

where the first vector is tangent to the Hopf fiber and the two other vectors have unit length. On $\mathbf{S}^2(1/2)$

$$\left\{\partial_r, \frac{2}{\sin(2r)}\partial_\theta\right\}$$

is an orthonormal frame. The Hopf map clearly maps

$$\begin{aligned}\partial_t &\longrightarrow \partial_r,\\ \frac{\cos^2 t\partial_{\theta_1}-\sin^2 t\partial_{\theta_2}}{\cos t\sin t} &\longrightarrow \frac{\cos^2 r\partial_\theta+\sin^2 r\partial_\theta}{\cos r\sin r}=\frac{2}{\sin(2r)}\cdot\partial_\theta.\end{aligned}$$

Hence, it is an isometry on vectors perpendicular to the fiber. ♠

The **Levi-Civita connection** $\nabla_g : C^\infty(T\mathcal{M})\times C^\infty(T\mathcal{M})\to C^\infty(T\mathcal{M})$ is the unique connection on $T\mathcal{M}$ that is compatible with the metric g and is torsion-free:

$$X\left(g(Y,Z)\right) = g\left((\nabla_g)_X Y, Z\right)+g\left(Y,(\nabla_g)_X Z\right), \tag{2.2.5}$$

$$(\nabla_g)_X Y-(\nabla_g)_Y X = [X,Y], \tag{2.2.6}$$

where $(\nabla_g)_X Y := (\nabla_g Y)(X) := \nabla_g(X,Y)$. The Levi-Civita connection is uniquely determined by the equation

$$\begin{aligned}2g((\nabla_g)_X Y, Z) &= X\left(g(Y,Z)\right)+Y(g(X,Z))-Z(g(X,Y)) \\ &+ g([X,Y],Z)-g([X,Z],Y)-g([Y,Z],X).\end{aligned} \tag{2.2.7}$$

Note 2.2

For any $c>0$ *we have* $\nabla_{cg}=\nabla_g$. ♣

Let $x^1,\cdots,x^m$ be a local coordinate system. The **Christoffel symbols** of the Levi-Civita connection are defined as

$$(\nabla_g)_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} := (\Gamma_g)^k_{ij}\frac{\partial}{\partial x^k}.$$

Then

$$(\Gamma_g)^k_{ij} = \frac{1}{2}g^{kl}\left(\partial_i g_{\ell j}+\partial_j g_{i\ell}-\partial_\ell g_{ij}\right) := \frac{1}{2}g^{k\ell}(\Gamma_g)_{ij,\ell}. \tag{2.2.8}$$

Here and throughout ∂_i stands for $\partial/\partial x^i$. We call $\{\Gamma_g\}_{ij,k}$ the Christoffel symbols of the first kind, while $(\Gamma_g)^k_{ij}$ the Christoffel symbols of the second kind. Classically the following notation has also been used

$$\left\{\begin{array}{c} k \\ ij \end{array}\right\}_g = (\Gamma_g)^k_{ij}, \quad [ij,k]_g = (\Gamma_g)_{ij,k}. \tag{2.2.9}$$

Note 2.3

Let $x^1, \cdots, x^m$ and $y^1, \cdots, y^m$ are two local coordinate systems. Show that

$$(\Gamma_g)^\gamma_{\alpha\beta}\frac{\partial x^k}{\partial y^\gamma} = (\Gamma_g)^k_{ij}\frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta} + \frac{\partial^2 x^k}{\partial y^\alpha \partial y^\beta}.$$

♣

Note 2.4

If $(\mathcal{M}, g)$ is an m-dimensional Riemannian manifold, $\varphi : \mathcal{N} \to \mathcal{M}$ is an immersion, and $y^1, \cdots, y^n$ and $x^1, \cdots, x^m$ are local coordinates on $\mathcal{N}$ and $\mathcal{M}$, respectively, then

$$(\Gamma_{\varphi^* g})^\gamma_{\alpha\beta}\frac{\partial \varphi^k}{\partial y^\gamma} = \left((\Gamma_g)^k_{ij}\circ\varphi\right)\frac{\partial \varphi^i}{\partial y^\alpha}\frac{\partial \varphi^j}{\partial y^\beta} + \frac{\partial^2 \varphi^k}{\partial y^\alpha \partial y^\beta},$$

where $\varphi^i := x^i \circ \varphi$.

♣

A vector field X along a path $\gamma : [a, b] \to \mathcal{M}$ is **parallel** if

$$(\nabla_g)_{\dot\gamma} X = 0$$

along γ; the vector field $X(\gamma(t))$ is called the **parallel translation** of $X(\gamma(a))$. We say that a path γ is a **geodesic** if the unit tangent vector field is parallel along γ:

$$(\nabla_g)_{\dot\gamma}\left(\frac{\dot\gamma}{|\dot\gamma|_g}\right) = 0.$$

A geodesic has **constant speed** if $|\dot\gamma|_g$ is constant along γ; in this case $(\nabla_g)_{\dot\gamma}\dot\gamma = 0$.

Note 2.5

If X is a parallel along a path γ, then $|X|^2_g$ is constant along γ. Since ∇_g is the Levi-Civita Connection, it follows that

$$(\nabla_g)_{\dot\gamma}|X|^2_g = (\nabla_g)_{\dot\gamma}\left(\langle X, X\rangle_g\right) = 2\left\langle(\nabla_g)_{\dot\gamma}X, X\right\rangle_g = 0$$

because of $(\nabla_g)_{\dot\gamma}X = 0$.

♣

2.2.2 Curvatures

The **Riemannian curvature $(3,1)$-tensor field** Rm_g is defined by

$$\mathrm{Rm}_g(X, Y)Z := (\nabla_g)_X(\nabla_g)_Y Z - (\nabla_g)_Y(\nabla_g)_X Z - (\nabla_g)_{[X,Y]}Z. \tag{2.2.10}$$

For any function f one has

$$\mathrm{Rm}_g(fX, Y)Z = \mathrm{Rm}_g(X, fY)Z = \mathrm{Rm}_g(X, Y)(fZ) = f\mathrm{Rm}_g(X, Y)Z. \tag{2.2.11}$$

If we define

$$(\nabla_g)^2_{X,Y}Z := (\nabla_g)_X(\nabla_g)_Y Z - (\nabla_g)_{(\nabla_g)_X Y}Z$$

so that

$$\mathrm{Rm}_g(X, Y)Z = (\nabla_g)^2_{X,Y}Z - (\nabla_g)^2_{Y,X}Z, \tag{2.2.12}$$

$$(\nabla_g)^2_{fX,Y}Z = (\nabla_g)^2_{X,fY}Z = f(\nabla_g)^2_{X,Y}Z, \tag{2.2.13}$$

and

$$
\begin{aligned}
(\nabla_g)^2_{X,Y}(fZ) &= f(\nabla_g)^2_{X,Y}Z + Y(f)(\nabla_g)_X Z + X(f)(\nabla_g)_Y Z \\
&\quad - \left(\left((\nabla_g)_X Y\right) f\right) Z + X\left(Y(f)\right) Z
\end{aligned} \tag{2.2.14}
$$

for any function f.

Note 2.6

Note that Rm_g *is a tensor field but* $(\nabla_g)^2_{\bullet,\bullet}$ *is not. Also note that* $(\nabla_g)^2_{\bullet,\bullet} \neq (\nabla_g)_\bullet(\nabla_g)_\bullet$. ♣

The components of the $(3,1)$-tensor field Rm_g are defined by

$$\mathrm{Rm}_g(\partial_i, \partial_j)\partial_k := R^\ell_{ijk}\partial_\ell$$

and $R_{ijk\ell} := g_{\ell p} R^p_{ijk}$. The quantities

$$R_{ijk\ell} := \mathrm{Rm}_g(\partial_i, \partial_j, \partial_k, \partial_\ell) := \langle \mathrm{Rm}_g(\partial_i, \partial_j)\partial_k, \partial_\ell \rangle_g$$

are the components of Rm_g as a $(4,0)$-tensor field Rm_g. Some basic symmetrics of the Riemann curvature tensor are

$$R_{ijk\ell} = -R_{jik\ell} = -R_{ij\ell k} = R_{k\ell ij}. \tag{2.2.15}$$

The metric g introduces the inner product on $C^\infty(\mathcal{M}, \wedge^2\mathcal{M})$:

$$g(X \wedge Y, V \wedge W) := \det \begin{pmatrix} g(X,V), & g(X,W) \\ g(Y,V), & g(Y,W) \end{pmatrix}$$

and then extend it by linearity to all of $C^\infty(\mathcal{M}, \wedge^2\mathcal{M})$. From the symmetry properties of the curvature tensor field we see that Rm_g actually defines a symmetric bilinear map

$$\mathbf{Rm}^\wedge_g : C^\infty(\mathcal{M}, \wedge^2\mathcal{M}) \times C^\infty(\mathcal{M}, \wedge^2\mathcal{M}) \longrightarrow C^\infty(\mathcal{M}) \tag{2.2.16}$$

given by

$$\mathbf{Rm}^\wedge_g(X \wedge Y, W \wedge V) := \mathrm{Rm}_g(X, Y, V, W). \tag{2.2.17}$$

The relation

$$g\left(\mathbf{Rm}_g(X \wedge Y), V \wedge W\right) := \mathbf{Rm}^\wedge_g\left(X \wedge Y, V \wedge W\right) \tag{2.2.18}$$

defines a self-adjoint operator

$$\mathbf{Rm}_g : C^\infty(\mathcal{M}, \wedge^2\mathcal{M}) \longrightarrow C^\infty(\mathcal{M}, \wedge^2\mathcal{M}). \tag{2.2.19}$$

This operator is called the **Riemann curvature operator**.

If $\Pi \subset T_p\mathcal{M}$ is a 2-plane, then the **sectional curvature of** Π is defined by

$$\mathrm{Sec}_g(\Pi) := \langle \mathrm{Rm}_g(e_1, e_2)e_2, e_1 \rangle_g = \mathrm{Rm}_g(e_1, e_2, e_2, e_1) = \mathbf{Rm}^\wedge_g(e_1 \wedge e_2, e_1 \wedge e_2)$$

where $\{e_1, e_2\}$ is an orthonormal basis of Π.

Note 2.7

If v and w are two vectors spanning Π, then

$$\mathrm{Sec}_g(\Pi) := \frac{\langle \mathrm{Rm}_g(v,w)w, v\rangle_g}{|v|_g^2|w|_g^2 - \langle v, w\rangle_g^2}.$$

Since v and 2 span the 2-plane Π, we can write

$$v = a^1 e_1 + a^2 e_2, \quad w = b^1 e_1 + b^2 e_2.$$

By symmetric properties of Rm_g, it follows that

$$\langle \mathrm{Rm}_g(v,w)w, v\rangle_g = \left(a^1b^2 - a^2b^1\right)^2 \langle \mathrm{Rm}_g(e_1,e_2)e_2, e_1\rangle_g .$$

On the other hand, we have

$$\begin{aligned} & |v|^2_{g(p)}|w|_g^2 - \langle v, w\rangle_g^2 \\ = \ & \left(\left(a^1\right)^2 + \left(a^2\right)^2\right)\left(\left(b^1\right)^2 + \left(b^2\right)^2\right) - \left(a^1b^2 + a^2b^2\right)^2 \ = \ \left(a^1b^2 - a^2b^1\right)^2 . \end{aligned}$$

Hence we prove the identity.

Geometrically, the sectional curvature of a 2-plane $\Pi \subset T_p\mathcal{M}$ is equal to the Gauss curvature at p of the surface spanned by the geodesics emanating from p and tangent to Π (this surface is smooth in a neighborhood of p). ♣

For any $v \in T_p\mathcal{M}$ let

$$\mathrm{Rm}_g(v) : T_p\mathcal{M} \longrightarrow T_p\mathcal{M}, \quad w \longmapsto \mathrm{Rm}_g(w,v)v \tag{2.2.20}$$

be the **directional curvature operator**. This operator is also known as the **tidal force operator**. Then

$$\begin{aligned} g\left(\mathrm{Rm}_g(w)v, v\right) &= g\left(\mathrm{Rm}_g(v,w)w, v\right) = \mathrm{Rm}_g(v,w,w,v) \\ &= \mathbf{Rm}_g^{\wedge}(v\wedge w, v\wedge w) = g\left(\mathbf{Rm}_g(v\wedge w), v\wedge w\right). \end{aligned}$$

Proposition 2.1. (Riemann, 1854)

The following properties are equivalent:

(1) $\mathrm{Sec}_g(\Pi) = k$ *for all 2-planes Π in $T_p\mathcal{M}$.*

(2) $\mathrm{Rm}_g(v_1,v_2)v_3 = k(v_1\wedge v_2)(v_3)$ *for all $v_1, v_2, v_3 \in T_p\mathcal{M}$.*

(3) $\mathrm{Rm}_g(w)(v) = k\left(v - \langle v, w\rangle_{g(p)} w\right)$ *for all $v \in T_p\mathcal{M}$ and $|w|_g = 1$.*

(4) $\mathbf{Rm}_g(\omega) = k\cdot\omega$ *for all $\omega \in \wedge^2 T_p\mathcal{M}$.* ♡

Proof. (2) $\Rightarrow$ (3): Calculate

$$\begin{aligned} \mathrm{Rm}_g(w)(v) &= \mathrm{Rm}_g(v,w)w = k(v\wedge w)w = k\left(\langle w,w\rangle_g v - \langle v,w\rangle_g w\right) \\ &= k\left(v - \langle v, w\rangle_g w\right). \end{aligned}$$

(3) $\Rightarrow$ (1): Calculate

$$\mathrm{Sec}_g(\Pi) = \frac{\langle k\left(v - \langle v,w\rangle_g w\right), v\rangle_g}{|v|_g - (v,w)_g^2} = k.$$

$(1) \Rightarrow (2)$: We introduce the multilinear maps:

$$\begin{aligned}\mathfrak{T}(v_1, v_2)v_3 &:= k(v_1 \wedge v_2)(v_3),\\ \mathrm{T}(v_1, v_2, v_3, v_4) &:= \langle \mathfrak{T}(v_1, v_2)v_3, v_4\rangle_g = k\langle (v_1 \wedge v_2)v_3, v_4\rangle_g.\end{aligned}$$

The basic symmetries are

$$\begin{aligned}\mathfrak{T}(v_1, v_2)v_3 + \mathfrak{T}(v_2, v_3)v_1 + \mathfrak{T}(v_3, v_1)v_2 &= 0,\\ \mathfrak{T}(v_1, v_2)v_3 &= -\mathfrak{T}(v_2, v_1)v_3,\\ \mathrm{T}(v_1, v_2, v_3, v_4) = -\mathrm{T}(v_2, v_1, v_3, v_4) &= -\mathrm{T}(v_1, v_2, v_4, v_3)\\ &= \mathrm{T}(v_3, v_4, v_1, v_2),\\ \mathrm{T}(v_1, v_2, v_3, v_4) + \mathrm{T}(v_2, v_3, v_1, v_4) + \mathrm{T}(v_3, v_2, v_1, v_4) &= 0.\end{aligned}$$

Now consider the map

$$\mathrm{S}(v_1, v_2, v_3, v_4) := \mathrm{Rm}_g(v_1, v_2, v_3, v_4) - \mathrm{T}(v_1, v_2, v_3, v_4)$$

which also satisfies the same symmetry properties. The assumption that $\mathrm{Sec}_g(\Pi) = k$ implies

$$\mathrm{S}(v, w, w, v) = 0$$

for all $v, w \in T_p\mathcal{M}$. Using polarization $w = w_1 + w_2$ we get

$$\begin{aligned}0 &= \mathrm{S}(v, w_1 + w_2, w_1 + w_2, v) = \mathrm{S}(v, w_1, w_2, v) + \mathrm{S}(v, w_2, w_1, v)\\ &= 2\mathrm{S}(v, w_1, w_2, v) = -2\mathrm{S}(v, w_1, v, w_2).\end{aligned}$$

Using the symmetric properties, S is alternating in all four variables. Hence $\mathrm{S} = 0$, which is exactly what we wish to prove.

$(2) \Rightarrow (4)$: Choose an orthonormal basis e_i for $T_p\mathcal{M}$; then $e_i \wedge e_j$, $i < j$, is a basis for $\wedge^2 T_p\mathcal{M}$. Using (2) we have

$$\begin{aligned}\langle \mathbf{Rm}_g(e_i \wedge e_j), e_\ell \wedge e_k\rangle_{g(p)} &= \mathrm{Rm}_g(e_i, e_j, e_k, e_\ell)\\ = k\left\langle \langle e_j, e_k\rangle_{g(p)} e_i - \langle e_i, e_k\rangle_{g(p)} e_j, e_\ell\right\rangle_{g(p)} &= k\langle e_i \wedge e_j, e_\ell \wedge e_k\rangle_{g(p)}\end{aligned}$$

that implies

$$\mathbf{Rm}_g(e_i \wedge e_j) = k\,(e_i \wedge e_j)\,.$$

$(4) \Rightarrow (1)$: If v, w are orthogonal unit vectors, then $k = \langle \mathbf{Rm}_g(v \wedge w), w \wedge v\rangle_g = \mathrm{Sec}_g(v, w)$. $\square$

A Riemannian manifold $(\mathcal{M}, g)$ has **constant sectional curvature** if the sectional curvature of every 2-plane is the same. So for we only know that $(\mathbf{R}^m, g_{\mathrm{can}})$ has sectional curvature zero. Later, we shall prove that $dr \otimes dr + \mathrm{sn}_k^2(r) ds_{n-1} \otimes ds_{n-1}$ has constant sectional curvature k.

Note 2.8

Show that

$$R_{ijk}^{\ell} = \partial_i \Gamma_{jk}^{\ell} - \partial_j \Gamma_{ik}^{\ell} + \Gamma_{jk}^{p}\Gamma_{ip}^{\ell} - \Gamma_{ik}^{p}\Gamma_{jp}^{\ell}. \tag{2.2.21}$$

♣

The **Ricci tensor field** Ric_g is the trace of the Riemann curvature tensor field:

$$\mathrm{Rc}_g(Y,Z) := \mathrm{tr}_g\left(X \longmapsto \mathrm{Rm}_g(X,Y)Z\right).$$

In terms of an orthonormal frame $e_1, \cdots, e_m$, we have

$$\mathrm{Rc}_g(Y,Z) = \sum_{1\le i\le m} \langle \mathrm{Rm}_g(e_i,Y)Z, e_i\rangle_g .$$

Its components, defined by

$$R_{ij} := \mathrm{Ric}_g(\partial_i, \partial_j)$$

are given by

$$R_{jk} = \sum_{1\le i\le m} R^i_{ijk}.$$

The **Ricci curvature** of a line $L \subset T_x\mathcal{M}$ is defined by

$$\mathrm{Rc}_g(L) := \mathrm{Ric}_g(e_1, e_1),$$

where $e_1 \in T_xM$ is a unit vector spanning L. The **scalar curvature** is the trace of the Ricci tensor field:

$$R_g := \sum_{1\le i\le m} \mathrm{Rc}_g(e_i, e_i).$$

Equivalently,

$$R_g = g^{ij}R_{ij}.$$

Here $(g^{ij}) := (g_{ij})^{-1}$ is the inverse matrix.

We say a metric has **constant Ricci curvature** if the Ricci curvature of every line is the same. We say $(\mathcal{M}, g)$ is an **Einstein manifold** with **Einstein constant k**, if

$$\mathrm{Rc}_g = kg.$$

If $(\mathcal{M}, g)$ has constant sectional curvature k, then $(\mathcal{M}, g)$ is also Einstein with Einstein constant $(m-1)k$. The converse may not be true; three basic types are

(1) $(\mathbf{S}^m \times \mathbf{S}^m, ds_m \otimes ds_m + ds_m \otimes ds_m)$ with Einstein constant $m-1$.

(2) The Fubini-Study metric on $\mathbf{CP}^m$ with Einstein constant $2m+2$.

(3) The Schwarzschild metric on $\mathbf{R}^2 \times \mathbf{S}^2$, which is a doubly warped product metric: $dr \otimes dr + \varphi^2(r)d\theta \otimes d\theta + \psi^2(r)ds_2 \otimes ds_2$ with Einstein constant 0.

Note 2.9

Given a metric g and a positive constant C, show that

$$\mathrm{Rm}^{(3,1)}_{Cg} = \mathrm{Rm}^{(3,1)}_g, \quad \mathrm{Rm}^{(4,0)}_{Cg} = C\mathrm{Rm}^{(4,0)}_g, \quad \mathrm{Rc}_{Cg} = \mathrm{Rc}_g, \quad R_{Cg} = C^{-1}R_g.$$

♣

Note 2.10. (Geometric interpretation of tracing)

The trace of a symmetric 2-tensor field α is given by the following formula:

$$\mathrm{tr}_g(\alpha) = \frac{1}{\omega_m}\int_{\mathbf{S}^{m-1}} \alpha(v,v)d\sigma(v),$$

where $\mathbf{S}^{m-1}$ *is the unit* $(m-1)$*-sphere,* $m\omega_m$ *its volume, and* $d\sigma$ *its volume form. For any unit vector* u*,* $\frac{1}{m-1}\mathrm{Rc}_g(u,u)$ *is the average of the sectional curvatures of planes containing the vector* u*. Similarly,* $\frac{1}{m}R_g(p)$ *is the average of* $\mathrm{Rc}_g(u,u)$ *over all unit vectors* $u \in \mathbf{S}^{m-1} \subset T_p\mathcal{M}$*.*

Choose an orthonormal basis $e_1, \cdots, e_m$ *such that* $\alpha = \sum_{1\le i\le m} \lambda_i e_i^* \otimes e_i^*$*. Then* $\mathrm{tr}_g(\alpha) = \sum_{1\le i\le m} \lambda_i$ *and*

$$\frac{1}{\omega_m}\int_{\mathbf{S}^{m-1}} \langle v, e_i\rangle_g^2 d\sigma(v) = 1.$$

Therefore

$$\begin{aligned}\frac{1}{\omega_m}\int_{\mathbf{S}^{m-1}} \alpha(v,v)d\sigma(v) &= \frac{1}{\omega_m}\int_{\mathbf{S}^{m-1}} \sum_{1\le i\le m} \lambda_i e_i^* \otimes e_i^*\,(v,v)\,d\sigma(v) \\ &= \sum_{1\le i\le m} \lambda_i \left(\frac{1}{\omega_m}\int_{\mathbf{S}^{m-1}} e_i^* \otimes e_i^*(v,v)d\sigma(v)\right) \\ &= \sum_{1\le i\le m} \lambda_i \left(\frac{1}{\omega_m}\int_{\mathbf{S}^{m-1}} \langle v, e_i\rangle_g^2 d\sigma(v)\right) = \sum_{1\le i\le m}\lambda_i = \mathrm{tr}_g(\alpha).\end{aligned}$$

Using this formula we can prove the rest facts. ♣

2.2.3 Covariant differentiation

Acting on $(0,s)$-tensor fields, we define covariant differentiation by

$$\left(\nabla_g^{(0,s)}\right)_X : C^\infty(\mathcal{M}, \otimes^s T\mathcal{M}) \longrightarrow C^\infty(\mathcal{M}, \otimes^s T\mathcal{M})$$

where

$$\left(\nabla_g^{(0,s)}\right)_X (Z_1 \otimes \cdots \otimes Z_s) := \sum_{1\le i\le s} Z_1 \otimes \cdots \otimes (\nabla_g)_X Z_i \otimes \cdots \otimes Z_s.$$

The covariant derivative of an (r,s)-tensor field α is defined by

$$\begin{aligned}\left(\left(\nabla_g^{(r,s)}\right)_X \alpha\right)(Y_1, \cdots, Y_r) &:= (\nabla_g)_X^{(0,s)}(\alpha(Y_1,\cdots,Y_r)) \\ &\quad - \sum_{1\le i\le r} \alpha(Y_1, \cdots, (\nabla_g)_X Y_i, \cdots, Y_r), \end{aligned} \tag{2.2.22}$$

Let $\otimes^{r,s}\mathcal{M} = \otimes^r T^*\mathcal{M} \otimes \otimes^s T\mathcal{M}$. The covariant derivative may be considered as

$$\nabla_g^{(r,s)} : C^\infty(\mathcal{M}, \otimes^{r,s}\mathcal{M}) \longrightarrow C^\infty(\mathcal{M}, \otimes^{r+1,s}\mathcal{M}),$$

where

$$\left(\left(\nabla_g^{(r,s)}\right)\alpha\right)(X, Z_1, \cdots, Z_r) := \left(\left(\nabla_g^{(r,s)}\right)_X \alpha\right)(Z_1, \cdots, Z_r),$$

or equivalently,

$$\nabla_g^{(r,s)}\alpha = \sum_{1\le i\le m} dx^i \otimes \left(\nabla_g^{(r,s)}\right)_i \alpha.$$

As an application, we prove

$$\nabla_g^{(2,0)} g = 0. \tag{2.2.23}$$

Indeed, by definition,

$$\begin{aligned}\left(\nabla_g^{(2,0)} g\right)(X, Z_1, Z_2) &= \left(\left(\nabla_g^{(2,0)}\right)_X g\right)(Z_1, Z_2) \\ &= \left(\nabla_g^{(0,0)}\right)_X (g(Z_1, Z_2)) - g\left((\nabla_g)_X Z_1, Z_2\right) - g\left(Z_1, (\nabla_g)_X Z_2\right) \\ &= X\left(g(Z_1, Z_2)\right) - X\left(g(Z_1, Z_2)\right) = 0.\end{aligned}$$

In general, we say a (r,s)-tensor field α is **parallel** if $\nabla_g^{(r,s)}\alpha = 0$. Thus, g is parallel.

We consider the composition of two covariant derivatives

$$\nabla_g^{(r+1,s)} \circ \nabla_g^{(r,s)} : C^\infty(\mathcal{M}, \otimes^{r,s}\mathcal{M}) \longrightarrow C^\infty(\mathcal{M}, \otimes^{r+2,s}\mathcal{M})$$

is given by

$$\begin{aligned}&\left(\left(\nabla_g^{(r+1,s)} \circ \nabla_g^{(r,s)}\right)\alpha\right)(X, Y, Z_1, \cdots, Z_r) = (\nabla_g^{(r+1,s)})_X \left(\nabla_g^{(r,s)}\alpha\right)(Y, Z_1, \cdots, Z_r) \\ &= (\nabla_g^{(0,s)})_X \left(\left(\nabla_g^{(r,s)}\alpha\right)(Y, Z_1, \cdots, Z_r)\right) - \left(\nabla_g^{(r,s)}\alpha\right)((\nabla_g)_X Y, Z_1, \cdots, Z_r) \\ &\qquad - \sum_{1\le i\le r}\left(\nabla_g^{(r,s)}\alpha\right)(Y, Z_1, \cdots, (\nabla_g)_X Z_i, \cdots, Z_r) \\ &= (\nabla_g^{(0,s)})_X \left(\left((\nabla_g^{(r,s)})_Y\alpha\right)(Z_1, \cdots, Z_r)\right) - \left((\nabla_g^{(r,s)})_{(\nabla_g)_X Y}\alpha\right)(Z_1, \cdots, Z_r) \\ &\qquad - \sum_{i=1}^{r}\left((\nabla_g^{(r,s)})_Y\alpha\right)(Z_1, \cdots, (\nabla_g)_X Z_i, \cdots, Z_r) \\ &= \left((\nabla_g^{(r,s)})_X \left((\nabla_g^{(r,s)})_Y\alpha\right)\right)(Z_1, \cdots, Z_r) - \left((\nabla_g^{(r,s)})_{(\nabla_g)_X Y}\alpha\right)(Z_1, \cdots, Z_r).\end{aligned}$$

If we write

$$\left(\left((\nabla_g^{(r+1,s)})_X \circ (\nabla_g^{(r,s)})_Y\right)\alpha\right)(Z_1, \cdots, Z_r) = \left(\left(\nabla_g^{(r+1,s)} \circ \nabla_g^{(r,s)}\right)\alpha\right)(X, Y, Z_1, \cdots, Z_r),$$

then

$$\left((\nabla_g^{(r+1,s)})_X \circ (\nabla_g^{(r,s)})_Y\right)\alpha = (\nabla_g^{(r,s)})_X \left((\nabla_g^{(r,s)})_Y\alpha\right) - (\nabla_g^{(r,s)})_{(\nabla_g)_X Y}\alpha.$$

Note 2.11

Throughout this note, we write ∇_g instead of $\nabla_g^{(r,s)}$. ♣

Hence

$$(\nabla_g^2)_{X,Y}\alpha = (\nabla_g)_X(\nabla_g)_Y\alpha - (\nabla_g)_{(\nabla_g)_X Y}\alpha.$$

If β is an (r,s)-tensor field, then we define the components $\nabla_i\beta_{j_1\cdots j_r}^{k_1\cdots k_s}$ of the covariant derivative $\nabla_g\beta$ by

$$\nabla_i\beta_{j_1\cdots j_r}^{k_1\cdots k_s}\partial_{k_1}\otimes\cdots\otimes\partial_{k_s} = \left((\nabla_g)_{\partial_i}\beta\right)(\partial_{j_1}, \cdots, \partial_{j_r}).$$

We then have

$$\begin{aligned}\nabla_i\beta_{j_1\cdots j_r}^{k_1\cdots k_s} &= \partial_i\beta_{j_1\cdots j_r}^{k_1\cdots k_s} - \sum_{1\le p\le r}\sum_{1\le q\le m}\Gamma_{ij_p}^{q}\beta_{j_1\cdots j_{p-1}\ell j_{p+1}\cdots j_r}^{k_1\cdots k_s} \\ &\quad + \sum_{1\le p\le s}\sum_{1\le q\le m}\Gamma_{iq}^{k_p}\beta_{j_1\cdots j_r}^{k_1\cdots k_{p-1}qk_{p+1}\cdots k_s}.\end{aligned} \tag{2.2.24}$$

For any 2-form β we have

$$\nabla_i \beta_{jk} = \partial_i \beta_{jk} - \Gamma^{\ell}_{ij}\beta_{\ell k} - \Gamma^{\ell}_{ik}\beta_{j\ell}.$$

If β is a 4-form, then

$$\nabla_i \beta_{jk\ell p} = \partial_i \beta_{jk\ell p} - \Gamma^{p}_{ij}\beta_{pk\ell p} - \Gamma^{p}_{ik}\beta_{jp\ell p} - \Gamma^{p}_{i\ell}\beta_{jkpp} - \Gamma^{p}_{ip}\beta_{jk\ell p}.$$

Note 2.12

We

$$\nabla_i \nabla_j f := (\nabla_g \nabla_g f)(\partial_i, \partial_j) = \partial_i \partial_j f - \Gamma^k_{ij}\partial_k f$$

and more generally for a 1*-form* $\alpha = \alpha_i dx^i$*,*

$$\nabla_i \alpha_j = \partial_i \alpha_j - \Gamma^k_{ij}\alpha_k.$$

Since $\nabla_g f = df = \partial_i f dx^i$*, it follows that* $\nabla_g f$ *can be viewed as a* 1*-form so that the two formulas coincide with each other in this case.* ♣

Let $\alpha_{i_1 \cdots i_r}$ denote the components of an $(r, 0)$-tensor field α:

$$\alpha_{i_1 \cdots i_r} := \alpha\left(\partial_{i_1}, \cdots, \partial_{i_r}\right).$$

We denote the components of $\nabla^k_g \alpha$ by $\nabla_{j_1} \cdots \nabla_{j_k} \alpha_{i_1 \cdots i_r}$, that is,

$$\nabla_{j_1} \cdots \nabla_{j_r} \alpha_{i_1 \cdots i_r} := \left(\nabla^k_g \alpha\right)\left(\partial_{j_1}, \cdots, \partial_{j_m}, \partial_{i_1}, \cdots, \partial_{i_r}\right).$$

For example, if α is a 1-form, then

$$\begin{aligned} \nabla_i \nabla_j \alpha_k &= \nabla_i \left(\partial_j \alpha_k - \Gamma^{\ell}_{jk}\alpha_\ell\right) \\ &= \partial_i \partial_j \alpha_k - \Gamma^{\ell}_{ij}\partial_\ell \alpha_k - \Gamma^{\ell}_{ik}\partial_j \alpha_\ell - \nabla_i \Gamma^{\ell}_{jk} \cdot \alpha_\ell - \Gamma^{\ell}_{jk}\left(\partial_i \alpha_\ell - \Gamma^p_{i\ell}\alpha_p\right); \end{aligned}$$

note that

$$\nabla_i \Gamma^{\ell}_{jk} = \partial_i \Gamma^{\ell}_{jk} - \Gamma^p_{ij}\Gamma^{\ell}_{pk} - \Gamma^p_{ik}\Gamma^{\ell}_{jp} + \Gamma^{\ell}_{ip}\Gamma^p_{jk}.$$

Therefore

$$\begin{aligned} \nabla_i \nabla_j \alpha_k &= \partial_i \partial_j \alpha_k - \left(\Gamma^{\ell}_{ij}\partial_\ell \alpha_k + \Gamma^{\ell}_{ik}\partial_j \alpha_\ell + \Gamma^{\ell}_{jk}\partial_i \alpha_\ell\right) \\ &\quad - \left(\partial_i \Gamma^{\ell}_{jk} - \Gamma^p_{ij}\Gamma^{\ell}_{pk} - \Gamma^p_{ik}\Gamma^{\ell}_{jp}\right)\alpha_\ell. \end{aligned}$$

Similarly, we can define the multiple covariant derivatives of an (r, s)-tensor field.

2.2.4 Holonomy

Given a path $\gamma : [a, b] \to \mathcal{M}$ from p to q, parallel translation along γ defines an isometry

$$\iota_\gamma : (T_p\mathcal{M}, g_p) \longrightarrow (T_q\mathcal{M}, g_q).$$

Given a point $p \in \mathcal{M}$, the set of isometries induced by parallel translation along contractible loop based at p is a group, called the **restricted holonomy group** $\mathrm{Hol}^0_p(\mathcal{M}, g)$.

Note 2.13

Show that

$$\mathrm{Hol}^0_{(p_1,p_2)}(\mathcal{M}_1^{m_1}\times\mathcal{M}_2^{m_2}, g_1+g_2)=\mathrm{Hol}^0_{p_1}(\mathcal{M}_1^{m_1}, g_1)+\mathrm{Hol}^0_{p_2}(\mathcal{M}_2^{m_2}, g_2).$$

♣

If $E_p\subset T_p\mathcal{M}$ is a subspace invariant under parallel translation, then $E_p^{\perp}$, the orthogonal complement in $T_p\mathcal{M}$, is also invariant. Consequently, the action of $\mathrm{Hol}^0_p(\mathcal{M},g)$ on $T\mathcal{M}$ induces a bundle decomposition

$$T\mathcal{M}=\mathcal{E}_1\oplus\cdots\oplus\mathcal{E}_k, \tag{2.2.25}$$

where $\mathcal{E}_i$ are subbundles invariant under parallel translation such that for each $p\in\mathcal{M}$,

$$T_p\mathcal{M}=(\mathcal{E}_1)_p\oplus\cdots\oplus(\mathcal{E}_k)_p$$

is the decomposition of $T_p\mathcal{M}$ into its irreducible invariant subspaces with respect to the action $\mathrm{Hol}^0_p(\mathcal{M},g)$. We call the splitting (2.2.25) the **irreducible holonomy decomposition** of $T\mathcal{M}$.

Theorem 2.1. (De Rham holonomy splitting theorem)

Let $(\mathcal{M},g)$ be a complete, simply-connected, m-dimensional Riemannian manifold. If $T\mathcal{M}=\mathcal{E}_1\oplus\cdots\oplus\mathcal{E}_k$ is the irreducible holonomoy decomposition of $T\mathcal{M}$, then $(\mathcal{M},g)$ splits as a Riemannian product, where $\mathcal{E}_i=T\mathcal{N}_i$,

$$(\mathcal{M},g)=(\mathcal{N}_1\times\cdots\mathcal{N}_k, g_1+\cdots+g_k)\,.$$

♡

2.3 Basic formulas and identities

Introduction

- *Bianchi identities*
- *Lie derivatives*
- *Commuting covariant derivatives*
- *The fundamental curvature equations*

2.3.1 Bianchi identities

The **first** and **second Bianchi identities** are

$$R_{ijk\ell}+R_{jki\ell}+R_{kij\ell}\ =\ 0, \tag{2.3.1}$$

$$\nabla_i R_{jk\ell p}+\nabla_j R_{ki\ell p}+\nabla_k R_{ij\ell p}\ =\ 0. \tag{2.3.2}$$

For any vector fields X, Y, and Z, we have (where we set $\nabla=\nabla_g$)

$$\begin{aligned}&\mathrm{Rm}_g(X,Y)Z+\mathrm{Rm}_g(Y,Z)X+\mathrm{Rm}_g(Z,X)Y\\ =\ &\nabla_X\nabla_Y Z-\nabla_Y\nabla_X Z-\nabla_{[X,Y]}Z+\nabla_Y\nabla_Z X-\nabla_Z\nabla_Y X-\nabla_{[Y,Z]}X\\ &+\nabla_Z\nabla_X Y-\nabla_X\nabla_Z Y-\nabla_{[Z,X]}Y\\ =\ &\nabla_X(\nabla_Y Z-\nabla_Z Y)+\nabla_Y(\nabla_Z X-\nabla_X Z)+\nabla_Z(\nabla_X Y-\nabla_Y X)\end{aligned}$$

$$
\begin{aligned}
&- \left(\nabla_{[X,Y]}Z + \nabla_{[Y,Z]}X + \nabla_{[Z,X]}Y\right) \\
= \ & \nabla_X[Y,Z] + \nabla_Y[Z,X] + \nabla_Z[X,Y] - \left(\nabla_{[X,Y]}Z + \nabla_{[Y,Z]}X + \nabla_{[Z,X]}Y\right) \\
= \ & [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] \ = \ 0.
\end{aligned}
$$

Setting $X = \partial_i$, $Y = \partial_j$, and $Z = \partial_k$, yields (2.3.1). For the second Bianchi identity, one can write down the expression of R_{ijkl} in terms of the metric components, then calculus the right side of (2.3.2).

The **twice contracted second Bianchi identity** is

$$2g^{ij}\nabla_i R_{jk} = \nabla_k R_g. \tag{2.3.3}$$

In fact, multiplying the second Bianchi identity (2.3.2) by $g^{ip}g^{j\ell}$ implies

$$0 = -g^{ip}\nabla_i R_{kp} - g^{j\ell}\nabla_j R_{k\ell} + \nabla_k R_g.$$

By rearranging the terms, we obtain (2.3.3). Using the convenient notation $\nabla^j := g^{ij}\nabla_i$, we can rewrite (2.3.3) as

$$\nabla^j R_{jk} = \frac{1}{2}\nabla_k R_g.$$

If we introduce the **Einstein tensor** $\mathrm{En}_g = \mathrm{Rc}_g - \frac{1}{2}R_g g$, then

$$\mathrm{div}_g\left(\mathrm{En}_g\right) = 0.$$

This is because

$$\left(\mathrm{div}_g(\mathrm{En}_g)\right)_k = g^{ij}\nabla_j\left(R_{jk} - \frac{R_g}{2}g_{jk}\right) = g^{ij}\nabla_i R_{jk} - \frac{1}{2}\nabla_k R_g = 0.$$

The **(once contracted second Bianchi identity)** is

$$-\left(\mathrm{div}_g(\mathrm{Rm}_g)\right)_{\ell jk} = \nabla^p R_{jk\ell p} = g^{im}\nabla_i R_{jk\ell p} = \nabla_j R_{k\ell} - \nabla_k R_{j\ell}. \tag{2.3.4}$$

Multiplying (2.3.2) by g^{ip}, we get $0 = \nabla^p R_{jk\ell p} - \nabla_j R_{k\ell} + \nabla_k R_{j\ell}$ that implies (2.3.4).

Theorem 2.2. (Schur, 1886)

(1) If g is an Einstein metric, i.e., $R_{ij} = \frac{1}{m}R_g g_{ij}$, and $m \geq 3$, then R is a constant. Note that any Riemannian metric on a surface is always Einstein.

(2) If $m \geq 3$ and the sectional curvatures at each point are independent of the 2-plane, that is, if

$$R_{ijk\ell} = \frac{R_g}{m(m-1)}(g_{i\ell}g_{jk} - g_{ik}g_{j\ell}),$$

then R_g is a constant. ♡

Proof. (1) Using (2.3.3) we obtain

$$\frac{1}{2}\nabla_k R_g = \nabla^j R_{jk} = \frac{1}{m}\nabla_k R_g.$$

If $m \geq 3$, it follows that $\nabla_k R_g = 0$ for any k and hence R_g is a constant.

(2) We use (1) by computing

$$R_{jk} = g^{i\ell} R_{ijk\ell} = \frac{R_g}{m(m-1)}(mg_{jk} - g_{jk}) = \frac{R_g}{m} g_{jk}.$$

Therefore R_g is a constant. □

2.3.2 Lie derivatives

A vector field X is **complete** if there is a 1-parameter group of diffeomorphisms $(\varphi_t)_{t\in\mathbf{R}}$ generated by X. If $\mathcal{M}$ is closed, then any smooth vector field is complete. Let α be an (r,s)-tensor field and let X be a complete vector field generating a global 1-parameter group of diffeomorphisms φ_t. The **Lie derivative** of α with respect to X is defined by

$$\mathscr{L}_X \alpha := \lim_{t\to 0} \frac{\alpha - (\varphi_t)_* \alpha}{t}. \tag{2.3.5}$$

Here $(\varphi_t)_{*,p} : T_p\mathcal{M} \to T_{\varphi(p)}\mathcal{M}$ is the differential of φ_t at $p \in \mathcal{M}$. It acts on the cotangent bundle by $(\varphi_t)_{*,p} := (\varphi_t^{-1})^*_{\varphi_t(p)} : T_p^*\mathcal{M} \to T^*_{\varphi_t(p)}\mathcal{M}$. We can then extend the action of $(\varphi_t)_*$ to the tensor bundles of M. The definition (2.3.5) extends to the case where X is not complete and only defines local 1-parameter groups of diffeomorphisms. Some basic properties of the Lie derivative are

(1) If f is a function, then $\mathscr{L}_X f = Xf$.
(2) If Y is a vector field, then $\mathscr{L}_X Y = [X,Y]$.
(3) If α and β are tensor fields, then $\mathscr{L}_X(\alpha\otimes\beta) = (\mathscr{L}_X\alpha)\otimes\beta + \alpha\otimes(\mathscr{L}_X\beta)$.
(4) If α is an $(r,0)$-tensor fields, then for any vector fields $X, Y_1, \cdots, Y_r$,

$$\begin{aligned}
(\mathscr{L}_X\alpha)(Y_1,\cdots,Y_r) &= X(\alpha(Y_1,\cdots,Y_r)) \\
&\quad - \sum_{1\le i\le m} \alpha(Y_1,\cdots,Y_{i-1},[X,Y_i],Y_{i+1},\cdots,Y_r) \\
&= (\nabla_X\alpha)(Y_1,\cdots,Y_r) \\
&\quad + \sum_{1\le i\le m} \alpha(Y_1,\cdots,Y_{i-1},\nabla_{Y_i}X,Y_{i+1},\cdots,Y_r).
\end{aligned} \tag{2.3.6}$$

If α is a 2-form, then

$$\begin{aligned}
(\mathscr{L}_X\alpha)(\partial_i,\partial_j) &= X(\alpha(\partial_i,\partial_j)) - \alpha([X,\partial_i],\partial_j) - \alpha(\partial_i,[X,\partial_j]) \\
&= X\alpha_{ij} - \alpha\left(-\partial_i X^\ell\cdot\partial_\ell,\partial_j\right) - \alpha\left(\partial_i, -\partial_j X^\ell\cdot\partial_\ell\right) \\
&= X\alpha_{ij} + \partial_i X^\ell\cdot\alpha_{\ell j} + \partial_j X^\ell\cdot\alpha_{i\ell} = \nabla_X\alpha_{ij} + \nabla_i X^\ell\cdot\alpha_{lj} + \nabla_j X^\ell\cdot\alpha_{i\ell}
\end{aligned}$$

where we use the formula $\partial_\ell\alpha_{ij} = \nabla_\ell\alpha_{ij} + \Gamma^p_{\ell i}\alpha_{pj} + \Gamma^p_{\ell j}\alpha_{ip}$.

Note 2.14

Given a diffeomorphism $\varphi : \mathcal{M}\to\mathcal{M}$, we have $\varphi_p^* : T^*_{\varphi(p)}\mathcal{M} \to T_p^*\mathcal{M}$. The pull-back acts on the tangent bundle by $\varphi_p^* := (\varphi^{-1})_{*,\varphi(p)} : T_{\varphi(p)}\mathcal{M}\to T_p\mathcal{M}$. These actions extend to the tensor bundles of M. Show that definition (2.3.5) is equivalent to

$$\mathscr{L}_X\alpha = \lim_{t\to 0}\frac{\varphi_t^*\alpha - \alpha}{t} = \frac{d}{dt}\Big|_{t=0}\varphi_t^*\alpha.$$

♣

The **gradient** of a function f with respect to the metric g is defined by

$$g\left(\mathrm{grad}_g f, X\right) := Xf = df(X).$$

We shall also use the notation $\nabla_g f$ to denote $\mathrm{grad}_g f$. In local coordinates,

$$df = \partial_i f dx^i, \quad \mathrm{grad}_g f = g^{ij}\partial_i f \partial_j.$$

Note 2.15. (Lie derivative of the metric)

The Lie derivative of the metric is given by

$$(\mathscr{L}_X g)(Y_1, Y_2) = g\left(\nabla_{Y_1} X, Y_2\right) + g\left(Y_1, \nabla_{Y_2} X\right) \tag{2.3.7}$$

and that in local coordinates this implies

$$(\mathscr{L}_X g)_{ij} = \nabla_i X_j + \nabla_j X_i, \quad X_i := g_{i\ell} X^\ell.$$

In particular, if f is a function, then

$$\left(\mathscr{L}_{\mathrm{grad}_g f}\, g\right)_{ij} = 2\nabla_i \nabla_j f. \tag{2.3.8}$$

Actually, (2.3.7) follows from $\nabla_g g = 0$. Let $X = \mathrm{grad}_g f$, by using (2.3.8) we have $X_i = \partial_i f$ so that we prove (2.3.8). ♣

Note 2.16

For any 1-form $\alpha = \alpha_i dx^i$, we write

$$\alpha^\sharp = g^{ij}\alpha_i \partial_j.$$

Then $(df)^\sharp = \mathrm{grad}_g f$. Also note that (2.3.8) has a short expression

$$\mathscr{L}_{(df)^\sharp} g = 2\nabla_g^2 f = 2\mathrm{Hess}_g(f), \quad \mathrm{Hess}_g(f) = \mathscr{L}_{\frac{1}{2}(df)^\sharp} g.$$

Correspondingly, to any vector field $X = X^i\partial_i$ we associate a 1-form $X^\flat$ defined by

$$X^\flat = g_{ij} X^i dx^j.$$

In terms of this notation, (2.3.7) equals $(\mathscr{L}_X)_{ij} = \nabla_i (X^\flat)_j + \nabla_j (X^\flat)_i$. ♣

Lemma 2.1

For any diffeomorphism $\varphi : \mathcal{M} \to \mathcal{M}$, tensor field α, and vector field X,

$$\varphi^*(\mathscr{L}_X \alpha) = \mathscr{L}_{\varphi^* X}(\varphi^*\alpha), \tag{2.3.9}$$

and if $f : \mathcal{M} \to \mathbf{R}$, then

$$\varphi^*(\mathrm{grad}_g f) = \mathrm{grad}_{\varphi^* g}(\varphi^* f) \tag{2.3.10}$$

where $\varphi^* f := f \circ \varphi$. ♡

Proof. Let ψ_t be a 1-parameter group of diffeomorphisms generated by X. Calculate

$$\begin{aligned}\varphi^*(\mathscr{L}_X\alpha) &= \varphi^*\left(\lim_{t\to 0}\frac{\psi_t^*\alpha - \alpha}{t}\right) = \lim_{t\to 0}\frac{\varphi^*(\psi_t^*\alpha) - \varphi^*\alpha}{t} \\ &= \lim_{t\to 0}\frac{(\psi_t\circ\varphi)^*\alpha - \varphi^*\alpha}{t} = \lim_{t\to 0}\frac{(\varphi^{-1}\circ\psi_t\circ\varphi)^*\varphi^*\alpha - \varphi^*\alpha}{t} = \mathscr{L}_Y(\varphi^*\alpha)\end{aligned}$$

where Y is the vector field generating the 1-parameter group of diffeomorphisms $\varphi^{-1} \circ \psi_t \circ \varphi$. For any point $p \in \mathcal{M}$, we have

$$\begin{aligned} Y(p) &= \frac{d}{dt}\Big|_{t=0} (\varphi^{-1} \circ \psi_t \circ \varphi)(p) = (\varphi^{-1})_* \frac{d}{dt}\Big|_{t=0} \psi_t \circ \varphi(p) \\ &= (\varphi^{-1})_*(X(\varphi(p))) = (\varphi^* X)(p). \end{aligned}$$

For any $p \in \mathcal{M}$ and $X \in T_{\varphi(p)}\mathcal{M}$, we have

$$\begin{aligned} & \left\langle \varphi^*(\mathrm{grad}_g f), \varphi * X \right\rangle_{\varphi^* g}(p) = \langle \mathrm{grad}_g f, X \rangle_g(\varphi(p)) \\ = & \ (Xf)(\varphi(p)) = (\varphi^* X)(\varphi^* f)(p) = \left\langle \mathrm{grad}_{\varphi^* g}(\varphi^* f), \varphi^* X \right\rangle_{\varphi^* g}(p). \end{aligned}$$

Thus we prove (2.3.10). □

Note 2.17

If $\varphi_t : \mathcal{M} \to \mathcal{M}$ is the 1-parameter family of diffeomorphisms and α is a tensor field, then

$$\partial_t \left(\varphi_t^* \alpha\right) = \mathscr{L}_{X_t} \left(\varphi_t^* \alpha\right), \tag{2.3.11}$$

where

$$X_{t_0} := \partial_{t|t=0} \left(\varphi_{t_0}^{-1} \circ \varphi_t\right) = \left(\varphi_{t_0}^{-1}\right)_* \partial_{t|t=0} \varphi_t.$$

Here we have not assumed that φ_t is a group. ♣

Definition 2.1

We say that a diffeomorphism $\psi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ is an **isometry** *if $\psi^* h = g$. If we do not require ψ to be a diffeomorphism, then ψ is called a* **local isometry**. *Two Riemannian manifolds are said to be* **isometric** *if there is an isometry from one to the other.* ♣

We say that a vector field X on $(\mathcal{M}, g)$ is **Killing** if $\mathscr{L}_X g = 0$. If X is a complete Killing vector field, then the 1-parameter group of diffeomorphisms φ_t that it generates is a 1-parameter group of isometries of $(\mathcal{M}, g)$. Indeed,

$$\partial_t \left(\varphi_t^* g\right) = \mathscr{L}_{\varphi_t^* X} (\varphi_t^* g) = \varphi_t^* \left(\mathscr{L}_X g\right) = 0.$$

Note 2.18

(1) Prove the **Jacobi identity**

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

for vector fields X, Y, Z as follows: Let $\varphi_t : \mathcal{M} \to \mathcal{M}$ be the 1-parameter group of diffeomorphisms generated by X and take the time derivative at $t = 0$ of the "invariance of the Lie bracket under diffeomorphism" equation

$$\varphi_t^*[Y, Z] = [\varphi_t^* Y, \varphi_t^* Z].$$

(2) (Kazdan, 1981) Prove the twice contracted second Bianchi identities by considering the diffeomorphism invariance of the scalar curvature and Riemannian curvature tensor.

(a) To obtain the twice contracted second Bianchi identity (2.3.3) we use the equation

$$\mathfrak{D}R_g(\mathscr{L}_X g) = \mathscr{L}_X R_g = X^i \nabla_i R_g \tag{2.3.12}$$

where $\mathfrak{D}R_g(\mathscr{L}_X g)$ denotes the linearization of R_g in the direction $\mathscr{L}_X g$.

(b) To prove the second Bianchi identity (2.3.2), we use

$$\mathfrak{D}\mathrm{Rm}_g(\mathscr{L}_X g) = \mathscr{L}_X \mathrm{Rm}_g. \tag{2.3.13}$$

♣

2.3.3 Commuting covariant derivatives

The **Ricci identities** are

$$(\nabla_i\nabla_j - \nabla_j\nabla_i)\alpha_{k_1\cdots k_r} = -\sum_{1\le\ell\le r} R_{ijk_r}^p \alpha_{k_1\cdots k_{\ell-1}pk_{\ell+1}\cdots k_r}. \tag{2.3.14}$$

If α is a 1-form, then

$$(\nabla_i\nabla_j - \nabla_j\nabla_i)\alpha_k = -R_{ijk}^\ell \alpha_\ell.$$

If β is a 2-form, then

$$\nabla_i\nabla_j\beta_{k\ell} - \nabla_j\nabla_i\beta_{k\ell} = -R_{ijk}^p\beta_{p\ell} - R_{ij\ell}^p\beta_{kp}.$$

Note 2.19

The vector space of Killing vector fields is a Lie algebra. ♣

Note 2.20

$$\begin{aligned}\nabla_i\nabla_j\alpha_{k_1\cdots k_r}^{\ell_1\cdots\ell_s} - \nabla_j\nabla_i\alpha_{k_1\cdots k_r}^{\ell_1\cdots\ell_s} &= -\sum_{1\le k\le r}\sum_{1\le p\le m} R_{ijk_h}^p \alpha_{k_1\cdots k_{h-1}pk_{h+1}\cdots k_r}^{\ell_1\cdots\ell_s} \\ &\quad + \sum_{1\le h\le s}\sum_{1\le p\le m} R_{ijp}^{\ell_h}\alpha_{k_1\cdots k_r}^{\ell_1\cdots\ell_{h-1}p\ell_{h+1}\cdots l_s}\end{aligned} \tag{2.3.15}$$

♣

2.3.4 The fundamental curvature equations in Riemannian geometry

Let $(\mathcal{M}, g)$ be an m-dimensional Riemannian manifold. If $f : \mathcal{M} \to \mathbf{R}$ is smooth, we define a self-adjoint $(1,1)$-tensor by

$$S_f : C^\infty(\mathcal{M}, T\mathcal{M}) \longrightarrow C^\infty(\mathcal{M}, T\mathcal{M}), \quad X \longmapsto (\nabla_g)_X\nabla_g f. \tag{2.3.16}$$

Then the Hessian of f can be written as

$$\mathrm{Hess}_g f(X,Y) = g\left(S_f(X), Y\right).$$

We say that a map $r : \mathcal{U} \to \mathbf{R}$, where $\mathcal{U} \subset \mathcal{M}$ is open, is a **distance function** if $|\nabla_g r|_g \equiv 1$ on $\mathcal{U}$. Distance functions are solutions to the **Hamilton-Jacobi equation**

$$|\nabla_g r|_g^2 = 1.$$

Example 2.7

(1) On $(\mathbf{R}^m, g_{\mathrm{can}})$ and a fixed point $y \in \mathbf{R}^m$, we define

$$r(x) := |x-y|_{g_{\mathrm{can}}}.$$

Then $r(x)$ is smooth on $\mathbf{R}^m \setminus \{y\}$ and has $|\nabla_{g_{\text{can}}} r|_{g_{\text{can}}} \equiv 1$.
(2) More generally, if $\mathcal{M} \subset \mathbf{R}^n$ is a submanifold, then it can be shown that

$$r(x) = d(x, \mathcal{M}) = \inf\{d(x, y) : y \in \mathcal{M}\}$$

is a distance function on some open set $\mathcal{U} \subset \mathbf{R}^n$.
(3) On $I \times \mathcal{M}$, where $I \subset \mathbf{R}$, is an integral we have metrics of the form $dr^2 + g_r$, where dr^2 is the standard metric on I and g_r is a metric on $\{r\} \times \mathcal{M}$ that depends on r. In this case the projection $I \times \mathcal{M} \to I$ is a distance function. ♠

Lemma 2.2
Given $r : \mathcal{U} \to I \subset \mathbf{R}$, then r is a distance function if and only if r is Riemannian submersion. ♡

Proof. From $dr(v) = g\left(\nabla_g r, v\right)$, we see that $v \perp \nabla_g r$ if and only if $Dr(v) := dr(v)\partial_t$, where ∂_t is the basis for TI. Thus, v is perpendicular to the kernel of Dr if and only if it is proportional to $\nabla_g r$. For $v = \alpha \nabla_g r$, we have

$$Dr(v) = \alpha Dr\left(\nabla_g r\right) = \alpha g\left(\nabla_g r, \nabla_g r\right)\partial_t.$$

Since ∂_t has length 1 in I, it follows that

$$|v|_g = |\alpha| \left|\nabla_g r\right|_g, \quad |Dr(v)| = |\alpha| \left|\nabla_g r\right|_g^2.$$

Thus, r is a Riemannian submersion if and only if $\left|\nabla_g r\right|_g = 1$. □

Let us fix a distance function $r : \mathcal{U} \to \mathbf{R}$ and an open subset $\mathcal{U} \subset \mathcal{M}$ of an m-dimensional Riemannian manifold $(\mathcal{M}, g)$. The dual of the gradient $\nabla_g r$ will usually be denoted by

$$\partial_r := g^{ij} \partial_i r \partial_j. \tag{2.3.17}$$

This is a tangent vector field over $\mathcal{U}$. The level sets for r are denoted

$$\mathcal{U}_r := \{x \in \mathcal{U} : r(x) = r\}, \tag{2.3.18}$$

and the induced metric on $\mathcal{U}_r$ is g_r. Set

$$S_{g,r}(\cdot) = \nabla_g \partial_r$$

so that

$$\text{Hess}_g r(X, Y) = g\left(S_{g,r}(X), Y\right).$$

$S_{g,r}$ stands for second derivative or **sharp operator** or **second fundamental form**. The last two terms are more or less synonymous and refer to the shape of $(\mathcal{U}_r, g_r)$ in $(\mathcal{U}, g) \subset (\mathcal{M}, g)$. The idea is that $S_{g,r} = \nabla_g \partial_r$ measures how the induced metric on $\mathcal{U}_r$ changes by computing how the unit normal to $\mathcal{U}_r$ changes.

Example 2.8

Let $\mathcal{M} \subset \mathbf{R}^{m+1}$ be an orientable hypersurface, ν the unit normal, and S_ν the sharp operator defined by

$$S_\nu(V) = \nabla_V \nu, \quad V \in C^\infty(\mathcal{M}, T\mathcal{M})$$

where $\nabla = \nabla_{g_{\text{can}}}$. If $S_\nu \equiv 0$ on $\mathcal{M}$ then ν must be a constant vector field on $\mathcal{M}$, and hence $\mathcal{M}$ is an open subset of the hyperplane

$$H = \{x + p \in \mathbf{R}^{m+1} : x \cdot \nu_p = 0\},$$

Recall our isometric immersion or embedding $(\mathbf{R}^m, g_{\text{can}}) \to (\mathbf{R}^{m+1}, g_{\text{can}})$ defined by

$$(x^1, \cdots, x^m) \longmapsto (\gamma(x^1), x^2, \cdots, x^m)$$

where $\gamma = (\gamma^1, \gamma^2) : \mathbf{R} \to \mathbf{R}^2$ is a unit speed curve. In this case,

$$\nu = (\nu(x^1), 0, \cdots, 0)$$

is a unit normal, where $\nu(x^1)$ is the unit normal to γ in $\mathbf{R}^2$. Then

$$\nu = \left(-\dot{\gamma}^2(x^1), \dot{\gamma}^1(x^1), 0, \cdots, 0\right)$$

in Cartesian coordinates. Calculate

$$\begin{aligned} \nabla \nu &= -d\left(\dot{\gamma}^2\right) \otimes \partial_1 + d\left(\dot{\gamma}^1\right) \otimes \partial_2 = -\ddot{\gamma}^2 dx^1 \otimes \partial_1 + \ddot{\gamma}^1 dx^1 \otimes \partial_2 \\ &= \left(-\ddot{\gamma}^2 \partial_1 + \ddot{\gamma}^1 \partial_2\right) \otimes dx^1. \end{aligned}$$

Thus, $S_\nu \equiv 0$ if and only if $\ddot{\gamma}^1 = \ddot{\gamma}^2 = 0$ if and only if γ is a straight line if and only if $\mathcal{M}$ is an open subset of a hyperplane. ♠

Theorem 2.3. (The radial curvature equation)

If $\mathcal{U} \subset (\mathcal{M}, g)$ is an open set and $r : \mathcal{U} \to \mathbf{R}$ a distance function, then

$$(\nabla_g)_{\partial_r} S_{g,r} + S_{g,r}^2 = -\mathrm{Rm}_g(\partial_r). \tag{2.3.19}$$

♡

Proof. If X is a vector field on $\mathcal{U}$, then

$$\begin{aligned} &\left((\nabla_g)_{\partial_r} S_{g,r}\right)(X) + S_{g,r}^2(X) = (\nabla_g)_{\partial_r}\left(S_{g,r}(X)\right) - S_{g,r}\left((\nabla_g)_{\partial_r} X\right) + S_{g,r}^2(X) \\ &= (\nabla_g)_{\partial_r}(\nabla_g)_X \partial_r - (\nabla_g)_{(\nabla_g)_{\partial_r} X}\partial_r + (\nabla_g)_{(\nabla_g)_X \partial_r}\partial_r = (\nabla_g)_{\partial_r}(\nabla_g)_X \partial_r - (\nabla_g)_{[\partial_r, X]}\partial_r; \end{aligned}$$

and

$$-\mathrm{Rm}_g(\partial_r)(X) = -\mathrm{Rm}_g(X, \partial_r)\partial_r = -(\nabla_g)_X(\nabla_g)_{\partial_r}\partial_r + (\nabla_g)_{\partial_r}(\nabla_g)_X \partial_r - (\nabla_g)_{[\partial_r, X]}\partial_r.$$

To finish the proof we shall check what happens to the term $-(\nabla_g)_X(\nabla_g)_{\partial_r}\partial_r$. By definition of the distance function, we have

$$\begin{aligned} g\left((\nabla_g)_{\partial_r}\partial_r, Y\right) &= g\left(S_{g,r}(\partial_r), Y\right) = \mathrm{Hess}_g r\left(Y, \partial_r\right) = \mathrm{Hess}_g r\left(Y, \partial_r\right) \\ &= g\left(S_{g,r}(Y), \partial_r\right) = g\left((\nabla_g)_Y \partial_r, \partial_r\right) = \frac{1}{2}(\nabla_g)_Y g\left(\partial_r, \partial_r\right) = \frac{1}{2}(\nabla_g)_Y 1 \end{aligned}$$

for any vector field Y on $\mathcal{U}$. In particular, $(\nabla_g)_{\partial_r}\partial_r = 0$ on $\mathcal{U}$. □

Each vector v on the level set $\mathcal{U}_r$ can be decomposed into normal and tangent components:

$$v = v^\top + v^\perp = (v - g(v, \partial_r)\partial_r) + g(v, \partial_r)\partial_r. \tag{2.3.20}$$

The decomposition is a direct sum since

$$\begin{aligned} g\left(v - g(v, \partial_r)\partial_r, g(v, \partial_r)\partial_r\right) &= g\left(v, g(v, \partial_r)\partial_r\right) - g\left(g(v, \partial_r)\partial_r, g(v, \partial_r)\partial_r\right) \\ &= g(v, \partial_r)^2 - g(v, \partial_r)^2 g(\partial_r, \partial_r) \quad = \quad 0. \end{aligned}$$

Theorem 2.4. (Tangent curvature equation)

For tangent vector fields X, Y, Z, W on the level set $\mathcal{U}_r$, we have

$$\begin{aligned} \left(\mathrm{Rm}_g(X, Y)Z\right)^\top &= \mathrm{Rm}_{g_r}(X, Y)Z - \left(S_{g,r}(X) \wedge S_{g,r}(Y)\right)(Z), \\ \mathrm{Rm}_g(X, Y, Z, W) &= \mathrm{Rm}_{g_r}(X, Y, Z, W) - \mathrm{II}_{g_r}(Y, Z)\mathrm{II}_{g_r}(X, W) \\ &\quad + \mathrm{II}_{g_r}(X, Z)\mathrm{II}_{g_r}(Y, W), \end{aligned}$$

where

$$\mathrm{II}_{g_r}(U, V) := \mathrm{Hess}_g r(U, V) = g\left(S_{g,r}(U), V\right)$$

is the **classical second fundamental form**. ♡

Proof. If X, Y are vector fields that are tangent to the level set $\mathcal{U}_r$, then we claim that

$$(\nabla_{g_r})_X Y = (\nabla_g)_X Y + \mathrm{II}_{g_r}(X, Y)\partial_r. \tag{2.3.21}$$

By definition, we have

$$(\nabla_{g_r})_X Y = \left((\nabla_g)_X Y\right)^\top = (\nabla_g)_X Y - g\left((\nabla_g)_X Y, \partial_r\right)\partial_r.$$

Since $Y \perp \partial_r$, it follows that

$$0 = (\nabla_g)_X g(Y, \partial_r) = g\left((\nabla_g)_X Y, \partial_r\right) + g\left(Y, S_{g,r}(X)\right)$$

and hence

$$(\nabla_{g_r})_X Y = (\nabla_g)_X Y + g\left(S_{g,r}(X), Y\right)\partial_r = (\nabla_g)_X Y + \mathrm{II}_{g_r}(X, Y)\partial_r.$$

Using (2.3.21) yields

$$\begin{aligned} \mathrm{Rm}_g(X, Y)Z &= \mathrm{Rm}_{g_r}(X, Y)Z - \left(S_{g,r}(X) \wedge S_{g,r}(Y)\right)(Z) \\ &\quad + g\left(-\left((\nabla_g)_X S_{g,r}\right)(Y) + \left((\nabla_g)_Y S_{g,r}\right)(X), Z\right) \cdot \partial_r. \end{aligned}$$

This establishes the first part of each formula. The second part follows from the definition. □

Theorem 2.5. (The normal or mixed curvature equation)

For tangent vector fields X, Y, Z on the level set $\mathcal{U}_r$, we have

$$\begin{aligned} \mathrm{Rm}_g(X, Y, Z, \partial_r) &= g\left(-\left((\nabla_g)_X S_{g,r}\right)(Y) + \left((\nabla_g)_Y S_{g,r}\right)(X), Z\right) \\ &= -\left((\nabla_g)_X \mathrm{II}_{g_r}\right)(Y, Z) + \left((\nabla_g)_Y \mathrm{II}_{g_r}\right)(X, Z). \end{aligned}$$

♡

Proof. Use the similar method in the proof of previous theorem. □

If $B(\cdot,\cdot)$ is a symmetric 2-form and $\mathscr{B}_g(\cdot)$ the corresponding self-adjoint $(1,1)$-tensor field defined via

$$g\left(\mathscr{B}_g(X),Y\right)=B(X,Y),$$

then the square of $\mathscr{B}$ is the symmetric bilinear form corresponding to $\mathscr{B}_g^2$

$$B^2(X,Y)=g\left(\mathscr{B}_g^2(X),Y\right)=g\left(\mathscr{B}_g(X),\mathscr{B}_g(Y)\right).$$

Note that this symmetric bilinear form is always nonnegative, i.e., $B^2(X,X)\geq 0$ for all X.

For example, if $B(\cdot,\cdot)=\mathrm{Hess}_g r(\cdot,\cdot)$, then $\mathscr{B}_g(\cdot)=S_{g,r}(\cdot)=\nabla_g\partial_r$.

Proposition 2.2

If we have a smooth distance function $r:(\mathcal{U},g)\to\mathbf{R}$ *and denote* $\mathrm{grad}_g r=\partial_r$, *then*

$$\begin{aligned}\mathscr{L}_{\frac{1}{2}\partial_r}g &= \mathrm{Hess}_g r,\\ \left((\nabla_g)_{\partial_r}\mathrm{Hess}_g r\right)(X,Y)+\mathrm{Hess}_g^2 r(X,Y) &= -\mathrm{Rm}_g(X,\partial_r,\partial_r,Y),\\ \left(\mathscr{L}_{\partial_r}\mathrm{Hess}_g r\right)(X,Y)-\mathrm{Hess}_g^2 r(X,Y) &= -\mathrm{Rm}_g(X,\partial_r,\partial_r,Y).\end{aligned}$$

♡

Proof. We have proved $(\nabla_g)_{\partial_r}\partial_r=0$ in the proof of **Theorem 2.3**. Keep in mind that $(\nabla_g)_X\partial_r$ is the self-adjoint operator corresponding to $\mathrm{Hess}_g r$. Using the radial curvature equation, we obtain

$$\begin{aligned}\left((\nabla_g)_{\partial_r}\mathrm{Hess}_g r\right)(X,Y) &= \partial_r\mathrm{Hess}_g r(X,Y)-\mathrm{Hess}_g r\left((\nabla_g)_{\partial_r}X,Y\right)-\mathrm{Hess}_g r\left(X,(\nabla_g)_{\partial_r}Y\right)\\ &= \partial_r g\left((\nabla_g)_X\partial_r,Y\right)-g\left((\nabla_g)_{(\nabla_g)_{\partial_r}X}\partial_r,Y\right)-g\left((\nabla_g)_X\partial_r,(\nabla_g)_{\partial_r}Y\right)\\ &= g\left((\nabla_g)_{\partial_r}(\nabla_g)_X\partial_r,Y\right)-g\left((\nabla_g)_{(\nabla_g)_{\partial_r}X}\partial_r,Y\right)\\ &\quad+g\left((\nabla_g)_X\partial_r,(\nabla_g)_{\partial_r}Y\right)-g\left((\nabla_g)_X\partial_r,(\nabla_g)_{\partial_r}Y\right)\\ &= g\left(\mathfrak{Rm}_g(\partial_r,X)\partial_r,Y\right)-g\left((\nabla_g)_{(\nabla_g)_X\partial_r}\partial_r,Y\right)\\ &= -\mathrm{Rm}_g(X,\partial_r,\partial_r,Y)-g\left((\nabla_g)_Y\partial_r,(\nabla_g)_X\partial_r\right) = -\mathrm{Rm}_g\left(X,\partial_r,\partial_r,Y\right)-\mathrm{Hess}_g^2 r(X,Y).\end{aligned}$$

Similarly, we can prove the third equation. □

A **Jacobi field** for a smooth distance function r is a smooth vector field J that does not depend on r, i.e., it satisfies the **Jacobi equation**

$$\mathscr{L}_{\partial_r}J=0. \tag{2.3.22}$$

This is a first-order linear PDE, which can be solved by the method of characteristics. Locally we select a coordinate system $(r,x^2,\cdots,x^m)$ where r is the first coordinate. Then $J=a^r\partial_r+a^i\partial_i$ and the Jacobi equation becomes

$$0=\mathscr{L}_{\partial_r}J=\mathscr{L}_{\partial_r}\left(a^r\partial_r+a^i\partial_i\right)=\partial_r(a^r)\partial_r+\partial_r(a^i)\partial_i.$$

Thus the coefficients a^r, a^i have to be independent of r. Since

$$\begin{aligned} -\mathrm{Rm}_g(J,\partial_r)\partial_r &= \mathrm{Rm}_g(\partial_r,J)\partial_r = (\nabla_g)_{\partial_r}(\nabla_g)_J\partial_r - (\nabla_g)_J(\nabla_g)_{\partial_r}\partial_r - (\nabla_g)_{[\partial_r,J]}\partial_r \\ (\nabla_g)_{\partial_r}(\nabla_g)_J\partial_r - (\nabla_g)_{[\partial_r,J]}\partial_r &= (\nabla_g)_{\partial_r}(\nabla_g)_{\partial_r}J - (\nabla_g)_{\partial_r}[\partial_r,J] - (\nabla_g)_{[\partial_r,J]}\partial_r \end{aligned}$$

it follows that (2.3.22) satisfies a more general second-order equation, also known as the Jacobi equation:

$$(\nabla_g)_{\partial_r}(\nabla_g)_{\partial_r}J + \mathrm{Rm}_g(J,\partial_r)\partial_r = 0. \tag{2.3.23}$$

If $J1$ and J_2 are Jacobi fields, then

$$\begin{aligned} \partial_r\left(g(J_1,J_2)\right) &= 2\mathrm{Hess}_g r(J_1,J_2), \\ \partial_r\left(\mathrm{Hess}_g r(J_1,J_2)\right) - \mathrm{Hess}_g^2 r(J_1,J_2) &= -\mathrm{Rm}_g(J_1,\partial_r,\partial_r,J_2) \end{aligned}$$

according to **Proposition 2.2**.

A **parallel field** for a smooth distance function r is a vector field X such that

$$(\nabla_g)_{\partial_r}X = 0. \tag{2.3.24}$$

This is a first-order linear PDE and can be solved in a similar manner. However, one crucial difference is that parallel fields are almost never Jacobi fields.

If X_1, X_2 are parallel fields for a smooth distance function r, then

$$\begin{aligned} \partial_r\left(g(X_1,X_2)\right) &= 0, \\ \partial_r\left(\mathrm{Hess}_g r(X_1,X_2)\right) + \mathrm{Hess}_g^2 r(X_1,X_2) &= -\mathrm{Rm}_g(X_1,\partial_r,\partial_r,X_2). \end{aligned}$$

2.4 Examples

Introduction

- *Warped products*
- *Hyperbolic spaces*
- *Metrics on Lie groups*
- *Riemannian submersions*

Let $(\mathcal{M},g)$ be an m-dimensional Riemannian manifold. Recall operators $\mathbf{Rm}_g^\wedge$ and $\mathbf{Rm}_g$:

$$\begin{aligned} \mathbf{Rm}_g^\wedge : C^\infty(\mathcal{M},\wedge^2\mathcal{M}) \times C^\infty(\mathcal{M},\wedge^2\mathcal{M}) &\longrightarrow C^\infty(\mathcal{M}), \\ (X\wedge Y, V\wedge W) &\longmapsto \mathrm{Rm}_g(X,Y,W,V), \\ \mathbf{Rm}_g : C^\infty(\mathcal{M},\wedge^2\mathcal{M}) &\longrightarrow C^\infty(\mathcal{M},\wedge^2\mathcal{M}), \\ g\left(\mathbf{Rm}_g\left(X\wedge Y\right), V\wedge W\right) &= \mathbf{Rm}_g^\wedge\left(X\wedge Y, V\wedge W\right) \end{aligned}$$

Proposition 2.3

Let $\{e_i\}_{1\le i\le m}$ be an orthonormal basis for $T_p\mathcal{M}$. If $\{e_i\wedge e_j\}_{1\le i,j\le m}$ diagonalize the Riemann curvature operator

$$\mathbf{Rm}_g(e_i\wedge e_j) = \lambda_{ij}e_i\wedge e_j$$

then for any plane Π in $T_p\mathcal{M}$ we have $\mathrm{Sec}_g(\Pi) \in [\min_{i,j} \lambda_{ij}, \max_{i,j} \lambda_{ij}]$. ♡

Proof. If v, w form an orthonormal basis for Π, then we have

$$\mathrm{Sec}_g(\Pi) = g\left(\mathbf{Rm}_g(v \wedge w), v \wedge w\right).$$

In this situation, the result is immediate. □

Proposition 2.4

Let $\{e_i\}_{1 \le i \le m}$ be an orthonormal basis for $T_p\mathcal{M}$ and suppose that $\mathrm{Rm}_g(e_i, e_j)e_k = 0$ if the indices are mutually distinct; then $e_i \wedge e_j$ diagonalize the Riemann curvature operator. ♡

Proof. If we use

$$g\left(\mathbf{Rm}_g(e_i \wedge e_j), e_k \wedge e_\ell\right) = -g\left(\mathrm{Rm}_g(e_i, e_j)e_k, e_\ell\right) = g\left(\mathrm{Rm}_g(e_i, e_j)e_\ell, e_k\right),$$

then we see this expression is zero when i, j, k are mutually distinct, or if i, j, ℓ are mutually distinct. Thus, the expression can only be nonzero when $\{k, \ell\} = \{i, j\}$. □

Proposition 2.5

Let $\{e_i\}_{1 \le i \le m}$ be an orthonormal basis for $T_p\mathcal{M}$ and suppose that

$$g\left(\mathrm{Rm}_g(e_i, e_j)e_k, e_\ell\right) = 0$$

if three of indices are mutually distinct, then e_i diagonalize Rc_g. ♡

Proof. By definition, we have

$$g\left(\mathrm{Rc}_g(e_i), e_j\right) = \sum_{1 \le k \le m} g\left(\mathrm{Rm}_g(e_i, e_k)e_k, e_j\right).$$

If $i \ne j$, then $g\left(\mathrm{Rm}_g(e_i, e_k)e_k, e_j\right) = 0$ unless k is either i or j. If $k = i, j$, then the expression is zero from the symmetry properties. Hence, e_i must diagonalize Rc_g. □

2.4.1 Warped products

In this subsection we consider the rotationally symmetric metrics, doubly warped products, and the Schwarzschild metric.

Example 2.9. (Spheres)

On $(\mathbf{R}^m, g_{\mathrm{can}})$ we have the standard distance function $r(\boldsymbol{x}) = |\boldsymbol{x}|$ and the polar coordinate:

$$g_{\mathrm{can}} = dr \otimes dr + g_r = dr \otimes dr + r^2 ds_{m-1} \otimes ds_{m-1},$$

where $ds_{m-1} \otimes ds_{m-1}$ is the canonical metric on $\mathbf{S}^{m-1}(1)$. The level sets are $U_r = \mathbf{S}^{m-1}(r)$ with the induced metric $g_r = r^2 ds_{m-1} \otimes ds_{m-1}$. The gradient is

$$\partial_r = \frac{1}{r} x^i \partial_i.$$

Since $ds_{m-1} \otimes ds_{m-1}$ is independent of r we compute the Hessian as follows:

$$\begin{aligned} 2\mathrm{Hess}_{g_{\mathrm{can}}} r &= \mathscr{L}_{\partial_r} g_{\mathrm{can}} = \mathscr{L}_{\partial_r}(dr \otimes dr) + \mathscr{L}_{\partial_r}\left(r^2 ds_{m-1} \otimes ds_{m-1}\right) \\ &= \partial_r(r^2) ds_{m-1} \otimes ds_{m-1} = 2r ds_{m-1} \otimes ds_{m-1} = 2\frac{1}{r} g_r. \end{aligned}$$

Hence $\mathrm{Hess}_{g_{\mathrm{can}}} r = \frac{1}{r} g_r$. *The tangent curvature equation then tells us that*

$$\mathrm{Rm}_{g_r}(X,Y)Z = \frac{1}{r^2}\left(g_r(Y,Z)X - g_r(X,Z)Y\right).$$

If particular, if e_i is any orthonormal basis, we see that $\mathrm{Rm}_{g_r}(e_i, e_j)e_k = 0$ when the indices are mutually distinct. Therefore, $(\mathbf{S}^{m-1}(r), g_{\mathrm{can}})$ has constant sectional curvature $\frac{1}{r^2}$, provided that $m \geq 3$.

Example 2.10. (Product spheres)

Consider the product spheres

$$\mathbf{S}^n_a \times \mathbf{S}^m_b := \mathbf{S}^n\left(\frac{1}{\sqrt{a}}\right) \times \mathbf{S}^m\left(\frac{1}{\sqrt{b}}\right) = \left(\mathbf{S}^n \times \mathbf{S}^m, \frac{1}{a} ds_n \otimes ds_n + \frac{1}{b} ds_m \otimes ds_m\right).$$

Let $g = \frac{1}{a} ds_n \otimes ds_n + \frac{1}{b} ds_m \otimes ds_m$. If X, Y are tangent to $\mathbf{S}^m$ and U, V tangent to $\mathbf{S}^m$,

$$\mathbf{Rm}_g(X \wedge V) = 0, \quad \mathbf{Rm}_g(X \wedge Y) = aX \wedge Y, \quad \mathbf{Rm}_g(U \wedge V) = bU \wedge V.$$

By **Proposition 2.3**, *all sectional curvatures lie in the interval $[0, \max\{a,b\}]$. Moreover,*

$$\mathrm{Rc}_g(X) = (n-1)aX, \quad \mathrm{Rc}_g(V) = (m-1)bV, \quad R_g = n(n-1)a + m(m-1)b.$$

Therefore, $\mathbf{S}^n_a \times \mathbf{S}^m_b$ always has constant scalar curvature, is an Einstein manifold exactly when $(n-1)a = (m-1)b$ (which requires $n, m \geq 2$ or $n = m = 1$), and has constant sectional curvature only when $n = m = 1$.

Example 2.11. (Rotationally symmetric metrics)

Consider the metric of the form

$$dr \otimes dr + \varphi^2(r) ds_{n-1} \otimes ds_{n-1} = dr \otimes dr + g_r$$

on $(a,b) \times \mathbf{S}^{n-1}$. Then the Hessian of the distance function r is

$$\mathrm{Hess}_g r = \frac{\partial_r \varphi}{\varphi} g_r.$$

If X is tangent to $\mathbf{S}^{n-1}$, then

$$\mathrm{Rc}_g(X) = \left[(n-2)\frac{1-\dot{\varphi}^2}{\varphi^2} - \frac{\ddot{\varphi}}{\varphi}\right] X;$$

if $X = \partial_r$, then

$$\mathrm{Rc}_g(\partial_r) = -(n-1)\frac{\ddot{\varphi}}{\varphi}\partial_r.$$

For the Riemann curvature operator, we have

$$\mathbf{Rm}_g(X \wedge \partial_r) = -\frac{\ddot{\varphi}}{\varphi} X \wedge \partial_r, \quad \mathbf{Rm}_g(X \wedge Y) = \frac{1-\dot{\varphi}^2}{\varphi^2} X \wedge Y$$

for any vector fields X, Y tangent to $\mathbf{S}^{n-1}$. Thus, all sectional curvatures lie between the two values $-\frac{\ddot{\varphi}}{\varphi}$ and $\frac{1-\dot{\varphi}^2}{\varphi^2}$. When $n = 2$, we have $\mathrm{Sec}_g = -\frac{\ddot{\varphi}}{\varphi}$, since there are no tangential curvatures.

(a) *For the metric*

$$dr \otimes dr + \mathrm{sn}_k^2(r) ds_{n-1} \otimes ds_{n-1}$$

on $\mathbf{S}^n\left(\frac{1}{\sqrt{k}}\right)$, since $\varphi(r) = \mathrm{sn}_k(r)$, we see all sectional curvatures are equal to k.

(b) *If the metric g is Ricci flat, then*

$$\frac{\ddot{\varphi}}{\varphi} = 0 = (n-2)\frac{1-\dot{\varphi}^2}{\varphi^2} - \frac{\ddot{\varphi}}{\varphi}.$$

If $n > 2$, we must have $\ddot{\varphi} \equiv 0$ and $\dot{\varphi}^2 \equiv 1$. Thus $\varphi(r) = a \pm r$. In case $n = 2$, we only need $\ddot{\varphi} \equiv 0$.

(c) *If the metric g is scalar flat ($n \geq 3$), then*

$$2(n-1)\left[-\frac{\ddot{\varphi}}{\varphi} + \frac{n-2}{2} \cdot \frac{1-\dot{\varphi}^2}{\varphi^2}\right] = 0.$$

Thus, we suffices to solve the equation

$$-\varphi\ddot{\varphi} + \frac{n-2}{2}\left(1-\dot{\varphi}^2\right) = 0.$$

Introducing the variables

$$\dot{\varphi} = G(\varphi)$$

we find that the above ODE reduces to the first-order equation

$$-\varphi G' G + \frac{n-2}{2}(1-G^2) = 0.$$

Using the separation of variables, we see that G and φ are related by

$$\dot{\varphi}^2 = G^2 = 1 + C\varphi^{2-n}.$$

♠

Example 2.12. (Doubly warped products)

Consider the metric

$$\left(I \times \mathbf{S}^p \times \mathbf{S}^q, dt \otimes dt + \varphi^2(r) ds_p \otimes ds_p + \psi^2(r) ds_q \otimes ds_q\right).$$

The Hessian is

$$\mathrm{Hess}_g r = (\partial_r \varphi)\, \varphi ds_p \otimes ds_p + (\partial_r \psi)\, \psi ds_q \otimes ds_q.$$

Let X, Y be tangent to $\mathbf{S}^p$ and V, W tangent to $\mathbf{S}^q$. Then

$$\mathbf{Rm}_g(\partial_r \wedge X) = -\frac{\ddot{\varphi}}{\varphi}\partial_r \wedge X, \quad \mathbf{Rm}_g(\partial_r \wedge V) = -\frac{\ddot{\psi}}{\psi}\partial_r \wedge V, \quad \mathbf{Rm}_g(X \wedge V) = -\frac{\dot{\varphi}\dot{\psi}}{\varphi\psi} X \wedge V$$

$$\mathbf{Rm}_g(X \wedge Y) = \frac{1-\dot{\varphi}^2}{\varphi^2} X \wedge Y, \quad \mathbf{Rm}_g(U \wedge V) = \frac{1-\dot{\psi}^2}{\psi^2} U \wedge V.$$

Moreover,

$$\mathrm{Rc}_g(\partial_r) = \left(-p\frac{\ddot{\varphi}}{\varphi} - q\frac{\ddot{\psi}}{\psi}\right)\partial_r, \quad \mathrm{Rc}_g(X) = \left(-\frac{\ddot{\varphi}}{\varphi} + (p-1)\frac{1-\dot{\varphi}^2}{\varphi^2} - q\frac{\dot{\varphi}\dot{\psi}}{\varphi\psi}\right) X,$$

$$\mathrm{Rc}_g(V) \quad = \quad \left(-\frac{\ddot{\psi}}{\psi} + (q-1)\frac{1-\dot{\psi}^2}{\psi^2} - p\frac{\dot{\varphi}\dot{\psi}}{\varphi\psi}\right) V.$$

Example 2.13. (The Schwarzschild metric)

We wish to find a Ricci flat metric on $\mathbf{R}^2 \times \mathbf{S}^2$. *Let* $p = 1$ *and* $q = 2$ *in the above doubly warped product case. This means we have to solve the following equstions*

$$\begin{aligned} -\frac{\ddot{\varphi}}{\varphi} - 2\frac{\ddot{\psi}}{\psi} &= 0, \\ -\frac{\ddot{\varphi}}{\varphi} - 2\frac{\dot{\varphi}\dot{\psi}}{\varphi\psi} &= 0, \\ -\frac{\ddot{\psi}}{\psi} + \frac{1-\dot{\psi}^2}{\psi^2} - \frac{\dot{\varphi}\dot{\psi}}{\varphi\psi} &= 0. \end{aligned}$$

The first equation yields $(\dot{\psi}/\dot{\varphi}) = \alpha$ *for some constant* α. *Hence the above system of equations reduces to*

$$\begin{aligned} 2\psi\ddot{\psi} - (1-\dot{\psi}^2) &= 0, \\ -\frac{\ddot{\varphi}}{\varphi} - \frac{4\alpha^2\dot{\varphi}^2}{1-\alpha^2\varphi^2} &= 0, \\ \dot{\psi} &= \alpha\varphi. \end{aligned}$$

One of solutions is $\psi(r) = r$ *and* $\varphi(r) = 1/a$. *To get more complicated solutions we can assume* $\dot{\psi}^2 = G(\psi)$. *Then* $G = 1 + C\psi^{-1}$ *for some constant* $C \in \mathbf{R}$. *Turning back to the system of equations, we obtain lots of solutions.*

2.4.2 Hyperbolic space

Example 2.14. (The rotationally symmetric model)

Define g *to be the rotationally symmetric metric* $dr \otimes dr + \sinh^2(r) ds_{m-1} \otimes ds_{m-1}$ *on* $\mathbf{R}^m$ *of constant sectional curvature* -1.

Example 2.15. (The upper half plane model)

Let

$$\mathfrak{H}^m = \{(x^1, \cdots, x^m) \in \mathbf{R}^m : x^m > 0\}$$

and let

$$g = \frac{dx^1 \otimes dx^1 + \cdots + dx^m \otimes dx^m}{(x^m)^2}.$$

Introducing $r := \ln(x^m)$, *we find the metric* g *can be written as*

$$g = dr \otimes dr + e^{-2r}\left(dx^1 \otimes dx^1 + \cdots + dx^m \otimes dx^m\right).$$

This metric has constant sectional curvature -1.

Example 2.16. (The Riemann model)

If $(\mathcal{M}, g)$ is an m-dimensional Riemannian manifold and φ is positive on $\mathcal{M}$, then we get a new Riemannian manifold $(\mathcal{M}, \varphi^2 g)$. Such a change in metric is called a **conformal change**, *and φ^2 is referred to as the* **conformal factor**. *On the upper half plane $\mathfrak{H}^m$ we can ask when*

$$\varphi^2 \cdot \left(dx^1 \otimes dx^1 + \cdots + dx^m \otimes dx^m\right)$$

has constant sectional curvature?

Note 2.21. (The Euler and Navier-Stokes equation on H^2)

Recently, Khesin and Misiolek[a] showed that non-uniqueness of the Leray-Hopf solutions of the Navier-Stokes equation on the hyperbolic plane $\mathbf{H}^2$ observed by Chan-Czubak is a consequence of the Hodge decomposition, and however, this phenomenon does not occur on $\mathbf{H}^m$ whenever $m \geq 3$.

Let $(\mathcal{M}, g)$ be a complete Riemannian manifold and consider the Lie algebra $\mathfrak{g}_{g,0} := \mathrm{Vect}_{g,0}(\mathcal{M})$ of (sufficiently smooth) divergence-free vector fields on $\mathcal{M}$ with finite L^2-norm with respect to g. Its dual space $\mathfrak{g}^*_{g,0}$ has a natural identification to the quotient space $\Omega^1_{L^2,g}(\mathcal{M}) \setminus \overline{d\Omega^0_{L^2,g}(\mathcal{M})}$ of the L^2 1-forms modulo (the L^2 closure of) the exact 1-forms on $\mathcal{M}$. The pairing between cosets $[\alpha] \in \Omega^1_{L^2,g}(\mathcal{M}) \setminus \overline{d\Omega^0_{L^2,g}(\mathcal{M})}$ of L^2 1-forms $\alpha \in \Omega^1_{L^2,g}(\mathcal{M})$ and vector fields $X \in \mathrm{Vect}_{g,0}(\mathcal{M})$ is given by

$$\langle [\alpha], w \rangle_g := \int_{\mathcal{M}} (\iota_X \alpha)\, dV_g,$$

where ι_X is the contraction of a differential form with a vector field X. Note that this definition is independent of the choice of representatives. Let

$$\mathfrak{g}_{g,0} \longrightarrow \mathfrak{g}^*_{g,0}$$

denote the inertia operator assigning a vector field $X \in \mathrm{Vect}_{g,0}(\mathcal{M})$ the coset $[X^\flat]$ of the corresponding 1-form $X^\flat$ via the Riemannian metric g.

The Euler equations read

$$\frac{\partial}{\partial t} X_t + \nabla_{X_t} X_t = -\mathrm{grad}_g p, \quad \mathrm{div}_g X_t = 0. \tag{2.4.1}$$

Thus, in the Hamiltonian framework (2.4.1) becomes

$$\frac{d}{dt}[X_t^\flat] = -\mathscr{L}_{X_t}[X_t^\flat]. \tag{2.4.2}$$

By Hodge decomposition, the space $\Omega^1_{L^2,g}(\mathcal{M})$ of the L^2 1-forms decomposes as

$$\Omega^1_{L^2,g}(\mathcal{M}) = \overline{d\Omega^0_{L^2,g}(\mathcal{M})} \oplus \overline{\delta_g \Omega^2_{L^2,g}(\mathcal{M})} \oplus \mathcal{H}^1_{L^2,g}(\mathcal{M}),$$

where δ_g denotes the adjoint operator of d relative to g. Therefore,

$$\mathfrak{g}^*_{g,0} = \overline{\delta_g \Omega^2_{L^2,g}(\mathcal{M})} \oplus \mathcal{H}^1_{L^2,g}(\mathcal{M}). \tag{2.4.3}$$

It turns out that the summand of the harmonic forms in (2.4.3) corresponds to steady solutions of the Euler equation.

(1) *Each harmonic* 1*-form on a complete Riemannian manifold* $(\mathcal{M}, g)$ *which belongs to* $L^2 \cap L^4$ *defines a steady solution of the Euler equation (2.4.1) or (2.4.2) on* $\mathcal{M}$*. Let* α_t *be a* L^2 *harmonic* 1*-form on* $\mathcal{M}$ *corresponding to* X_t*. According to Cartan's formula yields*

$$\frac{d}{dt}[\alpha_t] = -\mathscr{L}_{X_t}[\alpha_t] = -\left[(\iota_{X_t}d + d\iota_{X_t})\,\alpha_t\right] = -\left[d\,(\iota_{X_t}\alpha_t)\right].$$

By the assumption, it follows that

$$||\iota_{X_t}\alpha_t||^2_{L^2,g} = \int_{\mathcal{M}} |\iota_{X_t}\alpha_t|^2_g\, dV_g = \int_{\mathcal{M}} |\alpha_t(X_t)|^2_g dV_g = ||\alpha_t||^4_{L^4,g} < \infty;$$

thus $\iota_{X_t}\alpha_t \in \Omega^0_{L^2,g}(\mathcal{M})$*. Consequently,* $\frac{d}{dt}[\alpha_t] = 0 \in \mathfrak{g}^*_{g,0}$*. The latter means that the* 1*-form* α_t *defines a steady solution of the Euler equation.*

(2) **(L^2-form conjecture of Dodziuk-Singer, 1979)** *Let* $(\mathcal{M}, g)$ *be a complete simply-connected Riemannian manifold of sectional curvature* Sec_g *satisfying* $-a^2 \leq \mathrm{Sec}_g \leq -1, a \geq 1$*. Let* $\mathcal{H}^p_{L^2,g}(\mathcal{M})$ *denote the space of* L^2 *harmonic* p*-forms on* $\mathcal{M}$*, i.e.,* p*-forms* ω *on* $\mathcal{M}$ *such that*

$$\Delta_{H,g}\omega = 0, \quad \int_{\mathcal{M}} |\omega|^2_g dV_g < \infty.$$

It is clear $\mathcal{H}^p_{L^2,g}(\mathcal{M})$ *is naturally isomorphic to* $\mathcal{H}^{m-p}_{L^2,g}(\mathcal{M})$ *and* $\mathcal{H}^0_{L^2,g}(\mathcal{M}) = 0$*. Dodziuk and Singer conjectured that* $\mathcal{H}^p_{L^2,g}(\mathcal{M}) = 0$ *if* $p \neq m/2$ *and* $\dim(\mathcal{H}^{m/2}_{L^2,g}(\mathcal{M})) = \infty$ *if* m *is even. By means of the* L^2 *index theorem for regular covers of Atiyah, an affirmative solution of this conjecture implies a positive solution of the well-known Hopf conjecture: If* $\mathcal{M}^{2m}$ *is a compact manifold of negative sectional curvature, then* $(-1)^m\chi(\mathcal{M}^{2m}) > 0$*.*

Dodziuk has proved the L^2*-form conjecture for rotationally symmetric metrics – in particular for the space form* $\mathbf{H}^m(-a^2)$ *of curvature* $-a^2$*. However, this conjecture is in general not true.*

(3) **(Khesin-Misiolek, 2012)** *(i) There are no stationary* L^2 *harmonic solutions of the Euler equations on* $\mathbf{H}^m$ *for any* $m > 2$*. (ii) There exists an infinite-dimensional space of stationary* L^2 *harmonic solutions of the Euler equations on* $\mathbf{H}^2$*. The first result (i) follows from Dodziuk's result. To prove (ii) we note that the space of* L^2 *harmonic* 1*-forms on* $\mathbf{H}^2$ *is infinite-dimensional. Consider the subspace* $\mathcal{S} \subset \mathcal{H}^1_{L^2,g}(\mathbf{H}^2)$ *of* 1*-forms:*

$$\mathcal{S} := \{d\Phi : \Phi \textit{ is harmonic on } \mathbf{H}^2 \textit{ and } d\Phi \in L^2(\mathbf{H}^2)\}.$$

We claim that $\mathcal{S}$ *is infinite-dimensional. Indeed, let us consider the Poincaré model of* $\mathbf{H}^2$*, i.e., the unit disk* $\mathbf{D}$ *with the hyperbolic metric* g*, which we denote by* $\mathbf{D}_g$*. It is conformally equivalent to the standard unit disk with the Euclidean metric* e*, denoted by* $\mathbf{D}_e$*. Bounded harmonic functions on* $\mathbf{D}_g$ *can be obtained by solving the*

Dirichlet problem on $\mathbf{D}_e$. First, the 1-form $d\Phi$ is harmonic:

$$\Delta_{H,g} d\Phi = -d\delta_g d\Phi = d\Delta_{H,g}\Phi = 0.$$

Secondly,

$$\begin{aligned}
\|d\Phi\|^2_{L^2,g} &= \int_{\mathbf{D}} \langle d\Phi, d\Phi\rangle_g dV_g &= \int_{\mathbf{D}} \det(g^{ij})\langle d\Phi, d\Phi\rangle_e \det(g_{ij}) dV_e \\
&= \int_{\mathbf{D}} \langle d\Phi, d\Phi\rangle_e dV_e &= \|d\Phi\|^2_{L^2,e}, \\
\|d\Phi\|^4_{L^2,g} &= \int_{\mathbf{D}} \langle d\Phi, d\Phi\rangle^2_g dV_g &= \int_{\mathbf{D}} \det^2(g_{ij})\langle d\Phi, d\Phi\rangle^2_e \det(g_{ij}) dV_g \\
&= \int_{\mathbf{D}} (1-|z|^2)^2\langle d\Phi, d\Phi\rangle^2_e dV_e(z) &\le \|d\Phi\|^4_{L^2,e},
\end{aligned}$$

where $\det(g_{ij}) = 1/(1-|z|^2)^2$ is the determinant of the hyperbolic metric g. Furthermore, for sufficiently smooth boundary values $\varphi \in C^{1,\alpha}(\partial\mathbf{D})$ there is a uniform upper bound C for its harmonic extension inside the disk: $|d\Phi(x)| \le C\|\varphi\|_{C^{1,\alpha}(\partial\mathbf{D})}$ for any $x \in \mathbf{D}$ and $0 < \alpha < 1$. This implies that (for sufficiently smooth φ) the L^2 1-forms $d\Phi$ define an infinite-dimensional subspace $\mathcal{S}$ of harmonic forms in $L^2 \cap L^4$. By (1), they define an infinite-dimensional space of stationary solutions of the Euler equations on the hyperbolic plane $\mathbf{H}^2$.

(4) **(Chan-Czubak, 2010)** *Since suitably rescaled steady solutions of the Euler equations solve the Navier-Stokes equations on $(\mathcal{M}, g)$*

$$\frac{\partial}{\partial t}X_t + \nabla_{X_t}X_t - L_g X_t = -\mathrm{grad}_g p, \quad \mathrm{div}_g X_t = 0, \tag{2.4.4}$$

where $L_g := -\Delta_{H,g} - 2\mathrm{Rc}^\sharp_g$. Consider the hyperbolic plane $\mathbf{H}^2$ with $\mathrm{Rc}^\sharp_g = -1$. Letting $\alpha_t := X_t^\flat$ yields

$$\frac{\partial}{\partial t}\alpha_t + \nabla_{X_t}\alpha_t + \Delta_{H,g}\alpha_t - 2\alpha_t = -dp, \quad \delta_g\alpha_t = 0. \tag{2.4.5}$$

Let $\alpha_t = f(t)d\Phi$ for some $d\Phi \in \mathcal{S}$. Then (2.4.5) is equivalent to

$$dp = d\left[(2f(t) - f'(t))\Phi - \frac{1}{2}f^2(t)|d\Phi|^2_g\right];$$

consequently, the pair $(f(t)d\Phi, (2f(t) - f'(t))\Phi - \frac{1}{2}f^2(t)|d\Phi|^2_g)$, $\Phi \in \mathcal{S}$, solves (2.4.5). We say that X_t is a **Leray-Hopf solution** *of the Navier-Stokes equations if $X \in L^\infty([0,\infty), L^2) \cap L^2([0,\infty), H^1)$ and satisfies*

$$\|X_t\|^2_{L^2,g} + 4\int_0^t \|\mathrm{Def}_g X_s\|^2_{L^2,g}\, ds \le \|X_0\|^2_{L^2,g}, \quad \lim_{t\to 0}\|X_t - X_0\|_{L^2,g} = 0 \tag{2.4.6}$$

for any $0 \le t < \infty$ and where $(\mathrm{Def}_g X_t)_{ij} = \frac{1}{2}(\nabla_i(X_t)_j + \nabla_j(X_t)_i)$ is the deformation tensor field of X_t. In the case of surface, it was showed that the Leray-Hopf solutions are unique and regular. Any differentiable function $f(t)$ satisfying

$$f^2(t) + 4\int_0^t f^2(s) ds \frac{\|\nabla^2_g\Phi\|^2_{L^2,g}}{\|d\Phi\|^2_{L^2,g}} \le f^2(0)$$

yields a vector field X_t which satisfies (2.4.6). In summary, Chan and Czubak showed that There exist infinitely many real-valued functions $f(t)$ for which $X_t = f(t)(d\Phi)^\sharp$ is a Leray-Hopf solution of the Navier-Stokes equations.

[a] arXiv: math.AP/1205.5322

2.4.3 Metrics on Lie groups

Let G be a Lie group and $(\cdot,\cdot) := (\cdot,\cdot)_e$ the fixed Euclidean metric on T_eG. Using left translation $L_g(x) = gx$, we obtain the metric $(\cdot,\cdot)_g$ on T_gG for every $g \in G$. Since

$$(dL_g)_h = \left(dL_{ghh^{-1}}\right)_h = \left(d\left(L_{gh} \circ L_{h^{-1}}\right)\right)_h = \left(dL_{gh}\right)_e \circ \left(dL_{h^{-1}}\right)_h = \left(dL_{gh}\right)_e \circ \left(dL_h\right)_e^{-1}$$

it follows that L_g is an isometry for each $g \in G$.

Let $\mathfrak{g}$ be the space of all left-invariant vector fields ($dL_g \circ X = X \circ L_g$) on G. Then T_eG can be naturally identified with $\mathfrak{g}$. Note also that $\mathfrak{g}$ is a Lie algebra. If $X \in \mathfrak{g}$, then the integral curve $\gamma_X(t)$ through $e \in G$ is denoted by $\exp(tX)$:

$$\dot{\gamma}_X(t) = X_{\gamma_X(t)}, \quad \gamma_X(0) = e.$$

Letting $t = 0$ in above yields $\dot{\gamma}_X(0) = X_{\gamma_X(0)} = X_e$. By the uniqueness theorem in ODE, we have

$$\exp((t+s)X) = \exp(tX)\exp(sX), \quad t, s \in [0, \infty).$$

The entire flow for X can now be written as

$$F^t_X(x) := x\exp(tX) = L_x \exp(tX) = R_{\exp(tX)}(x). \tag{2.4.7}$$

This flow $F^t_X : G \to G$ don't act by isometries unless the metric is also invariant under right-translations, i.e., the metric is bi-invariant.

The inner automorphism

$$\mathbf{ad}_g : G \longrightarrow G, \quad x \longmapsto gxg^{-1} \tag{2.4.8}$$

is called the **adjoint action** of G on G. If we define

$$\mathbf{ad} : G \longrightarrow \mathrm{Aut}(G), \quad g \longmapsto \mathbf{ad}_g \tag{2.4.9}$$

then ad is a representation of the Lie group G. The differential of this action at $e \in G$ is a linear map

$$\mathbf{Ad}_g := (d\mathbf{ad}_g)_e : \mathfrak{g} \longrightarrow \mathfrak{g} \tag{2.4.10}$$

is called the **adjoint action** of G on $\mathfrak{g}$. This is a Lie algebra isomorphism. For each t, the integral curve $\exp(tX)$ gives a map

$$\exp(t\cdot) : \mathfrak{g} \longrightarrow G, \quad X \longmapsto \exp(tX). \tag{2.4.11}$$

Hence, we have the following commutative diagram:

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\exp(t\cdot)"] \arrow[d, "\mathbf{Ad}_g"'] & G \arrow[d, "\mathbf{ad}_g"] \\ \mathfrak{g} \arrow[r, "\exp(t\cdot)"'] & G \end{tikzcd}$$

Proposition 2.6

A left-invariant metric is bi-invariant if and only if the adjoint action on the Lie algebra is by isometries. ♡

Proof. In case the metric is bi-invariant we know both L_g and $R_{g^{-1}}$ act by isometries. Then also $\mathbf{ad}_g = L_g \circ R_{g^{-1}}$ acts by isometries. The differential is therefore a linear isometry on the Lie algebra.

Conversely, we assume that $\mathbf{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is an isometry. Using

$$(dR_g)_h = (dR_{hg})_e \circ ((dR_h)_e)^{-1}$$

it suffices to show that $(dR_g)_e$ is an isometry. This follows from

$$R_g = L_g \circ \mathrm{ad}_{g^{-1}}, \quad (dR_g)_e = (dL_g)_e \circ \mathrm{Ad}_{g^{-1}}.$$

Hence the metric is bi-invariant. □

Let G be a Lie group. We define the adjoint action $\mathbb{A}\mathbf{d} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ of the Lie algebra on the Lie algebra:

$$\mathbb{A}\mathbf{d} := (d\mathbf{Ad})_e, \tag{2.4.12}$$

where

$$\mathbf{Ad} : G \longrightarrow \mathrm{Aut}(\mathfrak{g}). \tag{2.4.13}$$

We claim that

$$\mathbb{A}\mathbf{d}_X Y = [X, Y]. \tag{2.4.14}$$

If we write $\mathbf{ad}_h = R_{h^{-1}} \circ L_h$, then

$$\mathbf{Ad}_h = d\left(\mathbf{ad}_h\right)_e = d\left(R_{h^{-1}} \circ L_h\right)_e = (dR_{h^{-1}})_h \circ (dL_h)_e.$$

Let F^t be the flow for X. Then $F^t(g) = gF^t(e) = L_g(F^t(e)) = R_{F^t(e)}g$ as both curves go through g at $t = 0$ and have X as tangent everywhere since X is a left-invariant vector field. This also shows that $dF^t = d\left(R_{F^t(e)}\right)$. Calculate

$$\begin{aligned} \left(\mathbb{A}\mathbf{d}_X Y\right)_e &= (d\mathbf{Ad})_e(X)(Y_e) = \left.\frac{d}{dt}\right|_{t=0} \left(\mathbf{Ad}_{F^t(e)}\right)(Y_e) \\ &= \left.\frac{d}{dt}\right|_{t=0} \left(\left(dR_{F^{-t}(e)}\right)_{F^t(e)} \circ \left(dL_{F^t(e)}\right)_e\right)(Y_e) \end{aligned}$$

$$= \left.\frac{d}{dt}\right|_{t=0} \left(dR_{F^{-t}(e)}\right)_{F^t(e)} \left(Y_{F^t(e)}\right) = \left.\frac{d}{dt}\right|_{t=0} \left(dF^{-t}(Y)\right)_{F^t(e)} = \left(\mathscr{L}_X Y\right)_e = [X, Y]_e.$$

By the left-invariance, we derive (2.4.14).

Proposition 2.7

Let G be a Lie group with a bi-invariant metric $(\cdot,\cdot)$. If $X, Y, Z, W \in \mathfrak{g}$, then

$$\begin{aligned} \nabla_Y X &= \frac{1}{2}[Y, X], \\ \mathrm{Rm}(X, Y)Z &= -\frac{1}{4}[[X, Y], Z], \\ \mathrm{Rm}(X, Y, Z, W) &= \frac{1}{4}\left([X, Y], [W, Z]\right). \end{aligned}$$

In particular, the sectional curvature is always nonnegative. ♡

Proof. The bi-invariance of the metric shows that the image $\mathbf{Ad}(G) \subset O(\mathfrak{g})$ lies in the group of orthogonal linear maps on $\mathfrak{g}$. This shows that the image of $\mathbb{A}\mathbf{d}$ lies in the set of skew-adjoint maps:

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_{t=0}(Y, Z) = \frac{d}{dt}\Big|_{t=0}\left(\mathbf{Ad}_{\exp(tX)}(Y), \mathbf{Ad}_{\exp(tX)}(Z)\right) \\ &= (\mathbb{A}\mathbf{d}_X Y, Z) + (Y, \mathbb{A}\mathbf{d}_X Z) = ([X, Y], Z) + (Y, [X, Z]). \end{aligned}$$

For $X, Y, Z \in \mathfrak{g}$, since the metric is bi-invariant, it follows that

$$(Y, Z)_g = (dL_g(Y_e), dL_g(Z_e))_g = (Y_e, Z_e)_e$$

so that (Y, Z) is constant and hence $X(Y, Z) \equiv 0$. Using the Koszul formula, we have

$$\begin{aligned} 2(\nabla_Y X, Z) &= X(Y, Z) + Y(Z, X) - Z(X, Y) - ([X, Y], Z) - ([Y, Z], X) + ([Z, X], Y) \\ &= -([X, Y], Z) - ([Y, Z], X) + ([Z, X], Y) \\ &= -([X, Y], Z) + ([Y, X], Z) + ([X, Y], Z) = ([Y, X], Z). \end{aligned}$$

As for the curvature we then have

$$\begin{aligned} \mathrm{Rm}(X, Y)Z &= \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z \\ &= \frac{1}{4}[X, [Y, Z]] - \frac{1}{4}[Y, [X, Z]] - \frac{1}{2}[[X, Y], Z] \\ &= \frac{1}{4}[X, [Y, Z]] + \frac{1}{4}[Y, [Z, X]] + \frac{1}{4}[Z, [X, Y]] - \frac{1}{4}[[X, Y], Z] \end{aligned}$$

which equals $-\frac{1}{4}[[X, Y], Z]$ because of the Jaboci identity. Finally, by the definition,

$$\begin{aligned} \mathrm{Rm}(X, Y, Z, W) &= (\mathrm{Rm}(X, Y)Z, W) = -\frac{1}{4}([[X, Y], Z], W) \\ &= \frac{1}{4}([Z, [X, Y]], W) = -\frac{1}{4}([Z, W], [X, Y]) = -\frac{1}{4}([X, Y], [Z, W]). \end{aligned}$$

In particular,

$$\begin{aligned} \mathbf{Rm}^{\wedge}(X \wedge Y, Z \wedge W) &= \frac{1}{4}([X, Y], [Z, W]) \\ (\mathbf{Rm}(X \wedge Y), Z \wedge W) &= \frac{1}{4}([X, Y], [Z, W]) \\ \mathrm{Sec}(X, Y) &= \frac{1}{4}\frac{([X, Y], [X, Y])}{(X, X)(Y, Y) - (X, Y)^2} = \frac{1}{4}\frac{|[X, Y]|^2}{|X \wedge Y|^2}. \end{aligned}$$

Thus the Lie groups with bi-invariant metrics always have non-negative sectional curvature and with a little work shows that the Rimann curvature operator is also non-negative. □

Example 2.17. e2.17

Let G be the 2-dimensional Lie group

$$G = \left\{ \begin{pmatrix} \alpha & \beta \\ 0 & 1 \end{pmatrix} : \alpha > 0,\ \beta \in \mathbf{R} \right\}.$$

The Lie algebra of G is

$$\mathfrak{g} = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} : a, b \in \mathbf{R} \right\}.$$

If we define

$$X := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

then

$$[X, Y] := XY - YX = Y.$$

We have the left-invariant metric where X, Y form an orthonormal frame on G. Then use the Koszul formula to compute

$$\nabla_X X = 0, \quad \nabla_Y Y = X, \quad \nabla_X Y = 0, \quad \nabla_Y X = -Y.$$

Hence

$$\mathrm{Rm}(X, Y)Y = \nabla_X \nabla_Y Y - \nabla_Y \nabla_X Y - \nabla_{[X,Y]} = -X,$$

which implies that G has constant sectional curvature -1. We can also compute $\mathbf{Ad}$:

$$\mathbf{Ad}_{\begin{pmatrix} \alpha & \beta \\ 0 & 1 \end{pmatrix}} \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & -a\beta + b\alpha \\ 0 & 0 \end{pmatrix} = aX + (-a\beta + b\alpha)Y.$$

The orthonormal basis X, Y is therefore mapped to the basis

$$\begin{pmatrix} 1 & -\beta \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}.$$

This, however, is not an orthonormal basis unless $\beta = 0$ and $\alpha = 1$. The metric is therefore not bi-invariant. ♠

Example 2.18. (Berger spheres)

The Lie algebra of

$$\begin{aligned} \mathbf{SU}(2) &= \{A \in \mathbf{M}_{2\times 2}(\mathbf{C}) : \det A = 1,\ A^* = A^{-1}\} \\ &= \left\{ \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} : |z|^2 + |w|^2 = 1 \right\} \end{aligned}$$

is

$$\mathfrak{su}(2) = \left\{ \begin{pmatrix} \sqrt{-1}\alpha & \beta + \sqrt{-1}\gamma \\ -\beta + \sqrt{-1}\gamma & -\sqrt{-1}\alpha \end{pmatrix} : \alpha, \beta, \gamma \in \mathbf{R} \right\}$$

and is spanned by

$$X_1 = \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}.$$

We have the left-invariant metric where $\lambda_1^{-1}X_1, \lambda_2^{-1}X_2, \lambda_3^{-1}X_3$ *is an orthonormal frame and* $[X_i, X_{i+1}] = 2X_{i+2}$ *(indices are mod* 3*). The Koszul formula is*

$$2(\nabla_{X_i}X_j, X_k) = ([X_i, X_j], X_k) + ([X_k, X_i], X_j) - ([X_j, X_k], X_i).$$

From this we obtain

$$\nabla_{X_i}X_i = 0.$$

On the other hand,

$$\begin{aligned} \nabla_{X_i}X_{i+1} &= \left(\frac{\lambda_{i+2}^2 + \lambda_{i+1}^2 - \lambda_i^2}{\lambda_{i+2}^2} \right) X_{i+2}, \\ \nabla_{X_{i+1}}X_i &= \left(\frac{-\lambda_{i+2}^2 + \lambda_{i+1}^2 - \lambda_i^2}{\lambda_{i+2}^2} \right) X_{i+2}. \end{aligned}$$

Therefore,

$$\mathrm{Rm}(X_i, X_{i+1})X_{i+2} = 0$$

that all curvatures between three distinct vectors vanish.

An interesting vase of the Berger spheres is $\lambda_1 = \epsilon < 1$, $\lambda_2 = \lambda_3 = 1$. *In this case*

$$\begin{aligned} \nabla_{X_1}X_2 &= (2-\epsilon^2)X_3, \quad \nabla_{X_2}X_1 = -\epsilon^2 X_3, \\ \nabla_{X_2}X_3 &= X_1, \quad \nabla_{X_3}X_2 = -X_1, \\ \nabla_{X_3}X_1 &= \epsilon^2 X_2, \quad \nabla_{X_1}X_3 = (\epsilon^2 - 2)X_2, \\ \mathrm{Rm}(X_1, X_2)X_2 &= \epsilon^2 X_1, \\ \mathrm{Rm}(X_3, X_1)X_1 &= \epsilon^4 X_3, \\ \mathrm{Rm}(X_2, X_3)X_3 &= (4 - 3\epsilon^2)X_2, \\ \mathbf{Rm}(X_1 \wedge X_2) &= \epsilon^2 X_1 \wedge X_2, \\ \mathbf{Rm}(X_3 \wedge X_1) &= \epsilon^2 X_3 \wedge X_1, \\ \mathbf{Rm}(X_2 \wedge X_3) &= (4 - 3\epsilon^2)X_2 \wedge X_3. \end{aligned}$$

Thus all sectional curvatures must lie in the interval $[\epsilon^2, 4 - 3\epsilon^2]$. *Letting* $\epsilon \to 0-$ *we find that all sectional curvatures equal* 1. *As* $\epsilon \to 0+$, *the sectional curvature* $\mathrm{Sec}(X_2, X_3) \to 4$, *which is the curvature of the base space* $\mathbf{S}^2(1/2)$ *in the Hopf fibration.*

The standard orthogonal basis X_1, X_2, X_3 is mapped to

$$\begin{aligned}
\mathbf{Ad}_{\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}} X_1 &= \left(|z|^2 - |w|^2\right) X_1 - 2\mathrm{Re}\left(wz\right) X_2 - 2\mathrm{Im}(wz) X_3, \\
\mathbf{Ad}_{\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}} X_2 &= 2\sqrt{-1}\mathrm{Im}(z\overline{w}) X_1 + \mathrm{Re}\left(w^2 + z^2\right) X_2 \\
&\quad + \mathrm{Im}(w^2 + z^2) X_3, \\
\mathbf{Ad}_{\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}} X_3 &= 2\mathrm{Re}(z\overline{w}) X_1 + \sqrt{-1}\mathrm{Re}(z^2 - w^2) X_2 \\
&\quad + \sqrt{-1}\mathrm{Im}(z^2 - w^2) X_3.
\end{aligned}$$

If the three vectors X_1, X_2, X_3 have the same length, then we see that the adjoint action is by isometries, otherwise it is not. ♠

2.4.4 Riemannian submersions

Let $\varphi : (\overline{\mathcal{M}}^{\overline{m}}, \overline{g}) \to (\mathcal{M}, g)$ be a Riemannian submersion. We say two points $\overline{p} \in \overline{\mathcal{M}}^{\overline{m}}$ and $p \in \mathcal{M}$ are φ-related if $\varphi(\overline{p}) = p$. We also say tao vector fields $\overline{X} \in C^\infty(T\overline{\mathcal{M}}^{\overline{m}})$ and $X \in C^\infty(T\mathcal{M})$ are φ-related if $d\varphi \circ \overline{X} = X \circ \varphi$.

For each point $\overline{p} \in \overline{\mathcal{M}}^{\overline{m}}$, the tangent space $T_{\overline{p}}\overline{\mathcal{M}}^{\overline{m}}$ can be decomposed into

$$T_{\overline{p}}\overline{\mathcal{M}}^{\overline{m}} = T^{\|}_{\overline{p}}\overline{\mathcal{M}}^{\overline{m}} \oplus T^{\perp}_{\overline{p}}\overline{\mathcal{M}}^{\overline{m}} \tag{2.4.15}$$

where

$$T^{\|}_{\overline{p}}\overline{\mathcal{M}}^{\overline{m}} := \mathrm{Ker}\left((d\varphi)_{\overline{p}}\right) \tag{2.4.16}$$

is the **vertical distribution at** $\overline{p}$, and $T^{\perp}_{\overline{p}}\overline{M}$ is the orthogonal complement and called the **horizontal distribution at** $\overline{p}$. Any vector $\overline{v}$ in $\overline{\mathcal{M}}^{\overline{m}}$ can be decomposed into horizontal and vertical parts:

$$v = v^{\|} + v^{\perp}. \tag{2.4.17}$$

Set

$$T^{\|}\overline{\mathcal{M}}^{\overline{m}} := \bigcup_{\overline{p}} T^{\|}_{\overline{p}}\overline{\mathcal{M}}^{\overline{m}}, \quad T^{\perp}\overline{\mathcal{M}}^{\overline{m}} := \bigcup_{\overline{p}} T^{\perp}_{\overline{p}}\overline{\mathcal{M}}^{\overline{m}}. \tag{2.4.18}$$

Hence, any vector field $\overline{X}$ in $\overline{\mathcal{M}}^{\overline{m}}$ can be written as

$$\overline{X} = \overline{X}^{\|} + \overline{X}^{\perp}. \tag{2.4.19}$$

The fact that φ is a Riemannian submersion means that $(d\varphi)_{\overline{p}} : T^{\perp}_{\overline{p}}\overline{\mathcal{M}}^{\overline{m}} \to T_p\mathcal{M}$, where $p = \varphi(\overline{p})$, is an isometry for all $\overline{p} \in \overline{\mathcal{M}}^{\overline{m}}$. Consequently, given a vector field X on $\mathcal{M}$ we can always find a unique horizontal vector field $\overline{X}^{\diamond} \in C^\infty(T^{\perp}\overline{\mathcal{M}}^{\overline{m}})$ on $\overline{\mathcal{M}}^{\overline{m}}$ that is φ-related to X. We say that $\overline{X}^{\diamond}$ is the **basic horizontal lift** of X.

Proposition 2.8

Let $\overline{V}$ be a vector field on $\overline{\mathcal{M}}^{\bar{m}}$ and X, Y, Z vector fields on $\mathcal{M}$ with basic horizontal lifts $\overline{X}^{\diamond}, \overline{Y}^{\diamond}, \overline{Z}^{\diamond}$.

(1) $\left[V^{\|}, \overline{X}^{\diamond}\right]$ *is vertical.*

(2) $\left(\mathscr{L}_{\overline{V}^{\|}}\overline{g}\right)\left(\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right) = \overline{V}^{\|}\overline{g}\left(\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right) = 0.$

(3) $\overline{g}\left(\left[\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right], \overline{V}^{\|}\right) = 2\overline{g}\left((\nabla_{\overline{g}})_{\overline{X}^{\diamond}}\overline{Y}^{\diamond}, \overline{V}^{\|}\right) = -2\overline{g}\left((\nabla_{\overline{g}})_{\overline{V}^{\|}}\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right)$ *and furthermore,* $\overline{g}\left(\left[\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right], \overline{V}^{\|}\right) = 2\overline{g}\left((\nabla_{\overline{g}})_{\overline{Y}^{\diamond}}\overline{V}^{\|}, \overline{X}^{\diamond}\right).$

(4) $(\nabla_{\overline{g}})_{\overline{X}^{\diamond}}\overline{Y}^{\diamond} = \overline{(\nabla_{g})_{X}Y}^{\diamond} + \frac{1}{2}\left(\left[\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right]\right)^{\|}.$ ♡

Proof. (1) Since $\overline{X}^{\diamond}$ is φ-related to X and $\overline{V}^{\perp}$ is φ-related to the zero vector field on M, it follows that

$$d\varphi\left(\left[\overline{V}^{\|}, \overline{X}^{\diamond}\right]\right) = \left[d\varphi\left(\overline{V}^{\|}\right), d\varphi\left(\overline{X}^{\diamond}\right)\right] = [0, X \circ \varphi] = 0.$$

(2) Using (1) yields

$$\begin{aligned}\left(\mathcal{L}_{\overline{V}^{\|}}\overline{g}\right)\left(\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right) &= \overline{V}^{\|}\left(\overline{g}\left(\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right)\right) - \overline{g}\left(\left[\overline{V}^{\|}, \overline{X}^{\diamond}\right], \overline{Y}^{\diamond}\right) \\ &\quad - \overline{g}\left(\overline{X}^{\diamond}, \left[\overline{V}^{\|}, \overline{Y}^{\diamond}\right]\right) \\ &= \overline{V}^{\|}\left(\overline{g}\left(\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right)\right) \;=\; \overline{V}^{\|}\left(g(X, Y)\right)\end{aligned}$$

since φ is a Riemannian submersion. But this implies that the inner product is constant in the direction of the vertical distribution.

(3) Recall the Koszul formula

$$\begin{aligned}2\overline{g}\left((\nabla_{\overline{g}})_{\overline{Y}}\overline{X}, \overline{Z}\right) &= \overline{X}\left(\overline{g}\left(\overline{Y}, \overline{Z}\right)\right) + \overline{Y}\left(\overline{g}\left(\overline{Z}, \overline{X}\right)\right) - \overline{Z}\left(\overline{g}\left(\overline{X}, \overline{Y}\right)\right) \\ &\quad - \overline{g}\left(\left[\overline{X}, \overline{Y}\right], \overline{Z}\right) - \overline{g}\left(\left[\overline{Y}, \overline{Z}\right], \overline{X}\right) + \overline{g}\left(\left[\overline{Z}, \overline{X}\right], \overline{Y}\right).\end{aligned}$$

In particular,

$$\begin{aligned}2\overline{g}\left((\nabla_{\overline{g}})_{\overline{X}^{\diamond}}\overline{Y}^{\diamond}, \overline{V}^{\|}\right) &= \overline{Y}^{\diamond}\left(\overline{g}\left(\overline{X}^{\diamond}, \overline{V}^{\|}\right)\right) + \overline{X}^{\diamond}\left(\overline{g}\left(\overline{V}^{\|}, \overline{Y}^{\diamond}\right)\right) \\ &\quad - \overline{V}^{\|}\left(\overline{g}\left(\overline{Y}^{\diamond}, \overline{X}^{\diamond}\right)\right) - \overline{g}\left(\left[\overline{Y}^{\diamond}, \overline{X}^{\diamond}\right], \overline{V}^{\|}\right) \\ &\quad - \overline{g}\left(\left[\overline{X}^{\diamond}, \overline{V}^{\|}\right], \overline{Y}^{\diamond}\right) + \overline{g}\left(\left[\overline{V}^{\|}, \overline{Y}^{\diamond}\right], \overline{X}^{\diamond}\right) \\ &= \overline{g}\left(\left[\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right], \overline{V}^{\|}\right).\end{aligned}$$

Similarly, we can prove other identities.

(4) From (3) we have showed that for any vector field $\overline{V}$ on $\overline{\mathcal{M}}^{\bar{m}}$

$$\overline{g}\left(\frac{1}{2}\left[\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right] - {}^{\overline{g}}\nabla_{\overline{X}^{\diamond}}\overline{Y}^{\diamond}, \overline{V}^{\|}\right) = 0$$

which implies

$$\left((\nabla_{\overline{g}})_{\overline{X}^{\diamond}}\overline{Y}^{\diamond}\right)^{\|} = \frac{1}{2}\left(\left[\overline{X}^{\diamond}, \overline{Y}^{\diamond}\right]\right)^{\|}.$$

Hence it suffices to show that the horizontal vector field $\overline{(\nabla_{g})_{X}Y}^{\diamond}$ is the horizontal component of $(\nabla_{\overline{g}})_{\overline{X}^{\diamond}}\overline{Y}^{\diamond}$. Using the Koszul formula, φ-relatedness, and the fact that inner products are the same in $\overline{\mathcal{M}}^{\bar{m}}$ and $\mathcal{M}$, shows that

$$2\overline{g}\left((\nabla_{\overline{g}})_{\overline{X}^\diamond}\overline{Y}^\diamond,\overline{Z}^\diamond\right)=2g\left((\nabla_g)_XY,Z\right)=2\overline{g}\left(\overline{(\nabla_g)_XY}^\diamond,\overline{Z}^\diamond\right).$$

Thus, we proved (4). □

The map

$$\begin{aligned} C^\infty(\overline{\mathcal{M}}^{m},T^\perp\overline{\mathcal{M}}^{\bar{m}})\times C^\infty(\overline{\mathcal{M}}^{m},T^\perp\overline{\mathcal{M}}^{\bar{m}}) &\longrightarrow C^\infty(\overline{\mathcal{M}}^{m},T^{\|}\overline{\mathcal{M}}^{\bar{m}}),\\ \left(\overline{X}^\perp,\overline{Y}^\perp\right) &\longmapsto \left[\overline{X}^\perp,\overline{Y}^\perp\right]^{\|} \end{aligned} \tag{2.4.20}$$

measures the extent to which the horizontal distribution in integrable in the sense of Frobenius. It is in fact tensorial as well as skew-symmetric since

$$\left[\overline{X}^\perp,f\overline{Y}^\perp\right]^{\|}=\left(f\left[\overline{X}^\perp,\overline{Y}^\perp\right]+\overline{X}^\perp f\cdot\overline{Y}^\perp\right)^{\|}=f\left[\overline{X}^\perp,\overline{Y}^\perp\right]^{\|}.$$

The map is called the **integrability tensor**.

Theorem 2.6. (O'Neill-Grey)

For vector fields X,Y on $(\mathcal{M},g)$, we have

$$g\left(\mathfrak{Rm}_g(X,Y)Y,X\right)=\overline{g}\left(\mathfrak{Rm}_{\overline{g}}\left(\overline{X}^\diamond,\overline{Y}^\diamond\right)\overline{Y}^\diamond,\overline{X}^\diamond\right)+\frac{3}{4}\left|\left[\overline{X}^\diamond,\overline{Y}^\diamond\right]^{\|}\right|^2_{\overline{g}}. \tag{2.4.21}$$

♡

Proof. By tensors properties we may assume that $[X,Y]=0$. Then in this case

$$0=d\varphi\left[\overline{X}^\diamond,\overline{Y}^\diamond\right]=\left[d\varphi\left(\overline{X}^\diamond\right),d\varphi\left(\overline{Y}^\diamond\right)\right]=[X\circ\varphi,Y\circ\varphi]=0.$$

Hence $\left[\overline{X}^\diamond,\overline{Y}^\diamond\right]$ is vertical. Calculate

$$\begin{aligned} \overline{g}\left(\overline{\mathrm{Rm}}_{\overline{g}}\left(\overline{X}^\diamond,\overline{Y}^\diamond\right)\overline{Y}^\diamond,\overline{X}^\diamond\right) &= \overline{g}\left(\overline{\nabla}_{\overline{X}^\diamond}\overline{\nabla}_{\overline{Y}^\diamond}\overline{Y}^\diamond-\overline{\nabla}_{\overline{Y}^\diamond}\overline{\nabla}_{\overline{X}^\diamond}\overline{Y}^\diamond-\overline{\nabla}_{[\overline{X}^\diamond,\overline{Y}^\diamond]}\overline{Y}^\diamond,\overline{X}^\diamond\right)\\ &= \overline{g}\left(\overline{\nabla}_{\overline{X}^\diamond}\left(\overline{\nabla_YY}^\diamond\right),\overline{X}^\diamond\right)-\overline{g}\left(\overline{\nabla}_{\overline{Y}^\diamond}\left(\overline{\nabla_XY}^\diamond+\frac{1}{2}\left[\overline{X}^\diamond,\overline{Y}^\diamond\right]\right),\overline{X}^\diamond\right)\\ &\quad + frac12\overline{g}\left(\left[\overline{Y}^\diamond,\overline{X}^\diamond\right],\left[\overline{X}^\diamond,\overline{Y}^\diamond\right]\right)\\ &= \overline{g}\left(\overline{\nabla_X\nabla_YY}^\diamond+\frac{1}{2}\left(\left[\overline{X}^\diamond,\overline{\nabla_YY}^\diamond\right]\right)^{\|},\overline{X}^\diamond\right)-\frac{1}{2}\left|\left[\overline{X}^\diamond,\overline{Y}^\diamond\right]\right|^2_{\overline{g}}\\ &\quad -\overline{g}\left(\overline{\nabla_Y\nabla_XY}^\diamond+\frac{1}{2}\left(\left[\overline{Y}^\diamond,\overline{\nabla_XY}^\diamond\right]\right)^{\|}+\frac{1}{2}\overline{\nabla}_{\overline{Y}^\diamond}\left[\overline{X}^\diamond,\overline{Y}^\diamond\right],\overline{X}^\diamond\right)\\ &= g\left(\mathrm{Rm}_g(X,Y)Y,X\right)-\frac{1}{2}\overline{g}\left(\overline{\nabla}_{\overline{Y}^\diamond}\left[\overline{X}^\diamond,\overline{Y}^\diamond\right],\overline{X}^\diamond\right)-\frac{1}{2}\left|\left[\overline{X}^\diamond,\overline{Y}^\diamond\right]\right|^2_{\overline{g}}\\ &= g\left(\mathrm{Rm}_g(X,Y)Y,X\right)-\frac{3}{4}\left|\left[\overline{X}^\diamond,\overline{Y}^\diamond\right]\right|^2_{\overline{g}}. \end{aligned}$$

More generally, one can find formulae for $\overline{\mathrm{Rm}}_{\overline{g}}$ where the variables are various combinations of basic horizontal and vertical fields. □

2.5 Exterior differential calculus and Bochner formulas

Introduction

- *Differential forms*
- *The rough Laplacian acting on tensor fields*
- *Bochner technique*
- *The Bochner technique in general and the Weitzenböck formula*

2.5.1 Differential forms

The **volume form** dV_g of an oriented m-dimensional Riemannian manifold $(\mathcal{M}, g)$ is defined in terms of a positively oriented orthonormal coframe $(\omega^i)_{i=1}^m$ by

$$dV_g = \omega^1 \wedge \cdots \wedge \omega^m.$$

The volume satisfies $m!(dV_g)(e_1, \cdots, e_m) = 1$, where $(e_i)_{i=1}^m$ is the orthonormal frame dual to $(\omega^i)_{i=1}^m$. In a positively oriented local coordinate system $x^1, \cdots, x^m$, we have

$$dV_g = \sqrt{\det(g_{ij})}\, dx^1 \wedge \cdots \wedge dx^m. \tag{2.5.1}$$

Note 2.22

The wedge product of a p-form α and a q-form β is defined by

$$(\alpha \wedge \beta)(X_1, \cdots, X_{p+q}) = \frac{1}{(p+q)!} \sum_{(J,K)} \operatorname{sign}(J,K) \alpha(X_{j_1}, \cdots, X_{j_p}) \beta(X_{k_1}, \cdots, X_{k_q}),$$

where $J := (j_1, \cdots, j_p)$ and $K := (k_1, \cdots, k_q)$ are multi-indices and $\operatorname{sign}(J,K)$ is the sign of the permutation $(1, \cdots, p+q) \mapsto (j_1, \cdots, j_p, k_1, \cdots, k_q)$. ♣

The **exterior derivative** of a p-form β satisfies

$$\begin{aligned}
& (d\beta)(X_0, \cdots, X_p) \\
= & \frac{1}{p+1} \sum_{j=0}^{p} (-1)^j X_j \left(\beta(X_0, \cdots, \widehat{X_j}, \cdots, X_p) \right) \\
& + \frac{1}{p+1} \sum_{0 \le i < j \le p} (-1)^{i+j} \beta \left([X_i, X_j], X_0, \cdots, \widehat{X_i}, \cdots, \widehat{X_j}, \cdots, X_p \right) \qquad (2.5.2) \\
= & \frac{1}{p+1} \sum_{j=0}^{p} (-1)^j \left(\nabla_{X_j} \beta \right) \left(X_0, \cdots, \widehat{X_j}, \cdots, X_p \right).
\end{aligned}$$

In local coordinates, this is

$$(d\beta)_{i_0 i_1 \cdots i_p} = \frac{1}{p+1} \sum_{j=0}^{p} (-1)^j \nabla_{i_j} \beta_{i_0 i_1 \cdots \widehat{i_j} \cdots i_p}, \tag{2.5.3}$$

where $\beta_{i_1 \cdots i_p} := \beta\left(\partial_{i_1}, \cdots, \partial_{i_p}\right)$. If β is a 1-form, then

$$(d\beta)_{ij} = \frac{1}{2}(\nabla_i \beta_j - \nabla_j \beta_i). \tag{2.5.4}$$

If β is a 2-form, then

$$(d\beta)_{ijk} = \frac{1}{3}(\nabla_i \beta_{jk} + \nabla_j \beta_{ki} + \nabla_k \beta_{ij}). \tag{2.5.5}$$

The **divergence** of a p-form α is

$$\mathrm{div}_g(\alpha)_{i_1\cdots i_{p-1}} := g^{jk}\nabla_j \alpha_{ki_1\cdots i_{p-1}} = \nabla^k \alpha_{ki_1\cdots i_{p-1}}. \tag{2.5.6}$$

In particular, if $\alpha = \alpha_i dx^i$ is a 1-form, then

$$\mathrm{div}_g(X) = g^{ij}\nabla_i X_j = \nabla^j X_j. \tag{2.5.7}$$

More generally, we define the divergence of a $(p,0)$-tensor field α with $p \geq 1$ by

$$\mathrm{div}_g(\alpha)_{i_1\cdots i_{p-1}} := g^{jk}\nabla_j \alpha_{ki_1\cdots i_{p-1}}.$$

Given a p-form β and a vector field X, we define the **interior product** by

$$(\iota_X\beta)(Y_1,\cdots,Y_{p-1}) := p\cdot\beta(X,Y_1,\cdots,Y_{p-1}) \tag{2.5.8}$$

for all vector fields $Y_1,\cdots,Y_{p-1}$. Recall the **Cartan formula**

$$\mathscr{L}_X = d\circ\iota_X + \iota_X\circ d. \tag{2.5.9}$$

The inner product on $\mathcal{A}^p(M) := C^\infty(\wedge^p T^*\mathcal{M})$ is defined by

$$\langle\alpha,\beta\rangle_g := p! g^{i_1j_1}\cdots g^{i_pj_p}\alpha_{i_1\cdots i_p}\beta_{j_1\cdots j_p}. \tag{2.5.10}$$

For example, for any positively orthonormal coframe $(\omega^i)_{i=1}^m$,

$$\left\langle\omega^{i_1}\wedge\cdots\wedge\omega^{i_p},\omega^{j_1}\wedge\cdots\wedge\omega^{j_p}\right\rangle_g = \det\left(\delta^{i_kj_\ell}\right).$$

For given p-forms α and β, their L^2**-inner product** is defined by

$$(\alpha,\beta)_{L^2,g} := \int_{\mathcal{M}}\langle\alpha,\beta\rangle_g dV_g. \tag{2.5.11}$$

The **Hodge star operator** $*_g : \mathcal{A}^p(\mathcal{M}) \to \mathcal{A}^{m-p}(\mathcal{M})$, $p = 0,\cdots,m$, is defined

$$\langle\alpha,\beta\rangle_g dV_g = \alpha\wedge *_g\beta \tag{2.5.12}$$

for any $\alpha,\beta\in\mathcal{A}^p(\mathcal{M})$. For instance,

$$*_g\left(\omega^1\wedge\cdots\wedge\omega^p\right) = \omega^{p+1}\wedge\cdots\wedge\omega^m$$

for a positively oriented orthonormal coframe $(\omega^i)_{i=1}^m$.

Note 2.23

Show that acting on $\mathcal{A}^p(\mathcal{M})$, we have $*_g^2 = (-1)^{p(m-p)}$. ♣

The **adjoint operator** δ_g of d acting on a p-form α is defined in terms of d and the Hodge star operator by the formula

$$\delta_g\alpha := (-1)^{mp+m+1} *_g d *_g \alpha. \tag{2.5.13}$$

In terms of covariant derivatives, the adjoint δ_g is given by

$$(\delta_g\alpha)(X_1,\cdots,X_{p-1}) = -p\sum_{i=1}^m(\nabla_{e_i}\alpha)(e_i,X_1,\cdots,X_{p-1}), \tag{2.5.14}$$

where $(e_i)_{i=1}^m$ is an orthonormal frame. That is, $\delta_g = -p\mathrm{div}_g\alpha$, or

$$(\delta_g\alpha)_{i_1\cdots i_{p-1}} = -pg^{jk}\nabla_j\alpha_{ki_1\cdots i_{p-1}}. \tag{2.5.15}$$

Note 2.24

Show that

$$(d\beta, \alpha)_{L^2,g} = (\beta, \delta_g \alpha)_{L^2,g}. \tag{2.5.16}$$

where $\alpha \in \mathcal{A}^p(\mathcal{M})$ *and* $\beta \in \mathcal{A}^{p-1}(\mathcal{M})$. *Directly calculate*

$$\begin{aligned}(\beta, \delta_g \alpha)_{L^2,g} &= \int_{\mathcal{M}} \beta \wedge *_g \delta_g \alpha dV_g = \int_{\mathcal{M}} \beta \wedge *_g (-1)^{mp+m+1} *_g d *_g \alpha dV_g \\ &= (-1)^{mp+m+1} \int_{\mathcal{M}} \beta (-1)^{(m-p+1)(p-1)} d *_g \alpha dV_g = (-1)^{p^2} \int_{\mathcal{M}} \beta \wedge d *_g \alpha dV_g \\ &= (-1)^{p^2}(-1)^p \left(\int_{\mathcal{M}} d\beta \wedge *_g \alpha dV_g - \int_{\mathcal{M}} d(\beta \wedge *_g \alpha) dV_g \right) = \int_{\mathcal{M}} d\beta \wedge *_g \alpha dV_g\end{aligned}$$

since $(\mathcal{M}, g)$ *is closed.* ♣

The **Hodge Laplacian** acting on differential p-forms is defined by

$$\Delta_{H,g} := -(d\delta_g + \delta_g d) \tag{2.5.17}$$

where we have adopted the opposite of the usual sign convention. Note that $\Delta_{H,g}$ is a self-adjoint operator. Acting on functions, it is the same as the usual Laplcian operator defined in (2.5.18).

2.5.2 The rough Laplacian acting on tensor fields

Let Δ_g denote the **Laplacian**, also called the **Laplace-Beltrami operator**, **acting on functions**, which is globally defined as the divergence of the gradient and is given in local coordinates by

$$\Delta_g := \mathrm{div}_g \nabla_g = g^{ij} \nabla_i \nabla_j = g^{ij} \left(\partial_i \partial_j - \Gamma^k_{ij} \partial_k \right). \tag{2.5.18}$$

If $\{e_i\}_{i=1}^m$ is an orthonormal frame, then

$$\Delta_g f = \sum_{i=1}^m e_i(e_i f) - (\nabla_{e_i} e_i) f. \tag{2.5.19}$$

Note 2.25

If $\mathcal{M}$ *is the Euclidean space* $\mathbf{R}^m$ *with the standard metric* g_{stand}, *then* $\Delta = \Delta_{\text{stand}} = \sum_{1 \le i \le m} \partial_i \partial_i$ *and the heat equation is* $(\partial_t - \Delta) u = 0$. ♣

Note 2.26

(1) For any function f *and any vector fields* X *and* Y *we define the* **Hessian** $\mathrm{Hess}_g f = \nabla_g \nabla_g f = \nabla_g^2 f$ *as follows:*

$$\nabla_g^2 f(X, Y) := X(Yf) - (\nabla_X Y) f.$$

Then $\Delta_g f = \mathrm{tr}_g \left(\nabla_g^2 f \right) = \Delta_{H,g} f$.

(2) If $|g| := \det(g_{ij})$, *then*

$$\Delta_g f = \frac{1}{\sqrt{|g|}} \partial_i \left(\sqrt{|g|} g^{ij} \partial_j f \right). \tag{2.5.20}$$

♣

More generally, the **rough Laplacian operator acting on tensor fields** is given by

$$\Delta_g := \mathrm{div}_g \nabla_g = \mathrm{tr}_g \nabla_g^2 = g^{ij}\nabla_i\nabla_j = \nabla^j\nabla_j. \tag{2.5.21}$$

More explicitly, given an (r,s)-tensor field β, $\nabla_g^2\beta$ is an $(r+2,s)$-tensor field, which we contract to

$$\Delta_g\beta(X_1,\cdots,X_r) = \sum_{i=1}^m \nabla_g^2\beta(e_i,e_i,X_1,\cdots,X_r) \tag{2.5.22}$$

for all vector fields $X_1,\cdots,X_r$.

2.5.3 Bochner technique

The **Bochner technique** was invented by Bochne. Yano fuether refined the Bochner technique, but it seems to be Lichnerowicz who really put things into gear, when around 1960 he presented his formulae for the Laplacian on forms and spinors. After this work, Bergerm Meyer, Gallot, Gromov-Lawson, Witten, and many others have made significant contributions to this tremendously important subject.

Prior to Bochners work Weitzenböck also developed a formula very similar to the Bochner formula.

Lemma 2.3. (Commutator of Δ_g and ∇_g on functions)

For any function f,

$$\Delta_g\nabla_i f = \nabla_i\Delta_g f + R_{ij}\nabla^j f. \tag{2.5.23}$$

♡

Proof. By definition, one has

$$\begin{aligned}\Delta_g\nabla_i f &= g^{k\ell}\nabla_k\nabla_\ell\nabla_i f = g^{k\ell}\nabla_k\nabla_i\nabla_\ell f = g^{k\ell}\left(\nabla_i\nabla_k\nabla_\ell f - R^p_{ki\ell}\nabla_p f\right)\\ &= \nabla_i\Delta_g f - g^{kl}R_{ki\ell p}\nabla^p f = \nabla_i\Delta_g f + g^{k\ell}R_{kip\ell}\nabla^p f = \nabla_i\Delta_g f + R_{ip}\nabla^p f.\end{aligned}$$

Here our convenience is that $g^{pq}R^p_{ki\ell} = R_{ki\ell q}$. □

Lemma 2.4. (Bochner formula for $|\nabla_g f|_g^2$)

Suppose that $(\mathcal{M},g)$ is a compact oriented Riemannian manifold. Show that for any function f,

$$\Delta_g|\nabla_g f|_g^2 = 2|\nabla_g^2 f|_g^2 + 2R_{ij}\nabla^i f\nabla^j f + 2\nabla^i f\nabla_i(\Delta_g f). \tag{2.5.24}$$

Conclude from this that if $\mathrm{Rc}_g \geq 0$, $\Delta_g f \equiv 0$ and $|\nabla_g f|_g \equiv 0$, then $\nabla_g^2 f \equiv 0$, and $\mathrm{Rc}_g(\nabla_g f, \nabla_g f) = 0$. (This lemma is interesting only in noncompact case.)

♡

Proof. Calculate

$$\begin{aligned}\Delta_g|\nabla_g f|_g^2 &= g^{ij}\nabla_i\nabla_j(g^{k\ell}\nabla_k f\nabla_\ell f) = g^{ij}g^{k\ell}\nabla_i(\nabla_j\nabla_k f\cdot\nabla_\ell f + \nabla_k f\cdot\nabla_j\nabla_\ell f)\\ &= g^{ij}g^{k\ell}(2\nabla_i\nabla_j\nabla_k f\cdot\nabla_\ell f + 2\nabla_j\nabla_k f\cdot\nabla_i\nabla_\ell f)\\ &= 2|\nabla_g^2 f|_g^2 + 2g^{ij}g^{k\ell}\nabla_i\nabla_k\nabla_j f\cdot\nabla_\ell f;\end{aligned}$$

since

$$\begin{aligned} 2g^{ij}g^{k\ell}\nabla_i\nabla_k\nabla_j f\cdot\nabla_\ell f &= 2g^{ij}g^{\ell}(\nabla_k\nabla_i\nabla_j f - R^p_{ikj}\nabla_p f)\cdot\nabla_\ell f \\ &= 2\langle\nabla_g\Delta_g f,\nabla_g f\rangle_g - 2g^{ij}R_{ikjp}\nabla^p f\cdot\nabla^k f \end{aligned}$$

we have

$$\Delta_g|\nabla_g f|_g^2 = 2|\nabla_g^2 f|_g^2 + 2\langle\nabla_g\Delta_g f,\nabla_g f\rangle_g + 2R_{kp}\nabla^p f\cdot\nabla^k f$$

which implies (2.5.24). If $\Delta_g f\equiv 0$, then

$$\int_{\mathcal{M}}\Delta_g|\nabla_g f|_g^2 dV_g = \int_{\mathcal{M}} 2|\nabla_g^2 f|_g^2 dV_g + \int_{\mathcal{M}} 2\mathrm{Rc}_g(\nabla_g f,\nabla_g f)\,dV_g.$$

By another hypothesis that $\mathrm{Rc}_g\geq 0$, we conclude that

$$\int_{\mathcal{M}}|\nabla_g^2 f|_g^2 dV_g = 0 = \int_{\mathcal{M}}\mathrm{Rc}_g(\nabla_g f,\nabla_g f)\,dV_g.$$

Both integrands are nonnegative so that they must vanish identically. □

Lemma 2.5

One has

$$\begin{aligned} \Delta_g|\nabla_g f|_g &= \frac{1}{|\nabla_g f|_g}\left(\langle\nabla_g f,\nabla_g(\Delta_g f)\rangle_g + \mathrm{Rc}_g(\nabla_g f,\nabla_g f) + |\nabla_g^2 f|_g^2\right) \\ &\quad - \frac{1}{|\nabla_g f|_g}\left|\left\langle\nabla_g^2 f,\frac{\nabla_g f}{|\nabla_g f|_g}\right\rangle_g\right|^2 \qquad (2.5.25) \end{aligned}$$

wherever $|\nabla_g f|_g\neq 0$, *and conclude that if* $\mathrm{Rc}_g\geq 0$, *then*

$$\Delta_g|\nabla_g f|_g \geq \left\langle\frac{\nabla_g f}{|\nabla_g f|_g},\nabla_g(\Delta_g f)\right\rangle_g.$$

In particular, if $\Delta_g f = 0$, *then*

$$\Delta_g|\nabla_g f|_g\geq 0. \qquad (2.5.26)$$

Since if $\mathcal{M}$ *is compact and oriented, the assumption* $\Delta_g f = 0$ *implies* $\nabla_g f\equiv 0$ *so (2.5.26) is automatically valid. Hence, this lemma is also interesting in the noncompact case.* ♡

Proof. Calculate

$$\Delta_g|\nabla_g f|_g = g^{ij}\nabla_i\nabla_j\left(g^{k\ell}\nabla_k f\cdot\nabla_\ell f\right)^{1/2} = g^{ij}g^{k\ell}\nabla_i\left(\frac{\nabla_j\nabla_k f\cdot\nabla_\ell f}{(g^{pq}\nabla_p f\cdot\nabla_q f)^{1/2}}\right)$$

$$= \frac{g^{ij}g^{k\ell}(\nabla_i\nabla_j\nabla_k f\cdot\nabla_\ell f + \nabla_j\nabla_k f\cdot\nabla_i\nabla_\ell f)}{|^g\nabla f|_g} - \frac{g^{ij}g^{kl}\nabla_j\nabla_k f\cdot\nabla_\ell f}{|\nabla_g f|_g^3}\left(g^{pq}\nabla_i\nabla_p f\cdot\nabla_q f\right).$$

Since

$$g^{ij}g^{k\ell}\nabla_i\nabla_j\nabla_k f\cdot\nabla_\ell f = \langle\nabla_g\Delta_g f,\nabla_g f\rangle_g + \mathrm{Rc}_g(\nabla_g f,\nabla_g f)$$

it follows that

$$\begin{aligned} \Delta_g|\nabla_g f|_g &= \frac{1}{|\nabla_g f|_g}\left(\langle\Delta_g f,\nabla_g f\rangle_g + \mathrm{Rc}_g(\nabla_g f,\nabla_g f) + |\nabla_g^2 f|_g^2\right) \\ &\quad - \frac{1}{|\nabla_g f|_g^3}g^{ij}\langle\nabla_j\nabla_g f,\nabla_g f\rangle_g\langle\nabla_i\nabla_g f,\nabla_g f\rangle_g \end{aligned}$$

that is (2.5.25). □

Note 2.27

In (2.5.25) $\langle\nabla^2_g f, \frac{\nabla_g f}{|\nabla_g f|_g}\rangle_g$ is the 1-form given by

$$\left(\left\langle\nabla^2_g f, \frac{\nabla_g f}{|\nabla_g f|_g}\right\rangle_g\right)_i := \left\langle\nabla_i\nabla_g f, \frac{\nabla_g f}{|\nabla_g f|_g}\right\rangle_g = g^{jk}\nabla_i\nabla_j f\frac{\nabla_k f}{|\nabla_g f|_g} = \nabla_i|\nabla_g f|_g.$$

Hence

$$\left|\left\langle\nabla^2_g f, \frac{\nabla_g f}{|\nabla_g f|_g}\right\rangle_g\right|^2 = |\nabla_g|\nabla_g f|_g|^2_g.$$

Related to this is the fact that if T_1 and T_2 are (r,s)-tensor fields, then

$$\nabla_g\langle T_1, T_2\rangle_g = \langle\nabla_g T_1, T_2\rangle_g + \langle T_1, \nabla_g T_2\rangle_g, \tag{2.5.27}$$

where

$$\langle\nabla_g T_1, T_2\rangle_g := (\nabla_i T_1)^{j_1\cdots j_r}_{k_1\cdots k_s}(T_2)^{k_1\cdots k_s}_{j_1\cdots j_r} \tag{2.5.28}$$

and similarly for $\langle T_1, \nabla_g T_2\rangle_g$. ♣

Note 2.28

If $\partial_t g(t) = -2\mathrm{Rc}_{g(t)}$, then

$$\begin{aligned}\left(\Delta_{g(t)} - \partial_t\right)|\nabla_{g(t)} f|^2_{g(t)} &= 2\left|\nabla^2_{g(t)} f\right|^2_{g(t)} \\ &\quad + 2\left\langle\nabla_{g(t)} f, \nabla_{g(t)}\left(\left(\Delta_{g(t)} - \partial_t\right) f\right)\right\rangle_{g(t)}.\end{aligned} \tag{2.5.29}$$

Calculate

$$\begin{aligned}\partial_t|\nabla_{g(t)} f|^2_{g(t)} &= \partial_t\left(g^{ij}\nabla_i f\nabla_j f\right) = -\partial_t g_{ij}\cdot\nabla^i f\cdot\nabla^j f + 2g^{ij}\nabla_i(\partial_t f)\cdot\nabla_j f \\ &= 2\mathrm{Rc}_{g(t)}\left(\nabla_{g(t)} f, \nabla_{g(t)} f\right) + 2\left\langle\nabla_{g(t)} f, \nabla_{g(t)}(\partial_t f)\right\rangle_{g(t)}.\end{aligned}$$

Combining it with (2.5.24), we complete the proof. ♣

Note 2.29

Here $|\nabla^2_g f|^2_g = g^{ik}g^{j\ell}\nabla_i\nabla_j f\cdot\nabla_k\nabla_\ell f$. Similarly we denote for a p-form α

$$|\alpha|^2 := g^{i_1j_1}\cdots g^{i_pj_p}\alpha_{i_1\cdots i_p}\alpha_{j_1\cdots j_p}. \tag{2.5.30}$$

♣

Note 2.30

Show that for any tensor field A

$$\nabla_g\Delta_g A - \Delta_g\nabla_g A = \mathrm{Rm}_g * \nabla_g A + (\nabla_g\mathrm{Rc}_g) * A. \tag{2.5.31}$$

Here, given tensor fields A and B, $A * B$ denotes some linear combination of contractions of $A\otimes B$. Calculate

$$\nabla_i\Delta_g A = g^{pq}\nabla_i\nabla_p\nabla_q A = g^{pq}\left(\nabla_p\nabla_i\nabla_q A + \mathrm{Rm}_g * \nabla_g A\right)$$

$$= g^{pq}\nabla_p(\nabla_q\nabla_i A + \mathrm{Rm}_g * A) + \mathrm{Rm}_g * \nabla_g A = \Delta_g\nabla_i A + \nabla_g\mathrm{Rm}_g * A + \mathrm{Rm}_g * \nabla_g A;$$

the formula (2.5.31) follows by $\nabla_g\mathrm{Rm}_g = \nabla_g\mathrm{Rc}_g$. ♣

Lemma 2.6. (Bochner, 1946)

Show that if X is a 1-form, then

$$\Delta_g X_i - g^{jk}R_{ik}X_j = \Delta_{H,g}X_i. \tag{2.5.32}$$

In particular, if the Ricci curvature of a closed manifold is positive, then there are no nontrivial harmonic 1-forms. By the Hodge theorem, this implies the first Betti number $b_1(\mathcal{M})$ is zero. ♡

Proof. By the definition, one has

$$\Delta_{H,g}X_i = -\left(d\left(\delta_g X\right)\right)_i - \left(\delta_g(dX)\right)_i$$

where

$$\begin{aligned}
\delta_g X &= -g^{jk}\nabla_j X_k, \\
-d\left(\delta_g X\right) &= d\left(g^{jk}\nabla_j X_k\right) = \nabla_\ell\left(g^{jk}\nabla_j X_k\right)dx^\ell = \left(g^{jk}\nabla_\ell\nabla_j X_k\right)dx^\ell, \\
(\delta_g dX)_i &= -2g^{jk}\nabla_j(dX)_{ki} = -g^{jk}\left(\nabla_j\nabla_k X_i - \nabla_j\nabla_i X_k\right).
\end{aligned}$$

Therefore,

$$\Delta_{H,g}X_i = g^{jk}\nabla_i\nabla_j X_k + g^{jk}\left(\nabla_j\nabla_k X_i - \nabla_j\nabla_i X_k\right) = \Delta_g X_i + g^{jk}R^\ell_{ijk}X_\ell$$

implying (2.5.32). If there is a nontrivial harmonic 1-form X, that is, $\Delta_{H,g}X = 0$, then $\Delta_g X_i \cdot X^i = R_{ij}X^iX^j = \mathrm{Rc}_g(X,X)$ and

$$-\int_{\mathcal{M}}|\nabla_g X|_g^2\,dV_g = \int_{\mathcal{M}}\mathrm{Rc}_g(X,X)\,dV_g \geq 0;$$

consequently, $\mathrm{Rc}_g(X,X) \equiv 0 = |\nabla_g X|_g$. We must have that $X_p = 0$ if the Ricci tensor is positive on $T_p\mathcal{M}$. But then $X \equiv 0$ since X is parallel. We get a contradiction. □

Note 2.31

In Lemma 2.6, we have proved that if $(\mathcal{M}, g)$ is compact, oriented, and has $\mathrm{Rc}_g \geq 0$, then every harmonic 1-form is parallel. By the Hodge theorem, $b_1(\mathcal{M}) = \dim\mathcal{H}^1(\mathcal{M})$. Now, all harmonic 1-forms are parallel, so the linear map

$$\mathcal{H}^1(\mathcal{M}) \longrightarrow T_p^*\mathcal{M}, \quad \omega \longmapsto \omega_p$$

is injective. In particular, $\dim\mathcal{H}^1(\mathcal{M}) \leq m$. Furthermore, we can show that $b_1(\mathcal{M}) = m$ if and only if (M, g) is a flat torus. ♣

If β is a 2-form, then

$$(\Delta_{H,g}\beta)_{ij} = \Delta_g\beta_{ij} + 2R_{ik\ell j}\beta^{k\ell} - R_{ik}\beta^k{}_j - R_{jk}\beta_i{}^k. \tag{2.5.33}$$

Let α be a p-form. In local coordinates we may write the Hodge Laplacian as

$$\begin{aligned}
(\Delta_{H,g}\alpha)_{i_1\cdots i_r} &= (-1)^{j+1}g^{kl}\nabla_{i_j}\nabla_k\alpha_{li_1\cdots i_{j-1}i_{j+1}\cdots i_r} + g^{k\ell}\nabla_k\nabla_\ell\alpha_{i_1\cdots i_r} \\
&\quad + (-1)^j g^{k\ell}\nabla_k\nabla_{i_j}\alpha_{\ell i_1\cdots i_{j-1}i_{j+1}\cdots i_r} \\
&= (\Delta_g\alpha)_{i_1\cdots i_r} + (-1)^j g^{k\ell}\left(\nabla_k\nabla_{i_j} - \nabla_{i_j}\nabla_k\right)\alpha_{\ell i_1\cdots i_{j-1}i_{j+1}\cdots i_r}.
\end{aligned} \tag{2.5.34}$$

Using the notation $*$, (2.5.34) can be written as

$$\Delta_{H,g}\alpha = \Delta_g\alpha + \mathrm{Rm}_g * \alpha. \tag{2.5.35}$$

2.5.4 The Bochner technique in general and the Weitzenböck formula

Let $(\mathcal{E}, h) \to (\mathcal{M}, g)$ be a vector bundle over an oriented closed Riemannian manifold. Let $\Gamma(\mathcal{M}, \mathcal{E})$ denote the sections $s : \mathcal{M} \to \mathcal{E}$. The connection on $\mathcal{E}$ is a map

$$\nabla^{\mathcal{E}} : \Gamma(\mathcal{M}, \mathcal{E}) \longrightarrow \Gamma(\mathcal{M}, \mathrm{Hom}(T\mathcal{M}, \mathcal{E})), \quad s \longmapsto \nabla^{\mathcal{E}} s. \tag{2.5.36}$$

We assume the connection is linear in s, tensorial in X, and compatible with the metric h:

$$X\langle s_1, s_2\rangle_h = \left\langle \nabla^{\mathcal{E}}_X s_1, s_2\right\rangle_h + \left\langle s_1, \nabla^{\mathcal{E}}_X s_2\right\rangle_h. \tag{2.5.37}$$

Since $\mathrm{Hom}(T\mathcal{M}, \mathcal{E}) \cong T^*\mathcal{M} \otimes \mathcal{E}$, the vector bundle has the induced metric $g^{T^*\mathcal{M}} \otimes h$, where $g^{T^*\mathcal{M}}$ is the bundle metric on the cotangent bundle $T^*\mathcal{M}$; usually we write it as $g \otimes h$, if there is no confusion. Using the pointwise inner product structures on $\Gamma(\mathcal{E}), \Gamma(T\mathcal{M})$, and integration, we get global L^2-inner product structures on $\Gamma(\mathcal{E})$ and $\Gamma(\mathrm{Hom}(T\mathcal{M}, \mathcal{E}))$:

$$(s_1, s_2)_{L^2(\Gamma(\mathcal{E}))} := \int_{\mathcal{M}} \langle s_1, s_2\rangle_h dV_g, \tag{2.5.38}$$

$$(S_1, S_2)_{L^2(\Gamma(\mathrm{Hom}(T\mathcal{M},\mathcal{E})))} := \int_M \langle S_1, S_2\rangle_{g\otimes h} dV_g. \tag{2.5.39}$$

From (2.5.36), we define the adjoint connection:

$$\nabla^{\mathcal{E},*} : \Gamma(\mathrm{Hom}(T\mathcal{M}, \mathcal{E})) \longrightarrow \Gamma(\mathcal{E}) \tag{2.5.40}$$

defined by

$$\int_{\mathcal{M}} \left\langle \nabla^{\mathcal{E},*} S, s\right\rangle_h dV_g = \int_{\mathcal{M}} \left\langle S, \nabla^{\mathcal{E}} s\right\rangle_{g\otimes h} dV_g. \tag{2.5.41}$$

Note 2.32

We use the notation $\nabla^{\mathcal{E},*}$ to denote the adjoint connection of $\nabla^{\mathcal{E}}$. The induced connection on the dual bundle $\mathcal{E}^\vee$ is denoted by $\nabla^{\mathcal{E}^\vee}$ or $\nabla^{\mathcal{E},\vee}$. Here we use the operation "$\vee$" to make the dual bundle; in this situation, the dual bundle of the tangent bundle $T\mathcal{M}$ is written as $T^\vee\mathcal{M}$, but we adopt the classical notation that is $T^*\mathcal{M}$. The metric g induces the Levi-Civita connection ∇_g on $\mathcal{M}$, which can be viewed as a bundle connection $\nabla^{T\mathcal{M}}$ on $T\mathcal{M}$. In the bundle setting, we use $\nabla^{T\mathcal{M}}$ to denote the induced connection rather than the classical notation ∇_g. ♣

The **connection Laplacian** of a section is defined as

$$\Delta^{\mathcal{E}}_{H,g} s := -\nabla^{\mathcal{E},*}\nabla^{\mathcal{E}} s. \tag{2.5.42}$$

There is a different way of defining the connection Laplacian. Consider the second covariant derivative

$$\Gamma(\mathcal{M}, \mathcal{E}) \xrightarrow{\nabla^{\mathcal{E}}} \Gamma(\mathrm{Hom}(\mathcal{M}, T\mathcal{M}, \mathcal{E})) \xrightarrow{\nabla^{T^*\mathcal{M}\otimes\mathcal{E}}} \Gamma(\mathcal{M}, T^*\mathcal{M} \otimes T^*\mathcal{M} \otimes \mathcal{E}).$$

We write it as

$$\left(\nabla^{\mathcal{E}}\right)^2_{X,Y} s := \nabla^{T^*\mathcal{M}\otimes\mathcal{E}} \circ \nabla^{\mathcal{E}}(s)(X,Y) - \nabla^{\mathcal{E}}_{\nabla^{T\mathcal{M}}_X Y} s. \tag{2.5.43}$$

Take the trace $\sum_{1\le i\le m} \left(\nabla^{\mathcal{E}}\right)^2_{\partial_i,\partial_i} s$ with respect to the orthonormal basis of $T\mathcal{M}$. This is easily seen to be invariantly defined. We shall use the notation

$$\Delta^{\mathcal{E}}_{L,g} s := \operatorname{tr}_g \left(\left(\nabla^{\mathcal{E}}\right)^2 s\right) := \sum_{1\le i\le m} \left(\nabla^{\mathcal{E}}\right)^2_{\partial_i,\partial_j} s. \tag{2.5.44}$$

Proposition 2.9

Let $(\mathcal{M},g)$ be an oriented closed Riemannian manifold, and $\mathcal{E}\to\mathcal{M}$ a vector bundle with an inner product h and compatible connection $\nabla^{\mathcal{E}}$, then

$$\Delta^{\mathcal{E}}_{H,g} s = \Delta^{\mathcal{E}}_{L,g} s \tag{2.5.45}$$

for all sections s of $\mathcal{E}$. ♡

Note 2.33

According to **Proposition 2.9***, we write*

$$\Delta^{\mathcal{E}}_g s := \Delta^{\mathcal{E}}_{H,g} s = \Delta^{\mathcal{E}}_{L,g} s. \tag{2.5.46}$$

♣

Proof. Let s_1, s_2 be two sections of $\mathcal{E}$ and $(e_i)_{i=1}^m$ be an orthonormal frame on $\mathcal{M}$. Calculate

$$\begin{aligned}\left(\Delta^{\mathcal{E}}_{H,g} s_1, s_2\right)_{L^2(\Gamma(E))} &= \int_{\mathcal{E}} \left\langle \Delta^{\mathcal{E}}_{H,g} s_1, s_2\right\rangle_h dV_g \\ = -\int_{\mathcal{M}} \left\langle \nabla^{\mathcal{E}} s_1, \nabla^{\mathcal{E}} s_2\right\rangle_{g\otimes h} dV_g &= -\sum_{1\le i\le m}\int_{\mathcal{M}} \left\langle \nabla^{\mathcal{E}}_{e_i} s_1, \nabla^{\mathcal{E}}_{e_i} s_2\right\rangle_h dV_g.\end{aligned}$$

The right hand side is equal to

$$\begin{aligned}\left(\Delta^{\mathcal{E}}_{L,g} s_1, s_2\right)_{L^2(\Gamma(E))} &= \int_{\mathcal{M}} \left\langle \Delta^{\mathcal{E}}_{L,g} s_1, s_2\right\rangle_h dV_g \\ &= \sum_{1\le i\le m}\int_{\mathcal{M}} \left\langle \nabla^{\mathcal{E}}_{e_i}\nabla^{\mathcal{E}}_{e_i} s_1 - \nabla^{\mathcal{E}}_{\nabla^{T\mathcal{M}}_{e_i} e_i} s_1, s_2\right\rangle_h dV_g \\ = -\sum_{1\le i\le m}\int_{\mathcal{M}} \left\langle \nabla^{\mathcal{E}}_{e_i} s_1, \nabla^{\mathcal{E}}_{e_i} s_2\right\rangle_h dV_g &+ \sum_{i=1}^m \int_{\mathcal{M}} \nabla^{T\mathcal{M}}_{e_i} \left\langle \nabla^{\mathcal{E}}_{e_i} s_1, s_2\right\rangle_h dV_g \\ &- \sum_{1\le i\le m}\int_{\mathcal{M}} \left\langle \nabla^{\mathcal{E}}_{\nabla^{T\mathcal{M}}_{e_i} e_i} s_1, s_2\right\rangle_h dV_g \\ &= \left(\Delta^{\mathcal{E}}_{H,g} s_1, s_2\right)_{L^2(\Gamma(E))} + \int_{\mathcal{M}} \operatorname{div}_g X\, dV_g\end{aligned}$$

where X is defined by $g(X,Y) := \left\langle \nabla^{\mathcal{E}}_Y s_1, s_2\right\rangle_h$. Setting $Y=\partial_i$, we have $X^i = g(X,\partial_i) = \left\langle \nabla^{\mathcal{E}}_{\partial_i} s_1, s_2\right\rangle_h$; hence $\operatorname{div}_g X = \nabla_i X^i = \nabla^{T\mathcal{M}}_{\partial_i}\left\langle \nabla^{\mathcal{E}}_{\partial_i} s_1, s_2\right\rangle_h$ which verifies the above identity. Since $\mathcal{M}$ is closed, it follows that $\left(\Delta^{\mathcal{E}}_{L,g} s_1, s_2\right)_{L^2(\Gamma(E))} = \left(\Delta^{\mathcal{E}}_{H,g} s_1, s_2\right)_{L^2(\Gamma(E))}$. Thus, we must have $\Delta^{\mathcal{E}}_{L,g} s_1 = \Delta^{\mathcal{E}}_{H,g} s_1$ for all sections s_1. □

By our notation, we have $\Delta^{T\mathcal{M}}_{L,g} = \Delta_{L,g}$ and $\Delta^{T\mathcal{M}}_{H,g} = \Delta_{H,g}$. For sections $s_1, s_2 \in \Gamma(\mathcal{E})$, we have

$$\begin{aligned}
\Delta_{L,g}\left(\frac{1}{2}\langle s_1, s_2\rangle_h\right) &= \sum_{1\le i\le m}\left(\nabla^{T\mathcal{M}}\right)^2_{e_i,e_i}\left(\frac{1}{2}\langle s_1, s_2\rangle_h\right)\\
&= \sum_{1\le i\le m}\left(\nabla^{T\mathcal{M}}_{e_i}\nabla^{T\mathcal{M}}_{e_i} - \nabla^{T\mathcal{M}}_{\nabla^{T\mathcal{M}}_{e}e_i}\right)\left(\frac{1}{2}\langle s_1, s_2\rangle_h\right)\\
&= \sum_{1\le i\le m}\frac{1}{2}\left[\nabla^{T\mathcal{M}}_{e_i}\left(\left\langle\nabla^{\mathcal{E}}_{e_i}s_1, s_2\right\rangle_h + \left\langle s_1, \nabla^{\mathcal{E}}_{e_i}s_2\right\rangle_h\right) - \left\langle\nabla^{\mathcal{E}}_{\nabla^{T\mathcal{M}}_{e_i}e_i}s_1, s_2\right\rangle_h - \left\langle s_1, \nabla^{\mathcal{E}}_{\nabla^{T\mathcal{M}}_{e_i}e_i}s_2\right\rangle_h\right]\\
&= \sum_{1\le i\le m}\left(\left\langle\nabla^{\mathcal{E}}_{e_i}\nabla^{\mathcal{E}}_{e_i}s_1, s_2\right\rangle_h + 2\left\langle\nabla^{\mathcal{E}}_{e_i}s_1, \nabla^{\mathcal{E}}_{e_2}s_2\right\rangle_h + \left\langle s_1, \nabla^{\mathcal{E}}_{e_i}\nabla^{\mathcal{E}}_{e_i}s_2\right\rangle_h\right)\\
&\quad - \frac{1}{2}\sum_{1\le i\le m}\left(\left\langle\nabla^{\mathcal{E}}_{\nabla^{T\mathcal{M}}_{e_i}e_i}s_1, s_2\right\rangle_h + \left\langle s_1, \nabla^{\mathcal{E}}_{\nabla^{T\mathcal{M}}_{e_i}e_i}s_2\right\rangle_h\right)\\
&= \left\langle\nabla^{\mathcal{E}}s_1, \nabla^{\mathcal{E}}s_2\right\rangle_{g\otimes h} + \frac{1}{2}\left\langle\Delta^{\mathcal{E}}_{L,g}s_1, s_2\right\rangle_h + \frac{1}{2}\left\langle s_1, \Delta^{\mathcal{E}}_{L,g}s_2\right\rangle_h\\
&= \left\langle\nabla^{\mathcal{E}}s_1, \nabla^{\mathcal{E}}s_2\right\rangle_{g\otimes h} + \frac{1}{2}\left\langle\Delta^{\mathcal{E}}_{H,g}s_1, s_2\right\rangle_h + \frac{1}{2}\left\langle s_1, \Delta^{\mathcal{E}}_{H,g}s_2\right\rangle_h.
\end{aligned}$$

In particular,

$$\Delta_{L,g}\left(\frac{1}{2}|s|^2_h\right) = \left|\nabla^{\mathcal{E}}s\right|^2_{g\otimes h} + \left\langle\Delta^{\mathcal{E}}_{H,g}s, s\right\rangle_h. \tag{2.5.47}$$

For the connection $\nabla^{\mathcal{E}}$ we define the curvature

$$R^{\mathcal{E}}: \Gamma(\mathcal{M}, T\mathcal{M})\otimes\Gamma(\mathcal{M}, T\mathcal{M})\otimes\Gamma(\mathcal{M}, \mathcal{E}) \longrightarrow \Gamma(\mathcal{M}, \mathcal{E})$$

by

$$R^{\mathcal{E}}(X,Y)s := \left(\nabla^{\mathcal{E}}\right)^2_{X,Y}s - \left(\nabla^{\mathcal{E}}\right)^2_{Y,X}s = \nabla^{\mathcal{E}}_X\nabla^{\mathcal{E}}_Y s - \nabla^{\mathcal{E}}_Y\nabla^{\mathcal{E}}_X s - \nabla^{\mathcal{E}}_{[X,Y]}s. \tag{2.5.48}$$

Then the operator

$$\Delta^{\mathcal{E}}_{H,g} + \mathscr{C}\left(R^{\mathcal{E}}\right): \Gamma(\mathcal{M}, \mathcal{E}) \longrightarrow \Gamma(\mathcal{M}, \mathcal{E})$$

is a map on $\Gamma(\mathcal{M}, \mathcal{E})$, where $\mathscr{C}\left(R^{\mathcal{E}}\right)$ is a trace of the curvature $R^{\mathcal{E}}$. We say a first-order operator $\mathbf{D}: \Gamma(\mathcal{M}, \mathcal{E}) \to \Gamma(\mathcal{M}, \mathcal{E})$ is the **Dirac-type operator** of the vector bundle $\mathcal{E} \to \mathcal{M}$ if

$$\mathbf{D}^2 = \Delta^{\mathcal{E}}_{H,g} + \mathscr{C}\left(R^{\mathcal{E}}\right). \tag{2.5.49}$$

Such a formulae are called **Weitzenböck formulae**. Here, we use the word "are" because there are lots of way to contract the curvature $R^{\mathcal{E}}$.

(1) **Riemannian geometry:** Let $(\mathcal{M}, g)$ be an m-dimensional Riemannian manifold. We take $\mathcal{E} := \oplus_{p=0}^m \wedge^p T^*\mathcal{M}$. In this case the Dirac-type operator $\mathbf{D}: \Gamma(\mathcal{M}, \mathcal{E}) \to \Gamma(\mathcal{M}, \mathcal{E})$ is $d + \delta_g$. Moreover, $-\mathbf{D}^2 = \Delta_{H,g}$, and the Weitzenböck formula now becomes

$$\Delta_{H,g} = \Delta^{\mathcal{E}}_{H,g} + \frac{1}{2}\mathrm{Rm}_g(e_i, e_j)\omega^i\omega^j. \tag{2.5.50}$$

Here we denote by $(\omega^i)_{i=1}^m$ the dual coframe of an orthonormal frame $(e_i)_{i=1}^m$ of $\mathcal{M}$. This was certainly known to both Bochner and Yano. However, in this case $\Delta^{\mathcal{E}}_{H,g} = \Delta^{\mathcal{E}}_{L,g} = \Delta_{L,g}$, so that (2.5.50) is exact the formula (2.5.32).

(2) **Spin geometry:** Let $(\mathcal{M}, g)$ be an m-dimentional spin manifold. It induces the spinor bundle $\mathcal{S}_{\mathcal{M}}$. In this case the Dirac-type operator is just the **Dirac operator** $\mathcal{D} : \Gamma(\mathcal{M}, \mathcal{S}_{\mathcal{M}}) \to \Gamma(\mathcal{M}, \mathcal{S}_{\mathcal{M}})$, and the Weitzenböck formula reads

$$-\mathcal{D}^2 = \Delta^{\mathcal{S}_{\mathcal{M}}}_{H,g} - \frac{1}{4} R_g. \tag{2.5.51}$$

The formula was discovered and used by Lichnerowicz, as well as Singer, to show that the $\widehat{A}$-genus vanishes for spin manifolds with positive scalar curvature. Using some generalization of this formula, Gromov-Lawson showed that any metric on a torus with nonnegative scalar curvature is in fact flat. Much of Witten's work, e.g., the positive mass conjecture, uses these ideas. Also, the work of Seiberg-Witten on 4-manifold geometry, is related to these ideas.

In the following we shall prove the Weitzenböck formula (2.5.50) for p-forms. As before, let $(\mathcal{M}, g)$ be an m-dimensional Riemannian manifold, and let $\mathcal{A}^*(\mathcal{M}) = \oplus_{p=0}^m \mathcal{A}^p(\mathcal{M})$ denotes the space of all forms on $\mathcal{M}$. On this space we can define a product structure that is different rom the wedge product. This product is called **Clifford multiplication**, and for $\omega \in \mathcal{A}^1(\mathcal{M})$ and $\theta \in \mathcal{A}^p(\mathcal{M})$, then

$$\omega \circledast \theta \quad := \quad \omega \wedge \theta - \iota_{\omega^\sharp}\theta, \tag{2.5.52}$$

$$\theta \circledast \omega \quad := \quad \theta \wedge \omega + (-1)^p \iota_{\omega^\sharp}\theta. \tag{2.5.53}$$

If ω_1, ω_2 are 1-forms, then by (2.5.52) we have

$$\omega_1 \circledast \omega_2 = \omega_1 \wedge \omega_2 - \iota_{\omega_1^\sharp}\omega_2.$$

On the other hand, using (2.5.53) yields

$$\omega_1 \circledast \omega_2 = \omega_1 \wedge \omega_2 - \iota_{\omega_2^\sharp}\omega_1.$$

To verify the well-defined operation $\circledast$, we shall check $\iota_{\omega_1^\sharp}\omega_2 = \iota_{\omega_2^\sharp}\omega_1$. By definition we have

$$\iota_{\omega_1^\sharp}\omega_2 = \omega_2\left(\omega_1^\sharp\right) = (\omega_2)_k\, dx^k\left(g^{ij}(\omega_1)_i\,\partial_j\right) = g^{ij}(\omega_1)_i(\omega_2)_j = \omega_1\left(\omega_2^\sharp\right) = \iota_{\omega_2^\sharp}\omega_1.$$

By declaring the product to be bilinear and associate, we can use these properties to define the product between any two forms. For example

$$(\omega_1 \wedge \omega_2) \circledast \theta := \omega_1 \circledast (\omega_2 \circledast \theta) + \iota_{\omega_1^\sharp}\omega_2 \cdot \theta.$$

Note that even when θ is a p-form, the Clifford product with a 1-form gives a mixed form.

For 1-form ω we have

$$\omega \circledast \omega = -|\omega|_g^2 \le 0. \tag{2.5.54}$$

In general, for 1-forms ω_1 and ω_2 we obtain

$$\omega_1 \circledast \omega_2 + \omega_2 \circledast \omega_1 = -2g(\omega_1, \omega_2). \tag{2.5.55}$$

If $(\omega_i)_{i=1}^m$ is an orthonormal coframe of $\mathcal{M}$, then

$$\omega_i \circledast \omega_j = -\omega_j \circledast \omega_i, \quad \omega_i \circledast \omega_j = \omega_i \wedge \omega_j, \quad i \neq j. \tag{2.5.56}$$

Hence, we see that Clifford multiplication not only depends on the inner product, wedge product, and interior product, but actually reproduces there three items.

Proposition 2.10

For $\theta_1, \theta_2 \in \mathcal{A}^*(\mathcal{M})$, $\omega \in \mathcal{A}^1(\mathcal{M})$, and $\psi \in \mathcal{A}^2(\mathcal{M})$, we have

$$g(\omega \divideontimes \theta_1, \theta_2) = -g(\theta_1, \omega_1 \divideontimes \theta_2), \tag{2.5.57}$$

$$g([\psi, \theta_1]_{\divideontimes}, \theta_2) = -g(\theta_1, [\psi, \theta_2]_{\divideontimes}) \tag{2.5.58}$$

where $[\theta_1, \theta_2]_{\divideontimes} := \theta_1 \divideontimes \theta_2 - \theta_2 \divideontimes \theta_1$. ♡

Proof. The proof is based on the definition of the Clifford multiplication and the fact that the two maps

$$\begin{aligned} \mathcal{A}^p(\mathcal{M}) &\longrightarrow \mathcal{A}^{p+1}(\mathcal{M}), \quad \theta \longmapsto \epsilon_\omega \theta := \omega \wedge \theta, \\ \mathcal{A}^{p+1}(\mathcal{M}) &\longrightarrow \mathcal{A}^p(\mathcal{M}), \quad \theta \longmapsto \iota_{\omega^\sharp}\theta \end{aligned}$$

are ajoint to each other. We write

$$\theta_1 = \sum_{0 \le p \le m} \theta_1^{(p)}, \quad \theta_2 = \sum_{0 \le p \le m} \theta_2^{(p)}.$$

Then

$$\omega \divideontimes \theta_1 = -\iota_{\omega^\sharp}\theta_1^{(1)} + \sum_{0 \le p \le m-2} \left(\epsilon_\omega \theta_1^{(p)} - \iota_{\omega^\sharp}\theta_1^{(p+2)}\right) + \epsilon_\omega \theta_1^{(m-1)}.$$

The left hand side of (2.5.57) becomes

$$\begin{aligned} g(\omega \divideontimes \theta_1, \theta_2) &= g\left(-\iota_{\omega^\sharp}\theta_1^{(1)}, \theta_2^{(0)}\right) + \sum_{0 \le p \le m-2} g\left(\epsilon_\omega \theta_1^{(p)} - \iota_{\omega^\sharp}\theta_1^{(p+2)}, \theta_2^{(p+1)}\right) \\ + g\left(\epsilon_\omega \theta_1^{(n-1)}, \theta_2^{(n)}\right) &= g\left(\theta_1^{(1)}, -\epsilon_\omega \theta_2^{(0)}\right) + \sum_{0 \le p \le m-2} g\left(\theta_1^{(p)}, \iota_{\omega^\sharp}\theta_2^{(p+1)}\right) \\ &- \sum_{0 \le p \le m-2} g\left(\theta_1^{(p+2)}, \epsilon_\omega \theta_2^{(p+1)}\right) + g\left(\theta_1^{(n-1)}, \iota_{\omega^\sharp}\theta_2^{(n)}\right). \end{aligned}$$

Rearranging the terms yields

$$-\left[g\left(\iota_{\omega^\sharp}\theta_2^{(1)}, \theta_1^{(0)}\right) + \sum_{0 \le p \le m-2} g\left(\epsilon_\omega \theta_2^{(p)} \iota_{\omega^\sharp}\theta_2^{(p+2)}, \theta_1^{(p+1)}\right) + g\left(\epsilon_\omega \theta_2^{(m-1)}, \theta_1^{(m)}\right)\right]$$

which equals the right hand side of (2.5.57). To prove the second formula (2.5.58), it suffices to prove

$$g([\psi, \theta]_{\divideontimes}, \theta) = 0 \tag{2.5.59}$$

for any form $\theta \in \mathcal{A}^*(\mathcal{M})$. Since ψ is a 2-form, we shall verify (2.5.59) for a special case that $\psi = \omega_1 \wedge \omega_2$ where $\omega_1, \omega_2 \in \mathcal{A}^1(\mathcal{M})$. By definition, one has

$$\begin{aligned} g([\omega_1 \wedge \omega_2, \theta]_{\divideontimes}, \theta) &= g((\omega_1 \wedge \omega_2) \divideontimes \theta, \theta) - g(\theta \divideontimes (\omega_1 \wedge \omega_2), \theta) \\ &= g\left(\left(\omega_1 \divideontimes \omega_2 + \iota_{\omega_1^\sharp}\omega_2\right) \divideontimes \theta, \theta\right) - g\left(\theta \divideontimes \left(\omega_1 \divideontimes \omega_2 + \iota_{\omega_1^\sharp}\omega_2\right), \theta\right) \\ &= g(\omega_1 \divideontimes (\omega_2 \divideontimes \theta), \theta) - g((\theta \divideontimes \omega_1) \divideontimes \omega_2, \theta) \end{aligned}$$

so that (2.5.59) holds for $\psi = \omega_1 \wedge \omega_2$ if and only if

$$g\left(\omega_1 \divideontimes \left(\omega_2 \divideontimes \theta\right), \theta\right) = g\left(\left(\theta \divideontimes \omega_1\right) \divideontimes \omega_2, \theta\right). \tag{2.5.60}$$

For convenience, we set

$$\psi := \omega_2 \divideontimes \theta = \sum_{0\le p\le m} \psi^{(p)}, \quad \varphi := \theta \divideontimes \omega_1 = \sum_{0\le p\le m} \varphi^{(p)}, \quad \theta = \sum_{0\le p\le m} \theta^{(p)}.$$

Since

$$\omega_2 \divideontimes \theta = -\iota_{\omega_2^\sharp}\theta^{(1)} + \sum_{0\le p\le m}\left(\epsilon_{\omega_2}\theta^{(p)} - \iota_{\omega_2^\sharp}\theta^{(p+2)}\right) + \epsilon_{\omega_2}\theta^{(m-1)}, \tag{2.5.61}$$

it follows that

$$\begin{aligned}
\psi^{(0)} &= -\iota_{\omega_2^\sharp}\theta^{(1)},\\
\psi^{(p)} &= \epsilon_{\omega_2}\theta^{(p-1)} - \iota_{\omega_2^\sharp}\theta^{(p+1)}, \quad p = 1, \cdots, m-1,\\
\psi^{(n)} &= \epsilon_{\omega_2}\theta^{(m-1)}.
\end{aligned}$$

Using formula (2.5.61) again implies

$$\begin{aligned}
\omega_1 \divideontimes (\omega_2 \divideontimes \theta) &= \omega_1 \divideontimes \psi = -\iota_{\omega_1^\sharp}\psi^{(1)} + \sum_{0\le p\le m-2}\left(\epsilon_{\omega_1}\psi^{(p)} - \iota_{\omega_1^\sharp}\psi^{(p+2)}\right) + \epsilon_{\omega_1}\psi^{(m-1)}\\
&= -\iota_{\omega_1^\sharp}\left(\epsilon_{\omega_2}\theta^{(0)} - \iota_{\omega_2^\sharp}\theta^{(2)}\right) + \epsilon_{\omega_1}\left(-\iota_{\omega_2^\sharp}\theta^{(1)}\right) - \iota_{\omega_1^\sharp}\left(\epsilon_{\omega_2}\theta^{(1)} - \iota_{\omega_2^\sharp}\theta^{(3)}\right)\\
&\quad + \sum_{1\le p\le m-3}\left[\epsilon_{\omega_1}\left(\epsilon_{\omega_2}\theta^{(p-1)} - \iota_{\omega_2^\sharp}\theta^{(p+1)}\right) - \iota_{\omega_1^\sharp}\left(\epsilon_{\omega_2}\theta^{(p+1)} - \iota_{\omega_2^\sharp}\theta^{(p+3)}\right)\right]\\
&\quad + \epsilon_{\omega_1}\left(\epsilon_{\omega_2}\theta^{(n-3)} - \iota_{\omega_2^\sharp}\theta^{(m-1)}\right) - \iota_{\omega_1^\sharp}\left(\epsilon_{\omega_2}\theta^{(m-1)}\right) + \epsilon_{\omega_1}\left(\epsilon_{\omega_2}\theta^{(m-2)} - \iota_{\omega_2^\sharp}\theta^{(m)}\right).
\end{aligned}$$

On the other hand, it is easy to see that

$$\theta \divideontimes \omega_1 = -\iota_{\omega_1^\sharp}\theta^{(1)} + \sum_{0\le p\le m-2}(-1)^p\left(\epsilon_{\omega_1}\theta^{(p)} + \iota_{\omega_1^\sharp}\theta^{(p+2)}\right) + (-1)^{m-1}\epsilon_{\omega_1}\theta^{(m-1)}. \tag{2.5.62}$$

Thus

$$\begin{aligned}
\varphi^{(0)} &= -\iota_{\omega_1^\sharp}\theta^{(1)},\\
\varphi^{(p)} &= (-1)^{p-1}\left(\epsilon_{\omega_1}\theta^{(p-1)} + \iota_{\omega_1^\sharp}\theta^{(p+1)}\right), \quad p = 1, \cdots, m-1,\\
\varphi^{(m)} &= (-1)^{m-1}\epsilon_{\omega_1}\theta^{(m-1)}.
\end{aligned}$$

Again, from (2.5.62) we deduce that

$$\begin{aligned}
(\theta \divideontimes \omega_1) \divideontimes \omega_2 &= \varphi \divideontimes \omega_2\\
= -\iota_{\omega_2^\sharp}\left(\epsilon_{\omega_1}\theta^{(0)} + \iota_{\omega_1^\sharp}\theta^{(2)}\right) &+ \epsilon_{\omega_2}\left(-\iota_{\omega_1^\sharp}\theta^{(1)}\right) - \iota_{\omega_2^\sharp}\left(\epsilon_{\omega_1}\theta^{(1)} + \iota_{\omega_1^\sharp}\theta^{(3)}\right)\\
- \sum_{1\le p\le m-3}&\left[\left(\epsilon_{\omega_2}\left(\epsilon_{\omega_1}\theta^{(p-1)}\right) + \iota_{\omega_1^\sharp}\theta^{(p+1)}\right) + \iota_{\omega_2^\sharp}\left(\epsilon_{\omega_1}\theta^{(p+1)} + \iota_{\omega_1^\sharp}\theta^{(p+3)}\right)\right]\\
- \epsilon_{\omega_2}\left(\epsilon_{\omega_1}\theta^{(n-3)} + \iota_{\omega_1^\sharp}\theta^{(n-1)}\right) &- \iota_{\omega_2^\sharp}\left(\epsilon_{\omega_1}\theta^{(n-1)}\right) - \epsilon_{\omega_2}\left(\epsilon_{\omega_1}\theta^{(n-2)} + \iota_{\omega_1^\sharp}\theta^{(n)}\right).
\end{aligned}$$

For degree 0, we have

$$g\left(\omega_1 \divideontimes \left(\omega_2 \divideontimes \theta\right), \theta\right)_0 = g\left(-\iota_{\omega_1^\sharp}\left(\epsilon_{\omega_2}\theta^{(0)} - \iota_{\omega_2^\sharp}\theta^{(2)}\right), \theta^{(0)}\right)$$

$$= -\, g\left(\epsilon_{\omega_2}\theta^{(0)} - \iota_{\omega_2^\sharp}\theta^{(2)}, \epsilon_{\omega_1}\theta^{(0)}\right) \;=\; -g\left(\epsilon_{\omega_1}\theta^{(0)}, \epsilon_{\omega_0}\theta^{(2)}\right) + g\left(\theta^{(2)}, \epsilon_{\omega_2}\epsilon_{\omega_1}\theta^{(0)}\right)$$

$$= -\, g\left(\iota_{\omega_2^\sharp}\epsilon_{\omega_1}\theta^{(0)}, \theta^{(0)}\right) - g\left(\theta^{(2)}, \epsilon_{\omega_1}\epsilon_{\omega_2}\theta^{(0)}\right)$$

$$= g\left(-\iota_{\omega_2^\sharp}\left(\epsilon_{\omega_1}\theta^{(0)} + \iota_{\omega_1^\sharp}\theta^{(2)}\right), \theta^{(0)}\right) \;=\; g\left((\theta \divideontimes \omega_1) \divideontimes \omega_2, \theta\right)_0.$$

Similarly, we can verify for other degrees. □

Proposition 2.11

For $\theta_1, \theta_2 \in \mathcal{A}^*(\mathcal{M})$ and vector fields X, Y we have the derivation properties:

$$\nabla_X(\theta_1 \divideontimes \theta_2) = \nabla_X\theta_1 \divideontimes \theta_2 + \theta_1 \divideontimes \nabla_X\theta_2, \tag{2.5.63}$$

and

$$\mathrm{Rm}_g(X,Y)(\theta_1 \divideontimes \theta_2) = (\mathrm{Rm}_g(X,Y)\theta_1) \divideontimes \theta_2 + \theta_1 \divideontimes (\mathrm{Rm}_g(X,Y)\theta_2). \tag{2.5.64}$$

♡

Proof. In case $\theta_1 = \theta_2 = \omega$ is a 1-form, we have

$$\nabla_X(\omega \divideontimes \theta) = -\nabla_X|\omega|_g^2 = -2g(\nabla_X\omega, \omega) = \nabla_X\omega \divideontimes \omega + \omega \divideontimes \nabla_X\omega.$$

In case $\theta_1 = \omega$ is a 1-form and $\theta_2 = \theta$ is a general form, we have

$$\begin{aligned}\nabla_X(\omega \divideontimes \theta) &= \nabla_X(\epsilon_\omega\theta - \iota_{\omega^\sharp}\theta) = \epsilon_{\nabla_X\omega}\theta + \epsilon_\omega\nabla_X\theta - \iota_{(\nabla_X\omega)^\sharp}\theta - \iota_{\omega^\sharp}\nabla_X\theta \\ &= \nabla_X\omega \divideontimes \theta + \omega \divideontimes \nabla_X\omega.\end{aligned}$$

The same formula holds for any forms. The second formula follows from the first formula. □

Let $\{e_i\}^{1\le i\le m}$ and $\{\omega^i\}_{1\le i\le m}$ denote the orthonormal frame and coframe, respectively. The **Dirac operator** on forms is given by

$$\mathbf{D}: \mathcal{A}^*(\mathcal{M}) \longrightarrow \mathcal{A}^*(\mathcal{M}), \quad \theta \longmapsto \sum_{1\le i\le m} \omega^i \divideontimes \nabla_{e_i}\theta. \tag{2.5.65}$$

The definition is independent of the choice of the frame fields and coframe fields.

Proposition 2.12

Given a frame $\{e_i\}_{1\le i\le m}$ and its dual coframe $\{\omega^i\}_{1\le i\le m}$, we have

$$d\theta = \epsilon_{\omega^i}\nabla_{e_i}\theta, \tag{2.5.66}$$

$$d_g^*\theta = -\iota_{(\omega^i)^\sharp}\nabla_{e_i}\theta, \tag{2.5.67}$$

$$\mathbf{D} = d + \delta_g. \tag{2.5.68}$$

♡

Proof. (2.5.66) and (2.5.67) hold for functions and 1-forms, so that it also holds for any forms. (2.5.68) is a direct consequence of the previous two formulas. □

The square of the Dirac operator satisfies

$$-\mathbf{D}^2 = -(d + \delta_g)^2 = -(d\delta_g + \delta_g d) = \Delta_{H,g}. \tag{2.5.69}$$

Corollary 2.1

If X be a vector field, then

$$\Delta_{H,g} X^\flat = \Delta_{H,g}^{\wedge^* T^* \mathcal{M}} X^\flat - \mathrm{Rc}_g(X)^\flat. \tag{2.5.70}$$

Proof. Calculate

$$\begin{aligned}
\Delta_{H,g} X^\flat(e_i) &= -d\left(\delta_g X^\flat\right)(e_i) - \delta_g\left(dX^\flat\right)(e_i) \\
&= -\nabla_{e_i}\delta_g X^\flat + 2\sum_{1\le j\le m}\left(\nabla_{e_j} dX^\flat\right)(e_i, e_j) \\
&= \nabla_{e_i}\sum_{1\le j\le m}\nabla_{e_j} X^\flat(e_j) + 2\sum_{1\le j\le m}\left(\nabla_{e_j} dX^\flat\right)(e_i, e_j) \\
&= \sum_{1\le j\le m}\left(\nabla^2_{e_i,e_j} X^\flat\right)(e_j) + \sum_{1\le j\le m}\nabla_{e_j}\left[\left(\nabla_{e_i} X^\flat\right)(e_j) - \left(\nabla_{e_j} X^\flat\right)(e_i)\right] \\
&= \sum_{1\le j\le m}\left(\nabla^2_{e_i,e_j} X^\flat - \nabla^2_{e_j,e_i} X^\flat\right)(e_j) - \sum_{1\le j\le m}\left(\nabla^2_{e_j,e_j} X^\flat\right)(e_i) \\
&= \Delta_{H,g}^{\wedge^* T^* \mathcal{M}} X^\flat + \sum_{1\le j\le m}\left(\mathrm{Rm}_g(e_i, e_j) X^\flat\right)(e_j).
\end{aligned}$$

By definition, one has

$$\begin{aligned}
\sum_{1\le j\le m}\left(\mathrm{Rm}_g(e_i,e_j)X^\flat\right)(e_j) &= \sum_{1\le j\le m}\left(\mathrm{Rm}_g(e_i,e_j)(X^\flat(e_j)) - X^\flat\left(\mathrm{Rm}_g(e_i,e_j)e_j\right)\right) \\
&= -\sum_{1\le j\le m} X^\flat\left(\mathrm{Rm}_g(e_i,e_j)e_j\right) = -X^\flat\left(\mathrm{Rc}_g(e_i)\right) \\
&= -g\left(X, \mathrm{Rc}_g(e_i)\right) = -g\left(\mathrm{Rc}_g(X), e_i\right) = -\mathrm{Rc}(X)^\flat(e_i).
\end{aligned}$$

Extending to any linear combinations of e_i, we prove the corollary. □

Proposition 2.13

For any form θ, we have

$$\begin{aligned}
\mathbf{D}^2\theta &= \omega^i * \omega^j * \nabla^2_{e_i,e_j}\theta \qquad (2.5.71)\\
&= \left(\nabla^2_{e_i,e_j}\theta\right) * \omega^j * \omega^i. \qquad (2.5.72)
\end{aligned}$$

Proof. Recall that

$$\nabla^2_{e_i,e_j} = \nabla_{e_i}\nabla_{e_j} - \nabla_{\nabla_{e_i} e_j}$$

is tensorial in both e_i and e_j, and thus the two expressions on the right hand side are invariantly defined. We may assume that $\nabla e_i = 0$ at a point and consequently $\nabla\omega^i = 0$. Calculate at this point,

$$\begin{aligned}
\mathbf{D}^2\theta &= \omega^i * \nabla_{e_i}(\mathbf{D}\theta) = \omega^i * \nabla_{e_i}\left(\omega^j * \nabla_{e_j}\theta\right) = \omega^i * \left(\nabla_{e_i}\omega^j * \nabla_{e_j}\theta + \omega^j * \nabla_{e_i}\nabla_{e_j}\theta\right) \\
&= \omega^i * \omega^j * \nabla_{e_i}\nabla_{e_j}\theta + \omega^i * \nabla_{e_i}\omega^j * \nabla_{e_j}\theta
\end{aligned}$$

$$= \omega^i \divideontimes \omega^j \divideontimes \nabla_{e_i}\nabla_{e_j}\theta - \omega^i \divideontimes \omega^j \divideontimes \nabla_{\nabla_{e_i}e_j}\theta \quad = \quad \omega^i \divideontimes \omega^j \divideontimes \nabla^2_{e_i,e_j}\theta.$$

Similarly, we can prove the second formula. □

We can now establish the relevant Weitzenböck formula.

Theorem 2.7

For any form θ, we have

$$\begin{aligned} -\mathbf{D}^2\theta &= \Delta_{H,g}^{\wedge^* T^*\mathcal{M}}\theta - \frac{1}{2}\omega^i \divideontimes \omega^j \divideontimes \mathrm{Rm}_g(e_i,e_j)\theta, && (2.5.73)\\ &= \Delta_{H,g}^{\wedge^* T^*\mathcal{M}}\theta - \frac{1}{2}\mathrm{Rm}_g(e_i,e_j)\theta \divideontimes \omega^j \divideontimes \omega^i. && (2.5.74)\end{aligned}$$

♡

Proof. Using **Proposition 2.13**, it suffices to check

$$\Delta_{H,g}^{\wedge^* T^*\mathcal{M}}\theta - \frac{1}{2}\omega^i \divideontimes \omega^j \divideontimes \mathrm{Rm}_g(e_i,e_j)\theta = -\omega^i \divideontimes \omega^j \divideontimes \nabla^2_{e_i,e_j}\theta. \tag{2.5.75}$$

The left hand side of (2.5.75) equals

$$\begin{aligned}&\sum_{1\le i\le m}\nabla^2_{e_i,e_i}\theta - \sum_{i\neq j}\omega^i \divideontimes \omega^j \divideontimes \nabla^2_{e_i,e_j}\theta \quad = \quad \Delta_{H,g}^{\wedge^* T^*\mathcal{M}}\theta - \sum_{i<j}\omega^i \divideontimes \omega^j \divideontimes \left(\nabla^2_{e_i,e_j}\theta\nabla^2_{e_j,e_i}\theta\right)\\ &= \Delta_{H,g}^{\wedge^* T^*\mathcal{M}}\theta - \sum_{i<j}\omega^i \divideontimes \omega^j \divideontimes \mathrm{Rm}_g(e_i,e_j)\theta \quad = \quad \Delta_{H,g}^{\wedge^* T^*\mathcal{M}}\theta - \frac{1}{2}\omega^i \divideontimes \omega^j \divideontimes \mathrm{Rm}_g(e_i,e_j)\theta\end{aligned}$$

where we use the fact that $\omega^i \divideontimes \omega^i = -1$ and $\omega^i \divideontimes \omega^j = -\omega^j \divideontimes \omega^i$. □

2.6 Integration and Hodge theory

Introduction

- *Integration by parts*
- *De Rham theorem and Hodge decomposition theorem*
- *Killing vector fields*
- *Affine vector fields*

2.6.1 Integration by parts

Let $(\mathcal{M}, g)$ be an oriented n-dimensional Riemannian manifold with boundary $\partial\mathcal{M}$. The orientation on $\mathcal{M}$ defines an orientation on $\partial\mathcal{M}$. Locally, on the boundary, choose a positively oriented frame field $\{e_i\}_{1\le i\le m}$ such that $e_1 = \nu$ is the unit outward normal. Then the frame field $\{e_i\}_{2\le i\le m}$ is positively oriented on $\partial\mathcal{M}$.

Theorem 2.8. (Stokes's theorem)

If α is an $(m-1)$-form on a compact oriented m-dimensional manifold $\mathcal{M}$ with (possibly empty) boundary $\partial\mathcal{M}$, then

$$\int_{\mathcal{M}} d\alpha = \int_{\partial\mathcal{M}}\alpha. \tag{2.6.1}$$

♡

Let $\{\omega^i\}_{1\le i\le m}$ denote the orthonormal coframe field dual to $\{e_i\}_{1\le i\le m}$. The volume form of $\mathcal{M}$ is $dV_g=\omega^1\wedge\cdots\wedge\omega^m$ and the volume form of ∂M is $dV_g':=\omega^2\wedge\cdots\wedge\omega^m$. According to (2.5.8), we have

$$(\iota_\nu dV_g)\left(e_2,\cdots,e_m\right)=m\cdot dV_g(e_1,e_2,\cdots,e_m)=\frac{m}{m!}=\frac{1}{(m-1)!}$$

and hence

$$dV_g'=\iota_V\left(dV_g\right). \tag{2.6.2}$$

Theorem 2.9. (Divergence theorem)

Let $(\mathcal{M},g)$ be a compact oriented m-dimensional Riemannian manifold. If X is a vector field, then

$$\int_{\mathcal{M}}\mathrm{div}_g(X)dV_g=\int_{\partial\mathcal{M}}\langle X,\nu\rangle_g dV_g', \tag{2.6.3}$$

where $\mathrm{div}_g(X)=\nabla_iX^i$.

Proof. Define the $(m-1)$-form α by $\alpha:=\iota_X\left(dV_g\right)$. Using (2.5.9) and $d^2=0$, we compute

$$d\alpha=d\circ\iota_X\left(dV_g\right)=\mathscr{L}_X\left(dV_g\right).$$

In an orthonormal frame $e_1,\cdots,e_m$, one has

$$\mathscr{L}_X\left(dV_g\right)\left(e_1,\cdots,e_m\right)=\sum_{1\le i\le m}dV_g\left(e_1,\cdots,\nabla_{e_i}X,\cdots,e_m\right)=\mathrm{div}_g(X)dV_g(e_1,\cdots,e_m)$$

so that $d\alpha=\mathrm{div}_g(X)dV_g$. Now **Theorem 2.8** implies

$$\int_{\mathcal{M}}\mathrm{div}_g(X)dV_g=\int_{\mathcal{M}}d\alpha=\int_{\partial\mathcal{M}}\alpha=\int_{\partial\mathcal{M}}\iota_X\left(dV_g\right).$$

Since

$$\left(\iota_X\left(dV_g\right)\right)\left(e_2,\cdots,e_m\right)=m\,dV_g(X,e_2,\cdots,e_m)=m\langle X,\nu\rangle_g dV_g(e_1,\cdots,e_m)=\frac{\langle X,\nu\rangle_g}{(m-1)!}$$

so that $\iota_X(dV_g)=\langle X,\nu\rangle_g dV_g'$. □

Lemma 2.7

Let $(\mathcal{M},g)$ be an m-dimensional compact oriented Riemannian manifold.

(1) If $\mathcal{M}$ is closed, then

$$\int_{\mathcal{M}}\Delta_g u\,dV_g=0.$$

(2) We have the following Green formula

$$\int_{\mathcal{M}}\left(u\Delta_g v-v\Delta_g u\right)dV_g=\int_{\partial\mathcal{M}}\left(u\frac{\partial v}{\partial\nu}-v\frac{\partial u}{\partial\nu}\right)dV_g'.$$

In particular, on a closed manifold

$$\int_{\mathcal{M}}u\Delta_g v\,dV_g=\int_{\mathcal{M}}v\Delta_g u\,dV_g.$$

(3) If f is a function and α is a 1-form, then

$$\int_{\mathcal{M}}f\mathrm{div}_g\alpha\,dV_g=-\int_M\langle\nabla_g f,\alpha\rangle_g dV_g+\int_{\partial\mathcal{M}}f\langle\alpha,\nu\rangle_g dV_g'.$$

Proof. Since $\Delta_g u = \mathrm{div}_g(\nabla_g u)$, (1) follows. For (2), we note $\mathrm{div}_g(u\nabla_g v - v\nabla_g u) = u\Delta_g v - v\Delta_g u$. For (3), we note $f\mathrm{div}_g\alpha = f\nabla^i\alpha_i = \nabla^i(f\alpha_i) - \nabla^i f\cdot\alpha_i$, where $\alpha = \alpha_i dx^i$. □

Corollary 2.2

Let $(\mathcal{M}, g)$ be a closed m-dimensional Riemannian manifold. If α is an (r,s)-tensor field and β is an $(r-1,s)$-tensor field, then

$$\int_{\mathcal{M}}\langle\alpha,\nabla_g\beta_g\rangle_g dV_g = -\int_{\mathcal{M}}\langle\mathrm{div}_g(\alpha),\beta\rangle_g dV_g. \tag{2.6.4}$$

♡

Proof. Let $\gamma_j := \alpha^{ji_2\cdots i_r}_{k_1\cdots k_s}\beta^{k_1\cdots k_s}_{i_2\cdots i_r}$. Since

$$\begin{aligned}\mathrm{div}_g(\gamma) &= \nabla^j\gamma_j = \nabla^j\left(\alpha^{ji_2\cdots i_r}_{k_1\cdots k_s}\beta^{k_1\cdots k_s}_{i_2\cdots i_r}\right)\\ &= \nabla^j\alpha^{ji_2\cdots i_r}_{k_1\cdots k_s}\cdot\beta^{i_2\cdots i_r}_{k_1\cdots k_s} + \alpha^{k_1\cdots k_s}_{ji_2\cdots i_r}\cdot\nabla^j\beta^{k_1\cdots k_s}_{i_2\cdots i_r} = \langle\mathrm{div}_g(\alpha),\beta\rangle_g + \langle\alpha,\nabla_g\beta\rangle_g.\end{aligned}$$

Hence, (2.6.5) follows by applying the divergence theorem. □

Note 2.34. (Norm of 2-tensor dominates trace)

Show that for any 2-tensor α on an m-dimensional Riemannian manifold $(\mathcal{M}, g)$, we have

$$|\alpha|_g^2 \ge \frac{1}{m}(\mathrm{tr}_g\alpha)^2. \tag{2.6.5}$$

More generally, for a p-tensor α $(p\ge 2)$,

$$|\alpha|_g^2 \ge \frac{1}{m}\left|g^{ij}\alpha_{ijk_3\cdots k_p}\right|_g^2. \tag{2.6.6}$$

Choose a normal coordinate system $x^1,\cdots,x^m$ so that $g_{ij}=\delta_{ij}$ at a point. Then

$$(\mathrm{tr}_g\alpha)^2 = \left(\sum_{1\le i\le m}\alpha_{ii}\right)^2 \le \left(\sum_{1\le i\le m}1^2\right)\left(\sum_{1\le i\le m}\alpha_{ii}^2\right) = m|\alpha|_g^2.$$

♣

Lemma 2.8. (Bochner formula and inequality)

On a closed oriented m-dimensional Riemannian manifold $(\mathcal{M}, g)$,

$$\int_{\mathcal{M}}\left|\nabla_g^2 f\right|_g^2 dV_g + \int_{\mathcal{M}}\mathrm{Rc}_g(\nabla_g f,\nabla_g f)\,dV_g = \int_{\mathcal{M}}(\Delta_g f)^2\,dV_g. \tag{2.6.7}$$

In particular,

$$\int_{\mathcal{M}}\mathrm{Rc}_g(\nabla_g f,\nabla_g f)\,dV_g \le \frac{m-1}{m}\int_M(\Delta_g f)^2\,dV_g. \tag{2.6.8}$$

♡

Proof. Taking the integral on both sides of (2.5.24) yields

$$0 = \int_{\mathcal{M}}\left|\nabla_g^2 f\right|_g^2 dV_g + \int_M\mathrm{Rc}_g(\nabla_g f,\nabla_g f)\,dV_g + \int_{\mathcal{M}}\langle\nabla_g f,\nabla_g(\Delta_g f)\rangle_g\,dV_g.$$

By (2.6.5), the last integral equals

$$-\int_M\langle\mathrm{div}_g(\nabla_g f),\Delta_g f\rangle_g\,dV_g = -\int_M|\Delta_g f|_g^2\,dV_g.$$

Thus we prove (2.6.7). For (2.6.8) we use the inequality

$$\left|\nabla_g^2 f\right|_g^2 \ge \frac{1}{m}(\Delta_g f)^2 \tag{2.6.9}$$

that is a consequence of (2.6.6). □

Lemma 2.9. (Lichnerowicz)

Suppose that f is a non-zero eigenfunction of the Laplacian with eigenvalue $\lambda > 0$,

$$\Delta_g f + \lambda f = 0,$$

on a closed oriented m-dimensional Riemannian manifold $(\mathcal{M}, g)$. If $\mathrm{Rc}_g \geq (m-1)Kg$, where $K > 0$ is a constant, then

$$\lambda \geq mK. \tag{2.6.10}$$

♡

Proof. By (2.6.8) we deduce that

$$(m-1)K \int_{\mathcal{M}} |\nabla_g f|_g^2 \, dV_g \leq \frac{m-1}{m} \int_{\mathcal{M}} (\Delta_g f)^2 \, dV_g = \frac{m-1}{m} \int_{\mathcal{M}} \lambda^2 f^2 dV_g.$$

On the other hand,

$$\int_{\mathcal{M}} |\nabla_g f|_g^2 \, dV_g = -\int_{\mathcal{M}} f \Delta_g f dV_g = \int_{\mathcal{M}} \lambda f^2 dV_g$$

so that $mK\lambda \leq \lambda^2$. Since $\lambda > 0$, we must have $\lambda \geq mK$. □

2.6.2 De Rham theorem and Hodge decomposition theorem

Let $\mathcal{M}$ be a closed oriented m-dimensional manifold. Consider the following complex induced from the exterior differentiation d,

$$d : 0 \to \mathcal{A}^0(\mathcal{M}) \to \mathcal{A}^1(\mathcal{M}) \to \cdots \to \mathcal{A}^{m-1}(\mathcal{M}) \to \mathcal{A}^m(\mathcal{M}) \to 0, \tag{2.6.11}$$

where $d^2 = 0$. Hence $\mathrm{Im}(d) \subset \mathrm{Ker}(d)$ and we define the ***p*-th de Rham cohomology group**

$$H^p_{\mathrm{deR}}(\mathcal{M}) := \frac{\mathrm{Ker}\left(d|_{\mathcal{A}^p(\mathcal{M})}\right)}{\mathrm{Im}\left(d|_{\mathcal{A}^{p-1}(\mathcal{M})}\right)}. \tag{2.6.12}$$

Theorem 2.10. (De Rham)

If $\mathcal{M}$ is a closed oriented n-dimensional Riemannian manifold, then the p-th de Rham cohomology group is isomorphic to the p-th singular real coholomogy group:

$$H^p_{\mathrm{deR}}(\mathcal{M}) \cong H^p_{\mathrm{sing}}(\mathcal{M}; \mathbf{R})$$

and consequently, the de Rham cohomology groups $H^p_{\mathrm{deR}}(\mathcal{M})$ are all finite.

♡

A differential form α is called **harmonic** if

$$\Delta_{H,g}\alpha = 0. \tag{2.6.13}$$

The space of harmonic p-forms is denoted by

$$\mathcal{H}^p_g(\mathcal{M}) := \{\alpha \in \mathcal{A}^p(\mathcal{M}) : \Delta_{H,g}\alpha = 0\}. \tag{2.6.14}$$

Since

$$\int_{\mathcal{M}} \langle \Delta_{H,g}\alpha, \alpha \rangle_g \, dV_g = -\int_{\mathcal{M}} \left(|d\alpha|_g^2 + |\delta_g \alpha|_g^2\right) dV_g,$$

we have that α is harmonic if and only if $d\alpha = 0 = \delta_g \alpha$. Therefore

$$\mathcal{H}^p_g(\mathcal{M}) = \{\alpha \in \mathcal{A}^p(\mathcal{M}) : d\alpha = \delta_g\alpha = 0\}.$$

Given a p-form γ, we want to find a condition that $\Delta_{H,g}\alpha = \gamma$ has a solution for some $\alpha \in \mathcal{A}^p(\mathcal{M})$. If $\beta \in \mathcal{H}_g^p(\mathcal{M})$ is a harmonic p-form, then

$$(\gamma, \beta)_{L^2,g} = (\Delta_{H,g}\alpha, \beta)_{L^2,g} = (\alpha, \Delta_{H,d}\beta)_{L^2,g} = 0.$$

Hence a necessary condition to solve $\Delta_{H,g}\alpha = \gamma$ is $(\gamma, \beta)_{L^2,g} = 0$ for all $\beta \in \mathcal{H}_g^p(\mathcal{M})$. The converse of this is also true.

Theorem 2.11. (Hodge decomposition theorem)

Let $(\mathcal{M}, g)$ be a closed oriented m-dimensional Riemannian manifold. Given $\gamma \in \mathcal{A}^p(\mathcal{M})$, the equation

$$\Delta_{H,g}\alpha = \gamma$$

has a solution $\alpha \in \mathcal{A}^p(\mathcal{M})$ if and only if $(\gamma, \beta)_{L^2,g} = 0$ for all $\beta \in \mathcal{H}^p(\mathcal{M})$. Consequently, we have the following decomposition of the space of p-forms

$$\begin{aligned} \mathcal{A}^p(\mathcal{M}) &= \Delta_{H,g}\left(\mathcal{A}^p(\mathcal{M})\right) \oplus \mathcal{H}_g^p(\mathcal{M}) \\ &= d\delta_g\left(\mathcal{A}^p(\mathcal{M})\right) \oplus \delta_g d\left(\mathcal{A}^p(\mathcal{M})\right) \oplus \mathcal{H}_g^p(\mathcal{M}). \end{aligned} \tag{2.6.15}$$

Moreover, the space $\mathcal{H}_g^p(\mathcal{M})$ is finite-dimensional. ♡

Corollary 2.3

In each de Rham cohomology class, there is a unique harmonic form representing the cohomology class. In particular, the p-th de Rham cohomology group $H^p_{\mathrm{deR}}(\mathcal{M})$ is isomorphic to the space of harmonic p-forms $\mathcal{H}_g^p(\mathcal{M})$. ♡

Corollary 2.4

If $(\mathcal{M}, g)$ is a closed oriented n-dimensional Riemannian manifold and if $f : \mathcal{M} \to \mathbf{R}$ is a smooth function with

$$\int_{\mathcal{M}} f\, dV_g = 0,$$

then there exists a smooth function $u : \mathcal{M} \to \mathbf{R}$ such that $\Delta_g u = f$. The function u is uniquely determined up to an additive constant. ♡

Proof. By the Hodge decomposition theorem, we have $f = \Delta_g u + h$ for some smooth function $u : M \to \mathbf{R}$ and $h \in \mathcal{H}_g^0(\mathcal{M})$. Thus $\Delta_g h = 0$ and hence $g = 0$ since $\mathcal{M}$ is compact. The uniqueness is obvious. □

The Hodge Laplacian $\Delta_{H,g}$ commutes with the Hodge star operator $*_g$:

$$\Delta_{H,g} \circ *_g = *_g \circ \Delta_{H,g}. \tag{2.6.16}$$

Thus, if $\alpha \in \mathcal{H}_g^p(\mathcal{M})$ is a harmonic p-form, then $*_g\alpha$ is a harmonic $(m-p)$-form, i.e.,

$$*_g : \mathcal{H}_g^p(\mathcal{M}) \longrightarrow \mathcal{H}_g^{m-p}(\mathcal{M})$$

is an isomorphism. By **Corollary 2.3**,

$$H^p_{\mathbf{DR}}(\mathcal{M}) \cong H^{m-p}_{\mathbf{DR}}(\mathcal{M}), \tag{2.6.17}$$

which is known as the **Poincaré duality theorem** for de Rham cohomology.

2.6.3 Killing fields

A vector field X on an m-dimensional Riemannian manifold $(\mathcal{M}, g)$ is called a **Killing field** if the local flows generated by X act by isometries. This translates into the following simple characterization.

Proposition 2.14

A vector field X is a Killing field if and only if $\mathscr{L}_X g = 0$. ♡

Proof. Let F^t be the local flow for X. Recall that

$$(\mathscr{L}_X g)(V, W) = \frac{d}{dt}\Big|_{t=0} g\left(dF^t(V), dF^t(W)\right).$$

Thus we have

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0} g\left(dF^t(V), dF^t(W)\right) &= \frac{d}{dt}\Big|_{t=0} g\left(dF^{t-t_0}dF^{t_0}(V), dF^{t-t_0}dF^{t_0}(W)\right)\\ = \frac{d}{ds}\Big|_{s=0} g\left(dF^s dF^{t_0}(V), dF^s dF^{t_0}(W)\right) &= \mathscr{L}_X g\left(dF^{t_0}(V), dF^{t_0}(W)\right).\end{aligned}$$

This shows that $\mathscr{L}_X g = 0$ if and only if $t \mapsto g(dF^t(V), dF^t(W))$ is constant. Since F^0 is the identity map this is equivalent to assuming the flow acts by isometries. □

Proposition 2.15

X is a Killing field if and only if $V \mapsto (\nabla_g)_V X$ is a skew-symmetric $(1,1)$-tensor. ♡

Proof. Recall that

$$\begin{aligned}dX^\flat(V, W) &= \frac{1}{2}\left((\nabla_g)_V X^\flat(W) - (\nabla_g)_W X^\flat(V)\right)\\ &= \frac{1}{2}\left(V(X^\flat(W)) - W(X^\flat(V)) - X^\flat([V, W])\right)\\ &\quad - \frac{1}{2}\left(Vg(X, W) - Wg(X, V) - g(X, [V, W])\right)\\ = \frac{1}{2}&\left[g\left((\nabla_g)_V X, W\right) + g\left(X, (\nabla_g)_V W\right) - g\left((\nabla_g)_W X, V\right) - g\left(X, (\nabla_g)_W V\right)\right.\\ \left.- g\left(X, (\nabla_g)_V W\right) + g\left(X, (\nabla_g)_W V\right)\right] &= \frac{1}{2}\left[g\left((\nabla_g)_V X, W\right) - g\left((\nabla_g)_W X, V\right)\right]\end{aligned}$$

and

$$\mathscr{L}_X g(V, W) = g\left((\nabla_g)_V X, W\right) + g\left(V, (\nabla_g)_W X\right).$$

Hence

$$dX^\flat(V, W) + \frac{1}{2}\mathscr{L}_X g(V, W) = g\left((\nabla_g)_V X, W\right).$$

Thus, $\mathscr{L}_X g \equiv 0$ if and only if $V \mapsto (\nabla_g)_V X$ is skew-symmetric. □

In the following we consider Killing fields in negative Ricci curvature and positive Ricci curvature. It is related to the **Hopf conjecture** which states that any even-dimensional manifold with positive sectional curvature has positive Euler characteristic.

Proposition 2.16

Let X be a Killing field and V any vector field. If we set $f := \frac{1}{2}|X|_g^2$, then

$$\begin{aligned}\nabla_g f &= -(\nabla_g)_X X,\\ \mathrm{Hess}_g f(V,V) &= g\left((\nabla_g)_V X, (\nabla_g)_V X\right) - \mathrm{Rm}_g(V,X,X,V),\\ \Delta_g f &= |\nabla_g X|_g^2 - \mathrm{Rc}_g(X,X).\end{aligned}$$

♡

Proof. Since X is a Killing field, it follows that $\mathscr{L}_X g(V,W) = 0$ for any vector fields V and W. Consequently, $g\left((\nabla_g)_V X, W\right) + g\left(V, (\nabla_g)_W X\right) = 0$. For (1),

$$g\left(V, \nabla_g f\right) = (\nabla_g)_V f = g\left((\nabla_g)_V X, X\right) = -g\left(V, (\nabla_g)_X X\right).$$

For (2), we use the fact that $g\left(V, (\nabla_g)_V X\right) = 0$ to derive

$$\mathrm{Hess}_g f(V,V) = g\left((\nabla_g)_V \nabla_g f, V\right) = g\left((\nabla_g)_V\left(-(\nabla_g)_X X\right), V\right)$$

$$= -g\left((\nabla_g)_X (\nabla_g)_V X, V\right) - g\left(\mathrm{Rm}_g(V,X)X, V\right) - g\left((\nabla_g)_{[V,X]} X, V\right)$$

$$= -\mathrm{Rm}_g(V,X,X,V) - g\left((\nabla_g)_X(\nabla_g)_V X, V\right) + g\left((\nabla_g)_{(\nabla_g)_X V} X, V\right) - g\left((\nabla_g)_{(\nabla_g)_V X} X, V\right)$$

$$= g\left((\nabla_g)_V X, (\nabla_g)_V X\right) - \mathrm{Rm}_g(V,X,X,V) - g\left((\nabla_g)_X(\nabla_g)_V X, V\right) - g\left((\nabla_g)_X V, (\nabla_g)_V X\right)$$

$$= -\mathrm{Rm}_g(V,X,X,V) + g\left((\nabla_g)_V X, (\nabla_g)_V X\right) - (\nabla_g)_X g\left(V, (\nabla_g)_V X\right)$$

$$= -\mathrm{Rm}_g(V,X,X,V) \,|(\nabla_g)_V X|_g^2.$$

For (3), we select an orthonormal frame $\{e_i\}_{1\le i\le m}$ and calculate

$$\begin{aligned}\Delta_g f &= \sum_{1\le i\le m} \mathrm{Hess}_g f(e_i,e_i) = \sum_{1\le i\le m} |\nabla_{e_i} X|_g^2 - \sum_{1\le i\le m} \mathrm{Rm}_g(e_i,X,X,e_i)\\ &= -\mathrm{Rc}_g(X,X) + |\nabla_g X|_g^2.\end{aligned}$$

Thus, we complete the proof. □

Theorem 2.12. (Bochner, 1946)

Suppose $(\mathcal{M}, g)$ is a compact and oriented m-dimensional Riemannian manifold and has non-positive Ricci curvature. Then every Killing field is parallel. Furthermore, if $\mathrm{Rc}_g < 0$, then there are no nontrivial Killing fields.

♡

Proof. Set $f = \frac{1}{2}|X|_g^2$. Using Proposition 2.16 yields

$$0 = \int_{\mathcal{M}} \Delta_g f \, dV_g = \int_{\mathcal{M}} \left(-\mathrm{Rc}_g(X,X) + |\nabla_g X|_g^2\right) dV_g \ge \int_{\mathcal{M}} |\nabla_g X|_g^2 \, dV_g \ge 0.$$

Thus, $|\nabla_g X| \equiv 0$ and X must be parallel. In addition

$$\int_{\mathcal{M}} -\mathrm{Rc}_g(X,X)dV_g = 0, \quad -\mathrm{Rc}_g(X,X) \geq 0.$$

So $\mathrm{Rc}_g(X,X) = 0$. If $\mathrm{Rc}_g < 0$ this implies $X \equiv 0$. □

Theorem 2.13

Suppose that $(\mathcal{M}, g)$ is a compact and oriented m-dimensional Riemannian manifold and has quasi-negative Ricci curvature, i.e., $\mathrm{Rc}_g \leq 0$ and $\mathrm{Rc}_g(V,V) \leq 0$ for all $V \in T_p\mathcal{M} \setminus \{0\}$ for some $p \in \mathcal{M}$. Then $(\mathcal{M}, g)$ admits no nontrivial Killing fields. ♡

Proof. We have proved in **Theorem 2.13** that every Killing field is parallel. Thus a Killing field is always zero or never zero. In the latter holds, then $\mathrm{Rc}_g(X,X)(p) < 0$, but this contracts with $0 = \Delta_g f(p) = -\mathrm{Rc}_g(X,X)(p) > 0$. □

Problem 2.1. (Hopf)

Any even-dimensional closed manifold with positive sectional curvature has positive Euler characteristic. ♠

We will show that $H_1(\mathcal{M}; \mathbf{R}) = 0$ provided the Ricci curvature is positive. Assume this, the Hopf conjecture holds in dimension 2; in dimension 4, Poincaré duality implies that $H_1(\mathcal{M}; \mathbf{R}) = H_3(\mathcal{M}; \mathbf{R}) = 0$. Hence $\chi(\mathcal{M}) = 1 + \dim(H_2(\mathcal{M}; \mathbf{R})) + 1 \geq 2$.

Theorem 2.14. (Berger, 1965)

If $(\mathcal{M}, g)$ is a closed, even-dimensional Riemannian manifold of positive sectional curvature, then every Killing field has a zero. ♡

Proof. Let X be a Killing field and consider the function $f = \frac{1}{2}|X|_g^2$. If X has no zeros, f will have a positive minimum at some point $p \in M$. Then $(\mathrm{Hess}_g f)_p \geq 0$. We also know that

$$\mathrm{Hess}_g f(V,V) = g\left((\nabla_g)_V X, (\nabla_g)_V X\right) - \mathrm{Rm}_g(V,X,X,V).$$

Since $(\nabla_g f)_p = -\left((\nabla_g)_X X\right)_p$ and f has a minimum at p, $\left((\nabla_g)_X X\right)_p = 0$. Thus, we have a skew-symmetric map $T_p\mathcal{M} \to T_p\mathcal{M}$ with at least one zero eigenvalue. But then, even dimensionality of $T_p\mathcal{M}$ ensures us that there must be at least one more zero eigenvector $v \in T_p\mathcal{M}$ linearly independent from X. Thus

$$\mathrm{Hess}_g f(v,v) = \langle (\nabla_g)_v X, (\nabla_g)_v X \rangle_{g(p)} - \mathrm{Rm}_g(v,X,X,v) = -\mathrm{Rm}_g(v,X,X,v) \leq 0$$

by assumption. □

Theorem 2.15. (Rong, 1995)

If a closed Riemannian m-manifold $(\mathcal{M}, g)$ admits a nontrivial Killing field, then the fundamental group has a cyclic subgroup of index $\leq c(m)$. ♡

We define the **Betti number** of an m-manifold $\mathcal{M}$ as

$$b_p(\mathcal{M}) := \dim(H_p(\mathcal{M};\mathbf{R})) = \dim(H^p(\mathcal{M};\mathbf{R})) \tag{2.6.18}$$

and the **Euler characteristic** as the alternating sum

$$\chi(\mathcal{M}) := \sum_{0\le p\le m} (-1)^p b_p(\mathcal{M}). \tag{2.6.19}$$

It is a key result in algebraic topology that $H_p(\mathcal{M};\mathbf{R})$ and $H^p(\mathcal{M};\mathbf{R})$ have the same dimension when we use real coefficients. Note that Poincaré duality implies that $b_p(\mathcal{M}) = b_{m-p}(\mathcal{M})$.

Theorem 2.16. (Conner, 1957)

Let X be a Killing field on a compact Riemannian manifold $\mathcal{M}$. If $\mathcal{N}_i \subset \mathcal{M}$ are the components of the zero set for X, then

$$\begin{aligned} \chi(\mathcal{M}) &= \sum_i \chi(\mathcal{N}_i), \\ \sum_p b_{2p}(\mathcal{M}) &\ge \sum_i \sum_p b_{2p}(\mathcal{N}_i), \\ \sum_p b_{2p+1}(M\mathcal{M}) &\ge \sum_i \sum_p b_{2p+1}(\mathcal{N}_i). \end{aligned}$$

Here we use the fact that the zero set of a Killing field is a disjoint union of totally geodesic submanifolds each of even codimension. ♡

Corollary 2.5

If $\mathcal{M}$ is a compact 6-manifold with positive sectional curvature that admits Killing field, then $\chi(M) > 0$. ♡

Proof. We know that the zero set for a Killing field is non-empty and that each component has even codimension. Thus each component is a 0, 2, or 4-dimensional manifold with positive sectional curvature. This shows that $\mathcal{M}$ has positive Euler characteristic. □

Theorem 2.17. (Hsiang-Kleiner, 1989)

If $\mathcal{M}$ is a compact orientable positively curved 4-manifold that admits a Killing field, then the Euler characteristic is ≤ 3. In particular, $\mathcal{M}$ is topologically equivalent to $\mathbf{S}^4$ or $\mathbf{CP}^2$. ♡

The **rank** of a compact Lie group is the maximal dimension of an Abelian subalgebra in the corresponding Lie algebra. The **symmetry rank** of a compact Riemannian manifold is the rank of the isometry group.

Theorem 2.18. (Grove-Searle, 1994)

(1)Let $\mathcal{M}$ be a compact m-manifold with positive sectional curvature and symmetry rank k. If $k \ge \frac{m}{2}$, then $\mathcal{M}$ is diffeomorphic to either a sphere, complex projective space or a cyclic quotient of a sphere $\mathbf{S}^m/\mathbf{Z}_q$, where $\mathbf{Z}_q$ is a cyclic group of order q acting by

isometries on the unit sphere.

(2) Let $\mathcal{M}$ be a closed m-manifold with positive sectional curvature. If $\mathcal{M}$ admits a Killing field such that the zero set has a component N of codimension 2, then $\mathcal{M}$ is diffeomorphic to $\mathbf{S}^m$, $\mathbf{CP}^{m/2}$, or a cyclic quotient of a sphere $\mathbf{S}^m/\mathbf{Z}_q$. ♡

Theorem 2.19. (Püttmann-Searle, 2002)

If $\mathcal{M}$ is a compact $2m$-manifold with positive sectional curvature and symmetry rank $k \geq \frac{2m-4}{4}$, then $\chi(\mathcal{M}) > 0$. ♡

Theorem 2.20. (Wilking, 2003)

Let $\mathcal{M}$ be a compact simply-connected positively curved m-manifold with symmetry rank k. If $k \geq \frac{m}{4} + 1$, then $\mathcal{M}$ has the topology of a sphere, complex projective space or quaternionic projective space. ♡

Theorem 2.21. (Rong-Su, 2005)

If $\mathcal{M}$ is a compact $2m$-manifold with positive sectional curvature and symmetry rank $k > \frac{2m-4}{8}$, then $\chi(\mathcal{M}) > 0$. ♡

The theorem also holds if we only assume that $k \geq \frac{2m-4}{8}$ as well as $k \geq 2$ when $2m = 12$.

2.6.4 Affine vector fields

Let $(\mathcal{M}, g)$ be an n-dimensional Riemannian manifold. For a vector field X we define the Lie derivative of the connection as

$$\begin{aligned}
(\mathscr{L}_X \nabla_g)(U, V) &:= \mathscr{L}_X((\nabla_g)_U V) - (\nabla_g)_{\mathscr{L}_X U} V - (\nabla_g)_U \mathscr{L}_X V \qquad (2.6.20)\\
&= [X, (\nabla_g)_U V] - (\nabla_g)_{[X,U]} V - (\nabla_g)_U [X, V]. \qquad (2.6.21)
\end{aligned}$$

Lemma 2.10

$\mathscr{L}_X \nabla_g$ is a $(2,1)$-tensor field. ♡

Proof. For any smooth function f on $\mathcal{M}$, we have

$$\begin{aligned}
(\mathscr{L}_X(\nabla_g)_g)(fU, V) &= [X, (\nabla_g)_{fU} V] - (\nabla_g)_{[X,fU]} V - (\nabla_g)_{fU}[X, V]\\
&= [X, f(\nabla_g)_U V] - (\nabla_g)_{[X,fU]} V - f(\nabla_g)_U [X, V]\\
&= f[X, (\nabla_g)_U V] + Xf \cdot (\nabla_g)_U V - (\nabla_g)_{f[X,U] + Xf \cdot U} V - f(\nabla_g)_U [X, V]\\
&= f\left([X, (\nabla_g)_U V] - (\nabla_g)_{[X,U]} V - (\nabla_g)_U [X, V]\right) = f(\mathscr{L}_X \nabla_g)(U, V).
\end{aligned}$$

For the second factor, one has

$$(\mathscr{L}_X \nabla_g)(U, fV) = [X, (\nabla_g)_U fV] - (\nabla_g)_{[X,U]} fV - (\nabla_g)_U [X, fV]$$

$$\begin{aligned}
&= [X, Uf \cdot V + f \cdot (\nabla_g)_U V] - \left([X,U]f \cdot V + f \cdot (\nabla_g)_{[X,U]} V\right) - (\nabla_g)_U \left(f[X,V] + Xf \cdot V\right) \\
&= [X, Uf \cdot V] + f\,[X, (\nabla_g)_U V] + Xf \cdot (\nabla_g)_U V - [X,V]f \cdot V - f(\nabla_g)_{[X,U]} V - Uf \cdot [X,V] \\
&\quad - f \cdot (\nabla_g)_U [X,V] - UXf \cdot V - Xf \cdot (\nabla_g)_U V \;=\; f\left(\mathscr{L}_X \nabla_g\right)(U,V).
\end{aligned}$$

Hence $\mathscr{L}_X \nabla_g$ is a $(2,1)$-tensor field. □

We say a vector field X is an **affine vector field** if $\mathscr{L}_X \nabla_g = 0$.

Lemma 2.11

(1) For an affine vector field X, we have

$$(\nabla_g)^2_{U,V} X = -\mathrm{Rm}_g(X,U)V.$$

(2) Show that Killing vector fields are affine. ♡

Proof. Calculate for any vector fields X, U, V,

$$\begin{aligned}
\left(\mathscr{L}_X \nabla_g\right)(U,V) - \mathrm{Rm}_g(X,U)V &= [X, (\nabla_g)_U V] - (\nabla_g)_{[X,U]} V - (\nabla_g)_U [X,V] \\
&\quad - (\nabla_g)_X (\nabla_g)_U V + (\nabla_g)_U (\nabla_g)_X V + (\nabla_g)_{[X,U]} V \\
&= (\nabla_g)_X (\nabla_g)_U V - (\nabla_g)_{(\nabla_g)_U V} X - (\nabla_g)_U \left((\nabla_g)_X V - (\nabla_g)_V X\right) \\
&\quad - (\nabla_g)_X (\nabla_g)_U V + (\nabla_g)_U (\nabla_g)_X V = (\nabla_g)_U (\nabla_g)_V X - (\nabla_g)_{\nabla_U V} X = (\nabla_g)^2_{U,V} X.
\end{aligned}$$

Hence, if X is affine, we obtain the desired result. □

Lemma 2.12. (Yano, 1958)

Let X be a vector field on a closed oriented m-dimensional Riemannian manifold $(\mathcal{M}, g)$. Then

$$\begin{aligned}
\int_{\mathcal{M}} \left[\mathrm{Rc}_g(X,X) + \mathrm{tr}_g\left((\nabla_g X)^2\right) - (\mathrm{div}_g X)^2\right] dV_g &= 0, \\
\int_{\mathcal{M}} \left[\mathrm{Rc}_g(X,X) + g\left(\mathrm{tr}_g \nabla^2_g X, X\right) + \frac{1}{2}\left|\mathscr{L}_X g\right|^2_g - (\mathrm{div}_g X)^2\right] dV_g &= 0.
\end{aligned}$$

♡

Proof. Calculate

$$\begin{aligned}
\mathrm{div}_g\left((\nabla_g)_X X\right) &= \nabla_i \left((\nabla_g)_X X\right)^i = \nabla_i \left(\nabla_j X^i \cdot X^j\right) \\
= \nabla_i \nabla_j X^i \cdot X^j + \nabla_j X^i \cdot \nabla_i X^j &= \left(\nabla_j \nabla_i X^i + R^i_{ijk} X^k\right) X^j + \nabla_j X^i \cdot \nabla_i X^j \\
&= \nabla_j \nabla_i X^i \cdot X^j + R_{jk} X^j X^k + \nabla_j X^i \cdot \nabla_i X^j.
\end{aligned}$$

The first term can be computed by

$$\mathrm{div}_g\left(\mathrm{div}_g X \cdot X\right) = \nabla_j\left(\nabla_i X^i \cdot X^j\right) = \nabla_j \nabla_i X^i \cdot X^j + \nabla_i X^i \cdot \nabla_j X^j.$$

Hence

$$\mathrm{div}_g\left((\nabla_g)_X X\right) - \mathrm{div}_g\left(\mathrm{div}_g X \cdot X\right) = \mathrm{Rc}_g(X,X) + \nabla_j X^i \cdot \nabla_i X^j - (\mathrm{div}_g X)^2.$$

Since X is compact and oriented, taking the integral on both sides implies the first integral formula. $\square$

Note 2.35. (Yano)

If X is an affine vector field then $\mathrm{tr}_g \nabla_g^2 X = -\mathrm{Rc}_g(X)$ and that $\mathrm{div}_g X$ is constant. Using **Lemma 2.12** *yields on closed oriented manifolds affine fields are Killing fields.* ♣

2.7 Curvature decomposition and LCF manifolds

Introduction

- *Decomposition of the curvature tensor field*
- *Locally conformally flat manifolds*

Let $(\mathcal{M}, g)$ be a Riemannian m-manifold. The Riemann curvature $(4,0)$-tensor field Rm_g is a section of the bundle $\odot^2 \wedge^2 T^*\mathcal{M} := \wedge^2 T^*\mathcal{M} \otimes_S \wedge^2 T^*\mathcal{M}$, where $\wedge^2 T^*\mathcal{M}$ denotes the vector bundle of 2-forms and $\otimes_S$ denotes the symmetric tensor product bundle.

2.7.1 Decomposition of the curvature tensor field

By the first Bianchi identity, Rm_g is a section of the subbundle $\mathrm{Ker}(\mathbf{b})$, the kernel of the linear map:

$$\mathbf{b} : \odot^2 \wedge^2 T^*\mathcal{M} \longrightarrow \wedge^3 T^*\mathcal{M} \otimes_S T^*\mathcal{M} \tag{2.7.1}$$

defined by

$$\mathbf{b}(\Omega)(X,Y,Z,W) := \frac{1}{3}\left(\Omega(X,Y,Z,W) + \Omega(Y,Z,X,W) + \Omega(Z,X,Y,W)\right). \tag{2.7.2}$$

We shall call $\mathbf{C}(\mathcal{M}) := \mathrm{Ker}(\mathbf{b})$ the bundle of curvature tensor fields. For every $x \in \mathcal{M}$, the fiber $\mathbf{C}_x(\mathcal{M})$ has the structure of an $\mathbf{O}(T_x^*\mathcal{M})$-module, given by

$$\times : \mathbf{O}(T_x^*\mathcal{M}) \times \mathbf{C}_x(\mathcal{M}) \longrightarrow \mathbf{C}_x(\mathcal{M}) \tag{2.7.3}$$

where

$$A \times ((\alpha \wedge \beta) \otimes (\gamma \wedge \delta)) := (A\alpha \wedge A\beta) \otimes (A\gamma \wedge A\delta) \tag{2.7.4}$$

for $A \in \mathbf{O}(T_x^*\mathcal{M})$ and $\alpha, \beta, \gamma, \delta \in T_x^*\mathcal{M}$. As an $\mathbf{O}(T_x^*\mathcal{M})$ representation space, $\mathbf{C}_x(M)$ has a natural decomposition into its irreducible components. Consider the **Kulkarni-Nomizu product**

$$\odot : \odot^2 T^*\mathcal{M} \times \odot^2 T^*\mathcal{M} \longrightarrow \mathbf{C}(\mathcal{M}) \tag{2.7.5}$$

defined by

$$(\alpha \odot \beta)_{ijk\ell} := \alpha_{i\ell}\beta_{jk} + \alpha_{jk}\beta_{i\ell} - \alpha_{ik}\beta_{j\ell} - \alpha_{j\ell}\beta_{ik}. \tag{2.7.6}$$

The irreducible decomposition of $\mathbf{C}_x(\mathcal{M})$ as an $\mathbf{O}(T_x^*\mathcal{M})$-module is given by

$$\mathbf{C}(\mathcal{M}) = (\mathbf{R}g \odot g) \oplus \left(\odot_0^2 T^*\mathcal{M} \odot g\right) \oplus \mathbf{W}(\mathcal{M}) \tag{2.7.7}$$

where $\odot_0^2 T^*\mathcal{M}$ is the bundle of symmetric, trace-free 2-forms and

$$\mathbf{W}(\mathcal{M}) := \mathrm{Ker}(\mathbf{b}) \cap \mathrm{Ker}(\mathbf{c}) \tag{2.7.8}$$

is the bundle of **Weyl curvature tensor fields**. Here

$$\mathbf{c} : \odot^2 \wedge^2 T^*\mathcal{M} \longrightarrow \odot^2 T^*\mathcal{M} \tag{2.7.9}$$

is the contraction map defined by

$$\mathbf{c}(\Omega)(X, Y) := \sum_{i=1}^{m} \Omega(e_i, X, Y, e_i). \tag{2.7.10}$$

Note that

$$(g \odot g)_{ijk\ell} = 2(g_{i\ell}g_{jk} - g_{ik}g_{j\ell}). \tag{2.7.11}$$

The irreducible decomposition of $\mathbf{C}(\mathcal{M})$ yields the following irreducible decomposition of the Riemann curvature tensor field:

$$\mathrm{Rm}_g = f \cdot g \odot g + h \odot g + W, \tag{2.7.12}$$

where $f \in C^\infty(\mathcal{M})$, $h \in C^\infty(\mathcal{M}, \odot_0^2 T^*\mathcal{M})$, and $W \in C^\infty(\mathcal{M}, \mathbf{W}(\mathcal{M}))$. Taking the contraction $\mathbf{c}$ of this equation implies

$$R_{jk} = 2(n-1) f g_{jk} + (m-2) h_{jk}; \tag{2.7.13}$$

taking the contraction again we have

$$R_g = 2(m-1)mf + (m-2)\mathrm{tr}_g h = 2m(m-1)f, \tag{2.7.14}$$

since h is trace-free. Hence

$$f = \frac{R_g}{2m(m-1)}, \quad h = \frac{1}{m-2}\mathrm{Rc}_g - \frac{1}{m(m-2)} R_g \cdot g. \tag{2.7.15}$$

Using (2.7.12) and (2.7.15) we deduce that for $m \geq 3$,

$$\begin{aligned} \mathrm{Rm}_g &= \frac{R_g}{2(m-1)(m-2)} g \odot g + \frac{1}{m-2}\mathrm{Rc}_g \odot g + \mathrm{Weyl}_g && (2.7.16) \\ &= \frac{R_g}{2m(m-1)} g \odot g + \frac{1}{m-2}\mathrm{Rc}_g^\circ \odot g + \mathrm{Weyl}_g, && (2.7.17) \end{aligned}$$

where $\mathrm{Rc}_g^\circ := \mathrm{Rc}_g - \frac{R_g}{m} g$ is the traceless Ricci tensor field and Weyl_g is the Weyl tensor field, which is defined by (2.7.16). We let

$$W_{ijkl} := \mathrm{Weyl}_g(\partial_i, \partial_j, \partial_k, \partial_l). \tag{2.7.18}$$

In local coordinates, (2.7.16) says that for $m \geq 3$,

$$\begin{aligned} W_{ijk\ell} &= R_{ijk\ell} + \frac{R_g}{(m-1)(m-2)}(g_{i\ell}g_{jk} - g_{ik}g_{j\ell}) && (2.7.19) \\ &\quad - \frac{1}{m-2}(R_{i\ell}g_{jk} + g_{i\ell}R_{jk} - R_{ik}g_{j\ell} - g_{ik}R_{j\ell}). \end{aligned}$$

Hence

$$W_{ijk\ell} = -W_{jik\ell} = -W_{ij\ell k} = W_{k\ell ij}. \tag{2.7.20}$$

We claim that

$$g^{ik} W_{ijk\ell} = 0. \tag{2.7.21}$$

Indeed,

$$
\begin{aligned}
g^{ik}W_{ijk\ell} &= -R_{j\ell} + \frac{R_g}{(m-1)(m-2)}(g_{j\ell} - ng_{j\ell}) - \frac{1}{m-2}(R_{j\ell} + R_{j\ell} - Rg_{j\ell} - nR_{j\ell}) \\
&= -R_{j\ell} - \frac{R_g}{m-2}g_{j\ell} - \frac{1}{m-2}((2-m)R_{j\ell} - Rg_{j\ell}) \\
&= -R_{j\ell} - \frac{R_g}{m-2}g_{j\ell} + R_{j\ell} + \frac{R_g}{m-2}g_{j\ell} \ = \ 0.
\end{aligned}
$$

Note 2.36

The Weyl tensor field vanishes when $m \leq 3$. For $m = 2$, the possibly non-trivial component is W_{1212}. Using (2.7.21), we have

$$0 = W_{1212} + W_{2222} = W_{1212}.$$

For $m = 3$ there are only two possible types of nonzero components of W. Either there are three distinct indices such as W_{1231} or there are two distinct indices such as W_{1221}. First we compute, using the trace-free property,

$$W_{1231} = -W_{2232} - W_{3233} = 0.$$

Next, we have

$$W_{1221} = -W_{2222} - W_{3223} = -W_{3223} = W_{3113} = -W_{2112} = -W_{1221}$$

which implies $W_{1221} = 0$. ♣

By **Note 2.36**, (2.7.12), (2.7.13), and (2.7.14), we conclude that for $m = 2$,

$$R_{ijk\ell} = \frac{R_g}{2}(g_{i\ell}g_{jk} - g_{ik}g_{j\ell}), \quad R_{jk} = \frac{R_g}{2}g_{jk}. \tag{2.7.22}$$

Similarly, for $m = 3$, we have

$$R_{ijk\ell} = R_{i\ell}g_{jk} + R_{jk}g_{i\ell} - R_{ik}g_{j\ell} - R_{j\ell}g_{ik} - \frac{R_g}{2}(g_{i\ell}g_{jk} - g_{ik}g_{j\ell}). \tag{2.7.23}$$

Lemma 2.13

If $\widetilde{g} = e^{2f}g$ for some function f, then

$$\widetilde{R}^{\ell}_{ijk} = R^{\ell}_{ijk} - a^{\ell}_i g_{jk} - a_{jk}\delta^{\ell}_i + a_{ik}\delta^{\ell}_j + a^{\ell}_j g_{ik}, \tag{2.7.24}$$

where

$$a_{ij} := \nabla_i\nabla_j f - \nabla_i f \nabla_j f + \frac{1}{2}|\nabla_g f|^2_g g_{ij}. \tag{2.7.25}$$

That is,

$$e^{-2f}\mathrm{Rm}_{\widetilde{g}} = \mathrm{Rm}_g - a \odot g. \tag{2.7.26}$$

From this deduce the Weyl tensor field is conformally invariant:

$$\mathrm{Weyl}_{e^{2f}g} = e^{2f}\mathrm{Weyl}_g. \tag{2.7.27}$$

♡

Proof. First we compute that

$$\widetilde{\Gamma}^k_{ij} = \Gamma^k_{ij} + \nabla_i f \cdot \delta^k_j + \nabla_j f \cdot \delta^k_i - \nabla^k f \cdot g_{ij}.$$

If we set $A_{ij}^k := \widetilde{\Gamma}_{ij}^k - \Gamma_{ij}^k$, then

$$\widetilde{R}_{ijk}^\ell = R_{ijk}^\ell + \nabla_i A_{jk}^\ell - \nabla_j A_{ik}^\ell + A_{jk}^m A_{im}^\ell - A_{ik}^m A_{jm}^\ell.$$

Simplifying it gives (2.7.24) and hence (2.7.26). □

From (2.7.7) we have the (reducible) decomposition:

$$\mathbf{C}(\mathcal{M}) \cong \left(\odot^2 T^*\mathcal{M} \odot g\right) \oplus \mathbf{W}(\mathcal{M}). \tag{2.7.28}$$

Then

$$\mathrm{Rm}_g = \frac{1}{m-2}\mathrm{S}_g \odot g + \mathrm{Weyl}_g \tag{2.7.29}$$

where $\mathrm{S}_g := \mathrm{Rc}_g - \frac{R_g}{2(m-1)}g$ is the **Weyl-Schouten tensor field**. If $m \geq 3$, then

$$\nabla^\ell W_{ijk\ell} = \frac{m-3}{m-2}C_{ijk} \tag{2.7.30}$$

where ($S_{ik} := \mathrm{S}_g(\partial_i, \partial_k)$)

$$\begin{aligned} C_{ijk} &:= \nabla_i S_{jk} - \nabla_j S_{ik} \\ &= \nabla_i R_{jk} - \nabla_j R_{ik} - \frac{1}{2(m-1)}(\nabla_i R_g \cdot g_{jk} - \nabla_j R_g \cdot g_{ik}) \end{aligned} \tag{2.7.31}$$

are the components of the **Cotton tensor field** C_g.

Note 2.37

(1) For $m \geq 4$, if the Weyl tensor field of an m-dimensional Riemannian manifold $(\mathcal{M}, g)$ vanishes, then the Cotton tensor field vanishes. When $m = 3$, the Weyl tensor field always vanishes but the Cotton tensor field does not vanishes in general.

(2) If $m = 3$, then

$$\mathrm{C}_{e^{2f}g} = e^{2f}\mathrm{C}_g. \tag{2.7.32}$$

♣

2.7.2 Locally conformally flat manifolds

We say that a m-dimensional Riemannian manifold $(\mathcal{M}, g)$ is **locally conformally flat** if for every point $p \in \mathcal{M}$, there exists a local coordinate system $\{x^i\}_{1\leq i\leq m}$ in a neighborhood $\mathcal{U}$ of p such that

$$g_{ij} = v \cdot \delta_{ij}$$

for some function v defined on $\mathcal{U}$, e.g., $v^{-1}g$ is a flat metric.

Note 2.38

Any Riemannian surface is locally conformally flat. Indeed, if $(\mathcal{M}, g)$ is a Riemannian surface and u is a function on $\mathcal{M}$, then

$$R_{e^u g} = e^{-u}\left(R_g - \Delta_g u\right).$$

To find u locally so that $R_{e^u g} = 0$ we need to solve the Poisson equation $\Delta_g u = R_g$ which is certainly possible.

♣

Proposition 2.17. (Weyl, Schouten)

An m-dimensional Riemannian manifold $(\mathcal{M}, g)$ is locally conformally flat if and only if

(1) for $m \geq 4$ the Weyl tensor field vanishes,

(2) for $m = 3$ the Cotton tensor field vanishes.

♡

Corollary 2.6

If a Riemannian manifold $(\mathcal{M}, g)$ has constant sectional curvature, then $(\mathcal{M}, g)$ is locally conformally flat.

♡

Proof. If the sectional curvature is constant, then

$$\mathrm{Rm}_g = \frac{R_g}{2m(m-1)} g \odot g$$

so that the Weyl tensor field vanishes. By **Proposition 2.17**, $(\mathcal{M}, g)$ is locally conformally flat. □

Corollary 2.7

(1) If $(\mathcal{N}, g_{\mathcal{N}})$ and $(\mathcal{P}, g_{\mathcal{P}})$ are Riemannian manifolds such that

$$\mathrm{Sec}_{g_{\mathcal{N}}} \equiv C, \quad \mathrm{Sec}_{g_{\mathcal{P}^{\mathcal{P}}}} \equiv -C, \quad \textit{for some } C \in \mathbf{R},$$

then their Riemannian product $(\mathcal{N} \times \mathcal{P}, g_{\mathcal{N}} + g_{\mathcal{P}})$ is locally conformally flat.

(2) If $(\mathcal{N}, g_{\mathcal{N}})$ has $\mathrm{Sec}_{g_{\mathcal{N}}} \equiv C$, then the Riemannian product $(\mathcal{N} \times \mathbf{R}, g_{\mathcal{N}} + dt^2)$ is locally conformally flat.

♡

Proof. (1) Since

$$\begin{aligned} \mathrm{Rm}_{g_{\mathcal{N} \times \mathcal{P}^{\mathcal{P}}}} &= \mathrm{Rm}_{g_{\mathcal{N}}} + \mathrm{Rm}_{g_{\mathcal{P}^{\mathcal{P}}}} = \frac{C}{2} g_{\mathcal{N}} \odot g_{\mathcal{N}} - \frac{C}{2} g_{\mathcal{P}^{\mathcal{P}}} \odot g_{\mathcal{P}^{\mathcal{P}}} \\ &= \frac{C}{2} \left(g_{\mathcal{N}} - g_{\mathcal{P}^{\mathcal{P}}}\right) \odot \left(g_{\mathcal{N}} + g_{\mathcal{P}^{\mathcal{P}}}\right), \end{aligned}$$

the uniqueness of the decomposition tells $\mathrm{Weyl}_{\mathcal{N} \times \mathcal{P}^{\mathcal{P}}} = 0$.

(2) The Riemann curvature tensor field of the product is

$$\mathrm{Rm}_{g_{\mathcal{N} \times \mathbf{R}}} = \frac{C}{2} g_{\mathcal{N}} \odot g_{\mathcal{N}} = \frac{C}{2} \left(g_{\mathcal{N}} - dt^2\right) \odot \left(g_{\mathcal{N}} + dt^2\right)$$

where we used the fact that $dt^2 \odot dt^2 = 0$. Therefore $\mathrm{Weyl}_{\mathcal{N} \times \mathbf{R}} = 0$. □

We say that two Riemannian manifolds $(\mathcal{M}, g)$ and $(\mathcal{N}, h)$ are **conformally equivalent** if there exist a diffeomorphism $\varphi : \mathcal{M} \to \mathcal{N}$ and a function $f : \mathcal{M} \to \mathbf{R}$ such that $g = e^f \varphi^* h$.

Theorem 2.22. (Kuiper)

If $(\mathcal{M}, g)$ is a simply-connected, locally conformally flat, closed m-dimensional Riemannian manifold, then $(\mathcal{M}, g)$ is conformally equivalent to the standard sphere $\mathbf{S}^m$.

♡

A map ψ from one Riemanian manifold $(\mathcal{M}, g)$ to another $(\mathcal{N}, h)$ is said to be **conformal** if there exists a function $f : \mathcal{M} \to \mathbf{R}$ such that $g = e^f \psi^* h$.

Theorem 2.23. (Schoen-Yau, 1988)

If $(\mathcal{M}, g)$ is a simply-connected, locally conformally flat, complete m-dimensional Riemannian manifold in the conformal class of a metric with nonnegative scalar curvature, then there exists a one-to-one conformal map of $(\mathcal{M}, g)$ into the standard sphere $\mathbf{S}^m$. ♡

When $(\mathcal{M}, g)$ is not simply-connected, we can apply the above results to the universal cover $(\widetilde{\mathcal{M}}^m, \tilde{g})$.

Note 2.39

If $(\mathcal{M}_1^{m_1}, g_1)$ and $(\mathcal{M}_2^{m_2}, g_2)$ are Riemannian manifolds, then the product Riemannian manifold $(\mathcal{M}_1^{m_1} \times \mathcal{M}_2^{m_2}, g_1 + g_2)$ satisfies

$$\begin{aligned} \mathrm{Rm}_{g_1+g_2}(X, Y, Z, W) &= \mathrm{Rm}_{g_1}(X_1, Y_1, Z_1, W_1) + \mathrm{Rm}_{g_2}(X_2, Y_2, Z_2, W_2), \\ \mathrm{Rc}_{g_1+g_2}(X, Y) &= \mathrm{Rc}_{g_1}(X_1, Y_1) + \mathrm{Rc}_{g_2}(X_2, Y_2) \end{aligned}$$

where $X = (X_1, X_2)$, etc. ♣

2.8 Moving frames and the Gauss-Bonnet formula

Introduction

- *Cartan structure equations*
- *Curvature under conformal change of metric*
- *The Gauss-Bonnet formula*
- *Moving frames adapted to hypersurfaces*

2.8.1 Cartan structure equations

Let $\{e_i\}_{1\le i\le m}$ be a local orthonormal frame field in an open set $\mathcal{U}$ of an m-dimensional Riemannian manifold $(\mathcal{M}, g)$. The dual orthonormal basis (or **coframe field**) $\{\omega^i\}_{1\le i\le m}$ of $C^\infty(\mathcal{M}, T^*\mathcal{M})$ is defined by $\omega^i(e_j) = \delta^i_j$ for all $i, j = 1, \cdots, m$. We can write the metric g as

$$g = \sum_{1\le i\le m} \omega^i \otimes \omega^i. \tag{2.8.1}$$

The **connection 1-forms** ω_i^j are the components of the Levi-Civita connection with respect to $\{e_i\}_{1\le i\le m}$:

$$(\nabla_g)_X e_i := \sum_{1\le j\le m} \omega_i^j(X) e_j, \tag{2.8.2}$$

for all $i, j = 1, \cdots, m$ and all vector fields X on $\mathcal{U}$. Since for all X

$$0 = X\langle e_i, e_j\rangle_g = \langle (\nabla_g)_X e_i, e_j\rangle_g + \langle e_i, (\nabla_g)_X e_j\rangle_g,$$

the connection 1-forms are anti-symmetric:

$$\omega_i^j = -\omega_j^i. \tag{2.8.3}$$

From $\omega^i(e_j) = \delta^i_j$ and the product rule we see that

$$(\nabla_g)_X \omega^i = -\omega^i_j(X)\omega^j. \tag{2.8.4}$$

The **curvature 2-forms** $\Omega^j_i := \mathrm{Rm}^j_i$ on $\mathcal{U}$ are defined by

$$\mathrm{Rm}^j_i(X,Y)e_j := \frac{1}{2}\mathrm{Rm}_g(X,Y)e_i \tag{2.8.5}$$

so that

$$\mathrm{Rm}^j_i(X,Y) := \frac{1}{2}\langle \mathrm{Rm}_g(X,Y)e_i, e_j\rangle_g. \tag{2.8.6}$$

Theorem 2.24. (Cartan structure equations)

The **first and second Cartan structure equations** *are*

$$\begin{aligned} d\omega^i &= \omega^j \wedge \omega^i_j, \quad (2.8.7)\\ \Omega^j_i &= d\omega^j_i - \omega^k_i \wedge \omega^j_k. \quad (2.8.8)\end{aligned}$$

♡

Proof. Calculate

$$\begin{aligned} d\omega^i(X,Y) &= \frac{1}{2}\left((\nabla_g)_X\omega^i\right)(Y) - \frac{1}{2}\left((\nabla_g)_Y\omega^i\right)(X) \\ &= -\frac{1}{2}\omega^i_j(X)\omega^j(Y) + \frac{1}{2}\omega^i_j(Y)\omega^j(X) \ = \ (\omega^j\wedge\omega^i_j)(X,Y)\end{aligned}$$

implying (2.8.7). From (2.2.8) and $\nabla^2_g e_i = \nabla_g\omega^k_i \cdot e_k + \omega^k_i\nabla_g e_k$, we have

$$\begin{aligned} \Omega^j_i(X,Y) &= \mathrm{Rm}^j_i(X,Y) = \frac{1}{2}\left\langle (\nabla_g)^2_{X,Y}e_i, (\nabla_g)^2_{Y,X}e_i, e_j\right\rangle_g \\ = \ & d\omega^k_i(X,Y)\langle e_k, e_j\rangle_g + \frac{1}{2}\left(\omega^k_i(Y)\omega^\ell_k(X) - \omega^k_i(X)\omega^\ell_k(Y)\right)\langle e_\ell, e_j\rangle_g\end{aligned}$$

that (2.8.8) follows. □

For a surface $\mathcal{M}$, we have

$$d\omega^1 = \omega^2\wedge\omega^1_2, \quad d\omega^2 = \omega^1\wedge\omega^2_1, \quad \Omega^1_2 = d\omega^1_2.$$

The **Gauss curvature** is defined by

$$K_g := 2\mathrm{Rm}^1_2(e_1,e_2) = 2d\omega^1_2(e_1,e_2). \tag{2.8.9}$$

Note 2.40

Show that

$$d\omega^k(e_i,e_j) = \frac{1}{2}\omega^k_i(e_j) - \frac{1}{2}\omega^k_j(e_i).$$

Consequently,

$$\omega^k_i(e_j) = d\omega^i(e_j,e_k) + d\omega^j(e_i,e_k) - d\omega^k(e_j,e_i). \tag{2.8.10}$$

By definition, we have

$$d\omega^k(e_i,e_j) = -\frac{1}{2}\omega^k_l(e_i)\omega^l(e_j) + \frac{1}{2}\omega^k_\ell(e_j)\omega^\ell(e_i) = -\frac{1}{2}\omega^k_j(e_i) + \frac{1}{2}\omega^k_i(e_j).$$

Therefore

$$d\omega^i(e_j,e_k) + d\omega^j(e_i,e_k) - d\omega^k(e_j,e_i)$$

$$= \frac{1}{2}\left(\omega^i_j(e_k) - \omega^i_k(e_j) + \omega^j_i(e_k) - \omega^j_k(e_i) - \omega^k_j(e_i) + \omega^k_i(e_j)\right)$$

$$= \frac{1}{2}\left(\omega^k_i(e_j) + \omega^k_i(e_j)\right) = \omega^k_i(e_j).$$

Note that the similarity between this and the formula for the Christoffel symbols. ♣

Note 2.41

Prove

$$d\Omega^j_i = \omega^k_i \wedge \Omega^j_k - \omega^j_k \wedge \Omega^k_i. \tag{2.8.11}$$

From (2.8.8), we have

$$\begin{aligned} d\Omega^j_i &= d\left(d\omega^j_i - \omega^k_i \wedge \omega^j_k\right) = -d\omega^k_i \wedge \omega^j_k + \omega^k_i \wedge d\omega^j_k \\ &= \omega^k_i \wedge \left(d\omega^j_k - \omega^\ell_k \wedge \omega^j_\ell\right) - \omega^j_k \wedge \left(d\omega^k_i - \omega^\ell_i \wedge \omega^k_\ell\right) = \omega^k_i \wedge \Omega^j_k - \omega^j_k \wedge \Omega^k_i. \end{aligned}$$

The identity (2.8.11) implies that $\nabla_g \Omega^j_i = 0$ that is equivalent to the second Bianchi identity. ♣

Note 2.42

If $(\mathcal{M}^2, g)$ is a Riemannian surface and $u : \mathcal{M}^2 \to \mathbf{R}$ is a function, then

$$R_{e^u g} = e^{-u}\left(R_g - \Delta_g u\right). \tag{2.8.12}$$

♣

2.8.2 Curvature under conformal change of metric

Let $\tilde{g} := e^{2u} g$ and let $\{\omega^i\}_{1\le i\le m}$ be a local orthonormal coframe field for g. Then $\{\widetilde{\omega}^i\}_{1\le i\le m}$, where $\widetilde{\omega}^i := e^u\omega^i$, is a local orthonormal coframe field for $\tilde{g}$. Also, let $\{e_i\}_{1\le i\le m}$ and $\{\tilde{e}_i\}_{1\le i\le m}$ denote the orthonormal frame fields dual to $\{\omega^i\}_{1\le i\le m}$ and $\{\widetilde{\omega}\}_{1\le i\le m}$, respectively, so that $\tilde{e}_i = e^{-u}e_i$.

Note 2.43

Show that

$$\begin{aligned} \widetilde{\Omega}^j_i &= \Omega^j_i + \nabla_{e_k}\nabla_{e_i} u \cdot \omega^k \wedge \omega^j - \nabla_{e_k}\nabla_{e_j} u \cdot \omega^k \wedge \omega^i \\ &\quad + |\nabla_g u|^2_g\, \omega^i \wedge \omega^j + du \wedge \left[e_j(u)\omega^i - e_i(u)\omega^j\right], \end{aligned} \tag{2.8.13}$$

where ∇_g denotes the Levi-Civita connection with respect to the metric g. ♣

According to (2.8.6) and (2.8.13), we calculate the Ricci curvatures by

$$\mathrm{Rc}_{\tilde{g}}\left(\tilde{e}_\ell, \tilde{e}_i\right) = 2\sum_{k=1}^n \left\langle \widetilde{\Omega}^j_i\left(\tilde{e}_k, \tilde{e}_\ell\right)\tilde{e}_j, \tilde{e}_k\right\rangle_g = 2\sum_{k=1}^n \widetilde{\Omega}^k_i\left(\tilde{e}_k, \tilde{e}_\ell\right)$$

$$= e^{-2u}\left[\mathrm{Rc}_g(e_\ell, e_i) + (2-m)\nabla_{e_l}\nabla_{e_i} u - \delta_{\ell i}\Delta_g u + (2-m)\left(|\nabla_g u|^2_g\,\delta_{i\ell} - e_\ell(u)e_i(u)\right)\right]$$

so, the scalar curvatures of g and $\widetilde{g}$ are related by

$$R_{e^{2u}g} = e^{-2u}\left[R_g - 2(m-1)\Delta_g u - (m-2)(m-1)\,|\nabla_g u|_g^2\right]. \tag{2.8.14}$$

If we let $u := -f/m$, where $f \in C^\infty(\mathcal{M})$, then

$$R_{e^{-2f/m}g} = e^{2f/m}\left[R_g + 2\left(1-\frac{1}{m}\right)\Delta_g f - \left(1-\frac{2}{m}\right)\left(1-\frac{1}{m}\right)|\nabla_g f|_g^2\right]. \tag{2.8.15}$$

If we take $m \to \infty$ then

$$\lim_{m\to\infty} R_{e^{-2f/m}g} = R_g + 2\Delta_g f - |\nabla_g f|_g^2. \tag{2.8.16}$$

2.8.3 The Gauss-Bonnet formula

The **Gauss-Bonnet formula** says that the integral of the Gaussian curvature (which is the half of the scalar curvature) on a closed Riemannian surface $(\mathcal{M}, g)$ is equal to 2π times the Euler characteristic of $\mathcal{M}$.

Theorem 2.25. (Gauss-Bonnet)

If $(\mathcal{M}, g)$ is a closed oriented Riemannian surface, then

$$\frac{1}{2\pi}\int_{\mathcal{M}^2} K_g\,dA_g = \chi(\mathcal{M}). \tag{2.8.17}$$

♡

Let e_1, e_2 be a local positively oriented orthonormal basis for $T\mathcal{M}$ in an open set $\mathcal{U} \subset \mathcal{M}$ so that

$$dA_g = \omega^1 \wedge \omega^2.$$

The Gauss-Bonnet integrand is locally the exterior derivative of the connection 1-form $-\omega_1^2$:

$$K_g dA_g = 2d\omega_2^1(e_1, e_2)(\omega^1 \wedge \omega^2) = d\omega_2^1. \tag{2.8.18}$$

For higher dimension m, we have the following **Gauss-Bonnet-Chern formula**:

$$\chi(\mathcal{M}) = \frac{2(m-1)!}{2^m \pi^{m/2}\left(\frac{m}{2}-1\right)!}\int_{\mathcal{M}} K_g\,dV_g \tag{2.8.19}$$

where m is even and

$$K := \frac{1}{m!}\sum_{i_1,\cdots,i_m} \mathrm{sign}\,(i_1,\cdots,i_m)\,\Omega_{i_2}^{i_1}\wedge\Omega_{i_4}^{i_3}\wedge\cdots\wedge\Omega_{i_m}^{i_{m-1}}.$$

For $m = 4$, it was shown by Allendoerfer and Weil that

$$\chi(\mathcal{M}) = \frac{1}{8\pi^2}\int_{\mathcal{M}}\left(|\mathrm{Rm}_g|_g^2 - \left|\mathrm{Rc}_g - \frac{R_g}{4}g\right|_g^2\right)dV_g, \quad m = 4. \tag{2.8.20}$$

2.8.4 Moving frames adapted to hypersurfaces

Let $(\overline{\mathcal{M}}, \bar{g})$ be an m-dimensional Riemannian manifold and let $\overline{\nabla}$ denote the associated Levi-Civita connection. Given a hypersurface $\mathcal{M} \to \overline{\mathcal{M}}$, let $\{e_i\}_{1\le i\le m}$ be a moving frame in a neighborhood $\mathcal{U} \subset \overline{\mathcal{M}}$ of a point in $\mathcal{M}$. The connection 1-form ω_i^j of $(\overline{\mathcal{M}}, \bar{g})$ satisfy $\overline{\nabla}_X e_i = \omega_i^j(X)e_j$. We assume that the frame is **adopted** to $\mathcal{M}$, that is, $e_m := \nu$ is normal to

$\mathcal{M}$. The **first fundamental form** is defined by

$$g(X,Y) := \bar{g}(X,Y) \tag{2.8.21}$$

for $X, Y \in C^\infty(T\mathcal{M})$. This is also the induced Riemannian metric on the hypersurface $\mathcal{M}$. The **second fundamental form** is

$$h(X,Y) := \left\langle \overline{\nabla}_X \nu, Y \right\rangle_{\bar{g}} = \omega_n^j(X)\langle Y, e_j\rangle_{\bar{g}} \tag{2.8.22}$$

for X and Y tangent to $\mathcal{M}$. The second fundamental form measures the extrinsic geometry of the hypersurfce, e.g., how nonparallel the normal is. Let

$$h_{ij} := h(e_i, e_j) = \omega_m^j(e_i) \tag{2.8.23}$$

so that

$$\omega_m^j = \sum_{1\le i\le m-1} h_{ij}\omega^i. \tag{2.8.24}$$

The **mean curvature** is the trace of the second fundamental form:

$$H := \sum_{1\le i\le m} h(e_i, e_i) = \sum_{1\le i\le m} h_{ii}. \tag{2.8.25}$$

The induced Levi-Civita connection ∇ of g satisfies

$$\nabla_X e_i := \left(\overline{\nabla}_X e_i\right)^\top = \sum_{1\le j\le m-1} \omega_i^j(X) e_j, \tag{2.8.26}$$

where $\top$ denotes the tangent component of a vector. Thus $\{\omega_i^j\}_{1\le i,j\le m-1}$ are the connection 1-forms of $(\mathcal{M}, g)$.

$$\overline{\Omega}_i^j = d\omega_i^j - \sum_{1\le k\le m} \omega_i^k \wedge \omega_k^j, \quad i,j = 1,\cdots,m, \tag{2.8.27}$$

$$\overline{\Omega}_i^j = d\omega_i^j - \sum_{1\le k\le m-1} \omega_i^k \wedge \omega_k^j, \quad i,j = 1,\cdots,m-1. \tag{2.8.28}$$

Thus, for $i,j = 1,\cdots,m-1$, we have $\Omega_i^j = \overline{\Omega}_i^j + \omega_i^m \wedge \omega_m^j = \overline{\Omega}_i^j - h_{ik}h_{j\ell}\omega^k \wedge \omega^\ell$. Hence, we obtain the **Gauss equation**

$$R_{ijk\ell} = \bar{R}_{ijk\ell} + h_{i\ell}h_{jk} - h_{ik}h_{j\ell}. \tag{2.8.29}$$

Lemma 2.14

One has

$$R_{jk} = \bar{R}_{jk} - \bar{R}_{mjkm} + Hh_{jk} - g^{i\ell}h_{j\ell}h_{ik}, \tag{2.8.30}$$

$$R_g = \bar{R}_{\bar{g}} - 2R_{mm} + H^2 - |h|_g^2. \tag{2.8.31}$$

♡

Proof. Calculate $R_{jk} = g^{i\ell}R_{ijk\ell} = \bar{R}_{jk} - \bar{R}_{mjkm} + Hh_{jk} - g^{i\ell}h_{j\ell}h_{ik}$. Hence $R_g = g^{jk}R_{jk} = \bar{R}_{\bar{g}} - 2\bar{R}_{mm} + H^2 - |h|_g^2$. □

For $j = 1,\cdots,m-1$, we have

$$\overline{\Omega}_m^j = d\omega_m^j - \sum_{1\le k\le m-1} \omega_m^k \wedge \omega_k^j. \tag{2.8.32}$$

The $(1,1)$-tensor field $W := \sum_{j=1}^{m-1} \omega_m^j e_j$ is the **Weingarten map**. Considering W as a 1-form with values in $T\mathcal{M}$, we have

$$\nabla_g W = \sum_{1 \le j \le m-1} \overline{\Omega}_m^j e_j, \tag{2.8.33}$$

which is a 2-form with values in $T\mathcal{M}$.

Note 2.44. (Codazzi equations)

Show that for X, Y, Z tangent to $\mathcal{M}$,

$$((\nabla_g)_X h)(Y,Z) - ((\nabla_g)_Y h)(X,Z) = -\left\langle \mathrm{Rm}_{\bar{g}}(X,Y)Z, \nu \right\rangle_{\bar{g}}. \tag{2.8.34}$$

♣

Consider a smooth function $f : \mathcal{M} \to \mathbf{R}$ on a $(m-1)$-dimensional manifold. For any regular value $c \in \mathbf{R}$ of f (i.e., $\nabla_g f(x) \neq 0$ for all $x \in \mathcal{M}$ such that $f(x) = c$), the level set $f^{-1}(c)$ is a smooth hypersurface by the implicit function theorem. The second fundamental form of the level set $f^{-1}(c)$ is the given by

$$h(V,W) := \frac{\mathrm{Hess}_g(f)(V,W)}{|\nabla_g f|_g}. \tag{2.8.35}$$

Indeed, $\nu := \nabla_g f / |\nabla_g f|_g$ is a unit normal vector for $f^{-1}(c)$. For V, W tangent to $f^{-1}(c)$ we have

$$\begin{aligned} h(V,W) &= \langle (\nabla_g)_V \nu, W \rangle_g = \left\langle (\nabla_g)_V \frac{\nabla_g f}{|\nabla_g f|_g}, W \right\rangle_g \\ &= \frac{1}{|\nabla_g f|_g} \langle (\nabla_g)_V \nabla_g f, W \rangle_g = \frac{1}{|\nabla_g f|_g} \mathrm{Hess}_g(f)(V,W) \end{aligned}$$

since $\langle \nabla_g f, W \rangle_g = 0$. In particular, if f is (strictly) convex $(\nabla_g^2 f \ge 0)$ $\nabla_g^2 f > 0$, then any smooth hypersurface $f^{-1}(c)$ is (strictly) convex $(h \ge 0)$ $h > 0$.

2.9 Variation of arc length, energy and area

Introduction

- *First variation of arc length*
- *Second variation of arc length*
- *Long stable geodesics*
- *Jacobi fields in relation to the index form*
- *First and second variation of energy*
- *First and second variation of area*

Let $(\mathcal{M}, g)$ be an m-dimensional Riemannian manifold.

2.9.1 First variation of arc length

Given a path $\gamma : [a,b] \to \mathcal{M}$, its **length** is defined by

$$L_g(\gamma) := \int_a^b |\dot{\gamma}(u)|_g du. \tag{2.9.1}$$

The **distance function** is defined by

$$d_{g,p}(x) := d_g(p,x) := \inf_{\gamma} L_g(\gamma), \tag{2.9.2}$$

where the infimum is taken over all paths $\gamma : [0,1] \to \mathcal{M}$ with $\gamma(0) = p$ and $\gamma(1) = x$. A geodesic segment is **minimal** if its length is equal to the distance between the two endpoints.

Let $\gamma_r : [a,b] \to \mathcal{M}$, $r \in \mathfrak{r} \subset \mathbf{R}$, be a 1-parameter family of paths. We define the map $\Upsilon : [a,b] \times \mathfrak{r} \to \mathcal{M}$ by

$$\Upsilon(s,r) := \gamma_r(s). \tag{2.9.3}$$

We define the vector fields R and S along Υ by

$$R := \Upsilon_* \left(\frac{\partial}{\partial r}\right), \quad S := \Upsilon_* \left(\frac{\partial}{\partial s}\right). \tag{2.9.4}$$

We call R the **variation vector field** and S the **tangent vector field**. More precisely, the map Υ induces a map between tangent spaces at each point $(s,r) \in [a,b] \times \mathfrak{r}$:

$$\Upsilon_{*,(s,r)} : T_{(s,r)}\left([a,b] \times \mathfrak{r}\right) \longrightarrow T_{\gamma_r(s)}\mathcal{M}.$$

Then

$$\begin{aligned} S(\gamma_r(s)) &= S_{\gamma_r(s)} = \Upsilon_{*,(s,r)}\left(\left(\left.\frac{\partial}{\partial s}\right|_s\right), 0\right) = \Upsilon_{*,(s,r)}\left(\left.\frac{\partial}{\partial s}\right|_s\right), \\ R(\gamma_r(s)) &= R_{\gamma_r(s)} = \Upsilon_{*,(s,r)}\left(0, \left(\left.\frac{\partial}{\partial r}\right|_r\right)\right) = \Upsilon_{*,(s,r)}\left(\left.\frac{\partial}{\partial r}\right|_r\right). \end{aligned}$$

Thus,

$$S, R \in C^\infty\left(\Upsilon\left([a,b] \times \mathfrak{r}\right), T\mathcal{M}\Big|_{\Upsilon([a,b]\times\mathfrak{r})}\right)$$

By the above notation, we have

$$\begin{aligned} \left.\frac{\partial}{\partial r}\right|_r \langle S(\gamma_r(s)), S(\gamma_r(s))\rangle_g &= \left.\frac{\partial}{\partial r}\right|_r \langle S(\Upsilon(s,r)), S(\Upsilon(s,r))\rangle_g \\ &= \left.\frac{\partial}{\partial r}\right|_r \left(\left(\langle S,S\rangle_g \circ \Upsilon\right)(s,r)\right) = \left(R\langle S,S\rangle_g\right)(\gamma_r(s)). \end{aligned} \tag{2.9.5}$$

We also note that

$$S(\gamma_0(s)) = \dot{\gamma}_0(s). \tag{2.9.6}$$

The length of γ_r is given by

$$L_g(\gamma_r) := \int_a^b |S(\gamma_r(s))|_g \, ds. \tag{2.9.7}$$

Lemma 2.15. (First variation of arc length)

Suppose $0 \in \mathfrak{r}$. If γ_0 is parametrized by arc length, that is, $|S(\gamma_0(s))|_g \equiv 1$, then

$$\left.\frac{d}{dr}\right|_{r=0} L_g(\gamma_r) = -\int_a^b \langle R, \nabla_S S\rangle_g (\gamma_0(s))\, ds + \left[\langle R, S\rangle_g(\gamma_0(s))\right]\Big|_a^b. \tag{2.9.8}$$

♡

Proof. Calculate (use the formula (2.9.5))

$$\begin{aligned}\frac{d}{dr}L_g(\gamma_r) &= \frac{1}{2}\int_a^b |S(\gamma_r(s))|_g^{-1}\frac{\partial}{\partial r}\Big|_r \langle S(\gamma_r(s)), S(\gamma_r(s))\rangle_g\, ds\\ &= \frac{1}{2}\int_a^b |S(\gamma_r(s))|_g^{-1}\left(R\langle S,S\rangle_g\right)(\gamma_r(s))\, ds.\end{aligned}$$

By the assumption that $|S(\gamma_0(s))|_g \equiv 1$, we conclude that

$$\frac{d}{dr}\Big|_{r=0} L_g(\gamma_r) = \frac{1}{2}\int_a^b \left(R\langle S,S\rangle_g\right)(\gamma_0(s))\, ds = \int_a^b \langle S, (\nabla_g)_R S\rangle_g (\gamma_0(s))\, ds.$$

However,

$$(\nabla_g)_R S - (\nabla_g)_S R = [R,S] = \Upsilon_*\left(\left[\frac{\partial}{\partial r}, \frac{\partial}{\partial s}\right]\right) = 0$$

which implies that

$$\frac{d}{dr}\Big|_{r=0} L_g(\gamma_r) = \int_a^b \langle S, (\nabla_g)_S R\rangle_g (\gamma_0(s))\, ds$$

Integrating by parts yields the formula (2.9.8). □

Corollary 2.8

If $\gamma_r : [0,b] \to \mathcal{M}$, $r \in \mathfrak{r} \subset \mathbf{R}$, is a 1-parameter family of paths emanating from a fixed point $p \in \mathcal{M}$ (i.e., $\gamma_r(0) = p$) and γ_0 is a geodesic parametrized by arc length, then

$$\frac{d}{dr}\Big|_{r=0} L_g(\gamma_r) = \left\langle \frac{\partial}{\partial r}\Big|_{r=0}\gamma_r(b), \dot{\gamma}_0(b)\right\rangle_g. \tag{2.9.9}$$

♡

Note 2.45

If we do not assume γ_0 is parametrized by arc length, then we have

$$\begin{aligned}\frac{d}{dr}\Big|_{r=0} L_g(\gamma_r) &= -\int_a^b \left\langle R, (\nabla_g)_S\left(\frac{S}{|S|_g}\right)\right\rangle_g (\gamma_0(s))ds \\ &+ \left[\left\langle R, \frac{S}{|S|_g}\right\rangle_g (\gamma_0(s))\right]\Big|_a^b.\end{aligned} \tag{2.9.10}$$

Hence, among all paths fixing two endpoints, the critical points of the length functional are the geodesics γ, which satisfy

$$(\nabla_g)_{\dot\gamma}\left(\frac{\dot\gamma}{|\dot\gamma|_g}\right) = 0.$$

♣

2.9.2 Second variation of arc length

Now we suppose that we have a 2-parameter family of paths $\gamma_{q,r} : [a,b] \longrightarrow \mathcal{M}$ with $q \in \mathfrak{q} \subset \mathbf{R}$ and $r \in \mathfrak{r} \subset \mathbf{R}$. Define $\Phi : [a,b]\times\mathfrak{q}\times\mathfrak{r} \to \mathcal{M}$ by

$$\Phi(s,q,r) := \gamma_{q,r}(s). \tag{2.9.11}$$

The map Φ induces a map between tangent spaces at each point $(s,q,r) \in [a,b]\times\mathfrak{q}\times\mathfrak{r}$: $\Phi_{*,(s,q,r)} : T_{(s,q,r)}([a,b]\times\mathfrak{q}\times\mathfrak{r}) \longrightarrow T_{\gamma_{q,r}(s)}\mathcal{M}$. We define vector fields Q, R, and S along Φ

as follows:

$$\begin{aligned}
S(\gamma_{q,r}(s)) &= S_{\gamma_{q,r}(s)} = \Phi_{*,(s,q,r)}\left(\left(\left.\frac{\partial}{\partial s}\right|_s\right),0,0\right) = \Phi_{*,(s,q,r)}\left(\left.\frac{\partial}{\partial s}\right|_s\right),\\
Q(\gamma_{q,r}(s)) &= Q_{\gamma_{q,r}(s)} = \Phi_{*,(s,q,r)}\left(0,\left(\left.\frac{\partial}{\partial s}\right|_q\right),0\right) = \Phi_{*,(s,q,r)}\left(\left.\frac{\partial}{\partial s}\right|_q\right),\\
R(\gamma_{q,r}(s)) &= R_{\gamma_{q,r}(s)} = \Phi_{*,(s,q,r)}\left(0,0,\left(\left.\frac{\partial}{\partial s}\right|_r\right)\right) = \Phi_{*,(s,q,r)}\left(\left.\frac{\partial}{\partial r}\right|_r\right).
\end{aligned}$$

Then $S, R, Q \in C^\infty(\Phi([a,b]\times\mathfrak{q}\times\mathfrak{r}), T\mathcal{M}|_{\Phi([a,b]\times\mathfrak{q}\times\mathfrak{r})})$.

Lemma 2.16. (Second variation of arc length)

Suppose $0\in\mathfrak{q}$ *and* $0\in\mathfrak{r}$*. If* $\gamma_{0,0}$ *is parametrized by arc length, then*

$$\begin{aligned}
&\left.\frac{\partial^2}{\partial q\partial r}\right|_{(q,r)=(0,0)} L_g(\gamma_{q,r})\\
=\;& \int_a^b \left(\langle(\nabla_g)_S Q,(\nabla_g)_S R\rangle_g - \langle(\nabla_g)_S Q, S\rangle_g \langle(\nabla_g)_S R, S\rangle_g\right)(\gamma_{0,0}(s))\,ds\\
&- \int_a^b \langle \mathrm{Rm}_g(Q,S)S,R\rangle_g(\gamma_{0,0}(s))\,ds \qquad (2.9.12)\\
&- \int_a^b \langle(\nabla_g)_Q R,(\nabla_g)_S S\rangle_g(\gamma_{0,0}(s))\,ds + \left[\langle(\nabla_g)_Q R, S\rangle_g(\gamma_{0,0}(s))\right]\Big|_a^b.
\end{aligned}$$

♡

Proof. Differentiating the first variation of arc length we have

$$\begin{aligned}
\left.\frac{\partial^2}{\partial q\partial r}\right|_{(q,r)=(0,0)} L_g(\gamma_{q,r}) &= \left.\frac{\partial}{\partial q}\right|_{(q,r)=(0,0)} \int_a^b \left\langle \frac{S}{|S|_g},(\nabla_g)_S R\right\rangle_g (\gamma_{q,r}(s))\,ds\\
&= \left.\frac{\partial}{\partial q}\right|_{(q,r)=(0,0)} \int_a^b \left(Q\left\langle \frac{S}{|S|_g},(\nabla_g)_S R\right\rangle_g\right)(\gamma_{q,r}(s))\,ds
\end{aligned}$$

$$= \int_a^b \left(\left\langle\frac{S}{|S|_g},(\nabla_g)_Q(\nabla_g)_S R\right\rangle_g + \left\langle(\nabla_g)_Q\left(\frac{S}{|S|_g}\right),(\nabla_g)_S R\right\rangle_g\right)(\gamma_{q,r}(s))\,ds\Big|_{(q,r)=(0,0)}$$

$$\begin{aligned}
&= \int_a^b \langle S,(\nabla_g)_S(\nabla_g)_Q R + \mathrm{Rm}_g(Q,S)R\rangle_g(\gamma_{0,0}(s))\,ds\\
&+ \int_a^b \left\langle \nabla_Q S - \langle S,(\nabla_g)_Q S\rangle_g S,(\nabla_g)_S R\right\rangle_g(\gamma_{0,0}(s))\,ds
\end{aligned}$$

where we use the identity that

$$(\nabla_g)_Q\left(\frac{S}{|S|_g}\right) = |S|_g^{-1}(\nabla_g)_Q S - |S|_g^{-3}\langle S,(\nabla_g)_Q S\rangle_g S. \qquad (2.9.13)$$

Then the result follows from an integration by parts. □

Corollary 2.9

If γ_r *is a* 1*-parameter family of piecewise smooth paths with fixed endpoints and such that* γ_0 *is a geodesic parametrized by arc length, then*

$$\left.\frac{d^2}{dr^2}\right|_{r=0} L_g(\gamma_r) = \int_a^b \left(\left|((\nabla_g)_S R)^\perp\right|_g^2 - \langle\mathrm{Rm}_g(R,S)S,R\rangle_g\right)(\gamma_0(s))\,ds, \qquad (2.9.14)$$

where $((\nabla_g)_S R)^\perp$ *is the projection of* $(\nabla_g)_S R$ *onto* $S^\perp$*, i.e.,* $((\nabla_g)_S R)^\perp := (\nabla_g)_S R - \langle(\nabla_g)_S R, S\rangle_g S$.

♡

Proof. It suffices to show that $|((\nabla_g)_S R)^\perp|_g^2 = \langle(\nabla_g)_S Q, ((\nabla_g)_S R)^\perp\rangle_g$ which is equivalent to prove $\left\langle\langle(\nabla_g)_S R, S\rangle_g S, ((\nabla_g)_S R)^\perp\right\rangle_g = 0$. By the definition, the left side of above equals $\left\langle\langle(\nabla_g)_S R, S\rangle_g S, (\nabla_g)_S R - \langle(\nabla_g)_S R, S\rangle_g S\right\rangle_g = \left|\langle(\nabla_g)_S R, S\rangle_g\right|_g^2 - \left|\langle(\nabla_g)_S R, S\rangle_g\right|_g^2 |S|_g^2$ which is zero, since γ_0 is parametrized by arc length. □

A geodesic is **stable** if the second variation of arc length, with respect to variation vector fields which vanish at the endpoints, is nonnegative.

Corollary 2.10

If, in addition, $(\mathcal{M}, g)$ *has nonnegative sectional curvature and the paths* γ_r *are smooth and closed, then*

$$\frac{d^2}{dr^2}\bigg|_{r=0} L_g(\gamma_r) \geq 0.$$

That is, any smooth closed geodesic γ_0 *is stable.*

♡

Theorem 2.26. (Synge)

If $(\mathcal{M}, g)$ *is an even-dimensional, orientable, closed Riemannian manifold with positive sectional curvature, then* $\mathcal{M}$ *is simply-connected.*

♡

If $\gamma_r : [0, b] \to \mathcal{M}$ is a 1-parameter family of paths, $r \in (-\epsilon, \epsilon)$, γ_0 is a unit speed geodesic, and $R(\gamma_0(0)) = 0$, then

$$\begin{aligned}\frac{d^2}{dr^2}\bigg|_{r=0} L_g(\gamma_r) &= \langle(\nabla_g)_R R, S\rangle_g (\gamma_0(b)) \\ &= \int_0^b \left(\left|((\nabla_g)_S R)^\perp\right|_g^2 - \langle \mathrm{Rm}_g(R, S)S, R\rangle_g\right)(\gamma_0(s))\, ds. \quad (2.9.15)\end{aligned}$$

Given a $V \in T_{\gamma_0(b)}\mathcal{M}$, we extend V along γ_0 by defining

$$\overline{V}(\gamma_0(s)) := \frac{b}{s}\mathcal{R}(\gamma_0(s)), \tag{2.9.16}$$

where $\overline{V}(\gamma_0(s))$ is the parallel translation of V along γ_0, i.e., $\left((\nabla_g)_S \overline{V}\right)(\gamma_0(s)) = 0$. Note that $V = \overline{V}(\gamma_0(b)) = R(\gamma_0(b))$. Then $((\nabla_g)_S R)(\gamma_0(s)) = (\nabla_g)_{S(\gamma_0(s))}(\frac{s}{b}\overline{V}(\gamma_0(s))) = \frac{1}{b}\overline{V}(\gamma_0(s))$ so that $(\nabla_g)_S R = \frac{1}{b}\overline{V}$. Since $\overline{V}(\gamma_0(s))$ is the parallel translation of V along γ_0, it follows that

$$\left|((\nabla_g)_S R)^\perp\right|_g^2 = \frac{1}{b^2}\left|\overline{V}^\perp\right|_g^2 = \frac{1}{b^2}\left|V^\perp\right|_g^2$$

where $\overline{V}^\perp := \overline{V} - \langle\overline{V}, S\rangle_g S$ and $V^\perp := V - \langle V, S(\gamma_0(b))\rangle_g S(\gamma_0(b)) = V - \langle V, \dot{\gamma}_0(b)\rangle_g \dot{\gamma}_0(b)$. Hence

$$\int_0^b \left|((\nabla_g)_S R)^\perp\right|_g^2 (\gamma_0(s))\, ds = \int_0^b \frac{1}{b^2}\left|V^\perp\right|_g^2 ds = \frac{1}{b}\left|V^\perp\right|_g^2$$

and

$$\begin{aligned}\frac{d^2}{dr^2}\Big|_{r=0} L_g(\gamma_r) &= \langle (\nabla_g)_R R, S\rangle_g (\gamma_0(b)) \\ &= \frac{1}{b}\left|V^\perp\right|_g^2 - \int_0^b \langle \mathrm{Rm}_g(R,S)S, R\rangle_g (\gamma_0(s))\,ds. \end{aligned} \tag{2.9.17}$$

Lemma 2.17

If $\gamma_r : [0,b] \to \mathcal{M}$, $r \in (-\epsilon, \epsilon)$, is a 1-parameter family of paths emanating from a fixed point $p \in \mathcal{M}$, i.e., $\gamma_r(0) = p$ and γ_0 is a geodesic parametrized by arc length, then

$$d_g(p, \beta_V(r)) \le L_g(\gamma_r), \quad d_g(p, \beta_V(0)) = L_g(\gamma_0) \tag{2.9.18}$$

where $\beta_V : (-\epsilon, \epsilon) \to \mathcal{M}$ where $\beta_V(r) := \gamma_r(b)$ so that $\dot\beta_V(0) = V \in T_{\gamma_0(b)}\mathcal{M}$. Thus, the function $r \mapsto d_g(p, \beta_V(r))$ is a lower support function for $r \mapsto L_g(\gamma_r)$ at $r = 0$. ♡

Proof. According to **Corollary 2.8**, we have

$$\frac{d}{dr}\Big|_{r=0} L_g(\gamma_r) = \left\langle \frac{\partial}{\partial r}\Big|_{(s,r)=(b,0)} \gamma_r(s), \frac{\partial}{\partial s}\Big|_{(s,r)=(b,0)} \gamma_r(s)\right\rangle_{g(\gamma_0(b))} = 0.$$

Hence $d_g(p, \beta(0)) = L_g(\gamma_0)$. □

Definition 2.2

Suppose that $u \in C^0(\mathcal{M})$ and $V \in T_p\mathcal{M}$. Let $\beta_V : (-\epsilon, \epsilon) \to \mathcal{M}$ be the constant speed geodesic with $\beta_V(0) = p$ and $\dot\beta_V(0) = V$. If $v : (-\epsilon, \epsilon) \to \mathbf{R}$ is a C^2-function such that

$$u(\beta_V(r)) \le v(r), \quad r \in (-\epsilon, \epsilon), \quad u(\beta_V(0)) = v(0),$$

then we say that

$$(\nabla_g)_V(\nabla_g)_V u \le v''(0) \tag{2.9.19}$$

in the sense of support functions with respect to p **and** V. *If (2.9.19) holds for all p and V, then we say (2.9.19)* **in the sense of support functions**. ♣

Note 2.46

Show that if $u : \mathcal{M} \to \mathbf{R}$ satisfies $\nabla_g^2 u \le 0$ in the sense of support functions, then u is concave; that is, for every unit speed geodesic $\beta_V : [a,b] \to \mathcal{M}$ we have

$$u\left(\beta_V((1-s)a + sb)\right) \ge (1-s)u(\beta_V(a)) + su(\beta_V(b)), \quad \textit{for all } s \in [0,1].$$ ♣

Returning to our discussion, we assume β_V is a geodesic so that $(\nabla_g)_Q Q = 0$. If the sectional curvatures are nonnegative and $d_{g,p}(x) := d_g(p, x)$ is the distance function, then

$$(\nabla_g)_V(\nabla_g)_V d_{g,p} \le \frac{d^2}{dr^2}\Big|_{r=0} L_g(\gamma_r) \le \left(\frac{1}{d_{g,p}}|V|_g^2\right)(\gamma_0(b)) \tag{2.9.20}$$

in the sense of support functions, since

$$b = \int_0^b |S(\gamma_0(s))|_g\,ds = L_g(\gamma_0) = d_g(\gamma_0(b), p) = d_{g,p}(\gamma_0(b))$$

and

$$\begin{aligned}\langle V, S(\gamma_0(b))\rangle_g &= \left\langle \frac{\partial}{\partial r}\Big|_{r=0}\gamma_r(b), \frac{\partial}{\partial s}\Big|_{s=b}\gamma_0(s)\right\rangle_g \\ &= \left\langle \frac{\partial}{\partial r}\Big|_{(s,r)=(b,0)}\gamma_r(s), \frac{\partial}{\partial s}\Big|_{(s,r)=(b,0)}\gamma_r(s)\right\rangle_g = 0.\end{aligned}$$

The inequality (2.9.20) is a special case ($K = 0$) of the Hessian comparison theorem. Note that this inequality holds in the usual C^2-sense at points where $d_{g,p}$ are smooth.

Lemma 2.18

Assuming the sectional curvature is nonnegative, one has

$$(\nabla_g)_V(\nabla_g)_V d^2_{g,p} \le 2|V|^2_g. \tag{2.9.21}$$

♡

Proof. Calculate

$$\frac{d^2}{dr^2}\Big|_{r=0} L_g(\gamma_r)^2 = \frac{d}{dr}\Big|_{r=0}\left(2L_g(\gamma_r)\cdot\frac{d}{dr}L_g(\gamma_r)\right)$$

$$= 2\left(\frac{d}{dr}\Big|_{r=0}L_g(\gamma_r)\right)^2 + 2L_g(\gamma_r)\cdot\frac{d^2}{dr^2}\Big|_{r=0}L_g(\gamma_r) \le 2\left(\frac{d}{dr}\Big|_{r=0}L_g(\gamma_r)\right)^2 + 2b\cdot\frac{1}{b}\left|V^\perp\right|^2_{g(\gamma_0(b))}$$

$$= 2\left|V^\perp\right|^2_{g(\gamma_0(b))} + 2\left(\frac{d}{dr}\Big|_{r=0}L_g(\gamma_r)\right)^2 = 2\left|V^\perp\right|^2_g = 2|V|^2_g.$$

So, $(\nabla_g)_V(\nabla_g)_V d^2_{g,p} \le 2|V|^2_g$. Equivalently, $\nabla^2_g d^2_{g,p} \le 2g(\gamma_0(b))$ for any $V \in T_{\gamma_0(b)}\mathcal{M}$. □

2.9.3 Long stable geodesics

Let $\gamma : [0, \bar{s}] \to \mathcal{M}$ be a stable unit speed geodesic in an m-dimensional Riemannian manifold $(\mathcal{M}, g)$ with $\mathrm{Rc}_g \le (m-1)K$ in $B_g(\gamma(0), r)$ and $B_g(\gamma(\bar{s}), r)$ where $K > 0$ and $2r < \bar{s}$. Let $\{E_i\}_{1\le i\le m-1}$ be a parallel orthonormal frame along γ perpendicular to $\dot\gamma$. By the second variation of arc length, we have

$$\begin{aligned} 0 &\le \sum_{1\le i\le m-1}\int_0^{\bar{s}}\left(\left|((\nabla_g)_{\dot\gamma}(\varphi E_i(\gamma)))^\perp\right|^2_g - \langle \mathrm{Rm}_g(\varphi E_i(\gamma), \dot\gamma)\dot\gamma, \varphi E_i(\gamma)\rangle_g\right) ds \\ &= \int_0^{\bar{s}}\left[(m-1)\left(\frac{d\varphi}{ds}\right)^2 - \varphi^2\mathrm{Rc}_g(\dot\gamma,\dot\gamma)\right] ds\end{aligned}$$

for any function $\varphi : [0, \bar{s}] \to \mathbf{R}$. Consider the piecewise smooth function

$$\varphi(s) := \begin{cases} \frac{s}{r}, & 0 \le s \le r, \\ 1, & r < s < \bar{s} - r, \\ \frac{\bar{s}-s}{r}, & \bar{s} - r \le s \le \bar{s}. \end{cases} \tag{2.9.22}$$

We then have

$$\begin{aligned}\int_0^{\bar{s}}\mathrm{Rc}_g(\dot\gamma,\dot\gamma)\,ds &\le \frac{2(m-1)}{r} + \int_0^{\bar{s}}(1-\varphi^2)\mathrm{Rc}_g(\dot\gamma,\dot\gamma)\,ds \\ &= \frac{2(m-1)}{r} + \int_0^{r}(1-\varphi^2)\mathrm{Rc}_g(\dot\gamma,\dot\gamma)\,ds + \int_{\bar{s}-r}^{\bar{s}}(1-\varphi^2)\mathrm{Rc}_g(\dot\gamma,\dot\gamma)\,ds\end{aligned}$$

$$\leq \quad \frac{2(m-1)}{r} + (m-1)K \cdot \frac{4r}{3} \quad \leq \quad 2(m-1)\left(\frac{1}{r} + Kr\right).$$

Proposition 2.18

If $\gamma : [0, L] \to \mathcal{M}$ is a stable unit speed geodesic in a Riemannian m-manifold with

$$\mathrm{Rc}_g \leq (m-1)K, \quad \textit{in} \quad B_g\left(\gamma(0), 1/\sqrt{K}\right) \cup B_g\left(\gamma(L), 1/\sqrt{K}\right),$$

where $K > 0$, then

$$\int_0^L \mathrm{Rc}_g(\dot\gamma, \dot\gamma) ds \leq 4(m-1)\sqrt{K}.$$

♡

The above computation is useful in obtaining an estimate for the rate of change of the distance function under the Ricci flow.

2.9.4 Jacobi fields in relation to the index form

Let $\gamma_r : [a, b] \to \mathcal{M}, r \in \mathfrak{r} \subset \mathbf{R}$, be a 1-parameter family of paths. Assume γ_0 is a geodesic. Then $S(\gamma_0(s)) = \dot\gamma_0(s)$, and hence $((\nabla_g)_S S)(\gamma_0(s)) = 0$. For the variation vector field R, we have

$$0 = (\nabla_g)_R(\nabla_g)_S S = (\nabla_g)_S(\nabla_g)_R S + \mathrm{Rm}_g(R, S)S = (\nabla_g)_S(\nabla_g)_S R + \mathrm{Rm}_g(R, S)S.$$

Thus

$$(\nabla_g)_{\dot\gamma_0(s)}(\nabla_g)_{\dot\gamma_0(s)} R(\gamma_0(s)) + \mathrm{Rm}_g\left(R(\gamma_0(s)), \dot\gamma_0(s)\right)\dot\gamma_0(s) = 0.$$

A **Jaboci field** J is a variation of geodesic and satisfies the **Jabobi equation**

$$(\nabla_g)_S(\nabla_g)_S J + \mathrm{Rm}_g(J, S)S = 0. \tag{2.9.23}$$

Given $p \in \mathcal{M}$ and $V, W \in T_p\mathcal{M}$, we define a 1-parameter family of geodesics $\gamma_r : [0, \infty) \to \mathcal{M}$ by

$$\gamma_r(s) := \exp_p\left(s(V + rW)\right) = \gamma_{V+rW}(s). \tag{2.9.24}$$

We may define a Jacobi field $J_{V,W}$ along $\gamma_0 = \gamma_V$ by

$$J_{V,W}(s) := \left.\frac{\partial}{\partial r}\right|_{r=0} \gamma_{V+rW}(s). \tag{2.9.25}$$

Definition 2.3

A point $x \in \mathcal{M}$ is a **conjugate point** *of $p \in \mathcal{M}$ if x is a singular value of $\exp_p : T_p\mathcal{M} \to \mathcal{M}$. That is, $x = \exp_p(V)$, for some $V \in T_p\mathcal{M}$, where $(\exp_p)_{*,V} : T_V(T_p\mathcal{M}) \to T_{\exp_p(V)}\mathcal{M}$ is singular (i.e., has nontrivial kernel).*

♣

Note 2.47

(1) Equivalently, $\gamma(r)$ is a conjugate point to p along γ if there is a nontrivial Jacobi field along γ vanishing at the endpoints.

(2) Given a geodesic $\gamma : [0, L] \to \mathcal{M}$ without conjugate points and vectors $A \in T_{\gamma(0)}\mathcal{M}$

and $B \in T_{\gamma(L)}\mathcal{M}$ with $\langle A, S\rangle_g = \langle B, S\rangle_g = 0$, there exists a unique Jacobi field J with $J(0) = A$ and $J(L) = B$. ♣

If $\gamma : [a, b] \to \mathcal{M}$ is a path and V and W are vector fields along γ, we define the **index form** of V and W by

$$\begin{aligned} I_{g,\gamma}(V, W) \quad &:= \int_a^b \Big(\langle (\nabla_g)_S V, (\nabla_g)_S W\rangle_g - \langle (\nabla_g)_S V, S\rangle_g \langle (\nabla_g)_S W, S\rangle_g \\ &- \langle \mathrm{Rm}_g(V, S)S, W\rangle_g \Big)\, ds, \end{aligned} \tag{2.9.26}$$

where $S := \frac{\partial}{\partial s}|_{s=0}\gamma_s$ and γ_s is a 1-parameter family of γ.

If $\gamma_{q,r}$ is a 1-parameter family of paths with fixed endpoints and if $\gamma_{0,0}$ is a unit speed geodesic, then by **Lemma 2.16**,

$$\frac{\partial^2}{\partial q \partial r}\Big|_{(q,r)=(0,0)} L_g(\gamma_{q,r}) = I_{g,\gamma_{0,0}}(Q, R).$$

Lemma 2.19. (Index lemma)

Suppose $\gamma : [0, L] \to \mathcal{M}$ is a geodesic without conjugate points. In the space $\mathrm{Vect}_{A,B}(\gamma)$ of vector fields X along γ with $\langle X, S\rangle_g \equiv 0$, $X(0) = A$ and $X(L) = B$, the Jacobi field minimizes the (modified) index form:

$$\mathcal{I}_{g,\gamma}(X) := \int_0^L \Big(|(\nabla_g)_S X|_g^2 - \langle \mathrm{Rm}_g(X, S)S, X\rangle_g \Big)\, ds. \tag{2.9.27}$$

♡

Proof. If X and Y are vector field along γ, then

$$\begin{aligned} \mathcal{I}_{g,\gamma}(X + tY) \;&=\; \int_0^L \Big(|(\nabla_g)_S(X + tY)|_g^2 - \langle \mathrm{Rm}_g(X + tY, S)S, X + tY\rangle_g \Big)\, ds \\ &= \int_0^L \Big(|(\nabla_g)_S X + t(\nabla_g)_S Y|_g^2 - \langle \mathrm{Rm}_g(X, S)S, X\rangle_g \Big)\, ds \\ &- \int_0^L \Big(2t \langle \mathrm{Rm}_g(X, S)S, Y\rangle_g + t^2 \langle \mathrm{Rm}_g(Y, S)S, Y\rangle_g \Big)\, ds \end{aligned}$$

so that

$$\frac{1}{2}\frac{d}{dt}\Big|_{t=0} \mathcal{I}_{g,\gamma}(X + tY) = \int_0^L \Big(\langle (\nabla_g)_S X, (\nabla_g)_S Y\rangle_g - \langle \mathfrak{Rm}_g(X, S)S, Y\rangle_g \Big)\, ds.$$

(Note that the tangent space $T_X \mathrm{Vect}_{A,B}(\gamma)$ is the space of all vector fields along γ which vanish at the endpoints) If furthermore Y satisfies $Y(0) = Y(L) = 0$, using

$$\frac{d}{ds}\langle Y, (\nabla_g)_S X\rangle_g = \langle (\nabla_g)_S Y, (\nabla_g)_S X\rangle_g + \langle Y, (\nabla_g)_S(\nabla_g)_S X\rangle_g,$$

we obtain

$$\frac{1}{2}\frac{d}{dt}\Big|_{t=0} \mathcal{I}_{g,\gamma}(X + tY) = -\int_0^L \langle (\nabla_g)_S(\nabla_g)_S X + \mathrm{Rm}_g(X, S)S, Y\rangle_g\, ds$$

and

$$\frac{1}{2}\frac{d^2}{dt^2}\Big|_{t=0} \mathcal{I}_{g,\gamma}(X) = \int_0^L \Big(|(\nabla_g)_S Y|_g^2 - \langle \mathrm{Rm}_g(Y, S)S, Y\rangle_g \Big)\, ds.$$

Hence the critical points of $\mathcal{I}_{g,\gamma}$ on $\mathrm{Vect}_{A,B}(\gamma)$ are the Jacobi fields.

We claim that

$$\int_0^L \left(|(\nabla_g)_S Y|_g^2 - \langle \mathrm{Rm}_g(Y,S)S, Y\rangle_g\right) ds > 0$$

for any nonzero vector field $Y \in T_X \mathrm{Vect}_{A,B}(\gamma)$, so that the index form $\mathcal{I}_{g,\gamma}$ is convex. Hence the Jacobi fields minimize $\mathcal{I}_{g,\gamma}$ in $\mathrm{Vect}_{A,B}(\gamma)$. We now give a variational proof of this inequality.

Normalize the index by defining

$$\iota(t) := \inf_{0 \neq Z \in T_X \mathrm{Vect}_{A,B}(\gamma)} \frac{\mathcal{I}_{g,\gamma,t}(Z)}{\int_0^t |Z(s)|_g^2 ds}$$

where

$$\mathcal{I}_{g,\gamma,t}(Z) := \int_0^t \left(|(\nabla_g)_S Y|_g^2 - \langle \mathfrak{Rm}_g(Y,S)S, Y\rangle_g\right) ds$$

for $t \in [0, L]$.

(i) First we have

$$\frac{d}{ds}|Z|_g \le |(\nabla_g)_S Z|_g, \quad \langle \mathrm{Rm}_g(Z,S)S, Z\rangle_g \le C|Z|_g^2$$

for some constant C depending only on g. Indeed,

$$\frac{d}{ds}|Z|_g = \frac{d}{ds}\langle Z, Z\rangle_g^{1/2} = \frac{1}{|Z|}\langle (\nabla_g)_S Z, Z\rangle_g \le \frac{1}{|Z|}|Z|_g |(\nabla_g)_S Z|_g = |(\nabla_g)_S Z|_g .$$

(ii) Second we have $\lambda_1([0,t]) = \frac{\pi^2}{t^2}$, where $\lambda_1([0,t])$ is the first eigenvalue of d^2/ds^2 with Dirichlet boundary conditions. Hence

$$\iota(t) \ge \inf_{0 \neq Z \in T_X \mathrm{Vect}_{A,B}(\gamma)} \frac{\int_0^t \left(\left(\frac{d}{ds}|Z|_g\right)^2 - C|Z|_g^2\right) ds}{\int_0^t |Z|_g^2 ds} \ge \frac{\pi^2}{t^2} - C.$$

For $t \in (0, L]$ where $\gamma_{[0,t]}$ is minimizing (e.g., for $t > 0$ small enough), we have $\iota(t) \ge 0$. Since $\mathcal{I}_{g,\gamma,t}$ is a second variation of $\gamma|_{[0,t]}$ vanishing at the endpoints 0 and t, and $\iota(t)$ is continuous, if the claim is not ture, we can find $t_0 \in (0, L]$ such that $\iota(t_0) = 0$. Then $\mathcal{I}_{g,\gamma,t_0}(Z_0) = 0$ for some vector field Z_0 with $Z_0(0) = 0$, $Z_0(t_0) = 0$, and $Z_0 \neq 0$. By considering the Euler-Lagrange equation for

$$E(Z) := \frac{\mathcal{I}_{g,\gamma,t_0}(Z)}{\int_0^{t_0} |Z(s)|_g^2 ds}$$

at Z_0, we have for all W vanishing at 0 and t_0,

$$0 = \frac{1}{2}\frac{d}{du}\Big|_{u=0} E(Z_0 + uW) = \frac{-1}{\int_0^{t_0} |Z_0(s)|_g^2 ds}\int_0^{t_0} \langle (\nabla_g)_S(\nabla_g)_S Z_0 + \mathrm{Rm}_g(Z_0, S)S, W\rangle_g \, ds$$

since $\mathcal{I}_{g,\gamma,t_0}(Z_0) = 0$. Thus Z_0 is a nontrivial Jacobi field along $\gamma|_{[0,t_0]}$ with $Z_0(0) = 0 = Z_0(t_0)$. This contradicts the assumption that there are no conjugate points along γ. Hence $\iota(t) > 0$ for all $t \in (0, L]$. □

2.9.5 First and second variation of energy

Given a path $\gamma : [a, b] \to \mathcal{M}$, its **energy** is defined by

$$E_g(\gamma) := \frac{1}{2}\int_a^b |\dot\gamma(s)|_g^2 \, ds. \tag{2.9.28}$$

Let $\gamma_r : [a,b] \to \mathcal{M}$ denote a 1-parameter family of paths, $r \in \mathfrak{r} \in \mathbf{R}$. We also use the variation vector field R and the tangent vector field S. The length of γ_r is given by

$$E_g(\gamma_r) := \frac{1}{2}\int_a^b |\dot{\gamma}_r(s)|_g^2\, ds.$$

Lemma 2.20. (First variation of energy)

Suppose $0 \in \mathfrak{r}$. The first variation of energy is

$$\frac{d}{dr}\Big|_{r=0} E_g(\gamma_r) = -\int_a^b \left(\langle R, (\nabla_g)_S S\rangle_g\right)(\gamma_0(s))\, ds + [\langle R, S\rangle_g(\gamma_0(s))]\Big|_a^b. \tag{2.9.29}$$

♡

Proof. Calculate

$$\frac{d}{dr}\Big|_{r=0} E_g(\gamma_r) = \frac{1}{2}\int_a^b \frac{\partial}{\partial r}\Big|_{r=0} \langle \dot{\gamma}_r(s), \dot{\gamma}_r(s)\rangle_g (\gamma_r(s))\, ds = \frac{1}{2}\int_a^b (R\langle S, S\rangle_g)(\gamma_0(s))\, ds$$
$$= \int_a^b \left(\langle (\nabla_g)_R S, S\rangle_g\right)(\gamma_0(s))\, ds = \int_a^b \left(\langle (\nabla_g)_S R, S\rangle_g\right)(\gamma_0(s))\, ds$$

that implies the lemma. □

Note 2.48

The critical points of the energy, among all paths fixing two endpoints, are the constant speed geodesics γ, which satisfy

$$(\nabla_g)_{\dot{\gamma}}\dot{\gamma} = 0.$$

The speed of γ is constant since $\dot{\gamma}|\dot{\gamma}|_g^2 = 2\left\langle (\nabla_g)_{\dot{\gamma}}\dot{\gamma}, \dot{\gamma}\right\rangle_g = 0$.

♣

Let $\gamma_{q,r} : [a,b] \to \mathcal{M}$ with $q \in \mathfrak{q} \subset \mathbf{R}$ and $r \in \mathfrak{r} \subset \mathbf{R}$, be a 2-parameter family of paths. Recall the definition of vector fields Q, R, and S.

Lemma 2.21. (Second variation of energy)

Suppose $0 \in \mathfrak{q}$ and $0 \in \mathfrak{r}$. Then the second variation of energy is

$$\begin{aligned} \frac{\partial^2}{\partial q \partial r}\Big|_{(q,r)=(0,0)} E_g(\gamma_{q,r}) &= \int_a^b \left(\langle (\nabla_g)_S Q, (\nabla_g)_S R\rangle_g\right)(\gamma_{0,0}(s))\, ds \\ &+ \int_a^b \langle \mathrm{Rm}_g(Q,S)R, S\rangle_g (\gamma_{0,0}(s))\, ds \\ &- \int_a^b \langle (\nabla_g)_Q R, (\nabla_g)_S S\rangle_g (\gamma_{0,0}(s)) ds + \left[\langle (\nabla_g)_Q R, S\rangle_g\right](\gamma_{0,0}(s))\Big|_a^b. \end{aligned} \tag{2.9.30}$$

♡

Proof. Since

$$\frac{\partial^2}{\partial q \partial r}\Big|_{(q,r)=(0,0)} E_g(\gamma_{q,r}) = \int_a^b \left(Q\,\langle S, (\nabla_g)_S R\rangle_g\right)(\gamma_{0,0}(s))\, ds$$
$$= \int_a^b \left(\langle (\nabla_g)_Q S, (\nabla_g)_S R\rangle_g + \langle S, (\nabla_g)_Q (\nabla_g)_S R\rangle_g\right)(\gamma_{0,0}(s))\, ds$$

we prove the lemma. □

2.9.6 First and second variation of area

Let $x_r : \mathcal{S}^{m-1} \to \mathcal{M}$ be a parametrized hypersurface in an m-dimensional Riemannian manifold $(\mathcal{M}, g)$ evolving by

$$\partial_r x_r = \beta_r \nu_r \tag{2.9.31}$$

where β_r is some function on $\mathcal{S}^{m-1}_r := x_r(\mathcal{S}^{m-1})$. In terms of local coordinates $(x^i)_{i=1}^{m-1}$ on $\mathcal{S}^{m-1}$, the area element of $\mathcal{S}^{m-1}_r$ is

$$dS'_V = \sqrt{\det(g_{ij})}\, dx^1 \wedge \cdots \wedge dx^{m-1}. \tag{2.9.32}$$

Then

$$\partial_r g_{ij} = 2\beta_r h_{ij}. \tag{2.9.33}$$

Hence

$$\partial_r dV'_r = \frac{1}{2} g^{ij} \left(\partial_r g_{ij}\right) dV'_r = \beta_r H_r dV'_r. \tag{2.9.34}$$

Thus the first variation of

$$A_g(\mathcal{S}^{m-1}_r) := \int_{\mathcal{S}^{m-1}_r} dV'_r \tag{2.9.35}$$

is

$$\frac{d}{dr} A_g(\mathcal{S}^{m-1}_r) = \int_{\mathcal{S}^{m-1}_r} \beta_r H_r\, dV'_r. \tag{2.9.36}$$

Under the hypersurface flow (2.9.31), we have

$$\partial_r H_r = -\Delta_{g_r} \beta_r - |h_r|^2_{g_r} \beta_r - \mathrm{Rc}_{g_r}(\nu_r, \nu_r)\beta_r. \tag{2.9.37}$$

When $\beta_r = -H_r$, the mean curvature flow, we have

$$\partial_r H_r = \Delta_{g_r} H_r + |h_r|^2_{g_r} H_r + \mathrm{Rc}_g\left(\nu_r, \nu_r\right) H_r. \tag{2.9.38}$$

Now, we can compute the second variation of area:

$$\begin{aligned} \frac{d^2}{dr^2} A_g(\mathcal{S}^{m-1}_r) &= \int_{\mathcal{S}^{m-1}_r} \beta_r \left(-\Delta_{g_r}\beta_r - |h_r|^2_{g_r}\beta_r - \mathrm{Rc}_g\left(\nu_r, \nu_r\right)\beta_r + H_r^2 \beta_r\right) dS'_r \\ &= \int_{\mathcal{S}^{m-1}_r} \left(|\nabla_{g_r}\beta_r|^2_{g_r} + \left(H_r^2 - |h_r|^2_{g_r} - \mathrm{Rc}_g\left(\nu_r, \nu_r\right)\right)\beta_r^2\right) dS'_r. \end{aligned}$$

If $\partial_r x_r = \nu_r$, then

$$\frac{d^2}{dr^2} A_g(\mathcal{S}^{m-1}_r) = -\int_{\mathcal{S}^{m-1}_r} \left(H_r^2 - |h_r|^2_{g_r} - \mathrm{Rc}_g(\nu_r, \nu_r)\right) dS'_r.$$

However,

$$R_{g_r} = R_g - 2\mathrm{Rc}_g(\nu_r, \nu_r) + H_r^2 - |h_r|^2_{g_r}. \tag{2.9.39}$$

Therefore, if $\partial_r x_r = \nu_r$, then the second variation of area is given by

$$\frac{d^2}{dr^2} A_g(\mathcal{S}^{m-1}_r) = \frac{1}{2} \int_{\mathcal{S}^{m-1}_r} \left(R_{g_r} - R_g + H_r^2 - |h_r|^2_{g_r}\right) dS'_r. \tag{2.9.40}$$

Theorem 2.27. (Schoen-Yau, 1979)

If $\mathcal{S}^2$ is an orientable closed stable minimal surface in a 3-manifold $(\mathcal{M}^3, g)$ with positive scalar curvature, then $\mathcal{S}^2$ is diffeomorphic to a 2-sphere. ♡

Proof. Let $\mathcal{S}^2_r$ be a variation of $\mathcal{S}^2$ with $\mathcal{S}^2_0 = \mathcal{S}^2$ and $\beta = 1$, by (2.9.35), $H \equiv 0$, and the Gauss-Bonnet formula, we have

$$0 \le 2\frac{d^2}{dr^2}\Big|_{r=0} A_g(\mathcal{S}^2_r) = \int_{\mathcal{S}^2}\left(R_{\mathcal{S}^2} - R_{\mathcal{M}^3} - |h|^2_g\right) dV'_g 4\pi\chi(\mathcal{S}^2) - \int_{\mathcal{S}^2}\left(R_{\mathcal{M}^3} + |h|^2_g\right) dV'_g.$$

Since $R_{\mathcal{M}^3} > 0$ and $|h|^2_g \ge 0$, it follows that $\chi(\mathcal{S}^2) > 0$. Since $\mathcal{M}^3$ is orientable, $\mathcal{S}^2 \cong \mathbf{S}^2$. □

2.10 Geodesics and the exponential maps

Introduction

- *Exponential maps*
- *Gauss lemma and the Hopf -Rinow theorem*
- *Cut locus and injectivity radius*

Let $(\mathcal{M}, g)$ be an m-dimensional Riemannian manifold and $p \in \mathcal{M}$. For $V \in T_p\mathcal{M}$, there is a unique constant speed geodesic $\gamma_V : [0, b_V) \to \mathcal{M}$ is the constant speed geodesic emanating from p with $\dot{\gamma}_V(0) = V$. Here $[0, b_V)$ is the maximal time interval on which γ_V is defined.

2.10.1 Exponential map

For all $\alpha > 0$ and $t < b_{\alpha V}$, we have

$$\gamma_{\alpha V}(t) = \gamma_V(\alpha t), \quad b_{\alpha V} = \alpha^{-1} b_V. \tag{2.10.1}$$

Let $O_p \subset T_p\mathcal{M}$ be the set of vectors V such that $1 < b_V$, so that $\gamma_V(t)$ is defined on $[0, 1]$. Then define the **exponential map** at p by

$$\exp_p : O_p \longrightarrow \mathcal{M}, \quad V \longmapsto \gamma_V(1). \tag{2.10.2}$$

If $b_V > t$, then $b_{tV} = t^{-1} b_V > 1$ and

$$\exp_p(tV) = \gamma_{tV}(1) = \gamma_V(t). \tag{2.10.3}$$

Note 2.49

If $\mathcal{M}$ is compact, then for each $p \in \mathcal{M}$ and $V \in T_p\mathcal{M}$, there is a unique constant speed geodesic $\gamma_V : [0, \infty) \to \mathcal{M}$ with $\gamma(0) = p$ and $\dot{\gamma}_V(0) = V$. ♣

Let $O := \cup_{p\in\mathcal{M}} O_p$. Then the exponential map $\exp_p$ induces a map

$$\exp : O \longrightarrow \mathcal{M} \tag{2.10.4}$$

by setting $\exp_{|O_p} = \exp_p$. This map is also called the exponential map. Furthermore, the set O is open in $T\mathcal{M}$ and $\exp : O \to \mathcal{M}$ is smooth. In addition, $O_p \subset T_p\mathcal{M}$ is open and $\exp_p : O_p \to \mathcal{M}$ is also smooth.

Proposition 2.19

(1)If $p \in \mathcal{M}$, then

$$d\exp_p : T_0(T_p\mathcal{M}) \longrightarrow T_p\mathcal{M} \tag{2.10.5}$$

is nonsingular at the origin of $T_p\mathcal{M}$. Consequently $\exp_p$ is a local diffeomorphsm.

(2) Define $\mathrm{Exp} : O \to \mathcal{M} \times \mathcal{M}$ by

$$\mathrm{Exp}(V) = \left(\pi(V), \exp_{\pi(V)} V\right),$$

where $\pi(V)$ is the base point of V, i.e., $V \in T_{\pi(V)}\mathcal{M}$. Then for each $p \in \mathcal{M}$ and with it the zero vector, $0_p \in T_p\mathcal{M}$,

$$d\mathrm{Exp}_{(p,0_p)} : T_{(p,0_p)}(T\mathcal{M}) \longrightarrow T_{(p,p)}(\mathcal{M} \times \mathcal{M})$$

is nonsingular. Consequently, Exp is a diffeomorphism from a neighborhood of the zero section of $T\mathcal{M}$ onto an open neighborhood of the diagonal in $\mathcal{M} \times \mathcal{M}$. ♡

Proof. Let $I_0 : T_p\mathcal{M} \to T_0(T_p\mathcal{M})$ be the canonical isomorphism, i.e.,

$$I_0(V) := \frac{d}{dt}(tV)\Big|_{t=0}. \tag{2.10.6}$$

Recall that if $V \in O_p$, then $\gamma_V(t) = \gamma_{tV}(1)$ for all $t \in [0, 1]$. Thus,

$$d\exp_p(I_0(V)) = \frac{d}{dt}\Big|_{t=0} \exp_p(tV) = \frac{d}{dt}\Big|_{t=0} \gamma_{tV}(1) = \frac{d}{dt}\Big|_{t=0} \gamma_V(t) = \dot{\gamma}_V(0) = V.$$

In other words, $d\exp_p \circ I_0$ is the identity map on $T_p\mathcal{M}$. This shows that $d\exp_p$ is nonsingular.

For (2), we note that the tangent space $T_{(p,p)}(\mathcal{M} \times \mathcal{M})$ is naturally identified with $T_p\mathcal{M} \times T_p\mathcal{M}$. The tangent space $T_{(p,0_p)}(T\mathcal{M})$ is also naturally identified to $T_pM \times T_{0_p}(T\mathcal{M}) \cong T_p\mathcal{M} \times T_p\mathcal{M}$. We know that $d\mathrm{Exp}_{(p,0_p)}$ takes (p, V) to $(p, \exp_p(V))$. Under above identification, if we consider the map $d\mathrm{Exp}_{(p,0_p)}$ as a linear map $T_p\mathcal{M} \times T_p\mathcal{M} \to T_p\mathcal{M} \times T_p\mathcal{M}$, then it looks like

$$\begin{bmatrix} I & 0 \\ * & I \end{bmatrix}$$

which is clearly nonsingular. □

Note 2.50

Suppose that $\mathcal{N}$ is an embedded submanifold of $\mathcal{M}$. The normal bundle of $\mathcal{N}$ in $\mathcal{M}$ is the vector bundle over $\mathcal{N}$ consisting of the orthogonal complements of the tangent spaces $T_p\mathcal{N} \subset T_p\mathcal{M}$:

$$T^{\perp}\mathcal{N} := \{(p, V) : V \in T_p\mathcal{M},\ p \in \mathcal{N},\ V \in (T_p\mathcal{N})^{\perp} \subset T_p\mathcal{M}\}. \tag{2.10.7}$$

So for each $p \in \mathcal{N}$,

$$T_p\mathcal{M} = T_p\mathcal{N} \oplus T_p^{\perp}\mathcal{N}$$

is an orthogonal direct sum. Define the **normal exponential map** $\exp^{\perp}$ *by restricting* exp *to* $O \cap T^{\perp}\mathcal{N}$*, so*

$$\exp^{\perp} : O \cap T^{\perp}\mathcal{N} \to \mathcal{M}.$$

As in part (2) of **Proposition 2.19**, $d\exp^{\perp}$ *is nonsingular at* 0_p, $p \in \mathcal{N}$. *Then it follows that there is an open neighborhood* $\mathcal{U}$ *of the zero section in* $T^{\perp}\mathcal{N}$ *on which* $\exp^{\perp}$ *is a diffeomorphism onto its image in* $\mathcal{M}$. *Such an image* $\exp^{\perp}(\mathcal{U})$ *is called a* **tubular neighborhood** *of* $\mathcal{N}$ *in* $\mathcal{M}$. ♣

Theorem 2.28

Suppose that $(\mathcal{M}, g)$ *is a Riemannian manifold,* $p \in \mathcal{M}$*, and* $\epsilon > 0$ *is such that*

$$\exp_p : B(0_p, \epsilon) \subset T_p\mathcal{M} \longrightarrow \mathcal{U} \subset \mathcal{M}$$

is a diffeomorphism onto its image $\mathcal{U} := \exp_p(B(0_p, \epsilon))$ *in* $\mathcal{M}$. *Then* $\mathcal{U} = B_g(p, \epsilon)$ *and for each* $V \in B(0_p, \epsilon)$*, the geodesic* $\gamma_V : [0, 1] \to \mathcal{M}$ *defined by*

$$\gamma_V(t) := \exp_p(tV)$$

is the unique minimal geodesic in $\mathcal{M}$ *from* p *to* $\exp_p(V)$. ♡

On $\mathcal{U}$ we have the function $r(x) := |\exp_p^{-1}(x)|$. That is, r is the Euclidean distance function from the origin on $B(0_p, \epsilon) \subset T_p\mathcal{M}$ in exponential coordinates.

2.10.2 Gauss lemma and the Hopf-Rinow theorem

Let $(\mathcal{M}, g)$ be a Riemannian manifold and $p \in \mathcal{M}$. Suppose that $V \in T_p\mathcal{M}$ and for some $L > 0$ the constant speed geodesic $\gamma_{\tilde{V}}$ with $\dot{\gamma}_{\tilde{V}}(0) = \widetilde{V}$ is defined on $[0, L]$ for every $\widetilde{V}$ in some neighborhood of V. Given $u \in (0, L)$, let

Lemma 2.22. (Gauss)

If $W \in T_{uV}(T_p\mathcal{M}) \cong T_p\mathcal{M}$ *is perpendicular to* V*, then the image* $(\exp_p)_{*,uV}(W)$ *of* W *is perpendicular to* $(\exp_p)_{*,p}(V) = \dot{\gamma}_V(u)$*:*

$$\left\langle (\exp_p)_{*,uV}(W), (\exp_p)_{*,uV}(V) \right\rangle_g = 0. \tag{2.10.8}$$

If the distance function $r := d_g(p, \cdot)$ *is smooth at a point* x*, we then have*

$$\nabla_g r(x) = \dot{\gamma}_0(b), \tag{2.10.9}$$

where $\gamma_0 : [0, b] \to \mathcal{M}$ *is the unique unit speed minimal geodesic from* p *to* x. *Thus, if* $\gamma_0 = \gamma_{V_0}$ *for some unit vector* V_0*, then* $\nabla_g r = (\exp_p)_{*,0}(V_0)$. ♡

Proof. Given $V, W \in T_p\mathcal{M}$, we define the family of geodesics

$$\gamma_r(s) := \exp_p(s(uV + rW)), \quad 0 \le s \le 1.$$

Then $L_g(\gamma_r) = |uV + rW|_{g(p)}$ and

$$\begin{aligned}\frac{d}{dr}\Big|_{r=0} L_g(\gamma_r) &= \frac{d}{dr}\Big|_{r=0} \langle uV + rW, uV + rW\rangle^{1/2}_{g(p)} \\ &= \frac{2u\langle V, W\rangle_{g(p)}}{2|uV + rW|_{g(p)}}\Big|_{r=0} = \frac{1}{|V|_{g(p)}}\langle V, W\rangle_{g(p)}.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\frac{d}{dr}\Big|_{r=0} L_g(\gamma_r) &= -\frac{1}{u|V|_{g(p)}}\int_0^1 \langle R, (\nabla_g)_S S\rangle_g ds + \frac{1}{u|V|_{g(p)}}\langle R, S\rangle_g\Big|_0^1 \\ &= \frac{1}{u|V|_{g(p)}}\left\langle (\exp_p)_{*,uV}(W), (\exp_p)_{*,uV}(uV)\right\rangle_g \\ &= \frac{1}{|V|_{g(p)}}\left\langle (\exp_p)_{*,uV}(W), (\exp_p)_{*,uV}(V)\right\rangle_g\end{aligned}$$

since γ_r are geodesics. Thus

$$\left\langle (\exp_p)_{*,uV}(W), (\exp_p)_{*,uV}(V)\right\rangle_g = \langle V, W\rangle_g = 0$$

which proves the first part.

Let $\gamma_r : [0, b] \to \mathcal{M}$, $r \in \mathfrak{r}$, be an arbitrary variation of γ_0 with $\gamma_r(0) = p$. Since γ_0 is a minimal geodesic, we have $L_g(\gamma_r) \geq d_g(p, \gamma_r(b))$ and $L_g(\gamma_0) = d_g(p, \gamma_0(b))$. Hence, since $\nabla_g r$ exists at x,

$$\langle \nabla_g r, X\rangle_{g(x)} = \frac{d}{dr}\Big|_{r=0} L_g(\gamma_r),$$

where $X := \frac{\partial}{\partial r}|_{r=0}\gamma_r(b)$. On the other hand, by the first variation formula (2.9.8),

$$\frac{d}{dr}\Big|_{r=0} L_g(\gamma_r) = \left\langle \frac{\partial}{\partial s}\Big|_{s=0}\gamma_s(b), X\right\rangle_g = \langle \dot{\gamma}_0(b), X\rangle_g.$$

Therefore, $\nabla_g r = \dot{\gamma}_0(b)$. □

Let $\partial/\partial r$ denote the radial unit outward pointing vector field on $T_p\mathcal{M} \setminus \{0\}$ and consider the map, where $X \in T_p\mathcal{M}$,

$$(\exp_p)_{*,X} : T_X(T_p\mathcal{M}) \longrightarrow T_{\exp_p(X)}\mathcal{M}.$$

We denote by ∂_X the canonical isomorphism

$$\partial_X : T_p\mathcal{M} \longrightarrow T_X(T_p\mathcal{M}), \quad Y \longmapsto \partial_X Y := \frac{d}{dt}\Big|_{t=0}(X + tY). \tag{2.10.10}$$

Hence we obtain

$$(\mathbf{exp}_p)_{*,X} := (\exp_p)_{*,X} \circ \partial_0 : T_p\mathcal{M} \longrightarrow T_{\exp_p(X)}\mathcal{M}. \tag{2.10.11}$$

Since $(\mathbf{exp}_p)_{*,0} = \mathrm{id}_{T_p\mathcal{M}}$ is invertible, there exists an $\epsilon > 0$ such that $\exp_p$ restricted on the punctured ball $B(0, \epsilon) \setminus \{0\} \subset T_p\mathcal{M}$ is an embedding. We denote

$$r_g(x) := \left|\exp_p^{-1}(x)\right|_{g(p)}, \quad x \in B_g(p, \epsilon) := \exp_p(B(0, \epsilon)).$$

If $\{r, \theta^1, \cdots, \theta^{m-1}\}$ are spherical coordinates in $T_p\mathcal{M}$, then we set

$$\frac{\partial}{\partial r_g}\Big|_{\exp_p(X)} := (\exp_p)_{*,X}\left(\frac{\partial}{\partial r}(X)\right), \quad \frac{\partial}{\partial \theta^i_g}\Big|_{\exp_p(X)} := (\exp_p)_{*,X}\left(\frac{\partial}{\partial \theta^i}(X)\right).$$

For every $X \in T_x\mathcal{M}$, we may write it as

$$X = a\frac{\partial}{\partial r_g}\Big|_x + \sum_{i=1}^{m-1} b^i \frac{\partial}{\partial \theta_g^i}\Big|_x;$$

By Gauss lemma, **Lemma 2.22**, one has

$$\left\langle \frac{\partial}{\partial r_g}\Big|_x, \frac{\partial}{\partial \theta_g^i}\Big|_x \right\rangle_g = 0$$

and hence

$$\left\langle \frac{\partial}{\partial r_g}\Big|_x, X \right\rangle_g = a = X(r_g) = \left\langle \mathrm{grad}_g(r_g(x)), X \right\rangle_g.$$

Thus

$$\frac{\partial}{\partial r_g}\Big|_x = \mathrm{grad}_g(r_g)(x), \quad x \in B_g(p, \epsilon). \tag{2.10.12}$$

Lemma 2.23. l2.23

or every $V \in B(0, \epsilon)$, $\gamma_V : [0, 1] \to \mathcal{M}$ is the unit path, up to reparametrization, joining p and $\gamma_V(1) = \exp_p(V)$ whose length realizes the distance $d_g(p, \exp_p(V)) = |V|_g$. In particular, short geodesics are minimal and $r_g(x) = d_g(p, x)$ for $x \in B_g(p, \epsilon)$. ♡

Proof. Since $\partial/\partial r$ is unit, it follows from (2.10.12) that

$$|\dot\beta(u)|_g \geq \left\langle \dot\beta(u), \frac{\partial}{\partial r_g}\Big|_{\beta(u)} \right\rangle_g = \left\langle \dot\beta(u), \mathrm{grad}_g(r_g)(\beta(u)) \right\rangle_g = \dot\beta(u)(r_g) = \frac{d}{du} r_g(\beta(u))$$

for any path β from p to $\exp_p(V)$ that stays inside $B_g(p, \epsilon)$, so that

$$|V|_g = r_g(\exp_p(V)) = \int_0^1 \frac{d}{du} r_g(\beta(u))\, du \leq \int_0^1 |\dot\beta(u)|_{g(\beta(u))}\, du \leq d_g(p, \exp_p(V)).$$

Hence γ_V realizes the distance from p to $\exp_p(V)$. □

Theorem 2.29. (Hopf-Rinow)

Let $(\mathcal{M}, g)$ be a Riemannian manifold. Then the following are equivalent:

(1) $(\mathcal{M}, d_g)$ is a complete metric space.

(2) There exists $p \in \mathcal{M}$ such that $\exp_p$ is defined on all of $T_p\mathcal{M}$.

(3) For all $p \in \mathcal{M}$, $\exp_p$ is defined on all of $T_p\mathcal{M}$.

Any one of these conditions implies

(4) For any $p, q \in \mathcal{M}$ there exists a smooth minimal geodesic form p to q. ♡

2.10.3 Cut locus and injectivity radius

Let $(\mathcal{M}, g)$ be a Riemannian manifold.

Definition 2.4. *Lipschitz functions)

A function $f : \mathcal{M} \to \mathbf{R}$ is a **globally Lipschitz function** *with Lipschitz constant C if for all $x, y \in \mathcal{M}$ we have*

$$|f(x) - f(y)| \leq C d_g(x, y).$$

If for every $z \in \mathcal{M}$ there exists a neighborhood $\mathcal{U}_z$ of z and a constant C_z such that

$$|f(x) - f(y)| \leq C_z \cdot d_g(x, y)$$

for all $x, y \in \mathcal{U}_z$, then we say that f is a **locally Lipschitz function***.* ♣

The distance function $d_g(p, \cdot)$ is a globally Lipschitz function with Lipschitz constant 1.

Given a point $p \in \mathcal{M}$ and a unit speed geodesic $\gamma : [0, \infty) \to \mathcal{M}$ with $\gamma(0) = p$, either γ is a **geodesic ray** (i.e., minimal on each finite subinterval) or there exists a unique $r_\gamma \in (0, \infty)$ such that $d_g(p, \gamma(r)) = r$ for $r \leq r_\gamma$ and $d_g(p, \gamma(r)) < r$ for $r > r_\gamma$. We say that $\gamma(r_\gamma)$ is a **cut point to p along γ**.

(i) If $\gamma(r)$ is a conjugate point to p along γ, then $r \geq r_\gamma$.

(ii) The **cut locus** $\mathrm{Cut}_g(p)$ of p in $\mathcal{M}$ is the set of all cur points of p.

(iii) Let

$$D_g(p) := \{V \in T_p\mathcal{M} : d_g(p, \exp_p(V)) = |V|_g\}, \tag{2.10.13}$$

which is a closed subset of $T_p\mathcal{M}$. We define $C_g(p) := \partial D_g(p)$ to be the cut locus of p in the tangent space. We have

$$\mathrm{Cut}_g(p) = \exp_p(C_g(p)) \tag{2.10.14}$$

and

$$\exp_p : \mathrm{int}(D_g(p)) \subset T_p\mathcal{M} \longrightarrow \mathcal{M} \setminus \mathrm{Cut}_g(p)$$

is a diffeomorphism. We call $\mathrm{int}(D_g(p))$ the interior to the cut locus in the tangent space $T_p\mathcal{M}$.

Lemma 2.24

A point $\gamma(r)$ is a cut point to p along γ if and only if r is the smallest positive number such that either $\gamma(r)$ is a conjugate point to p along γ or there exist two distinct minimal geodesics joining p and $\gamma(r)$. ♡

Given $V \in T_p\mathcal{M}$ and $r > 0$, we have $\gamma_V(r) = \exp_p(rV)$. For each unit vector $V \in T_p\mathcal{M}$ there exists at most a unique $r_V \in (0, \infty)$ such that $\gamma_V(r_V)$ is a cut point of p along γ_V. Furthermore, if we set $r_V = \infty$ when γ_V is a ray, then the map from the unit tangent space at p to $(0, \infty]$ given by $V \mapsto r_V$ is a continuous function. Hence we have

$$C_g(p) = \partial D_g(p) = \{r_V V : V \in T_p\mathcal{M},\ |V|_{g(p)} = 1,\ \gamma_V \text{ is not a ray}\} \tag{2.10.15}$$

has measure zero with respect to the Euclidean measure on $(T_p\mathcal{M}, g(p))$.

Lemma 2.25

$\mathrm{Cut}_g(p) = \exp_p(C_g(p))$ *has measure zero with respect to the Riemannian measure on* $(\mathcal{M}, g)$. ♡

If $x \notin \mathrm{Cut}_g(p)$ and $x \neq p$, then $d_g(p, \cdot)$ is smooth at x and $|\nabla_g d_g(p, x)|_{g(x)} = 1$ by (2.10.9). Since $\mathrm{Cut}_g(p)$ has measure zero, we have $|\nabla_g d_g(p, \cdot)|_g = 1$ a.e. on $\mathcal{M}$.

Definition 2.5. (Injectivity radius)

The **injectivity radius** $\mathrm{inj}_g(p)$ *of a point* $p \in \mathcal{M}$ *is defined to the supremum of all* $r > 0$ *such that* $\exp_p$ *is an embedding when restricted to* $B(0, r)$*. Equivalently,*

(1) $\mathrm{inj}_g(p)$ *is the distance from* 0 *to* $C_g(p)$ *with respect to* $g(p)$*.*

(2) $\mathrm{inj}_g(p)$ *is the Riemannian distance from* p *to* $\mathrm{Cut}_g(p)$*.*

The injectivity radius of a Riemannian manifold $(\mathcal{M}, g)$ *is defined to be*

$$\mathrm{inj}_g(\mathcal{M}) := \inf_{p \in \mathcal{M}} \mathrm{inj}_g(p). \tag{2.10.16}$$

When $\mathcal{M}$ *is compact, the injectivity radius is always positive.* ♣

Theorem 2.30. (Klingenberg)

(1) If $(\mathcal{M}, g)$ *is a compact Riemannian manifold with* $\mathrm{Sec}_g \leq K$*, then*

$$\mathrm{inj}_g(\mathcal{M}) \geq \min\left\{\frac{\pi}{\sqrt{K}}, \frac{1}{2} \cdot \textit{length of shortest closed geodesic}\right\}. \tag{2.10.17}$$

(2) If $(\mathcal{M}, g)$ *is a complete simply-connected Riemannian manifold with* $0 < \frac{1}{4}K < \mathrm{Sec}_g \leq K$*, then*

$$\mathrm{inj}_g(\mathcal{M}) \geq \frac{\pi}{\sqrt{K}}.$$

(3) If $(\mathcal{M}, g)$ *is a compact, even-dimensional, orientable Riemannian manifold with* $0 < \mathrm{Sec}_g \leq K$*, then*

$$\mathrm{inj}_g(\mathcal{M}) \geq \frac{\pi}{\sqrt{K}}.$$

♡

2.11 Second fundamental forms of geodesic spheres

Introduction

- *Geodesic coordinate expansion of the metric and volume form*
- *Geodesic spherical coordinates and the Jacobian*
- *The second fundamental form of distance spheres and the Ricatti equation*
- *Space form and rotationally symmetric metrics*
- *Mean curvature of geodesic spheres and the Bonnet-Myers theorem*

In this section we consider geodesic spherical coordinates and the second fundamental forms and mean curvatures of geodesic spheres. We also give the proofs of the Laplacian and Hessian comparison theorems for the distance function and the corresponding volume and Rauch comparison theorems.

2.11.1 Geodesic coordinate expansion of the metric and volume form

Let $(\mathcal{M}, g)$ be a Riemannian manifold. The exponential map $\exp_p : T_p\mathcal{M} \to \mathcal{M}$ is defined by $\exp_p(V) := \gamma_V(1)$, where $\gamma_V : [0, \infty) \to \mathcal{M}$ is the constant speed geodesic emanating from p with $\dot\gamma_V(0) = V$.

Given an orthonormal frame $\{e_i\}_{1\le i\le m}$ at p, let $\{X^i\}_{1\le i\le m}$ denote the standard Euclidean coordinates on $T_p\mathcal{M}$ defined by $V = \sum_{1\le i\le m} V^i e_i$. **Geodesic coordinates** are defined by

$$x^i := X^i \circ \exp_p^{-1} : \mathcal{M} \setminus \mathrm{Cut}_g(p) \longrightarrow \mathbf{R}. \tag{2.11.1}$$

In geodesic coordinates, we have

$$\begin{aligned} g_{ij} &= \delta_{ij} - \frac{1}{3}R_{ipqj}x^p x^q - \frac{1}{6}\nabla_r R_{ipqj}x^p x^q x^r \\ &\quad + \left(-\frac{1}{20}\nabla_r\nabla_s R_{ipqj} + \frac{2}{45}g^{uv}R_{ipqu}R_{jrsv}\right)x^p x^q x^r x^s + O(r_g^5) \end{aligned} \tag{2.11.2}$$

so that $g_{ij} = \delta_{ij} + O(r_g^2)$, and

$$\begin{aligned} \det(g) &= 1 - \frac{1}{3}R_{ij}x^i x^j - \frac{1}{6}\nabla_k R_{ij}x^i x^j x^k \\ &\quad - \left(\frac{1}{20}\nabla_\ell\nabla_k R_{ij} + \frac{1}{90}R_{pijq}R^p{}_{k\ell}{}^q - \frac{1}{18}R_{ij}R_{k\ell}\right)x^i x^j x^k x^\ell + O(r_g^5). \end{aligned} \tag{2.11.3}$$

Lemma 2.26. (Expansion for volumes of balls)

One has

$$\mathrm{Vol}(B_g(p,r)) = \omega_m r^m\left[1 - \frac{R_g(p)}{6(m+2)}r^2 + O(r^3)\right]. \tag{2.11.4}$$

♡

Proof. It follows from

$$\sqrt{\det(g)(x)} = 1 - \frac{1}{6}R_{ij}(p)x^i x^j + O(r_g^3(x))$$

by (2.11.3). □

Lemma 2.27

In geodesic coordinates centered at a point $p \in \mathcal{M}$ we have

$$g_{ij}(p) = \delta_{ij}, \quad \frac{\partial}{\partial x^i}g_{jk}(p) = 0. \tag{2.11.5}$$

♡

2.11.2 Geodesic spherical coordinates and the Jacobian

We say that the geometry is **bounded** or **controlled** if there is a curvature bound and an injectivity radius lower bound.

Given a point $p \in \mathcal{M}$, let $(X^i)_{i=1}^m$ be local spherical coordinates on $T_p\mathcal{M} \setminus \{p\}$. That is,

$$X^m(V) := r(V) = |V|_{g(p)}, \quad X^i(V) := \theta^i\left(\frac{V}{|V|_g}\right) \text{ for } 1 \le i \le m-1, \tag{2.11.6}$$

where $\{\theta^i\}_{1\le i\le m-1}$ are local coordinates on $S_p^{m-1} := \{V \in T_p\mathcal{M} : |V|_{g(p)} = 1\}$. Let

$\exp_p : T_p\mathcal{M} \to \mathcal{M}$ be the exponential map. We call the coordinate system

$$x := \{x^i := X^i \circ \exp_p^{-1}\} : B_g(p, \mathrm{inj}_g(p)) \setminus \{p\} \longrightarrow \mathbf{R}^m \tag{2.11.7}$$

a **geodesic spherical coordinate system**. Abusing notation, we let

$$r_g := x^m, \quad \theta_g^i := x^i \tag{2.11.8}$$

for $i = 1, \cdots, m-1$, so that

$$\frac{\partial}{\partial r_g} = (x^{-1})_* \frac{\partial}{\partial X^m}, \quad \frac{\partial}{\partial \theta_g^i} = (x^{-1})_* \frac{\partial}{\partial X^i}, \tag{2.11.9}$$

which form a basis of vector fields on $B_g(p, \mathrm{inj}_g(p)) \setminus \{p\}$. Recall from (2.10.12) that the Gauss lemma says that $\mathrm{grad}_g r_g = \frac{\partial}{\partial r_g}$ at all points outside the cut locus of p, so that

$$|\mathrm{grad}_g r_g|_g^2 = \left|\frac{\partial}{\partial r_g}\right|_g^2 = \left\langle \mathrm{grad}_g r_g, \frac{\partial}{\partial r_g}\right\rangle_g = \frac{\partial}{\partial r_g} r_g = 1 \tag{2.11.10}$$

and

$$g_{im} := g\left(\frac{\partial}{\partial r_g}, \frac{\partial}{\partial \theta_g^i}\right) = \left\langle \frac{\partial}{\partial X^m}, \frac{\partial}{\partial X^i}\right\rangle_g = 0$$

for $i = 1, \cdots, m-1$. We may then write the metric as

$$g = dr_g \otimes dr_g + g_{ij} d\theta_g^i \otimes d\theta_g^j, \tag{2.11.11}$$

where $g_{ij} := g(\partial/\partial\theta_g^i, \partial/\partial\theta_g^j)$.

Along each geodesic ray emanating from p, $\partial/\partial\theta_g^i$ is a Jacobi field, before the first conjugate point for each $i \le m-1$. We call

$$\mathrm{J}_g := \sqrt{\det(g_{ij})_{1\le i,j\le m-1}} \tag{2.11.12}$$

the **Jacobian of the exponential map**. The volume of g is

$$dV_g = \sqrt{\det(g)}\, d\theta_g^1 \wedge \cdots \wedge d\theta_g^{m-1} \wedge dr_g = \mathrm{J}_g\, d\Theta_g \wedge dr_g \tag{2.11.13}$$

in a positively oriented spherical coordinate system, where

$$d\Theta_g := d\theta_g^1 \wedge \cdots \wedge d\theta_g^{m-1}. \tag{2.11.14}$$

Hence the Jacobian of the exponential map is the volume density in spherical coordinates. If $\gamma(\bar{r})$ is a conjugate point to p along γ, then $\mathrm{J}_g(\gamma(r)) \to 0$ as $r \to \bar{r}$.

Note 2.51

Along a geodesic ray γ emanating from p we have that

$$\lim_{x\to p}\left((\nabla_g)_{\frac{\partial}{\partial r_g}} \frac{\partial}{\partial \theta_g^i}\right)(\gamma(x)) := E_i \in T_p\mathcal{M} \tag{2.11.15}$$

exists. Suppose $(E_i)_{1\le i\le m-1}$ is orthonormal (one can always choose such geodesic spherical coordinates and we shall often make this assumption in the sequel). Then

$$\lim_{r_g\to 0} \frac{\mathrm{J}_g(\gamma(r_g))}{r_g^{m-1}} = 1. \tag{2.11.16}$$

Intuitively, one way to see that (2.11.16) holds is to note that $(\mathcal{M}, cg, p)$ converges as $c \to \infty$ in the pointed limit $(\mathbf{R}^m, 0)$, so that the limit in (2.11.16) should equal the Euclidean value. ♣

2.11.3 The second fundamental form of distance spheres and the Ricatti equation

Consider the distance spheres

$$S_g(p,r) := \{x \in \mathcal{M} : d_g(p,x) = r\}. \tag{2.11.17}$$

Let h denote the second fundamental form of $S_g(p,r)$ as defined in (2.8.22). We have

$$\begin{aligned} h_{ij} &:= h\left(\frac{\partial}{\partial\theta^i_g}, \frac{\partial}{\partial\theta^j_g}\right) = \left\langle (\nabla_g)_{\frac{\partial}{\partial\theta^i_g}}\frac{\partial}{\partial r_g}, \frac{\partial}{\partial\theta^j_g}\right\rangle_g \\ &= -\left\langle \frac{\partial}{\partial r_g}, (\nabla_g)_{\frac{\partial}{\partial\theta^i_g}}\frac{\partial}{\partial\theta^j_g}\right\rangle_g = -\Gamma^m_{ij} = \frac{1}{2}\frac{\partial}{\partial r_g}g_{ij} \end{aligned} \tag{2.11.18}$$

since $\partial/\partial r_g$ is the unit normal to $S_g(p,r)$ and $g_{im} = g_{jm} = 0$. The mean curvature H of $S_g(p,r)$ is

$$H = -g^{ij}\Gamma^m_{ij} = \frac{1}{2}g^{ij}\frac{\partial}{\partial r_g}g_{ij} = \frac{\partial}{\partial r_g}\ln \mathrm{J}_g. \tag{2.11.19}$$

Note 2.52

(1) For r_g small enough,

$$\begin{aligned} h_{ij} &= \frac{1}{r_g}g_{ij} + O(r_g), & (2.11.20)\\ H &= \frac{m-1}{r_g} + O(r_g). & (2.11.21)\end{aligned}$$

(2) In spherical coordinates, the Laplacian is

$$\begin{aligned} \Delta_g &= g^{ab}\left(\frac{\partial^2}{\partial x^a\partial x^b} - \Gamma^c_{ab}\frac{\partial}{\partial x^c}\right) = \frac{\partial^2}{\partial r_g^2} + H\frac{\partial}{\partial r_g} + \Delta_{S_g(p,r)} \quad (2.11.22)\\ &= \frac{\partial^2}{\partial r_g^2} + \left(\frac{\partial}{\partial r_g}\ln\sqrt{\det(g)}\right)\frac{\partial}{\partial r_g} + \Delta_{S_g(p,r)} \end{aligned}$$

since $\Gamma^a_{mm} = 0$ for $a = 1, \cdots, m$ and where $\Delta_{S_g(p,r)}$ is the Laplacian with respect to the induced metric on $S_g(p,r)$. ♣

Lemma 2.28

We have the **Ricatti equation**

$$\frac{\partial}{\partial r}h_{ij} = -R_{mijm} + h_{ik}g^{k\ell}h_{\ell j} \tag{2.11.23}$$

where $R_{mijm} := \langle \mathrm{Rm}_g(\frac{\partial}{\partial r_g}, \frac{\partial}{\partial\theta^i_g})\frac{\partial}{\partial\theta^j_g}, \frac{\partial}{\partial r_g}\rangle_g$. In particular,

$$\frac{\partial}{\partial r_g}H = -\mathrm{Rc}_g\left(\frac{\partial}{\partial r_g}, \frac{\partial}{\partial r_g}\right) - |h|^2_g. \tag{2.11.24}$$

♡

Proof. Since $|\partial/\partial r_g|_g = 1$, it follows that $(\nabla_g)_{\partial/\partial r_g}(\partial/\partial r_g) = 0$. From (2.11.18), we have

$$\begin{aligned} \frac{\partial}{\partial r_g}h_{ij} &= -\frac{\partial}{\partial r_g}\left\langle \frac{\partial}{\partial r_g}, (\nabla_g)_{\frac{\partial}{\partial\theta^i_g}}\frac{\partial}{\partial\theta^j_g}\right\rangle_g = -\left\langle \frac{\partial}{\partial r_g}, (\nabla_g)_{\frac{\partial}{\partial r_g}}(\nabla_g)_{\frac{\partial}{\partial\theta^i_g}}\frac{\partial}{\partial\theta^j_g}\right\rangle_g \\ &= -\left\langle \frac{\partial}{\partial r_g}, (\nabla_g)_{\frac{\partial}{\partial\theta^i_g}}(\nabla_g)_{\frac{\partial}{\partial r_g}}\frac{\partial}{\partial\theta^j_g}\right\rangle_g - R_{mijm} \end{aligned}$$

$$= -\frac{\partial}{\partial\theta^i_g}\left\langle\frac{\partial}{\partial r_g}, (\nabla_g)_{\frac{\partial}{\partial r_g}}\frac{\partial}{\partial\theta^j_g}\right\rangle_g + \left\langle(\nabla_g)_{\frac{\partial}{\partial\theta^i_g}}\frac{\partial}{\partial r_g}, (\nabla_g)_{\frac{\partial}{\partial r_g}}\frac{\partial}{\partial\theta^j_g}\right\rangle_g - R_{mijm}$$

$$= -\frac{\partial}{\partial\theta^i_g}\left(\frac{\partial}{\partial r_g}\left\langle\frac{\partial}{\partial r_g}, \frac{\partial}{\partial\theta^i_g}\right\rangle_g - \left\langle(\nabla_g)_{\frac{\partial}{\partial r_g}}\frac{\partial}{\partial r_g}, \frac{\partial}{\partial\theta^i_g}\right\rangle_g\right)$$

$$+ \left\langle(\nabla_g)_{\frac{\partial}{\partial\theta^i_g}}\frac{\partial}{\partial r_g}, (\nabla_g)_{\frac{\partial}{\partial\theta^j_G}}\frac{\partial}{\partial r_g}\right\rangle_g - R_{mijm} = 0 + h_{ik}h_{j\ell}g^{k\ell} - R_{mijm}.$$

Since

$$\frac{\partial}{\partial r_g}H = g^{ij}\frac{\partial}{\partial r_g}h_{ij} - \frac{\partial}{\partial r_g}g_{ij}\cdot h^{ij}, \quad \frac{\partial}{\partial r_g}g_{ij} = h_{ij},$$

we obtain (2.11.24). □

Note 2.53

If $\mathrm{Rc}_g \geq (m-1)K$, *then*

$$\frac{\partial}{\partial r_g}\left(\frac{H}{m-1}\right) \leq -K - \left(\frac{H}{m-1}\right)^2. \tag{2.11.25}$$

From (2.11.21), one has

$$\lim_{r_g\to 0}\frac{r_g H}{m-1} = 1.$$

♣

In terms of the radial covariant derivative

$$\nabla_m h_{ij} := \left((\nabla_g)_{\frac{\partial}{\partial r_g}}h\right)_{ij} = \frac{\partial}{\partial r_g}h_{ij} - \Gamma^k_{mi}h_{kj} - \Gamma^k_{mj}h_{ik}$$

and $\Gamma^k_{mi} = h_i{}^k$, we deduce from (2.11.23) that

$$\nabla_m h_{ij} = -R_{mijm} - h_{ik}h_{j\ell}g^{k\ell}. \tag{2.11.26}$$

Invariantly, we write this as

$$\left((\nabla_g)_{\frac{\partial}{\partial r_g}}\right)(X,Y) = -\left\langle \mathrm{Rm}_g\left(\frac{\partial}{\partial r_g}, X\right)Y, \frac{\partial}{\partial r_g}\right\rangle_g - h^2(X,Y) \tag{2.11.27}$$

for $X, Y \in TS_g(p,r)$.

2.11.4 Space form and rotationally symmetric metrics

We consider the geodesic spheres in simply-connected space form $(\mathcal{M}_K, g_K)$. In this case the metric is given by

$$g_K := dr_g^2 + s_K^2(r_g)g_{\mathbf{S}^{m-1}}, \tag{2.11.28}$$

where

$$s_K(r_g) := \begin{cases} \frac{1}{\sqrt{K}}\sin(\sqrt{K}r_g), & K>0, \\ r_g, & K=0, \\ \frac{1}{\sqrt{K}}\sinh(\sqrt{|K|}r_g), & K<0. \end{cases} \tag{2.11.29}$$

Lemma 2.29. (Curvatures of a rotationally symmetric metric)

If

$$g = dr_g^2 + \phi^2(r_g) g_{\mathbf{S}^{m-1}} \tag{2.11.30}$$

for some function ϕ, which is called a **rotationally symmetric metric**, *then the sectional curvatures are*

$$K_{\text{rad}} = -\frac{\phi''}{\phi}, \quad K_{\text{sph}} = \frac{1-(\phi')^2}{\phi^2}, \tag{2.11.31}$$

where K_{rad} (rad for radial) or K_{sph} (sph for spherical) is the sectional curvature of planes containing or perpendicular to, respectively, the radial vector. As a consequence, we have

$$\mathrm{Rc}_g = -(m-1)\frac{\phi''}{\phi} dr_g^2 + \left[(m-2)\left(1-(\phi')^2\right) - \phi''\phi\right] g_{\mathbf{S}^{m-1}} \tag{2.11.32}$$

and

$$R_g = -2(m-1)\frac{\phi''}{\phi} + (m-1)(m-2)\frac{1-(\phi')^2}{\phi^2}. \tag{2.11.33}$$

Furthermore, the Laplacian of g is

$$\Delta_g = \frac{\partial^2}{\partial r_g^2} + (m-1)\frac{\phi'}{\phi}\frac{\partial}{\partial r_g} + \Delta_{S_g(p,r)}. \tag{2.11.34}$$

♡

Proof. One way is to use the Cartan structure equations: $\omega^m = dr_g$ and $\omega^i = \phi(r_g)\eta^i$, where $\{\eta^i\}_{1\le i\le m-1}$ is a local orthonormal coframe field for $(\mathbf{S}^{m-1}, g_{\mathbf{S}^{m-1}})$. Another way of deriving (2.11.31) is to consider the distance spheres. From (2.11.18), we have

$$h_{ij} = \frac{1}{2}\frac{\partial}{\partial r_g} g_{ij} = \phi\phi' g_{\mathbf{S}^{m-1}} = \frac{\phi'}{\phi} g_{ij}, \quad i,j = 1, \cdot, m-1. \tag{2.11.35}$$

That is, the distance spheres are **totally umbillic** with principal curvatures κ equal to ϕ'/ϕ. The intrinsic curvature of the hypersurface $S_g(p,r)$ is

$$K_{\text{in}} = \frac{1}{\phi^2}. \tag{2.11.36}$$

From the Gauss equations, we have $K_{\text{sph}} = K_{\text{in}} - \kappa^2$. □

Using (2.11.19) yields

$$H = g^{ij}h_{ij} = (m-1)\frac{\phi'}{\phi}. \tag{2.11.37}$$

Example 2.19

When $\phi(r_g) = s_K(r_g)$ given by (2.11.29), the mean curvature $H_K(r_g)$ of the distance sphere $S_K(p,r)$ is

$$H_K(r_g) := \begin{cases} (m-1)\sqrt{K}\cot(\sqrt{K}r_g), & K > 0, \\ \frac{m-1}{r_g}, & K = 0, \\ (m-1)\sqrt{|K|}\coth(\sqrt{|K|}r_g), & K < 0. \end{cases} \tag{2.11.38}$$

Note that $H_K(r_g)$ is a solution to the equality case of (2.11.25). That is,

$$\frac{\partial}{\partial r_g}\left(\frac{H_K(r_g)}{m-1}\right) = -K - \left(\frac{H_K(r_g)}{m-1}\right)^2 \tag{2.11.39}$$

and

$$\lim_{r_g\to 0+}\frac{r_g H_K(r_g)}{m-1}=1.$$

2.11.5 Mean curvature of geodesic spheres and the Bonnet-Myers theorem

By the ODE comparison theorem, we have

Lemma 2.30. (Mean curvature of distance spheres comparison)

If the Ricci curvature of $(\mathcal{M},g)$ satisfies the lower bound $\mathrm{Rc}_g\geq(m-1)K$ for some $K\in\mathbf{R}$, then the mean curvatures of the distance spheres $S_g(p,r)$ satisfy

$$H\leq H_K \tag{2.11.40}$$

at points where the distance function is smooth.

Proof. From (2.11.25) and (2.11.39), we have

$$\frac{\partial}{\partial r_g}(H-H_K)\leq-\frac{H+H_K}{m-1}(H-H_K).$$

Note that $(H-H_K)(r_g)=O(r_g)$. Integrating (2.11.40), we get that for any $r_g\geq\epsilon>0$,

$$(H-H_K)(r_g)\leq(H-H_K)(\epsilon)\cdot\exp\left[-\int_\epsilon^{r_g}\frac{H+H_K}{m-1}(s)ds\right]. \tag{2.11.41}$$

Letting $\epsilon\to 0$ yields $(H-H_K)(r_g)\leq 0$. □

Theorem 2.31. (Bonnet-Myers)

t2.31 If $(\mathcal{M},g)$ is a complete Riemannian manifold with $\mathrm{Rc}_g\geq(m-1)K$, where $K>0$, then $\operatorname{diam}(\mathcal{M},g)\leq\pi/\sqrt{K}$. In particular, $\mathcal{M}$ is compact and $\pi_1(\mathcal{M})<\infty$.

Proof. Consider any point $p\in\mathcal{M}$ and suppose $\gamma:[0,L]\to\mathcal{M}$ is a unit speed minimal geodesic emanating from p. Then $d_g(p,\cdot)$ is smooth on $\gamma((0,L))$ and for every $r\in(0,L)$, the distance sphere $S_g(p,r)$ is smooth in a neighborhood of $\gamma(r)$. By **Lemma 2.30**, we have

$$H(r_g)\leq(m-1)\sqrt{K}\cot(\sqrt{K}r_g)$$

along $\gamma|_{(0,L)}$. Since

$$\lim_{r_g\to(\pi/\sqrt{K})_+}\cot(\sqrt{K}r_g)=-\infty,$$

it forces that $L\leq\pi/\sqrt{K}$. Thus $\operatorname{diam}(\mathcal{M},g)\leq\pi/\sqrt{K}$. Now a complete Riemannian manifold with finite diameter is compact.

Furthermore, we may apply the diameter bound to the universal covering Riemannian manifold $(\widetilde{\mathcal{M}},\tilde{g})$, where $\tilde{g}$ is the lifted metric. Indeed, $\tilde{g}$ satisfies the same Ricci curvature lower bound as g. This implies $\widetilde{\mathcal{M}}$ is compact and we conclude $\pi_1(\mathcal{M})<\infty$. □

2.12 Comparison theorems

Introduction

- *Laplacian comparison theorem*
- *Volume comparison theorem*
- *Hessian comparison theorem*
- *Mean value inequalities*
- *Rauch comparision theorem*

Two fundamental results in Riemannian geometry are the **Laplacian** and **Hessian comparison theorems for the distance function**. They are directly related to the volume comparison theorem and a special case of the Rauch comparison theorem. The Hessian comparison theorem may also be used to prove the Toponogov triangle comparison theorem.

2.12.1 Laplacian comparison theorem

The idea of comparison theorem is to compare a geometric quantity on a Riemannian manifold with the corresponding quantity on a model space. In Riemannian geometry, model spaces have constant sectional curvature, while, model spaces for the Ricci flow are gradient Ricci solitons.

Theorem 2.32. (Laplacian comparison)

If $(\mathcal{M}, g)$ is a complete Riemannian manifold with $\mathrm{Rc}_g \geq (m-1)K$, where $K \in \mathbf{R}$, and if $p \in \mathcal{M}$, then for any $x \in \mathcal{M}$ where $d_g(x) := d_g(p, x)$ is smooth, we have

$$\Delta_g d_g(x) \leq \begin{cases} (m-1)\sqrt{K}\cot\left(\sqrt{K}d_g(x)\right), & K > 0, \\ \frac{m-1}{d_g(x)}, & K = 0, \\ (m-1)\sqrt{|K|}\coth\left(\sqrt{|K|}d_g(x)\right), & K < 0. \end{cases} \tag{2.12.1}$$

On the whole manifold, the Laplacian comparison theorem (2.12.1) holds in the sense of distributions. ♡

In general, we say that $\Delta_g f \leq F$ in the **sense of distributions** if for any nonnegative C^∞-function φ on $\mathcal{M}$ with compact support, we have

$$\int_{\mathcal{M}} f \Delta_g \varphi dV_g \leq \int_{\mathcal{M}} F \varphi dV_g.$$

Proof. If $r_g(x) := d_g(x)$ is the distance function to p, then since r_g is constant on each sphere $\Delta_{S_g(p,r)} = 0$, then from (2.11.22) we have that the Laplacian of the distance function is the radial derivative of the logarithm of the Jacobian (and is the mean curvature of the distance spheres)

$$\Delta_g r_g = H = \frac{\partial}{\partial r} \ln \mathrm{J}_g. \tag{2.12.2}$$

Hence, if $\mathrm{Rc}_g \geq (m-1)K$, then, by **Lemma 2.30**,

$$\Delta_g r_g \leq H_K(r_g). \tag{2.12.3}$$

This proves the Laplacian comparison theorem assuming we are within the cut locus.

To prove (2.12.1) holds in the sense of distributions on all of $\mathcal{M}$, we argue as follows. For any nonnegative $\varphi \in C^\infty(\mathcal{M})$ with compact support,

$$\int_{\mathcal{M}} \varphi(x) H_K(d_g(x)) dV_g(x) = \int_0^\infty \int_{C_g(r)} \varphi\left(\exp_p(\theta, r)\right) H_K(r) \mathrm{J}_g(\theta, r) d\Theta(\theta) dr.$$

Given a unit vector $\theta \in T_p\mathcal{M}$, let r_θ be the largest value of r such that $s \mapsto \gamma_\theta(s) = \exp_p(\theta, s)$ minimizes up to $s = r$. By the Fubini theorem, we have

$$\int_{\mathcal{M}} \varphi(x) H_K(d_g(x)) dV_g(x) = \int_{\mathbf{S}^{m-1}} \int_0^{r_\theta} \varphi\left(\exp_p(\theta, r)\right) H_K(r) \mathrm{J}_g(\theta, r) dr\, d\Theta(\theta).$$

Now for $0 < r < r_\theta$, by **Lemma 2.30** and (2.12.2),

$$H_K(r)\mathrm{J}_g(\theta, r) \geq H(\theta, r)\mathrm{J}_g(\theta, r) = \frac{\partial}{\partial r}\mathrm{J}_g(\theta, r).$$

Hence

$$\int_{\mathcal{M}} \varphi(x) H_K(d_g(x)) dV_g(x) \geq \int_{S_g^{m-1}} \int_0^{r_\theta} \varphi\left(\exp_p(\theta, r)\right) \frac{\partial}{\partial r}\mathrm{J}_g(\theta, r) dr\, d\Theta(\theta)$$

$$= -\int_{S_g^{m-1}} \int_0^{r_\theta} \frac{\partial}{\partial r}\left(\varphi \circ \exp_p\right)(\theta, r)\mathrm{J}_g(\theta, r) dr d\Theta(\theta) + \int_{S_g^{m-1}} \varphi\left(\exp_p(\theta, r_\theta)\right) \mathrm{J}_g(\theta, r_\theta) d\Theta(\theta)$$

$$\geq \; -\int_{S_g^{m-1}} \int_0^{r_\theta} \frac{\partial}{\partial r}\left(\varphi \circ \exp_p\right)(\theta, r)\mathrm{J}_g(\theta, r) dr\, d\Theta(\theta).$$

By the Gauss lemma we arrive at

$$\int_{\mathcal{M}} \varphi(x) H_K(d_g(x)) dV_g(x) \geq -\int_{\mathcal{M}} \langle \nabla_g \varphi, \nabla_g r_g \rangle_g dV_g = \int_{\mathcal{M}} r_g \Delta_g \varphi dV_g,$$

where the last equality follows from the fact that r_g is Lipschitz on $\mathcal{M}$ and the divergence theorem holds for Lipschitz functions. □

Using $x \coth x \leq 1 + x$ yields

Corollary 2.11

If $(\mathcal{M}, g)$ is a complete Riemannian manifold with $\mathrm{Rc}_g \geq (m-1)K$, where $K \leq 0$, then

$$\Delta_g d_g \leq \frac{m-1}{d_g} + (m-1)\sqrt{|K|} \tag{2.12.4}$$

in the sense of distributions. In particular, if $(\mathcal{M}, g)$ is a complete Riemannian manifold with $\mathrm{Rc}_g \geq 0$, then for any $p \in \mathcal{M}$

$$\Delta_g d_g \leq \frac{m-1}{d_g} \tag{2.12.5}$$

in the sense of distributions. ♡

Estimate (2.12.1) is sharp as can be seen from considering space forms of constant curvature $-K$. If $K = 0$, then (2.12.5) is sharp since on Euclidean space $\Delta|x| = \frac{m-1}{|x|}$.

2.12.2 Volume comparison theorem

A consequence of the Laplacian comparison theorem is

Theorem 2.33. (Bishop-Gromov volume comparison)

If $(\mathcal{M}, g)$ is a complete Riemannian manifold with $\mathrm{Rc}_g \geq (m-1)K$, where $K \in \mathbf{R}$, then for any $p \in \mathcal{M}$, the volume radio

$$\frac{\mathrm{Vol}_g(B_g(p,r))}{\mathrm{Vol}_K(B_K(p_K,r))}$$

is a nonincreasing function of r, where p_K is a point in the m-dimensional simply-connected space form of constant curvature K. In particular,

$$\mathrm{Vol}_g(B_g(p,r)) \leq \mathrm{Vol}_K(B_K(p_K,r)) \tag{2.12.6}$$

for all $r > 0$. Given $p \in \mathcal{M}$ and $r > 0$, equality holds in (2.12.6) if and only if $B_g(p,r)$ is isomorphic to $B_K(p_K,r)$. ♡

Proof. Given a point $p_K \in \mathcal{M}_K$, let $\psi_{p_K} : T_{p_K}\mathcal{M}_K \setminus \{0\} \to S^{m-1}_{p_K}$ be the standard projection $\psi_{p_k}(V) := V/|V|_{p_K}$. The volume element of the space form satisfies

$$dV_K := \sqrt{\det(g_K)}\, d\theta^1_K \wedge \cdots \wedge d\theta^{m-1}_K \wedge dr_K = s^{m-1}_K(r_g)\, d\sigma_K \wedge dr_K,$$

where $d\sigma_K$ is the pull-back by $\psi_{p_K} \circ \exp^{-1}_{p_K}$ of the standard volume form on the unit sphere $S^{m-1}_{p_K}$. If $(\theta^i)_{1\leq i\leq m-1}$ are coordinates on $S^{m-1}_{p_K}$, then

$$\theta^i_K = \theta^i \circ \psi \circ \exp^{-1}_{p_K}, \quad i = 1, \cdots, m-1.$$

From

$$d\sigma_K = (\psi \circ \exp^{-1}_{p_K})^* \left(d\theta^1 \wedge \cdots \wedge d\theta^{m-1}\right) = d\theta^1_K \wedge \cdots \wedge d\theta^{m-1}_K,$$

we get

$$J_K := \sqrt{\det(g_K)} = s^{m-1}_K(r_g).$$

When $K \leq 0$ the above formula holds for all $r_g > 0$ and when $K > 0$ we need to assume $r_g \in (0, \pi/\sqrt{K})$.

Now we consider a Riemannian manifold $(\mathcal{M}, g)$ with $\mathrm{Rc}_g \geq (m-1)K$. From (2.11.19) and (2.11.40) we obtain

$$\frac{\partial}{\partial r_g} \ln \frac{\sqrt{\det(g)}}{s^{m-1}_K(r_g)} \leq 0. \tag{2.12.7}$$

Assume that the coordinates $(\theta^i)^{m-1}_{i=1}$ on S^{m-1}_p are such that $\lim_{r\to 0+} \frac{1}{r}\frac{\partial}{\partial\theta^i} := e_i \in T_p\mathcal{M}$ are orthonormal. Then we have

$$\lim_{r_g\to 0+} \frac{\sqrt{\det(g)}}{s^{m-1}_K(r_g)} = 1$$

from which we conclude

$$\mathrm{J}_g \leq s^{m-1}_K(r_g). \tag{2.12.8}$$

Without making any normalizing assumption on the coordinates $(\theta^i)^{m-1}_{i=1}$, this says

$$\mathrm{J}_g(\theta_g, r_g)\, d\Theta_g(\theta_g) \leq s^{m-1}_K(r_g)\, d\sigma_{S_g(p,r)}(\theta_g).$$

Equivalently,

$$\mathrm{J}(\theta,r)d\Theta(\theta) \le s_K^{m-1}(r)d\sigma_{S_p^{m-1}}(\theta).$$

This is the infinitesimal are comparison formula which gives us

$$dV_g \le dV_K. \tag{2.12.9}$$

Integrating this proves (2.12.6), at least within the cut locus. To see that this result holds on the whole manifold, we argue as follows. Let

$$C_g(r) := \{V \in T_p\mathcal{M} : |V|_{g(p)} = 1 \text{ and } \gamma_V(s) = \exp_p(sV),\ s \in [0,r],\ \text{is minimizing}\}.$$

Note that $C_g(r_2) \subset C_g(r_1)$ for $r_1 \le r_2$. since the cut locus of p has measure zero and $\exp_p^*(dV_g) = \mathrm{J}\, d\Theta \wedge dr$ inside the cut locus of p, for any integrable function φ on a geodesic ball $B_g(p,\bar{r})$ we have

$$\int_{B_g(p,\bar{r})} \varphi(x)\,dV_g(x) = \int_0^{\bar{r}} \left(\int_{C_g(r)} \varphi\left(\exp_p(\theta,r)\right) \mathrm{J}(\theta,r)\,d\Theta(\theta)\right) dr.$$

In particular,

$$\begin{aligned}
\mathrm{Vol}_g(B_g(p,\bar{r})) &= \int_{B_g(p,\bar{r})} dV_g = \int_0^{\bar{r}} \int_{C_g(r)} \exp_p^*(dV_g) = \int_0^{\bar{r}} \left(\int_{C_g(r)} \mathrm{J}(\theta,r)\,d\Theta(\theta)\right) dr \\
&\le \int_0^{\bar{r}} \left(\int_{C_g(r)} s_K^{m-1}(r)\,d\sigma_{S_p^{m-1}}(\theta)\right) dr \le \int_0^{\bar{r}} \left(\int_{S_p^{m-1}} s_K^{m-1}(r)\,d\sigma_{S_p^{m-1}}(\theta)\right) dr \\
&= \int_0^{\bar{r}} \left(\int_{S_{p_K}^{m-1}} s_K^{m-1}(r)\,d\sigma_{S_{p_K}^{m-1}}(\theta)\right) dr = \mathrm{Vol}_K(B_K(p_K,\bar{r})).
\end{aligned}$$

This completes the proof of (2.12.6). □

Corollary 2.12

If $(\mathcal{M},g)$ is a complete Riemannian manifold with $\mathrm{Rc}_g \ge 0$, then for any $p \in \mathcal{M}$, the volume ratio

$$\frac{\mathrm{Vol}_g(B_g(p,r))}{r^m}$$

is a nonincreasing function of r. Since

$$\lim_{r\to 0} \frac{\mathrm{Vol}_g(B_g(p,r))}{r^m} = \omega_m,$$

we have

$$\frac{\mathrm{Vol}_g(B_g(p,r))}{r^m} \le \omega_m \tag{2.12.10}$$

for all $r > 0$, where ω_m is the volume of the Euclidean unit m-ball. ♡

Corollary 2.13. (Volume characterization of $\mathbb{R}^m$)

If $(\mathcal{M},g)$ is a complete noncompact Riemannian manifold with $\mathrm{Rc}_g \ge 0$ and if for some $p \in \mathcal{M}$

$$\lim_{r\to\infty} \frac{\mathrm{Vol}_g(B_g(p,r))}{r^m} = \omega_m,$$

then $(\mathcal{M}, g)$ is isomorphic to Euclidean space. ♡

Let $(\mathcal{M}, g)$ be a complete Riemannian manifold and $p \in \mathcal{M}$. Given a measurable subset Γ of the unit sphere $S_p^{m-1} \subset T_p\mathcal{M}$ and $0 < r \leq R < \infty$, define the **annular-type region**:

$$A^{\Gamma}_{g,r,R}(p) := \left\{ \begin{array}{c} x \in \mathcal{M} : r \leq d_g(p,x) \leq R \\ \text{and there exists a unit speed} \\ \text{minimal geodesic } \gamma \text{ from} \\ \gamma(0) = p \text{ to } x \\ \text{satisfying } \gamma'(0) \in \Gamma \end{array} \right\} \subset B_g(p,R) \setminus B_g(p,r). \tag{2.12.11}$$

Note that id $\Gamma = S_p^{m-1}$, then

$$A^{S_p^{m-1}}_{g,r,R}(p) = B_g(p,R) \setminus B_g(p,r).$$

Given $K \in \mathbf{R}$ and a point p_K in the m-dimensional simply-connected space form $(\mathcal{M}_K, g_K)$ of constant curvature K, let $A^{\Gamma}_{g_K,r,R}(p_K)$ denote the corresponding set in the space form.

Theorem 2.34. (Bishop-Gromov relative volume comparison theorem)

Suppose that $(\mathcal{M}, g)$ is a complete Riemannian manifold with $\mathrm{Rc}_g \geq (m-1)K$. If $0 \leq r \leq R \leq S$, $r \leq s \leq S$ and if $\Gamma \subset S_p^{m-1}$ is a measurable subset, then

$$\frac{\mathrm{Vol}_g\left(A^{\Gamma}_{g,s,S}(p)\right)}{\mathrm{Vol}_K\left(A^{\Gamma}_{g_K,s,S}(p_K)\right)} \leq \frac{\mathrm{Vol}_g\left(A^{\Gamma}_{g,r,R}(p)\right)}{\mathrm{Vol}_K\left(A^{\Gamma}_{g_K,r,R}(p_K)\right)}. \tag{2.12.12}$$

♡

Taking $r = s = 0$ and $\Gamma = S_p^{m-1}$ yields **Theorem 2.34**.

Corollary 2.14. (Yau)

Let $(\mathcal{M}, g)$ be a complete noncompact Riemannian manifold with nonnegative Ricci curvature. For any point $p \in \mathcal{M}$, there exists a constant $C = C(g,p,m) > 0$ such that for any $r \geq 1$

$$\mathrm{Vol}_g(B_g(p,r)) \geq Cr. \tag{2.12.13}$$

♡

Proof. Let $x \in \mathcal{M}$ be a point with $d_g(p,x) = r \geq 2$. By **Theorem 2.34** we have

$$\frac{\mathrm{Vol}_g\left(A^{S_x^{m-1}}_{g,r-1,r+1}(x)\right)}{\mathrm{Vol}_g\left(A^{S_x^{m-1}}_{g,0,r-1}(x)\right)} \leq \frac{\mathrm{Vol}_K\left(A^{S_x^{m-1}}_{g_K,r-1,r+1}(p_K)\right)}{\mathrm{Vol}_K\left(A^{S_x^{m-1}}_{g_K,0,r-1}(p_K)\right)}$$

giving us

$$\frac{\mathrm{Vol}_g(B_g(x,r+1)) - \mathrm{Vol}_g(B_g(x,r-1))}{\mathrm{Vol}_g(B_g(x,r-1))} \leq \frac{(r+1)^m - (r-1)^m}{(r-1)^m} \leq \frac{C(m)}{r}$$

for some constant $C(m)$ depending only on m. Since $B_g(p,1) \subset B_g(x,r+1) \setminus B_g(x,r-1)$ and $B_g(x,r-1) \subset B_g(p,2r-1)$, it follows that

$$\mathrm{Vol}_g(B_g(p,2r-1)) \geq \mathrm{Vol}_g(B_g(x,r-1)) \geq \frac{\mathrm{Vol}_g(B_g(p,1))}{C(m)} r.$$

We have proved the corollary for $r \geq 3$. Clearly it is then true for any $r \geq 1$. □

Example 2.20

A simple example of a complete Riemannian manifold with nonnegative sectional curvature and linear volume growth is $\mathbf{S}^{m-1} \times \mathbf{R}$ *(we may replace* $\mathbf{S}^{m-1}$ *by any closed manifold with nonnegative sectional curvature). If we want* $\mathcal{M}$ *to also have positive sectional curvature at least at one point, then we may take a cylinder* $\mathbf{S}^{m-1} \times [0, \infty)$*, attach a hemispherical cap, and then smooth out the metric.*

If $\mathrm{Rc}_g \geq 0$, then by (2.11.22) and (2.12.5) we have

$$H = \Delta_g d_g \leq \frac{m-1}{d_g}. \tag{2.12.14}$$

Hence the area $A_g(r)$ of the distance sphere $S_g(p,r)$ satisfies

$$\frac{d}{dr}A_g(r) = \int_{S_g(p,r)} H d\sigma \leq \int_{S_g(p,r)} \frac{m-1}{r_g} d\sigma = \frac{m-1}{r} A_g(r).$$

Integrating this yields

$$A_g(s) \leq A_g(r)\frac{s^{m-1}}{r^{m-1}}, \quad s \geq r.$$

Therefore

$$\mathrm{Vol}_g(B_g(p,r)) = \int_0^r A_g(\rho)d\rho \geq \int_0^r A_g(r)\frac{\rho^{m-1}}{r^{m-1}}d\rho = \frac{r}{m}A_g(r), \tag{2.12.15}$$

we obtain

$$A_g(s) \leq m\frac{\mathrm{Vol}_g(B_g(p,r))}{r^m}s^{m-1}, \quad s \geq r, \tag{2.12.16}$$

or

$$\frac{A_g(s)}{m\omega_m s^{m-1}} \leq \frac{\mathrm{Vol}_g(B_g(p,r))}{\omega_m r^m}, \quad s \geq r. \tag{2.12.17}$$

Let

$$\mathbf{AVR}_g(p) := \lim_{r\to\infty}\frac{\mathrm{Vol}_g(B_g(p,r))}{\omega_m r^m} \tag{2.12.18}$$

be the **asymptotic volume ratio**. The asymptotic volume ratio is an important invariant of the geometry at infinity of a complete noncompact manifold with nonnegative Ricci curvature.

Corollary 2.15

On a complete noncompact Riemannian manifold with nonnegative Ricci curvature we have

$$\frac{A_g(s)}{m\omega_m s^{m-1}} \geq \mathbf{AVR}_g(p) \tag{2.12.19}$$

for any $s \geq 0$.

Proof. Since $A_g(s)/s^{m-1}$ is nonincreasing, it follows that

$$\begin{aligned}\frac{\mathrm{Vol}_g(B_g(p,s)) - \mathrm{Vol}_g(B_g(p,r))}{\omega_m(s^m - r^m)} &= \frac{1}{\omega_m(s^m - r^m)}\int_r^s A_g(\rho)d\rho \\ \leq \frac{1}{\omega_m(r^m - s^m)}\int_r^s A_g(r)\frac{\rho^{m-1}}{r^{m-1}}d\rho &= \frac{A_g(r)}{\omega_m r^{m-1}}\frac{\int_r^s \rho^{m-1}d\rho}{r^m - s^m} = \frac{A_g(r)}{m\omega_m r^{m-1}}\end{aligned}$$

for any $s \geq r$. Letting $s \to \infty$ yields (2.12.19). ∎

2.12.3 Hessian comparison theorem

Proposition 2.20. (Hessian comparison theorem–general version)

Let $i = 1, 2$. Let $(\mathcal{M}_i, g_i)$ be complete Riemannian manifolds, let $\gamma_i : [0, L] \to \mathcal{M}_i$ be geodesics parametrized by arc length such that γ_i does not intersect the cur locus of $\gamma_i(0)$, and let $d_{g_i}(\cdot) := d_{g_i}(\gamma_i(0))$. If for all $t \in [0, L]$ we have

$$\mathrm{Sec}_{g_1}(V_1 \wedge \dot{\gamma}_1(t)) \geq \mathrm{Sec}_{g_2}(V_2 \wedge \dot{\gamma}_2(t))$$

for all unit vectors $V_i \in T_{\gamma_i(t)}\mathcal{M}_i$ perpendicular to $\dot{\gamma}_i(t)$, then

$$\nabla^2_{g_1} d_{g_1}(X_1, X_1) \leq \nabla^2_{g_2} d_{g_2}(X_2, X_2) \tag{2.12.20}$$

for all $X_i \in T_{\gamma_i(t)}\mathcal{M}_i$ perpendicular to $\dot{\gamma}_i(t)$ and $t \in (0, L]$. ♡

Theorem 2.35. (Hessian comparison theorem–special case)

Let $(\mathcal{M}, g)$ be a complete Riemannian manifold with $\mathrm{Sec}_g \geq K$. For any point $p \in \mathcal{M}$ the distance function $r_g(x) := d_g(p, x)$ satisfies

$$\nabla_i \nabla_j r_g = h_{ij} \leq \frac{1}{m-1} H_K(r_g) g_{ij} \tag{2.12.21}$$

at all points where r_g is smooth (i.e., away from p and the cut locus). On all of $\mathcal{M}$ the above inequality holds in the sense of support functions. ♡

Proof. From (2.11.27), we have

$$\nabla_{\frac{\partial}{\partial r_g}} h \leq -Kg - h^2$$

along a geodesic ray $\gamma : [0, L) \to \mathcal{M}$ emanating from p. We claim that

$$h(r_g, \theta_g) \leq \frac{1}{m-1} H_K(r_g) g(r_g, \theta_g). \tag{2.12.22}$$

Indeed, given any unit vector V at p, we parallel translate it along γ. Let $V(r_g) := V(\gamma(r_g))$; then $|V(r_g)|_{g(\gamma(r_g))} = 1$ and $\nabla_{\partial/\partial r_g} V(r_g) = 0$. Hence

$$\begin{aligned} \frac{d}{dr_g}\left[h(V(r_g), V(r_g))\right] &= \nabla_{\frac{\partial}{\partial r_g}} h(V(r_g), V(r_g)) + 2h\left(\nabla_{\frac{\partial}{\partial r_g}} V(r_g), V(r_g)\right) \\ &\leq -K|V(r_g)|_{g(\gamma(r_g))} - [h(V(r_g), V(r_g))]^2 \quad = \quad -K - [h(V(r_g), V(r_g))]^2. \end{aligned}$$

From (2.11.20) and (2.11.21) we have

$$h(V(r_g), V(r_g)) - \frac{H_K(r_g)}{m-1} = \frac{1}{r_g} + O(r_g) - \frac{1}{r_g} - O(r_g) = O(r_g).$$

Consequently

$$\begin{aligned} & h(V(r_g), V(r_g)) - \frac{H_K(r_g)}{m-1} \\ \leq \; & \left[h(V(\epsilon), V(\epsilon)) - \frac{H_K(\epsilon)}{m-1}\right] \exp\left[-\int_\epsilon^r \left(\frac{H_K(s)}{m-1} + h(V(s), V(s))\right) ds\right] \end{aligned}$$

which gives

$$h(V(r_g), V(r_g)) - \frac{1}{m-1} H_K(r_g) \leq 0$$

for all $r_g > 0$. Hence

$$\nabla_i \nabla_j r_g = h_{ij} \leq \frac{1}{m-1} H_K(r_g) g_{ij}$$

inside the cut locus when $\mathrm{Sec}_g \geq K$. □

Note that the Hessian of the distance function is the second fundamental form of the distance sphere, which in turn is the radial derivative of the metric. yielding information about the inner products of the Jacobi fields $\frac{\partial}{\partial \theta_g^i}$:

$$\nabla_i \nabla_j r_g = h_{ij} = \frac{1}{2} \frac{\partial}{\partial r_g} g_{ij} = \frac{1}{2} \frac{\partial}{\partial r_g} \left\langle \frac{\partial}{\partial \theta_g^i}, \frac{\partial}{\partial \theta_g^j} \right\rangle_g . \tag{2.12.23}$$

If J_1 and J_2 are Jacobi fields along a geodesic $\gamma : [0, L] \to \mathcal{M}$ without conjugate points and if $J_i(0) = 0$ and $\langle (\nabla_g)_{\dot\gamma} J_i(0), \dot\gamma(0) \rangle_{g(\gamma(0))} = 0$ for $i = 1, 2$, then we have

$$\frac{1}{2} \frac{\partial}{\partial r_g} \langle J_1, J_2 \rangle_g = (\nabla_g)_{J_1} (\nabla_g)_{J_2} r_g = h(J_1, J_2). \tag{2.12.24}$$

Corollary 2.16

Let $(\mathcal{M}, g)$ be Riemannian manifold with $\mathrm{Sec}_g \geq K$ and let $\gamma : [0, L] \to \mathcal{M}$ be a unit speed geodesic. If J is a Jacobi field along γ $J(0) = 0$ and $\langle (\nabla_g)_{\dot\gamma} J(0), \dot\gamma(0) \rangle_{g(\gamma(0))} = 0$, then

$$|J(r_g)|_{g(\gamma(r_g))} \leq \left| (\nabla_g)_{\dot\gamma(0)} J(0) \right|_{g(\gamma(0))} s_K(r_g). \tag{2.12.25}$$

♡

Proof. By our hypotheses, $\langle J(r_g), \dot\gamma(r_g) \rangle_{g(\gamma(r_g))} = 0$ for all $r_g \geq 0$. From (2.12.22) and (2.12.24),

$$\begin{aligned}
\frac{\partial}{\partial r_g} \left(\frac{|J(r_g)|_{g(\gamma(r_g))}}{s_K(r_g)} \right) &= \frac{\partial}{\partial r_g} \left(\frac{\langle J(r_g), J(r_g) \rangle^{1/2}_{g(\gamma(r_g))}}{s_K(r_g)} \right) \\
&= \frac{\frac{1}{2|J(r_g)|_{g(\gamma(r_g))}} \frac{\partial}{\partial r_g} \langle J(r_g), J(r_g) \rangle_{g(\gamma(r_g))} s_K(r_g) - |J(r_g)|_{g(\gamma(r_g))} s_K'(r_g)}{s_K^2(r_g)} \\
&= \frac{1}{|J(r_g)|_{g(\gamma(r_g))} s_K(r_g)} h(J(r_g), J(r_g)) - \frac{s_K'(r_g)}{s_K(r_g)} \frac{|J(r_g)|_{g(\gamma(r_g))}}{s_K(r_g)} \\
&= \left[h\left(\frac{J(r_g)}{|J(r_g)|_{g(\gamma(r_g))}}, \frac{J(r_g)}{|J(r_g)|_{g(\gamma(r_g))}} \right) - \frac{H_K(r_g)}{m-1} \right] \frac{|J(r_g)|_{g(\gamma(r_g))}}{s_K(r_g)} \leq 0.
\end{aligned}$$

The result now follows from $\lim_{r_g \to 0} |J(r_g)|_{g(\gamma(r_g))} / s_K(r_g) = |(\nabla_g)_{\dot\gamma(0)} J(0)|_{g(\gamma(0))}$. □

Note 2.54

Suppose that $(\mathcal{M}, g)$ is a Riemannian manifold with constant sectional curvature K. If J is a Jacobi field along a unit speed geodesic γ with $J(0) = 0$ and $\langle (\nabla_g)_{\dot\gamma} J(0), \dot\gamma(0) \rangle_{g(\gamma(0))} = 0$, then

$$|J(r_g)|_{g(\gamma(r_g))} = \left| (\nabla_g)_{\dot\gamma(0)} J(0) \right|_{g(\gamma(0))} s_K(r_g).$$

In general we have the expansion

$$\begin{aligned}|J(r_g)|^2_{g(\gamma(r_g))} &= r_g^2 - \frac{1}{3}\left\langle \mathrm{Rm}_g\left(\nabla_{\dot\gamma(0)}J(0), \dot\gamma(0)\right)\dot\gamma(0), \nabla_{\dot\gamma(0)}J(0)\right\rangle_{g(\gamma(0))} r_g^4 \\ &\quad + O(r_g^5). \end{aligned} \tag{2.12.26}$$

♣

Finally we consider the Hessian in spherical coordinates. We have

$$\begin{aligned}\nabla_m\nabla_m &= \frac{\partial^2}{\partial r_g^2} - \Gamma^a_{mm}\frac{\partial}{\partial x^a} = \frac{\partial^2}{\partial r_g^2}, \\ \nabla_m\nabla_i &= \frac{\partial^2}{\partial r_g \partial\theta_g^i} - \Gamma^a_{mi}\frac{\partial}{\partial x^a} = \frac{\partial^2}{\partial r_g\partial\theta^i_g} - h^j_i\frac{\partial}{\partial\theta^j_g}, \\ \nabla_i\nabla_j &= \frac{\partial^2}{\partial\theta^i_g\partial\theta^j_g} - \Gamma^a_{ij}\frac{\partial}{\partial x^a} = \nabla^S_i\nabla^S_j + h_{ij}\frac{\partial}{\partial r_g},\end{aligned}$$

where ∇^S is the intrinsic covariant derivative of the hypersurface $S_g(p,r)$. In particular, if $f = f(r_g)$ is a radial function, then

$$\nabla_m\nabla_m f = \frac{\partial^2 f}{\partial r_g^2}, \quad \nabla_m\nabla_i f = 0, \quad \nabla_i\nabla_j f = h_{ij}\frac{\partial f}{\partial r}.$$

2.12.4 Mean value inequalities

The following mean value inequality that is a consequence of the Laplacian comparison theorem, has an application in the proof of the splitting theorem.

Proposition 2.21. (Mean value inequality for $\mathrm{Rc}_g \geq 0$)

If $(\mathcal{M}, g)$ is a complete Riemannian manifold with $\mathrm{Rc}_g \geq 0$ and if $f \leq 0$ is a Lipschitz function with $\Delta_g f \geq 0$ in the sense of distribution, then for any $x \in \mathcal{M}$ and $0 < r < \mathrm{inj}_g(x)$,

$$f(x) \leq \frac{1}{\omega_m r^m}\int_{B_g(x,r)} f\, dV_g. \tag{2.12.27}$$

♡

Proof. By the divergence theorem for Lipschitz functions, we have

$$0 \leq \frac{1}{r^{m-1}}\int_{B_g(x,r)}\Delta_g f\, dV_g = \int_{\partial B_g(x,r)}\frac{\partial f}{\partial r}\frac{1}{r^{m-1}}\sqrt{\det(g)}\, d\Theta_g,$$

where $d\Theta_g = d\theta^1_g\wedge\cdots\wedge d\theta^{m-1}_g$. Since

$$\begin{aligned}\frac{\partial}{\partial r_g}\frac{\sqrt{\det(g)}}{r_g^{m-1}} &= \frac{\partial}{\partial r_g}\left(\frac{\mathrm{J}_g}{r_g^{m-1}}\right) = \frac{\frac{\partial}{\partial r_g}\mathrm{J}_g\cdot r_g^{m-1} - \mathrm{J}_g(m-1)r_g^{m-2}}{r_g^{2m-2}} \\ &= \frac{\frac{\partial}{\partial r_g}\mathrm{J}_g\cdot r_g - \mathrm{J}_g(m-1)}{r_g^m} \leq \frac{\frac{m-1}{r_g}\mathrm{J}_g r_g - \mathrm{J}_g(m-1)}{r_g^m} = 0\end{aligned}$$

from $\Delta_g r_g = H = \frac{\partial}{\partial r_g} \ln \mathrm{J}_g \le \frac{m-1}{r_g}$ and $f \le 0$, we have

$$\begin{aligned} 0 &\le \int_{\partial B_g(x,r)} \left(\frac{\partial f}{\partial r} \frac{\sqrt{\det(g)}}{r^{m-1}} + f \frac{\partial}{\partial r} \frac{\sqrt{\det(g)}}{r^{m-1}} \right) d\Theta_g \\ &= \int_{\partial B_g(x,r)} \frac{\partial}{\partial r} \left(f \frac{\sqrt{\det(g)}}{r^{m-1}} \right) d\Theta_g \;=\; \frac{d}{dr} \left(\frac{1}{r^{m-1}} \int_{\partial B_g(x,r)} f \, d\sigma_g \right), \end{aligned}$$

where $d\sigma_g = \sqrt{\det(g_{ij})} d\Theta_g$. Since

$$\lim_{r \to 0} \frac{1}{r^{m-1}} \int_{\partial B_g(x,r)} f \, d\sigma_g = m\omega_m f(x),$$

where $m\omega_m$ is the volume of the unit $(m-1)$-sphere, integrating the above inequality over $[0, s]$ yields

$$m\omega_m f(x) \le \frac{1}{s^{m-1}} \int_{\partial B_g(x,s)} f \, d\sigma_g.$$

Integrating this again over $[0, r]$ implies

$$f(x) \frac{r^m}{m} \le \frac{1}{m\omega_m} \int_{B_g(x,r)} f \, dV_g$$

which is the desired inequality (2.12.27). □

Proposition 2.22. (Mean value inequality for $\mathrm{Sec}_g \le H$)

Suppose that $(\mathcal{M}, g)$ is a complete Riemannian manifold with $\mathrm{Sec}_g \le H$ in a ball $B_g(x, r)$ where $r < \mathrm{inj}_g(\mathcal{M})$. If $f \in C^\infty(\mathcal{M})$ is subharmonic, i.e., if $\Delta_g f \ge 0$, and if $f \ge 0$ on $\mathcal{M}$, then

$$f(x) \le \frac{1}{V_H(r)} \int_{B_g(x,r)} f \, dV_g, \tag{2.12.28}$$

where $V_H(r)$ is the volume of a ball of radius r in the complete simply-connected manifold of constant sectional curvature H. ♡

2.12.5 Rauch comparison theorem

More generally, applying standard ODE comparison theory to the Jacobi equation, one has the following

Theorem 2.36. (Rauch comparison theorem)

Let $(\mathcal{M}, g)$ and $(\bar{\mathcal{M}}^m, \bar{g})$ be Riemannian manifolds and let $\gamma : [0, L] \to \mathcal{M}$ and $\bar{\gamma} : [0, L] \to \bar{\mathcal{M}}^m$ be unit speed geodesics. Suppose that $\bar{\gamma}$ has no conjugate points and for any $r \in [0, L]$ and any $X \in T_{\gamma(r)}\mathcal{M}$, $\bar{X} \in T_{\bar{\gamma}(r)}\bar{\mathcal{M}}^m$, we have

$$\mathrm{Sec}_g \left(X \wedge \dot{\gamma}(r) \right) \le \mathrm{Sec}_{\bar{g}} \left(\bar{X} \wedge \dot{\bar{\gamma}}(r) \right).$$

If J and $\bar{J}$ are Jacobi fields along γ and $\bar{\gamma}$ with $J(0)$ and $\bar{J}(0)$ tangent to γ and $\bar{\gamma}$, and if

$$\begin{aligned} |J(0)|_{g(\gamma(0))} &= |\bar{J}(0)|_{g_0(\bar{\gamma}(0))}, \\ \langle (\nabla_g)_{\dot{\gamma}(0)} J(0), \dot{\gamma}(0) \rangle_{g(\gamma(0))} &= \left\langle (\overline{\nabla}_{\bar{g}})_{\dot{\bar{\gamma}}(0)} \bar{J}(0), \dot{\bar{\gamma}}(0) \right\rangle_{\bar{g}(\bar{\gamma}(0))}, \\ \left| (\nabla_g)_{\dot{\gamma}(0)} J(0) \right|_{g(\gamma(0))} &= \left| (\overline{\nabla}_{\bar{g}})_{\dot{\bar{\gamma}}(0)} \bar{J}(0) \right|_{\bar{g}(\bar{\gamma}(0))}, \end{aligned}$$

then

$$|J(r)|_{g(\gamma(r))} \ge |\bar{J}(r)|_{\bar{g}(\bar{\gamma}(r))}. \tag{2.12.29}$$

♡

Corollary 2.17. (Cartan-Hadamard theorem)

If $(\mathcal{M}, g)$ is a complete Riemannian manifold with nonpositive sectional curvature, then for any $p \in \mathcal{M}$, the exponential map $\exp_p : T_p\mathcal{M} \to \mathcal{M}$ is a covering map. In particular, the universal cover of $\mathcal{M}$ is diffeomorphic to Euclidean space $\mathbf{R}^m$. ♡

2.13 Manifolds with nonnegative curvature

Introduction

- *The topological sphere theorem and soul theorem*
- *Cheeger-Gromoll splitting theorem*
- *Topological comparision theorem*

2.13.1 The topological sphere theorem

Given a Riemannian manifold $(\mathcal{M}, g)$, let $\mathrm{Sec}_g(\Pi)$ denote the sectional curvature of a 2-plane $\Pi \subset T_p\mathcal{M}$ where $p \in \mathcal{M}$. The Rauch-Klingenberg-Berger **topological sphere theorem** says the following.

Theorem 2.37. (Topological sphere theorem)

If $(\mathcal{M}, g)$ is a complete, simply-connected Riemannian manifold with $\frac{1}{4} < \mathrm{Sec}_g(\Pi) \le 1$ for all 2-planes Π, then $\mathcal{M}$ is homeomorphic to the m-sphere. In particular, if $m = 3$, then $\mathcal{M}^3$ is diffeomorphic to the 3-sphere. ♡

Recently, Brendle and Schoen showed that

Theorem 2.38. (Diffeomorphic sphere theorem)

If $(\mathcal{M}, g)$ is a complete, simply-connected Riemannian manifold with $\frac{1}{4} < \mathrm{Sec}_g(\Pi) \le 1$ for all 2-planes Π, then $\mathcal{M}$ is diffeomorphic to the m-sphere. ♡

There is not much known about general closed Riemannian manifolds with positive sectional curvature.

Problem 2.2. (Hopf conjecture I)

p2.2 Does there exist a Riemannian metric on $\mathbf{S}^2 \times \mathbf{S}^2$ with positive sectional curvature? ♠

Problem 2.3. (Hopf conjecture II)

Prove that if $(\mathcal{M}^{2m}, g)$ is a closed, even-dimensional Riemannian manifold with positive sectional curvature, then $\chi(\mathcal{M}^{2m}) > 0$. ♠

Note that any closed, odd-dimensional manifold has $\chi(\mathcal{M}^{2m+1}) = 0$. The case of complete noncompact manifolds with positive sectional curvature is simpler.

2.13.2 Cheeger-Gromoll splitting theorem and soul theorem

In the study of manifolds with nonnegative curvature, often (especially when the curvature is not strictly positive) the manifolds split as the product of a lower-dimensional manifold with a line.

A **geodesic line** is a unit speed geodesic $\gamma : (-\infty, \infty) \to \mathcal{M}$ such that the distance between any points on γ is the length of the arc of γ between those two points; that is, for any $s_1, s_2 \in (-\infty, \infty)$, $d_g(\gamma(s_1), \gamma(s_2)) = |s_2 - s_1|$. A unit speed geodesic $\beta : [0, \infty) \to \mathcal{M}$ is a **geodesic ray** if it satisfies the same condition as above. Given a geodesic ray $\beta : [0, \infty) \to \mathcal{M}$, the **Busemann function**

$$b_\beta : \mathcal{M} \longrightarrow \mathbf{R} \tag{2.13.1}$$

associated to β is defined by

$$b_\beta(x) := \lim_{s\to\infty} \left(s - d_g(\beta(s), x)\right). \tag{2.13.2}$$

Note 2.55

(1) In Euclidean space the Busemann function is linear. For any unit vector $V \in \mathbf{R}^m$, the Busemann function b_{γ_V} associated to the geodesic ray $\gamma_V : [0, \infty) \to \mathbf{R}^m$ defined by $\gamma_V(s) := sV$ is the linear function given by

$$b_{\gamma_V}(x) = \langle x, V\rangle$$

for all $x \in \mathbf{R}^m$.

(2) The Busemann function is well-defined, finite, and Lipschitz.

(3) $|\nabla_g b_\beta|_g = 1$ at points where it is C^1.

(4) If β is a geodesic ray in a Riemannian manifold with $\mathrm{Rc}_g \geq 0$, then $\Delta_g b_\beta \geq 0$ in the sense of distributions. Indeed, using (2.12.5) yields

$$\Delta_g b_\beta(x) \geq -\frac{m-1}{\lim_{s\to\infty} d_g(\beta(s), x)} = 0.$$

♣

Theorem 2.39. (Cheeger-Gromoll)

Suppose $(\mathcal{M}, g)$ is a complete noncompact Riemannian manifold with $\mathrm{Rc}_g \geq 0$ and suppose that there is a geodesic line in $\mathcal{M}$. Then $(\mathcal{M}, g)$ is isomorphic to $\mathbf{R} \times (\mathcal{N}^{m-1}, h)$ with the product metric, where $(\mathcal{N}^{m-1}, h)$ is a Riemannian manifold with $\mathrm{Rc}_h \geq 0$. ♡

Proof. Given a geodesic line γ, consider the two Busemann functions $b_{\gamma_\pm}$ associated to the geodesic rays $\gamma_\pm : [0, \infty) \to \mathcal{M}$ defined by $\gamma_\pm(s) = \gamma(\pm s)$ for $s \geq 0$. Since $\mathrm{Rc}_g \geq 0$, we have $\Delta_g b_{\gamma_\pm} \geq 0$ in the sense of distributions and hence $\Delta_g(b_{\gamma_+} + b_{\gamma_-}) \geq 0$. From $d_g(\gamma(s), \gamma(-s)) = 2s$, we note that for any $x \in \mathcal{M}$

$$\begin{aligned} b_{\gamma_+}(x) + b_{\gamma_-}(x) &= \lim_{s\to\infty} [2s - d_g(\gamma(s), x) - d_g(\gamma(-s), x)] \\ &\leq \lim_{s\to\infty} [2s - d_g(\gamma(s), \gamma(-s))] \;=\; 0. \end{aligned}$$

Using **Proposition 2.22**, we obtain

$$0 = b_{\gamma_+}(x) + b_{\gamma_-}(x) \leq \frac{1}{\omega_m r^m} \int_{B_g(x,r)} [b_{\gamma_+} + b_{\gamma_-}] dV_g \leq 0$$

for any $x \in \gamma$ and $0 < r < \mathrm{inj}_g(x)$. Hence $b_{\gamma_+} + b_{\gamma_-} \equiv 0$ in a neighborhood of γ.

By applying the mean value inequality again, we see that the set of points in $\mathcal{M}$ where $b_{\gamma_+} + b_{\gamma_-} = 0$ is open. Since this set is also closed and nonempty, we have $b_{\gamma_+} + b_{\gamma_-} \equiv 0$ on $\mathcal{M}$ and hence also $\Delta_g(b_{\gamma_+} + b_{\gamma_-}) \equiv 0$. Since $\Delta_g b_{\gamma_\pm} \geq 0$, this implies $\Delta_g b_{\gamma_\pm} = 0$ in the sense of distributions. Standard regularity theory of PDE now implies $b_{\gamma_\pm}$ is smooth. Therefore, $|\nabla_g b_{\gamma_\pm}|_g \equiv 1$. Since $\nabla_g b_{\gamma_\pm}$ is a nonzero parallel gradient vector field on $\mathcal{M}$, $(\mathcal{M}, g)$ splits as a Riemannian product $\mathbf{R} \times (\mathcal{N}, h)$ where $\mathcal{N} = \{x \in \mathcal{M} : b_{\gamma_+}(x) = 0\}$. □

Note 2.56

(1) The above result generalizes the Toponogov splitting theorem, which derives the same conclusion, under the stronger hypothesis of nonnegative sectional curvature.

(2) In the study of the Ricci flow on 3-manifolds one of the primary singularity models is the round cylinder $\mathbf{S}^2 \times \mathbf{R}$. This singularity model corresponds to neck pinching. We shall see that the splitting theorem has applications to the Ricci flow. ♣

A submanifold $\mathcal{S} \subset \mathcal{M}$ is **totally convex** if for every $x, y \in \mathcal{S}$ and any geodesic γ (not necessarily minimal) joining x to y we have $\gamma \subset \mathcal{S}$. We say that $\mathcal{S}$ is **totally geodesic** if its second fundamental form is zero. In particular, a path in a totally geodesic submanifold $\mathcal{S}$ is a geodesic in $\mathcal{S}$ if and only if it is a geodesic in $\mathcal{M}$.

Given a noncompact manifold $(\mathcal{M}, g)$, we say that a submanifold is a **soul** if it is a closed, totally convex, totally geodesic submanifold such that $\mathcal{M}$ is diffeomorphic to its normal bundle.

Theorem 2.40. (Cheeger-Gromoll, 1972)

Let $(\mathcal{M}, g)$ be a complete noncompact Riemannian manifold with nonnegative sectional curvature. Then there exists a soul. If the sectional curvature is positive, then the soul is *a point (e.g., $\mathcal{M}$ is diffeomorphic to $\mathbf{R}^m$).* ♡

Furthermore Sharafurdinov proved that any two souls are isometric. An important tool in the study of manifolds with nonnegative curvature is the **Sharafutdinov retraction**.

Theorem 2.41. (Soul conjecture; Perelman, 1994)

If $(\mathcal{M}, g)$ is a complete noncompact Riemannian manifold with nonnegative sectional curvature everywhere and positive sectional curvature at some point, then the soul is a point. ♡

Another fundamental result about noncompact manifolds with positive sectional curvature is the following.

Theorem 2.42. (Toponogov, 1959)

If $(\mathcal{M}, g)$ is a complete noncompact Riemannian manifold with positive sectional curvature bounded above by K, then

$$\mathrm{inj}(\mathcal{M}, g) \geq \frac{\pi}{\sqrt{K}}. \tag{2.13.3}$$

Moreover, $\mathcal{M}$ is diffeomorphic to Euclidean space. ♡

2.13.3 Topological comparison theorem

As a consequence of **Section** 2.9 we have the following

Lemma 2.31

Let $(\mathcal{M}, g)$ be a complete Riemannian manifold with nonnegative sectional curvature and $p \in \mathcal{M}$. If $\beta : (a, b) \to \mathcal{M}$ is a unit speed geodesic, then the function $\phi : (a, b) \to \mathbf{R}$ defined by

$$\phi(r) := r^2 - d_g^2(p, \beta(r))$$

is convex. ♡

Proof. Given $r_0 \in (a, b)$, let $\gamma_r : [0, L] \to \mathcal{M}$ be a 1-parameter family of paths from p to $\beta(r)$ with $\gamma_{r_0} : [0, L] \to \mathcal{M}$ a unit speed minimal geodesic from p to $\beta(r_0)$ and

$$\frac{\partial}{\partial r}\Big|_{r=r_0} \gamma_r(s) = \frac{s}{L} V(\gamma_{r_0}(s)),$$

where V is the parallel translation of $\dot{\beta}(r_0) \in T_{\gamma_{r_0}(L)}\mathcal{M}$ along γ. Since $|V|_g^2 = 1$, it follows from that (since $\mathrm{Sec}_g \geq 0$)

$$\frac{d^2}{dr^2}\Big|_{r=r_0} \left(r^2 - L_g^2(\gamma_r)\right) \geq 2 - 2 = 0.$$

Since

$$r^2 - d_g^2(p, \beta(r)) \geq r^2 - L_g^2(\gamma_r), \quad r_0^2 - d_g^2(p, \beta(r_0)) = r_0^2 - L_g^2(\gamma_{r_0}),$$

we conclude that ϕ is convex. □

Note 2.57

In general, if $\phi : (a,b) \to \mathbf{R}$ is a Lipschitz function such that for all $r_0 \in (a,b)$ there exists a C^2-function $\psi_{r_0}(r)$ defined in a neighborhood of r_0 with $\psi_{r_0}(r) \leq \phi(r)$, $\psi_{r_0}(r_0) = \phi(r_0)$ and $\frac{d^2}{dr^2}|_{r=r_0}\psi_{r_0}(r) \geq 0$, then ϕ is convex. ♣

Theorem 2.43. (Toponogov comparison theorem–$\mathrm{Sec}_g \geq 0$; Toponogov, 1959)

Let $(\mathcal{M}, g)$ be a complete Riemannian manifold with nonnegative sectional curvature and let $\alpha : [0, A] \to \mathcal{M}$ be a unit speed minimal geodesic joining p to q. If $\beta : [0, B] \to \mathcal{M}$ is a unit speed geodesic with $\beta(0) = q$ and if $\theta \in [0, \pi]$ is the angle between $\dot\beta(0)$ and $-\dot\alpha(A)$, then

$$d_g^2(p, \beta(r)) \leq r^2 + A^2 - 2rA \cdot \cos(\theta)$$

for all $r \in [0, B]$. In particular,

$$d_g^2(p, \beta(B)) \leq A^2 + B^2 - 2AB \cdot \cos(\theta). \tag{2.13.4}$$

By the law of cosines, equality is attained for Euclidean space. That is, the right-hand side of (2.13.4) is the length squared of the side in the corresponding Euclidean triangle with the same A, B and θ. ♡

Proof. For $\epsilon > 0$, let

$$f_\epsilon(r) := r^2 - d_g^2(p, \beta(r)) + A^2 - 2Ar \cdot \cos(\theta) + \epsilon r.$$

Then f_ϵ is convex. We also have

$$f_\epsilon(0) = -d_g^2(p, q) + A^2 = -L_g^2(\alpha) + A^2 = -A^2 + A^2 = 0$$

because α is a unit speed minimal geodesic. By a first variation argument (we may assume $d_g(p, \cdot)$ is smooth at q. Otherwise, we can apply Calabi's trick)

$$\begin{aligned}
\frac{\partial}{\partial r}\Big|_{r=0} f_\epsilon(r) &= \left(2r - 2d_g(p, \beta(r))\left\langle \nabla_g d_g(p, \beta(r)), \dot\beta(r)\right\rangle_g - 2A \cdot \cos(\theta) + \epsilon\right)_{r=0} \\
&= \epsilon - 2A\langle \dot\alpha(A), \dot\beta(0)\rangle_g - 2A \cdot \cos(\theta) \quad = \quad \epsilon > 0,
\end{aligned}$$

$f_\epsilon(r) > 0$ for $r > 0$ small enough, depending on ϵ. Since f_ϵ is convex, we conclude that $f_\epsilon(r) > 0$ for all $r \in (0, B]$. In particular, $\lim_{\epsilon\to 0} f_\epsilon(r) \geq 0$ for all $r \in (0, B]$, which proves the theorem. □

More generally, we have the following statements of the Toponogov comparison theorems for manifolds with a sectional curvature lower bound.

A **geodesic triangle** is a triangle (i.e., three vertices joined by three paths) whose sides are geodesics. The triangle version says that a triangle in a manifold has larger angles than the corresponding triangle with the same side lengths in the simply-connected constant curvature space. Given a triangle (p, q, r), $\measuredangle pqr$ denote the angle at q.

Theorem 2.44. (Toponogov comparison theorem–triangle version)

Let $(\mathcal{M}, g)$ be a complete Riemannian manifold with $\mathrm{Sec}_g \geq K$. Let $\triangle$ be a geodesic triangle with vertices (p, q, r), sides $\overline{qr}, \overline{rp}, \overline{pq}$, corresponding lengths $a = L_g(\overline{qr}), b = L_g(\overline{rp}), c = L_g(\overline{pq})$ such that $a \leq b + c, b \leq a + c, c \leq a + b$ (for example, when all of the geodesic sides are minimal), and interior angles $\alpha = \measuredangle rpq, \beta = \measuredangle pqr, \gamma = \measuredangle qrp$, where $\alpha, \beta, \gamma \in [0, \pi]$. If the geodesics $\overline{qr}$ and $\overline{rp}$ are minimal, and $c \leq \pi/\sqrt{K}$ in the case where $K > 0$ (no assumption on c when $K \leq 0$), then there exists a geodesic triangle $\overline{\triangle} = (\bar{p}, \bar{q}, \bar{r})$ in the complete, simply-connected 2-manifold of constant Gauss curvature K with the same side lengths (a, b, c) and such that we have the following comparison of the interior angles:

$$\alpha \geq \bar{\alpha} := \measuredangle \bar{r}\bar{p}\bar{q}, \quad \beta \geq \bar{\beta} := \measuredangle \bar{p}\bar{q}\bar{r}.$$

♡

A **geodesic hinge** consists of a pair of geodesic segments emanating from a point, called the **vertex**, making an angle at the vertex. The hinge version says that a hinge in a manifold has a smaller distance between its endpoints than the corresponding hinge in the constant curvature space with the same "side-angle-side".

Theorem 2.45. (Toponogov comparison theorem–hinge version)

Suppose $(\mathcal{M}, g)$ is a complete Riemannian manifold with $\mathrm{Sec}_g \geq K$. Let $\angle$ be a geodesic hinge with vertices (p, q, r), sides $\overline{qr}, \overline{rp}$, and interior angle $\measuredangle qrp \in [0, \pi]$ in $\mathcal{M}$. Suppose that $\overline{qr}$ is minimal and that $L_g(\overline{rp}) \leq \pi/\sqrt{K}$ if $K > 0$. Let $\angle'$ be a geodesic hinge with vertices (p', q', r') in the simply-connected space of constant curvature K with the same side lengths $L_g(\overline{q'r'}) = L_g(\overline{qr}), L_g(\overline{r'p'}) = L_g(\overline{rp})$ and same angle $\measuredangle q'r'p' = \measuredangle qrp$. Then we have the following comparison of the distance between the endpoints of the hinges:

$$d_g(p, q) \leq d_K(p', q')$$

where d_K denotes the distance in the simply-connected space of constant sectional curvature K.

♡

2.14 Lie groups and left-invariant metrics

Introduction

- *Lie groups and left-invariant metrics*
- *Left-invariant metrics*
- *Left-invariant vector fields*
- *Curvatures formulas for Lie groups with left-invariant metrics*

we have discussed the curvatures of bi-invariant metrics on Lie groups. In this section, we focus on the curvatures of left-invariant metrics on Lie groups.

2.14.1 Lie groups

A **Lie group** is a smooth manifold G with the structure of a group, such that the map $\mu : G \times G \to G$, defined by $\mu(\sigma, \tau) = \sigma \cdot \tau^{-1}$, is smooth.

Given $\sigma \in G$, we define **left multiplication** by σ:

$$L_\sigma : G \longrightarrow G, \quad \tau \longmapsto \sigma \cdot \tau. \tag{2.14.1}$$

2.14.2 Left-invariant metrics

A Riemannian metric g on G is **left-invariant** if for any $\sigma \in G$, L_σ is an isometry of (G, g):

$$(L_\sigma)^* g = g.$$

Theorem 2.46

Every Lie group admits a left-invariant metric. ♡

Proof. Since T_eG is a vector space, we can find an inner product g_e on T_eG. For any $\sigma \in G$, we set

$$g_\sigma := (L_{\sigma^{-1}})^* g_e.$$

Since G is smooth and the multiplication is also smooth, we obtain a smooth metric g on G. For $\tau \in G$,

$$(L_\tau)^* g_\sigma = (L_\tau)^* \left((L_{\sigma^{-1}})^* g_e\right) = (L_{\sigma^{-1}\cdot\tau})^* g_e = \left(L_{(\tau^{-1}\cdot\sigma)^{-1}}\right)^* g_e = g_{\tau^{-1}\cdot\sigma}.$$

Therefore g is left-invariant. □

2.14.3 Left-invariant vector fields

A vector field X is called left-invariant if

$$(L_\sigma)_* \circ X = X \circ L_\sigma \tag{2.14.2}$$

for any $\sigma \in G$. Let $\mathfrak{g}$ be the space of all left-invariant vector fields on G. Then T_eG, where e is the identity element of G, can be naturally identified with $\mathfrak{g}$. Note also that $\mathfrak{g}$ is a Lie algebra.

For any left-invariant vector fields $Y, Z \in \mathfrak{g}$, we claim that

$$\langle Y, Z\rangle_g = \text{const.} \tag{2.14.3}$$

Since Y and Z are left-invariant, it follows that for any $\sigma \in G$

$$Y_\sigma = (L_\sigma)_{*,e} Y_e, \quad Z_\sigma = (L_\sigma)_{*,e} Z_e.$$

Hence

$$\langle Y, Z\rangle_g(\sigma) = \langle Y_\sigma, Z_\sigma\rangle_{g_\sigma} = \langle (L_\sigma)_{*,e} Y_e, (L_\sigma)_{*,e} Z_e\rangle_{g_\sigma} = \langle Y_e, Z_e\rangle_{g_e} = \langle Y, Z\rangle_g(e).$$

2.14.4 Curvatures formulas for Lie groups with left-invariant metrics

The connection and curvature of a left-invariant metric may be computed algebraically and metrically using the following

Proposition 2.23

Let g be a left-invariant metric on G. If $X, Y, Z, W \in \mathfrak{g}$, then

$$\begin{aligned}
\langle (\nabla_g)_X Y, Z\rangle_g &= \frac{1}{2}\left(\langle [X,Y], Z\rangle_g - \langle [X,Z], Y\rangle_g - \langle [Y,Z], X\rangle_g\right), \\
\mathrm{Rm}_g(X,Y,Z,W) &= \langle (\nabla_g)_X Z, (\nabla_g)_Y W\rangle_g - \langle (\nabla_g)_Y Z, (\nabla_g)_X W\rangle_g \\
&\quad - \langle (\nabla_g)_{[X,Y]} Z, W\rangle_g, \\
\mathrm{Rm}_g(X,Y,Y,X) &= \langle (\nabla_g)_X Y, (\nabla_g)_Y X\rangle_g - \langle (\nabla_g)_Y Y, (\nabla_g)_X X\rangle_g \\
&\quad - \langle \nabla_{[X,Y]} Y, X\rangle_g.
\end{aligned}$$

♡

Lemma 2.32

If $(\mathcal{M}, g)$ is a Riemannian manifold and X, Y, and Z are Killing vector fields, then

$$\langle (\nabla_g)_X Y, Z\rangle_g = \frac{1}{2}\left(\langle [X,Y], Z\rangle_g + \langle [X,Z], Y\rangle_g + \langle [Y,Z], X\rangle_g\right). \tag{2.14.4}$$

♡

Proof. Since Y is a Killing vector field, we have

$$0 = \langle (\nabla_g)_X Y, Z\rangle + \langle (\nabla_g)_Z Y, X\rangle_g.$$

Hence

$$\begin{aligned}
\langle (\nabla_g)_X Y, Z\rangle_g &= -\langle (\nabla_g)_Z Y, X\rangle_g = \langle [Y,Z], X\rangle_g - \langle (\nabla_g)_Y Z, X\rangle_g \\
&= \langle [Y,Z], X\rangle_g + \langle (\nabla_g)_X Z, Y\rangle_g \\
&= \langle [Y,Z], X\rangle_g + \langle [X,Z], Y\rangle_g + \langle (\nabla_g)_Z X, Y\rangle_g \\
&= \langle [Y,Z], X\rangle_g + \langle [X,Z], Y\rangle_g - \langle (\nabla_g)_Y X, Z\rangle_g \\
&= \langle [Y,Z], X\rangle_g + \langle [X,Z], Y\rangle_g + \langle [X,Y], Z\rangle_g - \langle (\nabla_g)_X Y, Z\rangle_g
\end{aligned}$$

implying (2.14.4). □

A left-invariant metric is **bi-invariant** if it is also invariant under right multiplication.

Note 2.58

If (G, g) is bi-invariant and if X is a left-invariant vector field, then X is a Killing vector field. ♣

Lemma 2.33

Let g be a bi-invariant metric on a Lie group G. If X, Y, Z, W are left-invariant vector fields, then

(i) the Levi-Civita connection is given by

$$(\nabla_g)_X Y = \frac{1}{2}[X, Y], \tag{2.14.5}$$

(2) the Riemann curvature tensor field is given by

$$\mathrm{Rm}_g(X, Y, Z, W) = \frac{1}{4}\left(\langle [X, W], [Y, Z]\rangle_g - \langle [X, Z], [Y, W]\rangle_g\right). \tag{2.14.6}$$

♡

Proof. It follows from **Proposition 2.23** and **Lemma 2.32**. □

Corollary 2.18

A bi-invariant metric on a Lie group G has nonnegative sectional curvature. ♡

Proof. Letting $Z = X$ and $W = Y$ we have

$$\mathrm{Rm}_g(X, Y, Y, X) = -\frac{1}{4}\langle [X, Y], [Y, X]\rangle_g = \frac{1}{4}|[X, Y]|_g^2 \geq 0.$$

Hence $\mathrm{Sec}_g \geq 0$. □

Chapter 3 PDEs

Introduction

- *Ricci flow and Hamilton's theorem*
- *Ricci flow equation*
- *The maximum principle for heat-type equations*
- *The Einstein-Hilbert functional*
- *Evolution of geometric quantities*
- *De Turck's trick and short time existence*
- *Reaction-diffusion equation and Uhlenbeck's trick*

3.1 Ricci flow and Hamilton's theorem

Introduction

- *Ricci flow and geometrization*

3.1.1 Ricci flow and geometrization

Theorem 3.1. (Hamilton, 1982; 3-manifolds with positive Ricci curvature)

If $(\mathcal{M}^3, g)$ is a closed[a] 3-manifold with positive Ricci curvature, then it is diffeomorphic to a spherical space form. That is, $\mathcal{M}^3$ admits a metric with constant positive sectional curvature.

[a] Here, closed means compact without boundary.

♡

Problem 3.1. (Thurston geometrization conjecture)

Every closed 3-manifold admits a geometric decomposition.

♠

A corollary of the geometrization conjecture is the **Poincaré conjecture**, which says that every simply-connected closed topological 3-manifold is homeomorphic to the 3-sphere.

3.2 Ricci flow and the evolution of scalar curvature

Introduction

- *Ricci flow equation*
- *Simple examples*
- *Variation of scalar curvature*

3.2.1 Ricci flow equation

Given a 1-parameter family of metrics $g(t)$ on a Riemannian m-manifold $(\mathcal{M}^m, g_0)$, defined on a time interval $\mathfrak{t} \subset \mathbf{R}$, Hamilton's **Ricci flow equation** is

$$\partial_t g(t) = -2\mathrm{Ric}_{g(t)}, \quad \partial_t g_{ij} = -2R_{ij}. \tag{3.2.1}$$

For any C^∞ metric g_0 on a closed manifold $\mathcal{M}^m$, there exists a unique solution $g(t)$, $t \in [0, \epsilon)$, to the Ricci flow equation for some $\epsilon > 0$, with $g(0) = g_0$.

This was proved in Hamilton (1982) and shortly therefore a much simpler proof was given by De Turck (1983).

3.2.2 Simple examples

Let $\mathcal{M}^m = \mathbf{S}^m$ and let $g_{\mathbf{S}^m}$ denote the standard metric on the unit m-sphere in Euclidean space. If $g_0 := r_0^2 g_{\mathbf{S}^m}$ for some $r_0 > 0$ (r_0 is the radius), then

$$g(t) := \left[r_0^2 - 2(m-1)t\right] g_{\mathbf{S}^m} \tag{3.2.2}$$

is a solution to the Ricci flow (3.2.1) with $g(0) = g_0$ defined on the maximal time interval $(-\infty, T)$, where $T := \frac{r_0^2}{2(m-1)}$. That is, under the Ricci flow, the sphere stays round and shrinks at a steady rate.

Example 3.1. (Homothetic Einstein solutions)

Suppose that g_0 is an Einstein metric, i.e., $\mathrm{Ric}_{g_0} \equiv cg_0$ for some $c \in \mathbf{R}$. Derive the explicit formula for the solution $g(t)$ of the Ricci flow with $g(0) = g_0$. $g(t)$ is homothetic to the initial metric g_0 and that it shrinks, is stationary, or expands depending on whether c is positive, zero, or negative, respectively. In fact, $g(t) = a(t)g_0$ where $a(0) = 1$. We then have

$$a'(t)g_0 = \partial_t g(t) = -2\mathrm{Ric}_{g(t)} = -2\mathrm{Ric}_{g_0} = -2cg_0.$$

Hence $a(t) = 1 - 2ct$ so that $g(t) = (1 - 2ct)g_0$.

Example 3.2. (Product solutions)

Let $(\mathcal{M}_1^{m_1}, g_1(t))$ and $(\mathcal{M}_2^{m_2}, g_2(t))$ be solutions of the Ricci flow on a common time interval $\mathfrak{t}$. Show that

$$(\mathcal{M}_1^{m_1} \times \mathcal{M}_2^{m_2}, g_1(t) + g_2(t))$$

is a solution of the Ricci flow. In particular, if $(\mathcal{M}^m, g(t))$ is a solution of the Ricci flow, then so is $(\mathcal{M}^m \times \mathbf{R}, g(t) + dr^2)$.

Some other solutions are the cigar and Rosenau solutions on $\mathbf{R}^2$ and $\mathbf{S}^2$, respectively. In addition, some homogeneous solutions are explicit.

3.2.3 Variation of scalar curvature

Introduce

$$\square_{g(t)} = \partial_t - \Delta_{g(t)}.$$

The **evolution equation for the scalar curvature** is

$$\square_{g(t)} R_{g(t)} = 2 \left|\mathrm{Rc}_{g(t)}\right|^2_{g(t)}. \tag{3.2.3}$$

When $m = 2$, since then $\mathrm{Ric}_g = \frac{R_g}{2} g$, we have

$$\square_{g(t)} R_{g(t)} = R^2_{g(t)}. \tag{3.2.4}$$

Lemma 3.1. (Variation of scalar curvature)

If $\partial_s g_{ij} = v_{ij}$, then

$$\partial_s R_g = -\Delta_g V + \mathrm{div}_g(\mathrm{div}_g v) - \langle v, \mathrm{Ric}_g \rangle_g, \tag{3.2.5}$$

where $V = g^{ij} v_{ij} = \mathrm{tr}_g v$ is the trace of v. ♡

If $v = -2\mathrm{Ric}_g$, then

$$\mathrm{div}_g\left(\mathrm{div}_g v\right) = \nabla^p \nabla^q v_{pq} = -\nabla^p \nabla_p R_g = -\Delta_g R_g.$$

Hence we obtain (3.2.3).

Note 3.1

Let $(\mathcal{M}^2, h)$ be a Riemann surface. If $g = uh$ for some function u on $\mathcal{M}^2$, then

$$R_g = u^{-1}\left(R_h - \Delta_h \ln u\right). \tag{3.2.6}$$

Consequently, $g(t) = u(t)h$ is a solution of the Ricci flow if and only if $u = u(t)$ satisfies

$$\partial_t u = \Delta_h \ln u - R_h. \tag{3.2.7}$$

♣

3.3 The maximum principle for heat-type equations

Introduction

- *The maximum principle*
- *Heat equation and the Ricci flow*
- *Ricci flow on non-compact manifolds*

3.3.1 The maximum principle

For elliptic equations on a Riemannian manifold $(\mathcal{M}^m, g)$, the facts we use are that if a function $f : \mathcal{M}^m \to \mathbf{R}$ attains its minimum at a point $x_0 \in \mathcal{M}^m$, then

$$\nabla_g f(x_0) = 0, \quad \Delta_g f(x_0) \geq 0. \tag{3.3.1}$$

For equations of parabolic type, such as the heat equation, a simple version gives the following

Proposition 3.1. (Maximum principle for super-solutions of the heat equation)

Let $g(t)$ be a family of metrics on a closed m-dimensional manifold $\mathcal{M}^m$ and let $u : \mathcal{M}^m \times [0, T) \to \mathbf{R}$ satisfy

$$\Box_{g(t)} u \geq 0. \tag{3.3.2}$$

Then if $u \geq c$ at $t = 0$ for some $c \in \mathbf{R}$, then $u \geq c$ for all $t \geq 0$. ♡

Proof. Given any $\epsilon > 0$, define $u_\epsilon : \mathcal{M}^m \times [0, T) \to \mathbf{R}$ as

$$u_\epsilon := u + \epsilon(1 + t).$$

Since $u \geq c$ at $t = 0$, we have $u_\epsilon > c$ at $t = 0$. Suppose for some $\epsilon > 0$ we have $u_\epsilon \leq c$ somewhere in $\mathcal{M}^m \times [0, T)$. Since $\mathcal{M}^m$ is closed, there exists $(x_1, t_1) \in \mathcal{M}^m \times (0, T)$ such that $u_\epsilon(x_1, t_1) = c$ and $u_\epsilon(x, t) > c$ for all $x \in \mathcal{M}^m$ and $t \in [0, t_1)$. We then have at (x_1, t_1)

$$0 \geq \Box_{g(t)} u_\epsilon \geq \epsilon > 0,$$

which is a contradiction. Hence $u_\epsilon > c$ on $\mathcal{M}^m \times [0, T)$ for all $\epsilon > 0$ and by taking the limit as $\epsilon \to 0$, we get $u \geq c$ on $\mathcal{M}^m \times [0, T)$. □

Corollary 3.1. (Lower bound of scalar curvature is preserved under the Ricci flow)

If $g(t)$, $t \in [0, T)$, is a solution to the Ricci flow on a closed manifold with $R_{g(0)} \geq c$ at $t = 0$ for some $c \in \mathbf{R}$, then

$$R_{g(t)} \geq c$$

for all $t \in [0, T)$. In particular, nonnegative (positive) scalar curvature is preserved under the Ricci flow. ♡

Lemma 3.2. (Maximum principle)

Suppose $g(t)$ is a family of metrics on a closed manifold $\mathcal{M}^m$ and $u : M \times [0, T) \to \mathbf{R}$ satisfies

$$\Box_{g(t)} u \leq \langle X(t), \nabla_{g(t)} u \rangle_{g(t)} + F(u),$$

where $X(t)$ is a time-dependent vector field and F is a Lipschitz function. If $u \leq c$ at $t = 0$ for some $c \in \mathbf{R}$, then $u(x, t) \leq U(t)$ for all $x \in \mathcal{M}^m$ and $t \geq 0$, where $U(t)$ is the solution to the ODE

$$\frac{dU}{dt} = F(U)$$

with $U(0) = c$. ♡

Let $(\mathcal{M}^m, g(t))$, $t \in [0, T)$, be a solution to the Ricci flow on a closed manifold (or any solution where we can apply the maximum principle to the evolution equation for the scalar curvature). Since

$$|\mathrm{Ric}_g|_g^2 \geq \frac{1}{m} R_g^2$$

it follows that

$$\Box_{g(t)} R_{g(t)} \geq \frac{2}{m} R_{g(t)}^2. \tag{3.3.3}$$

By the maximum principle one has

$$R_{g(t)} \geq \frac{m}{\frac{m}{\inf_{\mathcal{M}^m} R_{g(0)}} - 2t} \tag{3.3.4}$$

over $\mathcal{M}^m$ for all $t \geq 0$. We let

$$R_{\min}(t) := \inf_{\mathcal{M}^m} R_{g(t)}. \tag{3.3.5}$$

We have

Corollary 3.2. (Finite singularity time for positive scalar curvature)

If $(\mathcal{M}^m, g)$ is a closed Riemannian manifold with positive scalar curvature, then for any solution $g(t)$, $t \in [0, T)$, to the Ricci flow with $g(0) = g$ we have

$$T \leq \frac{m}{2R_{\min}(0)} < \infty. \tag{3.3.6}$$

♡

3.3.2 Ricci flow on non-compact manifolds

Besides studying the Ricci flow on closed manifolds, we shall consider the Ricci flow on non-compact manifolds. This will be especially important in singularity analysis.

Definition 3.1

A solution $g(t)$, $t \in \mathfrak{t}$, of the Ricci flow is said to be **complete** *if for each $t \in \mathfrak{t}$, the Riemannian metric $g(t)$ is complete. We say a solution of the Ricci flow is* **ancient** *if it exists on the time interval $(-\infty, 0]$.*

♣

Lemma 3.3. (Ancient solutions have nonnegative scalar curvature)

If $(\mathcal{M}^m, g(t))$, $t \in (-\infty, 0]$, is a complete ancient solution to the Ricci flow with bounded curvature on compact time intervals, then either $R_{g(t)} > 0$ for all $t \in (-\infty, 0]$ or $\mathrm{Ric}_{g(t)} \equiv 0$ for all $t \in (-\infty, 0]$.

♡

Proof. If M is closed, then we can apply the maximum principle. If M is non-compact, then since the solution has bounded curvature on compact time intervals, we may still apply the maximum principle to the evolution equation for $R_{g(t)}$. For any solution of the Ricci flow for which we can apply the maximum principle on a time interval $[0, T]$, by (3.3.4), we have $R_{g(t)} \geq -\frac{m}{2t}$ for $t \in (0, T]$. Let α be any negative number. Since the solution is ancient, it exists on the time interval $[\alpha, 0]$. Then we have $R_{g(t)} \geq -\frac{m}{2(t-\alpha)}$ for all $x \in \mathcal{M}^m$ and $t \in (\alpha, 0]$. Taking the limit as $\alpha \to -\infty$, we conclude that $R_{g(t)} \geq 0$ for all $t \in (-\infty, 0]$. By Strong maximum principle, it implies that either $R_{g(t)} > 0$ always or $R_{g(t)} \equiv 0$ always. In the latter case, by the evolution equation for $R_{g(t)}$, we deduce $\mathrm{Rc}_{g(t)} \equiv 0$. □

3.3.3 Heat equation and the Ricci flow

Let u be a solution to the heat equation

$$\partial_t u = \Delta_g u$$

on a Riemannian manifold $(\mathcal{M}^m, g)$.

Lemma 3.4

One has

$$\partial_t |\nabla_g u|_g^2 = \Delta_g |\nabla_g u|_g^2 - 2\left|\nabla_g^2 u\right|_g^2 - 2\mathrm{Ric}_g(\nabla_g u, \nabla_g u). \tag{3.3.7}$$

If $\mathrm{Ric}_g \geq 0$*, then*

$$|\nabla_g u|_g \leq \frac{u_{\max}(0)}{\sqrt{2t}} \tag{3.3.8}$$

where $u_{\max}(0) := \max_{x\in\mathcal{M}^m} u(x,0)$*. If* $\mathrm{Ric}_g \geq -(m-1)K$ *for some positive constant* K*, then*

$$|\nabla_g u|_g \leq \frac{u_{\max}(0)}{\sqrt{2t}} e^{(m-1)Kt}. \tag{3.3.9}$$

Hence to get decay of $|\nabla_g u|_g$ *as* $t \to \infty$*, we should assume* $K = 0$*, i.e.,* Ric_g *is nonnegative.* ♡

Proof. Calculate

$$\begin{aligned}
\partial_t |\nabla_g u|_g^2 &= \partial_t\left(g^{ij}\nabla_i u \nabla_j u\right) = 2g^{ij}\nabla_j u\cdot\partial_t\nabla_i u = 2g^{ij}\nabla_j u\cdot\nabla_i\Delta_g u,\\
\Delta_g |\nabla_g u|_g^2 &= g^{k\ell}\nabla_k\nabla_\ell\left(g^{ij}\nabla_i u\nabla_j u\right) = g^{k\ell}g^{ij}\nabla_k\nabla_\ell(\nabla_i u\nabla_j u)\\
&= 2g^{k\ell}g^{ij}\nabla_k(\nabla_\ell\nabla_i u\cdot\nabla_j u) = 2g^{k\ell}g^{ij}(\nabla_k\nabla_\ell\nabla_i u\cdot\nabla_j + \nabla_\ell\nabla_i u\cdot\nabla_k\nabla_j u)\\
&= 2g^{k\ell}g^{ij}\nabla_k\nabla_i\nabla_\ell u\cdot\nabla_j u + 2\left|\nabla_g^2 u\right|_g^2 = 2g^{k\ell}g^{ij}\left(\nabla_i\nabla_k\nabla_\ell u - R^p_{ki\ell}\nabla_p u\right)\nabla_j u + 2\left|\nabla_g^2 u\right|_g^2\\
&= 2g^{ij}\nabla_i\Delta_g u\cdot\nabla_j u + 2|\nabla_g u|_g^2 + 2R_{ij}\nabla^i u\nabla^j u.
\end{aligned}$$

Hence

$$\partial_t |\nabla_g u|_g^2 = \Delta_g |\nabla_g u|_g^2 - 2\left|\nabla_g^2 u\right|_g^2 - 2R_{ij}\nabla^i u\nabla^j u.$$

Consequently,

$$\partial_t\left(t|\nabla_g u|_g^2 + \frac{1}{2}u^2\right) = \Delta_g\left(t|\nabla_g u|_g^2 + \frac{1}{2}u^2\right) - 2t\left|\nabla_g^2 u\right|_g^2 - 2tR_{ij}\nabla^i u\nabla^j u.$$

If $\mathrm{Rc}_g \geq -(m-1)K$, then

$$\begin{aligned}
(\partial_t - \Delta_g)\left(t|\nabla_g u|_g^2 + \frac{1}{2}u^2\right) &\leq 2t(m-1)K|\nabla_g u|_g^2\\
&\leq 2(m-1)K\left(t|\nabla_g u|_g^2 + \frac{1}{2}u^2\right).
\end{aligned}$$

Applying the maximum principle, we deduce the required result. □

Note 3.2

Let u be a solution to the heat equation with respect to a metric $g(t)$ evolving by the Ricci flow

$$\partial_t u = \Delta_{g(t)} u.$$

Then

$$\Box_{g(t)} \left|\nabla_{g(t)} u\right|^2_{g(t)} = -2\left|\nabla^2_{g(t)} u\right|^2_{g(t)}.$$

Similarly, we have

$$\left(\partial_t - \Delta_{g(t)}\right)\left(t\left|\nabla_{g(t)} u\right|^2_{g(t)} + \frac{1}{2}u^2\right) \le 0.$$

If $\mathcal{M}^m$ is closed, then

$$\left|\nabla_{g(t)} u\right|_{g(t)} \le \frac{u_{\max}(0)}{\sqrt{2t}}.$$

♣

3.4 The Einstein-Hilbert functional

Introduction

- *The Einstein-Hilbert functional*
- *The Perelman functional*

3.4.1 The Einstein-Hilbert functional

If $\partial_s g_{ij} = v_{ij}$, then

$$\partial_s dV_{g(s)} = \frac{1}{2} V dV_{g(s)} \tag{3.4.1}$$

where $V := g^{ij} v_{ij}$. On the other hand, we have

$$\partial_s g^{ij} = -g^{ik} g^{j\ell} \partial_s g_{k\ell}. \tag{3.4.2}$$

Consider the **Einstein-Hilbert functional**

$$\mathcal{E}(g) := \int_{\mathcal{M}^m} R_g dV_g. \tag{3.4.3}$$

If $\partial_s g_{ij} = v_{ij}$, then

$$\begin{aligned} \frac{d}{ds}\mathcal{E}(g(s)) &= \int_M \left(-\Delta_{g(s)} V + \nabla^p \nabla^q v_{pq} - \langle v, \mathrm{Rc}_{g(s)}\rangle_{g(s)} + \frac{1}{2} R_{g(s)} V\right) dV_{g(s)} \\ &= \int_M \left\langle v, \frac{1}{2} R_{g(s)} g(s) - \mathrm{Rc}_{g(s)}\right\rangle_{g(s)} dV_{g(s)}. \end{aligned} \tag{3.4.4}$$

The twice of the gradient flow of $\mathcal{E}$ is

$$\partial_t g_{ij} = -2R_{ij} + R_{g(t)} g_{ij}. \tag{3.4.5}$$

The equation (3.4.5) is not parabolic, and the short time existence is not expected as that for the Ricci flow. Dropping the $R_{g(t)} g_{ij}$ term yields the Ricci flow. The undesirable term $R_{g(t)} g_{ij}$ in (3.4.5) is due to the variation of the volume form dV_g in $\mathcal{E}(g)$. How can we get rid of this term? First, we should consider metrics up to geometric equivalence, that is, up to pull-back by diffeomorphisms. We consider the mere general class of flows

$$\partial_t g_{ij} = -2R_{ij} - 2\nabla_i \nabla_j f, \tag{3.4.6}$$

where f is a time-dependent function. The flow (3.4.6) is equivalent to the Ricci flow since $2\nabla_i \nabla_j f = \left(\mathscr{L}_{\mathrm{grad}_g f} g\right)_{ij}$.

Note 3.3

Define a 1*-parameter family of diffeomorphisms* $\Psi(t) : \mathcal{M}^m \to \mathcal{M}^m$ *by*

$$\partial_t \Psi(t) = \nabla_{g(t)} f(t), \quad \Psi(0) = \mathrm{id}.$$

Show that $\bar{g}(t) := [\Psi(t)]^* g(t)$ *and* $\bar{f}(t) := f \circ \Psi(t)$ *satisfy*

$$\begin{aligned} \partial_t \bar{g}(t) &= -2\mathrm{Ric}_{\bar{g}(t)}, & (3.4.7)\\ \partial_t \bar{f}(t) &= \left|\nabla_{\bar{g}(t)} \bar{f}(t)\right|^2_{\bar{g}(t)}. & (3.4.8) \end{aligned}$$

Since

$$\partial_t \bar{g}(t) = \partial_t \left(\Psi^*(t) g(t)\right) = \Psi^*(t)\left[\partial_t g(t) + \mathscr{L}_{\mathrm{grad}_{g(t)} f(t)} g(t)\right],$$

it follows that

$$\partial_t \bar{g}(t) = \Psi^*(t)\left(-2\mathrm{Ric}_{g(t)}\right) = -2\mathrm{Ric}_{\bar{g}(t)}.$$

(3.4.8) is obvious. ♣

3.4.2 The Perelman functional

Now we impose the condition

$$\partial_s \left(e^{-f(s)} dV_{g(s)}\right) = 0 \tag{3.4.9}$$

and consider the functional

$$\mathcal{E}(g, f) := \int_M R_g e^{-f} dV_g. \tag{3.4.10}$$

If $\partial_s g_{ij} = v_{ij}$, then

$$\begin{aligned} \frac{d}{ds}\mathcal{E}(g(s), f(s)) &= \int_M \partial_s R_{g(s)} e^{-f(s)} dV_{g(s)} = -\int_{\mathcal{M}^m} \left\langle v, \mathrm{Ric}_{g(s)}\right\rangle_{g(s)} e^{-f(s)} dV_{g(s)} \\ &+ \int_{\mathcal{M}^m} \left(-\Delta_{g(s)} V + \nabla_i \nabla_j v_{ij}\right) e^{-f(s)} dV_{g(s)} = -\int_{\mathcal{M}^m} \left\langle v, \mathrm{Ric}_{g(s)}\right\rangle_{g(s)} e^{-f(s)} dV_{g(s)} \\ &+ \int_{\mathcal{M}^m} v_{ij}\left[-\Delta_{g(s)}\left(e^{-f(s)}\right) g_{ij} + \nabla_i \nabla_j \left(e^{-f(s)}\right)\right] dV_{g(s)}. \end{aligned}$$

To cancel the undesirable term, the last integral, we introduce the Dirichlet energy-type term

$$\int_{\mathcal{M}^m} \left|\nabla_{g(s)} f(s)\right|^2_{g(s)} e^{-f(s)} dV_{g(s)} = 4\int_{\mathcal{M}^m} \left|\nabla_{g(s)}\left(e^{-f(s)/2}\right)\right|^2_{g(s)} dV_{g(s)}.$$

According to the assumption (3.4.9), we have

$$\frac{d}{ds}\int_{\mathcal{M}^m} \left|\nabla_{g(s)} f(s)\right|^2_{g(s)} e^{-f(s)} dV_{g(s)} = \int_{\mathcal{M}^m} \left(\partial_s \left|\nabla_{g(s)} f(s)\right|^2_{g(s)}\right) e^{-f(s)} dV_{g(s)}$$

$$= \int_{\mathcal{M}^m} \left[-v_{ij}\nabla^i f(s)\nabla^j f(s) + 2\nabla^i f(s)\nabla_i\left(\partial_s f(s)\right)\right] e^{-f(s)} dV_{g(s)}.$$

Using $\partial_s f(s) = \frac{1}{2}V$ (by (3.4.9)), we find that the above is equal to

$$\int_{\mathcal{M}^m} \left[-v_{ij}\nabla^i f(s)\nabla^j f(s) e^{-f(s)} + V\Delta_{g(s)}\left(e^{-f(s)}\right)\right] dV_{g(s)}.$$

Adding the above equations together, we obtain

$$\begin{aligned} & \frac{d}{ds}\int_{\mathcal{M}^m} \left(R_{g(s)} + |\nabla_{g(s)} f(s)|^2_{g(s)}\right) e^{-f(s)} dV_{g(s)} \\ = & -\int_M \left\langle v, \mathrm{Ric}_{g(s)} + \nabla^2_{g(s)} f(s)\right\rangle_{g(s)} e^{-f(s)} dV_{g(s)}. \end{aligned} \tag{3.4.11}$$

So if we define the **Perelman functional**

$$\mathcal{F}(g,f) := \int_{\mathcal{M}^m} \left(R_g + |\nabla_g f|^2_g\right) e^{-f} dV_g, \tag{3.4.12}$$

then the gradient flow for $\mathcal{F}$, under the constraint that $e^{-f}dV_g$ is fixed, is

$$\partial_t g_{ij} = -2R_{ij} - 2\nabla_i\nabla_j f, \tag{3.4.13}$$

$$\partial_t f = -R_{g(t)} - \Delta_{g(t)} f. \tag{3.4.14}$$

Under the flow (3.4.13)–(3.4.14), we have the monotonicity formula

$$\frac{d}{dt}\mathcal{F}(g(t), f(t)) = 2\int_M \left|\mathrm{Ric}_{g(t)} + \nabla^2_{g(t)} f(t)\right|^2_{g(t)} e^{-f(t)} dV_{g(t)} \geq 0. \tag{3.4.15}$$

Since

$$\int_M |\nabla_g f|^2_g e^{-f} dV_g = \int_M \Delta_g f \cdot e^{-f} dV_g$$

we can rewrite $\mathcal{F}$ as

$$\mathcal{F}(g,f) = \int_M \left(R_g + 2\Delta_g f - |\nabla_g f|^2_g\right) e^{-f} dV_g. \tag{3.4.16}$$

3.5 Evolution of geometric quantities

Introduction

- *Variation of the Christoffel symbols*
- *Commutator formulas*
- *Variation of Ricci formula*

3.5.1 Variation of the Christoffel symbols

Lemma 3.5. (Variation of Christoffel symbols)

If $g(s)$ is a 1-parameter family of metrics with $\partial_s g_{ij} = v_{ij}$, then

$$\partial_s \Gamma^k_{ij} = \frac{1}{2} g^{k\ell}\left(\nabla_i v_{j\ell} + \nabla_j v_{i\ell} - \nabla_\ell v_{ij}\right). \tag{3.5.1}$$

When we consider the Ricci flow, we have

Corollary 3.3. (Evolution of Christoffel symbols under the Ricci flow)

Under the Ricci flow $\partial_t g_{ij} = -2R_{ij}$, we have

$$\partial_t \Gamma^k_{ij} = -g^{k\ell}\left(\nabla_i R_{j\ell} + \nabla_j R_{i\ell} - \nabla_\ell R_{ij}\right). \tag{3.5.2}$$

♡

Lemma 3.6. (Evolution of Laplacian under the Ricci flow)

If $(\mathcal{M}^m, g(t))$ is a solution to the Ricci flow $\partial_t g_{ij} = -2R_{ij}$, then

$$\partial_t \Delta_{g(t)} = 2R_{ij}\nabla^i\nabla^j,$$

where $\Delta_{g(t)}$ is the Laplacian acting on functions. In particular, when $m = 2$, $\partial_t \Delta_{g(t)} = R_{g(t)}\Delta_{g(t)}$.

♡

Proof. Calculate $\partial_t \Delta_{g(t)} = \partial_t\left(g^{ij}\nabla_i\nabla_j\right) = -\partial_t g_{ij}\cdot\nabla^i\nabla^j - g^{ij}\partial_t\Gamma^k_{ij}\cdot\nabla_k$. But, $g^{ij}\partial_t\Gamma^k_{ij} = -g^{k\ell}\left(2g^{ij}\nabla_i R_{j\ell} - \nabla_\ell R_{g(t)}\right) = 0$, where we use the contracted second Bianchi identity. □

Note 3.4

If $\partial_s g_{ij} = v_{ij}$, then

$$\partial_s \Delta_{g(s)} = -v_{ij}\nabla^i\nabla^j - g^{k\ell}\left(\left(\operatorname{div}_{g(s)}v\right)_\ell - \frac{1}{2}\nabla_\ell V\right)\nabla_k. \tag{3.5.3}$$

♣

3.5.2 Variation of Ricci formula

Recall that the components of the Riemann curvature $(3,1)$-tensor field are defined by

$$R^\ell_{ijk} = \partial_i\Gamma^\ell_{jk} - \partial_j\Gamma^\ell_{ik} + \Gamma^p_{jk}\Gamma^\ell_{ip} - \Gamma^p_{ik}\Gamma^\ell_{jp} \tag{3.5.4}$$

and the Ricci tensor field is $R_{ij} = R^p_{pij} = R_{pijp}$. Then the **variation of the Ricci tensor field in terms of the variation of the connection**

$$\partial_s R_{ij} = \nabla_p\left(\partial_s\Gamma^p_{ij}\right) - \nabla_i\left(\partial_s\Gamma^p_{pj}\right). \tag{3.5.5}$$

If $\partial_s g_{ij} = v_{ij}$, then

$$\partial_s R_{ij} = \frac{1}{2}\nabla^\ell\left(\nabla_i v_{j\ell} + \nabla_j v_{i\ell} - \nabla_\ell v_{ij}\right) - \frac{1}{2}\nabla_i\nabla_j V. \tag{3.5.6}$$

Taking the trace, we obtain the **variation of the scalar curvature**

$$\partial_s R_{g(s)} = \nabla^i\nabla^j v_{ij} - \Delta_{g(s)}V - \left\langle v, \operatorname{Ric}_{g(s)}\right\rangle_{g(s)}. \tag{3.5.7}$$

If we introduce the **Lichnerowicz Laplacian**

$$\Delta_{L,g}v_{ij} := \Delta_g v_{ij} + 2R_{kij\ell}v^{k\ell} - R_{ik}{v_j}^k - R_{jk}{v_i}^k \tag{3.5.8}$$

acting on symmetric 2-tensor fields, then (3.5.6) becomes

$$\partial_s R_{ij} = -\frac{1}{2}\left(\Delta_{L,g(s)}v_{ij} + \nabla_i\nabla_j V - \nabla_i\left(\operatorname{div}_{g(s)}v\right)_j - \nabla_j\left(\operatorname{div}_{g(s)}v\right)_i\right). \tag{3.5.9}$$

If we set $X = \frac{1}{2}\nabla_{g(s)}V - \operatorname{div}_{g(s)}v$, then

$$\partial_s\left(-2R_{ij}\right) = \Delta_{L,g(s)}v_{ij} + \nabla_i X_j + \nabla_j X_i. \tag{3.5.10}$$

This is related to De Turck's trick in proving short time existence.

3.5.3 Commutator formulas

Lemma 3.7. (Commutator formula for the Hessian and the Lichnerowicz operator)
Under the Ricci flow, the Hessian and the Lichnerowicz Laplacian heat operator $\Box_{L,g(t)} := \partial_t - \Delta_{L,g(t)}$ *commute. That is, for any function* f *of space and time we have*

$$\nabla_i\nabla_j\Box_{L,g(t)}f = \Box_{L,g(t)}\nabla_i\nabla_j f. \tag{3.5.11}$$

♡

Proof. Calculate

$$\begin{aligned}
\nabla_i\nabla_j\Delta_g f &= g^{k\ell}\nabla_i\nabla_j\nabla_k\nabla_\ell f = g^{k\ell}\nabla_i\left(\nabla_k\nabla_j\nabla_\ell f - R^p_{jk\ell}\nabla_p f\right)\\
&= g^{k\ell}\nabla_i\nabla_k\nabla_\ell\nabla_j f - g^{k\ell}\nabla_i R^p_{jk\ell}\cdot\nabla_p f - g^{k\ell}R^p_{jk\ell}\nabla_i\nabla_p f
\end{aligned}$$

$$= g^{k\ell}\left(\nabla_k\nabla_i\nabla_\ell\nabla_j f - R^p_{ikl}\nabla_p\nabla_j f - R^p_{ikj}\nabla_\ell\nabla_p f\right) - \nabla_i R_{jp}\nabla^p f - R_{jp}\nabla_i\nabla^p f \tag{3.5.12}$$

$$= g^{k\ell}\nabla_k\left(\nabla_l\nabla_i\nabla_j f - R^p_{i\ell j}\nabla_p f\right) - R_{ip}\nabla^p\nabla_j f - R_{ikjp}\nabla^k\nabla^p f - \nabla_i R_{jp}\nabla^p f - R_{jp}\nabla_i\nabla^p f$$

$$= \Delta_g\nabla_i\nabla_j f - g^{k\ell}\nabla_k R_{i\ell jp}\cdot\nabla^p f - R_{ip}\nabla^p\nabla_j f - 2R_{ikjp}\nabla^k\nabla^p f - \nabla_i R_{jp}\cdot\nabla^p f - R_{jp}\nabla_i\nabla^p f$$

$$\begin{aligned}
&= \left(\Delta_g\nabla_i\nabla_j f + 2R_{kijp}\nabla^k\nabla^p f - R_{ip}\nabla_j\nabla^p f - R_{jp}\nabla_i\nabla^p f\right)\\
&- \nabla^\ell R_{i\ell jp}\cdot\nabla^p f - \nabla_i R_{jp}\cdot\nabla^p f = \Delta_{L,g}\nabla_i\nabla_j f - \left(\nabla_i R_{j\ell} + \nabla_j R_{i\ell} - \nabla_\ell R_{ij}\right)\nabla^\ell f
\end{aligned}$$

where we use $\nabla^\ell R_{ikj\ell} = \nabla_j R_{i\ell} - \nabla_\ell R_{ij}$. Using (3.5.2), we compute

$$\partial_t\nabla_i\nabla_j f = \nabla_i\nabla_j\partial_t f + \left(\nabla_i R_{j\ell} + \nabla_j R_{i\ell} - \nabla_\ell R_{ij}\right)\nabla^\ell f. \tag{3.5.13}$$

Formula (3.5.11) follows from combining the above two calculations. □

Corollary 3.4
If $g(t)$ *satisfies the Ricci flow and* $f(t)$ *satisfies the heat equation* $\partial_t f(t) = \Delta_{g(t)}f(t)$, *then the Hessian satisfies the Lichnerowicz Laplacian heat equation*

$$\partial_t\nabla^2_{g(t)}f(t) = \Delta_{L,g(t)}\nabla^2_{g(t)}f(t). \tag{3.5.14}$$

♡

Lemma 3.8. (Commutator of $\partial_t + \Delta_{L,g(t)}$ and $\nabla^2_{g(t)}$)
Under the Ricci flow, the Hessian and the Lichnerowicz Laplacian backward heat operator commute in the following sense

$$\begin{aligned}
\nabla_i\nabla_j\left(\partial_t f(t) + \Delta_{L,g(t)}f(t)\right) &= \left(\partial_t + \Delta_{L,g(t)}\right)\nabla_i\nabla_j f(t)\\
&\quad - 2\left(\nabla_i R_{j\ell} + \nabla_j R_{i\ell} - \nabla_\ell R_{ij}\right)\nabla^\ell f(t).
\end{aligned} \tag{3.5.15}$$

♡

Proof. It follows from (3.5.12) and (3.5.13). □

Note 3.5

Show that under the Ricci flow, for any 1-form α

$$\Box_{L,g(t)}\mathscr{L}_{\alpha^\sharp}g(t)=\mathscr{L}_{\left[\Box_{H,g(t)}\alpha\right]^\sharp}g(t),\quad \Box_{H,g(t)}:=\partial_t-\Delta_{H,g(t)},\tag{3.5.16}$$

that is,

$$\Box_{L,g(t)}\left(\nabla_i\alpha_j+\nabla_j\alpha_i\right)=\nabla_i\beta_j+\nabla_j\beta_i,\tag{3.5.17}$$

where $\beta:=\Box_{H,g(t)}\alpha$. From

$$\partial_t\nabla_i\alpha_j=\nabla_i\partial_t\alpha_j+\left(\nabla_iR_{jk}+\nabla_jR_{ik}-\nabla_kR_{ij}\right)\alpha^k$$

and

$$\begin{aligned}\nabla_i\left(\Delta_{g(t)}\alpha_j-R_{jk}\alpha^k\right) &= \Delta_{g(t)}\nabla_i\alpha_j+2R_{kij\ell}\nabla^k\alpha^\ell-R_{ik}\nabla^k\alpha_j-R_{jk}\nabla_i\alpha^k\\ &\quad-\left(\nabla_iR_{jk}+\nabla_jR_{ik}-\nabla_kR_{ij}\right)\alpha^k\end{aligned}$$

we conclude

$$\begin{aligned}&\nabla_i\left(\partial_t\alpha_j-\left(\Delta_{g(t)}\alpha_j-R_{jk}\alpha^k\right)\right)+\nabla_j\left(\partial_t\alpha_i-\left(\Delta_{g(t)}\alpha_i-R_{ik}\alpha^k\right)\right)\\ =\ &\Box_{L,g(t)}\left(\nabla_i\alpha_j+\nabla_j\alpha_i\right).\end{aligned}$$

Note that $\Delta_{g(t)}-R_{jk}\alpha^k=\Delta_{H,g(t)}\alpha_j=\beta_j$. ♣

Lemma 3.9

If $(\mathcal{M}^m,g(t))$ is a solution to the Ricci flow and if X is a vector field evolving by

$$\Box_{g(t)}X^i=R^{ik}X_k$$

then $h_{ij}:=\nabla_iX_j+\nabla_jX_i=\left(\mathscr{L}_Xg\right)_{ij}$ evolves by

$$\Box_{L,g(t)}h_{ij}=0.\tag{3.5.18}$$

In particular,

$$\Box_{g(t)}H=2\left\langle\mathrm{Rc}_{g(t)},h\right\rangle_g\tag{3.5.19}$$

where $H:=g^{ij}h_{ij}=\mathrm{tr}_{g(t)}h$. ♡

Proof. If we let $\alpha:=X^\flat$ the dual 1-form, then the lemma immediately follows from Note 3.5. □

Note 3.6

If $X=X^i\partial_i$ is a Killing vector field, then

$$\nabla_k\nabla_jX_i+R_{\ell kji}X^\ell=0.\tag{3.5.20}$$

If X is a Killing vector field, then $\nabla_iX_j+\nabla_jX_i=0$. Calculate

$$\begin{aligned}0 &= \nabla_k\left(\nabla_jX_i+\nabla_iX_j\right)+\nabla_j\left(\nabla_iX_k+\nabla_kX_i\right)+\nabla_i\left(\nabla_jX_k+\nabla_kX_j\right)\\ &= \left(\nabla_k\nabla_jX_i+\nabla_i\nabla_kX_j\right)+\left(\nabla_j\nabla_iX_k+\nabla_k\nabla_iX_j\right)+\left(\nabla_j\nabla_kX_i+\nabla_i\nabla_jX_k\right)\end{aligned}$$

$$= (\nabla_i\nabla_k X_j - \nabla_k\nabla_i X_j) - (\nabla_j\nabla_k X_i + \nabla_k\nabla_j X_i) + (\nabla_i\nabla_j X_k - \nabla_j\nabla_i X_k)$$

$$= R_{kij\ell}X^\ell - 2\nabla_k\nabla_j X_i + R_{jki\ell}X^\ell + R_{ijk\ell}X^\ell = -2\nabla_k\nabla_j X_i + 2R_{jik\ell}X^\ell$$

where we use the first Bianchi identity. ♣

Lemma 3.10

If $(\mathcal{M}^m, g)$ is oriented and closed, and the Ricci curvature is negative, then there are no nontrivial Killing vector fields. ♡

Proof. Tracing (3.5.10), we have

$$\Delta_g X_i + R_{\ell i}X^\ell = 0. \tag{3.5.21}$$

Then

$$\int_{\mathcal{M}^m} |\nabla_g X|_g^2 \, dV_g = \int_{\mathcal{M}^m} \mathrm{Ric}_g(X,X) dV_g.$$

If $\mathrm{Ric}_g \le 0$, then $\nabla_g X = 0$. Since $\mathcal{M}^m$ is closed, we must have $X = 0$. □

Since, by the contracted second Bianchi identity

$$\nabla_i\nabla_j R_g - \nabla_i \left(\mathrm{div}_g \mathrm{Ric}_g\right)_j - \nabla_j \left(\mathrm{div}_g \mathrm{Ric}_g\right)_j = 0,$$

equation (3.5.9) implies

Lemma 3.11. (Evolution of the Ricci tensor under the Ricci flow)

Under the Ricci flow,

$$\Box_{g(t)} R_{ij} = 2R_{kij\ell}R^{k\ell} - 2R_{ik}R_j{}^k. \tag{3.5.22}$$ ♡

Note 3.7

For any $\alpha \in \mathbf{R}$, show that

$$\Box_{g(t)}\left(R_{ij} - \alpha R_{g(t)} g_{ij}\right) = 2R_{kij\ell}R^{k\ell} - 2R_{ik}R_j{}^k - 2\alpha \left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} g_{ij} + 2\alpha R_{g(t)} g_{ij}.$$ ♣

3.6 De Turck's trick and short time existence

We use De Turck's trick to prove the short time existence for Ricci flow.

3.6.1 Symbol and Bianchi operator

Let $(\mathcal{M}^m, g)$ be an m-dimensional Riemannian manifold, and $\odot^2 T^*\mathcal{M}^m$ denote the vector bundle of symmmetric covariant 2-tensor fields. From (3.5.9), the linearization of $-2\mathrm{Ric}_g$ is given by

$$\left(\mathfrak{D}\left(-2\mathrm{Ric}_g\right)[v]\right)_{ij} = \Delta_{L,g} v_{ij} + \nabla_i\nabla_j V - \nabla_i \left(\mathrm{div}_g v\right)_i - \nabla_j \left(\mathrm{div}_g v\right)_i, \tag{3.6.1}$$

where $V := g^{ij}v_{ij}$. The **symbol** of the linearization of the Ricci tensor is obtained by replacing ∇_g by $\zeta \in C^\infty(\mathcal{M}^m, T^*\mathcal{M}^m)$ in the highest-order terms. Thus,

$$\sigma_\zeta := \sigma\left(\mathfrak{D}_\zeta\left(-2\mathrm{Rc}_g\right)\right) : C^\infty(\mathcal{M}^m, \odot^2 T^*\mathcal{M}^m) \longrightarrow C^\infty(\mathcal{M}^m, \odot^2 T^*\mathcal{M}^m) \tag{3.6.2}$$

where $\zeta \in C^\infty(\mathcal{M}^m, T^*\mathcal{M}^m)$ and

$$\sigma_\zeta(v)_{ij} = |\zeta|^2 v_{ij} + \zeta_i\zeta_j V - \zeta_i\zeta_k v^k{}_j - \zeta_j\zeta_k v^k{}_i. \tag{3.6.3}$$

Assuming that $\zeta_1 = 0$ and $\zeta_i \neq 0$ for $i \neq 1$, then for any symmetric 2-tensor field v

$$\begin{aligned} \sigma_\zeta(v)_{ij} &= v_{ij}, \quad i, j \neq 1, \\ \sigma_\zeta(v)_{ij} &= 0, \quad j \neq 1, \\ \sigma_\zeta(v)_{11} &= \sum_{2 \leq k \leq m} v_{kk}. \end{aligned}$$

When $m = 3$, σ_ζ is given by

$$\sigma_\zeta \begin{pmatrix} v_{11} \\ v_{12} \\ v_{13} \\ v_{22} \\ v_{33} \\ v_{23} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \\ v_{13} \\ v_{22} \\ v_{33} \\ v_{23} \end{pmatrix}. \tag{3.6.4}$$

In general, σ_ζ is given by a nonnegative $N \times N$ matrix, where $N = \frac{m(m+1)}{2}$. Its kernel is the m-dimensional subspace given by

$$\mathrm{Ker}\left(\sigma_\zeta\right) = \left\{ v \in C^\infty(\mathcal{M}^m, \odot^2 T^*\mathcal{M}^m) : v_{ij} = 0 \text{ for } i, j \neq 1 \text{ and } \sum_{2 \leq k \leq m} v_{kk} = 0 \right\}. \tag{3.6.5}$$

This kernel is due to the diffeomorphism invariance of the operator $g \mapsto -2\mathrm{Ric}_g$.

Define the linear **Bianchi operator**

$$\mathbf{B}_g : C^\infty(\mathcal{M}^m, \odot^2 T^*\mathcal{M}^m) \longrightarrow C^\infty(\mathcal{M}^m, T^*\mathcal{M}^m) \tag{3.6.6}$$

by

$$\mathbf{B}_g(h)_k := g^{ij}\left(\nabla_i h_{jk} - \frac{1}{2}\nabla_k h_{ij}\right) \tag{3.6.7}$$

so that $\mathbf{B}_g\left(-2\mathrm{Rc}_g\right) = 0$. We find that

$$K := \mathrm{Ker}\left(\sigma \mathbf{B}_g(\zeta)\right) = \mathrm{Im}\left(\sigma_\zeta\right) \subset C^\infty(\mathcal{M}^m, \odot^2 T^*\mathcal{M}^m)$$

is equal to

$$K = \left\{ v \in C^\infty(\mathcal{M}^m, \odot^2 T^*\mathcal{M}^m) : v_{1j} = 0 \text{ for } j \neq 1 \text{ and } v_{11} = \sum_{2 \leq k \leq m} v_{kk} \right\}. \tag{3.6.8}$$

Hence

$$\sigma_\zeta\Big|_K(v) = |\zeta|^2 v, \quad \sigma_\zeta\Big|_K = |\zeta|^2 \tag{3.6.9}$$

for any $\zeta \in C^\infty(\mathcal{M}^m, T^*\mathcal{M}^m)$.

3.6.2 De Turck's trick

The fundamental short time existence theorem for the Ricci flow on closed manifolds is the following

Theorem 3.2. (Hamilton, De Turck)

If $\mathcal{M}^m$ is a closed Riemannian manifold and if g is a smooth Riemannian metric, then there exists a unique smooth solution $\overline{g}(t)$ to the Ricci flow defined on some time interval $[0, \delta)$, $\delta > 0$, with $\overline{g}(0) = g$. ♡

From the previous subsection the principal symbol of the nonlinear partial differential operator $-2\mathrm{Ric}_g$ of the metric g is nonnegative define and has a nontrivial kernel. For this reason the Ricci flow equation is only weakly parabolic. We search for an equivalent flow which is strictly parabolic (i.e., where the principal symbol of the second-order operator on the RHS is positive definite).

Given a fixed background connection $\widetilde{\nabla}$, which for convenience we assume to be the Levi-Civita connection of a background metric $\tilde{g}$, we define the **Ricci-De Turck flow** by

$$\partial_t g_{ij} = -2R_{ij} + \nabla_i W_j + \nabla_j W_i, \quad g(0) = g_0, \tag{3.6.10}$$

where the time-dependent 1-form $W(t)$ is defined by

$$W_j := g_{jk} g^{pq} \left(\Gamma^k_{pq} - \widetilde{\Gamma}^k_{pq} \right). \tag{3.6.11}$$

Note 3.8

If $g(s)$ is a 1-parameter family of metrics with $g(0) = g$ and

$$\partial_s\Big|_{s=0} g_{ij} = v_{ij},$$

then

$$\partial_s\Big|_{s=0} W_j = -X_j + \mathcal{O}_0(v)_j \tag{3.6.12}$$

where $X = \frac{1}{2}{}^g\nabla V - \mathrm{div}_g v$ and $\mathcal{O}_0(v)_j$ is the zero-order term in v given by

$$\mathcal{O}_0(v)_j = v_{jk} W^k - g_{jk} v^{pq} \left(\Gamma^k_{pq} - \widetilde{\Gamma}^k_{pq} \right). \tag{3.6.13}$$

Consequently,

$$\partial_s\Big|_{s=0} \left(-2R_{ij} + \nabla_i W_j + \nabla_j W_i \right) = \Delta_{L,g} v_{ij} + \mathcal{O}_1(v)_{ij} \tag{3.6.14}$$

where $\mathcal{O}_1(v)_{ij}$ is the first-order term in v given by

$$\begin{aligned} \mathcal{O}_1(v)_{ij} &= \nabla_k v_{ij} \cdot W^k + v_{ik} \cdot \nabla_j W^k + v_{jk} \cdot \nabla_i W^k \\ &- \nabla_i \left(g_{jk} v^{pq} \left(\Gamma^k_{pq} - \widetilde{\Gamma}^k_{pq} \right) \right) - \nabla_j \left(g_{ik} v^{pq} \left(\Gamma^k_{pq} - \widetilde{\Gamma}^k_{pq} \right) \right). \end{aligned} \tag{3.6.15}$$

Calculate

$$\begin{aligned}\partial_s\Big|_{s=0} W_j &= v_{jk} g^{pq}\left(\Gamma^k_{pq} - \widetilde{\Gamma}^k_{pq}\right) - g_{jk} v^{pq}\left(\Gamma^k_{pq} - \widetilde{\Gamma}^k_{pq}\right) \\ &\quad + g_{jk} g^{pq} \frac{1}{2} g^{k\ell}\left(\nabla_p v_{q\ell} + \nabla_q v_{p\ell} - \nabla_\ell v_{pq}\right) \\ &= v_{jk} W^k - g_{jk} v^{pq}\left(\Gamma^k_{pq} - \widetilde{\Gamma}^k_{pq}\right) - \left(\frac{1}{2}\nabla_j V - (\mathrm{div}_g v)_j\right).\end{aligned}$$

Hence,

$$\begin{aligned}\partial_s\Big|_{s=0} \nabla_i W_j &= \partial_s\Big|_{s=0}\left(\partial_i W_j - \Gamma^k_{ij} W_k\right) \\ = \ & \nabla_i\left(\partial_s\Big|_{s=0} W_j\right) - \left(\partial_s\Big|_{s=0}\Gamma^k_{ij}\right) W_k \\ = \ & \nabla_i\left(-\frac{1}{2}\nabla_j V + (\mathrm{div}_g v)_j + \mathcal{O}_0(v)_j\right) - \frac{1}{2} g^{k\ell}\left(\nabla_i v_{j\ell} + \nabla_j v_{i\ell} - \nabla_\ell v_{ij}\right) W_k\end{aligned}$$

and

$$\begin{aligned}\partial_s\Big|_{s=0}\left(\nabla_i W_j + \nabla_j W_i\right) &= -\nabla_i\nabla_j V + \nabla_i\left(\mathrm{div}_g v\right)_j + \nabla_j\left(\mathrm{div}_g v\right)_i \\ + \nabla_i\left(\mathcal{O}_0(v)_j\right) + \nabla_j\left(\mathcal{O}_0(v)_i\right) &- g^{k\ell}\left(\nabla_i v_{j\ell} + \nabla_j v_{i\ell} - \nabla_\ell v_{ij}\right) W_k.\end{aligned}$$

Therefore the left hand side of (3.6.14) equals

$$\Delta_{L,g} v_{ij} + \nabla_j\left(\mathcal{O}_0(v)_i\right) + \nabla_i\left(\mathcal{O}_0(v)_j\right) - g^{k\ell}\left(\nabla_i v_{j\ell} + \nabla_j v_{i\ell} - \nabla_\ell v_{ij}\right) W_k.$$

Plugging (3.6.13) into above yields the right hand side of (3.6.14). ♣

Note 3.8 shows that the Ricci-De Turck flow is strictly parabolic and that given any smooth initial metric g on a closed manifold, there exists a unique solution $g(t)$ to the Ricci-De Turck flow. Another way of showing that the Ricci-De Turck flow is a strictly parabolic system is to compute an expression for the modified Ricci tensor of a metric g as an elliptic operator of g using the background metric $\tilde{g}$.

Note 3.9

(Another proof of the strict parabolicity of the Ricci-De Turck flow) *Define a tensor*

$$A^k_{ij} := \Gamma^k_{ij} - \widetilde{\Gamma}^k_{ij} = \frac{1}{2} g^{k\ell}\left(\widetilde{\nabla}_i g_{j\ell} + \widetilde{\nabla}_j g_{i\ell} - \widetilde{\nabla}_\ell g_{ij}\right). \tag{3.6.16}$$

At a point (x,t) in a local coordinate system where $\widetilde{\Gamma}^k_{ij}(p) = 0$, we have

$$R^\ell_{ijk} - \widetilde{R}^\ell_{ijk} = \widetilde{\nabla}_i A^\ell_{jk} - \widetilde{\nabla}_j A^\ell_{ik} + A^p_{jk} A^\ell_{ip} - A^p_{ik} A^\ell_{jp}. \tag{3.6.17}$$

From this we obtain

$$\begin{aligned}-2R_{jk} &= 2\widetilde{R}_{jk} - g^{\ell m}\widetilde{\nabla}_\ell\left(\widetilde{\nabla}_j g_{kp} + \widetilde{\nabla}_k g_{jp} - \widetilde{\nabla}_p g_{jk}\right) \\ &\quad + g^{\ell p}\widetilde{\nabla}_j\left(\widetilde{\nabla}_l g_{kp} + \widetilde{\nabla}_k g_{l\ell p} - \widetilde{\nabla}_m g_{\ell k}\right) + g^{-1} * g^{-1} * \widetilde{\nabla}_{\tilde{g}} g * \widetilde{\nabla}_{\tilde{g}} g.\end{aligned}$$

On the other hand, by definition of W_j, we get

$$\begin{aligned}\nabla_i W_j &= \frac{1}{2} g^{\ell p}\widetilde{\nabla}_i\left(\widetilde{\nabla}_\ell g_{pj} + \widetilde{\nabla}_p g_{\ell j} - \widetilde{\nabla}_j g_{\ell p}\right) \\ &\quad + g * g^{-1} * g^{-1} * g^{-1} * \widetilde{\nabla}_{\tilde{g}} g * \widetilde{\nabla}_{\tilde{g}} g.\end{aligned}$$

Hence

$$\begin{aligned}\partial_t g_{ij} &= -2R_{ij} + \nabla_i W_j + \nabla_j W_i = g^{\ell p}\widetilde{\nabla}_\ell \widetilde{\nabla}_p g_{ij} + g^{-1} * g * \widetilde{g}^{-1} * \widetilde{\mathrm{Rm}}_{\widetilde{g}} \\ &\quad + g * g^{-1} * g^{-1} * g^{-1} * \widetilde{\nabla}_{\widetilde{g}} g * \widetilde{\nabla}_{\widetilde{g}} g.\end{aligned}$$

Hence the Ricci-De Turck flow is strictly parabolic. ♣

Given a solution of the Ricci-De Turck flow, we can solve the following ODE at each point on $\mathcal{M}^m$:

$$\partial_t \varphi_t = -W^\flat(t) \circ \varphi(t), \quad \varphi_0 = \mathrm{id}. \tag{3.6.18}$$

The existence and uniqueness of (3.6.18) reduces to the harmonic map heat flow.

3.6.3 Harmonic map heat flow

Given a map $f : (\mathcal{M}^m, g) \to (\mathcal{N}^n, h)$, the **map Laplacian** of f is defined by

$$\begin{aligned}(\Delta_{g,h} f)^\gamma &= \Delta_g(f^\gamma) + g^{ij}\left({}^h\Gamma^\gamma_{\alpha\beta} \circ f\right)\frac{\partial f^\alpha}{\partial x^i}\frac{\partial f^\beta}{\partial x^j} \\ &= g^{ij}\left(\frac{\partial^2 f}{\partial x^i \partial x^j} - {}^g\Gamma^k_{ij}\frac{\partial f^\gamma}{\partial x^k} + \left({}^h\Gamma^\gamma_{\alpha\beta} \circ f\right)\frac{\partial f^\alpha}{\partial x^i}\frac{\partial f^\beta}{\partial x^j}\right),\end{aligned} \tag{3.6.19}$$

where $f^\gamma := y^\gamma \circ f$, and $(x^i)_{1 \le i \le m}$ and $(y^\alpha)_{1 \le \alpha \le n}$ are local coordinates on $\mathcal{M}^m$ and $\mathcal{N}^n$, respectively. Consider the map

$$\begin{array}{ccc} f^*(T\mathcal{N}^n) & & T\mathcal{N}^n \\ \downarrow & & \downarrow \\ \mathcal{M}^m & \xrightarrow{f} & \mathcal{N}^n \end{array}$$

We observe that $\Delta_{g,h} f \in C^\infty(f^*(T\mathcal{N}^n))$. In $\mathcal{M}^m = \mathcal{N}^n$ and f is the identity map, then

$$(\Delta_{g,h}\mathrm{id}_{\mathcal{M}^m})^k = g^{ij}\left(-{}^g\Gamma^k_{ij} + {}^h\Gamma^k_{ij}\right). \tag{3.6.20}$$

Note 3.10

For each $p \in \mathcal{M}^m$, we have $(df)_p : T_p\mathcal{M}^m \to T_{f(p)}\mathcal{N}^n$ and hence $(df)_p \in T^*_p\mathcal{M}^m \otimes T_{f(p)}\mathcal{N}^n$. Consequently, $df \in C^\infty(\mathcal{M}^m, T^*\mathcal{M}^m \otimes f^*T\mathcal{N}^n)$. The Levi-Civita connections ${}^g\nabla$ and ${}^h\nabla$ induce a canonical connection ${}^{g,h}\nabla$ on the vector bundle $T^*\mathcal{M}^m \otimes f^*T\mathcal{N}^n$:

$${}^{g,h}\nabla : C^\infty(\mathcal{M}^m, T^*\mathcal{M}^m \otimes f^*T\mathcal{N}^n) \longrightarrow C^\infty(\mathcal{M}^m, T^*\mathcal{M}^m \otimes T^*\mathcal{M}^m \otimes f^*T\mathcal{N}^n).$$

Since $df \in C^\infty(\mathcal{M}^m, T^*\mathcal{M}^m \otimes f^*T\mathcal{N}^n)$, it follows that

$$\left({}^{g,h}\nabla df\right)^\gamma_{ij} = \frac{\partial}{\partial x^i}\left(\frac{\partial f^\gamma}{\partial x^j}\right) - {}^g\Gamma^k_{ij}\frac{\partial f^\gamma}{\partial x^k} + \left({}^h\Gamma^\gamma_{\alpha\beta} \circ f\right)\frac{\partial f^\alpha}{\partial x^i}\frac{\partial f^\beta}{\partial x^j}.$$

Taking the trace with respect to g gives the map Laplacian of f

$$\Delta_{g,h} f = \mathrm{tr}_g\left({}^{g,h}\nabla f\right). \tag{3.6.21}$$

♣

A map $f : (\mathcal{M}^m, g) \to (\mathcal{N}^n, h)$ is called a **harmonic map** if $\Delta_{g,h} f = 0$. Harmonic maps are critical points of the **harmonic map energy**

$$\mathcal{E}_{g,h}(u) := \int_{\mathcal{M}^m} |du|^2_{g,h}\, dV_g \tag{3.6.22}$$

where

$$|du|^2_{g,h} := g^{ij} h_{\alpha\beta} \frac{\partial u^\alpha}{\partial x^i} \frac{\partial u^\beta}{\partial x^j}. \tag{3.6.23}$$

In the case that $\mathcal{N}^n = \mathbf{R}$, a harmonic map is the same as a harmonic function and the harmonic energy is the same as the Dirichlet energy. If M is 1-dimensional, then a harmonic map is the same as a constant speed geodesic.

Note 3.11

A harmonic map is the critical points of the harmonic map energy. ♣

Note 3.12

Suppose that $\mathcal{M}^m$ is an m-dimensional manifold, $(\mathcal{P}^m, k)$ is an m-dimensional Riemannian manifold, and $(\mathcal{N}^n, h)$ is an n-dimensional Riemannian manifold. If $F : (\mathcal{P}^m, k) \to (\mathcal{N}^n, h)$ is a map and $\varphi : \mathcal{M}^m \to \mathcal{P}^m$ is a diffeomorphism, then

$$\left(\Delta_{k,h} F\right)(\varphi(y)) = \left(\Delta_{\varphi^* k, h}(F \circ \varphi)\right)(y), \tag{3.6.24}$$

which corresponds to

$$(\mathcal{M}^m, \varphi^* k) \xrightarrow{\ \varphi\ } (\mathcal{P}^m, k) \xrightarrow{\ F\ } (\mathcal{N}^n, h).$$

In particular, given a diffeomorphism $f : (\mathcal{M}^m, g) \to (\mathcal{N}^m, h)$ between m-dimensional Riemannian manifolds, we consider

$$(\mathcal{M}^m, g) \xrightarrow{\ f\ } (\mathcal{N}^m, (f^{-1})^* g) \xrightarrow{\ \mathrm{id}_{\mathcal{N}^m}\ } (\mathcal{N}^m, h)$$

then

$$\left(\Delta_{g,h} f\right)(x) = \left(\Delta_{(f^{-1})^* g, h} \mathrm{id}_N\right)(f(x)) \in C^\infty(M, f^* TN). \tag{3.6.25}$$ ♣

If we set $\overline{g}(t) := \varphi_t^* g(t)$, then (3.6.18) is equivalent to

$$\partial_t \varphi_t = g^{pq}\left(-\Gamma^k_{pq} + \widetilde{\Gamma}^k_{pq}\right) \frac{\partial}{\partial x^k} \cdot \varphi_t = \left(\Delta_{g,\widetilde{g}} \mathrm{id}_{\mathcal{M}^m}\right)(\varphi_t) = \Delta_{\overline{g}(t), \widetilde{g}} \varphi_t.$$

This flow is called the **harmonic map heat flow**. If $\mathcal{M}^m$ is compact, then the short-time existence and uniqueness of this flow follow from the standard parabolic theory. Furthermore,

$$\begin{aligned}
\partial_t \overline{g}(t) &= \varphi_t^*\left(\partial_t g(t)\right) + \left.\frac{\partial}{\partial s}\right|_{s=0} \left(\varphi^*_{t+s} g(t)\right) \\
&= -2\mathrm{Ric}_{\varphi_t^* g(t)} + \varphi_t^*\left(\mathscr{L}_{W(t)} g(t)\right) - \mathscr{L}_{\left(\varphi_t^{-1}\right)_* W(t)}\left(\varphi_t^* g(t)\right) \\
&= -2\mathrm{Ric}_{\overline{g}(t)}.
\end{aligned}$$

3.6.4 Complete noncompact case

For any C^∞ complex metric g with bounded sectional curvature on a noncompact manifold $\mathcal{M}^m$, a short-time existence result for solutions to the Ricci flow was proved by W.-X. Shi in 1989.

Definition 3.2

We say that a solution $g(t)$, $t \in \mathfrak{t}$, of the Ricci flow has **bounded curvature** *(or* **bounded curvature on compact time intervals***) if on every compact subinterval $[a, b] \subset \mathfrak{t}$ the Riemann curvature tensor is bounded. In particular, we do not assume the curvature bound is uniform in time on noncompact time intervals.* ♣

Theorem 3.3. (W.-X. Shi, 1989)

t3.6.8 Given a complete metric g with bounded sectional curvature on a noncompact manifold $\mathcal{M}^m$, there exists a complete solution $g(t)$, $t \in [0, T)$, of the Ricci flow on $\mathcal{M}^m$ with $g(0) = g$ and bounded curvature such that either $\sup_{\mathcal{M}^m \times [0,T)} |\mathrm{Rm}_{g(t)}|_{g(t)} = \infty$ or $T = \infty$. ♡

Problem 3.2

Under what conditions does uniqueness hold for complete solutions to the Ricci flow on noncompact manifolds? ♠

Chen and Zhu proved the uniqueness of the Ricci flow on noncompact manifolds in the case of bounded curvature.

3.7 Reaction-diffusion equation for the curvature tensor field

In this section we discuss the evolution equation satisfied by the Riemann curvature tensor field.

3.7.1 Evolution equation for $R_{ijk\ell}$

Following Hamilton, we introduce the notation

$$B_{ijk\ell} := -g^{pr}g^{qs}R_{ipjq}R_{kr\ell s} = -g^{pr}g^{qs}R_{pijq}R_{rk\ell s}. \tag{3.7.1}$$

Note that

$$B_{ji\ell k} = B_{ijk\ell}, \quad B_{ijk\ell} = B_{k\ell ij}. \tag{3.7.2}$$

Note 3.13

If $\partial_s g_{ij} = v_{ij}$, then

$$\begin{aligned}
\partial_s R_{ijk}^{\ell} &= \frac{1}{2}g^{\ell p}\left[\nabla_i\nabla_j v_{kp} + \nabla_i\nabla_k v_{jp} - \nabla_i\nabla_p v_{jk}\right. \\
&\quad \left. - \nabla_j\nabla_i v_{kp} - \nabla_j\nabla_k v_{ip} + \nabla_j\nabla_p v_{ik}\right] \qquad (3.7.3)\\
&= \frac{1}{2}g^{\ell p}\left(\nabla_i\nabla_k v_{jp} - \nabla_i\nabla_p v_{jk} - \nabla_j\nabla_k v_{ip} + \nabla_j\nabla_p v_{ik}\right) \\
&\quad - \frac{1}{2}g^{\ell p}\left(R_{ijk}^{q} v_{qp} + R_{ijp}^{q} v_{kq}\right). \qquad (3.7.4)
\end{aligned}$$

♣

Lemma 3.12

The evolution equation of the Riemann curvature tensor field is given by

$$\begin{aligned} \Box_t R_{ijk\ell} &= 2\left(B_{ijk\ell} - B_{ij\ell k} + B_{ikj\ell} - B_{i\ell jk}\right) \\ &\quad - \left(R_i{}^p R_{pjk\ell} + R_j{}^p R_{ipk\ell} + R_k{}^p R_{ijp\ell} + R_\ell{}^p R_{ijkp}\right). \end{aligned} \tag{3.7.5}$$

In particular,

$$\Box_t R_{ijk\ell} = \mathrm{Rm}_g * \mathrm{Rm}_g + \mathrm{Ric}_g * \mathrm{Rm}_g. \tag{3.7.6}$$

♡

Proof. Use Note 3.13 and the second Bianchi identity. □

3.7.2 Riemann curvature operator

The Riemann curvature tensor field may be considered as an operator

$$\mathbf{Rm}_g : \mathcal{A}^2(\mathcal{M}^m) \longrightarrow \mathcal{A}^2(\mathcal{M}^m) \tag{3.7.7}$$

defined by

$$\left(\mathbf{Rm}_g(\alpha)\right)_{ij} := R_{ijk\ell}\alpha^{\ell k}. \tag{3.7.8}$$

Definition 3.3

We call $\mathbf{Rm}_g$ *the* **Riemann curvature operator** *or* **curvature operator**. *We say that* $(\mathcal{M}^m, g)$ *has* **positive (nonnegative) curvature operator** *if the eigenvalues of* $\mathbf{Rm}_g$ *are positive (nonnegative), and we denote this by* $\mathbf{Rm}_g > 0$ *(*$\mathbf{Rm}_g \geq 0$*).*

♣

We can define the square of $\mathbf{Rm}_g$ by

$$\mathbf{Rm}_g^2 := \mathbf{Rm}_g \circ \mathbf{Rm}_g : \mathcal{A}^2(\mathcal{M}^m) \longrightarrow \mathcal{A}^2(\mathcal{M}^m). \tag{3.7.9}$$

For $U, V \in \mathcal{A}^2(\mathcal{M}^m)$, we define $[U,V]_{ij} := g^{k\ell}(U_{ik}V_{\ell j} - V_{ik}U_{\ell j})$. Then $\mathcal{A}^2(\mathcal{M}^m) \cong \mathfrak{so}(m)$. Choose a basis $(\varphi^i)_{i=1}^{\frac{m(m-1)}{2}}$ of $\mathcal{A}^2(\mathcal{M}^m)$ and let C_k^{ij} denote the structure constants defined by

$$[\varphi^i, \varphi^j] := C_k^{ij}\varphi^k. \tag{3.7.10}$$

We define the **Lie algebra square**

$$\mathbf{Rm}_g^{\#} : \mathcal{A}^2(\mathcal{M}^m) \longrightarrow \mathcal{A}^2(\mathcal{M}^m) \tag{3.7.11}$$

by

$$\left(\mathbf{Rm}_g^{\#}(\alpha)\right)_{ij} := C_i^{k\ell} C_j^{pq} \left(\mathbf{Rm}_g(\alpha)\right)_{kp} \left(\mathbf{Rm}_g(\alpha)\right)_{\ell q}. \tag{3.7.12}$$

If we choose $(\varphi^i)_{i=1}^{\frac{m(m-1)}{2}}$ so that $\mathbf{Rm}_g$ is diagonal, then for any vector field $X = X^i\partial_i$, we have

$$\left(\mathbf{Rm}_g^{\#}(\alpha)\right)_{ij} X^i X^j = \left(C_i^{k\ell} X^i\right)^2 \left(\mathbf{Rm}_g(\alpha)\right)_{kk} \left(\mathbf{Rm}_g(\alpha)\right)_{\ell\ell}. \tag{3.7.13}$$

Lemma 3.13

If $\mathbf{Rm}_g \geq 0$*, then* $\mathbf{Rm}_g^{\#} \geq 0$*.*

♡

We have the following nice form for the evolution equation for $\mathbf{Rm}_{g(t)}$.

Lemma 3.14

The evolution equation of the curvature operator is

$$\Box_t \mathbf{Rm}_{g(t)} = \mathbf{Rm}^2_{g(t)} + \mathbf{Rm}^{\#}_{g(t)}. \tag{3.7.14}$$

♡

3.7.3 Uhlenbeck's trick

To prove (3.7.14), we use what is known as **Uhlenbeck's trick**. The idea is to choose a vector bundle $\mathcal{E} \to \mathcal{M}^m$ isomorphic to the tangent bundle $T\mathcal{M}^m \to \mathcal{M}^m$ and a bundle isomorphism $\iota : \mathcal{E} \to T\mathcal{M}^m$. Pulling back the initial metric, we get a bundle metric $h := \iota^* g$ on $\mathcal{E}$. By using the metric $g(t)$ to identify $T\mathcal{M}^m$ and $T^*\mathcal{M}^m$, we may consider the Ricci tensor field $\mathrm{Ric}_{g(t)}$ as a bundle map

$$\mathbf{Ric}_{g(t)} : T\mathcal{M}^m \longrightarrow T\mathcal{M}^m, \quad X \longmapsto \mathbf{Ric}_{g(t)}(X) = \left(\mathbf{Ric}_{g(t)}(X)\right)^i \partial_i \tag{3.7.15}$$

where

$$\left(\mathbf{Ric}_{g(t)}(X)\right)^i := g^{ij} X^k R_{jk}. \tag{3.7.16}$$

We define a 1-parameter family of bundle isomorphisms $\iota(t) : \mathcal{E} \to T\mathcal{M}^m$ by the ODE

$$\frac{d}{dt}\iota(t) = \mathbf{Ric}_{g(t)} \circ \iota(t), \quad \iota(0) = \iota. \tag{3.7.17}$$

Let $(e_a)_{1\le a\le m}$ be a local basis of sections of $\mathcal{E}$ and let $h_{ab} := h(e_a, e_b)$. Calculate

$$\begin{aligned} \partial_t \left[\iota(t)^* g(t)\right]_{ab} &= \partial_t \left[\iota(t)^i_a \iota(t)^j_b g_{ij}(t)\right] \\ = {} & \left[\partial_t \iota(t)\right]^i_a \iota(t)^j_b g_{ij}(t) + \iota(t)^i_a \left[\partial_t \iota(t)\right]^j_b g_{ij}(t) + \iota(t)^i_a \iota(t)^j_b \partial_t g_{ij}(t) \\ = {} & \left[\mathrm{Ric}_{g(t)}\right]^i_k \iota(t)^k_a \iota(t)^j_b g_{ij}(t) + \left[\mathrm{Ric}_{g(t)}\right]^j_k \iota(t)^i_a \iota(t)^k_b g_{ij}(t) - 2\iota(t)^i_a \iota(t)^j_b \left[\mathrm{Ric}_{g(t)}\right]_{ij}. \end{aligned}$$

Hence $\iota(t)^* g(t) = h$ is independent of t. Using the bundle isomorphisms $\iota(t)$, we can pull back tensor fields on $\mathcal{M}^m$. In particular, we consider $\iota(t)^*\mathrm{Rm}_{g(t)}$, which is a section of $\odot^2\mathcal{E}^\vee = \wedge^2\mathcal{E}^\vee \otimes_S \wedge^2\mathcal{E}^\vee$. Let $\boldsymbol{R}_{abcd}$ be the components of $\iota(t)^*\mathrm{Rm}_{g(t)}$, and

$$\boldsymbol{B}_{abcd} := \boldsymbol{h}^{ep}\boldsymbol{h}^{fq}\boldsymbol{R}_{eabf}\boldsymbol{R}_{pcdq}. \tag{3.7.18}$$

Lemma 3.15

One has

$$\left(\partial_t - \Delta_{\boldsymbol{h}(t)}\right)\boldsymbol{R}_{abcd} = 2\left(\boldsymbol{B}_{abcd} - \boldsymbol{B}_{abdc} + \boldsymbol{B}_{acbd} - \boldsymbol{B}_{adbc}\right). \tag{3.7.19}$$

♡

Proof. By the definition, we have

$$\begin{aligned} \partial_t \boldsymbol{R}_{abcd} &= \partial_t\left(\iota^i_a \iota^j_b \iota^k_c \iota^\ell_d R_{ijk\ell}\right) \\ &= \iota^i_a \iota^j_b \iota^k_c \iota^\ell_d \partial_t R_{ijk\ell} + \boldsymbol{R}_a{}^p \boldsymbol{R}_{pbcd} + \boldsymbol{R}_b{}^p \boldsymbol{R}_{apcd} + \boldsymbol{R}_c{}^p \boldsymbol{R}_{abpd} + \boldsymbol{R}_d{}^p \boldsymbol{R}_{abcp}. \end{aligned}$$

Therefore, (3.7.19) follows from (3.7.5). □

Later, we will show that (3.7.19) is equivalent to (3.7.14).

Definition 3.4

A Riemannian manifold has **2-positive curvature operator** *if*

$$\lambda_1\left(\mathbf{Rm}_g\right)+\lambda_2\left(\mathbf{Rm}_g\right)>0. \tag{3.7.20}$$

That is, the sum of the lowest two eigenvalues of $\mathbf{Rm}_g$ *is positive at every point.* ♣

Haiwen Chen showed that if $(\mathcal{M}^m, g)$ is a closed Riemannian manifold with 2-positive curvature operator, then under the Ricci flow $g(t)$, with the initial metric $g(0)=g$, has 2-positive curvature operator for all $t>0$.

3.7.4 The curvature operator in dimension 3

In dimension 3, $\mathfrak{so}(3)\cong\mathbf{R}^3$. The Lie algebra structure is $[U,V]=U\times V$, namely, the cross product. This implies $\mathbf{Rm}_g^{\#}$ is the adjoint of $\mathbf{Rm}_g$. If we diagonalize, i.e.,

$$\mathbf{Rm}_g=\begin{pmatrix}\lambda & 0 & 0\\ 0 & \mu & 0\\ 0 & 0 & \nu\end{pmatrix}, \tag{3.7.21}$$

then

$$\mathbf{Rm}_g^2+\mathbf{Rm}_g^{\#}=\begin{pmatrix}\lambda^2+\mu\nu & 0 & 0\\ 0 & \mu^2+\lambda\nu & 0\\ 0 & 0 & \nu^2+\lambda\mu\end{pmatrix} \tag{3.7.22}$$

Chose an orthonormal frame $\{e_1,e_2,e_3\}$ and its dual orthonormal coframe $\{\omega^1,\omega^2,\omega^3\}$ such that the 2-forms $\varphi^1:=\omega^2\wedge\omega^3, \varphi^2:=\omega^3\wedge\omega^1, \varphi^3:=\omega^1\wedge\omega^2$ are eigenvectors of $\mathbf{Rm}_g$. In this case

$$\begin{aligned}\lambda &= 2\mathrm{Rm}_g(e_2,e_3,e_3,e_2) &&= 2\mathrm{Sec}_g(e_2,e_3),\\ \mu &= 2\mathrm{Rm}_g(e_1,e_3,e_3,e_1) &&= 2\mathrm{Sec}_g(e_1,e_3),\\ \nu &= 2\mathrm{Rm}_g(e_1,e_2,e_2,e_1) &&= 2\mathrm{Sec}_g(e_1,e_2).\end{aligned}$$

Note 3.14

If $m\geq 3$ *and* g *has constant sectional curvature, then*

$$R_{ijk\ell}=\frac{R_g}{m(m-1)}(g_{i\ell}g_{jk}-g_{ik}g_{j\ell}).$$

Hence, for any $\alpha\in\mathcal{A}^2(\mathcal{M}^m)$*, we have*

$$\left(\mathbf{Rm}_g(\alpha)\right)_{ij}=\frac{R_g}{m(m-1)}\left(\alpha_{ij}-\alpha_{ji}\right)=\frac{2R_g}{m(m-1)}\alpha_{ij}.$$

Thus $\mathbf{Rm}_g=\frac{2R_g}{m(m-1)}\mathrm{id}_{\mathcal{A}^2(\mathcal{M}^m)}$. ♣

Chapter 4 Ricci flow

Introduction

- *Hamilton's theorem*
- *The maximum principle for tensor fields*
- *Curvature pinching estimates*
- *Gradient bounds for the scalar curvature*
- *Exponential convergence of the normalized Ricci flow*

4.1 Hamilton's theorem

Introduction

- *The normalized Ricci flow*
- *Hamilton's theorem*

4.1.1 The normalized Ricci flow

The evolution equation of the volume form for the Ricci flow

$$\partial_t g(t) = -2\mathrm{Ric}_{g(t)} \tag{4.1.1}$$

is

$$\partial_t dV_{g(t)} = -R_{g(t)} dV_{g(t)}. \tag{4.1.2}$$

Then the volume evolves by

$$\frac{d}{dt} V_{g(t)} = -\int_{\mathcal{M}^m} R_{g(t)} dV_{g(t)}. \tag{4.1.3}$$

Given a solution $g(t)$, $t \in [0, T)$, of the Ricci flow (4.1.1), we consider the metrics

$$\bar{g}(\bar{t}) := c(t) g(t), \tag{4.1.4}$$

where

$$c(t) := \exp\left(\frac{2}{m}\int_0^t \underline{R}_{g(\tau)} d\tau\right), \quad \bar{t}(t) := \int_0^t c(\tau) d\tau \tag{4.1.5}$$

and

$$\underline{R}_{g(t)} := \int_{\mathcal{M}^m} R_{g(t)} dV_{g(t)} \Big/ V_{g(t)}. \tag{4.1.6}$$

Then $\bar{g}(\bar{t})$ satisfies the **normalized Ricci flow**

$$\partial_{\bar{t}} \bar{g}(\bar{t}) = -2\mathrm{Ric}_{\bar{g}(\bar{t})} + \frac{2}{m}\underline{R}_{\bar{g}(\bar{t})} \bar{g}(\bar{t}). \tag{4.1.7}$$

Hence solutions of the normalized Ricci flow differ from solutions of the Ricci flow only by rescalings in spaces and time.

Note 4.1

Since

$$\frac{d\bar{t}}{dt} = c(t), \quad \frac{dc(t)}{dt} = \frac{2}{m}\underline{R}_{g(t)}c(t)$$

it follows that

$$\begin{aligned}\partial_{\bar{t}}\bar{g}(\bar{t}) &= \partial_t\bar{g}(\bar{t})\frac{dt}{d\bar{t}} = \left[\partial_t c(t)\cdot g(t) - 2c(t)\mathrm{Ric}_{g(t)}\right]\cdot c(t)^{-1} \\ &= \frac{2}{m}\underline{R}_{g(t)}g(t) - 2\mathrm{Ric}_{g(t)} = -2\mathrm{Ric}_{\bar{g}(\bar{t})} + \frac{2}{m}\underline{R}_{\bar{g}(\bar{t})}\bar{g}(\bar{t}).\end{aligned}$$

Thus we prove (4.1.7). ♣

4.1.2 Hamilton's theorem

The remainder of this chapter will be devoted to proving the following

Theorem 4.1. (Hamilton, 1982)

Let $(\mathcal{M}^3, g)$ be a closed Riemannian 3-manifold with positive Ricci curvature. Then there exists a unique solution $g(t)$ of the normalized Ricci flow with $g(0) = g$ for all $t \geq 0$. Furthermore, as $t \to \infty$, the metric $g(t)$ converge exponentially fast in every C^k-norm to a C^∞ metric g_∞ with constant positive sectional curvature. ♡

4.2 The maximum principle for tensor fields

Introduction

- *Hamilton's maximum principle for tensor fields*
- *Nonnegative Ricci curvature is preserved*
- *Ricci pinching is preserved*

4.2.1 Hamilton's maximum principle for tensor fields

Theorem 4.2. (Hamilton's maximum principle for tensor fields)

Let $g(t)$ be a smooth 1-parameter family of Riemannian metrics on a closed manifold $\mathcal{M}^m$. Let $\alpha(t)$ be a symmetric 2-tensor field satisfying

$$\square_t\alpha \geq \nabla_{X(t)}\alpha + \beta$$

where $X(t)$ is a time-dependent vector field and

$$\beta(x,t) = \beta(\alpha(x,t), g(x,t))$$

is a symmetric $(2,0)$-tensor field which is locally Lipschitz in all its arguments. Suppose that β satisfies the **null-eigenvector assumption** *that if A_{ij} is a nonnegative symmetric*

2-tensor at a point (x,t) and if V is such that $A_{ij}V^j = 0$, then

$$\beta_{ij}(A,g)V^iV^j \geq 0.$$

If $\alpha(0) \geq 0$, then $\alpha(t) \geq 0$ for all $t \geq 0$ as long as the soluton exists. ♡

Proof. Suppose that (x_1, t_1) is a point where there exists a vector V such that $(\alpha_{ij}V^j)(x_1,t_1) = 0$ for the first time so $(\alpha_{ij}W^iW^j)(x,t) \geq 0$ for all $W \in T_x\mathcal{M}^m$ and $t \leq t_1$. Choose V to be constant in time. We then have at (x_1, t_1)

$$\partial_t\left(\alpha_{ij}V^iV^j\right) = \left(\partial_t\alpha_{ij}\right)V^iV^j \geq \left(\Delta_{g(t)}\alpha_{ij}\right)V^iV^j + X^k\left(\nabla_k\alpha_{ij}\right)V^iV^j.$$

WE extend V in a neighborhood of x_1 by parallel translating it along geodesics, with respect to the metric $g(t_1)$, emanating from x_1. It is clear than $\nabla_{g(t_1)}V|_{x_1} = 0$ and $\Delta_{g(t_1)}V|_{x_1} = 0$. Thus we have

$$\Box_t\left(\alpha_{ij}V^iV^j\right) \geq X^k\nabla_k\left(\alpha_{ij}V^iV^j\right) \geq 0.$$

This shows that when α attains a zero eigenvalue for the first time, it wants to increase in the direction of any corresponding zero eigenvector. We can make the above argument rigorous by adding in an $\epsilon > 0$ just as for the scalar maximum principle. We can then show that there exists $\delta > 0$ such that $\alpha \geq 0$ on $[0,\delta]$ by applying the above argument to the symmetric 2-tensor $A_\epsilon(t) := \alpha(t) + \epsilon(\delta + t)g(t)$ for $\epsilon > 0$ sufficiently small and then letting $\epsilon \to 0$: We compute

$$\Box_t A_\epsilon(t) \geq X^k\nabla_k A_\epsilon(t) + \beta(\alpha(t), g(t)) + \epsilon(\delta+t)\partial_t g(t) + \epsilon g(t).$$

On any compact time interval, there exists $C < \infty$ such that $\partial_t g(t) \geq -Cg(t)$ and $\beta(\alpha(t), g(t)) - \beta(A_\epsilon(t), g(t)) \geq -C\epsilon(\delta+t)g(t)$. Thus

$$\Box_t A_\epsilon(t) \geq X^k\nabla_k A_\epsilon(t) + \beta(A_\epsilon(t), g(t)), \quad t \in [0,\delta],$$

by choosing $\delta < 1/4C$. Hence we can conclude that $A_\epsilon(t) > 0$ on $[0,\delta]$. Taking $\epsilon \to 0$ implies $\alpha(t) \geq 0$ on $[0,\delta]$. Continuing this way, we conclude that $\alpha(t) \geq 0$ on all of I. □

4.2.2 Nonnegative Ricci curvature is preserved

Recall that the evolution equation of the Ricci tensor field is given by

$$\partial_t R_{ij} = \Delta_{L,g(t)}R_{ij} = \Delta_{g(t)}R_{ij} + 2R_{kij\ell}R^{k\ell} - 2R_{ik}R_j{}^k. \tag{4.2.1}$$

When $m = 3$, the Weyl tensor field vanishes so that

$$R_{ijk\ell} = R_{i\ell}g_{jk} + R_{jk}g_{i\ell} - R_{ik}g_{j\ell} - R_{j\ell}g_{ik} - \frac{R_g}{2}\left(g_{i\ell}g_{jk} - g_{ik}g_{j\ell}\right). \tag{4.2.2}$$

Lemma 4.1

If $m = 3$, then under the Ricci flow we have

$$\Box_t R_{ij} = 3R_{g(t)}R_{ij} - 6R_{ip}R_j{}^p + \left(2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - R^2_{g(t)}\right)g_{ij}. \tag{4.2.3}$$

♡

Proof. Using (4.2.1) and (4.2.2) yields

$$\begin{aligned} 2R_{kij\ell}R^{k\ell} &= 2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} g_{ij} + 2R_{g(t)}R_{ij} - 4R_i{}^kR_{kj} - R_{g(t)}\left(R_{g(t)}g_{ij} - R_i{}^kg_{kj}\right) \\ &= 2\left|\mathrm{Rc}_{g(t)}\right|^2_{g(t)} g_{ij} + 3R_{g(t)}R_{ij} - 4R_i{}^kR_{kj} - R^2_{g(t)}g_{ij} \end{aligned}$$

which implies (4.2.3). □

Corollary 4.1. (Nonnegative Ricci is preserved)

If $(\mathcal{M}^3, g(t))$, $t \in [0, T)$ is a solution to the Ricci floe on a closed 3-manifold with $\mathrm{Ric}_{g(0)} \geq 0$, then $\mathrm{Ric}_{g(t)} \geq 0$ for all $t \in [0, T)$. ♡

Proof. Let

$$\beta_{ij} = 3R_{g(t)}R_{ij} - 6R_{ip}R_j{}^p + \left(2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - R^2_{g(t)}\right)g_{ij}.$$

If at a point and time $\mathrm{Ric}_{g(t)}$ has a null-eigenvector $V = V^i\partial_i$, then one of eigenvalues of $\mathrm{Ric}_{g(t)}$ is zero so that $2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - R^2_{g(t)} \geq 0$ and

$$\beta_{ij}V^iV^j = \left(2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - R^2_{g(t)}\right)|V|^2_{g(t)} \geq 0.$$

Therefore β satisfies the null-eigenvector assumption with respect to R_{ij}. Applying **Theorem 4.2** to this case, $\mathrm{Ric}_{g(t)} \geq 0$ as long as $\mathrm{Ric}_{g(0)} \geq 0$. □

4.2.3 Ricci pinching is preserved

Recall the Einstein tensor field

$$\mathrm{Ein}_g = \mathrm{Ric}_g - \frac{1}{2}R_g g. \tag{4.2.4}$$

In general, we consider the **ε-Einstein tensor field**

$$\mathrm{Ein}_{g,\varepsilon} := \mathrm{Ric}_g - \varepsilon R_g g. \tag{4.2.5}$$

In particular, $\mathrm{Ein}_{g,\frac{1}{2}} = \mathrm{Ein}_g$. Using (2.3.1) and (4.2.3) we have

$$\begin{aligned} \Box_t\left(R_{ij} - \epsilon R_{g(t)}g_{ij}\right) &= 3R_{g(t)}R_{ij} - 6R_i{}^pR_{jp} + \left(2\left|\mathrm{Rc}_{g(t)}\right|^2_{g(t)} - R^2_{g(t)}\right)g_{ij} \\ &\quad - 2\varepsilon\left|\mathrm{Rc}_{g(t)}\right|^2_{g(t)} g_{ij} + 2\varepsilon R_{g(t)}R_{ij}. \end{aligned}$$

Suppose

$$\begin{aligned} \beta_{ij} &= 3R_{g(t)}R_{ij} - 6R_i{}^pR_{jp} + \left(2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - R^2_{g(t)}\right)g_{ij} \\ &\quad - 2\varepsilon\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} g_{ij} + 2\varepsilon R_{g(t)}R_{ij} \end{aligned}$$

and $\left(R_{ij} - \varepsilon R_{g(t)}g_{ij}\right)V^j = 0$. Then

$$\begin{aligned} \beta_{ij}V^iV^j &= 3R_{g(t)}R_{ij}V^iV^j - 6R_i{}^pR_{jp}V^iV^j + \left(2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - R^2_{g(t)}\right)|V|^2_{g(t)} \\ &\quad - 2\varepsilon\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}|V|^2_{g(t)} + 2\varepsilon R_{g(t)}R_{ij}V^iV^j \\ &= \left[(3\varepsilon - 1 - 4\varepsilon^2)R^2_{g(t)} + (2 - 2\varepsilon)\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}\right]|V|^2_{g(t)}. \end{aligned}$$

Since one on eigenvalues of $\mathrm{Ric}_{g(t)}$ is $\varepsilon R_{g(t)}$, it follows that other eigenvalues $\lambda_2 + \lambda_3$ satisfy $\lambda_2 + \lambda_3 = (1-\varepsilon)R_{g(t)}$. By the elementary inequalities

$$\frac{(\lambda_2+\lambda_3)^2}{2} \le \lambda_2^2 + \lambda_3^2 \le (\lambda_2+\lambda_3)^2$$

we conclude that

$$\frac{(1-\varepsilon)^2}{2} R_{g(t)}^2 \le \lambda_2^2 + \lambda_3^2 \le (1-\varepsilon)^2 R_{g(t)}^2$$

and hence

$$\left(\varepsilon^2 + \frac{(1-\varepsilon)^2}{2}\right) R_{g(t)}^2 \le \left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^2 \le \left(\varepsilon^2 + (1-\varepsilon)^2\right) R_{g(t)}^2.$$

If $\epsilon < 1$, then

$$\begin{aligned} \beta_{ij}V^iV^j &\ge \left[(3\varepsilon - 1 - 4\varepsilon^2) + (2-2\varepsilon)\left(\varepsilon^2 + \frac{(1-\varepsilon)^2}{2}\right)\right] R_{g(t)}^2 |V|_{g(t)}^2 \\ &= \varepsilon(1-3\varepsilon)R_{g(t)}^2|V|_{g(t)}^2. \end{aligned}$$

Hence when $0 \le \varepsilon \le \frac{1}{3}$, we have $\beta_{ij}V^iV^j \ge 0$. On the other hand,

$$-\beta_{ij}V^iV^j = \left((2\varepsilon-2)\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^2 + (4\varepsilon^2 - 3\varepsilon + 1)R_{g(t)}^2\right)|V|_{g(t)}^2.$$

If $\varepsilon \ge 1$, then

$$\begin{aligned} -\beta_{ij}V^iV^j &\ge \left[(2\varepsilon-2)\left(\varepsilon^2 + \frac{(1-\varepsilon)^2}{2}\right) + 4\varepsilon^2 - 3\varepsilon + 1\right] R_{g(t)}^2|V|_{g(t)}^2 \\ &= \varepsilon(3\varepsilon-1)R_{g(t)}^2|V|_{g(t)}^2 \ \ge \ 0. \end{aligned}$$

If $0 \le \varepsilon \le 1$, then

$$\begin{aligned} -\beta_{ij}V^iV^j &\ge \left[(4\varepsilon^2 - 3\varepsilon + 1) + (2\varepsilon-2)(\varepsilon^2 + (1-\varepsilon)^2)\right] R_{g(t)}^2|V|_{g(t)}^2 \\ &= \left(\varepsilon - \frac{1}{2}\right)(4\varepsilon^2 - 2\varepsilon + 2)R_{g(t)}^2|V|_{g(t)}^2. \end{aligned}$$

In this case, $-\beta_{ij}V^iV^j \ge 0$ provided $\varepsilon \ge \frac{1}{2}$.

Corollary 4.2

Given an $\eta \in [1/2, +\infty)$. If $(\mathcal{M}^3, g(t))$, $t \in [0,T)$, is a solution to the Ricci flow on a closed 3-manifold with $\mathrm{Ein}_{g(0),\eta} \le 0$, then $\mathrm{Ein}_{g(t),\eta} \le 0$ for all $t \in [0,T)$. ♡

Corollary 4.3

Given an $\varepsilon \in [0, 1/3]$. If $(\mathcal{M}^3, g(t))$, $t \in [0,T)$, is a solution to the Ricci flow on a closed 3-manifold with $\mathrm{Ein}_{g(0),\eta} \ge 0$, then $\mathrm{Ein}_{g(t),\eta} \ge 0$ for all $t \in [0,T)$. ♡

Corollary 4.4

If $(\mathcal{M}^3, g(t))$, $t \in [0,T)$, is a solution to the Ricci flow on a closed 3-manifold with $\frac{1}{3}R_{g(0)}g(0) \le \mathrm{Ric}_{g(0)} \le \frac{1}{2}R_{g(0)}g(0)$, then $\frac{1}{3}R_{g(t)}g(t) \le \mathrm{Ric}_{g(t)} \le \frac{1}{2}R_{g(t)}g(t)$ for all $t \in [0,T)$. ♡

4.3 Curvature pinching estimates

Introduction

- *Maximum principle for systems*
- *Nonnegativity of scalar, sectional curvature, and Ricci are preserved*
- *Ricci pinching is preserved*
- *Ricci pinching improves*

4.3.1 Maximum principle for systems

Recall that Rm_g is a section of the bundle $\pi : \mathcal{E} \to \mathcal{M}^m$, where $\mathcal{E} := \wedge^2 T^*\mathcal{M}^m \otimes_S \wedge^2 T^*\mathcal{M}^m$. The Ricci tensor field Rm_g can be also considered as an operator

$$\mathbf{Rm}_g : C^\infty(\wedge^2 T^*\mathcal{M}^m) \to C^\infty(\wedge^2 T^*\mathcal{M}^m).$$

The bundle $\mathcal{E}$ has a natural bundle metric and Levi-Civita connection induced by the Riemanian metric and Levi-Civita connection on $T\mathcal{M}^m$. Let $\mathcal{E}_p := \pi^{-1}(p)$ be the fiber over p. For each $p \in \mathcal{M}^m$, consider the system of ODE on $\mathcal{E}_p$ corresponding to the PDE (3.7.14) obtained by dropping the Laplacian term:

$$\frac{d}{dt}\mathbf{M}(t) = \mathbf{M}^2(t) + \mathbf{M}^\#(t), \tag{4.3.1}$$

where $(\mathbf{M}(t))_p \in \mathcal{E}_p$ is a symmetric $N \times N$ matrix, where $N = \frac{m(m-1)}{2} = \dim(\mathfrak{so}(m))$. Actually,

$$\mathbf{M}(t)(X \wedge Y, W \wedge Z) = \left\langle \mathbf{Rm}_{g(t)}(X \wedge Y), Z \wedge W \right\rangle_{g(t)}. \tag{4.3.2}$$

A set K in a vector space is said to be **convex** if for any $X, Y \in K$, we have $sX + (1-s)Y \in K$ for all $s \in [0,1]$. A subset $\mathcal{K}$ of the vector bundle $\mathcal{E}$ is said to be **invariant under parallel translation** if for every path $\gamma : [a,b] \to \mathcal{M}^m$ and vector $X \in \mathcal{K} \cap \mathcal{E}_{\gamma(a)}$, the unique parallel section $X(s) \in \mathcal{E}_{\gamma(s)}$, $s \in [a,b]$, along $\gamma(s)$ with $X(a) = X$ is contained in $\mathcal{K}$.

Theorem 4.3. (Maximum principle applied to the curvature operator)

Let $g(t)$, $t \in [0,T)$, be a solution to the Ricci flow on a closed m-manifold $\mathcal{M}^m$. Let $\mathcal{K} \subset \mathcal{E}$ be a subset which is invariant under parallel translation and whose intersection $K_p := \mathcal{K} \cap \mathcal{E}_p$ with each fiber is closed and convex. Suppose the ODE *(4.3.1) has the property that for any $\mathbf{M}(0) \in \mathcal{K}$, we have $\mathbf{M}(t) \in \mathcal{K}$ for all $t \in [0,T)$. If $\mathbf{Rm}_{g(0)} \in \mathcal{K}$, then $\mathbf{Rm}_{g(t)} \in \mathcal{K}$ for all $t \in [0,T)$.* ♡

Corollary 4.5

If $(\mathcal{M}^m, g(t))$, $t \in [0,T)$, is a solution to the Ricci flow on a closed m-manifold with $\mathbf{Rm}_{g(0)} \geq 0$, then $\mathbf{Rm}_{g(t)} \geq 0$ for all $t \in [0,T)$. ♡

4.3.2 Nonnegativity of scalar, sectional curvature, and Ricci are preserved

In dimension 3, if $\mathbf{M}(0)$ is diagonal, then $\mathbf{M}(t)$ remains diagonal for all $t \in [0, T)$. Let $\lambda_1(\mathbf{M}(t)) \leq \lambda_2(\mathbf{M}(t)) \leq \lambda_3(\mathbf{M}(t))$ be the eigenvalues of $\mathbf{M}(t)$. Under the ODE the ordering of the eigenvalues is preserved and we have

$$\frac{d}{dt}\lambda_1(\mathbf{M}(t)) = \lambda_1(\mathbf{M}(t))^2 + \lambda_2(\mathbf{M}(t))\lambda_3(\mathbf{M}(t)), \tag{4.3.3}$$

$$\frac{d}{dt}\lambda_2(\mathbf{M}(t)) = \lambda_2(\mathbf{M}(t))^2 + \lambda_1(\mathbf{M}(t))\lambda_3(\mathbf{M}(t), \tag{4.3.4}$$

$$\frac{d}{dt}\lambda_3(\mathbf{M}(t)) = \lambda_3(\mathbf{M}(t))^2 + \lambda_1(\mathbf{M}(t))\lambda_2(\mathbf{M}(t)). \tag{4.3.5}$$

With this setup, we can come up with a number of closed, fiberwise convex sets $\mathcal{K}$, invariant under parallel translation, which are preserved by the ODE. Each such set corresponds to an a *priori* estimate for the curvature $\mathbf{Rm}_g$.

The following sets $\mathcal{K} \subset \mathcal{E}$ are invariant under parallel translation and for each $p \in \mathcal{M}^m$, $\mathcal{K}_p$ is closed, convex and preserved by the ODE.

(1) **Lower bound of scalar is preserved:** Given $C_0 \in \mathbf{R}$, let

$$\mathcal{K} = \{\mathbf{M} : \lambda_1(\mathbf{M}) + \lambda_2(\mathbf{M}) + \lambda_3(\mathbf{M}) \geq C_0\}.$$

The trace $(\lambda_1(\mathbf{M}))_p + (\lambda_2(\mathbf{M}))_p + (\lambda_3(\mathbf{M}))_p : \mathcal{E}_p \to \mathbf{R}$ is a linear function, which implies that K_p is closed and convex for each $p \in \mathcal{M}^m$. That $\mathcal{K}$ is preserved by the ODE follows from

$$\begin{aligned} & \frac{d}{dt}\left(\lambda_1(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t)) + \lambda_3(\mathbf{M}(t))\right) \\ = \; & \frac{1}{2}\left[(\lambda_1(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t)))^2 + (\lambda_1(\mathbf{M}(t)) + \lambda_3(\mathbf{M}(t)))^2 \right. \\ & \left. +(\lambda_2(\mathbf{M}(t)) + \lambda_3(\mathbf{M}(t)))^2\right] \\ \geq \; & \frac{2}{3}\left(\lambda_1(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t)) + \lambda_3(\mathbf{M}(t))\right)^2 \; \geq \; 0. \end{aligned}$$

Therefore, if $R_{g(0)} \geq C_0$ for some $C_0 \in \mathbf{R}$, then

$$R_{g(t)} \geq C_0 \tag{4.3.6}$$

for all $t \geq 0$.

(2) **Nonnegative sectional curvature is preserved:** Let $\mathcal{K} = \{\mathbf{M} : \lambda_1(\mathbf{M}) \geq 0\}$. Each K_p is closed and convex since $(\lambda_1(\mathbf{M}))_p : \mathcal{E}_p \to \mathbf{R}$ is a concave function. We see that $\mathcal{K}$ is preserved by the ODE since

$$\frac{d}{dt}\lambda_1(\mathbf{M}(t)) = \lambda_1(\mathbf{M}(t))^2 + \lambda_2(\mathbf{M}(t))\lambda_3(\mathbf{M}(t)) \geq 0$$

whenever $\lambda_1(\mathbf{M}(t)) \geq 0$ (Since $\lambda_1(\mathbf{M}(0)) \geq 0$, it follows that $\frac{d}{dt}|_{t=0}\lambda_1(\mathbf{M}(t)) \geq 0$. Thus $\lambda_1(\mathbf{M}(t)) \geq 0$ for t sufficiently close to 0. By continuity, $\lambda_1(\mathbf{M}(t)) \geq 0$ for all $t \geq 0$). This implies

$$\mathbf{Rm}_{g(t)} \geq 0 \tag{4.3.7}$$

for all $t \geq 0$ provided $\mathbf{Rm}_{g(0)} \geq 0$.

(3) **Nonnegative Ricci is preserved:** Let $\mathcal{K} = \{\mathbf{M} : \lambda_1(\mathbf{M}) + \lambda_2(\mathbf{M}) \ge 0\}$. Since $\lambda_1 + \lambda_2$ is concave, $\mathcal{K}$ is closed and convex. From

$$\begin{aligned}\frac{d}{dt}\left(\lambda_1(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t))\right) &= \lambda_1(\mathbf{M}(t))^2 + \lambda_2(\mathbf{M}(t))^2 \\ &+ \left(\lambda_1(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t))\right)\lambda_3(\mathbf{M}(t)) \ge 0\end{aligned}$$

whenever $\lambda_1(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t)) \ge 0$, we see that $\mathcal{K}$ is preserved by the ODE. From this we see that

$$\mathrm{Rc}_{g(t)} \ge 0 \tag{4.3.8}$$

for all $t \ge 0$ if $\mathrm{Rc}_{g(0)} \ge 0$, since the smallest eigenvalue of $\mathrm{Rc}_{g(t)}$ is

$$\frac{1}{2}\left[\lambda_1(\mathbf{Rm}_{g(t)}) + \lambda_2(\mathbf{Rm}_{g(t)})\right]$$

(see below).

4.3.3 Ricci pinching is preserved

Recall that

$$\lambda_1(\mathbf{M}) = \min_{|U\wedge V|_g=1} \mathbf{M}(U\wedge V, U\wedge V), \quad \lambda_3(\mathbf{M}) = \max_{|U\wedge V|_g=1} \mathbf{M}(U\wedge V, U\wedge V). \tag{4.3.9}$$

Hence, λ_1 is concave, λ_3 is convex, and $\lambda_1 + \lambda_2$ is concave. To compute the eigenvalues for Rc_g, we chose an orthonormal frame $\{e_1, e_2, e_3\}$ and its dual orthonormal coframe $\{\omega^1, \omega^2, \omega^3\}$ such that the 2-forms

$$\omega^2 \wedge \omega^3, \quad \omega^3 \wedge \omega^1, \quad \omega^1 \wedge \omega^2$$

are eigenvectors of $\mathbf{Rm}_g$. In this case

$$\begin{aligned}\lambda_1(\mathbf{Rm}_g) &= 2\mathrm{Rm}_g(e_2, e_3, e_3, e_2), \\ \lambda_2(\mathbf{Rm}_g) &= 2\mathrm{Rm}_g(e_1, e_3, e_3, e_1), \\ \lambda_3(\mathbf{Rm}_g) &= 2\mathrm{Rm}_g(e_1, e_2, e_2, e_1).\end{aligned}$$

The Ricci tensor Ric_g gives rise to an operator

$$\mathbf{Rc}_g : \mathcal{A}^1(\mathcal{M}^3) \longrightarrow \mathcal{A}^1(\mathcal{M}^3) \tag{4.3.10}$$

given by

$$\left(\mathbf{Rc}_g(\alpha)\right)_i := g^{jk} R_{ij}\alpha_k. \tag{4.3.11}$$

Hence its eigenvalues are

$$\lambda_j\left(\mathbf{Ric}_g\right) = \sum_{1\le i\le 3} \langle \mathrm{Rm}_g(e_j, e_i)e_i, e_j\rangle_g = \sum_{1\le i\le 3} \mathrm{Rm}_g(e_j, e_i, e_i, e_j).$$

Explicitly,

$$\begin{aligned}\lambda_1\left(\mathbf{Rc}_g\right) &= \frac{1}{2}\left[\lambda_2\left(\mathbf{Rm}_g\right) + \lambda_3\left(\mathbf{Rm}_g\right)\right], \\ \lambda_2\left(\mathbf{Rc}_g\right) &= \frac{1}{2}\left[\lambda_1\left(\mathbf{Rm}_g\right) + \lambda_3\left(\mathbf{Rm}_g\right)\right], \\ \lambda_3\left(\mathbf{Rc}_g\right) &= \frac{1}{2}\left[\lambda_1\left(\mathbf{Rm}_g\right) + \lambda_2\left(\mathbf{Rm}_g\right)\right].\end{aligned}$$

Since $\lambda_1(\mathbf{Rm}_g) \le \lambda_2(\mathbf{Rm}_g) \le \lambda_3(\mathbf{Rm}_g)$, it follows that $\lambda_3(\mathbf{Rc}_g)$ is the smallest eigenvalue of $\mathbf{Rc}_g$. Meanwhile, $\lambda_1(\mathbf{Rc}_g)$ is the largest eigenvalue of $\mathbf{Rc}_g$.

Given $C \ge \frac{1}{2}$, let

$$\mathcal{K} = \{\mathbf{M} : \lambda_3(\mathbf{M}) \le C[\lambda_1(\mathbf{M}) + \lambda_2(\mathbf{M})]\}.$$

Then $\mathcal{K}_p$ is convex for all $p \in \mathcal{M}^m$. That each $\mathcal{K}_p$ is preserved by the ODE follows from

$$\begin{aligned} & \frac{d}{dt}[\lambda_3(\mathbf{M})(t) - C(\lambda_1(\mathbf{M})(t) + \lambda_2(\mathbf{M})(t))] \\ = \ & \lambda_3(\mathbf{M}(t))[\lambda_3(\mathbf{M})(t) - C(\lambda_1(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t)))] \\ & - C\left[\lambda_1(\mathbf{M}(t))^2 - \frac{1}{C}\lambda_1(\mathbf{M}(t))\lambda_2(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t))^2\right]. \end{aligned}$$

Since $C \ge \frac{1}{2}$, it follows that

$$\begin{aligned} & \frac{d}{dt}[\lambda_3(\mathbf{M})(t) - C(\lambda_1(\mathbf{M})(t) + \lambda_2(\mathbf{M})(t))] \\ \le \ & \lambda_3(\mathbf{M}(t))[\lambda_3(\mathbf{M})(t) - C(\lambda_1(\mathbf{M}(t)) + \lambda_2(\mathbf{M}(t)))]. \end{aligned}$$

Set

$$f(t) := \lambda_3(\mathbf{M})(t) - C(\lambda_1(\mathbf{M})(t) + \lambda_2(\mathbf{M})(t)).$$

By assumption, $f(0) = 0$. If $f(t_1) > 0$ for some time t_1, then there exists a time t_0 so that $f(t_0) = 0$ and $f(t) < 0$ on $t \in (t_0 - \delta, t_0)$ for some small number $\delta > 0$. However, at the time t_0, we have

$$\left.\frac{d}{dt}\right|_{t=t_0} f(t) \le 0;$$

consequently, $f(t) \ge 0$ on some small neighborhood of t_0. This is a contradiction. Hence $f(t) \le 0$ for all $t \ge 0$.

Suppose $\mathrm{Rc}_{g(0)} > 0$. Then $\mathrm{Rc}_{g(t)} > 0$ for all times $t \ge 0$. Since M is compact, there exists $C \ge \frac{1}{2}$ such that

$$\lambda_3(\mathbf{Rm}_{g(0)}) \le C\left[\lambda_1(\mathbf{Rm}_{g(0)}) + \lambda_2(\mathbf{Rm}_{g(0)})\right].$$

That is $\mathbf{Rm}_{g(0)} \subset \mathcal{K}$. By the maximum principle for tensor fields, $\mathbf{Rm}_{g(t)} \in \mathcal{K}$ and hence

$$\lambda_3\left(\mathbf{Rm}_{g(t)}\right) \le C\left[\lambda_1\left(\mathbf{Rm}_{g(t)}\right) + \lambda_2\left(\mathbf{Rm}_{g(t)}\right)\right]$$

holds for all $t \ge 0$. Therefore

$$\begin{aligned} \mathrm{Rc}_{g(t)} & \ge \lambda_3\left(\mathbf{Rc}_{g(t)}\right) g(t) = \left(\frac{\lambda_1\left(\mathbf{Rm}_{g(t)}\right) + \lambda_2\left(\mathbf{Rm}_{g(t)}\right)}{2}\right) g(t) \\ & \ge \frac{\lambda_3\left(\mathbf{Rm}_{g(t)}\right)}{2C} g(t) \ge \frac{R_{g(t)}}{6C} g(t) \end{aligned}$$

since $\lambda_1\left(\mathbf{Rm}_{g(t)}\right) + \lambda_2\left(\mathbf{Rm}_{g(t)}\right) + \lambda_3\left(\mathbf{Rm}_{g(t)}\right) = R_{g(t)}$. From the assumption that $C \ge \frac{1}{2}$, we have $\frac{1}{6C} \le \frac{1}{3}$ and hence

$$\mathrm{Rc}_{g(t)} \ge \varepsilon R_{g(t)} g(t), \quad n = 3 \tag{4.3.12}$$

for any $\varepsilon \in (0, 1/3]$, is preserved under the Ricci flow.

Note 4.2

Consider the same $\mathcal{K}$. Suppose that $\mathbf{Rm}_{g(0)} \subset \mathcal{K}$ and

$$\left[\lambda_1\left(\mathbf{Rm}_{g(t)}\right)(x) + \lambda_2\left(\mathbf{Rm}_{g(t)}\right)(x)\right]\Big|_{(x,t)=(x_0,t_0)} < 0$$

for some $(x_0, t_0) \in \mathcal{M}^3 \times [0, T)$.

(1) *We claim that $\lambda_1\left(\mathbf{Rm}_{g(t)}\right) = \lambda_2\left(\mathbf{Rm}_{g(t)}\right) = \lambda_3\left(\mathbf{Rm}_{g(t)}\right)$ at (x_0, t_0). Indeed, at the point (x_0, t_0), we have*

$$\begin{aligned}
& \lambda_3\left(\mathbf{Rm}_{g(t)}\right) - \lambda_1\left(\mathbf{Rm}_{g(t)}\right) \\
\leq\ & C\lambda_1\left(\mathbf{Rm}_{g(t)}\right) + C\lambda_2\left(\mathbf{Rm}_{g(t)}\right) - \lambda_1\left(\mathbf{Rm}_{g(t)}\right) \\
\leq\ & C\lambda_3\left(\mathbf{Rm}_{g(t)}\right) + (C-1)\lambda_1\left(\mathbf{Rm}_{g(t)}\right) \\
=\ & C\left(\lambda_3\left(\mathbf{Rm}_{g(t)}\right) - \lambda_1\left(\mathbf{Rm}_{g(t)}\right)\right) + (2C-1)\lambda_1\left(\mathbf{Rm}_{g(t)}\right).
\end{aligned}$$

On the other hand, we note that

$$\lambda_3\left(\mathbf{Rm}_{g(t)}\right) \leq C\left(\lambda_1\left(\mathbf{Rm}_{g(t)}\right) + \lambda_2\left(\mathbf{Rm}_{g(t)}\right)\right) < 0$$

at the point (x_0, t_0). Therefore when $C \geq 1$, we obtain

$$\lambda_3\left(\mathbf{Rm}_{g(t)}\right) - \lambda_1\left(\mathbf{Rm}_{g(t)}\right) \leq 0;$$

consequently, $\lambda_1\left(\mathbf{Rm}_{g(t)}\right) = \lambda_2\left(\mathbf{Rm}_{g(t)}\right) = \lambda_3\left(\mathbf{Rm}_{g(t)}\right) := \lambda < 0$ at (x_0, t_0). When $\frac{1}{2} \leq C < 1$, we have

$$\lambda_3\left(\mathbf{Rm}_{g(x,t)}\right) - \lambda_1\left(\mathbf{Rm}_{g(t)}\right) \leq C\left(\lambda_3\left(\mathbf{Rm}_{g(t)}\right) - \lambda_1\left(\mathbf{Rm}_{g(t)}\right)\right);$$

consequently,

$$\lambda_3\left(\mathbf{Rm}_{g(t)}\right) - \lambda_1\left(\mathbf{Rm}_{g(t)}\right) \leq 0$$

and $\lambda_1\left(\mathbf{Rm}_{g(t)}\right) = \lambda_2\left(\mathbf{Rm}_{g(t)}\right) = \lambda_3\left(\mathbf{Rm}_{g(t)}\right) := \lambda < 0$ at (x_0, t_0).

(2) *But, in both cases, $\lambda \leq 2C\lambda$ so that $(2C-1)\lambda \geq 0$. From this, we deduce that $\lambda \geq 0$ when $C > \frac{1}{2}$, a contradiction. Therefore, $\mathbf{Rm}_{g(0)} \subset \mathcal{K}$ implies*

$$\lambda_1\left(\mathbf{Rm}_{g(t)}\right) + \lambda_2\left(\mathbf{Rm}_{g(t)}\right) \geq 0 \tag{4.3.13}$$

for all $t \in [0, T)$ provided $C > \frac{1}{2}$.

(3) *$C = \frac{1}{2}$. In this case, we have*

$$\lambda = \frac{1}{3}R_{g(t_0)}(x_0).$$

Since

$$\lambda_1\left(\mathbf{Rm}_{g(t)}\right) + \lambda_2\left(\mathbf{Rm}_{g(t)}\right) < 0$$

holds in a neighborhood of x_0 at time t_0, we must have

$$\lambda_1\left(\mathbf{Rm}_{g(t)}\right) = \lambda_2\left(\mathbf{Rm}_{g(t)}\right) = \lambda_3\left(\mathbf{Rm}_{g(t)}\right) = \frac{1}{3}R_{g(t)}$$

in U. By the contacted Bianchi identity, we then have that $R_{g(t)}$ is constant on U.

Since M is connected, it follows that

$$\lambda_1\left(\mathbf{Rm}_{g(t)}\right)=\lambda_2\left(\mathbf{Rm}_{g(t)}\right)=\lambda_3\left(\mathbf{Rm}_{g(t)}\right)=\frac{1}{3}R_{g(t)}=\frac{1}{3}R$$

on all of $M\times[0,T)$, where the scalar curvature $R_{g(t)}$ is a negative constant R. Thus, if $\mathbf{Rm}_{g(t_0)}\subset\mathcal{K}$ (where $C=\frac{1}{2}$) for some t_0 and if $g(t_0)$ does not have constant negative sectional curvature, then $\mathrm{Ric}_{g(t)}\geq 0$ for $t\in[0,T)$. ♣

There is an interesting question related to (4.3.12).

Problem 4.1. (Hamilton)

If $(\mathcal{M}^3,g)$ is a complete Riemannian 3-manifold with $\mathrm{Rc}_g\geq\varepsilon R_g g$, where $R_g>0$ and $\varepsilon>0$, then $\mathcal{M}^3$ is compact. ♠

Chen and Zhu proved that if $(\mathcal{M}^3,g)$ is a complete Riemannian 3-manifold with bounded nonnegative sectional curvature and $\mathrm{Rc}_g\geq\varepsilon R_g g$ with $\varepsilon>0$, then M is either compact or flat. A related question is

Problem 4.2

If $(\mathcal{M}^3,g(t))$ is a complete solution to the Ricci flow on a 3-manifold with nonnegative Ricci curvature which is bounded on compact time intervals, can one prove a trace differential Harnack inequality? ♠

Note that a result related to **Problem 4.1**, due to Hamilton, is

Theorem 4.4. (Hamilton, 1994)

If $\mathcal{M}^m\subset\mathbf{R}^{m+1}$ is a C^∞ complete, strictly convex hypersurface with $h_{ij}\geq\varepsilon Hg_{ij}$ for some $\varepsilon>0$, then $\mathcal{M}^m$ is compact. ♡

Problem 4.3

Does there exist a Harnack inequality for solutions to the mean curvature flow with nonnegative mean curvature and second fundamental form which is bounded on compact time intervals? ♠

4.3.4 Ricci pinching improves

Given $C_0>0$, $C_1\geq\frac{1}{2}$, $C_2>0$ and $0<\delta<1$, let

$$\mathcal{K}=\left\{\mathbf{M}:\begin{array}{c}\lambda_3(\mathbf{M})-\lambda_1(\mathbf{M})-C_2\left[\lambda_1(\mathbf{M})+\lambda_2(\mathbf{M})+\lambda_3(\mathbf{M})\right]^{1-\delta}\leq 0,\\ \lambda_3(\mathbf{M})\leq C_1\left[\lambda_1(\mathbf{M})+\lambda_2(\mathbf{M})\right],\\ \lambda_1(\mathbf{M})+\lambda_2(\mathbf{M})+\lambda_3(\mathbf{M})\geq C_0\end{array}\right\}. \tag{4.3.14}$$

$\mathcal{K}$ is a convex set since $\lambda_3-\lambda_1-C_2(\lambda_1+\lambda_2+\lambda_3)^{1-\delta}$ is a convex function for $C_2>0$.

If $\mathbf{M}\in\mathcal{K}$, then

$$C_0\leq\lambda_1(\mathbf{M})+\lambda_2(\mathbf{M})+C_1\left[\lambda_1(\mathbf{M})+\lambda_2(\mathbf{M})\right]$$

so that

$$\lambda_1(\mathbf{M}) + \lambda_2(\mathbf{M}) \ge \frac{C_0}{1+C_1} > 0.$$

We have proved (omit $\mathbf{M}$) that the inequalities $\lambda_1+\lambda_2+\lambda_3 \ge C_0$ and $\lambda_3 \le C_1(\lambda_1+\lambda_2)$ are preserved under the ODE. We need only to check the first inequality is also preserved under the ODE. Since $C_0 > 0$,

$$\begin{aligned}
& \frac{d}{dt}\left(\ln \frac{\lambda_3-\lambda_1}{(\lambda_1+\lambda_2+\lambda_3)^{1-\delta}}\right) \\
= \ & \frac{(\lambda_1+\lambda_2+\lambda_3)^{1-\delta}}{\lambda_3-\lambda_1} \cdot \frac{d}{dt}\left(\frac{\lambda_3-\lambda_1}{(\lambda_1+\lambda_2+\lambda_3)^{1-\delta}}\right) \\
= \ & \frac{\lambda_3^2+\lambda_1\lambda_2-\lambda_1^2-\lambda_2\lambda_3}{\lambda_3-\lambda_1} \\
& -\frac{1-\delta}{\lambda_1+\lambda_2+\lambda_3}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_1\lambda_3+\lambda_2\lambda_3+\lambda_1\lambda_2\right).
\end{aligned}$$

Note that

$$\begin{aligned}
& \lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_1\lambda_3+\lambda_2\lambda_3+\lambda_1\lambda_2 \\
= \ & (\lambda_1+\lambda_2)\lambda_2+(\lambda_2-\lambda_1)\lambda_3+\lambda_2^2+(\lambda_1-\lambda_3)^2-\lambda_2^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
& \frac{d}{dt}\ln\left(\frac{\lambda_3-\lambda_1}{(\lambda_1+\lambda_2+\lambda_3)^{1-\delta}}\right) \\
= \ & \delta(\lambda_1+\lambda_3-\lambda_2)-(1-\delta)\frac{(\lambda_1+\lambda_2)\lambda_2+(\lambda_2-\lambda_1)\lambda_3+\lambda_2^2}{\lambda_1+\lambda_2+\lambda_3}.
\end{aligned}$$

Because $\lambda_1+\lambda_2 > 0$, we have $\lambda_3 \ge \lambda_2 > 0$ and hence

$$\begin{aligned}
& \frac{d}{dt}\ln\left(\frac{\lambda_3-\lambda_1}{(\lambda_1+\lambda_2+\lambda_3)^{1-\delta}}\right) \\
\le \ & \delta(\lambda_1+\lambda_3-\lambda_2)-(1-\delta)\frac{\lambda_2^2}{\lambda_1+\lambda_2+\lambda_3}.
\end{aligned}$$

Note that

$$\frac{\lambda_2^2}{\lambda_1+\lambda_2+\lambda_3} \ge \frac{\frac{\lambda_1+\lambda_2}{2}}{3\lambda_3}\lambda_2 \ge \frac{(\lambda_1+\lambda_2)\lambda_2}{6\lambda_3} \ge \frac{1}{6C_1}\lambda_2,$$

and

$$\lambda_1+\lambda_3-\lambda_2 \le \lambda_3 \le C_1(\lambda_1+\lambda_2) \le 2C_1\lambda_2.$$

Combining those inequalities yields

$$\frac{d}{dt}\left(\ln\frac{\lambda_3-\lambda_1}{(\lambda_1+\lambda_2+\lambda_3)^{1-\delta}}\right) \le \left(2\delta C_1-\frac{1-\delta}{6C_1}\right)\lambda_2.$$

If we chose $\delta \in (0,1)$ small enough so that $\frac{\delta}{1-\delta} \le \frac{1}{12C_1^2}$ ($\delta \le \frac{1}{4}$), then

$$\frac{d}{dt}\left(\ln\frac{\lambda_3-\lambda_1}{(\lambda_1+\lambda_2+\lambda_3)^{1-\delta}}\right) \le 0.$$

Thus, K is preserved by the ODE provided $C_0 > 0$, $C_1 \ge \frac{1}{2}$, $C_2 > 0$, $0 < \delta < 1$, and $\frac{\delta}{1-\delta} \le \frac{1}{12C_1^2}$.

Since the largest eigenvalue of $\mathbf{Rc}_{g(t)}$ is $\frac{1}{2}\left[\lambda_2\left(\mathbf{M}(t)\right)+\lambda_3\left(\mathbf{M}(t)\right)\right] \le \lambda_3\left(\mathbf{M}(t)\right)$; meanwhile, the eigenvalue of $\frac{1}{3}R_{g(t)}g(t)$ is

$$\frac{1}{3}\left[\lambda_1(\mathbf{M}(t))+\lambda_2(\mathbf{M}(t))+\lambda_3(\mathbf{M}(t))\right]\ge\lambda_1(\mathbf{M}(t)).$$

Thus,

$$\left|\mathrm{Rc}_{g(t)}-\frac{1}{3}R_{g(t)}g(t)\right|_{g(t)}\le\lambda_3(\mathbf{M}(t))-\lambda_1(\mathbf{M}(t)).$$

Corollary 4.6

Suppose $\mathcal{M}^3$ is a closed 3-manifold and g_0 has positive Ricci curvature. There exist constants $C>0$ and $\delta\in\left(0,\frac{1}{4}\right]$ such that

$$\left|\mathrm{Rc}_{g(t)}-\frac{1}{3}R_{g(t)}g(t)\right|_{g(t)}\le CR_{g(t)}^{1-\delta} \tag{4.3.15}$$

holds for all time where the Ricci flow exists. ♡

Note 4.3

(1) The 2-tensor $\mathrm{Rc}^\circ_g:=\mathrm{Rc}_g-\frac{1}{3}R_g g$ is the trace-free part of the Ricci tensor Rc_g and

$$\left|\mathrm{Rc}_g-\frac{1}{3}R_g g\right|^2_g=|\mathrm{Rc}_g|^2_g-\frac{1}{3}R_g^2, \tag{4.3.16}$$

which vanishes everywhere exactly when g is Einstein. Indeed,

$$\mathrm{tr}_g\mathrm{Rc}^\circ_g=g^{ij}\left(R_{ij}-\frac{1}{3}R_g g_{ij}\right)=R_g-\frac{1}{3}R_g\cdot 3=R_g-R_g=0.$$

The formula (4.3.16) can be seen as follows:

$$\begin{aligned}
|\mathrm{Rc}^\circ_g|^2_g &= g^{ik}g^{j\ell}R^\circ_{ij}R^\circ_{k\ell} = g^{ik}g^{j\ell}\left(R_{ij}-\frac{1}{3}R_g g_{ij}\right)\left(R_{k\ell}-\frac{1}{3}R_g g_{k\ell}\right)\\
&= g^{ik}g^{j\ell}\left(R_{ij}R_{k\ell}-\frac{1}{3}R_gR_{k\ell}g_{ij}-\frac{1}{3}R_gR_{ij}g_{k\ell}+\frac{1}{9}R_g^2g_{ij}g_{k\ell}\right)\\
&= |\mathrm{Rc}_g|^2_g-\frac{1}{3}R_g^2-\frac{1}{3}R_g^2+\frac{1}{9}R_g^2\cdot 3 = |\mathrm{Rc}_g|^2_g-\frac{1}{3}R_g^2.
\end{aligned}$$

(2) We also have

$$\left|\mathbf{Rm}_g-\frac{1}{3}R_g\mathrm{id}_{\mathcal{A}^2(\mathcal{M}^3)}\right|^2_g=4\left|\mathrm{Rc}_g-\frac{1}{3}R_g g\right|^2_g. \tag{4.3.17}$$

Since

$$\mathbf{Rm}_g=\begin{pmatrix}\lambda_1&0&0\\0&\lambda_2&0\\0&0&\lambda_3\end{pmatrix},\quad R_g=\lambda_1+\lambda_2+\lambda_3,$$

we have

$$\left|\mathbf{Rm}_g-\frac{1}{3}R_g\mathrm{id}_{\mathcal{A}^2(\mathcal{M}^3)}\right|^2_g=\frac{6}{9}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2-\lambda_1\lambda_2-\lambda_1\lambda_3-\lambda_2\lambda_3\right).$$

Meanwhile,

$$\mathbf{Rc}_g=\begin{pmatrix}\frac{\lambda_2+\lambda_3}{2}&0&0\\0&\frac{\lambda_1+\lambda_3}{2}&0\\0&0&\frac{\lambda_1+\lambda_2}{2}\end{pmatrix}$$

and

$$\left|\mathrm{Rc}_g - \frac{1}{3}R_g g\right|_g^2 = \left(\frac{\lambda_2+\lambda_3}{2} - \frac{\lambda_1+\lambda_2+\lambda_3}{3}\right)^2 + \left(\frac{\lambda_1+\lambda_3}{2} - \frac{\lambda_1+\lambda_2+\lambda_3}{3}\right)^2 + \left(\frac{\lambda_1+\lambda_2}{2} - \frac{\lambda_1+\lambda_2+\lambda_3}{3}\right)^2 = \frac{1}{6}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2-\lambda_1\lambda_2-\lambda_1\lambda_3-\lambda_2\lambda_3\right).$$

So we verified (4.3.17). ♣

Let $[0, T)$ denote the maximum time interval of existence of our solution. Recall that

$$R_{\min}(t) \geq \frac{1}{R_{\min}(0)^{-1} - \frac{2}{3}t}.$$

If we assume that the initial metric g_0 has positive Ricci curvature, then $R_{\min}(0) > 0$ and we conclude that $T \leq \frac{3}{2}R_{\min}(0)^{-1} < \infty$. In later, we shall prove that

$$\sup_{M\times[0,T)} \left|\mathrm{Rm}_{g(t)}\right|_{g(t)} = \infty. \tag{4.3.18}$$

Intuitively speaking, we are in good shape now. Since the Ricci curvature is positive under the Ricci flow, the metric is shrinking: $\partial_t g(t) = -2\mathrm{Ric}_{g(t)} < 0$. If we can show an appropriate gradient estimate for the scalar curvature, then we could conclude that $\lim_{t\to T} R_{\min}(t) = \infty$. Assuming this, we then would have

$$\left|\frac{1}{R_{g(t)}}\mathrm{Ric}_{g(t)} - \frac{1}{3}g(t)\right|_{g(t)} \leq CR_{g(t)}^{-\delta},$$

which tends to 0 as $t \to T$ uniformly in x. To finish the proof of **Theorem 4.1**, we need to further show that the solution $\overline{g}(\overline{t})$ to the normalized Ricci flow exists for all time and the scalar invariant quantity $|\frac{1}{R_{\overline{g}(\overline{t})}}\mathrm{Ric}_{\overline{g}(\overline{t})} - \frac{1}{3}\overline{g}(\overline{t})|_{\overline{g}(\overline{t})}$ decays exponentially to zero as $\overline{t} \to \infty$. After that, we shall show that under the normalized Ricci flow the curvature tends to a constant. Finally we prove the long time existence and exponential convergence of the solution to a constant sectional curvature metric.

4.4 Gradient bounds for the scalar curvature

Introduction

- *Gradient estimate I*
- *Gradient estimate II*
- *Global scalar curvature pinching*
- *Global sectional curvature pinching*

Proposition 4.1. (Scalar curvature gradient estimate)

Let $(\mathcal{M}^3, g)$ be a closed 3-manifold with positive Ricci curvature. For any $\varepsilon > 0$, there exists $C(\varepsilon)$ depending only on ε and g such that

$$\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^2 \leq \varepsilon R_{g(t)}^3 + C(\varepsilon)R_{g(t)}$$

as long as the solution to the Ricci flow with initial metric g exists. ♡

4.4.1 Gradient estimate I

Let $(\mathcal{M}^m, g)$ be an m-dimensional Riemannian manifold. We decompose $\nabla_i R_{jk}$ into

$$\nabla_i R_{jk} := E_{ijk} + F_{ijk} \tag{4.4.1}$$

where

$$\begin{aligned} E_{ijk} &:= \frac{m-2}{2(m-1)(m+2)}\left(\nabla_j R_g \cdot g_{ik} + \nabla_k R_g \cdot g_{ij}\right) \\ &\quad + \frac{m}{(m-1)(m+2)}\nabla_i R_g \cdot g_{jk}. \end{aligned} \tag{4.4.2}$$

Lemma 4.2

For any $m \geq 3$, one has

$$\begin{aligned} \langle E_{ijk}, F_{ijk}\rangle_g &= 0, \\ |E_{ijk}|_g^2 &= \frac{3m-2}{2(m-1)(m+2)}|\nabla_g R_g|_g^2, \\ |\nabla_g \mathrm{Ric}_g|_g^2 &\geq \frac{3m-2}{2(m-1)(m+2)}|\nabla_g R_g|_g^2, \\ |\nabla_g \mathrm{Ric}_g|_g^2 - \frac{1}{m}|\nabla_g R_g|_g^2 &\geq \frac{(m-2)^2}{2m(m-1)(m+2)}|\nabla_g R_g|_g^2. \end{aligned}$$

♡

Proof. Calculate

$$\begin{aligned} \langle E_{ijk}, \nabla_i R_{jk}\rangle_g &= g^{ip}g^{jq}g^{ks}E_{ijk}\nabla_p R_{qs} \\ &= \frac{m-2}{2(m-1)(m+2)}\left(\frac{1}{2}|\nabla_g R_g|_g^2 + \frac{1}{2}|\nabla_g R_g|_g^2\right) + \frac{m}{(m-1)(m+2)}|\nabla_g R_g|_g^2 \\ &= \frac{3m-2}{2(m-1)(m+2)}|\nabla_g R_g|_g^2, \end{aligned}$$

and

$$\begin{aligned} |E_{ijk}|_g^2 &= \left(\frac{m-2}{2(m-1)(m+2)}\right)^2\left(2m|\nabla_g R_g|_g^2 + 2|\nabla_g R_g|_g^2\right) \\ &\quad + \frac{m^2}{(m-1)^2(m+2)^2}m|\nabla_g R_g|_g^2 + \frac{m(m-2)}{(m-1)^2(m+2)^2}2|\nabla_g R_g|_g^2 \\ &= \frac{(m-2)^2(m+1) + 2m^3 + 4m(m-2)}{2(m-1)^2(n+2)^2}|\nabla_g R_g|_g^2 \\ &= \frac{(3m-2)(m-1)(m+2)}{2(m-1)^2(m+2)^2}|\nabla_g R_g|_g^2 = \frac{3m-2}{2(m-1)(m+2)}|\nabla_g R_g|_g^2. \end{aligned}$$

The rest inequalities follows immediately. □

When $m = 3$, **Lemma 4.2** gives

$$|\nabla_g \mathrm{Ric}_g|_g^2 \geq \frac{7}{20}|\nabla_g R_g|_g^2.$$

If we use the inequality

$$|\nabla_g \mathrm{Ric}_g|_g^2 - \frac{1}{3}|\nabla_g R_g|_g^2 = \left|\nabla_i R_{jk} - \frac{1}{3}\nabla_i R_g g_{jk}\right|_g^2 \geq \frac{1}{3}\left|\mathrm{div}_g\left(\mathrm{Rc}_g - \frac{1}{3}R_g g\right)\right|_g^2$$

and the contracted second Bianchi identity, we have a worse constant

$$|\nabla_g \mathrm{Ric}_g|_g^2 \geq \frac{37}{108}|\nabla_g R_g|_g^2.$$

Lemma 4.3

Let $\Box_{g(t)} := \partial_t - \Delta_{g(t)}$. If f and h are functions of space and time and if $p, q \in \mathbf{R}$, then

$$\begin{aligned}\Box_{g(t)}\left(\frac{f^p}{h^q}\right) &= p\frac{f^{p-1}}{h^q}\Box_{g(t)}f - q\frac{f^p}{h^{q+1}}\Box_{g(t)}h - p(p-1)\frac{f^{p-2}}{h^q}\left|\nabla_{g(t)}f\right|^2_{g(t)}\\ &\quad - q(q+1)\frac{f^p}{h^{q+2}}\left|\nabla_{g(t)}h\right|^2_{g(t)} + 2pq\frac{f^{p-1}}{h^{q+1}}\left\langle\nabla_{g(t)}f, \nabla_{g(t)}h\right\rangle_{g(t)}\end{aligned} \tag{4.4.3}$$

In particular, taking $p = q = 1$, we obtain

$$\Box_{g(t)}\left(\frac{f}{h}\right) = \frac{1}{h}\Box_{g(t)}f - \frac{f}{h^2}\Box_{g(t)}h + \frac{2}{h}\left\langle\nabla_{g(t)}h, \nabla_{g(t)}\left(\frac{f}{h}\right)\right\rangle_{g(t)}. \tag{4.4.4}$$

♡

Proof. Calculate

$$\begin{aligned}\partial_t\left(\frac{f^p}{h^q}\right) &= \frac{pf^{p-1}\cdot\partial_t f\cdot h^q - f^p\cdot qh^{q-1}\cdot\partial_t h}{h^{2q}} = p\frac{f^{p-1}}{h^q}\partial_t f - q\frac{f^p}{h^{q+1}}\partial_t h,\\ \Delta_{g(t)}\left(\frac{f^p}{h^q}\right) &= \nabla^i\left(p\frac{f^{p-1}}{h^q}\nabla_i f - q\frac{f^p}{h^{q+1}}\nabla_i h\right)\\ &= \frac{p(p-1)f^{p-2}h^q\nabla^i f - pqf^{p-1}h^{q-1}\nabla^i h}{h^{2q}}\nabla_i f + p\frac{f^{p-1}}{h^q}\Delta_{g(t)}f\\ &\quad - \left(\frac{pqf^{p-1}h^{q+1}\nabla^i f - q(q+1)f^p h^q\nabla^i h}{h^{2q+2}}\nabla_i h + q\frac{f^p}{h^{q+1}}\Delta_{g(t)}h\right)\\ &= p(p-1)\frac{f^{p-2}}{h^q}\left|\nabla_{g(t)}f\right|^2_{g(t)} - 2pq\frac{f^{p-1}}{h^{q+1}}\left\langle\nabla_{g(t)}f, \nabla_{g(t)}h\right\rangle_{g(t)}\\ &\quad + p\frac{f^{p-1}}{h^q}\Delta_{g(t)}f + q(q+1)\frac{f^p}{h^{q+2}}\left|\nabla_{g(t)}h\right|^2_{g(t)} - q\frac{f^p}{h^{q+1}}\Delta_{g(t)}h.\end{aligned}$$

Combining those equations gives (4.4.3). □

As a consequence, we have

$$\begin{aligned}\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}}\right) &= \frac{1}{R_{g(t)}}\Box_{g(t)}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\\ &\quad - \frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R^2_{g(t)}}\Box_{g(t)}R_{g(t)} - 2\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R^3_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}.\end{aligned}$$

Since $\Box_{g(t)}R_{g(t)} = 2|\mathrm{Ric}_{g(t)}|^2_{g(t)}$, we compute the first term $\Box_{g(t)}|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}$:

$$\begin{aligned}\partial_t\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} &= -\partial_t g_{ij}\cdot\nabla^i R_{g(t)}\nabla^j R_{g(t)} + 2g^{ij}\nabla_j R_{g(t)}\nabla_i\partial_t R_{g(t)}\\ &= 2R_{ij}\nabla^i R_{g(t)}\nabla^j R_{g(t)} + 2\nabla^i R_{g(t)}\cdot\nabla_i\left(\Delta_{g(t)}R_{g(t)} + 2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}\right)\\ &= 2R_{ij}\nabla^i R_{g(t)}\nabla^j R_{g(t)} + 2\nabla^i R_{g(t)}\nabla_i\Delta_{g(t)}R_{g(t)} + 4\nabla^i R_{g(t)}\cdot\nabla_i\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)},\end{aligned}$$

and

$$\begin{aligned}\Delta_{g(t)}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} &= \Delta_{g(t)}\left(g^{ij}\nabla_i R_{g(t)}\nabla_j R_{g(t)}\right) = 2g^{ij}g^{k\ell}\nabla_k\left(\nabla_\ell\nabla_i R_{g(t)}\cdot\nabla_j R_{g(t)}\right)\\ &= 2g^{ij}g^{k\ell}\left(\nabla_k\nabla_i\nabla_\ell R_{g(t)}\cdot\nabla_j R_{g(t)}\right) + 2\left|\nabla^2_{g(t)}R_{g(t)}\right|^2_{g(t)}\end{aligned}$$

$$= 2g^{ij}g^{k\ell}\left(\nabla_i\nabla_k\nabla_\ell R_{g(t)} - R_{ki\ell p}\nabla^p R_{g(t)}\right) + 2\left|\nabla^2_{g(t)}R_{g(t)}\right|^2_{g(t)}$$

$$= 2\nabla_i\Delta_{g(t)}R_{g(t)}\cdot\nabla^i R_{g(t)} + 2R_{ip}\nabla^p R_{g(t)}\cdot\nabla^i R_{g(t)} + 2\left|\nabla^2_{g(t)}R_{g(t)}\right|^2_{g(t)}.$$

Hence

$$\Box_{g(t)}\left|\nabla_{g(t)}R_{g(t)}\right|^2 = 4\left\langle\nabla_{g(t)}R_{g(t)}, \nabla_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}\right\rangle_{g(t)} - 2\left|\nabla^2_{g(t)}R_{g(t)}\right|^2_{g(t)}. \tag{4.4.5}$$

Using (4.4.5) we obtain

$$\begin{aligned}&\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}}\right) \qquad (4.4.6)\\ = & -\frac{2}{R^3_{g(t)}}\left|R_{g(t)}\nabla^2_{g(t)}R_{g(t)} - \nabla_{g(t)}R_{g(t)}\otimes\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\\ & + \frac{4}{R_{g(t)}}\left\langle\nabla_{g(t)}R_{g(t)}, \nabla_{g(t)}|\mathrm{Ric}_{g(t)}|^2_{g(t)}\right\rangle_{g(t)} - 2\frac{|\mathrm{Ric}_{g(t)}|^2_{g(t)}}{R^2_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}.\end{aligned}$$

Proof of **Proposition 4.1**. The assumption that $\mathrm{Ric}_{g(0)} > 0$ yields that $\mathrm{Ric}_{g(t)} > 0$ for $t \in [0, T)$ where T is the maximal time. Hence $|\mathrm{Ric}_{g(t)}|_{g(t)} \leq R_{g(t)}$. On the other hand, from

$$\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} \geq \frac{1}{3}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)},$$

we obtain

$$\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)} \leq \sqrt{3}\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)} \leq 2\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)}.$$

Therefore, (4.4.6) becomes

$$\begin{aligned}&\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}}\right)\\ \leq & \frac{4}{R_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}\left|\nabla_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}\right|_{g(t)} - 2\frac{|\mathrm{Ric}_{g(t)}|^2_{g(t)}}{R^2_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\\ \leq & \frac{4}{R_{g(t)}}\cdot 2\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)}\cdot\left|2\mathrm{Ric}_{g(t)}\right|\cdot\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)}\\ & - 2\frac{|\mathrm{Ric}_{g(t)}|^2_{g(t)}}{R^2_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\\ \leq & 16\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - 2\frac{|\mathrm{Ric}_{g(t)}|^2_{g(t)}}{R^2_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}.\end{aligned}$$

It is clear that

$$\Box_{g(t)}R^2_{g(t)} = -2\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} + 4R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}. \tag{4.4.7}$$

Using (4.4.7) yields for any $\varepsilon > 0$

$$\begin{aligned}\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} - \varepsilon R^2_{g(t)}\right) &\leq 16\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - 2\frac{|\mathrm{Ric}_{g(t)}|^2_{g(t)}}{R^2_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\\ &\quad - \varepsilon\left(-2\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} + 4R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}\right)\\ &\leq 16\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{2}{3}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} + 2\varepsilon\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} - \frac{4}{3}\varepsilon R^3_{g(t)}\end{aligned}$$

$$= 16\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} + 2\left(\varepsilon - \frac{1}{3}\right)\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} - \frac{4}{3}\varepsilon R^3_{g(t)}.$$

If $0 < \varepsilon \le \frac{1}{3}$, then the above inequality reduces to

$$\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} - \varepsilon R^2_{g(t)}\right) \le 16\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{4}{3}\varepsilon R^3_{g(t)};$$

if $\varepsilon > \frac{1}{3}$, then

$$\begin{aligned}\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} - \varepsilon R^2_{g(t)}\right) &\le \left(16 + 6\left(\varepsilon - \frac{1}{3}\right)\right)\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{4}{3}\varepsilon R^3_{g(t)} \\ &= (6\varepsilon + 14)\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{4}{3}\varepsilon R^3_{g(t)}.\end{aligned}$$

In both cases, we derive

$$\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} - \varepsilon R^2_{g(t)}\right) \le \left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{4}{3}\varepsilon R^3_{g(t)}. \tag{4.4.8}$$

To deal with the "bad" term $|\nabla_{g(t)}\mathrm{Ric}_{g(t)}|^2_{g(t)}$ we consider the evolution equation for $|\mathrm{Rc}_{g(t)}|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}$. Calculate using **Lemma 4.1**,

$$\begin{aligned}\partial_t\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} &= \partial_t\left(g^{ik}g^{j\ell}R_{ij}R_{k\ell}\right) = -2\partial_t g_{ik}R^{ij}R^k{}_j + 2R^{ij}\partial_t R_{ij} \\ &= 4R_{ik}R^i{}_\ell R^{k\ell} + 2R^{ij}\left(\Delta_{g(t)}R_{ij} + 3R_{g(t)}R_{ij} - 6R_{ip}R_j{}^p\right) + 2R\left(2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - R^2_{g(t)}\right) \\ &= 2R^{ij}\Delta_{g(t)}R_{ij} - 2R^3_{g(t)} + 10R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - 8\mathrm{tr}_{g(t)}\mathrm{Ric}^3_{g(t)}\end{aligned}$$

where $\mathrm{tr}_{g(t)}\mathrm{Ric}^3_{g(t)} := g^{ip}g^{kr}g^{\ell s}R_{ik}R_{p\ell}R_{rs}$. Meanwhile,

$$\Delta_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} = g^{ik}g^{j\ell}g^{pq}\nabla_p\nabla_q\left(R_{ij}R_{k\ell}\right) = 2\Delta_{g(t)}R_{ij}\cdot R^{ij} + 2\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)}.$$

Hence

$$\begin{aligned}\Box_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} &= -2\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - 2R^3_{g(t)} \\ &\quad + 10R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - 8\mathrm{tr}_{g(t)}\mathrm{Ric}^3_{g(t)},\end{aligned} \tag{4.4.9}$$

and

$$\begin{aligned}\Box_{g(t)}\left(\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}\right) &= -2\left(\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{1}{3}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\right) \\ &\quad - 2R^3_{g(t)} + \frac{26}{3}R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - 8\mathrm{tr}_{g(t)}\mathrm{Ric}^3_{g(t)}.\end{aligned} \tag{4.4.10}$$

From (4.4.2) for $m = 3$ and the assumption that $\mathrm{Ric}_{g(0)} > 0$, we have

$$\begin{aligned}\Box_{g(t)}\left(\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}\right) &\le -\frac{2}{21}\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} \\ &\quad + 4R_{g(t)}\left(\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}\right).\end{aligned} \tag{4.4.11}$$

For any positive constant a we obtain

$$\begin{aligned}&\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} - \varepsilon R^2_{g(t)} + a\left(|\mathrm{Ric}|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}\right)\right) \\ \le\ &\left(6\varepsilon + 14 - \frac{2}{21}a\right)\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} + R^{3-2\delta}\left(4aC - \frac{4}{3}\varepsilon R^{2\delta}_{g(t)}\right)\end{aligned}$$

where we use the estimate (4.4.8). If we pick $a = 21(3\varepsilon + 7)$, then

$$\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} - \varepsilon R^2_{g(t)} + a\,|\mathrm{Ric}|^2_{g(t)} - \frac{aR^2_{g(t)}}{3}\right) \le R^{3-2\delta}_{g(t)}\left[84(3\varepsilon+7)C - \frac{4\varepsilon R^{2\delta}_{g(t)}}{3}\right]$$

When $84(3\varepsilon+7)C - \frac{4}{3}\varepsilon R^{2\delta}_{g(t)} < 0$, the right side is negative; when $84(3\varepsilon+7)C - \frac{4}{3}\varepsilon R^{2\delta}_{g(t)} \ge 0$, we have $R^{2\delta}_{g(t)} \le 63(3\varepsilon+7)C/\varepsilon$ so that the right side is bounded by a constant depending on C and ε. In both case, it follows that

$$\Box_{g(t)}\left(\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} - \varepsilon R^2_{g(t)} + a\left(|\mathrm{Ric}|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}\right)\right) \le C(\varepsilon).$$

By the maximum principle, we deduce that

$$\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} \le C(\varepsilon) + \varepsilon R^2_{g(t)}$$

as long as the solution exists.

4.4.2 Gradient estimate II

The paths toward obtaining various estimates are often/usually not unique.

Proposition 4.2

Let $(\mathcal{M}^3, g)$ be a closed 3-manifold with positive Ricci curvature and $g(t)$ is the solution to the Ricci flow with initial metric g. We have the following variant of the gradient of scalar curvature estimate. There exists a constant $\delta \in (0, 1/4]$ depending only on $g(0)$ such that for any $\beta > 0$

$$\frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R^3_{g(t)}} \le \beta R^{-\delta}_{g(t)} + C(\beta)R^{-2}_{g(t)}, \tag{4.4.12}$$

where $C(\beta) < +\infty$ depends only on β and $g(0)$. ♡

Proof. Let

$$V := \frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} + a\left(\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}\right), \tag{4.4.13}$$

where a is a positive constant. By previous calculus we get

$$\begin{aligned}
\Box_{g(t)}V \;&=\; -\frac{2}{R^3_{g(t)}}\left|R_{g(t)}\nabla^2_{g(t)}R_{g(t)} - \nabla_{g(t)}R_{g(t)}\otimes\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} \\
&+ \frac{4}{R_{g(t)}}\left\langle\nabla_{g(t)}R_{g(t)}, \nabla_{g(t)}\left|\mathrm{Rc}_{g(t)}\right|^2\right\rangle_{g(t)} - 2\frac{|\mathrm{Ric}_{g(t)}|^2_{g(t)}}{R^2_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)} \\
&- 2a\left(|\nabla_{g(t)}\mathrm{Ric}_{g(t)}|^2_{g(t)} - \frac{1}{3}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\right) \\
&+ a\left(-2R^3_{g(t)} + \frac{26}{3}R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - 8\mathrm{tr}_{g(t)}\mathrm{Ric}^3_{g(t)}\right) \\
\le\;& \frac{8|\mathrm{Ric}_{g(t)}|_{g(t)}}{R_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)} - \frac{2a}{21}\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)} \\
&+ a\left(-2R^3_{g(t)} + \frac{26}{3}R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - 8\mathrm{tr}_{g(t)}\mathrm{Ric}^3_{g(t)}\right).
\end{aligned}$$

Since

$$-2R_{g(t)}^3+\frac{26}{3}R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^2-8\mathrm{tr}_{g(t)}\mathrm{Ric}_{g(t)}^3\leq\frac{50}{3}R_{g(t)}\left(\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^2-\frac{1}{3}R_{g(t)}^2\right),$$

and $\frac{1}{3}\leq\frac{|\mathrm{Rc}_{g(t)}|_{g(t)}}{R_{g(t)}}\leq 1$, $|\nabla_{g(t)}R_{g(t)}|_{g(t)}\leq\sqrt{3}|\nabla_{g(t)}\mathrm{Rc}_{g(t)}|_{g(t)}$, and the pinching estimate $|\mathrm{Rc}_{g(t)}-\frac{1}{3}R_{g(t)}g(t)|_{g(t)}\leq CR_{g(t)}^{1-\delta}$ for some $\delta\in\left(0,\frac{1}{4}\right]$, it follows that

$$\Box_{g(t)}V\leq\left(8\sqrt{3}-\frac{2a}{21\sqrt{3}}\right)|\nabla_{g(t)}\mathrm{Ric}_{g(t)}|_{g(t)}^2+\frac{50a}{3}CR_{g(t)}^{3-2\delta}.$$

We choose a satisfying

$$8\sqrt{3}-\frac{2a}{21\sqrt{3}}\leq-1;$$

thus,

$$a\geq\frac{21\sqrt{3}}{2}\left(8\sqrt{3}+1\right).$$

For example, we may choose $a=\frac{37}{2}\left(8\sqrt{3}+1\right)$, and hence

$$\Box_{g(t)}V\leq-\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)}^2+CR_{g(t)}^{3-2\delta}. \tag{4.4.14}$$

From $\Box_{g(t)}R_{g(t)}=2|\mathrm{Rc}_{g(t)}|_{g(t)}^2$, we have

$$\begin{aligned}\partial_t\left(R_{g(t)}^{2-\delta}\right)&=(2-\delta)R_{g(t)}^{1-\delta}\partial_tR_{g(t)}\\&=(2-\delta)R_{g(t)}^{1-\delta}\left(\Delta_{g(t)}R_{g(t)}+2\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^2\right),\\\Delta_{g(t)}\left(R_{g(t)}^{2-\delta}\right)&=\nabla^i\left((2-\delta)R_{g(t)}^{1-\delta}\nabla_iR_{g(t)}\right)\\&=(2-\delta)\left((1-\delta)R_{g(t)}^{-\delta}\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^2+R_{g(t)}^{1-\delta}\Delta_{g(t)}R_{g(t)}\right)\\&=(2-\delta)R_{g(t)}^{1-\delta}\Delta_{g(t)}R_{g(t)}+(2-\delta)(1-\delta)R_{g(t)}^{-\delta}\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^2.\end{aligned}$$

Therefore

$$\Box_{g(t)}R_{g(t)}^{2-\delta}=-(1-\delta)(2-\delta)R_{g(t)}^{-\delta}\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^2+2(2-\delta)R^{1-\delta}\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^2.$$

Now, the evolution inequality of $V-R_{g(t)}^{2-\delta}$ is

$$\begin{aligned}\Box_{g(t)}\left(V-\beta R_{g(t)}^{2-\delta}\right)&\leq-\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)}^2\\+&\beta(1-\delta)(2-\delta)R^{-\delta}\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^2+CR_{g(t)}^{3-2\delta}-2\beta(2-\delta)R_{g(t)}^{1-\delta}\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^2.\end{aligned}$$

Since $R_{g(0)}>0$ and $\mathcal{M}^3$ is compact, we have $R_{g(0)}\geq c:=\min_{\mathcal{M}^3}R_{g(0)}>0$ and hence $R_{g(t)}\geq c$. Calculate

$$\begin{aligned}&-\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)}^2+\beta(1-\delta)(2-\delta)\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^2\\\leq&\left[-\frac{1}{3}+\beta(1-\delta)(2-\delta)c^{-\delta}\right]\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^2.\end{aligned}$$

If $\beta\in[0,\beta_0]$, where $\beta_0:=\frac{c^\delta}{3(1-\delta)(2-\delta)}$, then

$$-\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|_{g(t)}^2+\beta(1-\delta)(2-\delta)\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^2\leq 0$$

On the other hand,

$$
\begin{aligned}
&CR_{g(t)}^{3-2\delta} - 2\beta(2-\delta)R_{g(t)}^{1-\delta}\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2} \\
\leq\ & CR_{g(t)}^{3-2\delta} - \frac{2(2-\delta)}{9}\beta R_{g(t)}^{3-\delta} \ = \ R_{g(t)}^{3-2\delta}\left[C - \frac{2}{9}(2-\delta)\beta R_{g(t)}^{\delta}\right].
\end{aligned}
$$

If $R_{g(t)} > (\frac{9C}{2\beta(2-\delta)})^{1/\delta}$, then the right side is less than zero; if $R_{g(t)} \leq \frac{9C}{2\beta(2-\delta)})^{1/\delta}$, then the right side is less than $C\,[9C/2\beta(2-\delta)]^{\frac{3-2\delta}{\delta}}$. Hence in both cases, we have

$$
CR_{g(t)}^{3-2\delta} - 2\beta(2-\delta)R_{g(t)}^{1-\delta}\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2} \leq C'(\beta)
$$

for some constant $C'(\beta)$ depending only on β and $g(0)$. Combining the above two estimates yields

$$
\Box_{g(t)}\left(V - \beta R_{g(t)}^{2-\delta}\right) \leq C'(\beta)
$$

for any $\beta \in [0, \beta_0]$. By the maximum principle, we conclude that

$$
V - \beta R_{g(t)}^{2-\delta} \leq C + C'(\beta)t \leq C(\beta) \tag{4.4.15}
$$

holds for any $\beta \in [0, \beta_0]$, since $t \leq T_{\max} \leq \frac{1}{c}$. However, when $\beta \geq \beta_0$, from

$$
V - \beta R_{g(t)}^{2-\delta} = V - \beta_0 R_{g(t)}^{2-\delta} + (\beta_0 - \beta)R_{(t)}^{2-\delta}
$$

the inequality (4.4.15) is valid. Finally by the definition of V we prove (4.4.12). □

Lemma 4.4

The function

$$
F := \frac{\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2} - \frac{1}{3}R_{g(t)}^{2}}{R_{g(t)}^{2-\delta}} \tag{4.4.16}
$$

satisfies

$$
\begin{aligned}
\Box_{g(t)}F \ &= \ \frac{2(1-\delta)}{R_{g(t)}}\left\langle \nabla_{g(t)}R_{g(t)}, \nabla_{g(t)}F\right\rangle_{g(t)} \\
&\quad - \frac{2}{R_{g(t)}^{4-\delta}}\left|R_{g(t)}\nabla_i R_{jk} - \nabla_i R_{g(t)}\cdot R_{jk}\right|_{g(t)}^{2} \\
&\quad - \frac{\delta(1-\delta)}{R_{g(t)}^{4-\delta}}\left(\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2} - \frac{1}{3}R_{g(t)}^{2}\right)\left|\nabla_{g(t)}R_{g(t)}\right|_{g(t)}^{2} \\
&\quad + \frac{2}{R^{3-\delta}}\left[\delta\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2}\left(\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2} - \frac{1}{3}R_{g(t)}^{2}\right) - J\right]
\end{aligned} \tag{4.4.17}
$$

where

$$
J := 2\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{4} + R_{g(t)}\left[R_{g(t)}^{3} - 5R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2} + 4\mathrm{tr}_{g(t)}\mathrm{Ric}_{g(t)}^{3}\right]. \tag{4.4.18}
$$

♡

Proof. Using (4.4.3) yields

$$
\begin{aligned}
\Box_{g(t)}\left(\frac{\left|\mathrm{Rc}_{g(t)}\right|_{g(t)}^{2} - \frac{1}{3}R_{g(t)}^{2}}{R_{g(t)}^{2-\delta}}\right) \ &= \ \frac{1}{R_{g(t)}^{2-\delta}}\Box_{g(t)}\left(\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2} - \frac{1}{3}R_{g(t)}^{2}\right) \\
&\quad - (2-\delta)\frac{\left|\mathrm{Ric}_{g(t)}\right|_{g(t)}^{2} - \frac{1}{3}R_{g(t)}^{2}}{R^{3-\delta}}\Box_{g(t)}R_{g(t)}
\end{aligned}
$$

$$-\frac{2(2-\delta)}{R^{3-\delta}_{g(t)}}\left\langle\nabla_{g(t)}\left(\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}-\frac{1}{3}R^2_{g(t)}\right),\nabla_{g(t)}R_{g(t)}\right\rangle_{g(t)}.$$

Applying the evolution equation (4.4.10) to above, we obtain

$$\Box_{g(t)}F=A+B$$

where

$$\begin{aligned}A\ &:=\ -\frac{2}{R^{2-\delta}_{g(t)}}\left(\left|\nabla_{g(t)}\mathrm{Ric}_{g(t)}\right|^2_{g(t)}-\frac{1}{3}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\right)\\&\quad-(2-\delta)(3-\delta)\frac{\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}-\frac{1}{3}R^2_{g(t)}}{R^{4-\delta}_{g(t)}}\left|\nabla_{g(t)}R_{g(t)}\right|^2_{g(t)}\\&\quad+2(2-\delta)\frac{1}{R^{3-\delta}_{g(t)}}\left\langle\nabla_{g(t)}\left(\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}-\frac{1}{3}R^2_{g(t)}\right),\nabla_{g(t)}R_{g(t)}\right\rangle_{g(t)}\end{aligned}$$

are the gradient terms and

$$\begin{aligned}B\ &:=\ \frac{1}{R^{2-\delta}_{g(t)}}\left(-2R^3_{g(t)}+\frac{26}{3}R_{g(t)}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}-8\mathrm{tr}_{g(t)}\mathrm{Ric}^3_{g(t)}\right)\\&\quad-2(2-\delta)\frac{\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}-\frac{1}{3}R^2_{g(t)}}{R^{3-\delta}_{g(t)}}\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)}\end{aligned}$$

are the curvature terms. Simplifying A and B, we prove (4.4.17). □

Corollary 4.7

Let $(\mathcal{M}^3,g)$ be a closed 3-manifold with positive Ricci curvature and $g(t)$ is the solution to the Ricci flow with initial metric g. Then there exists constants $C=C(g(0))<+\infty$ and $\delta\in\left(0,\frac{2}{9}\right]$ such that

$$\left|\mathrm{Ric}_{g(t)}-\frac{1}{3}R_{g(t)}g(t)\right|_{g(t)}\le CR^{1-\delta}_{g(t)}.$$

♡

Proof. Since $\mathcal{M}^3$ is closed, we can always find a positive constant $\varepsilon\in\left(0,\frac{1}{3}\right]$ so that $\mathrm{Ric}_{g(0)}\ge\varepsilon R_{g(0)}g(0)$. (4.3.12) tells us that $\mathrm{Ric}_{g(t)}\ge\varepsilon R_{g(t)}g(t)$ for $t\in[0,T_{\max}]$. Note that

$$J\ge 2\varepsilon^2\left|\mathrm{Rc}_{g(t)}\right|^2_{g(t)}\left(\left|\mathrm{Rc}_{g(t)}\right|^2_{g(t)}-\frac{1}{3}R^2_{g(t)}\right).$$

If we choose $\delta\le 2\varepsilon^2\le\frac{2}{9}<\frac{1}{4}$, then from (4.4.17) we get

$$\Box_{g(t)}F\le\frac{2(1-\delta)}{F}\left\langle\nabla_{g(t)}R_{g(t)},\nabla_{g(t)}F\right\rangle_{g(t)}.$$

therefore the estimate immediately follows from the maximum principle. □

4.4.3 Global scalar curvature pinching

In this subsection we apply the gradient estimate and the Bonnet-Myers theorem to show that the global pinching of the scalar curvature tends to 1 as we approach the singularity time. For any solution $g(t)$ of the Ricci flow, we set

$$R_{\max}(t):=\max_{\mathcal{M}^3}R_{g(t)},\quad R_{\min}(t):=\min_{\mathcal{M}^3}R_{g(t)}.$$

Lemma 4.5. (Global scalar curvature pinching)

Let $(\mathcal{M}^3, g)$ be a closed 3-manifold with positive Ricci curvature. Suppose that $g(t)$, $t \in [0, T_{\max})$ is the solution to the Ricci flow with initial value g. Then we have

$$\lim_{t \to T_{\max}} \frac{R_{\min}(t)}{R_{\max}(t)} = 1. \tag{4.4.19}$$

In fact, there exist constants $C < +\infty$ and $\gamma > 0$ depending only on g_0 such that

$$1 \geq \frac{R_{\min}(t)}{R_{\max}(t)} \geq 1 - \frac{C}{R_{\max}^{\gamma}(t)}. \tag{4.4.20}$$

♡

Note 4.4

In later we will prove $\lim_{t \to T_{\max}} \max_{\mathcal{M}^3} |\mathrm{Rm}_{g(t)}|_{g(t)} = +\infty$. In our case $m = 3$, $\mathrm{Ric}_{g(t)} > 0$ since $\mathrm{Ric}_{g_0} > 0$, and $R_{g(t)} \geq |\mathrm{Ric}_{g(t)}|_{g(t)}$, we must have $\lim_{t \to T_{\max}} R_{\max}(t) = +\infty$. Together (4.4.20) we prove (4.4.19). Meanwhile,

$$\lim_{t \to T_{\max}} R_{\min}(t) = +\infty.$$

♣

Proof. By (4.4.12) and $\lim_{t \to T_{\max}} R_{\max}(t) = +\infty$, there exist constants $C < +\infty$ and $\delta > 0$ such that

$$\left|\nabla_{g(t)} R_{g(t)}\right|_{g(t)} \leq C R_{\max}(t)^{\frac{3}{2} - \delta}$$

on $\mathcal{M}^3$ for all $t \in [0, T_{\max})$. Given $t \in [0, T_{\max})$ there exists $x_t \in \mathcal{M}^3$ such that $R_{\max}(t) = R_{g(t)}(x_t)$. Given $\eta > 0$, to be chosen sufficiently small later, for any point

$$x \in B_{g(t)}(x_t, 1/\eta\sqrt{R_{\max}(t)})$$

we have $R_{\max}(t) - R_{g(t)}(x) \leq \frac{\max_{\mathcal{M}^3} |\nabla_{g(t)} R_{g(t)}|_{g(t)}}{\eta\sqrt{R_{\max}(t)}} \leq \frac{C}{\eta} R_{\max}(t)^{1-\delta}$. So that

$$R_{g(t)}(x) \geq R_{\max}(t)\left(1 - \frac{C}{\eta} R_{\max}(t)^{-\delta}\right) \tag{4.4.21}$$

for all $x \in B_{g(t)}(x_t, 1/\eta\sqrt{R_{\max}(t)})$. We claim that this ball is all of $\mathcal{M}^3$. Since

$$\lim_{t \to T_{\max}} R_{\max}(t) = +\infty,$$

by (4.4.21), there exists $\tau < T_{\max}$ such that for $t \in [\tau, T_{\max})$ we have

$$R_{g(t)}(x) \geq R_{\max}(t)(1 - \eta)$$

for all $x \in B_{g(t)}(x_t, 1/\eta\sqrt{R_{\max}(t)})$. Now the Bonnet-Myers theorem and the pinching estimate $\mathrm{Ric}_{g(t)} \geq \varepsilon R_{g(t)} g(t)$, where $\varepsilon > 0$, show that for $\eta > 0$ sufficiently small $\mathcal{M}^3 = B_{g(t)}(x_t, 1/\eta\sqrt{R_{\max}(t)})$. □

4.4.4 Global sectional curvature pinching

The Rauch-Klingenberg-Berger topological sphere theorem states that if $(\mathcal{M}^m, g)$ is a complete, simply-connected m-dimensional Riemanian manifold with $\frac{1}{4} < \mathrm{Sec}_g \leq 1$, then $\mathcal{M}^m$ is homeomorphic to the m-sphere. In particular, if $m = 3$, then $\mathcal{M}^m$ is diffeomorphic to the 3-sphere (since in dimension 3, the differential and topological categories are the same).

Lemma 4.6. (Global sectional curvature pinching)

Let $(\mathcal{M}^3, g)$ be a closed 3-manifold with positive Ricci curvature. Suppose that $g(t)$, $t \in [0, T_{\max})$, is the solution to the Ricci flow with initial value g. For every $\varepsilon \in (0,1)$, there exists $\tau(\varepsilon) < T_{\max}$ such that for all $t \in [\tau(\varepsilon), T_{\max})$ the sectional curvatures of $g(t)$ are positive and

$$\min_{\mathcal{M}^3} \lambda_1\left(\mathbf{Rm}_{g(t)}\right) \geq (1-\varepsilon) \max_{\mathcal{M}^3} \lambda_3\left(\mathbf{Rm}_{g(t)}\right).$$

♡

Note 4.5

The Rauch-Klingenberg-Berger topological sphere theorem and **Lemma 4.6** *implies that if $(\mathcal{N}^3, g)$ is a simply-connected Riemannian manifold satisfies*

$$\min_{\mathcal{N}^3} \lambda_1\left(\mathbf{Rm}_g\right) > \frac{1}{4} \max_{\mathcal{N}^3} \lambda_3\left(\mathbf{Rm}_g\right)$$

then $(\mathcal{N}^3, g)$ is diffeomorphic to the 3-sphere. Hence the universal cover $(\widetilde{\mathcal{M}}^3, \tilde{g}(t))$ of $(\mathcal{M}^3, g(t))$ is diffeomorphic to the 3-sphere for t sufficiently closed to $T_{\max}$.

♣

Proof. Recall $\lambda_1\left(\mathbf{Rm}_{g(t)}\right) \leq \lambda_2\left(\mathbf{Rm}_{g(t)}\right) \leq \lambda_3\left(\mathbf{Rm}_{g(t)}\right)$. By (4.3.15), there exist $C < +\infty$ and $\delta \in (0, 1/4]$ such that

$$\frac{\lambda_1\left(\mathbf{Rm}_{g(t)}\right)}{\lambda_3\left(\mathbf{Rm}_{g(t)}\right)} \geq 1 - C\frac{R_{g(t)}^{1-\delta}}{\lambda_3\left(\mathbf{Rm}_{g(t)}\right)} \geq 1 - 3CR_{\min}(t)^{-\delta}$$

on $x \in \mathcal{M}^3$ for $t \in [0, T_{\max})$. Hence for any $\varepsilon > 0$, there exists $\tau(\varepsilon) < T_{\max}$ such that for all $t \in [\tau(\varepsilon), T_{\max})$ we have

$$\lambda_1\left(\mathbf{Rm}_{g(t)}\right) \geq (1-\varepsilon)\lambda_3\left(\mathbf{Rm}_{g(t)}\right) \tag{4.4.22}$$

on $\mathcal{M}^3$. Hence, by (4.4.20),

$$\begin{aligned}
\lambda_1\left(\mathbf{Rm}_{g(t)}\right)(x) \ &\geq \ (1-\varepsilon)\lambda_3\left(\mathbf{Rm}_{g(t)}\right)(x) \ \geq \ \frac{1-\varepsilon}{3}R_{g(t)} \ \geq \ \frac{(1-\varepsilon)^2}{3}R_{\max}(t)(x) \\
\geq \ \frac{(1-\varepsilon)^2}{3}R_{g(t)}(y) \ &= \ \frac{(1-\varepsilon)^2}{3}\left[\lambda_1\left(\mathbf{Rm}_{g(t)}\right)(y) + \lambda_2\left(\mathbf{Rm}_{g(t)}\right)(y) + \lambda_3\left(\mathbf{Rm}_{g(t)}\right)(y)\right] \\
\geq \ \frac{(1-\varepsilon)^2}{3}&\left[\lambda_3\left(\mathbf{Rm}_{g(t)}\right)(y) + 2(1-\varepsilon)\lambda_3\left(\mathbf{Rm}_{g(t)}\right)(y)\right] \ \geq \ (1-\varepsilon)^3\lambda_3\left(\mathbf{Rm}_{g(t)}\right)(y).
\end{aligned}$$

Thus proves the lemma. □

4.5 Exponential convergence of the normalized Ricci flow

Introduction

- *Degree of tensor fields*
- *Maximum and average scalar curvatures under the Ricci flow*
- *Interpolation inequalities for tensor fields*
- *Higher derivatives of the curvature and the proof of* **Theorem 4.1**

Now we return back to the normalized Ricci flow

$$\partial_t g(t) = -2\mathrm{Ric}_{g(t)} + \frac{2}{3}\underline{R}_{g(t)} g(t)$$

where $\underline{R}_{g(t)} := \int_M R_{g(t)} dV_{g(t)} / \int_M dV_{g(t)}$ is the average of the scalar curvature. Since the volume is preserved by the normalized flow, it follows that

$$\underline{R}_{g(t)} = \frac{1}{V_g}\int_{\mathcal{M}^3} R_{g(t)} dV_{g(t)}$$

with

$$V_g = \int_M dV_g.$$

4.5.1 Degree of tensor fields

Let $(\mathcal{M}^m, g)$ be an m-dimensional Riemannian manifold. We say that a tensor quantity α_g depending on the metric g has **degree k in g** if

$$\alpha_{cg} = c^k \alpha_g \tag{4.5.1}$$

for any $c > 0$.

Note 4.6

$\mathrm{Rm}_g^{(3,1)}$ *has degree* 0, $\mathrm{Rm}_g^{(4,0)}$ *has degree* 1, Ric_g *has degree* 0, R_g *has degree* -1, *and* dV_g *has degree* $\frac{m}{2}$. ♣

Lemma 4.7

If an expression $X_{g(t)}$ formed algebraically from the metric and the Riemann curvature tensor by contractions has degree k and if under the Ricci flow

$$\Box_{g(t)} X_{g(t)} = Y_{g(t)}, \tag{4.5.2}$$

then the degree of $Y_{g(t)}$ is $k-1$ and the evolution under the normalized Ricci flow

$$\partial_{\bar{t}}\bar{g}(\bar{t}) = -2\mathrm{Ric}_{\bar{g}(\bar{t})} + \frac{2}{m}\underline{R}_{\bar{g}(\bar{t})}\bar{g}(\bar{t})$$

of $X_{\bar{g}(\bar{t})}$ is given by

$$\Box_{\bar{g}(\bar{t})} X_{\bar{g}(\bar{t})} = Y_{\bar{g}(\bar{t})} + k\frac{2}{m}\underline{R}_{\bar{g}(\bar{t})} X_{\bar{g}(\bar{t})}. \tag{4.5.3}$$

♡

Note 4.7

The above lemma also holds when the equalities in (4.5.2) and (4.5.3) are replaced by inequalities going the same way. ♣

Proof. It is clear that the degree of $Y_{g(t)}$ is $k-1$. Recall that $\bar{g}(\bar{t}) = c(t)g(t)$, where

$$c(t) = \exp\left(\frac{2}{m}\int_0^t \underline{R}_{g(\tau)}\,d\tau\right), \quad \bar{t}(t) = \int_0^t c(\tau)\,d\tau.$$

Then

$$\partial_{\bar{t}} X_{\bar{g}(\bar{t})} \;=\; \partial_t X_{\bar{g}(\bar{t})}\cdot\frac{dt}{d\bar{t}} \;=\; \partial_t\left([c(t)]^k X_{g(t)}\right)\frac{1}{c(t)}$$

$$= \frac{1}{c(t)}\left(k[c(t)]^{k-1}\frac{dc(t)}{dt}X_{g(t)} + [c(t)]^k\partial_t X_{g(t)}\right)$$

$$= [c(t)]^{k-1}\left(\Delta_{g(t)}X_{g(t)} + Y_{g(t)}\right) + k[c(t)]^{k-1}\frac{2}{m}\underline{R}_{g(t)}X_{g(t)}$$

$$= \Delta_{\bar{g}(\bar{t})}X_{\bar{g}(\bar{t})} + Y_{\bar{g}(\bar{t})} + k\frac{2}{m}\underline{R}_{\bar{g}(\bar{t})}X_{\bar{g}(\bar{t})}$$

where we use the fact that $\underline{R}_{\bar{g}(\bar{t})} = \frac{1}{c(t)}\underline{R}_{g(t)}$. □

4.5.2 Maximum and average scalar curvatures under the Ricci flow

Next we study the maximum and average scalar curvatures under the Ricci flow.

Lemma 4.8

Let $(\mathcal{M}^3, g)$ be a closed 3-manifold with positive Ricci curvature. Suppose that $g(t)$, $t \in [0, T_{\max})$, is the solution to the Ricci flow with initial value g. Then

$$R_{\max}(t) \geq \frac{1}{2(T_{\max} - t)} \tag{4.5.4}$$

and in particular

$$\int_0^{T_{\max}} R_{\max}(t)\,dt = +\infty. \tag{4.5.5}$$

♡

Proof. We have

$$R'_{\max}(t) \leq \sup_{\mathcal{M}^3} 2\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} \leq 2R_{\max}(t)^2.$$

Because $\lim_{t\to T_{\max}} R_{\max}(t) = \infty$, we have $R_{\max}(t) \geq \frac{1}{2(T_{\max}-t)}$. (4.5.5) is an immediate consequence of (4.5.4). □

Note 4.8

If $[0, \overline{T}_{\max})$ is the maximal time interval of existence of the normalized Ricci flow, then

$$\int_0^{\overline{T}_{\max}} \underline{R}_{\bar{g}(\bar{t})}\,d\bar{t} = +\infty.$$

Indeed,

$$\int_0^{\bar{t}_0} \underline{R}_{\bar{g}(\bar{t})}\,d\bar{t} = \int_0^{t_0} c(t)^{-1}\underline{R}_{g(t)}\frac{d\bar{t}}{dt}dt = \int_0^{t_0} \underline{R}_{g(t)}\,dt.$$

Lemma 4.5 *and* **Lemma 4.8** *imply*

$$\int_0^{T_{\max}} \underline{R}_{g(t)}\,dt = +\infty.$$

♣

Lemma 4.9. (Estimates for the normalized Ricci flow)

Let $(\mathcal{M}^3, g)$ be a closed 3-manifold with positive Ricci curvature. Suppose that $\bar{g}(\bar{t})$, $\bar{t} \in [0, \overline{T}_{\max})$, is the solution to the normalized Ricci flow with initial value g. Then there *exist constants $C < +\infty$ and $\delta > 0$ such that*

$$\begin{aligned}
\lim_{\bar{t}\to\overline{T}_{\max}} \frac{R_{\max}(\bar{t})}{R_{\min}(\bar{t})} &= 1, &(4.5.6)\\
\mathrm{Ric}_{\bar{g}(\bar{t})} &\geq \delta R_{\bar{g}(\bar{t})}\bar{g}(\bar{t}), &(4.5.7)\\
R_{\max}(\bar{t}) &\leq C, \quad \mathrm{diam}(\mathcal{M}^3, \bar{g}(\bar{t})) \geq \frac{1}{C} &(4.5.8)\\
\overline{T}_{\max} &= +\infty, &(4.5.9)\\
R_{\min}(\bar{t}) &\geq \frac{1}{C}, \quad \mathrm{diam}(\mathcal{M}^3, \bar{g}(\bar{t})) \leq C, &(4.5.10)\\
\left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{1}{3}R_{\bar{g}(\bar{t})}\bar{g}(\bar{t})\right|_{\bar{g}(\bar{t})} &\leq Ce^{-\delta\bar{t}}, &(4.5.11)\\
R_{\max}(\bar{t}) - R_{\min}(\bar{t}) &\leq Ce^{-\delta\bar{t}}, &(4.5.12)\\
\left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{1}{3}\underline{R}_{\bar{g}(\bar{t})}\bar{g}(\bar{t})\right|_{\bar{g}(\bar{t})} &\leq Ce^{-\delta\bar{t}}. &(4.5.13)
\end{aligned}$$

In particular, there exists a constant $C < +\infty$ such that

$$\frac{1}{C}\bar{g}(0) \leq \bar{g}(\bar{t}) \leq C\bar{g}(0) \tag{4.5.14}$$

for all $\bar{t} \in [0, \infty)$, and the metrics $\bar{g}(\bar{t})$ converge uniformly on compact sets to a continuous metric $\bar{g}(+\infty)$ as $\bar{t} \to +\infty$. ♡

Proof. (4.5.6) and (4.4.15) immediately follow from the corresponding results for the unnormalized Ricci flow and the scalar-invariant properties.

Since under the normalized Ricci flow the volume is invariant, we have $V_{\bar{g}(\bar{t})}$ is constant. On the other hand, applying the Bishop-Gromov volume comparison theorem to our case that $\mathrm{Ric}_{\bar{g}(\bar{t})} \geq 0$, we have

$$V_{\bar{g}(\bar{t})} \leq C\left[\mathrm{diam}(\mathcal{M}^3, \bar{g}(\bar{t}))\right]^3$$

for some universal constant C. Byproduct we have $\mathrm{diam}(\bar{g}(\bar{t})) \geq C > 0$. Since $\mathrm{Rc}_{\bar{g}(\bar{t})} \geq \varepsilon R_{\max}(t)\bar{g}(t)$ for some $\varepsilon > 0$, by the Bonnet-Myers theorem, we have

$$\mathrm{diam}(\mathcal{M}^3, \bar{g}(\bar{t})) \leq \frac{C}{\sqrt{R_{\max}(t)}}$$

and we conclude $R_{\max}(\bar{t}) \leq C$. (4.5.8) implies

$$\underline{R}_{\bar{g}(\bar{t})} = \frac{1}{V_{\bar{g}(\bar{t})}}\int_{\mathcal{M}^3} R_{\bar{g}(\bar{t})}dV_{\bar{g}(\bar{t})} \leq C.$$

If $\overline{T}_{\max} < +\infty$, then **Note 4.8** shows that $+\infty \leq C\overline{T}_{\max}$, a contradiction. Hence $\overline{T}_{\max} = +\infty$.

By Klingenberg's injectivity radius estimate and replacing $(\mathcal{M}^3, \bar{g}(\bar{t}))$ by their universal covering Riemannian manifolds $(\widetilde{\mathcal{M}}^3, \tilde{\bar{g}}(\bar{t}))$, we have

$$\mathrm{inj}_{\tilde{\bar{g}}(\bar{t})} \geq \varepsilon R_{\max}(\bar{t})^{-1/2}$$

for some universal constant $\varepsilon > 0$. Since $\mathrm{Sec}_{\tilde{\bar{g}}(\bar{t})} \leq CR_{\max}(\bar{t})$, this implies

$$V_{\tilde{\bar{g}}(\bar{t})} \geq \varepsilon R_{\max}(\bar{t})^{-3/2}$$

for some other constant $\varepsilon > 0$. Hence we have

$$V_{\bar{g}(\bar{t})} \geq \delta' R_{\max}(\bar{t})^{-3/2}$$

where $\delta' > 0$ depends also on $|\pi_1(\mathcal{M}^3)| < +\infty$. Hence $R_{\max}(\bar{t}) \geq \frac{1}{C}$ and the same estimate holds for $R_{\min}(\bar{t})$ by (4.5.6). Note that we also obtained a uniform upper bound for $\operatorname{diam}(\mathcal{M}^3, \bar{g}(\bar{t}))$.

Let

$$\bar{f}(\bar{t}) := \frac{|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{1}{3}R_{\bar{g}(\bar{t})}\bar{g}(\bar{t})|^2_{\bar{g}(\bar{t})}}{R^2_{\bar{g}(\bar{t})}}.$$

Then $\bar{f}(\bar{t})$ satisfies the following equation

$$\begin{aligned}\Box_{\bar{g}(\bar{t})}\bar{f}(\bar{t}) &= 2\left\langle \nabla_{\bar{g}(\bar{t})} \ln R_{\bar{g}(\bar{t})}, \nabla_{\bar{g}(\bar{t})}\bar{f}(\bar{t})\right\rangle_{\bar{g}(\bar{t})} \\ &\quad - \frac{2}{R^4_{\bar{g}(\bar{t})}}\left|R_{\bar{g}(\bar{t})}\nabla_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})} - \nabla_{\bar{g}(\bar{t})}R_{\bar{g}(\bar{t})} \otimes \mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} + 4\bar{P}(\bar{t})\end{aligned}$$

where

$$\overline{P}(\bar{t}) = \frac{1}{R^3_{\bar{g}}(\bar{t})}\left[\frac{5}{2}R^2_{\bar{g}(\bar{t})}\left|\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} - 2R_{\bar{g}(\bar{t})}\mathrm{tr}_{\bar{g}(\bar{t})}\mathrm{Ric}^3_{\bar{g}(\bar{t})} - \frac{1}{2}R^4_{\bar{g}(\bar{t})} - \left|\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^4_{\bar{g}(\bar{t})}\right].$$

We claim that

$$\overline{P}(\bar{t}) \leq -\delta^2\frac{|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{1}{3}R_{\bar{g}(\bar{t})}\bar{g}(\bar{t})|^2_{\bar{g}(\bar{t})}}{R_{\bar{g}(\bar{t})}}.$$

Let $\lambda_1 \geq \lambda_2 \geq \lambda_3$ denote the eigenvalues of $\mathbf{Rc}_{\bar{g}(\bar{t})}$. Then

$$\begin{aligned}R^3_{\bar{g}(\bar{t})}\overline{P}(\bar{t}) &= \frac{5}{2}(\lambda_1+\lambda_2+\lambda_3)^2\left(\lambda_1^2+\lambda_1^2+\lambda_3^2\right) \\ &\quad - 2(\lambda_1+\lambda_2+\lambda_3)\left(\lambda_1^3+\lambda_2^3+\lambda_3^3\right) - \frac{1}{2}(\lambda_1+\lambda_2+\lambda_3)^4 - \left(\lambda_1^2+\lambda_2^2+\lambda_3^2\right)^2 \\ &= -(\lambda_1-\lambda_2)^2\left[\lambda_1^2+(\lambda_1+\lambda_2)(\lambda_2-\lambda_3)\right] - \lambda_3^2(\lambda_1-\lambda_3)(\lambda_2-\lambda_3) \\ &\leq -\lambda_1^2(\lambda_1-\lambda_2)^2 - \lambda_3^2(\lambda_2-\lambda_3)^2 \leq -\delta^2R^2_{\bar{g}(\bar{t})}\left[(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^2\right] \\ &\leq -\delta^2R^2_{\bar{g}(\bar{t})}\frac{1}{3}\left[(\lambda_1-\lambda_2)^2+(\lambda_1-\lambda_3)^2+(\lambda_2-\lambda_3)^2\right] \\ &= -\delta^2R^2_{\bar{g}(\bar{t})}\left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{1}{3}R_{\bar{g}(\bar{t})}\bar{g}(\bar{t})\right|^2_{\bar{g}(\bar{t})}.\end{aligned}$$

Plugging this estimate into the evolution equation of $\bar{f}(\bar{t})$ yields

$$\begin{aligned}\Box_{\bar{g}(\bar{t})}\bar{f}(\bar{t}) &\leq 2\left\langle \nabla_{\bar{g}(\bar{t})} \ln R_{\bar{g}(\bar{t})}, \nabla_{\bar{g}(\bar{t})}\bar{f}(\bar{t})\right\rangle_{\bar{g}(\bar{t})} - 4\delta^2R_{\bar{g}(\bar{t})}\bar{f}(\bar{t}) \\ &\leq 2\left\langle \nabla_{\bar{g}(\bar{t})} \ln R_{\bar{g}(\bar{t})}, \nabla_{\bar{g}(\bar{t})}\bar{f}(\bar{t})\right\rangle_{\bar{g}(\bar{t})} - \frac{4\delta^2}{C}\bar{f}(\bar{t}).\end{aligned}$$

The maximum principle shows that $\bar{f}(\bar{t}) \leq Ce^{-\delta t}$ for some universal constant C. From (4.5.8) we derive (4.5.11).

(4.4.6) and (4.4.11) imply that

$$\psi(t) := \frac{|\nabla_{g(t)}R_{g(t)}|^2_{g(t)}}{R_{g(t)}} + 168\left(\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}\right)$$

satisfies, under the un-normalized Ricci flow

$$\Box_{g(t)}\psi(t) \leq 672R_{g(t)}\left(\left|\mathrm{Ric}_{g(t)}\right|^2_{g(t)} - \frac{1}{3}R^2_{g(t)}\right).$$

Hence, for the normalized Ricci flow, the corresponding quantity $\overline{\psi}(\bar{t})$, since ψ has degree -2, satisfies

$$\Box_{\bar{g}(\bar{t})}\overline{\psi}(\bar{t}) \le 672 R_{\bar{g}(\bar{t})}\left(\left|\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} - \frac{1}{3}R^2_{\bar{g}(\bar{t})}\right) - \frac{4}{3}\underline{R}_{\bar{g}(\bar{t})}\overline{\psi}(\bar{t}) \le Ce^{-\delta\bar{t}} - \delta_1\overline{\psi}(\bar{t})$$

for some $\delta_1 \in (0,\delta]$. We can conclude that

$$\Box_{\bar{g}(\bar{t})}\left(e^{\delta_1\bar{t}}\overline{\psi}(\bar{t}) - C\bar{t}\right) \le 0$$

and hence $\overline{\psi}(\bar{t}) \le Ce^{-\delta_1\bar{t}}(1+\bar{t}) \le Ce^{-\delta_2\bar{t}}$ for some $\delta_2 \in (0,\delta_1]$. This gives us the gradient estimate

$$\left|\nabla_{\bar{g}(\bar{t})}R_{\bar{g}(\bar{t})}\right|_{\bar{g}(\bar{t})} \le Ce^{-\delta_2\bar{t}}.$$

Since the diameters of $\bar{g}(\bar{t})$ are uniformly bounded, we obtain (4.5.12) by integrating the gradient estimate along minimal geodesics. Calculate

$$\begin{aligned}\left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{1}{3}\underline{R}_{\bar{g}(\bar{t})}\bar{g}(\bar{t})\right|_{\bar{g}(\bar{t})} &\le \left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{1}{3}R_{\bar{g}(\bar{t})}\bar{g}(\bar{t})\right|_{\bar{g}(\bar{t})} + \frac{1}{3}\left|R_{\bar{g}(\bar{t})} - \underline{R}_{\bar{g}(\bar{t})}\right|_{\bar{g}(\bar{t})}\\ &\le \left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{1}{3}R_{\bar{g}(\bar{t})}\bar{g}(\bar{t})\right|_{\bar{g}(\bar{t})} + \frac{1}{3}\left[R_{\max}(\bar{t}) - R_{\min}(\bar{t})\right].\end{aligned}$$

It is clear that (4.5.13) follows from (4.5.11) and (4.5.12). The last result will be proved later. □

4.5.3 Interpolation inequalities for tensor fields

Let $T = (T_{i_1\cdots i_\ell})$ denote a ℓ-form or a covariant tensor of degree ℓ, on a compact Riemmanian manifold $(\mathcal{M}^m, g)$ of dimension m. The L^p-norm is denoted by $||\cdot||_{L^p,g}$:

$$||\cdot||_{L^p,g} := \left(\int_{\mathcal{M}^m}|\cdot|^p_g dV_g\right)^{1/p}.$$

Theorem 4.5

Suppose

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r}, \quad r \ge 1.$$

Then

$$||\nabla_g T||^2_{L^{2r},g} \le (2r-2+m)\left|\left|\nabla^2_g T\right|\right|_{L^p,g}||T||_{L^q,g}. \tag{4.5.15}$$

♡

Proof. Calculate

$$\begin{aligned}\int_{\mathcal{M}^m}|\nabla_g T|^{2r}_g dV_g &= \int_{\mathcal{M}^m}\nabla^j T\nabla_j T\,|\nabla_g T|^{2r-2}_g dV_g\\ &= -\int_{\mathcal{M}^m}T\nabla^j\left(\nabla_j T\,|\nabla_g T|^{2r-2}_g\right)dV_g\\ &= -\int_{\mathcal{M}^m}T\left(\Delta_g T\,|\nabla_g T|^{2r-2}_g + \nabla_j T\nabla^j\,|\nabla_g T|^{2r-2}_g\right)dV_g.\end{aligned}$$

Since

$$\begin{aligned}\nabla^j|\nabla_g T|^{2r-2}_g &= \nabla^j\left(\langle\nabla_g T, \nabla_g T\rangle_g\right)^{r-1} = (r-1)\,|\nabla_g T|^{2r-4}_g\nabla^j\langle\nabla_g T,\nabla_g T\rangle_g\\ &= 2(r-1)\left\langle\nabla^j\nabla_g T, \nabla_g T\right\rangle_g|\nabla_g T|^{2r-4}_g,\end{aligned}$$

it follows that

$$\begin{aligned}\int_M |\nabla_g T|_g^{2r}\, dV_g &= -\int_M T\Delta_g T\, |\nabla_g T|_g^{2r-2}\, dV_g \\ &\quad - 2(r-1)\int_M \langle T, \nabla_j T\rangle_g \left\langle \nabla_g T, \nabla^j \nabla_g T\right\rangle_g |\nabla_g T|_g^{2r-4}\, dV_g.\end{aligned}$$

Now

$$\begin{aligned}|T\Delta_g T|_g &\le m\,|T|_g \left|\nabla_g^2 T\right|_g, \\ \langle T, \nabla_j T\rangle_g \left\langle \nabla_g T, \nabla^j \nabla_g T\right\rangle_g &\le |T|_g\, |\nabla_g T|_g^2 \left|\nabla_g^2 T\right|_g\end{aligned}$$

and therefore

$$\int_{\mathcal{M}^m} |\nabla_g T|_g^{2r}\, dV_g \le (2r-2+m)\int_{\mathcal{M}^m} |T|_g \left|\nabla_g^2 T\right|_g |\nabla_g T|_g^{2r-2}\, dV_g.$$

We can estimate the last integral using Hölder inequality with

$$\frac{1}{p}+\frac{1}{q}+\frac{r-1}{r}=1,$$

and we get

$$\begin{aligned}\int_{\mathcal{M}^m} |\nabla_g T|_g^{2r}\, dV_g &\le (2r-2+m)\left(\int_{\mathcal{M}^m} \left|\nabla_g^2 T\right|_g^2 dV_g\right) \\ &\quad \left(\int_{\mathcal{M}^m} |T|_g^q\, dV_g\right)^{1/q} \cdot \left(\int_M |\nabla_g T|_g^{2r}\, dV_g\right)^{1-\frac{1}{r}}.\end{aligned}$$

Simplifying above inequality yields the required result. □

Corollary 4.8

If $p \ge 1$ then

$$\|\nabla_g T\|_{L^{2p},g}^2 \le (2p-2+n)\max_{\mathcal{M}^m} |T|_g \cdot \left\|\nabla_g^2 T\right\|_{L^p,g}. \tag{4.5.16}$$

♡

Next we need a result on convexity, which is geometrically obvious.

Lemma 4.10

Let $f(i)$ be real-valued function of the integer k for $i = 0, 1, \cdots, k$. If

$$f(i) \le \frac{f(i-1)+f(i+1)}{2}, \quad i = 1, \cdots, k-1.$$

then

$$f(i) \le \left(1-\frac{i}{k}\right) f(0) + \frac{i}{k} f(k). \tag{4.5.17}$$

Thus, the value $f(i)$ is determined by the values of f at the endpoints. ♡

Proof. Set

$$\overline{f}(i) := f(i) - \left(1-\frac{i}{k}\right) f(0) - \frac{i}{k} f(k)$$

and $g(i) := \overline{f}(i) - \overline{f}(i-1)$ for $i = 1, \cdots, n$. We can check that the hypothesis for f also holds for $\overline{f}$:

$$\overline{f}(i-1)+\overline{f}(i+1) = f(i-1) - \left(1-\frac{i-1}{k}\right) f(0) - \frac{i-1}{k} f(k)$$

$$
\begin{aligned}
&+ f(i+1) - \left(1 - \frac{i+1}{k}\right) f(0) - \frac{i+1}{k} f(k) \\
&= f(i-1) + f(i+1) - \left(2 - \frac{2i}{k}\right) f(0) - \frac{2i}{k} f(k) \\
&\geq 2f(i) - 2\left(1 - \frac{i}{k}\right) f(0) - \frac{2i}{k} f(k) \quad = \quad 2\overline{f}(i).
\end{aligned}
$$

Moreover $\overline{f}(0) = 0 = \overline{f}(k)$, and $g(i) \leq g(i+1)$. Therefore we can choose an integer j so that

$$g(1) \leq \cdots \leq g(j) \leq 0 \leq g(j+1) \leq \cdots \leq g(k).$$

For any i,

$$\overline{f}(i) = \sum_{1 \leq \ell \leq i} g(\ell) = - \sum_{i+1 \leq \ell \leq k} g(\ell).$$

When $i \leq j$, the first representation is negative and when $i \geq j$ the second is. This proves $\overline{f}(i) \leq 0$ for $0 \leq i \leq k$. □

Corollary 4.9

If $f(i)$ satisfies

$$f(i) \leq \frac{f(i-1) + f(i+1)}{2} + C$$

for some constant C, then

$$f(i) \leq \left(1 - \frac{i}{k}\right) f(0) + \frac{i}{k} f(k) + Ci(k-i).$$

♡

Proof. Let

$$g(i) := f(i) + Ci^2.$$

Then $2g(i) \leq g(i-1) + g(i+1)$. **Lemma 4.10** implies

$$g(i) \leq \left(1 - \frac{i}{k}\right) g(0) + \frac{i}{k} g(k)$$

that is the desired result. □

Corollary 4.10

If $f(i)$ is nonnegative and satisfies

$$f(i) \leq C f(i-1)^{1/2} f(i+1)^{1/2}$$

for some positive constant $C > 0$, then

$$f(i) \leq C^{i(k-i)} f(0)^{1-\frac{i}{k}} f(k)^{\frac{i}{k}}.$$

♡

Proof. Without loss of generality, we may assume $f(i) > 0$ for all i. Set

$$g(i) := \ln f(i).$$

Hence

$$g(i) \leq \frac{g(i-1) + g(i+1)}{2} + \ln C.$$

By **Corollary 4.10**, $g(i) \le \left(1 - \frac{i}{k}\right) g(0) + \frac{i}{k} g(k) + \ln C \cdot i(k-i)$. □

Lemma 4.11

For any $k \in \mathbf{N}$ there exists a constant $C = C(k,m)$ depending only on k and m such that for any tensor field T

(1) if $j = 1, \cdots, k-1$, then

$$\int_{\mathcal{M}^m} \left|\nabla_g^j T\right|_g^{2k/j} dV_g \le C(k,m) \max_{\mathcal{M}^m} |T|_g^{2\left(\frac{k}{j}-1\right)} \int_M \left|\nabla_g^k T\right|_g^2 dV_g.$$

Here we can choose $C(k,m) = (2k-2+m)^{2k^2}$.

(2) if $j = 0, \cdots, k$, then

$$\int_{\mathcal{M}^m} \left|\nabla_g^j T\right|_g^2 dV_g \le C(k,m) \left(\int_{\mathcal{M}^m} \left|\nabla_g^k T\right|_g^2 dV_g\right)^{\frac{j}{k}} \left(\int_{\mathcal{M}^m} |T|_g^2 dV_g\right)^{1-\frac{j}{k}}.$$

Here we can choose $C(k,m) = m^{2k^2}$. ♡

Proof. Pick $p = \frac{2k}{i+1}$, $q = \frac{2k}{i-1}$, and $r = \frac{k}{i} \ge 1$ in **Theorem 4.5**. Then

$$\left|\left|\nabla_g^i T\right|\right|_{L^{\frac{2k}{i}},g}^2 \le \left(\frac{2k}{i} - 2 + m\right) \left|\left|\nabla_g^{i+1} \nabla_g T\right|\right|_{L^{\frac{2k}{i+1}},g} \left|\left|\nabla_g^{i-1} T\right|\right|_{L^{\frac{2k}{i-1}},g}.$$

For $i = 1$, we use **Corollary 4.8** to get

$$||\nabla_g T||_{L^{2k},g}^2 \le (2k-2+m) \max_{\mathcal{M}^m} |T|_g \left|\left|\nabla_g^2 T\right|\right|_{L^k,g}.$$

Let

$$C = 2k-2+n, \quad f(0) = \max_{\mathcal{M}^m} |T|_g, \quad f(i) = \left|\left|\nabla_g^i T\right|\right|_{L^{\frac{2k}{i}},g}, \quad i = 1, \cdots, k.$$

So $f(i) \le C f(i+1)^{1/2} f(i-1)^{1/2}$. By **Corollary 4.10**,

$$f(i) \le C^{i(k-i)} f(0)^{1-\frac{i}{k}} f(k)^{\frac{i}{k}}.$$

Thus,

$$\left|\left|\nabla_g^i T\right|\right|_{L^{\frac{2k}{i}},g} \le (2k-2+m)^{i(k-i)} \left(\max_{\mathcal{M}^m} |T|_g\right)^{1-\frac{i}{k}} \left|\left|\nabla_g^k T\right|\right|_{L^2,g}^{\frac{i}{k}}.$$

For (2), we choose $p = q = 2$ and $r = 1$ in **Theorem 4.5** and obtain

$$\left|\left|\nabla_g^i T\right|\right|_{L^2,g}^2 \le m \left|\left|\nabla_g^{i+1} T\right|\right|_{L^2,g} \left|\left|\nabla_g^{i-1} T\right|\right|_{L^2,g}.$$

Applying **Corollary 4.10** to above implies

$$\left|\left|\nabla_g^i T\right|\right|_{L^2,g} \le m^{i(k-i)} ||T||_{L^2,g}^{1-\frac{i}{k}} \left|\left|\nabla_g^k T\right|\right|_{L^2,g}^{\frac{i}{k}}.$$

This completes the proof. □

4.5.4 Higher derivatives of the curvature and the proof of Theorem 4.1

Next we study the higher derivatives of the curvature. It is easy to see that (or see later)

$$\begin{aligned}\Box_{g(t)} \left|\nabla_{g(t)}^k \mathrm{Rm}_{g(t)}\right|_{g(t)}^2 &\le -2\left|\nabla_{g(t)}^{k+1} \mathrm{Rm}_{g(t)}\right|_{g(t)}^2 \\ &+ C \sum_{0 \le \ell \le k} \left|\nabla_{g(t)}^\ell \mathrm{Rm}_{g(t)}\right|_{g(t)} \left|\nabla_{g(t)}^{k-\ell} \mathrm{Rm}_{g(t)}\right|_{g(t)} \left|\nabla_{g(t)}^k \mathrm{Rm}_{g(t)}\right|_{g(t)}.\end{aligned}$$

Lemma 4.12

Let $(\mathcal{M}^3, g)$ be a closed 3-manifold with positive Ricci curvature. Suppose that $\bar{g}(\bar{t})$, $\bar{t} \in [0, +\infty)$, is the solution to the normalized Ricci flow with initial value g_0. Then there exists positive constants C and δ depending on k and n such that

$$\left|\nabla^k_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|_{\bar{g}(\bar{t})} \le Ce^{-\delta\bar{t}} \tag{4.5.18}$$

for all $k \in \mathbb{N}$. ♡

Proof. First we consider the un-normalized Ricci flow. Calculate

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathcal{M}^3}\left|\nabla^k_{g(t)}\mathrm{Rm}_{g(t)}\right|^2_{g(t)} dV_{g(t)} + 2\int_{\mathcal{M}^3}\left|\nabla^{k+1}_{g(t)}\mathrm{Rm}_{g(t)}\right|^2_{g(t)} dV_{g(t)} \\
\le\ & C\sum_{0\le\ell\le k}\int_{\mathcal{M}^3}\left|\nabla^\ell_{g(t)}\mathrm{Rm}_{g(t)}\right|_{g(t)}\left|\nabla^{k-\ell}_{g(t)}\mathrm{Rm}_{g(t)}\right|_{g(t)}\left|\nabla_{g(t)}\mathrm{Rm}_{g(t)}\right|_{g(t)} dV_{g(t)} \\
\le\ & C\sum_{0\le\ell\le k}\left(\int_{\mathcal{M}^3}\left|\nabla^\ell_{g(t)}\mathrm{Rm}_{g(t)}\right|^{\frac{2k}{\ell}}_{g(t)} dV_{g(t)}\right)^{\frac{\ell}{2k}} \\
&\left(\int_{\mathcal{M}^3}\left|\nabla^{k-\ell}_{g(t)}\mathrm{Rm}_{g(t)}\right|^{\frac{2k}{k-\ell}}_{g(t)} dV_{g(t)}\right)^{\frac{k-\ell}{2k}}\left(\int_{\mathcal{M}^3}\left|\nabla^k_{g(t)}\mathrm{Rm}_{g(t)}\right|^2_{g(t)} dV_{g(t)}\right)^{1/2}.
\end{aligned}$$

By **Lemma 4.11**, we have

$$\begin{aligned}
\left(\int_{\mathcal{M}^3}\left|\nabla^\ell_{g(t)}\mathrm{Rm}_{g(t)}\right|^{\frac{2k}{\ell}}_{g(t)} dV_{g(t)}\right)^{\frac{\ell}{2k}} &\le C\max_{\mathcal{M}^3}\left|\mathrm{Rm}_{g(t)}\right|^{1-\frac{\ell}{k}}_{g(t)} \\
&\quad \left(\int_{\mathcal{M}^3}\left|\nabla^k_{g(t)}\mathrm{Rm}_{g(t)}\right|^2_{g(t)} dV_{g(t)}\right)^{\frac{\ell}{2k}} \\
\left(\int_{\mathcal{M}^3}\left|\nabla^{k-\ell}_{g(t)}\mathrm{Rm}_{g(t)}\right|^{\frac{2k}{k-\ell}}_{g(t)} dV_{g(t)}\right)^{\frac{k-\ell}{2k}} &\le C\max_{\mathcal{M}^3}\left|\mathrm{Rm}_{g(t)}\right|^{\frac{\ell}{k}}_{g(t)} \\
&\quad \left(\int_{\mathcal{M}^3}\left|\nabla^k_{g(t)}\mathrm{Rm}_{g(t)}\right|^2_{g(t)} dV_{g(t)}\right)^{\frac{k-\ell}{2k}}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\frac{d}{dt}\int_{\mathcal{M}^3}\left|\nabla^k_{g(t)}\mathrm{Rm}_{g(t)}\right|^2_{g(t)} dV_{g(t)} &\le -2\int_{\mathcal{M}^3}\left|\nabla^{k+1}_{g(t)}\mathrm{Rm}_{g(t)}\right|^2_{g(t)} dV_{g(t)} \\
+ C\max_{\mathcal{M}^3}\left|\mathrm{Rm}_{g(t)}\right|_{g(t)}\int_{\mathcal{M}^3}\left|\nabla^k_{g(t)}\mathrm{Rm}_{g(t)}\right|^2_{g(t)} dV_{g(t)}.
\end{aligned}$$

Lemma 4.7 and (4.5.8), together with above inequality, shows that

$$\begin{aligned}
\frac{d}{d\bar{t}}\int_{\mathcal{M}^3}\left|\nabla^k_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})} &\le -2\int_{\mathcal{M}^3}\left|\nabla^{k+1}_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})} \\
+ C\int_{\mathcal{M}^3}\left|\nabla^k_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})}
\end{aligned}$$

where we use the fact that in dimension three Rm is equivalent to Ric_g. For any $k > 0$ we have

$$\left|\nabla^k_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} = \left|\nabla^k_{\bar{g}(\bar{t})}\left(\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{R_{\bar{g}(\bar{t})}}{3}\bar{g}(\bar{t})\right)\right|^2_{\bar{g}(\bar{t})}.$$

Applying **Lemma 4.11** to the tensor $\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{R_{\bar{g}(\bar{t})}}{3}\bar{g}(\bar{t})$ yields

$$\begin{aligned}\int_{\mathcal{M}^3}\left|\nabla^k_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})} \ &\leq \ C\left(\int_{\mathcal{M}^3}\left|\nabla^{k+1}_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})}\right)^{\frac{k}{k+1}} \\ &\quad \left(\int_{\mathcal{M}^3}\left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{R_{\bar{g}(\bar{t})}}{3}\bar{g}(\bar{t})\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})}\right)^{\frac{1}{k+1}}.\end{aligned}$$

Now we claim that for any $x, y, \varepsilon > 0$,

$$x^n y \leq \varepsilon x^{n+1} + \frac{1}{\varepsilon^n}y^{n+1}. \tag{4.5.19}$$

The proof of (4.5.19) is based on the following elementary inequality $t^n \leq t^{n+1} + 1$ for any $t > 0$. Picking $t = \varepsilon x/y$ implies (4.5.19). For simplicity, we define

$$f_i(\bar{t}) := \int_{\mathcal{M}^3}\left|\nabla^i_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})}.$$

Using (4.5.19), we have, choosing $\varepsilon = \frac{2}{C}$,

$$\frac{d}{d\bar{t}}f_k(\bar{t}) \leq -2f_{k+1}(\bar{t}) + C\varepsilon f_{k+1}(\bar{t}) + \frac{C}{\varepsilon^n}\int_M\left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{R_{\bar{g}(\bar{t})}}{3}\bar{g}(\bar{t})\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})} \leq Ce^{-\delta\bar{t}}$$

for some $\delta > 0$. Hence

$$\frac{d}{d\bar{t}}\left(e^{\delta\bar{t}}f_k(\bar{t})\right) \leq C + \delta e^{\delta\bar{t}}$$

which implies $f_k(\bar{t}) \leq C$ for some uniform constant C depending on k and n. We can apply **Lemma 4.11** again to obtain

$$\begin{aligned}\int_M\left|\nabla^j_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^{\frac{2k}{j}} dV_{\bar{g}(\bar{t})} \ &\leq \ C\max_{\mathcal{M}^3}\left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{R_{\bar{g}(\bar{t})}}{3}\bar{g}(\bar{t})\right|^{2\left(\frac{k}{j}-1\right)} \\ &\quad \int_{\mathcal{M}^3}\left|\nabla^k_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})}\end{aligned}$$

for any $j = 1, \cdots, k-1$. Hence, given $j, p \in \mathbf{N}$, we may choose $k = pj$ to conclude

$$\begin{aligned}\int_{\mathcal{M}^3}\left|\nabla^j_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^{2p}_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})} \ &\leq \ C\max_{\mathcal{M}^3}\left|\mathrm{Ric}_{\bar{g}(\bar{t})} - \frac{R_{\bar{g}(\bar{t})}}{3}\bar{g}(\bar{t})\right|^{2(p-1)} \\ &\quad \int_{\mathcal{M}^3}\left|\nabla^{pj}_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|^2_{\bar{g}(\bar{t})} dV_{\bar{g}(\bar{t})} \ \leq \ Ce^{-\delta\bar{t}}\end{aligned}$$

for some $C < +\infty$ and $\delta > 0$, depending on j and p. Since all metrics $\bar{g}(\bar{t})$ are uniformly equivalent for $t \in [0, +\infty)$, the Sobolev constant is uniformly bounded and it follows from the Sobolev inequality that for any $k \in \mathbf{N}$, there exists $C < +\infty$ and $\delta > 0$, depending on k and n, such that

$$\left|\nabla^k_{\bar{g}(\bar{t})}\mathrm{Ric}_{\bar{g}(\bar{t})}\right|_{\bar{g}(\bar{t})} \leq Ce^{-\delta\bar{t}}$$

that completes the proof. □

From the above lemma and the fact that we can estimate the derivatives of the metrics in terms of the estimates for the derivatives of the Ricci tensor, we can finish the proof of **Theorem 4.1**:

Proof of **Theorem 4.1**. By (4.5.14) the metrics $\bar{g}(\bar{t})$ are uniformly equivalent and converge uniformly on compact sets to a continuous metric $\bar{g}(\infty)$ as $\bar{t} \to +\infty$. On the other hand, the estimates (4.5.18) imply the exponential convergence in each C^k-norm of $\bar{g}(\bar{t})$ to $\bar{g}(\infty)$. This implies $\bar{g}(\infty)$ is C^∞. By (4.5.13) we conclude

$$\left|\mathrm{Ric}_{\bar{g}(\infty)} - \frac{R_{\bar{g}(\infty)}}{3}\bar{g}(\infty)\right|_{\bar{g}(\infty)} \equiv 0.$$

That is, $\bar{g}(\infty)$ has constant positive sectional curvature.