Jacobi fields, conjugate points and some applications

Michael Law

1 Introduction

Geodesics tell us a great deal about the global structure of a Riemannian manifold. We have already seen an example of this with the Hopf-Rinow Theorem, and in this report, we do further justice to this claim by exploring yet another aspect of this. Namely, we will see how one determines, to some extent, which geodesics are length-minimising.[1] The main results we build up to are Theorem 4.1 and Corollary 4.2, which roughly say that when a manifold satisfies a uniform sectional curvature upper bound C, then

- If $C < 0$, then all geodesics are minimising; if $C = 1/R^2 > 0$ with $R > 0$, then all geodesics are minimising for at least length πR;
- Geodesics on manifolds with negative sectional curvatures tend to spread out; geodesics on manifolds with positive sectional curvatures cannot converge too quickly.

These results are formulated using the notion of *conjugate points* along geodesics, which are defined using the machinery of *Jacobi fields*. We will introduce these in Sections 2 and 3.

Using curvature bounds to discover properties about a Riemannian manifold is a key theme in Riemannian geometry. To demonstrate how successful this approach has been, we devote Section 5 to explaining some applications of injectivity radius estimates that originate from the results of Section 4.

Our main references are [1] and [2]. To suit our discussion, we alter the presentation and rebalance the level of detail as appropriate.

1.1 Notation and conventions

In all that follows, (M, g) is a Riemannian manifold. We will write $\langle u, v\rangle$ to mean $g(u, v)$, which causes no ambiguity since the metric will always be fixed. Also $|v|^2 := \langle v, v\rangle$. The Levi-Civita connection on M is ∇. The Riemann curvature tensor is defined by

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z,$$

which differs by a sign from the convention of the course. We let $\mathrm{Rm}(X, Y, Z, W) = \langle R(X, Y)Z, W\rangle$. The sectional curvature is defined for $X, Y \in TM$ in the same tangent space (alternatively $X, Y \in \Gamma(TM)$) by

$$K(X \wedge Y) = \frac{\mathrm{Rm}(X, Y, Y, X)}{\langle X, X\rangle\langle Y, Y\rangle - \langle X, Y\rangle^2}.$$

Date: June 25, 2021.

[1] We already know geodesics are locally minimising, but exactly *how* local remains a mystery!

2 Jacobi fields

The main piece of machinery driving the analysis to come is that of Jacobi fields, and we introduce these in this section. We first recall some notions from variational calculus.

Definition 2.1. *Let $\gamma : [a,b] \to M$ be a smooth curve in M. A **variation** of γ is a map $\Gamma : (-\varepsilon, \varepsilon) \times [a,b] \to M$ such that the curve $\Gamma_s(t) := \Gamma(s,t)$ is smooth for each $s \in (-\varepsilon, \varepsilon)$, the curve $\Gamma^{(t)}(s) := \Gamma(s,t)$ is smooth for each $t \in [a,b]$, and $\Gamma(0,\cdot) = \gamma$. A variation Γ is called **proper** if additionally $\Gamma_s(a) = \gamma(a)$ and $\Gamma_s(b) = \gamma(b)$ for all s. The **variation field** of the variation Γ is the vector field $V : [a,b] \to TM|_{\gamma([a,b])}$ along γ defined by $V(t) = \partial_s \Gamma(0,t)$.*

Note that the variation field V is a vector field on the pullback bundle $\gamma^* TM$ over $[a,b]$. Evidently, if Γ is a proper variation then its variation field is proper, i.e. $V(a) = V(b) = 0$.

Before proceeding, we need the following key identities.

Lemma 2.2. *Let Γ be a variation of γ as above. Write $S(s,t) = \partial_s \Gamma(s,t)$ and $T(s,t) = \partial_t \Gamma(s,t)$. Then*

(a) (Symmetry Lemma) $\nabla_s T = \nabla_t S$.

(b) If V is a smooth vector field along Γ, then $\nabla_s \nabla_t V - \nabla_t \nabla_s V = R(S,T)V$, where ∇_s and ∇_t are covariant derivative operators for vector fields along Γ.

Proof. In local coordinates $\{x^i\}$ around an arbitrary point $(s,t) \in (-\varepsilon, \varepsilon) \times [a,b]$, write

$$\Gamma(s,t) = (x^1(s,t), \ldots, x^n(s,t)).$$

Then $S = \partial_s \Gamma = \frac{\partial x^k}{\partial s} \partial_k$ and $T = \partial_t \Gamma = \frac{\partial x^k}{\partial t} \partial_k$, and the covariant derivative formulas now read[2]

$$\nabla_s T = \left(\frac{\partial^2 x^k}{\partial s \partial t} + \frac{\partial x^i}{\partial t} \frac{\partial x^j}{\partial s} \Gamma^k_{ji} \right) \partial_k \qquad \text{and} \qquad \nabla_t S = \left(\frac{\partial^2 x^k}{\partial t \partial s} + \frac{\partial x^i}{\partial s} \frac{\partial x^j}{\partial t} \Gamma^k_{ji} \right) \partial_k.$$

Swapping the roles of i and j, and using the fact that $\Gamma^k_{ij} = \Gamma^k_{ji}$ by symmetry of the Levi-Civita connection on M, we see that these two equations are exactly the same. This proves (a). For (b), recall that ∇_s and ∇_t are formally the pullback connection on $\Gamma^* TM$ over $(-\varepsilon, \varepsilon) \times [a,b]$, obtained from the Levi-Civita connection on M by pulling back with Γ. Noting that $S = D\Gamma(\partial_s)$ and $T = D\Gamma(\partial_t)$, applying the definition of the pullback connection gives

$$R(S,T)V = \nabla_s \nabla_t V - \nabla_t \nabla_s V - \nabla_{[\partial_s, \partial_t]} V = \nabla_s \nabla_t V - \nabla_t \nabla_s V,$$

since $[\partial_s, \partial_t] = 0$. □

If γ has a unit-speed parametrisation and Γ is a proper variation of γ, then we get a family of curves $\Gamma_s(t) := \Gamma(s,t)$. Since we will mostly be dealing with geodesics only, we give the second variation formula for geodesics, which expresses the second derivative (with respect to s) of the length of Γ_s at $s = 0$ in terms of the variation field and γ itself.

[2] cf. our Lecture 10A, for example.

Proposition 2.3. *Let $\gamma : [a, b] \to M$ be a unit-speed geodesic in M, and Γ a proper variation of γ with variation field V. Let $V^\perp$ be the component of V orthogonal to $\dot{\gamma}$. The second variation formula is*

$$\frac{d^2}{ds^2}\Big|_{s=0} L[\Gamma_s] = \int_a^b \left(|\nabla_t V^\perp|^2 - \mathrm{Rm}(V^\perp, \dot{\gamma}, \dot{\gamma}, V^\perp) \right) dt. \tag{1}$$

Remark. *There is also the first variation formula, given for an arbitrary smooth curve γ by*

$$\frac{d}{ds}\Big|_{s=0} L[\Gamma_s] = -\int_a^b \langle V, \nabla_t \dot{\gamma} \rangle \, dt.$$

We see that γ is a geodesic, i.e. $\nabla_t \dot{\gamma} = 0$, if and only if its first variation with respect to any variation field V vanishes, which is why geodesics are 'locally minimising'.

The proof of Proposition 2.3 uses the identities of Lemma 2.2, but we omit it since it is just a long calculation. Anyhow, the proof of Proposition 2.4 below will also see the use of Lemma 2.2.

The overarching principle of our analysis later is to investigate how nearby geodesics on M relate to each other, and infer from this substantive information about M itself. As one may expect, this can be done using a variation Γ of a geodesic, but we also want each curve Γ_s of the variation to be a geodesic, i.e. Γ is a **geodesic variation**. The following lemma gives an explicit equation which geodesic variations satisfy.

Proposition 2.4. *Let γ be a geodesic, and Γ a variation of γ with variation field V. If Γ is a geodesic variation, then V satisfies the **Jacobi equation***

$$\nabla_t \nabla_t V + R(V, \dot{\gamma})\dot{\gamma} = 0. \tag{2}$$

*Conversely, if a vector field J along γ satisfies the Jacobi equation, i.e. J is a **Jacobi field**, then J is the variation field of a geodesic variation of γ.*

Proof. For the forward direction, let S and T be defined as in Lemma 2.2. Note that $\nabla_t T = 0$ since Γ is a geodesic variation i.e. each Γ_s is a geodesic. Then using the results of Lemma 2.2,

$$0 = \nabla_s \nabla_t T = \nabla_t \nabla_s T + R(S, T)T = \nabla_t \nabla_t S + R(S, T)T.$$

Evaluating at $s = 0$ (and keeping t free), we have $S(0, t) = V(t)$ and $T(0, t) = \dot{\gamma}(t)$ from the definitions of S and T. Substituting into the above gives (2), where the LHS is a function of t.

For the converse direction, let $J = J(t)$ be a given Jacobi field, starting at $t = 0$ say. We find a geodesic variation Γ of γ whose variation field is J. Define the curve $\sigma(s) = \exp_{\gamma(0)}(sJ(0))$, and let W be a vector field along σ with initial conditions $W(0) = \dot{\gamma}(0)$ and $\nabla_s W(0) = \nabla_t J(0)$. Such W exists by the theory of ODEs, e.g. by passing to a chart and formulating this as an initial value problem. Then let

$$\Gamma(s, t) = \exp_{\sigma(s)}(tW(s)),$$

defined for small enough s. This is indeed a geodesic variation, since fixing any s gives a geodesic parametrised by t. The variation field $V = V(t)$ of Γ is then a Jacobi field by the first part of this proposition. We compute

$$V(0) = \partial_s \Gamma(s, t)|_{s,t=0} = \partial_s \Gamma(s, 0)|_{s=0} = \partial_s \left(\exp_{\sigma(s)}(0) \right) \Big|_{s=0} = \partial_s \sigma(s)|_{s=0} = \sigma'(0) = J(0), \text{ and}$$

$$\nabla_t V(0) = \nabla_t V(t)|_{t=0} = \nabla_t S|_{s,t=0} = \nabla_s T|_{s,t=0} = \nabla_s \left(\partial_t \Gamma(s, t)|_{t=0} \right) \Big|_{s=0} = \nabla_s W(s)|_{s=0} = \nabla_t J(0).$$

Here we used Lemma 2.2 and the defining properties of W. Hence, the Jacobi fields V and J agree at $t = 0$, and their covariant derivatives along γ agree at $t = 0$ also. By the uniqueness of Jacobi fields with initial conditions (see below), we have $V = J$. So J is the variation field of the geodesic variation Γ. □

Let us state some basic facts about Jacobi fields. If $\gamma : [a, b] \to M$ is a geodesic segment, then for any $X, Y \in T_{\gamma(a)}M$ there exists a unique Jacobi field J along γ satisfying the initial conditions

$$J(a) = X, \quad \nabla_t J(a) = Y.$$

This follows from converting the Jacobi equation (2), a system of n second-order linear ODEs, into a system of $2n$ first-order linear ODEs, and applying the Picard-Lindelof theorem for solutions to initial value problems like what we did for the geodesic equation. Here of course Y is really an element of $T_X T_{\gamma(a)}M$, which we identify directly with $T_{\gamma(a)}M$. As a result of this, we have an isomorphism from the $\mathbb{R}$-vector space $\mathcal{J}(\gamma)$ of Jacobi fields along γ to $T_{\gamma(a)}M \oplus T_{\gamma(a)}M$ given by $J \mapsto (J(a), \nabla_t J(a))$. So $\mathcal{J}(\gamma)$ is a $2n$-dimensional $\mathbb{R}$-linear subspace of the space of smooth vector fields along γ.

Along any geodesic γ there are always two trivial Jacobi fields, $J(t) = \dot\gamma(t)$ and $J(t) = t\dot\gamma(t)$. They are 'trivial' because they are the variation fields of the variations $\Gamma(s, t) = \gamma(s+t)$ and $\Gamma(s, t) = \gamma(e^s t)$ respectively, and the curves Γ_s generated by these variations are merely reparametrisations of γ. Since these tell us nothing about how nearby geodesics behave, we are only interested in the **normal** Jacobi fields J, which satisfy $J(t) \perp \dot\gamma(t)$ for all t. Meanwhile, the two trivial Jacobi fields turn out to be a basis for the **tangential** Jacobi fields i.e. $J(t)$ is a scalar multiple of $\dot\gamma(t)$. So the tangential Jacobi fields span a 2-dimensional subspace of $\mathcal{J}(\gamma)$, while the normal ones span a $(2n - 2)$-dimensional subspace of $\mathcal{J}(\gamma)$.

A useful example for later concerns the situation of constant sectional curvature. In this case, we can derive all normal Jacobi fields along a geodesic $\gamma : [0, T] \to M$ which vanish at $t = 0$.

Lemma 2.5. *Suppose M has constant sectional curvature C, and $\gamma : [0, T] \to M$ is a unit speed geodesic segment. The normal Jacobi fields vanishing at $t = 0$ are precisely the vector fields $J(t) = u(t)E(t)$ where E is a parallel normal vector field along γ, and $u(t)$ is given by*

$$u(t) = \begin{cases} t & C = 0; \\ R\sin\left(\frac{t}{R}\right) & C = 1/R^2 > 0; \\ R\sinh\left(\frac{t}{R}\right) & C = -1/R^2 < 0. \end{cases} \tag{3}$$

Proof. We quote a formula (see e.g. [2]) for the Riemann curvature tensor when the sectional curvature is a constant C:

$$R(X, Y)Z = C(\langle Y, Z\rangle X - \langle X, Z\rangle Y).$$

Substituting this into the Jacobi equation (2), any Jacobi field J satisfies

$$0 = \nabla_t\nabla_t J + C(\langle\dot\gamma, \dot\gamma\rangle J - \langle J, \dot\gamma\rangle\dot\gamma) = \nabla_t\nabla_t J + CJ,$$

since $|\dot\gamma|^2 = 1$ and $\langle J, \dot\gamma\rangle = 0$. We make an ansatz that $J(t) = u(t)E(t)$ for a parallel normal vector field E along γ and some real-valued function u; plugging this into the above equation yields

$$0 = \nabla_t\nabla_t(uE) + CuE = \nabla_t(\dot uE + u\nabla_t E) + CuE = \ddot uE + \dot u\nabla_t E + CuE = (\ddot u + Cu)E,$$

using the fact that $\nabla_t E = 0$ by parallelity. Hence, as long as u is a solution to the initial value problem

$$\ddot{u} + Cu = 0, \quad u(0) = 0,$$

the vector field $J(t) = u(t)E(t)$ along γ is a normal Jacobi field with $J(0) = 0$. One checks that (3) uniquely solves this differential equation. Since the possibilities for E span an $(n-1)$-dimensional $\mathbb{R}$-vector space (corresponding to the $n-1$ normal directions at each point on γ), and the $\mathbb{R}$-vector space of normal Jacobi fields along γ vanishing at $t = 0$ is also $(n-1)$-dimensional,[3] it follows that the Jacobi fields of the form $J(t) = u(t)E(t)$ generate all the normal Jacobi fields vanishing at $t = 0$. □

3 Conjugate points: when do geodesics fail to minimise?

In this section, we introduce the concept of conjugate points and show how they relate to the failure of geodesics to minimise lengths on a nonlocal scale.

Definition 3.1. *Let γ be a geodesic segment joining $p, q \in M$, $p \neq q$. We say q is* ***conjugate to p along*** *γ if there is a Jacobi field along γ vanishing at p and q, but is not identically zero. Note that conjugacy is a symmetric relation.*

Example 3.2. *Let S^n_R be the round sphere, and take $\gamma : [0, \pi R] \to S^n_R$ to be a unit-speed geodesic from the north to south pole on S^n_R. There is a parallel, nonzero normal vector field E along γ (e.g. the 'rotation' vector field), so by Lemma 2.5 there is a Jacobi field along γ given by $J(t) = R\sin(t/R)$, using the fact that S^n_R has constant sectional curvature $C = 1/R^2 > 0$. Since J vanishes at both $t = 0$ and $t = \pi R$, it follows that the south pole is conjugate to the north pole. By symmetry, every point on S^n_R is conjugate to its antipode. In fact, one can show that the only point conjugate to a given point on S^n_R is its antipode.*

Combining Definition 3.1 with Proposition 2.4, we see that conjugacy between p and q implies the existence of a proper geodesic variation of γ keeping p and q fixed. Hence, geodesics 'close' to γ emanating from p also meet at q. This is indeed what happens on the round sphere; in fact the sphere is more special since *all* geodesics through a point meet at its antipode.

There are two important interpretations of conjugate points. The first one says that conjugate points are precisely those points where the exponential map fails to be a local diffeomorphism.

Theorem 3.3. *Fix $p \in M$ and $V \in T_pM$, and let $q = \exp_p(V)$. Then $\exp_p$ is a local diffeomorphism at V if and only if q is not conjugate to p along $\exp_p(tV)$, $t \in [0, 1]$.*

We will not pursue this interpretation further, as the second interpretation hits the nail on the head in terms of what we set out to explore: it says that geodesics do not minimise past conjugate points.

Theorem 3.4. *If a geodesic segment γ starting at p has an interior point conjugate to p, then γ is not minimising.*

[3] From Page 3, the normal Jacobi fields span a $(2n-2)$-dimensional space. In view of the isomorphism $\mathcal{J}(\gamma) \to T_{\gamma(0)}M \oplus T_{\gamma(0)}M$ mentioned there, it follows immediately that the normal Jacobi fields vanishing at $t = 0$ span an $(n-1)$-dimensional space.

We aim to prove Theorem 3.4, but we first need a definition. Recall that a vector field V along a curve $\gamma : [a, b] \to M$ is *proper* if $V(a) = V(b) = 0$, and *normal* if $\langle V(t), \dot{\gamma}(t)\rangle = 0$ for all t.

Definition 3.5. *Let $\gamma : [a, b] \in M$ be a geodesic segment in M. The* ***index form*** *I is a symmetric bilinear form defined on the space of proper, normal vector fields (not necessarily smooth) along γ by*

$$I(V, W) = \int_a^b (\langle \nabla_t V, \nabla_t W\rangle - \mathrm{Rm}(V, \dot{\gamma}, \dot{\gamma}, W))\, dt.$$

When γ has a unit-speed parametrisation and Γ is a proper variation of γ with variation field V, the quantity $I(V^\perp, V^\perp)$ is nothing but the second variation of $L(\Gamma_s)$, given in equation (2.3). The first variation of $L(\Gamma_s)$ vanishes because γ is a geodesic, so the following is immediate.

Proposition 3.6. *With the setup of the above paragraph, the geodesic segment γ is minimising if $I(V^\perp, V^\perp) \geq 0$ for any proper vector field V along γ. Equivalently, since the tangential component of a proper variation field does not affect length, γ is minimising if $I(V, V) \geq 0$ for any proper* normal *vector field V along γ.*

Proof of Theorem 3.4. Suppose $\gamma : [0, T] \to M$ is a unit-speed geodesic with $\gamma(0) = p$, and suppose $\gamma(a)$ is conjugate to p along γ for some $0 < a < T$. We exhibit a proper normal vector field X along γ such that $I(X, X) < 0$, and the result follows at once from Proposition 3.6.

The conjugacy between p and $\gamma(a)$ implies the existence of a nontrivial normal Jacobi field J along $\gamma|_{[0,a]}$ vanishing at $t = 0$ and $t = a$.[4] We define a vector field V along γ by

$$V(t) = \begin{cases} J(t), & 0 \leq t \leq a; \\ 0, & a \leq t \leq T. \end{cases}$$

This is obviously proper, normal and piecewise smooth. Now construct a new vector field W along γ as follows. Assign $W(a) = -\nabla_t J(a)$, which is orthogonal to $\dot{\gamma}(a)$ since J is normal. Extend to a normal and smooth (not just piecewise smooth) vector field along γ within a chart around $\gamma(a)$ using an adapted orthonormal frame, and finally extend to a proper, normal and smooth vector field W along all of γ using a bump function. Note that $W(a) \neq 0$, otherwise $\nabla_t J(a) = 0$ and $J(a) = 0$ which implies $J \equiv 0$ by uniqueness of Jacobi fields, a contradiction.

Let $\varepsilon > 0$ and set $X_\varepsilon = V + \varepsilon W$. Since I is bilinear,

$$I(X_\varepsilon, X_\varepsilon) = I(V, V) + 2\varepsilon I(V, W) + \varepsilon^2 I(W, W). \tag{4}$$

Now we compute $I(V, W)$ using Definition 3.5. In the following computation, we use the following facts: V is piecewise smooth for $t \in [0, a]$ and $t \in [a, T]$; V satisfies the Jacobi equation (2) on each

[4] The reason J is normal is as follows. Recall from Page 3 that the tangential Jacobi fields along γ are spanned by $\dot{\gamma}(t)$ and $t\dot{\gamma}(t)$. Since γ is parametrised with unit speed, neither of these can vanish at both $t = 0$ and $t = a$. So J cannot have a tangential component.

of these two intervals; and V vanishes at $t = 0, T$. And so we compute

$$\begin{aligned}
I(V, W) &= \int_0^T (\langle \nabla_t V, \nabla_t W \rangle - \mathrm{Rm}(V, \dot\gamma, \dot\gamma, W))\, dt \\
&= \left(\int_0^a + \int_a^T \right) (\langle \nabla_t V, \nabla_t W \rangle - \langle R(V, \dot\gamma)\dot\gamma, W \rangle)\, dt \\
&= \left(\int_0^a + \int_a^T \right) \left(\frac{d}{dt} \langle \nabla_t V, W \rangle - \langle \nabla_t \nabla_t V + R(V, \dot\gamma)\dot\gamma, W \rangle \right) dt \\
&= \langle \nabla_t V, W \rangle \Big|_{t=0}^{t=a} + \langle \nabla_t V, W \rangle \Big|_{t=a}^{t=T} - \left(\int_0^a + \int_a^T \right) \langle \nabla_t \nabla_t V + R(V, \dot\gamma)\dot\gamma, W \rangle\, dt \\
&= \left\langle \lim_{t \nearrow a} \nabla_t V(t) - \lim_{t \searrow a} \nabla_t V(t), W(a) \right\rangle \\
&= \left\langle \lim_{t \nearrow a} \nabla_t V(t), W(a) \right\rangle \\
&= \langle \nabla_t J(a), W(a) \rangle \\
&= -|W(a)|^2.
\end{aligned}$$

By virtually the same calculation we get

$$I(V, V) = \overset{\text{(as above)}}{\cdots} = \langle \nabla_t J(a), V(a) \rangle = 0,$$

since $V(a) = 0$. Substituting these computations into (4) now gives

$$I(X_\varepsilon, X_\varepsilon) = -2\varepsilon |W(a)|^2 + \varepsilon^2 I(W, W).$$

If we take ε small enough, the second term on the right is negligible and we get $I(X_\varepsilon, X_\varepsilon) < 0$. Since X_ε is also proper and normal along γ, this gives the desired vector field mentioned at the start of the proof. □

This interpretation of conjugate points is consistent with the sphere example from Example 3.2. We stated there that antipodes on a sphere S^n_R are conjugate. Indeed, all geodesic segments crossing the antipode are not minimising; it is shorter to go the 'other way'.

Remark. *The converse of Theorem 3.4 is not true. No two points on the cylinder $S^1 \times \mathbb{R}$ are conjugate, but clearly not all geodesic segments are minimising. To see the nonexistence of conjugate points, suppose γ is a geodesic segment on the cylinder, and J is a Jacobi field along γ. Lifting to the universal cover $\mathbb{R}^2$, the lifted Jacobi field $\widetilde{J}$ is a Jacobi field along $\widetilde{\gamma}$.*[5] *Since the Christoffel symbols vanish in $\mathbb{R}^2$, the Jacobi equation (2) reads $\widetilde{J}''(0)$. So $\widetilde{J}$ is a linear function, and for it to vanish at two distinct points it must be identically zero. But this means J is identically zero.*

[5] We do not prove this here, but we shall take it for granted!

4 Comparison theory of conjugate points

Theorem 3.4 says that conjugate points are obstructions to minimising geodesics, but can we say anything about the existence (or nonexistence) of conjugate points in the first place? The answer is yes – at least to some degree – when all sectional curvatures of M are bounded above by a constant $C \in \mathbb{R}$. Instead of building up the tension, we state and explain the result now, then prove it afterwards.

Theorem 4.1 (Jacobi Field Comparison Theorem)**.** *Suppose all sectional curvatures of (M, g) are bounded above by a constant C. If $\gamma : [0, T] \to M$ is a unit-speed geodesic and J is a normal Jacobi field along γ with $J(0) = 0$, then*

$$|J(t)| \geq \begin{cases} t|\nabla_t J(0)| & \text{whenever } t \geq 0, & \text{if } C = 0; \\ R\sin\left(\frac{t}{R}\right)|\nabla_t J(0)| & \text{whenever } 0 \leq t \leq \pi R, & \text{if } C = 1/R^2 > 0; \\ R\sinh\left(\frac{t}{R}\right)|\nabla_t J(0)| & \text{whenever } t \geq 0, & \text{if } C = -1/R^2 < 0. \end{cases} \tag{5}$$

It is interesting to note that the different cases for J correspond to the different cases for u in Proposition 2.5. This theorem implies obstructions to conjugate points in the following manner.

Corollary 4.2. *If all sectional curvatures of (M, g) are nonpositive, then there are no conjugate points along any geodesic. If all sectional curvatures are bounded above by a positive constant $C = 1/R^2 > 0$ where $R > 0$, then any pair of conjugate points along a geodesic must be separated by a distance of at least πR along that geodesic.*

Proof. If all sectional curvatures of (M, g) are nonpositive, then according to (5), the only way J can be zero at both $t = 0$ and some other $t > 0$ is if $\nabla_t J(0) = 0$. But then $J(0) = 0$ and $\nabla_t J(0) = 0$, which implies J is identically zero by the existence and uniqueness of Jacobi fields. So there are no conjugate points along γ. If all sectional curvatures are bounded above by $C = 1/R^2 > 0$, then since $\sin(t/R) > 0$ for $0 < t < \pi R$, the same reasoning holds. That is, by (5), the only way J can vanish at both $t = 0$ and some other $t \in (0, \pi R)$ is if $\nabla_t J(0) = 0$, so any conjugate point to $\gamma(0)$ along γ must occur at $t \geq \pi R$. □

Theorem 4.1 is called a *comparison theorem* is because it compares the geometry of M to the geometries of model spaces with constant sectional curvatures. These model spaces are $\mathbb{R}^n$, the n-spheres S^n_R and the hyperbolic n-spaces H^n_R, corresponding to $C = 0, C > 0$ and $C < 0$ respectively. It is known (again using Jacobi field techniques!) that all simply connected Riemannian n-manifolds with constant sectional curvatures are isometric to one of these three spaces.

Corollary 4.2 also admits an interpretation purely from the point of view of Jacobi fields and infinitesimally close geodesics. It says that if M has uniformly nonpositive sectional curvatures, nearby geodesics tend to spread out. Otherwise, we could construct a normal geodesic variation with distinct fixed points, i.e. conjugate points, which cannot exist by the Corollary. On the other hand, a positive upper bound on sectional curvature prevents nearby geodesics from meeting too quickly.

We now prove Theorem 4.1. We need the following technical ODE comparison result:

Lemma 4.3 (Sturm Comparison Theorem)**.** *Suppose u and v are differentiable real-valued functions on $[0, T]$, twice-differentiable on $(0, T)$, and $u > 0$ on $(0, T)$. Suppose further that u and v satisfy*

$$\ddot{u}(t) + a(t)u(t) = 0, \qquad \ddot{v}(t) + a(t)u(t) \geq 0, \qquad u(0) = v(0) = 0, \qquad \dot{u}(0) = \dot{v}(0) > 0, \tag{6}$$

for some function $a : [0,T] \to \mathbb{R}$. Then $v(t) \geq u(t)$ on $[0,T]$.

Proof of Theorem 4.1. Note that J is smooth, being the variation field of a geodesic variation which is necessarily smooth. So $|J(t)|$ is smooth whenever $J(t) \neq 0$. At such a point, we compute

$$\begin{aligned}
\frac{d^2}{dt^2}|J| &= \frac{d^2}{dt^2}\langle J, J\rangle^{1/2} = \frac{d}{dt}\frac{\langle \nabla_t J, J\rangle}{\langle J, J\rangle^{1/2}} \\
&= \frac{\langle \nabla_t\nabla_t J, J\rangle + \langle \nabla_t J, \nabla_t J\rangle}{\langle J, J\rangle^{1/2}} - \frac{\langle \nabla_t J, J\rangle^2}{\langle J, J\rangle^{3/2}} \\
&= \frac{-\langle R(J,\dot\gamma)\dot\gamma, J\rangle + |\nabla_t J|^2}{|J|} - \frac{\langle \nabla_t J, J\rangle^2}{|J|^3} && \text{(Jacobi equation for } J\text{)} \\
&\geq -\frac{\langle R(J,\dot\gamma)\dot\gamma, J\rangle}{|J|} + \frac{|\nabla_t J|^2}{|J|} - \frac{|\nabla_t J|^2|J|^2}{|J|^3} && \text{(Cauchy-Schwarz)} \\
&= -\frac{\mathrm{Rm}(J,\dot\gamma,\dot\gamma,J)}{|J|} \\
&= -|J|\cdot K(J\wedge\dot\gamma), && (\text{as } \langle J,\dot\gamma\rangle = 0, \langle\dot\gamma,\dot\gamma\rangle = 1)
\end{aligned}$$

where $K(J \wedge \dot\gamma)$ is the sectional curvature of the plane spanned by J and $\dot\gamma$. But the sectional curvatures are all bounded above by C by assumption, so

$$\frac{d^2}{dt^2}|J| \geq -C|J| \tag{7}$$

wherever $J \neq 0$. Meanwhile, the solutions u from Lemma 2.5 satisfy $\ddot u + C\dot u = 0$ for the same C, so the idea now is to use Lemma 4.3 (with $|J|$ playing the role of v) to establish a lower bound for $|J|$ in terms of u. Note that the first three conditions in (6) are automatically satisfied. For the last one to hold, we need $d|J|/dt = 1$ at $t = 0$, since $\dot u(0) = 1$ as one easily checks. This is in general not true, but we can replace J with $\widetilde{J}(t) := J(t)/|\nabla_t J(0)|$ and start again.[6] Being a positive scalar multiple of J, $\widetilde{J}$ is still a Jacobi field satisfying (7) and the first three conditions of (6). The last condition of (6) also holds, since by passing to normal coordinates around $\gamma(0)$ we get (treating J as a function in the normal chart)

$$\frac{d}{dt}\Big|_{t=0}|\widetilde{J}(t)| = \lim_{t\to 0}\frac{|J(t)| - |J(0)|}{t|\nabla_t J(0)|} = \lim_{t\to 0}\frac{|J(0) + t\partial_t J(0) + \mathcal{O}(t^2)| - |J(0)|}{t|\partial_t J(0)|} = 1.$$

Here we used the fact that the Christoffel symbols in normal coordinates around $\gamma(0)$ vanish at $\gamma(0)$, so $\nabla_t J(0) = \partial_t J(0)$ in these coordinates. Also $J(0) = 0$ by assumption.

Thus, Lemma 4.3 applies to $|\widetilde{J}|$, and we conclude $|\widetilde{J}| \geq u$ whenever $J \neq 0$. Since $d|\widetilde{J}|/dt = 1$ and $\widetilde{J} = 0$ at $t = 0$, it follows that $|\widetilde{J}| > 0$ on some interval $(0, \varepsilon)$. Now $|\widetilde{J}|$ cannot attain its first zero before u does, otherwise $0 < |\widetilde{J}| < u$ for some t close to the first zero, a contradiction. Therefore, $|\widetilde{J}| \geq u$ for all nonnegative t before u attains its first positive zero. This translates directly into the theorem, substituting u according to the three different cases in Lemma 2.5. □

5 Injectivity radius estimates and applications

The remainder of this report will outline some further applications, just to demonstrate how far this kind of analysis can go. Our starting point is a relatively straightforward consequence of Corollary

[6] Here we assumed $|\nabla_t J(0)|$ is nonzero. But if it were zero, then the theorem holds trivially.

4.2, which gives a lower bound on the *injectivity radius* of M when its sectional curvatures are bounded above by $C > 0$.

Definition 5.1. *The* ***injectivity radius*** *at a point* $p \in M$*, denoted* $\mathrm{inj}(p)$*, is the supremum of all* $R > 0$ *such that* $\exp_p : T_pM \to M$ *restricts to a diffeomorphism on* $B_R(0) \subset T_pM$*, the ball of radius* R *centred at 0. The* ***injectivity radius*** *of* M *is* $\mathrm{inj}(M) := \inf_{p \in M} \mathrm{inj}(p)$*.*

Corollary 5.2. *Suppose* M *has sectional curvatures bounded above by a positive constant* $C > 0$*. Then*

$$\mathrm{inj}(M) \geq \min \left\{ \frac{\pi}{\sqrt{C}}, \frac{\textit{length of shortest closed geodesic in } M}{2} \right\}.$$

Proof. Let $p \in M$. In the absence of conjugate points to p, it is clear that

$$\mathrm{inj}(p) \geq \frac{1}{2} \cdot (\text{length of shortest closed geodesic loop based at } p). \tag{8}$$

But if conjugate points exist, the injectivity radius can be forced to be smaller. Suppose $q \in M$ is a conjugate point to p along some geodesic, then $\exp_p$ is not a local diffeomorphism at q by Theorem 3.3. Hence, $\exp_p$ fails to be a diffeomorphism when restricted to the ball $B_{d(p,q)+\varepsilon}(0) \subset T_pM$, where $d(p, q)$ is the intrinsic distance between p and q in M, and $\varepsilon > 0$ is arbitrary. By Corollary 4.2, $d(p, q) \geq \pi/\sqrt{C}$. Therefore, $\mathrm{inj}(p) \geq \pi/\sqrt{C}$ when conjugate points exist, and this supersedes (8). Since we do not know whether conjugate points exist in general, we therefore settle for

$$\mathrm{inj}(p) \geq \min \left\{ \frac{\pi}{\sqrt{C}}, \frac{\text{length of shortest closed geodesic loop based at } p}{2} \right\}.$$

The result now follows from the definition of $\mathrm{inj}(M)$. □

Successive developments of this result lead to *Klingenberg's Lemma*:

Theorem 5.3 (Klingenberg)**.** *Suppose* (M, g) *is simply-connected and compact with dimension* ≥ 3*. If all sectional curvatures lie in the interval* $(1/4, 1]$*, then* $\mathrm{inj}(M) \geq \pi$*.*

Injectivity radius estimates find widespread use in geometric analysis. By combining this with the Rauch Comparison Theorem (see [2]), Berger and Klingenberg proved the celebrated Sphere Theorem in the 1960s:

Theorem 5.4 (Topological Sphere Theorem)**.** *Suppose* (M, g) *is simply-connected and compact with dimension* n*. If all sectional curvatures of* M *lie in the interval* $(1/4, 1]$*, then* M *is homeomorphic to an* n*-sphere.*

This is quite remarkable, as it greatly restricts the topology of M subject to mild uniform 'pinching' of sectional curvatures. However, there exist exotic spheres with differentiable structures nonequivalent to the standard one, and one would wonder whether Theorem 5.4 still holds if 'homeomorphic' is replaced with 'diffeomorphic'. This turns out to be true, but was not proved until 2007 by Brendle and Schoen [3]. Furthermore, their proof only requires a weaker condition of *pointwise* pinched sectional curvatures:

Theorem 5.5 (Differentiable Sphere Theorem)**.** *Suppose* (M, g) *is simply-connected and compact with dimension* n*. Suppose further that for each* $p \in M$*, there exists* $C > 0$ *such that all sectional curvatures of 2-planes in* T_pM *lie in the interval* $(C/4, C]$*. Then* M *is diffeomorphic to an* n*-sphere with the round metric.*

Their proof uses the *Ricci flow*, which is a one-parameter family of Riemannian metrics $g(t)$ on a manifold M satisfying

$$\frac{\partial}{\partial t} g(t) = -2\mathrm{Ric}_{\mathrm{g(t)}}.$$

The Ricci flow seeks to make the metric on M 'rounder' over time. It therefore makes sense to use Ricci flow to study geometries related to round spheres.

Speaking of Ricci flow, we describe yet another application of injectivity radius estimates which ties in with geometric flows more generally. When studying these flows, one is often interested in analysing the formation of singularities, classifying them, and perhaps developing robust methods to continue the flow past singularities. One technique is to 'magnify' around the singularity as the singular time T is approached; this is carried out by scaling the metrics $g(t)$ by a factor such as $1/(T-t)$. One hopes that this new flow, $(M, \tilde{g}(t))$ say, converges to a limiting Riemannian manifold $(M, \tilde{g})$ as $t \to T$, which we can use to study the singularity.

While this is not always possible in general, there is an Arzelà-Ascoli type theorem for the existence of convergent subsequences of Riemannian manifolds in the so-called *smooth Cheeger-Gromov topology*. This topology is defined on the set of *pointed Riemannian manifolds*; a pointed Riemannian manifold (M, g, p) is just a Riemannian manifold (M, g) with a choice of point $p \in M$. The exact definition is out of scope, but we emphasise the injectivity radius condition needed for this crucial convergence result to hold. This version of the theorem is Theorem 11.4 in [4].

Theorem 5.6 (Compactness theorem for Riemannian manifolds)**.** *Suppose that* $\{(M_k, g_k, p_k)\}_{k\in\mathbb{N}}$ *is a sequence of connected, complete, pointed Riemannian manifolds such that*

(i) For each $m \in \mathbb{N}$ *and each* $A < \infty$*, there exists* $C_{m,A} < \infty$ *and* $k_A \in \mathbb{N}$ *such that*

$$\sup_{B_A(x_k)} |{}^k\nabla^m \, \mathrm{Rm}_k \, | \leq C_{m,A} \quad \textit{for all } k \geq k_A,$$

where ${}^k\nabla$ *and* Rm_k *are the Levi-Civita connection and Riemann curvature tensor of* (M_k, g_k) *respectively, and* $B_A(x_k)$ *is the radius-A geodesic ball around* x_k *in* M_k*.* (This is analogous to the equibounded derivatives condition in Arzelà-Ascoli)*;*

(ii) There exists $\delta > 0$ *with* $\mathrm{inj}_k(x_k) \geq \delta$ *for every* $k \in \mathbb{N}$*, where* inj_k *is the injectivity radius function of* (M_k, g_k)*.*

Then $\{(M_k, g_k, p_k)\}_{k\in\mathbb{N}}$ *converges in the smooth Cheeger-Gromov topology to a complete pointed Riemannian manifold* $(M_\infty, g_\infty, p_\infty)$*.*

At least in the case of mean curvature flow, this is the formal foundation for singularity analysis (see e.g. [5]). The takeaway is that injectivity radius estimates, along with other concepts to do with Jacobi fields that we didn't cover, are absolutely indispensable in Riemannian geometry and geometric analysis.

References

[1] J. M. Lee. *Riemannian Manifolds: An Introduction to Curvature*, Graduate Texts in Mathematics, no. 176, Springer-Verlag, 1997.

[2] M. do Carmo. *Riemannian Geometry*, Birkhäuser, Basel, 1992.

[3] S. Brendle, R. Schoen. *Manifolds with* 1/4*-pinched curvature are space forms*, J. Amer. Math. Soc **22** (2009), no. 1, 287-307.

[4] B. Andrews, B. Chow, C. Guenther, M. Langford. *Extrinsic Geometric Flows*, Graduate Studies in Mathematics, no. 206, American Mathematical Society, 2020.

[5] G. Huisken. *Asymptotic behaviour for singularities of the mean curvature flow*, J. Diff. Geom. **31** (1990), 285-299.