Soul theorem And More

Guo Li

Contents

1 Soul Theorem: Statement and Remarks

Theorem 1.1 (Gromoll-Meyer, 1969 and Cheeger-Gromoll, 1972).
If (M, g) is a complete noncompact Riemannian manifold with sectional curvature $K \geq 0$, then M contains a soul $S \subset M$. The soul S is a closed totally convex submanifold and M is diffeomorphic to the normal bundle over S. Moreover, when $K > 0$, the soul is a point.

Here we explain the terms in the statement. A subset $A \subset M$ of a Riemannian manifold is said to be *totally convex* if any geodesic in M joining two points in A also lies in A. We see in Euclidean space, this agrees with the usual definition of convexity. And if a point is a totally convex subset, then there are no geodesic loops containing this point. Thus we see this is a strong condition.[①] A *normal bundle* over a submanifold S is the vector bundle whose fiber at each point $p \in S$ is the orthogonal complement of the tangent space T_pS in T_pM.

Example 1. *If M is $\mathbb{R}^n$, then any point $p \in M$ is totally convex, and actually is a soul of M.*

Example 2. *If M is a cylinder $S^1 \times \mathbb{R}$ in $\mathbb{R}^3$, then each circle $S^1 \times \{p\}$ is totally convex, and actually is a soul of M.*

Example 3. *If M is a paraboloid $\{z = x^2 + y^2\}$ in $\mathbb{R}^3$, then we see $\{(0, 0, 0)\}$ is totally convex, and actually is the only soul of M.*

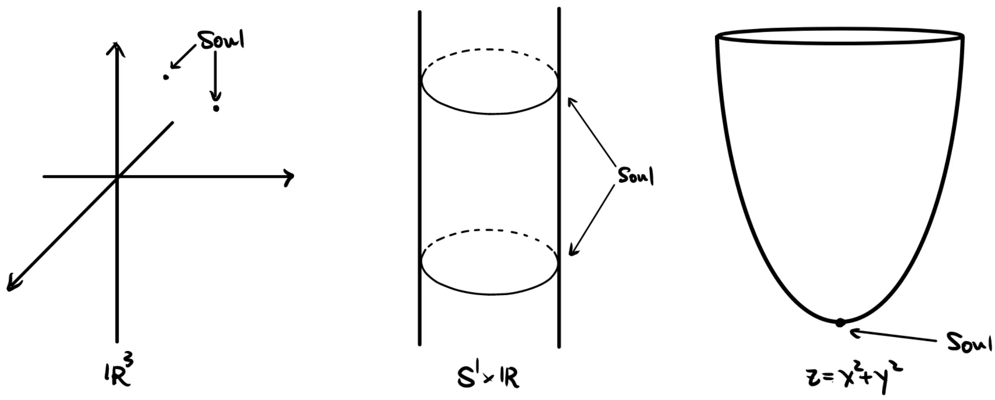

Figure 1: Soul of some manifolds

In the above examples, we see that soul may be not unique but be unique up to isometry. This is right generally, and contained in the following theorem.

Theorem 1.2. *Let S be a soul of a complete Riemannian manifold with $K \geq 0$, arriving from the above construction.*

(1) (Sharafudtinov, 1978) There is a distance nonincreasing map $\mathrm{Sh} : M \to S$ s.t. $\mathrm{Sh}|_S = \mathrm{id}$. In particular, all souls must be isometric to each other.

(2) (Perelman, 1993) The map $\mathrm{Sh} : M \to S$ is a submetry[②]. In particular, S must be a point if all sectional curvatures based at just one point in M are positive.

①In fact, if M is closed, then M is the only totally convex subset. So every compact manifold is its own soul. For this reason, the theorem is often stated only for non-compact manifolds.

②Recall $f : X \to Y$ is a submetry if $f(B_r(x)) = B_r(f(x))$ for $\forall x \in X, r \geq 0$.

Remark 1.3 (The History of Soul Theorem)**.**

(1) In 1969, Gromoll and Meyer proved that If $K > 0$ on M, then M is diffeomorphic to $\mathbb{R}^m$ (So S is a single point).

(2) Soon the Soul theorem stated here was proved by Cheeger-Gromoll in 1972.

(3) Cheeger and Gromoll also conjectured that this theorem would hold if the strict positivity of K holds only at a single point. This is the famous Soul conjecture and proved by Perelman in 1994, see [PG]: The whole paper is just 4 pages and even establishs some more powerful facts!

Remark 1.4. *There is also a beautiful theorem on the structure of manifold whose Ric ≥ 0:*

> *[The Splitting Theorem, Cheeger and Gromoll, 1971]*
> *Let (M, g) be a complete Riemannian manifold with nonnegative Ricci curvature. Suppose M contains a geodesic line, then M can be decomposed as a Riemannian product $M' \times \mathbb{R}$.*

It also has far-reaching consequences for compact manifolds with nonnegative Ricci curvature. For instance, it can be used to show that $S^3 \times S^1$ does not admit a Ricci flat metric. One can see Theorem 7.3.5 of [P] for more details.

2 Soul Theorem: Proof

2.1 Critical Point Theory of Distance Functions

In differential geometry, Morse theory is a basic tool that studies the topology of manifolds by analyzing the critical points of differentiable functions on those manifolds. The following is an inspiring result (See Lemma 12.1.1 of [P]):

> Let (M, g) be a Riemannian manifold and $f : M \to \mathbb{R}$ a proper smooth function. If f has no critical values in the closed interval $[a, b]$, then the pre-images $f^{-1}([-\infty, b])$ and $f^{-1}([-\infty, a])$ are diffeomorphic.

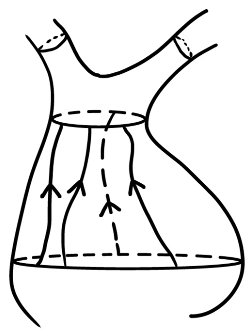

Figure 2: An inspiring result

Here we will specifically consider the distance function, which was developed by Grove-Shiohama first.

Definition 2.1. *A point $q \neq p$ is called a critical point of d_p if for all $X_q \in T_qM$, there exists a minimizing geodesic γ from $q = \gamma(0)$ to p s.t.*

$$\langle \dot{\gamma}(0), X_q \rangle \geq 0,$$

One can also replace p with a compact set K and only consider the segments from q to K with length $|qK|$, and define the critical point of d_K in the same way.

We see p is a trivial critical point of d_p, and we usually do not consider it. If q is not a critical point of d_p, then the tangent vector of all minimizing geodesics from q to p lie in an open half space of T_qM. Similarly if q is not a critical point of d_K, then $\overset{\Rightarrow}{xK}$ lies in an open half space of T_qM where $\overset{\Rightarrow}{xK}$ is the set of all unit tangent vectors of minimizing geodesics from q to K.

Example 4. *The following are some examples of critical points of* d_p*:*

(1) S^2*: the only critical point of* p *is its antipodal* $\bar{p}$*.*

(2) $S^1 \times \mathbb{R}^1$*: the only critical point of* (x, y) *is* $(-x, y)$*.*

(3) $S^1 \times S^1$ *the flat torus with fundamental domain a rectangle centered at* p*: the critical points are the two barycenters of the sides and the corner point.*

(4) If M *is compact, and* q *is a farthest point from* p*, then* q *is a critical point of* p*.*

We define d_K is α-regular at $x \in M$ if $\overset{\Rightarrow}{xK}$ is contained in a 2α sector surface of T_xM, say the center $v \in T_xM$. And denote $R_\alpha(x, K) \subset T_xM$ be the set of all such unit directions v at α-regular points x.The following proposition is a basic result.

Proposition 2.2. *(Berger) Suppose* (M, g) *and* $r(x) = |xK|$ *are as above. Then:*

1. $\overset{\Rightarrow}{xK}$ *is closed and hence compact for all* x*.*
2. *The set of* α*-regular points is open in* M*.*
3. *The set* $R_\alpha(x, K)$ *is convex for all* $\alpha \leq \pi/2$*.*
4. *If* U *is an open set of* α*-regular points for* r*, then there is a unit vector field* Y *on* U *s.t.* $Y|_x \in R_\alpha(x, K)$ *for all* $x \in U$*. Furthermore, if* c *is an integral curve for* Y *and* $s < t$*, then* $|c(s)K| - |c(t)K| > \cos(\alpha)(t - s)$*.*

Proof. This is exactly Proposition 12.1.2 of [P], whose proof is direct. □

Lemma 2.3 (Grove-Shiohama)**.** *Let* (M, g) *and* $r(x) = |xK|$ *be as above. If all points in* $r^{-1}([a, b])$ *are* α*-regular for* $\alpha < \pi/2$*, then* $r^{-1}([-\infty, a])$ *is homeomorphic to* $r^{-1}([-\infty, b])$*, and* $r^{-1}([-\infty, b])$ *deformation retracts onto* $r^{-1}([-\infty, a])$*.*

Proof. The construction is similar Lemma 12.1.1 of [P]. By choose such Y in Proposition 2.2 and a bump function $\phi : M \to [0, 1]$ which take value 1 on $r^{-1}([a, b])$ and take value 0 outside some neighborhood of $r^{-1}([a, b])$, we set $X = -\phi \cdot Y$ and consider the flow F^t for X. For $q \in M$, $\frac{\mathrm{d}}{\mathrm{d}t} r(F^t(q)) = -\phi g(\nabla r, Y) \leq -\phi$. Again by proposition 2.2, we see $r(p) - r(F^t(p)) > t \cdot \cos(\alpha)(t \geq 0)$ if $p, F^t(p) \in r^{-1}([a, b])$.

For each $p \in r^{-1}(b)$ there is a first time $t_p \leq \frac{b-a}{\cos \alpha}$ for which $F^{t_p}(p) \in r^{-1}(a)$. The function $p \mapsto t_p$ is continuous and thus we get the desired retraction

$$\forall t \in [0, 1],\ r_t : r^{-1}([-\infty, b]) \to r^{-1}([-\infty, b]),\ p \mapsto \begin{cases} p & \text{if } r(p) \leq a \\ F^{t \cdot (t_p - 0)}(p) & \text{if } a \leq r(p) \leq b. \end{cases}$$

□

The next corollary is what we want to use in the proof of the soul theorem.

Corollary 2.4. *Suppose* K *is a compact submanifold of a complete Riemannian manifold* (M, g) *and that the distance function* $|xK|$ *is regular everywhere on* $M - K$*. Then* M *is diffeomorphic to the normal bundle of* K *in* M*.*

Proof. The normal exponential map $\exp^{\perp} : T^{\perp}K \to M$ is a diffeomorphism from a neighborhood of the zero section in $T^{\perp}K$ onto a neighborhood of K. We see there is a vector field X on $M - K$ s.t. $|xK|$ increases along the integral curves for $-X$. Near K X is just $\overrightarrow{xK}$. For each $v \in T^{\perp}K$ there is a unique integral curve for $-X$ denoted $c_v(t) : (0, \infty) \to M$ s.t. $\lim_{t\to 0} \dot{c}_v(t) = v$. Define

$$F : T^{\perp}K \to M, \quad \begin{cases} F(0_p) = p & \text{for the origin in } T_p^{\perp}K, \\ F(tv) = c_v(t) & \text{where } |v| = 1. \end{cases}$$

For small t this is just the exponential map, one can show this is the desired diffeomorphism. □

2.2 Toponogov comparison theorem

Let (M, g) be complete. We call a curve σ connecting $p, q \in M$ is a *segment* if $L(\sigma) = |pq|$. And a *geodesic triangle* $\triangle ABC$ consists of three points A, B, C in M and three segments γ_{AB}, γ_{BC}, γ_{CA} joining each two of them. If just one geodesic side may be not minimal, but still satisfies the triangle inequality

$$L(\gamma_{BC}) \le L(\gamma_{AB}) + L(\gamma_{AC}),$$

then we will call $\triangle ABC$ a *generalized geodesic triangle*. Similarly we can define a *geodesic hinge* $\angle BAC$, which consists of a point A in M and two segmants γ_{AB}, γ_{AC}. If just one geodesic side is not minimal, then we call $\angle BAC$ a *generalized geodesic hinge*. From now on, when we say hinge or triangle, we always mean generalized geodesic hinge or generalized geodesic triangle.

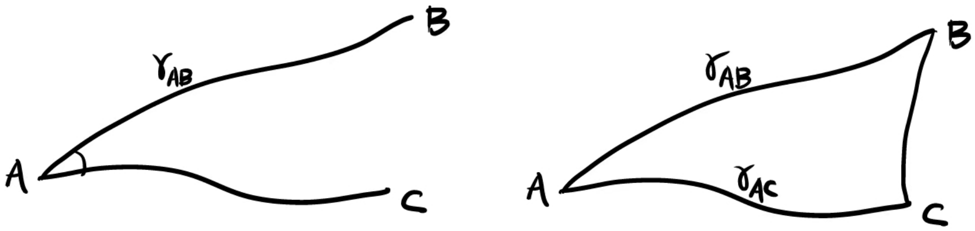

Figure 3: Hinge and triangle in M

Remark 2.5. *When $k > 0$, we have the following basic result*[3]*:*

> *If (M, g) is complete and satisfies $K \ge k > 0$, then M is compact and* $\operatorname{diam}(M, g) \le \frac{\pi}{\sqrt{k}}$. *Thus, the fundamental group is finite.*

The result can not be extended to $K > 0$, one may see the paraboloid $\{z = x^2 + y^2\}$ which has $K > 0$ and is complete but non-compact.

---

[3] In particular, M has finite fundamental group: Note that the universal cover of M has the same curvature condition to conclude that it must also be compact.

This is a theorem that we proved in our course and one also can see Corollary 6.3.2 (Hopf and Rinow, 1931 and Myers, 1932) of [P]. There is also a generalized version on manifolds with positive Ricci curvature (See Theorem 3.1 below of Chapter 9 of [dC] or Theorem 6.3.3 (Myers, 1941) of [P]):

> *[Bonnet-Meyers]*
> *Let (M, g) be a complete Riemannian manifold with Ricci curvature* $\mathrm{Ric} \geq k > 0$. *Then M is compact and* $\mathrm{diam}(M, g) \leq \frac{\pi}{\sqrt{k}}$.

The following lemma allows us to compare the hinges and triangles in M with those in *the constant sectional curvature k spaces S_k^n*, which is $\mathbb{R}^n$ if $k = 0$, the hyperbolic space $\mathbb{H}^n$ if $k < 0$, and the sphere S_k^n if $k > 0$. One can see Chapter 8: *Spaces of constant curvature* of [dC] for more details.

Lemma 2.6. *Let (M, g) be a complete Riemannian manifold of dimension n whose sectional curvature $K \geq k$. Then for each hinge and triangle in M, there is a corresponding hinge and triangle in S_k^n where the corresponding segments have the same length and the angle is the same or all corresponding segments have the same length.*

Proof. For $k = 0$ and $k < 0$, there is a unique geodesic between any two points in S_k^n, and the lemma is clear. If $k > 0$, then $\mathrm{diam}\, M \leq \frac{\pi}{\sqrt{k}}$. Moreover, if $\mathrm{diam}\, M = \frac{\pi}{\sqrt{k}}$, then according to the following theorem (See Theorem 7.2.5 of [P])

> [Cheng's maximal diameter theorem, 1975]
> If (M, g) is a complete Riemannian manifold with $\mathrm{Ric} \geq (n-1)k > 0$ and $\mathrm{diam}(M, g) = \frac{\pi}{\sqrt{k}}$, then (M, g) is isometric to S_k^n.

The conclusion holds. Hence we may assume that any minimizing geodesic has length less than $\frac{\pi}{\sqrt{k}}$:

- In the hinge case, we can fix any $\widetilde{A} \in S_k^n$, and choose any $\widetilde{B}$ s.t. $|\widetilde{A}\widetilde{B}| = L(\gamma_{AB}) < \frac{\pi}{\sqrt{k}}$ and join them by the unique segment. Then we choose a direction at $\widetilde{A}$ s.t. the angle with the geodesic $\tilde{\gamma}_{\widetilde{A}\widetilde{B}}$ is $\angle A$. Use this direction one can generate a unitspeed geodesic emanating from $\widetilde{A}$. Take the point $\widetilde{C}$ to be the point on this geodesic with length $L(\gamma_{AC})$. This is the desired comparison hinge and one may note that $\tilde{\gamma}_{\widetilde{A}\widetilde{C}}$ is minimizing if γ_{AC} is minimizing.
- In the triangle case, we pick $\widetilde{B}$ and $\widetilde{C}$ as above, and then consider the two distance spheres $\partial B_{\widetilde{B}}(|AB|)$ and $\partial B_{\widetilde{C}}(|AC|)$. One can easily check that these two spheres are non-empty and must intersect. We then take any point $\widetilde{A}$ from the intersection and connect it to $\widetilde{B}$ and $\widetilde{C}$ by minimizing geodesics.

□

Now we state the following important theorem.

Theorem 2.7 (Toponogov Comparison Theorem).
Let (M, g) be a complete Riemannian manifold with sectional curvature $K \geq k$. Then

1. *(Hinge Version) Let $\angle BAC$ be a hinge in M and $\angle \widetilde{B}\widetilde{A}\widetilde{C}$ a comparing hinge in S^n_k. Then* $\operatorname{dist}(B, C) \leq \operatorname{dist}(\widetilde{B}, \widetilde{C})$.
2. *(Triangle Version) Let $\triangle ABC$ be a triangle in M and $\triangle \widetilde{A}\widetilde{B}\widetilde{C}$ a comparing triangle in S^n_k. Then the three angles in $\triangle ABC$ are greater than the corresponding angles in $\triangle \widetilde{A}\widetilde{B}\widetilde{C}$.*

Proof. One could see Section 3.1: *Proof of Toponogov comparison theorem* of this Note. □

Remark 2.8.
(1) One can see we could replace S^n_k by S^2_k.
(2) The hinge version is equivalent to the triangle version.
(3) There seems no analogous theorem for the case $K \leq k$.

It is interesting to view the Toponogov comparison theorem 2.7 is a "global" version of the Hessian comparison theorem,

> [The Hessian Comparison Theorem]
> Let (M, g), $(\widetilde{M}, \tilde{g})$ be complete Riemannian manifolds, $\gamma : [0, b] \to M$ and $\tilde{\gamma} : [0, b] \to \widetilde{M}$ be minimizing unitspeed geodesics in M and $\widetilde{M}$ respectively s.t.
>
> $$\max\{K(\widetilde{\Pi}_{\tilde{\gamma}(t)}) \mid \dot{\tilde{\gamma}}(t) \in \widetilde{\Pi}_{\tilde{\gamma}(t)}\} =: \widetilde{K}^+(t) \leq K^-(t) := \min\{K(\Pi_{\gamma(t)}) \mid \dot{\gamma}(t) \in \Pi_{\gamma(t)}\}$$
>
> holds for all $t \in [0, b]$. Denote $q = \gamma(a)$ and $\tilde{q} = \tilde{\gamma}(a)$ $(a < b)$. Suppose $X_q \in T_q M$ and $\widetilde{X}_{\tilde{q}} \in T_{\tilde{q}} \widetilde{M}$ satisfy $\langle X_q, \dot{\gamma}(a)\rangle = \langle \widetilde{X}_{\tilde{q}}, \dot{\tilde{\gamma}}(a)\rangle, |X_q| = |\widetilde{X}_{\tilde{q}}|$. Then we have
>
> $$\nabla^2 d_p(X_q, X_q) \leq \widetilde{\nabla}^2 \tilde{d}_{\tilde{p}}(\widetilde{X}_{\tilde{q}}, \widetilde{X}_{\tilde{q}}).$$

since the norm of gradient of distance functions is always 1 and the former compare functions while the latter compare their Hessian.

Remark 2.9 (Law of Cosines in S^n_k).
This is a useful result. Let a triangle be given in S^n_k with side lengths a, b, c. If α denotes the angle opposite to a, then

$$\begin{aligned} k = 0: &\quad a^2 = b^2 + c^2 - 2bc\cos\alpha. \\ k = -1: &\quad \cosh a = \cosh b \cosh c - \sinh b \sinh c \cos\alpha. \\ k = 1: &\quad \cos a = \cos b \cos c + \sin b \sin c \cos\alpha. \end{aligned}$$

A proof of this result lies in Section 3.1:Proof of Toponogov comparison theorem of this Note.

2.3 Totally Convex Sets

To prove the soul theorem, we need to understand the properties of totally convex sets and know how to find totally convex sets in a complete Riemannian manifold.

2.3.1 How to find: by Busemann Function

From the following lemma, we see we can find totally convex sets via convexity of functions.

Lemma 2.10. *If $f : (M, g) \to \mathbb{R}$ is concave, in the sense that the Hessian is weakly nonpositive everywhere, then every superlevel set $A = \{x \in M \mid f(x) \geq a\}$ is totally convex.*

Proof. Note $f \circ c$ has nonpositive weak second derivative for each geodesic c in M and thus $f \circ c$ is concave as a function on $\mathbb{R}$. Hence the minimum of this function on any compact interval is obtained at one of the endpoints and the result follows. □

Now we need to construct a proper concave function on a complete Riemannian manifold with nonnegative sectional curvature. The idea is to use the Busemann function, which is defined in section 7.3.1 and 7.3.2 of [P].

In the rest of the subsection, we fix (M, g) is a complete, noncompact Riemannian manifold with $K \geq 0$ unless otherwise specified. Recall a *ray* $r(t) : [0, +\infty) \to (M, g)$ is a unit speed geodesic s.t. $|r(s)r(t)| = |t - s|$.

Definition 2.11. *The Busemann function $b_c : M \to \mathbb{R}$ associated to a ray c is defined as*

$$b_c(x) := \lim_{t \to +\infty} (|xc(t)| - t).$$

And given a ray c and a point $p \in M$, we consider *asymptotes* $\tilde{c}$ for c from p as a ray which is limit of a subsequence of a family of unit speed segments $\sigma_t : [0, |pc(t)|] \to (M, g)$ from p to $c(t)$. Now we state a important lemma, From this we can see many properties of the Busemann function.

Lemma 2.12. *If* $\mathrm{Ric} \geq 0$*, then* $\Delta b_c \leq 0$ *everywhere.*

Proof. First we claim that

$$b_c(x) \leq b_c(p) + b_{\tilde{c}}(x)$$

Note $\sigma_i(s) \to \tilde{c}(s)$ since $|pc(t_i)| = |p\sigma_i(s)| + |\sigma_i(s)c(t_i)|$, then

$$b_c(p) = \lim_{i \to \infty} (|pc(t_i)| - t_i) = \lim_{i \to \infty} (|p\tilde{c}(s)| + |\tilde{c}(s)c(t_i)| - t_i) = s + b_c(\tilde{c}(s))$$

Thus we have

$$\begin{aligned} |xc(s)| - s \leq |x\tilde{c}(t)| + |\tilde{c}(t)c(s)| - s &= |x\tilde{c}(t)| - t + t + |\tilde{c}(t)c(s)| - s \\ &\to b_{\tilde{c}}(x) + t + b_c(\tilde{c}(t)) = b_{\tilde{c}}(x) + b_c(p). \end{aligned}$$

Hence we only need to show that $\Delta b_{\tilde{c}} \leq 0$ at p. And if we denote $b_t(x) := |x\tilde{c}(t)| - t$, then $\Delta b_t \leq \frac{n-1}{b_t + t}$ everywhere since $b_t(x) + t$ is the distance from $c(t)$ and the Laplace comparison theorem. Hence

$$\Delta b_t(p) \leq \frac{n-1}{0+t} \to 0$$

The proof actually combines Proposition 7.3.6, Proposition 7.3.8 and Lemma 7.3.9 of [P]. □

Remark 2.13 (The Laplace Comparison Theorem).

> *Let (M, g) be a Riemannian manifold of dimension n and $\mathrm{Ric} \geq (n-1)k$. Let $\gamma : [0, b] \to M$, $\tilde{\gamma} : [0, b] \to S_k^n$ be minimizing unit-speed geodesics in M and S_k^n starting from $\gamma(0) = p$ and $\tilde{\gamma}(0) = \tilde{p}$ respectively. Then we have*
>
> $$\Delta d_p(\gamma(t)) \leq \widetilde{\Delta} \tilde{d}_{\tilde{p}}(\tilde{\gamma}(t)).$$

In our case of Lemma 2.12, $k = 0$, and $\widetilde{\Delta} \tilde{d}_{\tilde{p}} = (n-1)/\tilde{d}_{\tilde{p}}$. Moreover, One can replace S_k^n by a Riemannian manifold $\widetilde{M}$ which has constant sectional curvature k along the geodesic $\tilde{\gamma}$, and assume that for any $t \in [0, b]$, $\mathrm{Ric}(\dot{\gamma}(t), \dot{\gamma}(t)) \geq (n-1)k = \widetilde{\mathrm{Ric}}(\dot{\tilde{\gamma}}(t), \dot{\tilde{\gamma}}(t))$. The proof is identically the same.

Now we can construct a proper concave function on M which is used to find the soul.

Lemma 2.14. *Let (M, g) be complete, noncompact, have $K \geq 0$, and $p \in M$. If we take all rays $R_p = \{c : [0, \infty) \to M \mid c(0) = p\}$ and construct*

$$f = \inf_{c \in R_p} b_c,$$

*where b_c denotes the Busemann function, then f is both proper*④ *and concave.*

Proof. Recall from Lemma 2.12 that in nonnegative Ricci curvature Busemann functions are superharmonic. The proof of its concavity is almost identical. If $r(x) = |xp|$, then $\mathrm{Hess} r$ vanishes on radial directions $\partial_r = \nabla r$ because $|\nabla r| = 1$.⑤ And on vectors perpendicular to the radial direction, we have

$$\mathrm{Hess}\, r \leq \frac{g}{r}$$

This can be deduced by the following theorem (See Theorem 6.4.3 of [P])⑥:

> [Rauch Comparison Theorem]
> Assume that (M, g) satisfies $k \leq K \leq s$. If $g = dr^2 + g_r$ represents the metric in the polar coordinates, then
>
> $$\frac{\mathrm{sn}_s'(r)}{\mathrm{sn}_s(r)} g_r \leq \mathrm{Hess}\, r \leq \frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)} g_r.$$

④We call a function $f : M \to \mathbb{R}$ is *proper* if f pulls compact sets back to compact sets.

⑤Recall $\mathrm{Hess}\, f(X, Y) = g(\nabla_X \nabla f, Y) = X(Y(f)) - (\nabla_X Y) f$.

⑥The function $\mathrm{sn}_k(r)$ is a standard function, defined as:

$$\mathrm{sn}_k(r) = \begin{cases} \frac{\sin(\sqrt{k} r)}{\sqrt{k}} & \text{if } k > 0; \\ r & \text{if } k = 0; \\ \frac{\sinh(\sqrt{-k} r)}{\sqrt{-k}} & \text{if } k < 0. \end{cases} \quad \text{and} \quad \mathrm{cn}_k(t) = \mathrm{sn}_k'(t) = \begin{cases} \cos(\sqrt{k} t), & k > 0, \\ 1, & k = 0, \\ \cosh(\sqrt{-k} t), & k < 0. \end{cases}$$

In particular, $\operatorname{Hess} r \leq r^{-1} g$ at all smooth points. We can now proceed as in the Ricci curvature case to show that Busemann functions have nonpositive Hessians in the weak sense.[7]

The infimum of a collection of concave functions is clearly also concave. We left to show that f is proper. If not, some superlevel set $A = \{x \in M \mid f(x) \geq a\}$ is noncompact. If $a > 0$, then $\{x \in M \mid f(x) \geq 0\}$ is also noncompact. So we can assume that $a \leq 0$. Note all b_c are zero at p then $f(p) = 0$, so $p \in A$. Using noncompactness select a sequence $p_k \in A$ that goes to infinity. Then consider segments $\overline{pp_k}$, and as in the construction of rays, choose a subsequence s.t. $\overline{pp_k}$ converges. This forces the segments to converge to a ray, say c, emanating from p (This is a usual technique, similar to Exercise 6 of Chapter 7 of [dC]). Note $c \in R_p$ since A is totally convex and closed. This leads to a contradiction as

$$a \leq f(c(t)) \leq b_c(c(t)) = -t \to -\infty.$$

□

2.3.2 Some properties

Now we must analysis some basic properties of Totally Convex Sets.

Lemma 2.15. *If $A \subset (M, g)$ is totally convex, then A has an interior*[8]*, denoted by* $\operatorname{int}A$*, and a boundary ∂A. The interior is a totally convex submanifold of M, and the boundary has the supporting hyperplane property: for each $x \in \partial A$ there is an inward pointing vector $w \in T_x M$ s.t. any segment $\overline{xy}$ with $y \in \operatorname{int}A$ has the property that $\angle(w, \overline{xy}) < \pi/2$.*

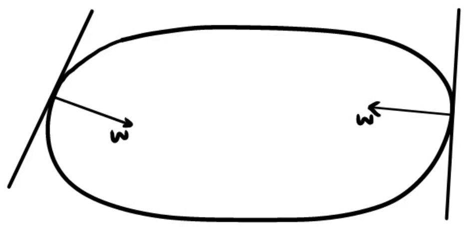

Figure 4: supporting hyperplane property

[7] Section 7.1.3 of [P] gives what "weak" means. For a continuous function $f : (M, g) \to \mathbb{R}$ we say that $\operatorname{Hess} f|_p \geq B$, where B is a symmetric bilinear map on $T_p M$ (or $\Delta f(p) \geq a$) holds in the *weak/support/barrier sense* $\iff$ for all $\varepsilon > 0$ there exist smooth support functions f_ε satisfying the following conditions.

1. $f_\varepsilon(p) = f(p)$.
2. $f(x) \geq f_\varepsilon(x)$ in some neighborhood of p.
3. $\operatorname{Hess} f_\varepsilon|_p \geq B - \varepsilon \cdot g|_p$ (or $\Delta f_\varepsilon(p) \geq a - \varepsilon$).

Lemma 7.1.5 of [P] implies that for a C^2 function $f : M \to \mathbb{R}$ the "weak" is equivalent to the "strong" condition we are familiar with. Such functions f_ε are called *support functions from below*. One can analogously use *support functions from above* to find upper bounds for $\operatorname{Hess} f$ and Δf.The notion of barrier sense was first introduced by Calabi in 1958.

[8] Here we define *the interior of the convex set* is the set which lies on one side of a hyperplane at any of the boundary points. Thus the distance function to a subset of $\operatorname{int}A$ cannot have any critical points on ∂A.

Before we prove this lemma, we state a result about "the convexity radius". Define *the injectiviy radius* at a point $p \in M$ as

$$\operatorname{inj}(p) = \sup\{R > 0 \mid \exp|_{B_R(p)} \text{ is a diffeomorphism}\} = |p\mathrm{Cut}(p)|.$$

and $\operatorname{inj}(M) = \inf_{p\in M} \operatorname{inj}(p)$. We also say the largest R s.t. $r(x) = |xp|$ is convex on $B(p,R)$ and any two points in $B(p,R)$ are joined by unique segments in $B(p,R)$ is called the *convexity radius* at p. And set $\operatorname{conv.rad}(M,g) = \inf_{p\in M} \operatorname{conv.rad}(p)$.

Proposition 2.16.

$$\operatorname{conv.rad}(M,g) \geq \min\left\{\frac{\operatorname{inj}(M,g)}{2}, \frac{\pi}{2\sqrt{\sup K(M,g)}}\right\}.$$

Proof. Set the right hand side of the inequality to be R. $R \leq \frac{\operatorname{inj}(M,g)}{2}$ means any two points in $B(p,R)$ are joined by a unique segment in M and $R \leq \frac{\pi}{2\sqrt{\sup K(M,g)}}$ means $\operatorname{Hess} r \geq 0$ on $B(p,R)$. $\forall x \in B(p,R)$ define $C_x \subset B(p,R)$ to be the set of y s.t. if $c : [0,1] \to M$ is the unique segment joining them, then $c \subset B(p,R)$. Note $x \in C_x$ and C_x is open. If $y \in B(p,R) \cap \partial C_x$, then the segment $c : [0,1] \to M$ joining x to y must lie in $\overline{B(p,R)}$ by continuity. Now consider $\varphi(t) = r(c(t)) := |c(t)p|$. We have $\varphi(0), \varphi(1) < R, \ddot{\varphi}(t) = \operatorname{Hess} r(\dot{c}(t), \dot{c}(t)) \geq 0$. Thus, φ is convex and then $\max \varphi(t) \leq \max\{\varphi(0), \varphi(1)\} < R$, showing that $c \subset B(p,R)$. □

Proof of Lemma 2.15 . By Proposition 2.16, there is a positive function $\varepsilon(p) : M \to (0,\infty)$ s.t. $r_p(x) = |xp|$ is smooth and strictly convex on $B(p,\varepsilon(p))\backslash\{p\}$. Here we assume that A is closed.

Now we seek the maximal integer k s.t. A contains a k-dimensional submanifold of M and define $N \subset A$ as being the union of all k-dimensional submanifolds in M that are contained in A. We claim that N is a k-dimensional totally convex submanifold whose closure is A.

(1) N is a submanifold: pick $p \in N$ and let $N_p \subset A$ be a k-dimensional submanifold of M containing p. By shrinking N_p if necessary, we can also assume that it is embedded. Thus there exists $\delta \in (0, \varepsilon(p))$ s.t. $B(p,\delta) \cap N_p = N_p$. The claim is that also $B(p,\delta) \cap A = N_p$. If this were not true, then we could find $q \in A \cap B(p,\delta) - N_p$. Now assume that δ is so small that also $\delta < \operatorname{inj}_q$. Then we can join each point in $B(p,\delta) \cap N_p$ to q by a unique segment. The union of these segments will, away from q, form a cone that is a $(k+1)$-dimensional submanifold contained in A, thus contradicting maximality of k.

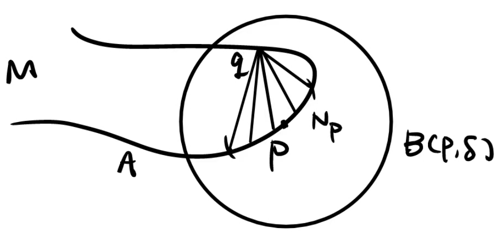

Figure 5: N is a submanifold

(2) N is dense in A: What we have just proved can easily be modified to show that for points $p \in N$ and $q \in A$ with the property that $|pq| < \operatorname{inj}(q)$ there is a k-dimensional

submanifold $N_p \subset N$ s.t. $q \in \bar{N}_p$. Specifically, choose a $(k-1)$-dimensional submanifold through p in N perpendicular to the segment from p to q, and consider the cone over this submanifold with vertex q. From this statement we get the property that any segment $\overline{xy}$ with $y \in N$ must, except possibly for x, lie in N. In particular, N is dense in A.

Hence we define $\text{int}A = N$ and $\partial A = A - N$ and now establish the supporting hyperplane property. For $p \in \partial A$, we write $T_pA \subset T_pM$ which parallel translates the tangent spaces to N along curves in N that end at p and is well defined by noting since N is totally geodesic[9], T_qN are preserved by this operation. Define

$$C_pA = \left\{ v \in T_pM \mid \exp_p(tv) \in N \text{ for some } t > 0 \right\}.$$

One can easily see that C_pA is a cone, open in T_pA and does not contain pair $\pm v \in T_pM$.

For $p \in \partial A$ and $\varepsilon > 0$ assume that there are $q \in A_\varepsilon = \{x \in A \mid |x\partial A| \geq \varepsilon\}$ with $|qp| = \varepsilon$. Choose small ε s.t. $r_q(x) = |xq|$ is smooth and convex on a neighborhood containing p. Then $\angle(-\nabla r_q, v) < \frac{\pi}{2}$ for all $v \in C_pA$, Otherwise then openness of C_pA in T_pA implies C_pA contains antipodal vectors, a contradiction.

□

The following lemma allows us to analysis the distance function to the boundary of a totally convex set.

Lemma 2.17. *Let (M, g) have $K \geq 0$. If $A \subset M$ is totally convex, then the distance function $r : A \to \mathbb{R}$ defined by $r(x) = |x\partial A|$ is concave on A. When $K > 0$, then any maximum for r is unique.*

Proof. Fix $q \in \text{int}A$, take $p \in \partial A$ s.t. $|pq| = |q\partial A|$. Then select a segment $\overline{pq}$ in A and get a hypersurface H which is perpendicular to $\overrightarrow{pq}$ and $H \cap \text{int}A = \emptyset$. We have that

$$f(x) = |xH|$$

is a support function from above for $r(x) = |x\partial A|$ at all points on $\overline{pq}$.

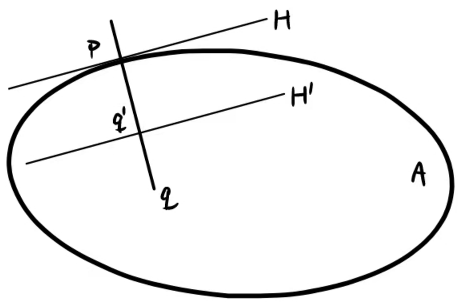

Figure 6: Support function for $|x\partial A|$

[9]Recall N is totally geodesic if $\forall p \in N, \forall \eta \in (T_pM)^\perp$, the second fundamental form $H_\eta : T_pM \times T_pM \to \mathbb{R}, (x, y) \mapsto \langle(-\nabla_x F)^T, y\rangle$ is zero, where F is a local extension of η normal to N (See Proposition 2.3 of Chapter 6 of [dC]). By Proposition 2.9 of Chapter 6 of [dC], this $\iff$ each geodesic of N is also a geodesic in M.

Select a point $q' \neq p, q$ on the segment $\overline{pq}$, we now show f is concave at $q' \neq q$. Here we need the following result about the Hessian of a function (See Theorem 3.2.2 of [P])[10]:

> [The Radial Curvature Equation]
> When $H \subset f^{-1}(a)$ consists of regular points for f we have:
>
> $$\nabla_{\nabla f} \operatorname{Hess} f + \operatorname{Hess}^2 f - \operatorname{Hess}\left(\frac{1}{2}|\nabla f|^2\right) = -R(\nabla f, \cdot, \nabla f, \cdot)$$

Apply this on a parallel field E along $\overline{pq}$, that starts out being perpendicular to $\overrightarrow{pq}$, then we get[11]:

$$\begin{aligned}\frac{d}{dt}\operatorname{Hess} f(E,E) &= -R(\nabla f, E, \nabla f, E) - \operatorname{Hess}^2 f(E,E)\\ &= -K(\nabla f, E) - \operatorname{Hess} f(\nabla_E \nabla f, E)\\ &= -K(\nabla f, E) - |\nabla_E \nabla f|^2 \leq 0.\end{aligned}$$

Since $\overline{pq}$ is an integral curve for ∇f. Combining $\operatorname{Hess} f(E,E) = 0$ at p we see that $\operatorname{Hess} f(E,E)$ is nonpositive along $\overline{pq}$ (and < 0 if $K > 0$). This shows that we have a smooth support function for $|x\partial A|$ on an open and dense subset in A. If f is not smooth at q, then choose a hypersurface H' which is perpendicular to $\overrightarrow{q'q}$ at q'. For q' close to q, $|xH'|$ is smooth at q and has nonpositive (negative) Hessian at q. We claim that

$$|pq'| + |xH'| \text{ is a support function for } |x\partial A|.$$

We only notice such x where $|x\partial A| > |pq'|$. Select a $z \in H'$ with $|x\partial A| = |z\partial A| + |xH'|$ then we just need to show $|z\partial A| \leq |pq'|, \forall z \in H'$. Choose a segment $\overline{q'z} \subset H'$, concavity of $x \mapsto |x\partial A|$ along the segment then shows $|z\partial A| \leq |q'\partial A|$ as it lies under the tangent through q'.

□

2.4 Proof: By induction

We are now ready to prove the soul theorem.[12] The maximum level set of the function $f = \inf_{c \in R_p} b_c$ constructed from Lemma 2.14

$$C_1 = \{x \in M \mid f(x) = \max f\}$$

is nonempty and totally convex since f is proper and concave. Now we find Soul of M as follows.

---

[10] For a function f, we use $S(X) = \nabla_X \nabla f$ for the $(1,1)$-tensor that corresponds to Hess f and $\operatorname{Hess}^2 f$ for the $(0,2)$-tensor that corresponds to $S^2 = S \circ S$.

[11] There may be some confused notations here. The curvature tensor define by Section 3.1.1 of [P] is $R_{[P]}(X,Y) := [\nabla_X, \nabla_Y] - \nabla_{[X,Y]}$ whereas the curvature tensor of [dc] is $R_{[dC]} = -R_{[P]}$. Here we follows the notation of [dC] and then the right hand side of the equation we quote is different from the original form in [P]. The definition of sectional curvature K and the Ricci curvature Ric are consistent.

[12] I think the proof lies in Section 12.4 of [P] is a bit strange and possibly even wrong, so the proof here is based on my understanding of the soul theorem.

- If C_1 is a submanifold, we set $S = C_1$.
- If C_1 is a convex set with nonempty boundary. $|x\partial C_1|$ is concave on C_1 by Lemma 2.17 and The maximum set C_2 is again nonempty and convex. If it is a submanifold, then we set $C_2 = S$. Otherwise we can iterate this process to obtain a sequence of convex sets $C_1 \supset C_2 \supset \cdots \supset C_k$.

We claim that by finite steps we arrive at a point or submanifold S that we call the soul, since we always have

$$\dim C_i > \dim C_{i+1}.$$

Otherwise we assume $\dim C_i = \dim C_{i+1}$ for some i. $\mathrm{int}C_{i+1}$ will be open in $\mathrm{int}C_i$ and for $p \in \mathrm{int}C_{i+1}$, there is a δ s.t. $B(p,\delta) \cap \mathrm{int}C_{i+1} = B(p,\delta) \cap \mathrm{int}C_i$. Choose a segment c from p to ∂C_i and $|x\partial C_i|$ is strictly increasing along c, however c runs through $B(p,\delta) \cap \mathrm{int}C_i$, thus $|x\partial C_i|$ must be constant on some part of c, a contradiction.

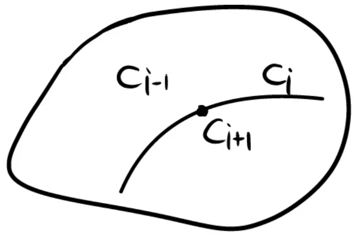

Figure 7: The way to find S

We shall show M is diffeomorphic to the normal bundle of S. Let $q \in M\backslash S$, we can find a compact totally convex set C s.t. $q \in \partial C$ and $S \subset \mathrm{int}\ C$ by similar above construction of finite steps and any geodesic from q to S has its initial tangent vector in the tangent cone T_qC. All such initial vectors are contained in an open half space of T_qM by Lemma 2.15, therefore dist_S has no critical points on $M \quad S$ and the whole result follows by Corollary 2.4.

Example 5. *We now repeat the entire proof above on $M = S^1 \times \mathbb{R}$ to better understand what we have done. Set $p = (0,0)$ and note the rays start at p are only $c(t) = (0,t)$ and $\tilde{c}(t) = (0,-t)$. Hence for $q = (\theta, s)$, we have*

$$f(q) = \inf_{c \in R_p} b_c(q) = \inf\left\{\lim_{t\to\infty}\left(\sqrt{\theta^2 + (t \pm s)^2} - t\right)\right\} = -|s|.$$

The maximum set is $C_1 = \{(\theta, 0)\}$, which is a submanifold and we see C_1 is a soul of $S^1 \times \mathbb{R}$.

It is worth mentioning that there is a lecture note of Wolfgang Meyer[13] which gives a different proof (actually a different construction of C_1 above) of the soul theorem, as an application of the Toponogov comparison theorem. This lecture is basically self-consistent and has many interesting remarks so is worth reading.

3 Appendix

3.1 Proof of Toponogov comparison theorem

The subsection mainly gives the proof of the Cosines Law for S^n_k and the Toponogov comparison theorem 2.7, by investigativing the Hessian of the distance function d_p.

Let $\gamma : [0,l] \to S^n_k$ be a unitspeed geodesic from p to $q \notin \mathrm{Cut}(p) \cup \{p\}$. By the calculation on our course or the example 2.3 of Chapter 5 of [dC], for any $X_q \in (\dot{\gamma}(l))^\perp$, if we set $X_q(t)$

[13]Available at https://www2.math.upenn.edu/~wziller/math660/TopogonovTheorem-Myer.pdf

is the parallel vector field along γ with $X_q(l) = X_q$, then the Jacobi field V along γ with $V(0) = 0$ and $V(l) = X_q$ is

$$V(t) = \frac{\mathrm{sn}_k(t)}{\mathrm{sn}_k(l)} X_q(t).$$

For any $Y_q \in T_qM$, consider a flow $\sigma(s,t)$ where $\sigma(0,t) = \gamma(t)$, $\frac{\partial}{\partial s}\sigma(0,t) = V(t)$, then

$$(\nabla^2 d_p)_q(X_q, Y_q) = \langle \nabla_{X_q} \nabla d_p, Y_q \rangle = \langle \nabla_{\frac{\partial}{\partial s}} \frac{\partial \sigma}{\partial t} \mid_{s=0,t=l}, Y_q \rangle = \langle \nabla_{\dot{\gamma}(l)} V, Y_q \rangle = \frac{\mathrm{cn}_k(l)}{\mathrm{sn}_k(l)} \langle X_q, Y_q \rangle.$$

since $\nabla_{\dot{\gamma}} X_q = 0$. So with respect to an orthonormal basis $e_1(l) = \dot{\gamma}(l), e_2(l), \cdots, e_m(l)$, we have

$$(\nabla^2 d_p)_q = \mathrm{diag}\left\{0, \frac{\mathrm{cn}_k(l)}{\mathrm{sn}_k(l)}, \cdots, \frac{\mathrm{cn}_k(l)}{\mathrm{sn}_k(l)}\right\}$$

Just like Section 4.3: *Warped Products in General* of [P], we wish to modify d_p s.t. its Hessian is indeed a constant matrix. For this purpose, we consider the function $f \circ d_p$, where f is a smooth function on $\mathbb{R}$ that we will choose later. Now for any $X_q, Y_q \in T_qM$,

$$\begin{aligned}
\nabla^2(f \circ d_p)(X_q, Y_q) &= \langle \nabla_{X_q} \nabla(f \circ d_p), Y_q \rangle = \langle \nabla_{X_q}(f'(d_p)\nabla d_p), Y_q \rangle \\
&= f'(d_p)\langle \nabla_{X_q} \nabla d_p, Y_q \rangle + f''(d_p)\langle \nabla d_p, X_q \rangle \langle \nabla d_p, Y_q \rangle \\
&= f'(d_p)\nabla^2 d_p(X_q, Y_q) + f''(d_p)\langle \dot{\gamma}(l), X_q \rangle \langle \dot{\gamma}(l), Y_q \rangle,
\end{aligned}$$

Thus we should choose f s.t. $f'(t)\frac{\mathrm{cn}_k(t)}{\mathrm{sn}_k(t)} = f''(t)$. an it is enough to take f to satisfy $f'(t) = \mathrm{sn}_k(t)$, i.e., take f to be

$$\mathrm{md}_k(r) := \int_0^r \mathrm{sn}_k(t)dt = \begin{cases} \frac{1-\cos(\sqrt{k}r)}{k}, & \text{if } k > 0, \\ \frac{r^2}{2}, & \text{if } k = 0, \\ \frac{1-\cosh(\sqrt{-k}r)}{k}, & \text{if } k < 0. \end{cases}$$

Then we have $\nabla^2(\mathrm{md}_k \circ d_p)_q = \mathrm{cn}_k(l)\,\mathrm{Id}$. Now we start fulfilling our promise in this section.

Proof of the Cosine Law in S_k^n.

If we let $\varphi(t) = \mathrm{md}_k \circ d_p \circ \gamma(t)$, then

$$\varphi''(t) = \mathrm{cn}_k \circ d_p \circ \gamma(t) = 1 - k\mathrm{md}_k(d_p \circ \gamma(t)) = 1 - k\varphi(t).$$

Consider a geodesic triangle $\triangle ABC$ in S_k^n with side lengths a, b, c and angles A, B, C, where for $k > 0$ we require $a, b, c < \pi/\sqrt{k}$. Take γ_{BC} and γ_{CA} as the unitspeed geodesics from B to C and from C to A, respectively. Now consider $\varphi(t) = \mathrm{md}_k \circ d_B \circ \gamma_{CA}(t)$, then $\varphi(0) = \mathrm{md}_k(a), \varphi(b) = \mathrm{md}_k(c)$ and $\varphi'(0) = \mathrm{sn}_k(a)\langle \dot{\gamma}_1(a), \dot{\gamma}_2(0) \rangle = -\mathrm{sn}_k(a)\cos C$.

- if $k = 0$, we get $\varphi(t) = \frac{a^2}{2} - a\cos Ct + \frac{1}{2}t^2$ and thus $\varphi(b) = \mathrm{md}_k(c)$ becomes

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

- If $k \neq 0$, we get $\varphi(t) = \frac{1}{k} + c_1\mathrm{sn}_k(t) + c_2\mathrm{cn}_k(t)$. And by the initial condition, $c_1 = -\mathrm{sn}_k(a)\cos C$, $c_2 = -\frac{1}{k}\mathrm{cn}_k(a)$. Hence $\varphi(b) = \mathrm{md}_k(c)$ becomes the following cosine law in S_k^n

$$\mathrm{cn}_k(c) = \mathrm{cn}_k(a)\,\mathrm{cn}_k(b) + k\,\mathrm{sn}_k(a)\,\mathrm{sn}_k(b)\cos C.$$

□

Proof of The Toponogov comparison theorem.

By the cosine law in S_k^2, we see for a hinge with sides γ_0, γ_1 and angle α, then the the distance between the end points is increasing in α, i.e. the Hinge version implies the Triangle version.

Denote $\gamma_0 = \gamma_{AB}$, $\gamma_1 = \gamma_{AC}$ with length l_0, l_1 and $\alpha = \angle BAC$. Assume γ_0 is minimizing. For $\varepsilon > 0$, let $\rho_\varepsilon(t) = d(\gamma_0(l_0 - \varepsilon), \gamma_1(t)), t \in [0, l_1]$, then ρ_ε is smooth for $t > 0$ small enough, $\rho_\varepsilon(0) = l_0 - \varepsilon$ and $\rho_\varepsilon'(0) = \langle -\dot\gamma_0(0), \dot\gamma_1(0)\rangle = -\cos\alpha$. By the global Hessian comparison theorem, the following inequality holds in the barrier sense:

$$(\mathrm{md}_k \circ \rho_\varepsilon)''(t) \leq \mathrm{cn}_k \circ \rho_\varepsilon(t) = 1 - k\mathrm{md}_k \circ \rho_\varepsilon(t)$$

For the comparison hinge in S_k^2, we have $(\mathrm{md}_k \circ \tilde\rho_\varepsilon)''(t) = 1 - k\mathrm{md}_k \circ \tilde\rho_\varepsilon(t)$. So if we let $f(t) = \mathrm{md}_k \circ \rho_\varepsilon - \mathrm{md}_k \circ \tilde\rho_\varepsilon$, then in the barrier sense, $f''(t) + kf(t) \leq 0$ where $f(0) = 0$, $f'(0) = sn_k(\rho_\varepsilon(0))\rho_\varepsilon'(0) - sn_k(\tilde\rho_\varepsilon'(0))\tilde\rho_\varepsilon'(0) = 0$. We claim that

$$f(t) \leq 0, \forall t \in [0, l_1].$$

It follows that $\rho_\varepsilon(t) \leq \tilde\rho_\varepsilon(t)$ for all $t \in [0, l_1]$. Letting $\varepsilon \to 0$ we get the desired conclusion.

Now all is reduced to porve the claim. Let $f_\varepsilon(t) = f(t) - \varepsilon sn_k(t)$, then $f_\varepsilon(0) = 0$ and $f_\varepsilon'(0) \leq -\varepsilon < 0$. So $f_\varepsilon(t) < 0$ for $t > 0$ small enough. We claim $f_\varepsilon(t) \leq 0$ for all $t \in [0, l]$ and let $\varepsilon \to 0$ we get the desired conclusion.

By contradiction assume t_0 be the smallest positive root of f_ε.

- $k \leq 0$. Suppose $f_\varepsilon|_{[0,t_0]}$ takes its minimum at t_1. Then we get,

  $$f_\varepsilon''(t_1) + kf_\varepsilon(t_1) = f''(t_1) + kf(t_1) - \varepsilon(sn_k''(t_1) + ksn_k(t_1)) \leq 0$$

  in the barrier sense. Thus $-kf_\varepsilon(t_1) \geq 0$, i.e. $f_\varepsilon(t_1) \geq 0$, a contradiction.

- $k > 0$. We may assume $t_0 < l$, otherwise we are done. Take $\delta > 0$ small s.t. $[-\delta, \frac{\pi}{\sqrt{k+\delta}} - \delta] \supset [0, t_0]$. Let $\phi(t) = -\sin(\sqrt{k+\delta}(t+\delta))$ and suppose $\frac{f_\varepsilon}{\phi}|_{[0,t_0]}$ takes its maximum at t_1. Let $g_{\varepsilon,\varepsilon'}$ be an upper barrier function of f_ε at t_1, i.e.

  $$g_{\varepsilon,\varepsilon'}(t_1) = f_\varepsilon(t_1),\ g_{\varepsilon,\varepsilon'}(t) \geq f_\varepsilon(t) \text{ near } t_1 \text{ and } g_{\varepsilon,\varepsilon'}''(t_1) \leq (-k)f_\varepsilon(t_1) + \varepsilon'.$$

  Then t_1 is a maximum for $\frac{g_{\varepsilon,\varepsilon'}}{\phi}$ since $\phi < 0$. It follows $\left(\frac{g_{\varepsilon,\varepsilon'}}{\phi}\right)'(t) = \frac{g_{\varepsilon,\varepsilon'}'(t)\phi(t) - g_{\varepsilon,\varepsilon'}(t)\phi'(t)}{\phi^2(t)}$ equals 0 at t_1, and thus

  $$0 \geq \left(\frac{g_{\varepsilon,\varepsilon'}}{\phi}\right)''(t_1) = \frac{g_{\varepsilon,\varepsilon'}''(t_1)\phi(t_1) - g_{\varepsilon,\varepsilon'}(t_1)\phi''(t_1)}{\phi^2(t_1)} \geq \frac{\varepsilon' + \delta g_{\varepsilon,\varepsilon'}(t_1)}{\phi(t_1)}.$$

  Take $\varepsilon' \to 0$ we get $f_\varepsilon(t_1) = g_{\varepsilon,\varepsilon'}(t_1) \geq 0$, a contradiction.

□

3.2 Some Sphere theorems

In 1926 Hopf proved that any compact simply connected Riemannian manifold with constant curvature 1 must be the standard round sphere S^m.[14] Motivated by this result, Hopf posed the question whether a compact, simply connected manifold with suitably pinched curvature is topologically a sphere. This is exactly the celebrated Rauch-Berger-Klingenberg sphere theorem below.

Theorem 3.1 (Rauch-Berger-Klingenberg sphere theorem)**.**
Let (M, g) be a complete simply connected Riemannian manifold with $\frac{1}{4} < K \leq 1$. Then M is homeomorphic to a sphere.

Proof. By Bonnet-Myers theorem, M is compact. So there exists $k > \frac{1}{4}$ s.t. $k \leq K \leq 1$. Then by Klingenberg's estimate (One can see Theorem 6.5.5 of [P]):

> [Klingenberg's estimate, 1961]
> Let (M, g) be a complete simply connected Riemannian manifold s.t. $\frac{1}{4} < K \leq 1$. Then $\mathrm{inj}(M, g) \geq \pi$.

We have $\frac{\pi}{\sqrt{k}} \geq l := \mathrm{diam}(M, g) \geq \mathrm{inj}(M, g) \geq \pi > \frac{\pi}{2\sqrt{k}}$. Take $p, q \in M$ s.t. $d(p, q) = \mathrm{diam}(M, g)$. Let $q_0 \in M$ be any point in M s.t. $l_1 := d(p, q_0) > \frac{\pi}{2\sqrt{k}}$, and let γ_1 be a minimizing unitspeed geodesic connecting $p = \gamma_1(0)$ to $q_0 = \gamma_1(l_1)$. According to the following lemma:

> [Berger]
> Let (M, g) be a compact Riemannian manifold, $p, q \in M$ s.t. $d(p, q) = \mathrm{diam}(M, g)$. Then q is a critical point of p.

Hence one can find a minimizing unitspeed geodesic γ_2 from $p = \gamma_2(0)$ to $q = \gamma_2(l)$ s.t. $\langle \dot{\gamma}_1(0), \dot{\gamma}_2(0) \rangle \geq 0$, i.e. the angle α between $\dot{\gamma}_1(0)$ and $\dot{\gamma}_2(0)$ is no more than $\frac{\pi}{2}$. Applying the Toponogov comparison theorem 2.7 to the hinge $\angle q_0 p q$, we get

$$\frac{\pi}{2\sqrt{k}} < d(q_0, q) \leq d(\tilde{q}_0, \tilde{q}) \leq \frac{\pi}{\sqrt{k}},$$

where $\angle \tilde{q}_0 \tilde{p} \tilde{q}$ is a comparison hinge in S^n_k. By the cosine law in S^n_k,

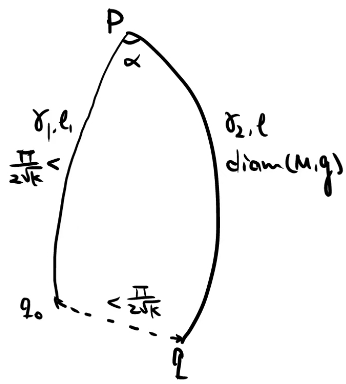

Figure 8: Balls of p and q cover M

$$\cos(\sqrt{k} d(q_0, q)) \geq \cos(\sqrt{k} d(\tilde{q}_0, \tilde{q})) = \cos(\sqrt{k} l) \cos(\sqrt{k} l_1) + \sin(\sqrt{k} l) \sin(\sqrt{k} l_1) \cos(\alpha) > 0.$$

It follows that $d(q_1, q) < \frac{\pi}{2\sqrt{k}}$. In other words, we proved $\overline{B_{\frac{\pi}{2\sqrt{k}}}(p)} \cup \overline{B_{\frac{\pi}{2\sqrt{k}}}(q)} = M$. So if we denote $r = \frac{1}{2}\left(\mathrm{inj}(M, g) + \frac{\pi}{2\sqrt{k}}\right) > \frac{\pi}{2\sqrt{k}}$, then $M = B_r(p) \cup B_r(q)$. Since $r < \mathrm{inj}(M, g)$, both

[14] In dimension $n = 2$, this is a standard result by Gauss-Bonnet theorem, which can deduce that a compact surface of positive Gaussian curvature is diffeomorphic to S^2 or $\mathbb{RP}^2$.

$B_r(p)$ and $B_r(q)$ are homeomorphic to $\mathbb{R}^m$. Now the sphere theorem follows from the following well-known theorem in topology:

> [Brown]
> Let M be a smooth compact manifold. If $M = U_1 \cup U_2$, where U_1, U_2 are open subsets in M that are homeomorphic to $\mathbb{R}^m$, then M is homeomorphic to the sphere S^m.

One can also construct the homeomorphism explicitly instead of using the Brown theorem. □

Remark 3.2. *The history of the sphere theorem (where δ-pinched means $\delta < K \leq 1$):*
(1) Rauch, 1951: all dimension, $\frac{3}{4}$-pinched.
(2) Klingenberg 1959: m even and 0.55-pinched.
(3) Berger 1960: m even and $\frac{1}{4}$-pinched.
(4) Klingenberg 1961: all dimension, $\frac{1}{4}$-pinched.
(5) Gromov 1982: Reprove.[15]

Remark 3.3. *There is a generalized sphere theorem*

> *[Berger, 1962 and Grove-Shiohama, 1977]*
> *Let (M, g) be a complete simply connected Riemannian manifold with $K > \frac{1}{4}$ and $\mathrm{diam}(M, g) \geq \pi$, then M is homeomorphic to S^m.*

One can see Theorem 12.3.2 of [P] for more details where two different proofs are given.

Remark 3.4. *The pinching constant $1/4$ is optimal.*

in fact, any compact rank one symmetric space (CROSS) admits a metric whose sectional curvatures lie in the interval $[1, 4]$. The list of these spaces includes the following examples:
(1) The complex projective space $\mathbb{CP}^m$ (dimension $2m \geq 4$).
(2) The quaternionic projective space $\mathbb{HP}^m$ (dimension $4m \geq 8$).
(3) The projective plane over the octonions $\mathbb{OP}^2$ (dimension 16).

The manifold $\mathbb{HP}^m$ can similarly be defined as the space of left (we need to distinguish left and right quaternionic lines since $\mathbb{H}$ is not commutative.) quaternionic lines through the origin in $\mathbb{H}^{m+1}$. One can see $\mathbb{HP}^m = S^{4m+3}/S^3$ since the group of unit quaternions acts freely on the unit sphere $S^{4m+3} \subset \mathbb{H}^{m+1}$ by $\alpha.(z_1, \dots, z_{m+1}) = (\alpha z_1, \dots, \alpha z_{m+1})$ and left multiplication by a unit quaternion preserves the standard metric on S^{4m+3}. The concept of an octonionic projective space $\mathbb{OP}^n$ only makes sense for $n \leq 2$, due to the nonassociativity of $\mathbb{O}$.

It is natural and nontrivial to ask whether "homeomorphic to a sphere" can be replaced by "diffeomorphic to a sphere", as we all know there are *Exotic Spheres* which is a differentiable manifold M that is homeomorphic but not diffeomorphic to the standard Euclidean n-sphere, e.g. the famous Milnor's exotic 7-sphere. And The answer is YES.

Theorem 3.5. *Let (M, g) be a complete simply connected n-Riemannian manifold,*

[15] J. H. Eschenburg, *Local convexity and nonnegative curvature - Gromov' s proof of the sphere theorem*, Invent. Math. 84, 507–522 (1986). Available at
https://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/deliver/index/docId/57849/file/57849.pdf

- *(Brendle-Schoen, 2009) If it satisfies $0 < \sup K(\pi) < 4 \inf K(\pi)$ for $\forall p \in M$ and plane $\pi \subset T_pM$, then M is diffeomorphic to the sphere S^n.*

- *(Petersen-Tao, 2010) There is $\varepsilon(n) > 0$ s.t. if $1 \leq K \leq 4 + \varepsilon$, then M is diffeomorphic to a sphere or a CROSS.*

One could see the survey paper [BS], or Section 12.3 of [P] for more details. They are the main reference materials for this subsection.

3.3 A finiteness theorem of Gromov

As an interesting application of the Toponogov comparison theorem 2.7, we discuss a finiteness theorem of Gromov, which refers to a paper of Raman Aliakseyeu[16] and Section 12.5: *Finiteness of Betti Numbers* of [P].

Theorem 3.6 (Gromov, 1978 and 1981)**.** *There is a constant $C(n)$ s.t. any complete manifold (M, g) with $\sec \geq 0$ satisfies*

(1) $\pi_1(M)$ can be generated by $\leq C(n)$ generators.

(2) For any field $\mathbb{F}$ of coefficients the Betti numbers are bounded:

$$\sum_{i=0}^{n} b_i(M, \mathbb{F}) = \sum_{i=0}^{n} \dim H_i(M, \mathbb{F}) \leq C(n).$$

Prooof of (1). We will consider $\pi_1(M)$ as the group of Deck transformations on the universal covering $\widetilde{M}$. Fix $p \in \widetilde{M}$ and choose inductively a generating set of $\pi_1(M)$ as follows:

$|pg_1(p)| \leq |pg(p)|$ for all $g \in \pi_1(M) \backslash \{e\}$ and $|pg_k(p)| \leq |pg(p)|$ for all $g \in \pi_1(M) \backslash \langle g_1, \cdots, g_{k-1} \rangle$.

Now let $\tilde{\gamma}_k$ be a minimal geodesic in $(\widetilde{M}, \tilde{g})$ from p to $g_k \cdot p$ and $l_k = |pg_k(p)|$. Suppose for some k and l, the angle between $\tilde{\gamma}_k$ and $\tilde{\gamma}_{k+l}$ is less than $\frac{\pi}{3}$, by the Toponogov comparison theorem 2.7,

$$|g_{k+l}(p) g_k(p)|^2 < l_k^2 + l_{k+l}^2 - l_k l_{k+l} \leq l_{k+l}^2.$$

Then we have $|p(g_{k+l}^{-1} g_k)(p)| = |g_{k+l}(p) g_k(p)| < l_{k+l} = |pg_{k+l}(p)|$ which contradicts with the choice of g_{k+l}. Hence we have done. □

Remark 3.7. *The part (2) of Theorem 3.6 is considered as the "most beautiful theorem" in Riemannian geometry. One can see a proof of Section 12.5: Finiteness of Betti numbers of [P]. Gromov also conjecture that*

$$C(m) = 2^m, \text{ which the right hand side is the total Betti number for } T^m.$$

---

[16] Available at https://math.uchicago.edu/~may/REU2024/REUPapers/Aliakseyeu.pdf

By the same way one can prove the case k negative, but the diameter of M must be bounded, since we have examples of surface of genus g:

> *[Gromov]*
> *For k negative, there is a constant $C = C(m, k, D)$ s.t. for any complete Riemannian manifold (M, g) with $K \geq k$ and $\operatorname{diam}(M, g) \leq D$, the fundamental group $\pi_1(M)$ is generated by no more than $C(m, k, D)$ generators.*

A famous conjecture of H. Hopf asks if $S^2 \times S^2$ admits a metric of positive sectional curvature. Theorem 3.6 gives some insight into this question, which shows that the connect sum $(S^2 \times S^2)\# \cdots \#(S^2 \times S^2)$ with sufficiently many terms does not admit a metric of nonnegative sectional curvature. Moreover, since Sha-Yang exhibited a metric of positive Ricci curvature on this space, it follows there exist simply connected manifolds that admit positive Ricci curvature but not nonnegative sectional curvature.

Acknowledgements

感谢授课教师王健对微分流形课程一学期的辛勤付出以及教学热情。通过该课程以及最后对灵魂定理的了解和学习，我逐渐理解为何老师说各种比较定理是黎曼几何的精髓，以及流形上度量的一些性质如何与流形本身的拓扑相互限制。在这门课程中的收获如此丰富，已然超出了我最初的期待。

References

[BS] Brendle, S., & Schoen, R. M. (2010). *Curvature, sphere theorems, and the Ricci flow.* arXiv. https://arxiv.org/abs/1001.2278

[dC] do Carmo, M. P. (1992). *Riemannian geometry.* Translated by Francis Flaherty. Birkhäuser.

[P] Petersen, Peter (2016). *Riemannian geometry.* Graduate Texts in Mathematics. Vol. 171 (Third edition of 1998 original ed.). Springer

[PG] Perelman, Grigori (1994). *Proof of the soul conjecture of Cheeger and Gromoll.* Journal of Differential Geometry. 40 (1): 209–212.