Perelman's proof of the soul conjecture

Jack Ceroni[1, *]

[1] *Department of Mathematics, University of Toronto*

(Dated: Friday 6th December, 2024)

Contents

I. Introduction

The goal of this paper is to outline Perelman's proof of the *soul conjecture*, originally formulated by Cheeger and Gromoll [1]. Perelman's proof is stunningly concise, so I will attempt to provide a fair amount of detail in this write-up. I will also attempt to keep this essay as self-contained as possible, mostly relying on basic Riemannian geometry knowledge and a few elementary results from Lee's book [2] (with a few exceptions), which I will cite throughout. In additon to the original paper of Perelman [1], the notes by Keshu [3] were helpful in preparing this writeup.

II. Background

We begin by stating the original result of Cheeger and Gromoll which led to the formulation of the soul conjecture discussed in this paper. First, recall that a Riemannian submanifold S of a Riemannian manifold (M, g) is said to be *totally geodesic* if the second fundamental form Π^S on S vanishes. In addition, a Riemannian submanifold S is said to be *totally convex* if for every pair of points $p, q \in S$, S contains every geodesic in M between these points.

Theorem II.1 (Cheeger-Gromoll, Ref.[4])**.** Let (M, g) be a complete non-compact Riemannian manifold of non-negative sectional curvature. Then M contains a (perhaps non-unique) totally convex, totally geodesic compact submanifold without boundary, with $0 \leq \dim(S) < \dim(M)$, such that M is diffeomorphic to the total space of the normal bundle of S in M. We call S a *soul* of M.

---

* jack.ceroni@mail.utoronto.ca

Remark II.1. Similar to how the exponential map $\exp_p : T_pM \to M$ (where $\exp_p$ is defined on all of T_pM for a complete manifold) is a local diffeomorphism away from points conjugate to p, associating velocity vectors at p to points lying away from p, the mapping given in the theorem of Cheeger-Gromoll associates velocity vectors pointing away from $p \in S$ to every point of M, in a diffeomorphic way.

In the particular case that M has strictly positive sectional curvature, Cheeger and Gromoll also proved that a soul of M will always be a single point s. Of course, in this case, the tangent bundle of s is empty, so the normal bundle is simply $T_sM \simeq \mathbb{R}^{\dim(M)}$, thus implying that M is diffeomorphic to Euclidean space. This result led to their posing of what was known as the *soul conjecture*, before it was proved by Perelman. In particular, Cheeger and Gromoll conjectured that a soul of M would always be a point under the weaker assumption that M has strictly positive sectional curvatures at a single point. This is precisely the result which was proved by Perelman.

III. The proof

The main theorem proved by Perelman is the following:

Theorem III.1 (Perelman, Ref. [1])**.** Let (M, g) be a complete, non-compact Riemannian manifold of non-negative sectional curvature, let S be a soul of M. Recall that a *retraction* $f : X \to A$, where X is a Riemannian manifold and $A \subset X$ is a submanifold, is a continuous map which fixes A. Let $P : M \to S$ be a distance non-increasing retraction. We have the following:

1. Let UNS denote the unit normal bundle of S: the subbundle of the normal bundle consisting of all unit-length vectors. Then, for any $p \in S$ and $V \in UN_pS \subset T_pM$, we have

$$P(\exp_p(tV)) = p \quad \text{for all } t \geq 0 \tag{1}$$

   Note that since M is complete, we can always assume the exponential map is defined on the entire tangent space.

2. For any geodesic $\gamma : [a, b] \to M$ in S and any vector field $V \in \Gamma(UNS)$ (i.e. V is a smooth section of the unit normal bundle of S) which is parallel along γ, the curves $\gamma_t : [a, b] \to M$ given by $\gamma_t(s) = \exp_{\gamma(s)}(tV)$ are geodesics, and the union of their images form a flat (zero sectional curvature), totally geodesic "strip" as we vary t over $[0, \infty)$. Moreover, if some curve segment $\gamma|_{[u_0,u_1]}$ is minimizing, then all of the curves $\gamma_t|_{[u_0,u_1]}$ are minimizing.

This theorem, as will be explained later, will imply that the soul conjecture is true.

The proof strategy is to show that if (1) and (2) hold for $t \in [0, \ell]$ for some $\ell \geq 0$, then they also holds for $t \in [0, \ell + \varepsilon(\ell)]$, where $\varepsilon(\ell) > 0$. If we had this fact, then we can immediately argue by contradiction: let t_0 be the supremum over all ℓ for which (1) and (2) hold on $[0, \ell]$, assume $t_0 < \infty$. Since $P(\exp_p(tV)) = p$ for every $t < t_0$, we will have $P(\exp_p(t_0V)) = p$, by continuity, so (1) holds at t_0. In addition, we will have $D_s\dot{\gamma}_t(s) = 0$ for all $t < t_0$, so again from continuity, we get $D_s\dot{\gamma}_{t_0}(s) = 0$. The same logic implies that the sectional curvatures and the second fundamental form continue to vanish at vectors tangent to points in the image of γ_{t_0}, and that $\gamma_{t_0}|_{[u_0,u_1]}$ is minimizing, so (2) holds at t_0. But then, (1) and (2) hold on $[0, t_0 + \varepsilon(t_0)]$, where $t_0 + \varepsilon(t_0) > t_0$, a contradiction. Thus, (1) and (2) must hold on all of $[0, \infty)$.

So, we assume (1) and (2) hold for $t \in [0, \ell]$. We define the function $f : \mathbb{R} \times UNS \to \mathbb{R}$ as

$$f(r, p, V) = d(p, P(\exp_p((\ell + r)V)) \tag{2}$$

and we define $Mf : \mathbb{R} \to \mathbb{R}$ as

$$Mf(r) = \max\{f(r, p, V) \mid p \in S, V \in UN_pS\} \tag{3}$$

Note that the function $(p, V) \mapsto f(r, p, V)$ from the unit normal bundle UNS is continuous, for fixed r and ℓ. Since the unit normal bundle of a compact space is compact, if follows that it achieves its maximum, so Mf is well-defined.

Claim III.1. The function Mf is Lipschitz non-negative, and $Mf(0) = 0$ when we restrict r to a sufficiently small interval $[0, \delta]$.

Proof. We have assumed that (1) holds for $t \in [0, \ell]$, so we immediately have

$$P(\exp_p(\ell V)) = p \tag{4}$$

which immediately gives $Mf(0) = 0$. Let us consider r and r' in $[0, \delta]$, where we choose δ according to Lem. A.1, with respect to the fixed ℓ. Note that

$$\begin{aligned} |Mf(r) - Mf(r')| &= \left|\max d(p, P(\exp_p((\ell + r)V))) - \max d(p, P(\exp_p((\ell + r')V)))\right| &(5)\\ &\leq \left|d(p, P(\exp_p((\ell + r)V))) - d(p, P(\exp_p((\ell + r')V)))\right| &(6)\end{aligned}$$

where we taking p and V on the second line to be the pair which maximizes $(q, W) \mapsto f(r, q, W)$. The pair (p', V') which maximizes $(q, W) \mapsto f(r', q, W)$ will necessarily satisfy $f(r', p, V) \leq f(r', p', V')$, which gives the above inequality. From here, we have by triangle inequality followed by the fact that P is distance non-increasing, as well as Lem. A.1:

$$\begin{aligned} \left|d(p, P(\exp_p((\ell + r)V))) - d(p, P(\exp_p((\ell + r')V)))\right| &\leq d(P(\exp_p((\ell + r)V))), P(\exp_p((\ell + r')V)))) \\ &\leq d(\exp_p((\ell + r)V)), \exp_p((\ell + r')V))) &(7)\\ &\leq |r - r'| &(8)\end{aligned}$$

which immediately establishes the desired Lipschitz property of Mf on the interval $[0, \delta]$. □

From here, the idea is to prove that $Mf = 0$ on $[0, \delta]$, which will immediately establish (1) for the interval $[0, \ell + \delta]$, where $\delta > 0$, as we will have

$$d(p, P(\exp_p((\ell + r)V))) = 0 \Longrightarrow P(\exp_p((\ell + r)V)) = p \tag{9}$$

for all $p \in S$ and $V \in UN_pS$.

To accomplish this goal, we will demonstrate that Mf is monotone decreasing on $[0, \delta]$: it is non-negative, so this will imply it is 0, as $Mf(0) = 0$. To begin, we fix ℓ, and use the result above to fix some δ so that f is Lipschitz. In fact, since Mf is continuous (it is Lipschitz), we can always choose δ sufficiently small so that $Mf(r) < \text{inj}(S)$ as well, which we know is positive as S is compact (see Def. A.1). Since we established that the maximum of $f(r, p, V)$ over $(p, V) \in UNS$ is achieved, there exists some (p, V) such that $Mf(r) = f(r, p, V)$. Let $p' = P(\exp_p((\ell + r)V))$. The normal geodesic ball B in S of radius $\text{inj}(S)$ around p' is well-defined, and is equal to the metric ball of the same radius around p (this is Corollary 6.12 in Lee 2nd Edition [2]). Moreover,

$$d(p, P(\exp_p((\ell + r)V))) = f(r, p, V) = Mf(r) < \text{inj}(S) \tag{10}$$

so it follows that $p \in B$. Therefore, we can choose some $q \in B$ such that $q \neq p$ and p lies inside the minimizing, radial geodesic from p' to q in B. We let $\gamma : [a, b] \to M$ denote this geodesic, and let u_0 and u_1 denote the times such that $\gamma(u_0) = p$ (not p'!) and $\gamma(u_1) = q$ (note that $b = u_1$). Also note that because S is totally convex, we can guarantee that γ is a minimizing geodesic not just in S, but in the ambient manifold M. We also extend the vector $V \in UN_pS$ to a parallel vector field along all of γ, which we also denote by V. By Lem. A.2, and the fact that S is totally-geodesic, the vector field V will remain normal to S, so we can say that $V(s) \in UN_{\gamma(s)}S$ for all s.

Let us consider the variation through geodesics $\Gamma : [a, b] \times \mathbb{R} \to M$ given by

$$\Gamma(s, t) = \exp_{\gamma(s)}(tV(s)) \tag{11}$$

Note that the curves $s \mapsto \Gamma(s,t)$ precisely satisfy the criterion for (2) of Thm. III.1. Thus, by our assumption, for all $t \in [0,\ell]$, the curves $s \mapsto \Gamma(s,t)$ are geodesics and the union of their images forms a flat, totally-geodesic strip, with each curve being minimizing on the interval $s \in [a,b]$. In particular, $s \mapsto \Gamma(s,\ell)$ is minimizing. The idea is to use our knowledge of the behaviour of these geodesics in the time interval $t \in [0,\ell]$ to extend our knowledge to the case that $t \in [\ell, \ell+r]$. To do this, we first prove the following claim.

Claim III.2. Let Γ be the variation through geodesics of Eq. (11). The vector field $s \mapsto \partial_t\Gamma(s,t)$ along each curve $s \mapsto \Gamma(s,t)$ (where $t \in [0,\ell]$ is fixed) is parallel. Moreover, it is perpendicular to the velocity vector $s \mapsto \partial_s\Gamma(s,t)$ along each of these curves. In addition, the length of each curve $s \mapsto \Gamma(s,t)$ (for $t \in [0,\ell]$) is constant.

Proof. Recall the result in Lee (Proposition 7.5 Lee 2nd Edition [2]) which relates the commutator of covariant derivatives of a variation field to curvature. In particular, we have

$$D_t D_s \partial_t \Gamma = D_s D_t \partial_t \Gamma - R(\partial_s\Gamma, \partial_t\Gamma)\partial_t\Gamma \tag{12}$$

We have already assumed that for $t \in [0,\ell]$, the curves $s \mapsto \Gamma(s,t)$ are contained in a region of 0 curvature. Moreover, since each curve $t \mapsto \Gamma(s,t)$ is itself a geodesic, it follows that $D_t\partial_t\Gamma = 0$. Therefore, the right-hand side of Eq. (12) is 0, implying $D_t D_s \partial_t \Gamma = 0$. This immediately implies that

$$D_s\partial_t\Gamma(s,t) = D_s\partial_t\Gamma(s,0) = D_s V(s) = 0 \tag{13}$$

as V is parallel along γ. This immediately implies our first claim, that the vector fields $s \mapsto \partial_t\Gamma(s,t)$ are parallel. As for the second claim, note that

$$\frac{d}{dt}\langle \partial_t\Gamma(s,t), \partial_s\Gamma(s,t)\rangle = \langle D_t\partial_t\Gamma(s,t)\rangle + \langle \partial_t\Gamma(s,t), D_t\partial_s\Gamma(s,t)\rangle \tag{14}$$

We know from the symmetry lemma that

$$D_t\partial_s\Gamma(s,t) = D_s\partial_t\Gamma(s,t) = 0 \tag{15}$$

Therefore, Eq. (14) reduces to 0, which means that

$$\langle \partial_t\Gamma(s,t), \partial_s\Gamma(s,t)\rangle = \langle \partial_t\Gamma(s,0), \partial_s\Gamma(s,0)\rangle = \langle V(s), \dot\gamma(s)\rangle = 0 \tag{16}$$

as γ lies in S and V takes values in the normal bundle to S. To prove the last claim, we simply note that since $D_t\partial_s\Gamma(s,t) = 0$, the length of the velocity vectors $\partial_s\Gamma(s,t)$ remain constant as we vary t, so the length of these curves will not change either. □

This result, while somewhat innocent, allows us to apply a powerful result of the Berger-Rauch comparison theorem, which is presented as Corollary 3.4 in Ref. [5]. For further details, see Appx. B. For now, let us state the claim:

Claim III.3. Let M be a complete manifold of non-negative curvature, let $p, q \in M$, and let $\gamma : [a,b] \to M$ be a minimizing geodesic connecting them. Suppose W is a parallel vector field along γ which is perpendicular to $\dot\gamma$. If we let $\Gamma : [a,b] \times \mathbb{R} \to M$ be a variation through geodesics of the form $\Gamma(s,t) = \exp_{\gamma(s)}(tW(s))$, then for all t,

$$L(\Gamma(\cdot,t)) \leq L(\Gamma(\cdot,0)) \tag{17}$$

with equality at some $t_0 > 0$ implying that the points $p, q, \Gamma(a,t_0), \Gamma(b,t_0)$ bound a flat, totally-geodesic rectangle, and that the curves $s \mapsto \Gamma(s,t)$ are minimizing geodesics for $t \in [0,t_0]$.

Let us apply Claim. III.3 to the geodesic $s \mapsto \Gamma(s,\ell)$ for $s \in [u_0,u_1]$ (note that this is a segment of the minimizing geodesic on $[a,b]$, so it is a minimizing geodesic), and the vector field $W(s) = \partial_t\Gamma(s,\ell)$ (as a result of Claim. III.2, we can in fact use this vector field). In particular, we can define a variation $\Lambda(s,t) = \exp_{\Gamma(s,\ell)}(t\partial_t\Gamma(s,\ell))$ to which we can apply the above claim, and note that $\Lambda(s,0) = \Gamma(s,\ell)$ and $\partial_t\Lambda(s,0) = \partial_t\Gamma(s,\ell)$, which means that $\Lambda(s,t) = \Gamma(s,t+\ell)$, by uniqueness of geodesics. It then it immediately follows from the result that for $t \in [\ell, \ell+r]$, $L(\Gamma(\cdot,t)) \leq L(\Gamma(\cdot,\ell))$, with equality implying the curves $s \mapsto \Gamma(s,t)$ (for $t \in [\ell,\ell+r]$) are in fact geodesics filling a flat, totally geodesic strip.

Let $q' = P(\exp_q((\ell+r)V))$. It follows that we have:

$$\begin{aligned} d(p',q') &\leq d(\exp_p((\ell+r)V), \exp_q((\ell+r)V)) && (18)\\ &\leq L(\Gamma(\cdot,\ell+r)|_{[u_0,u_1]}) \leq L(\Gamma(\cdot,\ell)|_{[u_0,u_1]}) && (19)\\ &= L(\Gamma(\cdot,0)|_{[u_0,u_1]}) = d(p,q) && (20)\end{aligned}$$

where, on the last line, we are using the fact that the lengths $L(\Gamma(\cdot,t))$ are constant for $t \in [0,\ell]$. On the other hand, we also know that

$$d(q,q') \leq f(r) = d(p,p') \tag{21}$$

And finally, since p', p and q all lie on a minimizing geodesic, the triangle inequality is saturated, and we have

$$d(p,p') + d(p,q) = d(p',q) \leq d(p',q') + d(q,q') \tag{22}$$

If we take the sum of Eq. (18) and Eq. (21), we get $d(p',q') + d(q,q') \leq d(p,p') + d(p,q)$. Therefore, comparing with Eq. (22), these inequalities must be equalities. In particular, we have

$$L(\Gamma(\cdot,\ell+r)|_{[u_0,u_1]}) = L(\Gamma(\cdot,\ell)|_{[u_0,u_1]}) \tag{23}$$

so the curves $s \mapsto \Gamma(s,t)$, for $t \in [\ell,\ell+r]$, on the interval $[u_0,u_1]$, are minimizing geodesics filling a flat, totally geodesic strip.

From here, consider small δ, and note that

$$\begin{aligned} Mf(r-\delta) &\geq d(q, P(\exp_q((\ell+r-\delta)V))) && (24)\\ &\geq d(p',q) - d(p', P(\exp_q((\ell+r-\delta)V))) && (25)\\ &\geq d(p',q) - d(\exp_q((\ell+r-\delta)V), \exp_p((\ell+r)V)) && (26)\end{aligned}$$

where we are using the reverse triangle inequality and the fact that P is distance non-increasing. From here, let us consider the quantity $d(\exp_q((\ell+r-\delta)V), \exp_p((\ell+r)V))$. The following inequality is a consequence of Toponogov's comparison theorem, which we prove in Appx. C. For now, let us simply state it:

$$d(\exp_q((\ell+r-\delta)V), \exp_p((\ell+r)V)) \leq \sqrt{d(\exp_q((\ell+r)V), \exp_p((\ell+r)V))^2 + \delta^2} \tag{27}$$

Using this result, we note that (using $d(p,q) > 0$ as $p \neq q$):

$$\begin{aligned} \sqrt{d(\exp_q((\ell+r)V), \exp_p((\ell+r)V))^2 + \delta^2} &= \sqrt{d(p,q)^2 + \delta^2} && (28)\\ &= d(p,q)\sqrt{1 + \frac{\delta^2}{d(p,q)^2}} && (29)\\ &\leq d(p,q) + \frac{\delta^2}{d(p,q)} && (30)\end{aligned}$$

so if we return to Eq. (24), we find that

$$Mf(r-\delta) \geq d(p',q) - d(p,q) - \frac{\delta^2}{d(p,q)} = d(p',p) - \frac{\delta^2}{d(p,q)} = Mf(r) - \frac{\delta^2}{d(p,q)} \tag{31}$$

where we are once again using the fact that p, p' and q all lie on a minimizing geodesic to saturate the triangle inequality. It follows from this fact that Mf is monotonely decreasing on $[0,\delta]$. Thus, $Mf = 0$ on this interval.

Remark III.1. We have proved that point (1) of Thm. III.1 holds on the interval $[0, \ell + \varepsilon(\ell)]$ for some $\varepsilon(\ell) > 0$. In order to show that (2) also holds on such an interval, consider some geodesic $\gamma : [a, b] \to M$ in S and a parallel unit normal vector field $V \in \Gamma(UNS)$ along γ. Since V takes values in the normal bundle to S, it will be perpendicular to $\dot{\gamma}$. We have that (2) holds for $t \in [0, \ell]$. Let p be a point in the geodesic, define p' exactly as we did before, with our vector field V, noting that *any* choice of unit normal will maximize Mf, as it is identically 0. We know that $p = p'$, so we can simply choose any q such that $\gamma_{[u_0,u_1]}$ is minimizing, with $\gamma(u_0) = p$ and $\gamma(u_1) = q$. We can then essentially re-do the above proof, applying Claim. III.3 to the minimizing geodesic $\sigma = \gamma|_{[u_0,u_1]}$. In particular, we define q' as we did before: we now know that $q = q'$ (as we have proved (1)), so by the collection of inequalities from Eq. (18) to Eq. (23), we conclude that $L(\sigma) = L(\exp_{\sigma(\cdot)}(tV))$ for $t \in [0, \ell + r]$. Thus, by Claim. III.3, the curves $s \mapsto \exp_{\sigma(s)}(tV)$ are minimizing geodesics filling a flat totally geodesic strip for $t \in [0, \ell + r]$. We can apply this local argument to any point p in γ, which gives the desired result. In particular, we have shown that (2) holds on an interval $[0, \ell + \varepsilon(\ell)]$ for some $\varepsilon(\ell) > 0$.

We have proved Thm. III.1! Let us now finish by proving the soul conjecture:

Corollary III.1.1 (Perelman)**.** The soul conjecture is true.

Proof. It was shown by Sharafutdinov in Ref. [6] that a distance non-increasing retraction $P : M \to S$ from M to a soul of M, S, does exist. If S is a soul, note that $\exp : NS \to M$ will be a surjection (see Ref. [7] for an explanation). Therefore, if $\dim(S) > 0$, we may choose a normal parallel vector field V along some non-trivial geodesic γ lying in S such that at point p in the geodesic, $\exp_p(tV) = p_0$: the point of M with positive sectional curvatures, for some $t \in [0, \infty)$. However, note from (2) in Thm. III.1 that the variation $\gamma_t(s) = \exp_{\gamma(s)}(tV)$ will fill a flat, two-dimensional plane containing p_0. Thus, *some* sectional curvature at p_0 will be 0, which is a contradiction.

It follows that $\dim(S) = 0$, so the soul is a point, and the proof is complete. □

A. Miscellaneous definitions and results

Lemma A.1. If M is a complete Riemannian manifold, S is a compact submanifold, and $\ell \in \mathbb{R}$ is fixed, then there exists some interval $[0, \delta]$ such that if $r, r' \in [0, \delta]$, then for any $(p, V) \in UNS$,

$$d(\exp_p((\ell + r)V), \exp_p((\ell + r')V)) \leq |r - r'| \tag{A1}$$

Proof. Fix ℓ, note that the set $[0, 1] \times UNS$ is compact, so the image of the map $(r, p, V) \mapsto \exp_p((\ell + r)V)$ is compact in M as well. We call this set K. For each $p \in K$, pick a uniformly normal geodesic ball around p, pick a Lebesgue number δ' for this cover, so that if $d(\exp_p((\ell+r)V), \exp_p((\ell+r')V)) < \delta'$, then $\exp_p((\ell+r)V)$ and $\exp((\ell + r')V)$ will both be contained in one of the geodesic balls.

Since $[0, 1] \times UNS$ is a compact metric space, $(r, p, V) \mapsto \exp_p((\ell + r)V)$ is uniformly continuous, so we can choose some δ where

$$|r - r'| < \delta \implies d(\exp_p((\ell + r)V), \exp_p((\ell + r')V)) < \delta' \tag{A2}$$

as the metric distance between (r, p, V) and (r', p, V) will reduce to $|r - r'|$ if we endow $[0, 1] \times UNS$ with a standard product metric. Thus, if $r, r' \in [0, \delta]$, then $\exp_p((\ell + r')V)$ is in a geodesic ball centred at $\exp_p((\ell + r)V)$. In particular, $\exp_p((\ell + r')V)$ is given by $(r' - r)W$ in normal coordinates for this geodesic ball, where $W = \dot{\gamma}_V(\ell + r)$ with γ_V the geodesic with initial velocity V and position p. Clearly, this is a unit vector as geodesics are constant speed. Thus,

$$d(\exp_p((\ell + r)V), \exp_p((\ell + r')V)) = ||(r' - r)W|| = |r' - r| \tag{A3}$$

as desired. □

Definition A.1 (Injectivity radius)**.** Let (M, g) be a Riemannian manifold, we define $\text{inj}(p)$, the injectivity radius at point $p \in M$ to be the supremum over all $r > 0$ such that $\exp_p : B_r(0) \to \exp_p(B_r(0))$ is a diffeomorphism. We define $\text{inj}(M)$ to be $\inf_{p \in M} \text{inj}(p)$. Note that if S is compact, then the injectivity radius of S will be positive.

Lemma A.2. Let (M, g) be a Riemannian manifold, let S be a totally geodesic submanifold. Then if $V_0 \in N_pS$, the normal bundle of S, its parallel transport along any curve in N will remain normal to N.

Proof. We can write $T_pM = T_pS \oplus N_pS$. Let us pick an orthonormal frame for T_pS, which we call $E_1, \ldots, E_k$. Let $V \in N_pS$. We can parallel transport the vectors $E_1, \ldots, E_k$ relative to the connection on S, yielding $E_1^S(t), \ldots, E_k^S(t)$, which will be an orthonormal frame for S at each point along the curve. We can parallel transport V relative to the connection on M to get $V(t)$. Note that because M is totally geodesic, from Gauss' formula, $D_t^S = D_t$ (i.e. the covariant derivative along γ relative to the connection on S is the same as the covariant derivative along γ relative to the connection on M). It follows that

$$\frac{d}{dt}\langle V(t), E_j^S(t)\rangle = \langle D_t V, E_j^S(t)\rangle + \langle V, D_t E_j^S(t)\rangle = \langle D_t V, E_j^S(t)\rangle + \langle V, D_t^S E_j^S(t)\rangle = 0 \tag{A4}$$

so that V remains orthogonal to all E_j^S at subsequent times. In particular, V remains normal to S, as it is orthogonal to a basis for $T_{\gamma(t)}S$ at each $\gamma(t)$ along the curve. □

B. Berger-Rauch comparison theorem

In this section, we will review the crucial comparison theorem of Berger and Rauch. The quoted theorems are taken from Ref. [5]. To begin, let us make a few remarks about the Jacobi equation:

$$D_t^2 J + R(J, \dot\gamma)\dot\gamma = 0 \tag{B1}$$

where γ is a geodesic and J is a vector field along γ. It is standard result that this equation precisely characterizes variations through geodesics (i.e. the variation field of a variation through geodesics will satisfy this equation, and if a vector field satisfies this equation, it is the variation field of some variation through geodesics centred at γ). Define $A(t) \in T^*M \otimes TM$, a $\binom{1}{1}$-tensor, as

$$A(t)V = R(V, \dot\gamma(t))\dot\gamma(t) \tag{B2}$$

so that we can rewrite the Jacobi equation as $D_t^2 J + A(t)J = 0$. This form of the Jacobi equation will be useful to us in stating the Berger-Rauch comparison theorem. However, before we do this, let us briefly comment on $A(t)$, thought of as a section of the endomorphism bundle of TM along the curve γ. Note that

$$\langle V, A(t)V\rangle = \langle V, R(V, \dot\gamma)\dot\gamma\rangle = Rm(V, \dot\gamma, \dot\gamma, V) = K(V, \dot\gamma) \tag{B3}$$

where Rm is the curvature tensor and K is the sectional curvature. In the case that we know the sectional curvature is always non-negative, we have $\langle V, A(t)V\rangle \geq 0$, implying $A(t)$ is a *positive operator*.

Let us now state Berger-Rauch in full generality:

Theorem B.1 (Berger-Rauch)**.** Let M be a Riemannian manifold, let $A_j(t)$ be sections of the endomorphism bundle of TM along the curve γ, for $j = 1, 2$. Suppose $\lambda_-(A_1(t)) \geq \lambda_+(A_2(t))$ for all t, that is, the smallest eigenvalue of $A_1(t)$ is greater than or equal to the largest eigenvalue of $A_2(t)$ for all t. Then, if J_j are solutions to the ODEs $D_t^2 J_j + A_j(t)J_j = 0$, and we have

$$D_t J_j(0) = 0 \quad \text{and} \quad ||J_1(0)|| = ||J_2(0)|| \tag{B4}$$

it follows that $||J_1(t)|| \leq ||J_2(t)||$ for all t up until the first 0 of J_1.

This result will allow us to immediately arrive at the following corollary:

Corollary B.1.1. If M is a manifold of non-negative sectional curvature and J is a solution to the Jacobi equation with $D_tJ(0) = 0$, then $||J(t)|| \leq ||J(0)||$ up to the first 0 of J.

Proof. Let $A(t) = R(\cdot, \dot\gamma(t))\dot\gamma(t)$ be the Jacobi endomorphism discussed previously, note that by our remarks, since M has non-negative sectional curvature, it follows that $A(t)$ is a positive operator, so $\lambda_-(A(t)) \geq 0$. Note that J satisfies $D_t^2J + A(t)J = 0$. The constant I clearly satisfies the necessary ODE with $A_j(t) = 0$, so we immediately have the required inequality of eigenvalues. The result follows immediately from Berger-Rauch. □

To conclude this section, we may provide a partial proof of Claim. III.3:

Partial Proof of Claim III.3. Each curve $t \mapsto \Gamma(s,t) = \exp_{\gamma(s)}(tW(s))$ is a geodesic, so that Γ is a variation through geodesics. Clearly, for each fixed s, $t \mapsto \partial_s\Gamma(s,t)$ is the variation field of a variation through geodesics, so it solves the Jacobi equation. Moreover,

$$D_t\partial_s\Gamma(s,t)|_{t=0} = D_s\partial_t\Gamma(s,t)|_{t=0} = D_sW(s) = 0 \tag{B5}$$

as W is parallel. It follows that $||\partial_s\Gamma(s,t)|| = ||\partial_s\Gamma(s,0)|| = ||\dot\gamma(s)||$. Since M is a manfold of non-negative sectional curvature, we may apply the previous corollary to conclude that

$$L(\Gamma(\cdot,t)) = \int_a^b ||\partial_s\Gamma(s,t)||\; ds \leq \int_a^b ||\dot\gamma(s)||\; ds = L(\gamma) \tag{B6}$$

which proves the first part of the claim. The second part of this claim is somewhat more arduous to prove, and relies on some other comparison results presented in Ref. [5]. Therefore, I will omit this part of the proof and refer readers to Section 3 of Ref. [5] for more details. □

C. Toponogov's comparison theorem

In this section, we will review Toponogov's comparison theorem: a very powerful result, and will use it to prove Eq. (27). To begin, let us state Toponogov's comparison theorem for the case we wish to compare to Euclidean space (i.e. constant 0-curvature):

Theorem C.1 (Euclidean Toponogov)**.** Let M be a Riemannian manifold with sectional curvature K satisfying $K \geq 0$ everywhere. Let $p, q, r \in M$ be points connected pairwise by geodesics, where γ_{pq}, the geodesic connecting p and q, is minimal. Let $x, y, z \in \mathbb{R}^n$ be points forming a triangle in Euclidean space, such that the lengths of the sides xy and xz are equal to the lengths of γ_{pq} and γ_{pr} respectively, and such that the angle at p is equal to the angle at x. Then $d(q,r) \leq ||y - z||$.

This will allow us to prove Eq. 27:

Proof of Eq. (27). We will apply Euclidean Toponogov to the points $\exp_q((\ell+r)V), \exp_q((\ell+r-\delta)V)$, and $\exp_p((\ell+r)V)$. In particular, we proved that the curve $s \mapsto \Gamma(s, \ell+r)$ on the interval $s \in [u_0, u_1]$ is a geodesic connecting points $\exp_q((\ell+r)V)$ and $\exp_p((\ell+r)V)$. Moreover, for δ small, we know that the radial geodesic connecting $\exp_q((\ell+r)V)$ and $\exp_q((\ell+r-\delta)V)$ is minimizing, and has length δ. Finally, to check the angle at point $\exp_q((\ell+r)V)$, we must compute the angle between the velocity vectors of these two geodesics. However, Claim. III.2 immediately implies that they are orthogonal.

Thus, from Euclidean Toponogov and the Pythagorean theorem for a right-angled triangle,

$$d(\exp_q((\ell+r-\delta)V), \exp_p((\ell+r)V)) \leq \sqrt{d(\exp_q((\ell+r)V), \exp_p((\ell+r)V))^2 + \delta^2} \tag{C1}$$

which is precisely the desired inequality. □

[1] Grisha Perelman. Proof of the soul conjecture of cheeger and gromoll. *Journal of Differential Geometry*, 40(1):209–212, 1994.

[2] John M Lee. *Introduction to Riemannian manifolds*, volume 2. Springer, 2018.

[3] Zhou Keshu. Soul theorem and soul conjecture. 2021.

[4] Jeff Cheeger and Detlef Gromoll. On the structure of complete manifolds of nonnegative curvature. *Annals of Mathematics*, 96(3):413–443, 1972.

[5] Jost-Hinrich Eschenburg. Comparison theorems in riemannian geometry. 1994.

[6] Vladimir Al'tafovich Sharafutdinov. The pogorelov-klingenberg theorem for manifolds homeomorphic to rn. *Siberian Mathematical Journal*, 18(4):649–657, 1977.

[7] Peter Michor (https://mathoverflow.net/users/26935/peter michor). Surjectivity of the normal exponential map. MathOverflow. URL:https://mathoverflow.net/q/131489 (version: 2013-05-22).