An Introduction to Curve-Shortening and the Ricci Flow

Connor Mooney

May 17, 2011

Contents

1 Introduction

In this paper, we study the evolution of plane curves by curve-shortening and the evolution of Riemannian metrics on surfaces and three-manifolds under the Ricci flow. We focus on global theorems concerning the long-time existence and asymptotic shape of solutions. The results we describe have all appeared in research papers, and they are included because the author believes them integral to the exposition.

A one-parameter family of closed curves in the plane $\gamma(t)$ is said to evolve by the curve shortening flow for $0 \leq t < T$ if a parametrization $X(\cdot, t) : S^1 \times [0, T) \to \mathbb{R}^2$ satisfies the equation

$$\frac{\partial}{\partial t} X = \kappa N \tag{1}$$

where N is the inward pointing unit normal and κ is the signed curvature. In view of the first variation formula for arclength in the direction of a variation η, given by $-\int \langle \eta, \kappa N \rangle ds$, we see that the curve shortening flow is the gradient flow of length. We derive several basic and useful facts about curve-shortening in Section 2.

It is easy to verify that a one-parameter family of circles with radius $r(t) = \sqrt{2(T-t)}$, shrinking to a point in finite time $T = \frac{1}{2} r(0)^2$, form a self-similar solution to the flow. However, a self-intersecting curve like a lop-sided figure eight will generally develop a cusp before it can shrink to a point. We ask under what conditions solutions to the curve-shortening flow will shrink to points without developing singularities, and what is the limiting shape of such solutions. The following theorem of Grayson [8] provides a satisfying answer:

Theorem 1. *Let ρ and R denote the inradius and circumradius of a simple closed curve, and let M be an embedded closed curve in the plane. Then the curve-shortening flow will shrink M to a point in finite time, and M will become asymptotically circular in the sense that the ratio $\frac{\rho}{R}$ converges to* 1.

We discuss two approaches. The first approach, outlined in Section 3, is to prove the theorem assuming that M is convex and then prove that embedded curves become convex in finite time. These together imply Theorem 1. The convex case was proven by Gage and Hamilton in [6],[5], and [7], and Grayson showed that embedded curves become convex. This method relies on integral estimates and an interesting isoperimetric inequality known as the Bonneson inequality. In Section 4 we discuss an alternative approach, taken by Huisken [13], which is to show that a particular isoperimetric inequality "improves" under the flow thus ruling out any bad behavior.

In Section 5 we provide the necessary Riemannian geometry background for our treatment of the Ricci flow. Let M be a smooth manifold of dimension n. A one-parameter family of Riemannian metrics $g(t)$ on M is said to evolve by the normalized Ricci flow if

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

where $r(t)$ is the mean of the scalar curvature R at time t. We will see that on surfaces, this equation reduces to

$$\frac{\partial}{\partial t} g(t) = (r - R) g(t). \tag{3}$$

In Section 6, we discuss basic properties of the Ricci flow and derive the evolution equations it implies for the curvature quantities. We can then address long-time existence and asymptotic roundness results for the Ricci flow on the two sphere:

Theorem 2. *Under the normalized Ricci flow, any metric on S^2 converges to a metric of constant curvature.*

There is a striking similarity in the historical development of the curve-shortening flow and Ricci flow on S^2. Again, two approaches were taken. The first approach, outlined in Section 7, was to prove the theorem assuming that the scalar curvature R is initially positive, and then to show that R becomes positive in finite time. Hamilton [10] proved the positive curvature case, and Chow [4] subsequently showed that the curvature becomes positive. This approach relies heavily on the scalar maximum principle and a peculiar entropy monotonicity formula. In Section 8 we discuss another approach taken by Hamilton [11] which is again to show the improvement of an isoperimetric inequality which rules out the formation of singularities.

Finally, in Section 9 we discuss the Ricci flow on 3-manifolds. In higher dimensions, it is computationally convenient to consider the unnormalized Ricci flow

$$\frac{\partial}{\partial t} g(t) = -2\mathrm{Ric}_{g(t)}. \tag{4}$$

This equation differs from the normalized flow only by a scaling in space and a reparametrization in time, so it is not difficult to infer results about one from results about the other. We discuss three critical geometric estimates needed for the following famous theorem of Hamilton [9]:

Theorem 3. *Let M be a compact Riemannian 3-manifold which admits a metric of strictly positive Ricci curvature. Then M also admits a metric of constant positive curvature.*

In three-dimensional geometry, the scalar curvature is no longer the only meaningful curvature quantity, so one must obtain estimates for solutions to systems of equations. The essential tool developed for this purpose is the tensor maximum principle. Using this, we prove the preservation of positive Ricci curvature. This enables us to prove a pointwise pinching estimate for the eigenvalues of the Ricci tensor, and then obtain a bound on the gradient of the scalar curvature. Equipped with this bound, we can get global pinching of the eigenvalues of the Ricci tensor, and then sketch the remainder of the proof.

2 Curve-Shortening Basics

In this section, we establish notation and derive several basic facts about the curve-shortening flow, including the evolution equations for various geometric quantities and some explicit solutions. Recall that a family of closed, simple curves $\gamma(t)$ parametrized by $X(u,t) : S^1 \times [0,\tau) \to \mathbb{R}^2$ is a solution to the curve-shortening flow if it moves by its curvature vector, i.e.

$$\frac{\partial}{\partial t}X = \kappa N$$

where κ is the signed curvature and N is the inward-pointing unit normal. This implies evolution equations for various geometric quantities. The quantities of interest for us are length $L(t)$, the enclosed area $A(t)$ and the curvature $\kappa(u,t)$. We begin by deriving the evolution equation for length. Let s be the arclength parameter and let T be the unit tangent vector, so that $\frac{\partial}{\partial u}X = s'(u)T$. Then the length is

$$L(t) = \int ds = \int_{S^1} \left\langle \frac{\partial}{\partial u}X, \frac{\partial}{\partial u}X \right\rangle^{\frac{1}{2}} du.$$

Differentiating with respect to time and applying the evolution equation, we obtain

$$\begin{aligned} \frac{\partial}{\partial t}L &= \int_{S^1} \left\langle \frac{\partial^2}{\partial t \partial u}X, T \right\rangle du \\ &= \int_{S^1} \left\langle \frac{\partial^2}{\partial u \partial t}X, T \right\rangle du \\ &= \int_{S^1} \left\langle \frac{\partial}{\partial u}(\kappa N), T \right\rangle du \end{aligned}$$

where we commuted derivatives because u and t are independent variables. Recall that T and N are related by the Frenet formulas $\frac{\partial}{\partial u}T = s'(u)\kappa N$ and $\frac{\partial}{\partial u}N = -s'(u)\kappa T$. The tangential component of $\frac{\partial}{\partial u}(\kappa N)$ is thus $-\kappa^2 s'(u)$. Substituting, we get the evolution equation for length:

$$\frac{\partial}{\partial t}L = -\int \kappa^2 ds. \tag{5}$$

As remarked in the introduction, this is also easily derived from the first variation formula for length, which implies that if our curves move with normal velocity v the length changes like $-\int \kappa v ds$. We now compute the evolution of curvature. To that end, we take the time derivative of $\kappa N = \frac{\partial^2}{\partial s^2}X$ and commute derivatives of t and s in order to use the evolution equation for X. By the above computation, $\frac{\partial}{\partial t}ds = -\kappa^2 ds$. It follows that

$$\frac{\partial}{\partial t}\frac{\partial}{\partial s} = \frac{\partial}{\partial s}\frac{\partial}{\partial t} + \kappa^2 \frac{\partial}{\partial s}.$$

Using the commutator rule, we obtain

$$\begin{aligned}
\frac{\partial}{\partial t}(\kappa N) &= \frac{\partial}{\partial t}\frac{\partial^2}{\partial s^2}X \\
&= \frac{\partial}{\partial s}\frac{\partial^2}{\partial t \partial s}X + \kappa^2 \frac{\partial^2}{\partial s^2}X \\
&= \frac{\partial}{\partial s}\left(\frac{\partial}{\partial s}\frac{\partial}{\partial t}X + \kappa^2 \frac{\partial}{\partial s}X\right) + \kappa^3 N \\
&= \frac{\partial^2}{\partial s^2}(\kappa N) + \frac{\partial}{\partial s}(\kappa^2 T) + \kappa^3 N \\
&= \frac{\partial}{\partial s}\left(\frac{\partial \kappa}{\partial s}N - \kappa^2 T\right) + \frac{\partial}{\partial s}(\kappa^2 T) + \kappa^3 N \\
&= \left(\frac{\partial^2}{\partial s^2}\kappa + \kappa^3\right) N - \kappa \frac{\partial \kappa}{\partial s} T.
\end{aligned}$$

The left hand side can also be written $\frac{\partial \kappa}{\partial t}N + \kappa \frac{\partial}{\partial t}N$, and the time derivative of N is tangential because N is a unit vector. Equating components, we get the curvature evolution equation

$$\frac{\partial \kappa}{\partial t} = \frac{\partial^2 \kappa}{\partial s^2} + \kappa^3. \tag{6}$$

We also get the useful information

$$\frac{\partial}{\partial t}N = -\frac{\partial \kappa}{\partial s}T. \tag{7}$$

We use this to compute the evolution of area. By Green's theorem, we have $2A = \int \langle X, N\rangle ds$. Thus, we compute

$$\begin{aligned}
2\frac{\partial}{\partial t}A &= -\int \left(\left\langle \frac{\partial}{\partial t}X, N\right\rangle + \left\langle X, \frac{\partial}{\partial t}N\right\rangle - \kappa^2 \langle X, N\rangle\right) ds \\
&= -\int \left(\kappa - \frac{\partial \kappa}{\partial s}\langle X, T\rangle - \kappa^2 \langle X, N\rangle\right) ds \\
&= -\int \big(\kappa + \kappa(\langle T, T\rangle + \langle X, \kappa N\rangle) - \kappa^2 \langle X, N\rangle\big)\, ds \\
&= -2\int \kappa ds \\
&= -4\pi
\end{aligned}$$

where we integrated by parts in the third step so that the third and fourth terms cancel. Hence, area evolves as

$$\frac{\partial}{\partial t}A = -2\pi. \tag{8}$$

Remark 1. *Suppose that a family of curves parametrized by $X(u,t)$ satisfies the equation $\frac{\partial}{\partial t}X = \kappa(u,t)N + g(u,t)T$, so the the normal velocity is still κN. Then we may transform away the tangential component via a spatial reparametrization $u \to v(u,t)$, under which κ will not change, being a geometric quantity.*

With this remark in mind, we derive an interesting graph solution to the curve-shortening flow moving by translation with constant speed. Such a solution has the form $(x, t + y(x))$, with velocity $(0, 1)$. For a graph we have $N = \frac{(-y'(x),1)}{\sqrt{1+y'(x)^2}}$ and $\kappa = \frac{y''(x)}{(1+y'(x)^2)^{\frac{3}{2}}}$. We only require that the normal component of the velocity is κ, yielding the ODE

$$y''(x) = 1 + y'(x)^2.$$

Integrating twice, we get a particular solution $y(x) = -\log(\cos(x))$, so our solution to the curve-shortening flow is

$$y(x, t) = t - \log(\cos(x)). \tag{9}$$

This solution is aptly named the Grim Reaper, and we will use it for comparison purposes in Section 4.

3 Convex Curves Shrink to Points

The main theorem of this section is the following:

Theorem 4. *If M is a convex curve embedded in the plane, the curve-shortening flow shrinks M to a point in finite time. M remains convex and becomes circular in the sense that the ratio $\frac{\rho}{R}$ approaches* 1.

We divide the proof into two parts. In the first part, we prove longtime existence following the arguments in Gage and Hamilton [7]. This relies mostly on integral estimates. In the second part, we prove the asymptotic circularity of solutions following Gage [6],[5]. This section relies critically on an interesting isoperimetric inequality known as the Bonneson inequality.

3.1 Long-time Existence

The primary advantage of working with convex curves is that we can parametrize by the angle θ our tangent vector makes with the x-axis. This is the content of the following theorem.

Theorem 5. *The curve-shortening flow for convex curves is equivalent to the following PDE problem: Find $\kappa : S^1 \times [0,T) \to \mathbb{R}$ satisfying*

1. $\frac{\partial \kappa}{\partial t} = \kappa^2 \frac{\partial^2 \kappa}{\partial \theta^2} + \kappa^3$
2. $\kappa(\theta, 0) = \psi(\theta)$ *where* $\psi > 0$ *and* $\int_0^{2\pi} \frac{\cos(\theta)}{\psi(\theta)} d\theta = \int_0^{2\pi} \frac{\sin(\theta)}{\psi(\theta)} d\theta = 0.$

Proof. Assume first that we have a solution to the curve-shortening flow, where the initial curve is convex and closed. Then condition 2 is automatically satisfied. To prove condition 1, we change variables from (s,t) to $(\theta(s,t), \tau = t)$. Differentiating with respect to t corresponds to keeping s fixed, and likewise differentiating with respect to τ corresponds to keeping θ fixed. By equation 6 the curvature evolves according to the equation

$$\begin{aligned}
\frac{\partial \kappa}{\partial t} &= \frac{\partial^2 \kappa}{\partial s^2} + \kappa^3 \\
&= \frac{\partial^2 \kappa}{\partial \theta^2} \left(\frac{\partial \theta}{\partial s} \right)^2 + \frac{\partial \kappa}{\partial \theta} \frac{\partial^2 \theta}{\partial s^2} + \kappa^3 \\
&= \kappa^2 \frac{\partial^2 \kappa}{\partial \theta^2} + \frac{\partial \kappa}{\partial \theta} \frac{\partial \kappa}{\partial s} + \kappa^3.
\end{aligned}$$

We can write $N = (-\sin(\theta(t)), \cos(\theta(t)))$, so that $\frac{\partial}{\partial t} N = -\frac{\partial \theta}{\partial t} T$. Equation 7 then implies that $\frac{\partial \theta}{\partial t} = \frac{\partial \kappa}{\partial s}$. Substituting, we obtain

$$\frac{\partial \kappa}{\partial t} = \kappa^2 \frac{\partial^2 \kappa}{\partial \theta^2} + \kappa^3 + \frac{\partial \kappa}{\partial \theta} \frac{\partial \theta}{\partial t}.$$

Finally, by the chain rule, the last term is $\frac{\partial \kappa}{\partial t} - \frac{\partial \kappa}{\partial \tau}$, which gives us the desired evolution equation

$$\frac{\partial \kappa}{\partial \tau} = \kappa^2 \frac{\partial^2 \kappa}{\partial \theta^2} + \kappa^3. \tag{10}$$

Conversely, given the function $\psi(\theta)$ in condition 2, we may define a curve by $x(\theta) = \int_0^\theta \frac{\cos(z)}{\psi(z)} dz$ and $y(\theta) = \int_0^\theta \frac{\sin(z)}{\psi(z)} dz$. It is straightforward to check that this curve is closed ($x(0) = y(0) = x(2\pi) = y(2\pi) = 0$), has curvature ψ, and is convex ($\psi > 0$). One may easily check that the curves represented by $\kappa(\theta, t)$ remain closed by computing

$$\frac{\partial}{\partial t} \int_{S^1} \frac{\cos(\theta)}{\kappa(\theta, t)} d\theta = \frac{\partial}{\partial t} \int_{S^1} \frac{\sin(\theta)}{\kappa(\theta, t)} d\theta = 0,$$

a consequence of the evolution equation for κ. Finally, one checks that the curves represented by κ have normal velocity equal to κ, completing the proof. See [7] for details. □

Corollary 1. *Convexity is preserved under the curve-shortening flow.*

Proof. This is a direct application of the maximum principle. Let $\epsilon = \inf\{\kappa(\theta, 0) : \theta \in S^1\}$. By assumption, $\epsilon > 0$. Let $0 < \delta < \epsilon$. Suppose by way of contradiction that κ achieves the value δ for a first time t_0 at a point θ_0. Then $\kappa(\theta_0, t) \geq \delta$ for $t < t_0$, so $\frac{\partial \kappa}{\partial t}(\theta_0, t_0) \leq 0$. In addition, because this is the first time δ is achieved, $\kappa(\theta_0, t_0)$ is a local spatial minimum, so $\frac{\partial^2 \kappa}{\partial \theta^2}(\theta_0, t_0) \geq 0$. Applying the evolution equation, we obtain

$$\begin{aligned} 0 &\geq \frac{\partial \kappa}{\partial t}(\theta_0, t_0) \\ &= \delta^2 \frac{\partial^2 \kappa}{\partial \theta^2}(\theta_0, t_0) + \delta^3 \\ &> 0, \end{aligned}$$

giving us the desired contradiction. □

Theorem 6. *If $\kappa : S^1 \times [0, T)$ satisfies the conditions of Theorem 5 and the area of the curve it represents is bounded away from 0, then κ is uniformly bounded.*

Proof. It will be useful to consider the quantity

$$\kappa^*(t) = \sup\{b : \kappa(\theta, t) > b \text{ on an interval of length } \pi\}.$$

Lemma 1. *If A is bounded away from 0 on $[0, T)$, then κ^* is bounded on $[0, T)$.*

Intuitively, if curvature were as high as we liked on an interval of length π, then since the curvature has a positive lower bound, we could make the area arbitrarily small. Take any $M < \kappa^*(t)$ and translate our domain so that $\kappa(\theta, t) > M$ on $[0, \pi]$. Then the distance d between the points on a curve when $\theta = 0$ and $\theta = \pi$ is given by $\int_0^\pi \sin\theta ds = \int_0^\pi \frac{\sin\theta}{\kappa} d\theta \leq \frac{2}{M}$. It is clear by convexity that $A \leq dL$, so combining these two we have $M \leq \frac{2L}{A}$; since L is decreasing and A is bounded away from 0, we have that M is bounded. Since M may be chosen arbitrarily close to κ^*, this proves the lemma. We can now control the integral of the curvature in terms of κ^*:

Lemma 2. *Suppose κ^* is bounded on $[0, T)$. Then $\int_{S^1} \log\kappa(\theta, t) d\theta$ is bounded on $[0, T)$.*

The proof of this lemma is an integral estimate relying on the Wirtinger inequality. Applying the evolution equation and using integration by parts, we obtain

$$\frac{\partial}{\partial t}\int_{S^1} \log \kappa d\theta = \int_{S^1} \{\kappa^2 - \left(\frac{\partial \kappa}{\partial \theta}\right)^2\} d\theta.$$

Fix t, and let $O = \{\theta : \kappa(\theta, t) > \kappa^*(t)\}$. Note that O is open, so we can write it as a countable disjoint union of open intervals I_k of length less that π. Note that $\kappa - \kappa^* = 0$ at the endpoints of the I_k. Extending this function to be 0 on an interval of length π containing I_k and translating, we can take its Fourier series $\sum \sin n\theta$ and apply Parseval's identity to get the estimate

$$\int_{I_k} (\kappa - \kappa^*)^2 d\theta \leq \int_{I_k} \left(\frac{\partial \kappa}{\partial \theta}\right)^2 d\theta$$

which we can rearrange to obtain

$$\begin{aligned}
\int_O \left(\kappa^2 - \left(\frac{\partial \kappa}{\partial \theta}\right)^2\right) d\theta &\leq 2\kappa^* \int_O \kappa d\theta \\
&\leq 2\kappa^* \int_{S^1} \kappa^2 ds \\
&= -2\kappa^* \frac{\partial L}{\partial t}.
\end{aligned}$$

Away from O, the curvature is bounded above by κ^*. Using this and the above estimate, we obtain

$$\begin{aligned}
\frac{\partial}{\partial t}\int_{S^1} \log \kappa d\theta &= \int_{S^1} \{\kappa^2 - \left(\frac{\partial \kappa}{\partial \theta}\right)^2\} d\theta \\
&\leq 2\pi(\kappa^*)^2 - 2\kappa^* \frac{\partial L}{\partial t}
\end{aligned}$$

Integrating and using that κ^* is bounded, we see that $\int_{S^1} \log \kappa d\theta$ remains bounded on $[0, T)$.

This integral estimate tells us that κ can only be large on small intervals; more precisely, for every $\delta > 0$, there exists a constant $C(\delta)$ such that $\kappa(\theta, t) \leq C$ except on intervals of length less that δ. Indeed, if this were not true, we could find some δ a sequence of intervals I_n of length δ on which $\kappa \geq n$, which would imply $\int_{S^1} \log \kappa d\theta \geq \delta \log n + (2\pi - \delta) \log \kappa_{min}(0)$, using the above maximum principle estimate. This is a contradiction for n large.

Knowing that κ can only be large on small intervals, we can control κ_{max} if we bound the derivative $\frac{\partial \kappa}{\partial \theta}$. For $\delta > 0$, we know that $\kappa \leq C(\delta)$ except on intervals of length less than δ; fixing t, suppose $\kappa_{max}(t) = \kappa(\theta_0, t)$ lies on such an interval, i.e. $\theta_0 \in [a, b]$ with $b - a \leq \delta$. Extend $[a, b]$ so that $\kappa(a, t) = C$. Then we have $\kappa_{max}(t) = C(\delta) + \int_a^{\theta_0} \left(\frac{\partial \kappa}{\partial \theta}\right) d\theta$. We may estimate the right hand side using the following computation:

$$\begin{aligned}\frac{\partial}{\partial t}\int_{S^1}\left(\kappa^2-\left(\frac{\partial\kappa}{\partial\theta}\right)^2\right) &= \int_{S^1}\left(2\kappa\frac{\partial\kappa}{\partial t}-2\frac{\partial\kappa}{\partial\theta}\frac{\partial^2\kappa}{\partial t\partial\theta}\right)d\theta\\ &= 2\int_{s^1}\frac{\partial\kappa}{\partial t}\left(\kappa+\frac{\partial^2\kappa}{\partial\theta^2}\right)d\theta\\ &= 2\int_{s^1}\kappa^2\left(\kappa+\frac{\partial^2\kappa}{\partial\theta^2}\right)^2 d\theta\\ &\geq 0.\end{aligned}$$

Integrating gives us $\int_{S^1}\left(\frac{\partial\kappa}{\partial\theta}\right)^2 d\theta \leq \int_{S^1}\kappa^2 d\theta + D$ for some $D>0$. We can now control the integral of $\frac{\partial\kappa}{\partial\theta}$ in terms of $\kappa_{max}(t)$, which does not depend on δ. Applying Cauchy-Schwartz, we obtain

$$\begin{aligned}\kappa_{max}(t) &= C(\delta)+\int_a^{\theta_0}\left(\frac{\partial\kappa}{\partial\theta}\right)d\theta\\ &\leq C(\delta)+\sqrt{\delta}\left(\int_{S^1}\left(\frac{\partial\kappa}{\partial\theta}\right)^2 d\theta\right)^{\frac{1}{2}}\\ &\leq C(\delta)+\sqrt{\delta}\left(\int_{S^1}\kappa^2 d\theta + D\right)^{\frac{1}{2}}\\ &\leq C(\delta)+\sqrt{\delta}\left(2\pi\kappa_{max}(t)^2+D\right)^{\frac{1}{2}}\\ &\leq C(\delta)+\sqrt{2\pi\delta}\kappa_{max}(t)+\sqrt{\delta D}.\end{aligned}$$

For δ sufficiently small we thus have $\kappa_{max}(t)\leq 2C(\delta)$ for all $t\in[0,T)$, proving the theorem. □

Equipped with an upper bound on κ, we can control all the higher derivatives of κ:

Theorem 7. *If κ is bounded, then all higher derivatives of κ are bounded.*

The proof of this theorem relies on integral estimates similar to those used above; for details, refer to [7]. We can now show that convex curves shrink to points under the curve-shortening flow without developing singularities.

Theorem 8. *Convex curves shrink to points.*

Proof. We assume short-time existence and uniqueness of solutions; see for instance Gage and Hamilton. If A is bounded away from 0 on $[0,T)$, then the previous two theorems imply that κ has a smooth limit as t approaches T, so we can extend the solution past T. □

3.2 Asymptotic Circularity

Knowing that convex curves shrink to points, we turn our attention to the asymptotic global shape of convex solutions to the curve-shortening flow. It is reasonable to believe, based on examples, that solutions become circular. The idea is to compare the inradius ρ and circumradius R of our curves. The following isoperimetric inequality, known as the Bonneson inequality, is particularly useful for such comparisons:

Theorem 9. *A closed, convex curve γ in the plane satisfies*

$$rL - A - \pi r^2 \geq 0$$

for $r \in [\rho, R]$ and

$$L^2 - 4\pi A \geq \pi^2 (R - \rho)^2.$$

Proof. We first give a proof of the theorem for convex polygons, due to Osserman [16]. Let P be a convex polygon of length L in the plane bounding a region D of area A. Now let $P(t)$ be the curve consisting of points a distance t from D, and denote by $D(t)$ the region it bounds, $A(t)$ its area, and $L(t)$ its length. Since P is convex, we obtain $P(t)$ by translating the sides of P out by distance t and connecting them by circular arcs. Hence,

$$A(t) = \pi t^2 + Lt + A.$$

If $t \in [\rho, R]$, then by the definition of ρ and R, any circle of radius t with center in $D(t)$ must have nonempty intersection with P. Consider the regions

$$E_k = \{p \in D(t) : C_t(p) \text{ intersects P in k points}\}$$

with areas A_k. We claim that

$$\sum_k kA_k = 4tL.$$

Observe that for a closed polygon, the centers of those circles of radius t which intersect P an odd number of times must have tangential intersections, and these form a set of measure zero in the plane, so we may ignore odd indices. It would follow, then, from this equation that

$$2A(t) = 2(A_2 + A_4 + ...) \leq \sum_k kA_k = 4tL$$

and, substituting our equation for $A(t)$, we can simplify to get

$$tL - A - \pi t^2 \geq 0$$

as desired.

The equation $\sum_k kA_k = 4tL$ actually holds for all polygons. We proceed inductively. If P is a line segment, then $D(t)$ decomposes into three parts: First, the region E_1 consisting of two disks of radius t centered at the endpoints of P minus their intersection E_0, and the region E_2 consisting of those points outside both disks but inside the rectangle of length L and width $2t$ bisected by P. We then have $A_1 = 2\pi t^2 - 2A_0$ (the region E_0 is double counted) and $A_2 = 2tL - \pi t^2 + A_0$ (the region E_0 is subtracted twice), giving the desired conclusion. For the inductive step, notice that by adding on another line segment, we can simply add the segment inequality to the assumed inequality. The right hand side $4tL$ is correct, and where the new regions intersect the old ones, they add 0, 1 or 2 to k, yielding the correct formula on the left hand side.

Finally, note that every smooth convex curve γ can be approximated arbitrarily well by convex polygons P_n with lengths $L_n \to L(\gamma)$ and areas $A_n \to A(\gamma)$. Since the Bonneson inequality holds for all of these polygons, it also holds for γ.

Finally, the second Bonneson inequality follows easily by manipulating the first Bonneson inequality to get

$$2\pi\rho \geq L - \sqrt{L^2 - 4\pi A}$$

and

$$2\pi R \leq L + \sqrt{L^2 - 4\pi A},$$

(implicitly using the isoperimetric inequality), and then subtracting the two. □

Remark 2. *The above theorem also holds for non-convex curves; it is easy to show the Bonneson inequality for $r = R$ using the fact that taking the convex hull of a curve doesn't change R, but the above technique will not work for ρ, which generally increases when we take the convex hull. See Osserman [16] for details on how to prove the general case. The second Bonneson inequality also shows that the isoperimetric inequality becomes an equality only when $R = \rho$, i.e. in the case of a circle.*

The first theorem hinting towards asymptotic circularity for convex curves is the improvement of the standard isoperimetric ratio.

Theorem 10. *The isoperimetric ratio $\frac{L^2}{A}$ is decreasing under the curve-shortening flow for convex curves.*

Proof. In view of the evolution equations for L and A, we compute

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{L^2}{A}\right) &= \frac{2L\frac{\partial L}{\partial t}}{A} - \frac{L^2\frac{\partial A}{\partial t}}{A^2} \\ &= -\frac{2L}{A}\left(\int \kappa^2 ds - \frac{\pi L}{A}\right).\end{aligned}$$

Thus, the proof of this theorem boils down to the following lemma;

Lemma 3. *A closed, convex curve γ in the plane satisfies*

$$\frac{\pi L}{A} \leq \int_\gamma \kappa^2 ds.$$

The quantity that will allow us to apply the Bonneson inequality is the length of the projection of X onto its tangent line, given by $p = -\langle X, N\rangle$. We can relate integrals of p and κ using the following relation, obtained via integration by parts:

$$\begin{aligned}\int p\kappa ds &= -\int \langle X, \kappa N\rangle ds \\ &= \int \langle T, T\rangle ds \\ &= L.\end{aligned}$$

Applying Cauchy-Schwartz, we have $L^2 \leq \left(\int p^2 ds\right)\left(\int \kappa^2 ds\right)$, so the desired conclusion will hold if we can prove that

$$\int p^2 ds \leq \frac{LA}{\pi}. \tag{11}$$

The strategy, used in [6], is to prove the conclusion first for convex curves which are symmetric through the origin, for which the quantity p can be controlled, and then extend this result to any convex curve using a clever trick. Observe that for convex curves symmetric through the origin, the breadth of the curve is twice the projection of X onto its tangent line, which in turn is $2p$. By convexity the breadth lies in the interval $[2\rho, 2R]$, so p lies in the interval on which we can apply the Bonneson inequality:

$$\pi p^2 \leq pL - A. \tag{12}$$

Integrating, and noting that $\int p ds = 2A$ by Green's theorem, the right hand side becomes LA, proving the theorem for convex curves symmetric through the origin.

Finally, we generalize to all convex curves. In this case, the object is to choose the origin cleverly. This amounts to choosing it so that we can apply our result for convex curves symmetric through the origin. Consider the set of line segments that bisect the lamina bounded by our curve. Each divides γ into two parts, γ_1 and γ_2, which together with the segment bound regions of area $\frac{A}{2}$. Suppose first that if we choose the center of one of these segments as our origin, the reflections of γ_1 and γ_2 through the origin are both convex. Then by the previous argument,

$$\int_{\gamma_i} p^2 ds \leq \frac{L_i A}{\pi}$$

for $i = 1,2$. Adding the two inequalities together gives the result.

We would like to find a bisecting segment such that the tangent vectors to γ at its endpoints are parallel, and choose our origin as the center of this segment; this way, we guarantee that the reflections of both γ_1 and γ_2 through the origin are convex. The convexity of γ ensures that we can parametrize by the angle θ of its tangent line. Let $X(\theta)$ parametrize γ, and let $A(\theta)$ be the area bounded by $X([\theta, \theta+\pi])$ and the line segment connecting $X(\theta)$ and $X(\theta+\pi)$. Suppose without loss of generality that $A(0) \leq \frac{A}{2}$. Then we must have $A(\pi) \geq \frac{A}{2}$, and by the continuity of $A(\theta)$ and the intermediate value theorem, there exists θ_0 such that $A(\theta_0) = \frac{A}{2}$, proving the claim. □

In the previous section we showed that that convex solutions to the curve-shortening flow shrink to points. Using this fact and improving the estimates used to prove the monotonicity of $\frac{L^2}{A}$, we can show [5] that the isoperimetric ratio must in fact go to that of a circle:

Theorem 11. *Let $\gamma(t)$ be a family of convex curves evolving by the curve-shortening flow for $0 < t < T$. If $\lim_{t \to T} A(t) = 0$, then the isoperimetric ratio $\frac{L^2}{A}$ approaches 4π.*

Knowing that area goes to 0 linearly and that the isoperimetric ratio is bounded below, it is intuitive in light of the evolution equation for $\frac{L^2}{A}$ that the quantity

$$L\left(\int \kappa^2 ds - \frac{\pi L}{A}\right) \tag{13}$$

should approach 0 along a subsequence of times $\{t_i\}$ approaching T. Indeed, if not, then in some δ-neighborhood of T, it is bounded below by some $\epsilon > 0$. It would follow that

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{L^2}{A}\right) &< -\frac{2\epsilon}{A} \\ &= -\frac{2\epsilon}{A(0) - 2\pi t}.\end{aligned}$$

Integrating, we have the isoperimetric ratio bounded above by a multiple of $\log(A(0) - 2\pi t)$, which blows down as t approaches T, contradicting the isoperimetric inequality. The idea is to exploit this information by obtaining a measure of circularity bounded by the difference between $L \int \kappa^2 ds$ and $\frac{\pi L^2}{A}$.

To that end, we sharpen our estimate on $\int p^2 ds$. Last time, we obtained our estimate by simply integrating the Bonneson inequality. Note that the function $rL - A - \pi r^2$ is concave, so it lies above the secant line connecting the graph at ρ and R. Using this fact, we obtain

$$\begin{aligned}rL - A - \pi r^2 &\geq (\rho L - A - \pi \rho^2) + (r - \rho)\left(\frac{(R-\rho)L - \pi(R^2 - \rho^2)}{R - \rho}\right) \\ &= \rho L - A - \pi\rho^2 + (r - \rho)\left(L - \pi(R + \rho)\right) \\ &= rL - A + \pi\rho R - \pi r(R + \rho) \\ &\geq 0\end{aligned}$$

for $r \in [\rho, R]$. For convex curves symmetric through the origin, p satisfies this condition. Integrating and noting that $\int p ds = 2A$, we get a better estimate for the integral of p^2:

$$LA - \pi \int p^2 ds \geq AL\left(1 + \frac{\pi\rho R}{L} - \frac{2\pi(\rho + R)}{L}\right).$$

Denote by $G(\gamma)$ the expression in parentheses, so that $\int p^2 ds \leq \frac{AL}{\pi}(1 - G)$. Observe that if $G(\gamma) = 0$ then our secant line is on the r-axis. This tells us that Bonneson's inequality is sharp at ρ and R, from which it follows that γ is a circle. Like before, we apply Cauchy-Schwartz to estimate κ, getting an inequality better by a factor of $1 - G$:

$$(1 - G)\int \kappa^2 ds \geq \frac{\pi L}{A}.$$

We can rearrange this and multiply by L to take advantage of the geometric information contained in G, getting

$$\begin{aligned}L\left(\int \kappa^2 ds - \frac{\pi L}{A}\right) &\geq GL\int \kappa^2 \\ &\geq G\frac{\pi L^2}{A} \\ &\geq 4\pi^2 G\end{aligned}$$

where we applied the weaker inequality for $\int \kappa^2$ in the second line and the isoperimetric inequality in the third. By our previous result, G goes to 0 at a subsequence of times approaching T. Since G is invariant under rescaling of the plane, it is not difficult to show (see, for example, Gage [5]) that the normalized laminae $\sqrt{\frac{\pi}{A}}\gamma(t)$ converge to the unit disk. Since the isoperimetric $\frac{L^2}{A}$ is also invariant under rescaling, it must converge to 4π.

To extend this result to arbitrary convex curves, one may use the same origin-choosing trick as before, i.e. take the center of a bisecting segment chosen such that the tangent vectors at its endpoints are parallel, and reduce to the previous case by reflecting the resulting regions over this origin.

Corollary 2. *The ratio* $\frac{\rho}{R}$ *converges to* 1.

Proof. By the Bonneson inequality and the inequality $\pi R^2 > A$,

$$\begin{aligned}\frac{L^2}{A} - 4\pi &\geq \frac{\pi^2(R-\rho)^2}{A} \\ &= \frac{\pi^2 R^2}{A}\left(1 - \frac{\rho}{R}\right)^2 \\ &\geq \left(1 - \frac{\rho}{R}\right)^2.\end{aligned}$$

The left hand side converges to 0 by the previous theorem, proving asymptotic circularity. □

Remark 3. *The convergence of* $\frac{\rho}{R}$ *to* 1 *can be considered "*C^0*" convergence to the circle. By generalizing our definition of* κ^* *from the longtime existence section to any angle and obtaining a geometric estimate, one can also show that the ratio of maximum to minimum curvature approaches* 1*, a type of "*C^2*" convergence. See [7] for a proof.*

4 Embedded Curves Shrink to Points

Grayson [8] proved that under the curve-shortening flow, embedded curves become convex. This result combined with the previous section would prove Theorem 1. However, here we discuss a shorter approach, taken by Huisken [13]. The idea is to identify an isoperimetric ratio which improves under the flow, ruling out all the ways singularities can form. We do not address the classification of singularities, which are closely connected to self-similar solutions such as the Grim Reaper, but we give the relevant results and references where they are needed.

Let $X : S^1 \times [0,T] \to \mathbb{R}^2$ be a smooth embedded solution of the curve-shortening flow. Let d be the extrinsic distance between two points on a curve, and let ψ be the distance between two points on a round circle of the same length with the same intrinsic distance l between points. Note that ψ is the same regardless of which intrinsic distance we choose, so it is well-defined; by elementary geometry, one obtains $\psi = \frac{L}{\pi}\sin(\frac{l}{L}\pi)$.

Theorem 12. *The minimum of $\frac{d}{\psi}$ on S^1 is nondecreasing, and is strictly increasing unless $\frac{d}{\psi}$ is identically* 1.

The strategy is to control the time derivative by applying the vanishing of first variations and nonnegativity of second variations at a local spatial minimum. To that end, we derive the spatial variation formulas. Since these are computed at a fixed time, we can parametrize our curve by arclength, $F(s)$. Suppose our minimum occurs at (p,q), and define ω as the unit vector in the direction of $F(p) - F(q)$. Let

$$d_{a,b,\epsilon} = \|F(p + a\epsilon) - F(q + b\epsilon)\|$$

and

$$\psi_{a,b,\epsilon} = \frac{L}{\pi}\sin\left(\frac{l(p + a\epsilon, q + b\epsilon)}{L}\pi\right).$$

We then compute

$$\frac{d}{d\epsilon}|_{\epsilon=0} d_{a,b,\epsilon} = \langle \omega, aT(p) - bT(q)\rangle,$$

$$\frac{d^2}{d\epsilon^2}|_{\epsilon=0} d_{a,b,\epsilon} = \langle \omega, a^2\kappa N(p) - b^2\kappa N(q)\rangle + \frac{1}{d}\|aT(p) - bT(q)\|^2 - \frac{1}{d}\langle \omega, aT(p) - bT(q)\rangle,$$

$$\frac{d}{d\epsilon}|_{\epsilon=0}\psi_{a,b,\epsilon} = (a-b)\cos\left(\frac{l}{L}\pi\right),$$

and

$$\frac{d^2}{d\epsilon^2}|_{\epsilon=0}\psi_{a,b,\epsilon} = -\frac{\pi}{L}(a-b)^2\sin\left(\frac{l}{L}\pi\right) = -\frac{(a-b)^2\pi^2}{L^2}\psi.$$

For the first variation, take $a = 1$ and $b = 0$ to get

$$\begin{aligned} 0 &= \frac{d}{d\epsilon}|_{\epsilon=0}\frac{d}{\psi} \\ &= \langle \omega, T(p)\rangle - \frac{d}{\psi}\cos\left(\frac{l}{L}\pi\right). \end{aligned}$$

By taking $a = 0$ and $b = 1$, one obtains the same formula replacing $T(p)$ by $T(q)$, so we have the useful relation

$$\langle \omega, T(p) \rangle = \langle \omega, T(q) \rangle = \frac{d}{\psi} \cos\left(\frac{l}{L}\pi\right). \tag{14}$$

The tangent vectors have unit length, so $T(p) + T(q)$ is perpendicular to $T(p) - T(q)$ which is in turn perpendicular to ω, hence $T(p) + T(q)$ is parallel to ω. We exploit this fact in the second variation by taking $a = 1$ and $b = -1$. In this case, the last two terms in the second variation formula for d cancel because $\langle \omega, T(p) + T(q) \rangle = \|T(p) + T(q)\|$, so

$$\frac{d^2}{d\epsilon^2}|_{\epsilon=0} d = \langle \omega, \kappa N(p) - \kappa N(q) \rangle.$$

By equation 14, we have

$$\frac{d}{d\epsilon}|_{\epsilon=0} d = \frac{2d}{\psi} \cos\left(\frac{l}{L}\pi\right)$$

and using the variation formulas above we have

$$\frac{d}{d\epsilon}|_{\epsilon=0} \psi = 2\cos\left(\frac{l}{L}\pi\right)$$

and

$$\frac{d^2}{d\epsilon^2}|_{\epsilon=0} \psi = -\frac{4\pi^2}{L^2}\psi.$$

Putting it all together and applying the nonnegativity of the second variation, we obtain

$$\begin{aligned}
0 &\leq \frac{d^2}{d\epsilon^2}|_{\epsilon=0} \frac{d}{\psi} \\
&= \frac{1}{\psi}\frac{d^2}{d\epsilon^2} d - \frac{2}{\psi^2}\frac{d}{d\epsilon} d \frac{d}{d\epsilon}\psi + \frac{2d}{\psi^3}\left(\frac{d}{d\epsilon}\psi\right)^2 - \frac{d}{\psi^2}\frac{d^2}{d\epsilon^2}\psi \\
&= \frac{1}{\psi}(\langle \omega, \kappa N(p) - \kappa N(q) \rangle) - \frac{2}{\psi^2}\left(\frac{4d}{\psi}\cos^2\left(\frac{l}{L}\pi\right)\right) \\
&+ \frac{2d}{\psi^3}\left(4\cos^2\left(\frac{l}{L}\pi\right)\right) - \frac{d}{\psi^2}\left(-\frac{4\pi^2}{L^2}\psi\right) \\
&= \frac{1}{\psi}\langle \omega, \kappa N(p) - \kappa N(q) \rangle + \frac{4\pi^2}{L^2}\frac{d}{\psi}.
\end{aligned}$$

We will use these estimates to control the time derivative of $\frac{d}{\psi}$ at the local spatial minimum and prove Theorem 12.

Proof. Using the evolution equations for our curve, l, and L, we obtain

$$\frac{d}{dt} d = \langle \omega, \kappa N(p) - \kappa N(q) \rangle$$

and

$$\frac{d}{dt}\psi = -\frac{\psi}{L}\int_{S^1} \kappa^2 ds - \cos\left(\frac{l}{L}\pi\right)\int_p^q \kappa^2 ds + \frac{l}{L}\cos\left(\frac{l}{L}\pi\right)\int_{S^1} \kappa^2 ds.$$

Putting these together, we get the evolution equation

$$\frac{d}{dt}\left(\frac{d}{\psi}\right) = \frac{1}{\psi}\langle \omega, \kappa N(p) - \kappa N(q)\rangle$$
$$+ \frac{d}{\psi^2}\left(\frac{\psi}{L}\int_{S^1}\kappa^2 ds + \cos\left(\frac{l}{L}\pi\right)\int_p^q \kappa^2 ds - \frac{l}{L}\cos\left(\frac{l}{L}\pi\right)\int_{S^1}\kappa^2 ds\right).$$

For the first term, we apply our second variation inequality and rearrange terms to get

$$\frac{d}{dt}\left(\frac{d}{\psi}\right) \geq -\frac{4\pi^2 d}{L^2\psi} + \frac{d}{L\psi}\int_{S^1}\kappa^2 ds\left(1 - \frac{l}{\psi}\cos\left(\frac{l}{L}\pi\right)\right)$$
$$+ \frac{d}{\psi^2}\cos\left(\frac{l}{L}\pi\right)\int_p^q \kappa^2 ds.$$

Now, without loss of generality, $\frac{l}{L}\pi < \frac{\pi}{2}$, choosing the correct direction from p to q. It follows that $\frac{l}{\psi}\cos\left(\frac{l}{L}\pi\right) = \frac{l\pi/L}{\tan(l\pi/L)} < 1$, so the second term is positive. Applying Cauchy-Schwartz and the fact that $\int_{S^1}\kappa ds = 2\pi$, we have $\int_{S^1}\kappa^2 ds \geq \frac{4\pi^2}{L}$. When we substitute this inequality into the above differential inequality, the first term cancels with part of the second, yielding

$$\frac{d}{dt}\left(\frac{d}{\psi}\right) \geq \frac{d}{\psi^2}\cos\left(\frac{l}{L}\pi\right)\left(\int_p^q \kappa^2 ds - \frac{4\pi^2 l}{L^2}\right).$$

We can apply Cauchy-Schwartz again to get $\int_p^q \kappa^2 ds \geq \frac{\left(\int_p^q \kappa ds\right)^2}{l}$. The integral $\int_p^q \kappa ds$ is the angle β between $T(p)$ and $T(q)$. This yields

$$\frac{d}{dt}\left(\frac{d}{\psi}\right) \geq \frac{4d}{l\psi^2}\cos\left(\frac{l}{L}\pi\right)\left(\left(\frac{\beta}{2}\right)^2 - \frac{\pi^2 l^2}{L^2}\right).$$

Finally, by our first variation formula, $\cos\left(\frac{\beta}{2}\right) = \frac{d}{\psi}\cos\left(\frac{l}{L}\pi\right)$. At a spatial minimum, we must have $\frac{d}{\psi} \leq 1$, where equality holds if and only if our curve is a circle. Note that $\beta \leq \pi$, because otherwise, we could perturb p and q to reduce d and increase l, contradicting the assumption that we are at a spatial minimum of $\frac{d}{\psi}$; the previous two inequalities then imply that

$$\frac{\beta}{2} \geq \frac{l}{L}\pi.$$

Hence, our differential inequality becomes

$$\frac{d}{dt}\left(\frac{d}{\psi}\right) \geq 0$$

with equality only if our curve is a circle, as desired. □

Initial embeddedness gives us a positive lower bound on $\frac{d}{\psi}$, which rules out all of the ways singularities can form. Let T be the maximal time interval on which a solution exists. In the case that the curvature times the time to blowup remains finite, Huisken [12] proved the following:

Theorem 13. *If* $\limsup_{t\to T} \|\kappa(\cdot,t)\|_\infty (T-t) < \infty$ *then our curve is asymptotic to a homothetically shrinking solution.*

The following theorem of Abresch and Langer [1] then proves that the curve is asymptotic to a circle:

Theorem 14. *The only embedded homothetically shrinking solution to the curve-shortening flow is the circle.*

In the other case, Altschuler [2] obtained the following characterization:

Theorem 15. *If* $\limsup_{t\to T} \|\kappa(\cdot,t)\|_\infty (T-t) = \infty$ *then there exists a sequence of points and times* (p_n, t_n)*, with* $t_n \to T$*, such that a limit of rescalings along* (p_n, t_n) *converges to the grim reaper.*

Note that the quantity $\frac{d}{\psi}$ is invariant under rescaling of the plane. The Grim Reaper can be written $y(x,t) = t - \log\cos(x)$; by taking horizontal secant lines progressively higher, we can make $\frac{d}{\psi}$ arbitrarily small. The above result rules out this type of behavior, proving theorem 1.

5 Riemannian Geometry Basics

In this section we introduce the basic concepts from Riemannian geometry necessary for our treatment of the Ricci flow. We emphasize results useful for the intense computations involved in Ricci flow theory. Most of this material was gleaned from Brendle [3], Milnor [14], and O'Neill [15].

Let M be a smooth manifold of dimension n. Denote the space of smooth vector fields on M by $\mathcal{X}(M)$, its dual space of one-forms by $\mathcal{X}^*(M)$, and the space of smooth functions on M by $\mathcal{F}(M)$.

5.1 Tensor Calculus

In this section we develop tensor calculus on Riemannian manifolds. We begin by defining tensors, and discuss a natural way to differentiate these objects when a metric structure is introduced.

Definition 1. *An (r,s) tensor is a map $T : \mathcal{X}^*(M)^r \times \mathcal{X}(M)^s \to \mathcal{F}(M)$ which is multilinear over $\mathcal{F}(M)$, i.e. for any one-forms $\alpha_1, ..., \alpha_r$ and vector fields $X_1, ..., X_s$,*

$$\begin{aligned} fT(\alpha_1, ..., \alpha_r, X_1, ..., X_s) &= T(\alpha_1, ..., f\alpha_i, ..., \alpha_r, X_1, ..., X_s) \\ &= T(\alpha_1, ..., \alpha_r, X_1, ..., fX_j, ..., X_s). \end{aligned}$$

Linearity over functions makes an (r,s) tensor T a pointwise object in the sense that at a point p, $T(\alpha_1, ..., \alpha_r, X_1, ..., X_s)(p)$ depends only on the values $\alpha_i(p)$ and $X_j(p)$. We refer to this property as tensoriality. In local coordinates, we denote by $T^{i_1 \dots i_r}_{j_1 \dots j_s}$ the tensor components $T(dx^{i_1}, ..., dx^{i_r}, \partial_{j_1}, ..., \partial_{j_n})$.

A Riemannian metric tensor g on a smooth manifold M is a symmetric, positive-definite tensor of type $(0,2)$. From now on, we equip M with a Riemannian metric and denote it by (M, g). Note that one-forms are $(0,1)$ tensors and vector fields are $(1,0)$ tensors. We will now discuss how to differentiate vector fields and one forms in a way that interacts well with the metric structure on M.

Definition 2. *Let X and Y be smooth vector fields on M. A connection for M is a mapping D from $\mathcal{X}(M) \times \mathcal{X}(M)$ to $\mathcal{X}(M)$, sending (X,Y) to a vector field we denote by $D_X Y$, with the following properties:*

1. *$\mathbb{R}$-linearity in Y: $D_X(aY + bZ) = aD_X Y + bD_X Z$*
2. *$\mathcal{F}(M)$-linearity in X: $D_{fX+gZ} Y = fD_X Y + gD_Z Y$*
3. *Leibniz rule: $D_X(fY) = fD_X Y + (Xf)Y$*

A priori, there are many possibilities for a connection on M. The Riemannian metric g allows us to select a unique preferred connection.

Theorem 16. *Let (M, g) be a Riemannian manifold. Then there exists a unique connection D on M, called the Levi-Civita connection, satisfying*

1. *Metric Compatibility:* $V(g(X,Y)) = g(D_V X, Y) + g(X, D_V Y)$ *for all* $V, X, Y \in \mathcal{X}(M)$.

2. *Zero Torsion:* $D_X Y - D_Y X = [X, Y]$.

Proof. We begin with uniqueness. Assume D is a metric-compatible and torsion-free connection. Then for vector fields V, W and X, one obtains the formula

$$\begin{aligned} 2g(D_V W, X) = &V(g(W,X)) + W(g(X,V)) - X(g(V,W)) \\ &- g(V,[W,X]) + g(W,[X,V]) + g(X,[V,W]) \end{aligned}$$

by applying metric compatibility to the first three terms and symmetry to the last three. This is called the Koszul formula. It is clear that if D and $\widetilde{D}$ are any two such connections, then $g(D_V W, X)$ and $g(\widetilde{D}_V W, X)$ agree for arbitrary X, proving uniqueness.

To prove existence, define the map Γ from $\mathcal{X}(M)^3$ to $\mathcal{F}(M)$ by the Koszul formula:

$$\begin{aligned} \Gamma(V,W,X) = \frac{1}{2}\{&V(g(W,X)) + W(g(X,V)) - X(g(V,W)) \\ &- g(V,[W,X]) + g(W,[X,V]) + g(X,[V,W])\} \end{aligned}$$

It is straightforward to show that Γ is tensorial in the third slot, so by fixing V and W, Γ becomes a covector field which is dual to some vector field Z by the finite dimensional representation theorem, i.e. $\Gamma(V,W,X) = g(Z,X)$. We define $D_V W$ to be Z. Again, by a simple computation, Γ is tensorial in the first slot and satisfies the Leibniz rule $\Gamma(V, fW, X) = V(f)g(W,X) + f\Gamma(V,W,X)$ in the second slot, which proves that D is indeed a connection.

Finally, note that the first three terms in the Koszul formula are symmetric in V and W, the fourth and fifth term are symmetric in V and W, and the last term is antisymmetric in V and W, so

$$g(D_V W, X) - g(D_W V, X) = 2(\frac{1}{2}g(X,[V,W]) = g(X,[V,W])$$

for arbitrary X, proving symmetry. Metric compatibility is proved by a simple computation as well. □

From now on, D refers to the Levi-Civita connection. It is locally defined by n^3 functions Γ_{ij}^k known as the Christoffel symbols. Letting ∂_i and ∂_j denote coordinate vector fields, we define $D_{\partial i}\partial_j = \sum_{i=1}^{n} \Gamma_{ij}^k \partial_k$. Henceforth, we will use the Einstein summation notation. Writing X as $x^i\partial_i$ and Y as $y^i\partial_i$ with respect to the coodinate basis, we compute

$$D_X Y = x^i D_{\partial_i}(y^j \partial_j) = (x^i \partial_i y^k + x^i y^j \Gamma_{ij}^k)\partial_k.$$

It is also convenient to compute the Christoffel symbols. Using the Koszul formula, we have

$$g(D_{\partial_i}\partial_j, \partial_l) = \Gamma_{ij}^k g_{kl} = \frac{1}{2}\{\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\}.$$

Multiplying by the inverse of the metric tensor, we get

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl}\{\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\}.$$

We now define how to differentiate tensors in the direction of a vector field V.

Definition 3. *Let α be a one-form and let V be a vector field. We define the one-form $D_V\alpha$ by the formula*

$$(D_V\alpha)(W) = V(\alpha(W)) - \alpha(D_V W).$$

It is easy to check that $D_V\alpha$ is well-defined, i.e. is tensorial in W. These definitions induce a covariant derivative on tensors as follows. Let T be an (r, s) tensor. Given a vector field V, the (r, s) tensor $D_V T$ is given by

$$\begin{aligned}(D_V T)(\alpha_1, ..., \alpha_r, X_1, ..., X_s) &= V(T(\alpha_1, ..., \alpha_r, X_1, ..., X_s)) \\ &\quad - \sum_{k=1}^{r} T(\alpha_1, ..., D_V\alpha_k, ..., \alpha_r, X_1, ..., X_s) \\ &\quad - \sum_{j=1}^{s} T(\alpha_1, ..., \alpha_r, X_1, ..., D_V X_j, ...X_s).\end{aligned}$$

Again, it is easy to verify that $D_V(T)$ is tensorial in all slots. Note also that $D_{fV}T = fD_V T$, so we can define an $(r, s + 1)$ tensor DT by

$$DT(\alpha_1, ..., \alpha_r, X_1, ..., X_s, V) = (D_V T)(\alpha_1, ..., \alpha_r, X_1, ..., X_s),$$

interpreted as the total derivative of T. Observe that metric compatibility is equivalent to the condition $Dg = 0$.

We now discuss how to take second order covariant derivatives of an arbitrary tensor. Let T be a type (r, s) tensor. We define the (r, s) tensor $D^2_{V,W}T$ by

$$D^2_{V,W}T(\alpha_1, ..., \alpha_r, X_1, ..., X_s) = DDT(\alpha_1, ..., \alpha_r, X_1, ..., X_s, W, V).$$

This definition is manifestly tensorial in V and W. Using the rules above, one easily computes

$$D^2_{V,W}T = D_V D_W T - D_{D_V W}T.$$

As a special case, given a function f we obtain its Hessian as the $(0, 2)$ tensor defined by

$$\begin{aligned}\mathrm{Hess} f(X, Y) &= D^2_{X,Y} f \\ &= D_X D_Y f - D_{D_X Y} f \\ &= X(Y(f)) - D_X Y(f).\end{aligned}$$

Finally, we discuss contraction and define the Laplacian of a tensor. Let T be a $(0, 2)$ tensor, and define the function $f = \sum_{i=1}^{n} T(E_i, E_i)$, where E_i form an orthonormal basis for T_pM. To show that this is independent of the basis chosen, take another orthonormal basis $E'_i = a_{ji}E_j$, where (a_{ij}) is an orthogonal transformation. We then have

$$\begin{aligned}\sum_{i=1}^{n} T(E'_i, E'_i) &= \sum_{i,j,k=1}^{n} a_{ji}a_{ki}T(E_j, E_k) \\ &= \sum_{j,k=1}^{n} \delta_{jk}T(E_j, E_k) \\ &= \sum_{j=1}^{n} T(E_j, E_j).\end{aligned}$$

We call this operation on T contraction. It is convenient to consider how to contract in local coordinates. To do so, we employ a nice trick that allows us to select a natural orthonormal basis given a coordinate basis $\{\partial_1, ..., \partial_n\}$. Since the metric is symmetric and positive-definite, there exists an orthogonal transformation O on T_pM such that

$$g^{-1} = O\mathrm{diag}(\lambda_1, ..., \lambda_n)O^T.$$

Define $g^{-\frac{1}{2}} = O\mathrm{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n})O^T$. Note that this is symmetric. Let

$$E_i = \left(g^{-\frac{1}{2}}\right)_{ki} \partial_k = \left(g^{-\frac{1}{2}}\right)_{ik} \partial_k.$$

It is easy to verify directly that these are orthonormal. Thus, the contraction of T is given by

$$\begin{aligned}\sum_{i=1}^{n} T(E_i, E_i) &= \sum_{i,j,k=1}^{n} \left(g^{-\frac{1}{2}}\right)_{ji} \left(g^{-\frac{1}{2}}\right)_{ik} T(\partial_j, \partial_k) \\ &= \sum_{j,k=1}^{n} \left(O\mathrm{diag}(\lambda_1, ..., \lambda_n)O^T\right)_{jk} T_{jk} \\ &= g^{jk} T_{jk}.\end{aligned}$$

The Laplacian of a tensor is defined as the contraction of its second covariant derivative, i.e.

$$\triangle T = \sum_i D^2_{e_i,e_i} T = g^{jk} D^2_{j,k} T.$$

5.2 Geodesics

In this section we discuss how to differentiate vector fields along curves. This allows us to define geodesics, which are very natural objects in Riemannian geometry because they locally minimize arclength. Finally, we introduce geodesic normal coordinates, which will be very useful for computations in the upcoming sections.

Recall that the vector field $D_V W$ is tensorial in V, and thus only depends on the values of V at a point. This allows us to make sense of item 3 in the following proposition.

Proposition 1. *Let $c(t)$ be a smooth parametrized curve in M, and let $V(t)$ be a vector field on c. The covariant derivative of V along c, denoted by $\frac{DV}{dt}$, is uniquely characterized by the following properties:*

1. $\frac{D(V+W)}{dt} = \frac{DV}{dt} + \frac{DW}{dt}$
2. $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$
3. *If V is induced by a vector field Y, then* $\frac{DV}{dt} = D_{c'(t)}Y$

Proof. Suppose first that such an operation exists. Choose local coordinates u^i and let $u^i(t)$ denote the coordinates of $c(t)$. Write $V = v^i(t)\partial_i$. Applying the above rules, we get

$$\frac{DV}{dt} = (\frac{dv^k}{dt} + \frac{du^i}{dt}v^j\Gamma_{ij}^k)\partial_k$$

proving uniqueness. Conversely, defining $\frac{DV}{dt}$ using this formula gives us all of the desired properties, proving existence. □

Equipped with the notion of differentiation along a curve, we can now understand parallel transport.

Definition 4. *A vector field V along a curve c is called parallel if $\frac{DV}{dt}$ is identically* 0.

Setting the above formula for the covariant derivative to 0, we get a system of first-order linear ODE, which is guaranteed to have a unique solution given $V(0)$. The following proposition tells us that parallel transport preserves inner product:

Proposition 2. *Let $V(t)$ and $W(t)$ be vector fields along a curve $c(t)$. Then*

$$\frac{d}{dt}g(V(t), W(t)) = g(\frac{DV}{dt}, W) + g(V, \frac{DW}{dt}).$$

Proof. This is an immediate consequence of metric compatibility. □

Definition 5. *Let $\gamma : I \to M$ be a smooth parametrized curve. We say γ is geodesic if $\frac{D}{dt}\gamma'(t) = 0$.*

Note first that if γ is geodesic, then $\frac{d}{dt}g(\gamma'(t), \gamma'(t)) = 2g(\frac{D}{dt}\gamma'(t), \gamma'(t)) = 0$. Thus, the parameter of a geodesic is always linearly related to arclength. In a coordinate system $(u^1, ..., u^n)$, finding a geodesic $(u^1(t), ..., u^n(t))$ is equivalent to solving following system of ODE, obtained using the parallel translation formula derived earlier:

$$\frac{d^2u^k}{dt^2} + \frac{du^i}{dt}\frac{du^j}{dt}\Gamma_{ij}^k(t) = 0.$$

For a fixed point $p \in M$ and $v \in T_pM$, ODE theory guarantees the existence of a unique geodesic $\gamma_v(t)$ with $\gamma(0) = p$ and $\gamma'(0) = v$, defined for t in some interval $(-\epsilon, \epsilon)$. We can now define the exponential map:

Definition 6. *The exponential map $\exp_p : T_p(M) \to M$ is defined by $\exp_p(v) = \gamma_v(1)$, where $\gamma_v(t)$ is the unique geodesic with $\gamma_v(0) = p$ and $\gamma_v'(0) = v$.*

It is clear that $\exp_p$ is not necessarily defined on all of $T_p(M)$. However, it is certainly defined for sufficiently small intial velocities, i.e. in a neighborhood of 0, which follows from the simple observation that $\gamma_v(at) = \gamma_{av}(t)$ where both exist.

Let $e_1, ..., e_n$ be an orthonormal basis of T_pM, and consider the map $\Phi : \mathbb{R}^n \to M$ sending $(x^1, ..., x^n)$ to $\exp_p(\sum_{k=1}^{n} x^k e_k)$. It is easy to verify that the differential of the exponential map at the origin is the identity, so by the inverse function theorem, Φ is is a local coordinate system near 0. These coordinates are referred to as geodesic normal coordinates. In these coordinates, geodesics through p correspond to straight lines through the origin, and the Christoffel symbols vanish at p:

Proposition 3. *Let* $(x^1, ..., x^n)$ *be geodesic normal coordinates coordinates around* p. *Then the Christoffel symbols vanish at* p.

Proof. Fix $v \in \mathbb{R}^n$ and let $\alpha(t) = \exp_p(tv)$. Then $\alpha(t)$ is a geodesic for small t, with coordinate expression tv. Thus, in local coordinates $\alpha'(t) = v$ and $\alpha''(t) = 0$. Applying the geodesic equation and setting $t = 0$, we get

$$\Gamma^k_{ij}(0)v_i v_j = 0$$

for all $k \in \{1, ..., n\}$. Since v is arbitary, this implies the antisymmetry $\Gamma^k_{ij}(0) + \Gamma^k_{ji}(0) = 0$. Symmetry of the connection then gives $\Gamma^k_{ij}(0) = 0$. □

This property of geodesic normal coordinates makes them very useful for pointwise computations. In particular, in normal coordinates the covariant derivative at p reduces to the usual directional derivative, i.e.

$$D_i T^{i_1 \dots i_r}_{j_1 \dots j_s} = \partial_i T^{i_1 \dots i_r}_{j_1 \dots j_s}$$

and the Laplacian reduces to

$$\triangle T = g^{ij} D_i D_j T.$$

Finally, we touch on the path length minimizing properties of geodesics. One may equip a Riemannian manifold with a distance function by letting $d(p, q)$ be the infimum of the lengths of all piecewise smooth paths joining p and q. It is a standard result that geodesics locally minimize arclength, which proves that d is in fact a distance function. Furthermore, by another standard result known as the Hopf-Rinow theorem, any two points on a compact manifold can be joined by a minimizing geodesic. For a proof of these theorems, see for example Milnor [14].

5.3 Curvature

In this section, we introduce the Riemann curvature tensor, the Ricci curvature tensor and the scalar curvature.

Definition 7. *The Riemann curvature tensor of* (M, g) *is the* $(0, 4)$ *tensor given by*

$$\begin{aligned} R(X, Y, Z, W) &= -g(D^2_{X,Y} Z - D^2_{Y,X} Z, W) \\ &= -g(D_X D_Y Z - D_Y D_X Z - D_{[X,Y]} Z, W). \end{aligned}$$

R is manifestly tensorial in X, Y and W, and one easily checks that $R(X, Y, fZ, W) = fR(X, Y, Z, W)$ from the definition. To compute in local coordinates, first note that

$$\begin{aligned} D_i D_j \partial_k &= D_i \left(\Gamma^l_{jk} \partial_l \right) \\ &= \{ \partial_i \Gamma^l_{jk} + \Gamma^m_{jk} \Gamma^l_{im} \} \partial_l. \end{aligned}$$

Since the Lie bracket of coordinate vector fields vanishes, using the above computation we obtain

$$\begin{aligned} R_{ijkp} &= -g(D_i D_j \partial_k - D_j D_i \partial_k - D_{[\partial_i, \partial_j]} \partial_k, \partial_p) \\ &= -g_{lp} \{ \partial_i \Gamma^l_{jk} - \partial_j \Gamma^l_{ik} + \Gamma^m_{jk} \Gamma^l_{im} - \Gamma^m_{ik} \Gamma^l_{jm} \} \end{aligned}$$

The Riemann curvature possesses algebraic symmetries that will be extremely important for computations. We discuss these next.

Proposition 4. *The Riemann curvature tensor satisfies the following algebraic identities:*

1. $R(X,Y,Z,W) + R(Y,X,Z,W) = 0$
2. $R(X,Y,Z,W) + R(Y,Z,X,W) + R(Z,X,Y,W) = 0$ *(the first Bianchi identity)*
3. $R(X,Y,Z,W) + R(X,Y,W,Z) = 0$
4. $R(X,Y,Z,W) = R(Z,W,X,Y)$

Proof. The first symmetry follows directly from the definition of R. By tensoriality, when computing R at a point we may without loss of generality extend to vector fields with constant components so that Lie brackets vanish. It follows from symmetry that $D_X Y - D_Y X = [X,Y] = 0$. To prove the first Bianchi identity, we compute

$$\begin{aligned}
&R(X,Y,Z,W) + R(Y,Z,X,W) + R(Z,X,Y,W) \\
&= -g(D_X D_Y Z - D_Y D_X Z + D_Y D_Z X - D_Z D_Y X + D_Z D_X Y - D_X D_Z Y, W) \\
&= -g(D_X[Y,Z] + D_Y[Z,X] + D_Z[X,Y], W) \\
&= 0
\end{aligned}$$

as desired. The third identity is equivalent to the condition $R(X,Y,Z,Z) = 0$, so it suffices to prove that $g(D_X D_Y Z, Z)$ is symmetric in X and Y. By metric compatibility, $XY g(Z,Z) - 2g(D_X Z, D_Y Z) = 2g(D_X D_Y Z, Z)$. Using that the Lie brackets vanish, the left hand side is symmetric in X and Y, proving the third identity. Finally, the last identity is obtained by applying the others. Using antisymmetry in the first and last pairs of slots and the first Bianchi identity,

$$\begin{aligned}
R(X,Y,Z,W) &= -R(Y,X,Z,W) \\
&= R(X,Z,Y,W) + R(Z,Y,X,W) \\
R(X,Y,Z,W) &= -R(X,Y,W,Z) \\
&= R(Y,W,X,Z) + R(W,X,Y,Z).
\end{aligned}$$

Adding these two equations together, we get

$$\begin{aligned}
2R(X,Y,Z,W) &= R(X,Z,Y,W) + R(Z,Y,X,W) \\
&\quad + R(Y,W,X,Z) + R(W,X,Y,Z).
\end{aligned}$$

By switching the roles of X and Z, likewise Y and W, we obtain

$$\begin{aligned}
2R(Z,W,X,Y) &= R(Z,X,W,Y) + R(X,W,Y,Z) \\
&\quad + R(W,Y,Z,X) + R(Y,Z,X,W).
\end{aligned}$$

It is easy to verify using antisymmetry in the first two and last two slots that the right hand sides of these equations agree, finishing the proof. $\square$

Note that with respect to an orthonormal basis $\{e_k\}$, we have the formula

$$D^2_{X,Y}Z - D^2_{Y,X}Z = -\sum_{k=1}^{n} R(X,Y,Z,e_k)e_k.$$

We would like to obtain similar commutator formulas for arbitrary tensors of type (r,s). For motivation, we begin with one-forms. Observe that

$$\begin{aligned} X(Y(\alpha(Z))) &= X\{(D_Y\alpha)(Z) + \alpha(D_Y Z)\} \\ &= (D_X D_Y \alpha)(Z) + (D_Y\alpha)(D_X Z) + (D_X\alpha)(D_Y Z) + \alpha(D_X D_Y Z). \end{aligned}$$

Note that the second two terms are symmetric in X and Y; Antisymmetrizing in X and Y, we get

$$[X,Y](\alpha(Z)) = (D_X D_Y\alpha - D_Y D_X\alpha)(Z) + \alpha(D_X D_Y Z - D_Y D_X Z).$$

By definition, the left hand side is $(D_{[X,Y]}\alpha)(Z) + \alpha(D_{[X,Y]}Z)$. Rearranging, we obtain

$$\begin{aligned} (D^2_{X,Y}\alpha - D^2_{Y,X}\alpha)(Z) &= -\alpha(D^2_{X,Y}Z - D^2_{Y,X}Z) \\ &= \sum_{k=1}^{n} R(X,Y,Z,e_k)\alpha(e_k). \end{aligned}$$

It is straightforward to extend this to arbitrary tensors T of type (r,s). For example, for a $(0,2)$ tensor, we have

$$\begin{aligned} (D^2_{X,Y}T - D^2_{Y,X}T)(V,W) &= -T(D^2_{X,Y}V - D^2_{Y,X}V, W) - T(V, D^2_{X,Y}W - D^2_{Y,X}W) \\ &= \sum_{k=1}^{n} R(X,Y,V,e_k)T(e_k,W) + \sum_{k=1}^{n} R(X,Y,W,e_k)T(V,e_k) \end{aligned}$$

From this, we get the classical coordinate formula for the commutator of covariant derivatives:

$$[D_{\partial_i}, D_{\partial_j}]T_{kl} = g^{pq}R_{ijkp}T_{ql} + g^{pq}R_{ijlp}T_{kp}.$$

The first covariant derivative of the Riemann curvature tensor also satisfies a very interesting and important identity known as the second Bianchi identity:

Proposition 5.

$$(D_X R)(Y,Z,V,W) + (D_Y R)(Z,X,V,W) + (D_Z R)(X,Y,V,W) = 0.$$

Proof. It suffices to prove this identity at an arbitrary point p. Choose geodesic normal coordinates around p; by tensoriality, it is enough to check this identity for the coordinate vector fields at p. Note that the lie brackets of the coordinate vector fields vanish in a neighborhood of p, and that the Christoffel symbols vanish at p. Since the covariant derivative at p reduces to the usual derivative, we have

$$\begin{aligned} D_{\partial_i}R(\partial_j,\partial_k,\partial_l,\partial_m) &= \partial_i R(\partial_j,\partial_k,\partial_l,\partial_m) \\ &= -\partial_i g(D_{\partial_j}D_{\partial_k}\partial_l - D_{\partial_k}D_{\partial_j}\partial_l, \partial_m) \\ &= -g(D_{\partial_i}D_{\partial_j}D_{\partial_k}\partial_l - D_{\partial_i}D_{\partial_k}D_{\partial_j}\partial_l, \partial_m) \end{aligned}$$

where in the second line we used the vanishing of lie brackets in a neighborhood of p and in the third we applied metric compatibility and used that the Christoffel symbols vanish at p. Permuting the i,j and k indices cyclically, we get

$$\begin{aligned}
&(D_{\partial_i}R)(\partial_j, \partial_k, \partial_l, \partial_m) + (D_{\partial_j}R)(\partial_k, \partial_i, \partial_l, \partial_m) + (D_{\partial_k}R)(\partial_i, \partial_j, \partial_l, \partial_m) \\
&= -g(\partial_m, D_{\partial_i}D_{\partial_j}D_{\partial_k}\partial_l - D_{\partial_i}D_{\partial_k}D_{\partial_j}\partial_l \\
&+ D_{\partial_j}D_{\partial_k}D_{\partial_i}\partial_l - D_{\partial_j}D_{\partial_i}D_{\partial_k}\partial_l \\
&+ D_{\partial_k}D_{\partial_i}D_{\partial_j}\partial_l - D_{\partial_k}D_{\partial_j}D_{\partial_i}\partial_l).
\end{aligned}$$

Finally, notice that

$$-g(\partial_m, D_{\partial_i}D_{\partial_j}D_{\partial_k}\partial_l - D_{\partial_j}D_{\partial_i}D_{\partial_k}\partial_l) = R(\partial_i, \partial_j, D_{\partial_k}\partial_l, \partial_m),$$

and this vanishes at p because the christoffel symbols vanish at p. The other terms cancel similarly, completing the proof. □

We can now discuss the other notions of curvature derived from the Riemann curvature tensor.

Definition 8. *Let $p \in M$ and let π be a 2-dimensional subspace of T_pM spanned by v and w. We define the sectional curvature $K(\pi)$ of this plane as the ratio*

$$\frac{R(v, w, v, w)}{g(v, v)g(w, w) - g(v, w)^2}.$$

It is easy to check, using antisymmetry of R in the first pair and last pair of slots, that this definition is independent of the choice of basis. Using the symmetries of the Riemann curvature tensor, one can also show that the sectional curvatures determine R at a point p; in particular, if we have any $(0, 4)$ tensor $\widetilde{R}$ with the same symmetries as the Riemann curvature tensor such that $\widetilde{R}(X, Y, X, Y) = R(X, Y, X, Y)$ for any $X, Y \in T_pM$, we can conclude that $R(X, Y, Z, W) = \widetilde{R}(X, Y, Z, W)$ for any $X, Y, Z, W \in T_pM$. The other way we can obtain curvature quantities are by taking the nonzero traces of the curvature tensor:

Definition 9. *The Ricci curvature tensor is the symmetric $(0, 2)$-tensor defined by*

$$Ric(X, Y) = \sum_{k=1}^{n} R(X, e_k, Y, e_k)$$

where e_k form an orthonormal basis for T_pM.

As shown in the tensor calculus section, this definition is independent of the orthonormal basis chosen, and in local coordinates

$$R_{ij} = g^{kl}R_{ikjl}.$$

Taking one more trace, we obtain the scalar curvature:

Definition 10. *The scalar curvature is the smooth function R given by*

$$\sum_{k=1}^{n} Ric(e_k, e_k).$$

In local coordinates, $R = g^{ij} R_{ij}$. Finally, the second Bianchi identity can be contracted to give us a relation between the covariant derivative of the Ricci tensor and the directional derivative of the scalar curvature:

Proposition 6. $\displaystyle\sum_{k=1}^{n} (D_{e_k} Ric)(e_k, X) = \frac{1}{2} X(R).$

Proof. We start from the second Bianchi identity

$$0 = D_X R(Y, Z, V, W) + D_Y R(Z, X, V, W) + D_Z R(X, Y, V, W).$$

Taking the trace over Z and W, and then over Y and V, we obtain

$$0 = \sum_{k,l=1}^{n} (D_X R)(e_l, e_k, e_l, e_k) + \sum_{k,l=1}^{n} (D_{e_l} R)(e_k, X, e_l, e_k) + \sum_{k,l=1}^{n} (D_{e_k} R)(X, e_l, e_l, e_k).$$

Noting that $\sum_k R(e_k, X, e_l, e_k) = -\mathrm{Ric}(X, e_l)$, the second term becomes $-\displaystyle\sum_{l=1}^{n} D_{e_l} \mathrm{Ric}(X, e_l)$. Similarly, the last term becomes $-\displaystyle\sum_{k=1}^{n} D_{e_k} \mathrm{Ric}(X, e_k)$. Finally, the sum $\displaystyle\sum_{k,l=1}^{n} R(e_l, e_k, e_l, e_k) = R$, so the first term is $X(R)$, completing the proof. □

In local coordinates, we may write this identity as

$$g^{ij} D_i R_{jk} = \frac{1}{2} \partial_k R.$$

6 Ricci Flow Basics

In this section we introduce the Ricci flow and derive basic facts. We begin in dimension n, and later specialize these results to dimensions 2 and 3.

6.1 Normalized and Unnormalized Ricci Flow

Let $g(t)$ be a one-parameter family of Riemannian metrics on M. We say $g(t)$ is a solution of the (unnormalized) Ricci flow if

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

The short-time existence and uniqueness of solutions to this equation is proven in, for example, Brendle [3]. To study convergence theory for the Ricci flow, it is convenient to adjust the equation so that volume is preserved. Denote by $r(t)$ the mean scalar curvature of M at time t. We say that $g(t)$ is a solution of the normalized Ricci flow if

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

The volume of M remains constant under the normalized flow. To see this, observe that

$$\begin{aligned}
\frac{\partial}{\partial t}(\det(g)) &= \sum_{i,j} \frac{\partial(\det(g))}{\partial g_{ij}} \frac{\partial}{\partial t} g_{ij} \\
&= \sum_{i,j} (\mathrm{Adj}(g))_{ij} \frac{\partial}{\partial t} g_{ij} \\
&= \det(g) g^{ij} \frac{\partial}{\partial t} g_{ij} \\
&= 2(r - R)\det(g).
\end{aligned}$$

It follows that the volume element $d\mu = \sqrt{\det(g)}dx$ satisfies $\frac{\partial}{\partial t} d\mu = (r - R)d\mu$. Integrating gives 0 by the definition of r, as desired.

The following theorem relates the two equations, making it straightforward to back and forth between the two.

Theorem 17. *The normalized and unnormalized Ricci flow equations differ only by a scaling in space and a reparametrization in time.*

Proof. Begin with the unnormalized equation $\frac{\partial}{\partial t} g(t) = -2\mathrm{Ric}_{g(t)}$. Let $\widetilde{g}(t) = \psi(t) g(t)$, where $\psi(t)$ is chosen so that $\mathrm{vol}_{\widetilde{g}(t)} M$ remains constant. Denote by $d\mu$ the volume element of g and by $d\widetilde{\mu}$ the volume element of $\widetilde{g}$. These are related by the equation $d\widetilde{\mu} = \psi^{\frac{n}{2}} d\mu$. We have $\frac{\partial}{\partial t} d\mu = -R d\mu$ by a similar computation to the one above, so the condition that $\mathrm{vol}_{\widetilde{g}(t)} M$

remains constant is equivalent to

$$\begin{aligned}
0 &= \frac{\partial}{\partial t}\int d\widetilde{\mu} \\
&= \int \frac{\partial}{\partial t}(\psi^{\frac{n}{2}} d\mu) \\
&= \psi^{\frac{n}{2}} \int \left(\frac{n}{2}\frac{\partial}{\partial t}\log\psi - R\right) d\mu.
\end{aligned}$$

Thus, we have the relation $\frac{\partial}{\partial t}\log\psi = \frac{2}{n}r$. We can now get an evolution equation for $\widetilde{g}$:

$$\begin{aligned}
\frac{\partial}{\partial t}\widetilde{g}(t) &= \frac{\partial}{\partial t}(\psi(t)g(t)) \\
&= \frac{\partial}{\partial t}\log\psi\widetilde{g}(t) - 2\psi(t)\mathrm{Ric}_{g(t)} \\
&= \psi\left(\frac{2}{n}\widetilde{r}(t)\widetilde{g}(t) - 2\mathrm{Ric}_{\widetilde{g}(t)}\right)
\end{aligned}$$

where in from the second to third line we used the above computation, and the fact that $\mathrm{Ric}_{\widetilde{g}} = \mathrm{Ric}_g$ and $\widetilde{r} = \frac{1}{\psi}r$. By taking the new time parameter $\widetilde{t} = \int \psi(t)dt$, we recover the normalized Ricci flow. □

6.2 The Evolution Equations for Curvature

We now derive the evolution equations for the curvatures in arbitrary dimension under the unnormalized Ricci flow. As a general rule, we will compare the time derivative to the Laplacian to get a reaction-diffusion equation. This will allow us to apply powerful maximum principle techniques in future sections.

Theorem 18. *The Riemann curvature tensor satisfies the evolution equation*

$$\begin{aligned}
\frac{\partial}{\partial t}R(X,Y,Z,W) &= (\triangle R)(X,Y,Z,W) \\
&+ \sum_{j,k} R(X,Y,e_j,e_k)R(Z,W,e_j,e_k) \\
&+ 2\sum_{j,k} R(X,e_k,Z,e_j)R(Y,e_k,W,e_j) \\
&- 2\sum_{j,k} R(X,e_k,W,e_j)R(Y,e_k,Z,e_j) \\
&- \sum_j Ric(X,e_j)R(e_j,Y,Z,W) - \sum_j Ric(Y,e_j)R(X,e_j,Z,W) \\
&- \sum_j Ric(Z,e_j)R(X,Y,e_j,W) - \sum_j Ric(W,e_j)R(X,Y,Z,e_j).
\end{aligned}$$

Proof. The proof is a rather tedious exercise in the symmetries of the curvature tensor. We first concern ourselves with the evolution of the connection. Define $A(X,Y) = \frac{\partial}{\partial t} D_X Y$. Note that the difference of two connections is tensorial, so A is tensorial. Observe that

$$g(A(X,Y),Z) = \frac{\partial}{\partial t}(g(D_X Y, Z)) + 2Ric(D_X Y, Z).$$

Applying the Koszul formula to the right hand side, we obtain

$$g(A(X,Y),Z) = -(D_X Ric)(Y,Z) - (D_Y Ric)(X,Z) + (D_Z Ric)(X,Y). \tag{15}$$

Next, observing that

$$\begin{aligned}\frac{\partial}{\partial t}(D_X D_Y Z - D_Y D_X Z - D_{[X,Y]} Z) &= A(X, D_Y Z) + D_X(A(Y,Z)) \\ &\quad - A(Y, D_X Z) - D_Y(A(X,Z)) - A([X,Y],Z) \\ &= (D_X A)(Y,Z) - (D_Y A)(X,Z),\end{aligned}$$

one computes

$$\begin{aligned}\frac{\partial}{\partial t} R(X,Y,Z,W) = &-2\sum_k R(X,Y,Z,e_k) Ric(e_k, W) \\ &- g((D_X A)(Y,Z) - (D_Y A)(X,Z), W).\end{aligned}$$

Applying the above formula (see numbering), we get

$$\begin{aligned}\frac{\partial}{\partial t} R(X,Y,Z,W) = &(D^2_{X,Z} Ric)(Y,W) - (D^2_{X,W} Ric)(Y,Z) \\ &- (D^2_{Y,Z} Ric)(X,W) + (D^2_{Y,W} Ric)(X,Z) \\ &- \sum_k R(X,Y,e_k,W) Ric(Z,e_k) - \sum_k R(X,Y,Z,e_k) Ric(e_k,W).\end{aligned}$$

The Laplacian of the curvature tensor is given by

$$(\triangle R)(X,Y,Z,W) = \sum_k (D^2_{e_k,e_k} R)(X,Y,Z,W).$$

We would like to get the twice-differentiated Ricci terms we see in the time derivative of the curvature tensor. The agenda is to first move one of the e_k derivates into the curvature tensor via the second Bianchi identity, commute derivatives, and repeat. The first application of the second Bianchi identity yields

$$(\triangle R)(X,Y,Z,W) = \sum_k (D^2_{e_k,X} R)(e_k,Y,Z,W) - \sum_k (D^2_{e_k,Y} R)(e_k,X,Z,W).$$

We focus on the first term and then antisymmetrize in X and Y. Commuting derivatives, we obtain

$$\begin{aligned}\sum_k (D^2_{e_k,X}R)(e_k,Y,Z,W) &= \sum_k (D^2_{X,e_k}R)(e_k,Y,Z,W)\\ &\quad - \sum_{j,k} R(X,e_k,e_k,e_j)R(e_j,Y,Z,W)\\ &\quad - \sum_{j,k} R(X,e_k,Y,e_j)R(e_k,e_j,Z,W)\\ &\quad - \sum_{j,k} R(X,e_k,Z,e_j)R(e_k,Y,e_j,W)\\ &\quad - \sum_{j,k} R(X,e_k,W,e_j)R(e_k,Y,Z,e_j).\end{aligned}$$

Using the second Bianchi identity on the first term, we get the desired second derivatives of the Ricci tensor, i.e. $(D^2_{X,Z}Ric)(Y,W) - (D^2_{X,W}Ric)(Y,Z)$. The second term is $\sum_j Ric(X,e_j)R(e_j,Y,Z,W)$. Before antisymmetrizing in X and Y, we make two observations. First, looking at the third term, antisymmetrizing $-R(X,e_k,Y,e_j)$ gives us $-R(X,Y,e_k,e_j)$. Second, the last two terms are antisymmetric in X and Y. This gives us

$$\begin{aligned}(\triangle R)(X,Y,Z,W) &= (D^2_{X,Z}Ric)(Y,W) - (D^2_{X,W}Ric)(Y,Z)\\ &\quad - (D^2_{Y,Z}Ric)(X,W) + (D^2_{Y,W}Ric)(X,Z)\\ &\quad - \sum_{j,k} R(X,Y,e_j,e_k)R(Z,W,e_j,e_k)\\ &\quad - 2\sum_{j,k} R(X,e_k,Z,e_j)R(Y,e_k,W,e_j)\\ &\quad + 2\sum_{j,k} R(X,e_k,W,e_j)R(Y,e_k,Z,e_j)\\ &\quad + \sum_j Ric(X,e_j)R(e_j,Y,Z,W) + \sum_j Ric(Y,e_j)R(X,e_j,Z,W).\end{aligned}$$

Combining this with the time derivative gives us the desired conclusion. □

Theorem 19. *The Ricci curvature tensor satisfies the evolution equation*

$$\begin{aligned}\frac{\partial}{\partial t}Ric(X,Z) &= (\triangle Ric)(X,Z) + 2\sum_{i,j} Ric(e_i,e_j)R(X,e_i,Z,e_j)\\ &\quad - 2\sum_j Ric(X,e_j)Ric(Z,e_j).\end{aligned}$$

Proof. It is tempting to simply contract the evolution equation for the Riemann curvature, but orthonormal vectors will not necessarily remain so while the metric is evolving. Take normal coordinates with $e_i = \partial_i$. One derives

$$\frac{\partial}{\partial t}g^{ij} = -g^{ip}g^{jq}\frac{\partial}{\partial t}g_{pq} = 2g^{ip}g^{jq}R_{pq},$$

by noting that $g_{pq}g^{ip}$ is constant and differentiating. Thus, we compute

$$\begin{aligned}\frac{\partial}{\partial t}Ric(X,Z) &= \frac{\partial}{\partial t}\left(g^{ij}R(X,\partial_i,Z,\partial_j)\right)\\ &= \sum_i \frac{\partial}{\partial t}R(X,e_i,X,e_i) + 2\sum_{i,j} Ric(e_i,e_j)R(X,e_i,Z,e_j).\end{aligned}$$

Using our formula for the evolution of the Riemann curvature, this becomes

$$\begin{aligned}\frac{\partial}{\partial t}Ric(X,Z) &= (\triangle Ric)(X,Z)\\ &+ \sum_{i,j,k} R(X,e_i,e_j,e_k)\left(R(X,e_i,e_j,e_k) - 2R(Z,e_k,e_j,e_i)\right)\\ &+ 2\sum_{i,j} Ric(e_i,e_j)R(X,e_i,Z,e_j)\\ &- 2\sum_j Ric(X,e_j)Ric(Z,e_j).\end{aligned}$$

To complete the proof, all we must show is that the second term is zero. By the antisymmetry of the Riemann curvature in the last two slots, we have

$$2\sum_{i,j,k} R(X,e_i,e_j,e_k)R(Z,e_k,e_j,e_i) = \sum_{i,j,k} R(X,e_i,e_j,e_k)\left(R(Z,e_k,e_j,e_i) - R(Z,e_j,e_k,e_i)\right).$$

The first Bianchi identity then tells us that this is

$$\sum_{i,j,k} R(X,e_i,e_j,e_k)R(Z,e_i,e_j,e_k)$$

as desired. □

Theorem 20. *The scalar curvature R evolves according to the equation*

$$\frac{\partial}{\partial t}R = \triangle R + 2|Ric|^2.$$

Proof. We compute

$$\begin{aligned}\frac{\partial}{\partial t}R &= \frac{\partial}{\partial t}\left(g^{ij}Ric(\partial_i,\partial_j)\right)\\ &= \sum_i \frac{\partial}{\partial t}Ric(e_i,e_i) + 2\sum_{i,j} Ric(e_i,e_j)Ric(e_i,e_j).\end{aligned}$$

The second term is the desired reaction term, so we must show that the first is simply $\triangle R$. Contracting our evolution equation for the Ricci tensor, we get

$$\sum_i \frac{\partial}{\partial t}Ric(e_i,e_i) = \triangle R + 2|Ric|^2 - 2|Ric|^2 = \triangle R,$$

completing the proof. □

6.3 Evolution Equations in Dimension 3

In 3 dimensions, the Riemann curvature tensor can be recovered from the Ricci tensor. Define the $(0,4)$ tensor W by

$$\begin{aligned} W(X,Y,Z,V) &= g(X,Z)Ric(Y,V) - g(X,V)Ric(Y,Z) - g(Y,Z)Ric(X,V) \\ &\quad + g(Y,V)Ric(X,Z) - \frac{1}{2}R\left(g(X,Z)g(Y,V) - g(X,V)g(Y,Z)\right). \end{aligned}$$

Proposition 7. *In dimension* 3*, we have* $R(X,Y,Z,V) = W(X,Y,Z,V)$.

Proof. One easily checks that W satisfies the same symmetries as the Riemann curvature tensors. It thus suffices to check that for any 2-dimensional subspace π of T_pM, we have $K(\pi) = \frac{W(x,y,x,y)}{g(x,x)g(y,y)-g(x,y)^2}$, where x and y form a basis for π. The expression on the right hand side is independent of the basis chosen because W satisfies the requisite symmetries. Choose an orthonormal basis e_1, e_2 of π, and extend it to an orthonormal basis $\{e_1, e_2, e_3\}$ of T_pM. We then compute

$$\begin{aligned} W(e_1,e_2,e_1,e_2) &= Ric(e_1,e_1) + Ric(e_2,e_2) - \frac{1}{2}R \\ &= \frac{1}{2}\left(Ric(e_1,e_1) + Ric(e_2,e_2) - Ric(e_3,e_3)\right) \\ &= \frac{1}{2}\sum_{k=1}^{3}\left(R(e_1,e_k,e_1,e_k) + R(e_2,e_k,e_2,e_k) - R(e_3,e_k,e_3,e_k)\right) \\ &= \frac{1}{2}\left(2R(e_1,e_2,e_1,e_2)\right) \\ &= K(\pi), \end{aligned}$$

completing the proof. ☐

With this information, we can simplify our equation for the evolution of the Ricci curvature.

Proposition 8. *In dimension* 3*, the Ricci curvature satisfies the evolution equation*

$$\begin{aligned} \frac{\partial}{\partial t}Ric(X,Y) &= \triangle Ric(X,Y) + 3RRic(X,Y) \\ &\quad - 6\sum_k Ric(X,e_k)Ric(Y,e_k) - g(X,Y)\left(R^2 - 2|Ric|^2\right). \end{aligned}$$

Proof. We begin with the evolution equation for Ric in n dimensions,

$$\begin{aligned} \frac{\partial}{\partial t}Ric(X,Y) &= \triangle Ric(X,Y) + 2\sum_{i,k} R(e_i,X,e_k,Y)Ric(e_i,e_k) \\ &\quad - 2\sum_k Ric(e_k,X)Ric(e_k,Y). \end{aligned}$$

In dimension 3, we can simplify the second term by applying the previous proposition, which tells us that

$$
\begin{aligned}
&\sum_{i,k} R(e_i, X, e_i, Y) Ric(e_i, e_k) \\
&\qquad = \sum_{i,k} \delta_{ik} Ric(X, Y) Ric(e_i, e_k) - \sum_{i,k} g(Y, e_i) Ric(X, e_k) Ric(e_i, e_k) \\
&\qquad - \sum_{i,k} g(X, e_k) Ric(Y, e_i) Ric(e_i, e_k) + \sum_{i,k} g(X, Y) Ric(e_i, e_k) Ric(e_i, e_k) \\
&\qquad - \frac{1}{2} R \left(\sum_{i,k} \delta_{ik} g(X, Y) Ric(e_i, e_k) - \sum_{i,k} g(e_i, Y) g(X, e_k) Ric(e_i, e_k) \right) \\
&\qquad = \frac{3}{2} R Ric(X, Y) - 2 \sum_k Ric(X, e_k) Ric(Y, e_k) \\
&\qquad + g(X, Y) |Ric|^2 - \frac{1}{2} R^2 g(X, Y).
\end{aligned}
$$

Substituting, we get the desired equation. □

6.4 Evolution Equations on Surfaces

In 2 dimensions it is reasonable to deal with the normalized Ricci flow because the evolution equations simplify greatly. Let K be the Gaussian curvature of our surface. The curvature tensor is

$$R(X, Y, Z, W) = K \left(g(X, Z) g(Y, W) - g(X, W) g(Y, Z) \right).$$

Indeed, one easily checks that the right hand side is curvature-like with the same sectional curvature as R. We then obtain

$$
\begin{aligned}
\mathrm{Ric}(X, Z) &= \sum_{i=1}^{2} R(X, e_i, Y, e_i) \\
&= \sum_{i=1}^{2} K \left(g(X, Z) g(e_i, e_i) - g(X, e_i) g(Y, e_i) \right) \\
&= K g(X, Z)
\end{aligned}
$$

and

$$
\begin{aligned}
R &= \sum_{i=1}^{2} \mathrm{Ric}(e_i, e_i) \\
&= 2K.
\end{aligned}
$$

One immediate consequence is that r is constant. Indeed, by the Gauss-Bonnet theorem, $\int R d\mu = 4\pi\chi(M)$ remains constant, so by volume preservation, $r = \int R d\mu / \int d\mu$ is constant. A second consequence is that scalar curvature is the only meaningful curvature quantity, so we only concern ourselves with the evolution of R. In normal coordinates, the formula for the Riemann curvature tensor is

$$R_{ijkl} = -\frac{1}{2}\left(\partial_i\partial_k g_{jl} - \partial_i\partial_l g_{jk} - \partial_j\partial_k g_{il} + \partial_j\partial_l g_{ik}\right).$$

We have ignored terms quadratic in the Christoffel symbols because those will remain zero when we take a time derivative. On a surface, $R = 2K = \frac{2R_{1212}}{g_{11}g_{22}-g_{12}^2}$. Using this relation, we obtain

$$R = \frac{2\partial_1\partial_2 g_{12} - \partial_1^2 g_{22} - \partial_2^2 g_{11}}{g_{11}g_{22} - g_{12}^2}.$$

For surfaces the Ricci curvature is $\frac{R}{2}g$, so the normalized Ricci flow equation $\partial_t g = \frac{2r}{n}g - 2\mathrm{Ric}$ becomes

$$\frac{\partial}{\partial t}g = (r - R)g.$$

Thus, we have

$$\begin{aligned}\frac{\partial}{\partial t}R = {} & \frac{2\partial_1\partial_2(r-R)g_{12} - \partial_1^2(r-R)g_{22} - \partial_2^2(r-R)g_{11}}{g_{11}g_{22} - g_{12}^2} \\ & - 2(r-R)R \\ & = (r-R)R + \triangle R - 2R(r-R)\end{aligned}$$

where we have used the fact that $g_{ij} = \delta_{ij}$ and that in normal coordinates, $\triangle R = g^{ij}\partial_i\partial_j R$. Thus, the evolution equation for R is

$$\frac{\partial}{\partial t}R = \triangle R + R(R - r). \tag{16}$$

We conclude this section by producing an explicit self-similar solution to the Ricci flow on $\mathbb{R}^2$ which is analagous to the Grim reaper for the curve-shortening flow. We seek a metric which is the euclidean metric scaled by a positive radially symmetric function, i.e.

$$g_{ij} = \psi(r,t)\delta_{ij}.$$

It is a well-known fact under a conformal change of metric $g = \psi h$ on a surface, the scalar curvatures are related by

$$R_g = \frac{1}{\psi}\left(R_h - \triangle_h(\log\psi)\right).$$

As a corollary, we find that g is a solution to the (unnormalized) Ricci flow if and only if

$$\begin{aligned}\frac{\partial}{\partial t}\psi h &= \frac{\partial}{\partial t}g(t) \\ &= -R_g g \\ &= -\left(R_h - \triangle_h(\log\psi)\right)h,\end{aligned}$$

which holds if and only if

$$\frac{\partial}{\partial t}\psi = \triangle_h(\log\psi) - R_h.$$

In our case, h is the euclidean metric, so R_h vanishes and we seek a solution to the PDE

$$\frac{\partial}{\partial t}\psi = \triangle_{eucl}(\log\psi).$$

We claim that the function

$$\psi(r,t) = \frac{4}{e^t + r^2}$$

satisfies this condition. Indeed, we have

$$\frac{\partial\psi}{\partial t} = -\frac{4e^t}{\left(e^t + r^2\right)^2}$$

and

$$\begin{aligned}
\triangle_{eucl}(\log\psi) &= -\triangle_{eucl}\left(\log(e^t + r^2)\right)\\
&= -\left(\partial_r^2 + \frac{1}{r}\partial_r\right)(\log(e^t + r^2))\\
&= -\left(\frac{2}{e^t + r^2} - \frac{4r^2}{\left(e^t + r^2\right)^2}\right) - \frac{2}{e^t + r^2}\\
&= -\frac{4e^t}{\left(e^t + r^2\right)^2}
\end{aligned}$$

as desired. The solution

$$\left(\mathbb{R}^2, \frac{4}{e^t + r^2}\delta_{ij}\right)$$

is known as the Cigar soliton.

7 Ricci Flow on 2-Spheres with Positive Scalar Curvature

In this section, we discuss Ricci flow theory for compact surfaces with positive scalar curvature developed by Hamilton [10]. We first discuss several interesting techniques used to obtain positive uniform bounds for R. These bounds then imply that the limit metric must be a member of a class of self-similar solutions to the Ricci flow known as gradient Ricci solitons. Finally, we show that such metrics on S^2 must be round.

7.1 The Scalar Maximum Principle

The most important tool in this discussion will be the scalar maximum principle for solutions to heat-type equations. The following ODE comparison version will be the most useful for our purposes.

Theorem 21. *Suppose $u(x,t) : M \times [0,T] \to \mathbb{R}$ satisfies the differential inequality*

$$\frac{\partial}{\partial t}u \geq \triangle u + \langle X(t), \nabla u \rangle + F(u)$$

where $X(t)$ is a time-dependent vector field and F is a locally Lipschitz function. Let $h(t)$ be the solution to the ODE $\frac{d}{dt}h = F(h)$, with $u(\cdot, 0) \geq h(0)$. Then $u \geq h$ for all $x \in M$ and $t \in [0,T]$.

Proof. This theorem is a consequence of the simple fact that at a local minimum, the Laplacian is nonnegative and the gradient vanishes. We first give a heuristic argument. If $u = h$ at a first time $t_0 > 0$ at a point x_0, then for times leading up to t_0 we have $(u-h)(t, x_0) > 0$ so our time derivative is nonpositive, and we are at a spatial minimum. It would follow that

$$\begin{aligned}
0 &\geq \frac{\partial}{\partial t}(u-h)(x_0,t_0) \\
&\geq \triangle(u-h)(x_0,t_0) + \langle X(t_0), \nabla(u-h)(x_0,t_0)\rangle + F(u(x_0,t_0)) - F(h(x_0,t_0)) \\
&\geq F(u(x_0,t_0)) - F(u(x_0,t_0)) \\
&= 0.
\end{aligned}$$

To get the strict inequality we desire, we instead consider the function $u_\epsilon = u + \epsilon(\delta + t)$. Note that $M \times [0,T]$ is compact, so we can choose a uniform Lipschitz constant K for F. We will choose a small δ depending only on K so that $u_\epsilon - h > 0$ for $t \in [0, \delta]$; we can let ϵ go to 0 to prove the claim on $[0,\delta]$ and then repeat the argument with the same δ to cover the entire time interval.

Note that $u_\epsilon > h$ at $t = 0$. Suppose by way of contradiction that $u_\epsilon = h$ for a first time t_0 at a point x_0. Then at (x_0, t_0) we would have

$$\begin{aligned}
0 &\geq \frac{\partial}{\partial t}(u_\epsilon - h) \\
&\geq \epsilon + \triangle(u_\epsilon - h) + \langle X, \nabla(u_\epsilon - h)\rangle + F(u_\epsilon - \epsilon(\delta+t)) - F(h) \\
&\geq \epsilon - K|u_\epsilon - h - \epsilon(\delta + t)| \\
&= \epsilon\left(1 - K|\delta + t|\right).
\end{aligned}$$

Letting $\delta < \frac{1}{2K}$, this last expression is strictly positive on $[0, \delta]$, proving the theorem. $\square$

It is worth noting that the analagous conclusion holds if

$$\frac{\partial}{\partial t} u \leq \triangle u + \langle X(t), \nabla u \rangle + F(u)$$

and h satisfies the same ODE with $h(0) \geq u(\cdot, 0)$. Namely, $u \leq h$ for all subsequent times.

Remark 4. *Recall that on a surface, the scalar curvature evolution equation is*

$$\frac{\partial}{\partial t} R - \triangle R + R(R - r).$$

The function $R(R-r)$ is locally Lipschitz, so we may apply the scalar maximum principle. Since 0 is a solution to the ODE $\frac{dh}{dt} = h(h-r)$, we immediately have preservation of positive or negative scalar curvature. The case where $R < 0$ to begin with is especially easy because $R = r$ is an attractive fixed point for the ODE. Solving the above ODE for h, we obtain the general solution

$$h(t) = \frac{r}{1 - ke^{rt}} \tag{17}$$

where k is the integration constant. If $R \in [-C, -\epsilon]$ at time $t = 0$, then by the scalar maximum principle we have

$$\frac{r}{1 - \left(1 + \frac{r}{C}\right) e^{rt}} \leq R \leq \frac{r}{1 - \left(1 + \frac{r}{\epsilon}\right) e^{rt}}.$$

Since $r < 0$ and $1 + \frac{r}{C}$, $1 + \frac{r}{\epsilon}$ are bounded above by 1, this gives us long-time existence of solutions and exponential convergence to a metric of constant curvature r.

Remark 5. *Applying the maximum principle to R is not sufficient to conclude longtime existence in all cases. Instead, one may consider the evolution of a potential function f satisfying $\triangle f = R - r$ and use the maximum principle to obtain the result*

$$-C \leq R \leq Ce^{rt} + r$$

which gives longtime existence for any compact surface. This also implies that for $r < 0$, the scalar curvature becomes negative in finite time, and we can apply the above remark to get convergence of the metric to one of constant curvature.

7.2 The Harnack Inequality

Let M be a compact Riemannian manifold of dimension n with positive Ricci curvature, and let f be a real-valued function on M satisfying the standard heat equation, i.e. $\frac{\partial}{\partial t} f = \triangle f$. The Harnack inequality for f puts allows us to compare values of f at neighboring points and times:

Theorem 22. *Let f be as above. Then*

$$f(\xi, \tau) \leq \left(\frac{T}{\tau}\right)^{\frac{n}{2}} e^{\frac{d(X,\xi)^2}{4(T-\tau)}} f(X, T)$$

for any times $\tau < T$ and points ξ and X.

The approach of the proof is to get a lower bound for $\frac{\partial}{\partial t}\log f$ via a maximum principle argument, and integrate along a minimizing geodesic connecting ξ and X, the existence of which is guaranteed by the Hopf-Rinow theorem. (For a proof of this fact, see for instance Milnor [14].)

Proof. Let $L = \log f$. We have $\frac{\partial L}{\partial t} = \frac{\triangle f}{f}$, and $\triangle L = \frac{\triangle f}{f} - \frac{|\nabla f|^2}{f^2}$. Combining these gives the evolution equation for L:

$$\frac{\partial L}{\partial t} = \triangle L + |\nabla L|^2.$$

Now, we define $Q = \triangle L$ and derive an evolution equation that will bound Q below. We compute $\frac{\partial Q}{\partial t} = \triangle(\triangle L + |\nabla L|^2) = \triangle Q + \triangle|\nabla L|^2$. We revert to coordinates to compute the second term:

$$\begin{aligned} g^{pq}g^{ij}D_pD_q(D_iLD_jL) &= 2g^{pq}g^{ij}D_p(D_qD_iLD_jL) \\ &= 2|\text{Hess}L|^2 + 2g^{pq}g^{ij}(D_pD_qD_iLD_jL) \end{aligned}$$

Commuting i and q in the last term, applying the identity $[D_p, D_i]D_qL = g^{lm}R_{piql}D_mL$ and contracting, we get $2g^{ij}g^{lm}R_{il}D_jLD_mL + 2\langle\nabla Q, \nabla L\rangle \geq 2\langle\nabla Q, \nabla L\rangle$ since we are assuming the Ricci curvature is positive. Furthermore, by the Cauchy-Schwartz inequality, we have $\frac{1}{n}Q^2 \leq |\text{Hess}L|^2$. This gives us the differential inequality

$$\frac{\partial Q}{\partial t} \geq \triangle Q + 2\langle\nabla Q, \nabla L\rangle + \frac{2}{n}Q^2.$$

Note that the general solution to the ODE $\frac{dh}{dt} = \frac{2}{n}h^2$ is

$$h(t) = \frac{-n}{2t+k}$$

where k is the integration constant. If we choose $k = 0$, then h blows down as t goes to 0, so Q is greater than h for small times. Applying the standard maximum principle, we obtain the estimate

$$Q \geq -\frac{n}{2t}.$$

Thus, we have the desired lower bound

$$\frac{\partial L}{\partial t} \geq |\nabla L|^2 - \frac{n}{2t}.$$

We parametrize a minimizing geodesic connecting ξ and X by time, $t \in [\tau, T]$, with parameter proportional to arclength so that $d(\xi, X) = \frac{ds}{dt}(T - \tau)$. We have

$$\begin{aligned} \frac{dL}{dt} &= \frac{\partial L}{\partial t} + \frac{\partial L}{\partial s}\frac{ds}{dt} \\ &\geq -\frac{n}{2t} + \left(\frac{\partial L}{\partial s}\right)^2 + \frac{\partial L}{\partial s}\frac{ds}{dt} \\ &\geq -\frac{n}{2t} - \frac{1}{4}\left(\frac{ds}{dt}\right)^2 \\ &= -\frac{n}{2t} - \frac{d(\xi, X)^2}{4(T-\tau)^2}. \end{aligned}$$

Integrating, we get

$$L(X,T) - L(\xi,\tau) \geq \log\left(\frac{\tau}{T}\right)^{\frac{n}{2}} - \frac{d(X,\xi)^2}{4(T-\tau)}.$$

Exponentiating gives us the Harnack inequality. □

We now prove a Harnack inequality for the curvature R of a compact surface evolving by the Ricci flow using essentially the same technique as above. However, we must first define a new notion of space-time distance between points, since the metric is changing. For this, we define

$$\rho(\xi,\tau,X,T) = \inf_{\gamma} \int_{\tau}^{T} \left(\frac{ds}{dt}\right)^2, \tag{18}$$

where the infimum is taken over all curves parametrized by time connecting ξ and X, and $\frac{ds}{dt}$ is the space velocity with respect to the metric at time t.

Theorem 23. *The scalar curvature R of a compact surface evolving by the Ricci flow satisfies*

$$R(\xi,\tau) \leq \left(\frac{e^{rT}-1}{e^{rt}-1}\right) e^{\frac{\rho}{4}} R(X,T)$$

for any times $\tau < T$ and points ξ and X.

We will be commuting time derivatives and Laplacian operators, so the following lemma is very useful:

Lemma 4. *Suppose $\frac{\partial}{\partial t} g = fg$ on a manifold of dimension n. Then we have*

$$\left[\frac{\partial}{\partial t}, \triangle\right]\phi = -f \triangle \phi - \left(1 - \frac{n}{2}\right)\langle \nabla f, \nabla \phi\rangle.$$

Proof. We can make the computation easier using integration by parts. Let ψ be any test function. We then have

$$\int_M (\triangle\phi)\psi d\mu = -\int_M g^{ij}\partial_i\phi\partial_j\psi d\mu.$$

Note that $\frac{\partial}{\partial t} d\mu = \frac{n}{2} f d\mu$ and that $\frac{\partial}{\partial t} g^{ij} = -f g^{ij}$. Differentiating the above identity with respect to time, we obtain

$$\begin{aligned}
\int \left(\frac{\partial}{\partial t}\triangle\phi\right)\psi d\mu + \int \frac{n}{2} f\,(\triangle\phi)\,\psi d\mu &= -\int g^{ij}\partial_i(\frac{\partial}{\partial t}\phi)\partial_j\psi d\mu + \int (1-\frac{n}{2}) f g^{ij}\partial_i\phi\partial_j\psi d\mu \\
&= \int \triangle\left(\frac{\partial}{\partial t}\phi\right)\psi d\mu - \int \left(1-\frac{n}{2}\right) f(\triangle\phi)\psi d\mu \\
&\quad - \int \left(1-\frac{n}{2}\right) g^{ij}\partial_i f\partial_j\phi\psi d\mu.
\end{aligned}$$

Rearranging, we get

$$\int \left(\left[\frac{\partial}{\partial t}, \triangle\right]\phi\right)\psi d\mu = \int \left(-f\triangle\phi - (1-\frac{n}{2})\langle\nabla f, \nabla\phi\rangle\right)\psi d\mu.$$

This holds for all test functions ψ, completing the proof. □

Note that in dimension 2, the term involving the gradient of f disappears, making the computation of evolution equations a bit easier. An immediate corollary is that for the Ricci flow on surfaces we have

$$\frac{\partial}{\partial t} \triangle \phi = \triangle \left(\frac{\partial}{\partial t} \phi \right) + (R - r) \triangle \phi. \tag{19}$$

We proceed to the proof of the Harnack inequality.

Proof. As before, let $L = \log R$. Then, using the evolution equation for R, we get $\frac{\partial L}{\partial t} = \frac{\triangle R}{R} + R - r$ and $\triangle L = \frac{\triangle R}{R} - |\nabla L|^2$. Combining these two computations gives

$$\frac{\partial L}{\partial t} = \triangle L + |\nabla L|^2 + R - r.$$

Now, we define $Q = \triangle L + R - r$, and we derive an evolution equation for Q which will give us a lower bound for $\frac{\partial L}{\partial t}$. First, using our commutator rule we obtain

$$\begin{aligned} \frac{\partial Q}{\partial t} &= \frac{\partial}{\partial t} \triangle L + \frac{\partial R}{\partial t} \\ &= (R - r) \triangle L + \triangle(\triangle L + |\nabla L|^2 + R - r) + \triangle R + R(R - r). \end{aligned}$$

Next, we compute $\triangle Q = \triangle(\triangle L) + \triangle R$. Putting these together, we get

$$\frac{\partial Q}{\partial t} = \triangle Q + \triangle |\nabla L|^2 + (R - r) \triangle L + \triangle R + R(R - r).$$

Recall that

$$\triangle |\nabla L|^2 = 2|\text{Hess} L|^2 + g^{ij} g^{lm} R_{il} D_j L D_m L + 2\langle \nabla(\triangle L), \nabla L \rangle.$$

We will get rid of the Hessian term first; By Cauchy-Schwarz, we have $2|\text{Hess}L|^2 \geq (\triangle L)^2$. Next, using that $R_{il} = \frac{R}{2} g_{il}$, the second term becomes $R|\nabla L|^2$. We also want to get a gradient term, so we write the last one as $2\langle \nabla Q, \nabla L \rangle - 2R|\nabla L|^2$. This yields

$$\triangle |\nabla L|^2 \geq (\triangle L)^2 - R|\nabla L|^2 + 2\langle \nabla Q, \nabla L \rangle.$$

Finally, we want to transform R derivatives to L derivatives via the identity

$$\triangle R = R \triangle L + R|\nabla L|^2.$$

Substituting the previous two expressions into our formula for $\frac{\partial Q}{\partial t}$, we obtain

$$\frac{\partial Q}{\partial t} \geq \triangle Q + 2\langle \nabla Q, \nabla L \rangle + (\triangle L)^2 + 2(R - r) \triangle L + (R - r)^2 + r(\triangle L + R - r).$$

The last several terms are $Q^2 + rQ$, finally leaving us with

$$\frac{\partial Q}{\partial t} \geq \triangle Q + 2\langle \nabla Q, \nabla L \rangle + Q(Q + r).$$

Note that the general solution of the ODE $\frac{dh}{dt} = h(h + r)$ is

$$h(t) = \frac{-re^{rt}}{e^{rt} + k},$$

where k is our integration constant. By choosing $k = -1$, we guarantee that h blows down as t goes to zero, so Q is certainly greater than $h(t)$ for small t. Applying the scalar maximum principle, we get

$$Q \geq \frac{-re^{rt}}{e^{rt} - 1},$$

and thus have the desired lower bound for $\frac{\partial L}{\partial t}$.

As before, along a curve connecting (ξ, τ) and (X, T), we have

$$\begin{aligned}\frac{dL}{dt} &= \frac{\partial L}{\partial t} + \frac{\partial L}{\partial s}\frac{ds}{dt} \\ &\geq \frac{-re^{rt}}{e^{rt} - 1} + \left(\frac{\partial L}{\partial s}\right)^2 + \frac{\partial L}{\partial s}\frac{ds}{dt} \\ &\geq \frac{-re^{rt}}{e^{rt} - 1} - \frac{1}{4}\left(\frac{ds}{dt}\right)^2.\end{aligned}$$

Integrating from τ to T, we get

$$L(X, T) - L(\xi, \tau) \geq \log\left(\frac{e^{r\tau} - 1}{e^{rT} - 1}\right) - \frac{1}{4}\int_\tau^T \left(\frac{ds}{dt}\right)^2 dt.$$

Since this holds for all paths, we can replace the last term by $\frac{\rho}{4}$; exponentiating us gives us the Harnack inequality. □

7.3 Entropy Monotonicity

The Harnack inequality gives us a tool to compare curvatures in a neighborhood a point on the sphere. In this section, we prove that the entropy functional $\int R \log R d\mu$ is monotonically decreasing. Since $R \log R$ is bounded below by $\frac{1}{e}$, this tells us that R can only be large on regions of small area.

Theorem 24. *Under the Ricci flow on a compact surface, the quantity $\int R \log R d\mu$ is monotonically decreasing.*

Proof. First, note that

$$\begin{aligned}\frac{d}{dt}\int R \log R d\mu &= \int \left((\triangle R + R(R - r))(\log R + 1) + R \log R(r - R)\right) d\mu \\ &= \int \left(\triangle R \log R + \triangle R + R(R - r)\right) d\mu \\ &= \int \left(R(R - r) - \frac{|\nabla R|^2}{R}\right) d\mu\end{aligned}$$

where we have used the computation

$$\begin{aligned}\triangle(R\log R) &= g^{pq}D_pD_q(R\log R)\\ &= g^{pq}D_p(D_qR\log R + D_qR)\\ &= \triangle R\log R + \frac{|\nabla R|^2}{R} + \triangle R.\end{aligned}$$

Hence, it suffices to prove that $\int\left(R(R-r) - \frac{|\nabla R|^2}{R}\right)d\mu$ is nonpositive.

Note that $QR = \triangle R + R(R-r) - \frac{|\nabla R|^2}{R}$, so $\int QRd\mu = \frac{d}{dt}\int R\log Rd\mu$. We will show that the quantity $Z = \int QRd\mu / \int Rd\mu$ satisfies the differential inequality

$$\frac{dZ}{dt} \geq Z^2 + rZ,$$

which means that if Z were ever positive, it would blow up in finite time, a contradiction of longtime existence. Hence, $\frac{d}{dt}\int R\log Rd\mu = \int QRd\mu \leq 0$, proving the claim.

Since $\int Rd\mu$ is constant by Gauss-Bonnet,

$$\frac{dZ}{dt}\int Rd\mu = \int (Q\frac{\partial R}{\partial t} + R\frac{\partial Q}{\partial t} + QR(r-R))d\mu.$$

Using the nonnegativity of R and the computations from last section, we have

$$R\frac{\partial Q}{\partial t} \geq R\triangle Q + 2\langle\nabla Q, \nabla R\rangle + RQ^2 + RQr$$

and

$$Q\frac{\partial R}{\partial t} = Q\triangle R + QR^2 - RQr.$$

Adding all terms together, we get

$$R\frac{\partial Q}{\partial t} + Q\frac{\partial R}{\partial t} \geq \triangle(RQ) + RQ^2 + QR^2.$$

Substituting into our equation for $\frac{dZ}{dt}$, and applying Cauchy-Schwarz, we obtain

$$\begin{aligned}\frac{dZ}{dt}\int Rd\mu &\geq \int RQ^2d\mu + rZ\int Rd\mu\\ &\geq \left(\int QRd\mu\right)^2 \Big/ \left(\int Rd\mu\right) + rZ\int Rd\mu\\ &= \left(Z^2 + rZ\right)\int Rd\mu,\end{aligned}$$

completing the proof. $\square$

7.4 Scalar Curvature is Bounded

Theorem 25. *If $R(\cdot, 0) > 0$ on S^2 then R is uniformly bounded between some positive constants c and C.*

Proof. We first produce an upper bound for R. The idea is to use the Harnack inequality to show that $R(t)$ is comparable to $R_{max}(t)$ in a sufficiently large ball around the point at which R_{max} is attained, and then to apply entropy monotonicity to to conclude that R_{max} cannot be arbitrarily large. First, note that at a local maximum the Laplacian is nonpositive, so by our evolution equation

$$\frac{\partial}{\partial t} R_{max} \leq R_{max}^2.$$

Let $1 < \tau < T$, and observe that if $T - \tau = \frac{1}{2R_{max}(\tau)}$ we can integrate the above differential inequality to obtain

$$\frac{1}{2} R_{max}(T) \leq R_{max}(\tau).$$

Suppose $R_{max}(\tau) = R(\xi, \tau)$. We can control the values of R around ξ at time T using the Harnack inequality, which requires an esimate on the spacetime distance between X and ξ. By our evolution equation for the metric, we have for any vector V and time $t_0 \in (\tau, T)$ the relation

$$g(V, V)|_T = e^{\int_{t_0}^T (r-R)dt} g(V, V)|t_0.$$

Noting that $R_{max}(t) \leq 2R_{max}(\tau)$ for $t \in (t_0, T)$ and that $T - t_0 \leq \frac{1}{2R_{max}(\tau)}$, we have

$$\begin{aligned} \int_{t_0}^T (r - R)dt &\geq - \int_{t_0}^T R_{max}(t)dt \\ &\geq -2(T - t_0)R_{max}(\tau) \\ &\geq -1. \end{aligned}$$

This tells us that distances cannot change too much over a bounded time interval, i.e.

$$g(V, V)|_T \geq \frac{1}{e} g(V, V)_{t_0}.$$

Hence, if we take a minimizing geodesic γ connecting X and ξ at time T, parametrized with speed $\frac{d(X,\xi)|_T}{T-\tau}$, we have

$$\begin{aligned} \rho(\xi, \tau, X, T) &\leq \int_\tau^T \left(\frac{ds}{dt}\right)^2 |_t dt \\ &\leq e^2 \int_\tau^T \left(\frac{ds}{dt}\right)^2 |_T dt \\ &= \frac{e^2 d(X, \xi)^2|_T}{T - \tau}. \end{aligned}$$

We conclude that ρ is bounded for X in a time T ball of radius $\delta = \frac{\pi}{\sqrt{R_{max}(T)/2}}$ around ξ. Using this fact, that $\tau > 1$ and that $T - \tau$ is bounded, we can apply the Harnack inequality: there is some positive constant C_0 (independent of τ) such that

$$R(\xi, \tau) \leq C_0 R(X, T)$$

for X in the time T ball $B_\delta(\xi)$. Using this and that $R_{max}(T) \le 2R_{max}(\tau)$, we have that $R(X,T) \ge C_1 R_{max}(T)$ on $B_\delta(\xi)$ for some uniform positive constant C_1.

The injectivity radius is at least δ at time T by an estimate of Klingenberg, so the area of $B_\delta(\xi)$ is bounded below by $\frac{C_2}{R_{max}(T)}$ for some uniform positive constant C_2. Using entropy monotonicity, we thus have some positive constant K so that

$$\begin{aligned} K &\ge \int_{B_\delta(\xi)} R \log R d\mu \\ &\ge C_1 R_{max}(T) \log(C_1 R_{max}(T)) \frac{C_2}{R_{max}(T)} \\ &= C_1 C_2 \left(\log(C_1) + \log(R_{max}(T))\right). \end{aligned}$$

This estimate holds for all $\tau \ge 1$, so if R_{max} got arbitrarily large, we would have a contradiction, proving that R remains bounded above.

Equipped with an upper bound on R, we have a uniform lower bound for the injectivity radius by Klingenberg's estimate. Thus, if the diameter of our sphere got arbitrarily large, we could take some small ϵ less than the injectivity radius and arbitrarily long geodesics and then construct arbitrarily many disjoint balls of radius ϵ centered along these geodesics. If a geodesic is sufficiently long, the areas of these balls sum to something larger than the constant area of our sphere, a contradiction. This means we have an upper bound on the diameter of our sphere.

Suppose now that we have any time $t \ge 1$. Equipped with our diameter bound, the spacetime distance between any two points (x,t) and $(y,t+1)$ is bounded by some uniform positive constant. Thus, by the Harnack inequality (and using that $t \ge 1$) we have

$$R(x,t) \le C R(y,t+1).$$

If R could get arbitrarily close to 0, there would be some time t_0 such that $R_{min}(t_0) < \frac{r}{C}$; but then, we would have $R_{max}(t_0 - 1) < r$, a contradiction, completing the proof. $\square$

We briefly discuss an important consequence of these bounds on R. Let $f : M \to \mathbb{R}$ satisfy

$$\triangle f = R - r.$$

Denote by M the trace-free part of the Hessian of f, i.e.

$$M = D^2 f - \frac{1}{2} \triangle f g.$$

With positive uniform bounds on the scalar curvature, one can show via the maximum principle that the quantity $|M|^2$ converges exponentially to 0. See Hamilton [10] for details. Modifying the flow by a one-parameter family of diffeomorphisms ϕ_t generated by $\xi = \nabla f$, so that $\tilde{g} = \phi_t^* g$, one obtains the evolution equation

$$\begin{aligned}
\frac{\partial}{\partial t}\tilde{g}(X,Y) &= (L_\xi \tilde{g})(X,Y) + (r-R)\tilde{g}(X,Y) \\
&= \xi(\tilde{g}(X,Y)) - \tilde{g}(L_\xi X, Y) - \tilde{g}(X, L_\xi Y) - \triangle f \tilde{g}(X,Y) \\
&= \tilde{g}(D_\xi X - [\xi, X], Y) + \tilde{g}(X, D_\xi Y - [\xi, Y]) - \triangle f \tilde{g}(X,Y) \\
&= \tilde{g}(D_X \xi, Y) + \tilde{g}(X, D_Y \xi) - \triangle f \tilde{g}(X,Y) \\
&= XY(f) - D_X Y(f) + YX(f) - D_Y X(f) - \triangle f \tilde{g}(X,Y) \\
&= 2D^2 f(X,Y) - \triangle f \tilde{g}(X,Y) \\
&= 2M(X,Y).
\end{aligned}$$

Note that M is diffeomorphism invariant, so it still converges exponentially to 0. By obtaining similar estimates for the derivatives of M, one gets smooth convergence of $\tilde{g}$ to a metric for which M vanishes identically. We show in the next section such metrics, known as gradient Ricci solitons, must be round.

7.5 Gradient Ricci Solitons on S^2 are Round

Assume there exists a function f on (S^2, g_0) such that

$$2D^2 f = (R - r) g_0.$$

Such a metric is called a gradient Ricci soliton. Let $\xi = -\nabla f$, and let ψ_t be the one-parameter family of diffeomorphisms generated by ξ. Then we can obtain a self-similar solution to the normalized Ricci flow by defining

$$g(t) = \psi_t^* g_0.$$

This is proven by the following easy computation:

$$\begin{aligned}
\frac{\partial}{\partial t}(\psi_t^* g_0)(X,Y) &= (L_\xi g)(X,Y) \\
&= g(D_X \xi, Y) + g(X, D_Y \xi) \\
&= -2D^2 f(X,Y) \\
&= (r - R) g(X,Y).
\end{aligned}$$

In this section, we will show that such metrics are round:

Theorem 26. *Suppose (S^2, g) is a gradient Ricci soliton. Then $R \equiv r$.*

We follow an argument due to Chen, Lu and Tian [17] in Brendle [3]. For the proof, we equip S^2 with an almost-complex structure J obtained by defining Jv to be the vector of the same length as v such that $g(v, Jv) = 0$ and (v, Jv) is positively oriented. One easily checks that $J^2 = -Id$, that $g(v,w) = g(Jv, Jw)$ and that J is parallel.

A first consequence of the soliton equation is that $J\xi$ generates a one-parameter family of isometries ϕ_τ. Indeed, we compute

$$\begin{aligned}(L_{J\xi}g)(X,Y) &= g(JD_X\xi, Y) + g(X, JD_Y\xi)\\ &= -g(D_X\xi, JY) - g(JX, D_Y\xi)\\ &= D^2f(X,JY) + D^2f(JX,Y)\\ &= \frac{1}{2}(R-r)\left(g(X,JY) + g(JX,Y)\right)\\ &= 0.\end{aligned}$$

We would like to show that ϕ_τ is the identity for all τ, from which it would follow that f is constant, and hence that $R \equiv r$. Let p and q be two distinct critical points of f, so that $J\xi|_p = J\xi|_q = 0$. Then the flow ϕ_τ of $J\xi$ fixes these points. One way to show that ϕ_τ is the identity, then, is to show that $(d\phi_\tau)_p$ is the identity on T_pM for all τ. If this were the case, then ϕ_τ would map the geodesic $\exp_p(tv)$ to another geodesic $\alpha(t)$ such that $\alpha(0) = p$ and $\alpha'(0) = (d\phi_\tau)_p(v) = v$. By ODE uniqueness, these geodesics must coincide everywhere, i.e.

$$\phi_\tau(\exp_p(v)) = \exp_p(v).$$

Since the exponential map is surjective, ϕ_τ would then be the identity.

We apply the soliton equation to find the τ for which $(d\phi_\tau)_p$ is the identity. Define

$$\alpha = \frac{1}{2}(r-R)(p)$$

and

$$\beta = \frac{1}{2}(r-R)(q).$$

Then $-(D^2f)_p(v,v) = \alpha g(v,v)$ and $-(D^2f)_q(w,w) = \beta g(w,w)$ for all $v \in T_pM$ and $w \in T_qM$. It follows that

$$(d\phi_\tau)_p(v) = \cos(\alpha\tau)v + \sin(\alpha\tau)Jv$$

and

$$(d\phi_\tau)_q(w) = \cos(\beta\tau)w + \sin(\beta\tau)Jw.$$

If $\frac{\alpha\tau}{2\pi}$ is an integer, then $(d\phi_\tau)_p$ is the identity on $T_p(M)$ and hence ϕ_τ is the identity. One has the same result if $\frac{\beta\tau}{2\pi}$ is an integer. Conversely, if ϕ_τ is the identity, then $\frac{\alpha\tau}{2\pi}$ and $\frac{\beta\tau}{2\pi}$ are integers. We may summarize this relation by

$$\frac{\alpha\tau}{2\pi} \in \mathbb{Z} \Longleftrightarrow \phi_\tau = \mathrm{id} \Longleftrightarrow \frac{\beta\tau}{2\pi} \in \mathbb{Z}.$$

Two immediate corollaries are that $\alpha^2 = \beta^2$ and that if $\alpha = \beta = 0$, then ϕ_τ is the identity for all τ. We can now prove that gradient Ricci solitons are round.

Proof. Let $\sigma = d(p,q)$ and join p and q by a minimizing unit-speed geodesic $\gamma : [0,\sigma] \to S^2$. Observe that the curves $\phi_\tau(\gamma(s))$ are geodesics because the ϕ_τ are isometries. Thus, this family of curves generates a Jacobi field $V(s)$ defined by

$$V(s) = \frac{\partial}{\partial\tau}\phi_\tau(\gamma(s))|_{\tau=0} = J\xi|_{\gamma(s)}.$$

Note that $V(0) = V(\sigma) = 0$, so that p and q are conjugate points. Furthermore, we have

$$\begin{aligned}\frac{d}{ds}g(V(s), \gamma'(s)) &= g(D_{\gamma'(s)}V(s), \gamma'(s))\\ &= \frac{1}{2}(L_{J\xi}g)(\gamma'(s), \gamma'(s))\\ &= 0,\end{aligned}$$

so V is always perpendicular to $\gamma'(s)$. We can thus find some function $u(s)$ so that

$$V(s) = u(s)J\gamma'(s).$$

Applying the Jacobi equation, we obtain

$$u''(s) + \frac{1}{2}R(\gamma(s))u(s) = 0.$$

The soliton equation allows us to replace the scalar curvature by a derivative of u. Note that

$$\begin{aligned}u(s) &= g(V(s), J\gamma'(s))\\ &= g(J\xi|_{\gamma(s)}, J\gamma'(s))\\ &= g(\xi|_{\gamma(s)}, \gamma'(s)).\end{aligned}$$

Differentiating both sides with respect to s, we get

$$\begin{aligned}u'(s) &= g(D_{\gamma'(s)}\xi, \gamma'(s))\\ &= -(D^2f)(\gamma'(s), \gamma'(s))\\ &= \frac{1}{2}(r - R).\end{aligned}$$

This immediately tells us that $u'(0) = \alpha$ and $u'(\sigma) = \beta$, so that $u'(0)^2 = u'(\sigma)^2$. Furthermore, $\frac{1}{2}R = \frac{1}{2}(r - 2u')$. Substituting this into the Jacobi equation, we get

$$u''(s) + \frac{1}{2}ru(s) = u(s)u'(s).$$

Using this equation, we compute

$$\begin{aligned}0 &= u'(\sigma)^2 - u'(0)^2\\ &= \int_0^\sigma \frac{d}{ds}\left(u'(s)^2\right) ds\\ &= 2\int_0^\sigma u'(s)u''(s)ds\\ &= 2\int_0^\sigma u'(s)^2u(s)ds - \frac{r}{2}\int_0^\sigma \frac{d}{ds}(u(s)^2)ds\\ &= 2\int_0^\sigma u'(s)^2u(s)ds,\end{aligned}$$

where the second term drops out because $u(0) = u(\sigma) = 0$. Because u vanishes at its endpoints, this equation tells us that $u(s_0)$ must be zero for some $s_0 \in (0, \sigma)$. Since γ is minimizing, it cannot have any conjugate points, so u must vanish identically. We conclude that $\alpha = u'(0) = u'(\sigma) = \beta = 0$, so ϕ_τ is the identity map for all τ, completing the proof. □

Using similar methods to those introduced above, Chow [4] proved that the scalar curvature becomes positive in finite time which proves Theorem 2. We outline an alternative method in the next section.

8 Ricci Flow on Arbitrary 2-Spheres

In this section, we prove Theorem 2 in much the same way we proved Theorem 1. Namely, we show that an isoperimetric ratio increases under the flow, ruling out singular behavior. We follow the arguments in Hamilton [11].

Let Γ be a smooth embedded curve of length L on the S^2, dividing it into two regions of area A_1 and A_2. Now, take the curve of latitude $\bar{\Gamma}$ of length $\bar{L}$ on the round sphere of equal area, dividing it into regions of area A_1 and A_2. Let $\bar{C}$ denote the infimum over all curves Γ of the ratio $C = \frac{L^2}{\bar{L}^2}$. The main theorem of this section is the following:

Theorem 27. *The ratio $\bar{C}$ increases under the unnormalized Ricci flow on S^2.*

In his paper [11], Hamilton proves that the infimum is attained on a smooth, embedded curve Γ. For simplicity, will assume this result and focus on the improvement of the isoperimetric ratio (which is analagous to that used by Huisken to prove Grayson's theorem). The quantity C can easily be expressed in terms of L, A_1 and A_2. Parametrizing the round sphere of radius R using spherical coordinates, with the circle of latitude at angle ϕ, we have $A_1 = 2\pi R^2(1-\cos\phi) = 4\pi R^2 - A_2$. We then compute

$$\begin{aligned} A_1 A_2 &= (2\pi R^2)^2 \sin^2\phi \\ &= R^2 \bar{L}^2 \\ &= \frac{A\bar{L}^2}{4\pi}. \end{aligned}$$

We solve for $\bar{L}^2$ and apply the definition of C to get the formula

$$C = \frac{L^2 A}{4\pi A_1 A_2}.$$

Construct a family of curves $\gamma(r)$ a distance r from $\Gamma = \gamma(0)$ with respect to the metric at time t, for $r \in (-\epsilon, \epsilon)$. Keeping these curves fixed while the metric evolves, we see that L, A_1 and A_2 become functions of r and t. We will derive an evolution equation for $\log(C)$, and show that C is increasing for this loop by a standard maximum principle argument.

Proof. We compute the space derivatives first. Suppose that r is increasing into region 2. It is clear that

$$\frac{\partial A_1}{\partial r} = -\frac{\partial A_2}{\partial r} = L.$$

Next, note that $\frac{\partial L}{\partial r}$ is just the first variation of L in the normal direction, i.e.

$$\frac{\partial L}{\partial r} = \int \kappa ds$$

where κ is the geodesic curvature. Finally, applying the Gauss-Bonnet theorem, we have $\frac{\partial L}{\partial r} = 2\pi - \int_1 K da$ so that

$$\frac{\partial^2 L}{\partial r^2} = -\int K ds.$$

We now treat the time derivatives using the Ricci flow equation. Letting v be the velocity vector for a curve with parameter u, we have

$$\begin{aligned}\frac{\partial L}{\partial t} &= \int \frac{\partial}{\partial t}(g(v,v))^{\frac{1}{2}}du \\ &= -\int \frac{1}{2}(g(v,v))^{-\frac{1}{2}}\frac{\partial}{\partial t}g(v,v)du \\ &= -\int K ds.\end{aligned}$$

Notice that this is the same as $\frac{\partial^2 L}{\partial r^2}$, so L evolves according to a heat equation. Next, recall that volume element da evolves as $\frac{\partial}{\partial t}da = -Rda = -2Kda$ so that

$$\frac{\partial A_1}{\partial t} = -2\int_1 K da = -4\pi + 2\int \kappa ds.$$

Similarly,

$$\frac{\partial A_2}{\partial t} = -2\int_2 K da = -4\pi - 2\int \kappa ds,$$

with the sign of the geodesic curvature term switched because the curve is negatively oriented with respect to region 2. Finally, adding these gives

$$\frac{\partial A}{\partial t} = -8\pi.$$

We are now equipped to find an evolution equation for $\log C = 2\log L + \log A - \log A_1 - \log A_2 - \log 4\pi$. First,

$$\frac{\partial}{\partial r}\log C = \frac{2}{L}\frac{\partial L}{\partial r} - L\left(\frac{1}{A_1} - \frac{1}{A_2}\right).$$

We will use the information that this first variation vanishes at $r = 0$, i.e.

$$\frac{\partial L}{\partial r} = \frac{L^2}{2}\left(\frac{1}{A_1} - \frac{1}{A_2}\right).$$

Taking another derivative, applying $\frac{\partial A_1}{\partial r} = -\frac{\partial A_2}{\partial r} = L$ and using the vanishing of the first variation, we obtain

$$\begin{aligned}\frac{\partial^2 \log C}{\partial r^2} &= \frac{2}{L}\frac{\partial^2 L}{\partial r^2} - \frac{2}{L^2}\left(\frac{\partial L}{\partial r}\right)^2 - \frac{\partial L}{\partial r}\left(\frac{1}{A_1} - \frac{1}{A_2}\right) + L\left(\frac{1}{A_1^2}\frac{\partial A_1}{\partial r} - \frac{1}{A_2^2}\frac{\partial A_2}{\partial r}\right) \\ &= \frac{2}{L}\frac{\partial^2 L}{\partial r^2} - L^2\left(\frac{1}{A_1} - \frac{1}{A_2}\right)^2 + L^2\left(\frac{1}{A_1^2} + \frac{1}{A_2^2}\right).\end{aligned}$$

We now compute the time derivative:

$$\begin{aligned}\frac{\partial \log C}{\partial t} &= \frac{2}{L}\frac{\partial L}{\partial t} - \frac{1}{A}\frac{\partial A}{\partial t} - \frac{1}{A_1}\frac{\partial A_1}{\partial t} - \frac{1}{A_2}\frac{\partial A_2}{\partial t} \\ &= \frac{2}{L}\frac{\partial^2 L}{\partial r^2} - \frac{8\pi}{A} - (2\frac{\partial L}{\partial r} - 4\pi)/A_1 + (2\frac{\partial L}{\partial r} + 4\pi)/A_2.\end{aligned}$$

Using the vanishing of the first r-variation and the computation

$$4\pi\left(\frac{1}{A_1}+\frac{1}{A_2}\right)-\frac{8\pi}{A}=\frac{4\pi(A_1^2+A_2^2)}{A_1A_2(A_1+A_2)},$$

we get

$$\frac{\partial \log C}{\partial t}=\frac{\partial^2 \log C}{\partial r^2}+\frac{4\pi(A_1^2+A_2^2)}{A_1A_2(A_1+A_2)}(1-C).$$

We can get C arbitrarily close to 1 by considering tiny circles around any point, so in any case, $\bar{C}\leq 1$. Hence, the reaction term in our differential equation is positive, and the theorem follows from a straightforward maximum principle argument. □

For the details of the remainder of the proof of Theorem 2, we refer the reader to Hamilton [11]. Roughly, in direct analogy with the curve-shortening argument, one considers two possibilities for singularity formation. Let T be the maximal time interval on which a solution exists, and consider a limit of rescalings around points where the curvature is larger than previous times. In the first case, $\sup_M R(T-t)=\infty$. Then a subsequence of these rescalings must converge to a complete, non-compact solution with positive curvature which exists for all time. The only such solution is the Cigar soliton, i.e.

$$\left(\mathbb{R}^2, \frac{4}{1+|x|^2}\delta_{ij}\right).$$

One easily computes that the length of circles of radius R around 0 grow like $\frac{r}{\sqrt{1+r^2}}$, so the cigar is asymptotic to a cylinder. The area bounded by these circles grows like $\log(1+r^2)$, so by taking arbitrarily large circles, we can make the isoperimetric ratio as small as we like. The above theorem thus rules out this type of singularity.

In the second case, $\sup_M R(T-t)<\infty$. It follows from an entropy monotonicity formula proven by Chow [4] that the limit of dilations must be compact and homothetically shrinking, and the only such solution is the round sphere.

9 Ricci Flow on 3-Manifolds with Positive Ricci Curvature

We devote the final section of this paper to sketching a proof of Theorem 3, first given by Hamilton in [9]. The proof relies on three critical geometric estimates, all of which follow from maximum principle arguments. In dimension 2, the scalar maximum principle suffices because constant scalar curvature is constant curvature. This is false in 3 dimensions, where all the curvature information is contained in the Ricci tensor. To handle tensor heat equations, we introduce one last tool known as the tensor maximum principle, from which preservation of positive Ricci curvature can be inferred. We prove this estimate in detail.

The second estimate tells us that the eigenvalues of the Ricci tensor get pinched pointwise, and the final estimate gives us a bound on the gradient of the scalar curvature. These proofs rely on long derivations of evolution equations to which we apply the standard maximum principle; we will discuss the more intricate parts of these arguments, and refer the reader to Hamilton for the derivations.

9.1 The Tensor Maximum Principle

Theorem 28. *Let $g(t)$ be a smooth 1-parameter family of Riemannian metrics on a closed manifold M of dimension n. Let $\alpha(t)$ be a symmetric $(0,2)$ tensor satisfying the evolution equation*

$$\frac{\partial}{\partial t}\alpha(t) = \triangle_{g(t)}\alpha(t) + \beta(\alpha, g, t)$$

where β is a symmetric $(0,2)$ tensor which is locally Lipschitz in all arguments and satisfies a null eigenvector condition, namely if v is a null-eigenvector of α, then $\beta(v,v) \geq 0$. Then if $\alpha \geq 0$ at $t = 0$, it remains so for all later times at which a solution exists.

Proof. This proof is in the spirit of a scalar maximum principle proof. Let α solve the above equation on $M \times [0,T]$. Define $A_\epsilon(t) = \alpha(t) + \epsilon(\delta + t)g(t)$. We will choose δ independently of ϵ so that for all $t_0 \in [0, T-\delta]$, we have $A_\epsilon > 0$ on $[t_0, t_0+\delta]$. Since δ is chosen independently of t_0 and ϵ, A_ϵ is positive on $[0,T]$, and we can let ϵ go to 0 to get $\alpha \geq 0$ on $[0,T]$.

We choose δ to get a differential inequality for A_ϵ. First, suppressing notation for g and t dependence,

$$\begin{aligned}
\frac{\partial}{\partial t}A_\epsilon &= \triangle\alpha + \beta(\alpha) + \epsilon g + \epsilon(\delta+t)\frac{\partial}{\partial t}g \\
&= \triangle A_\epsilon + \beta(A_\epsilon) + \epsilon g + (\beta(\alpha) - \beta(A_\epsilon)) + \epsilon(\delta+t)\frac{\partial}{\partial t}g
\end{aligned}$$

We will choose δ small enough that the last three terms together are positive. Choose δ_0 so that for $t \in [0, \delta_0]$, we have $\epsilon(\delta_0 + t)\frac{\partial}{\partial t}g > -\frac{1}{2}\epsilon g$, i.e. so that $\frac{\partial}{\partial t}g > -\frac{1}{4\delta_0}g$, which is possible because g is positive. Next, choose a Lipschitz constant K for β and choose δ_1 so that for $t \in [0, \delta_1]$, we have $\beta(\alpha) - \beta(A_\epsilon) \geq -K\epsilon(\delta_1 + t)g > -\frac{1}{2}\epsilon g$, i.e. so that $\delta_1 < \frac{1}{4K}$. Taking $\delta = min\{\delta_1, \delta_2\}$, each of the last two terms exceeds $-\frac{1}{2}\epsilon g$, giving us the differential inequality

$$\frac{\partial}{\partial t}A_\epsilon > \triangle A_\epsilon + \beta(A_\epsilon).$$

We are now equipped to run the maximum principle argument. Note that $A_\epsilon(x,0) > 0$ by hypothesis. Suppose by way of contradiction that A_ϵ acquires a null eigenvector v for a first time $t_0 \in [0,\delta]$. We have $v \in T_{x_0}(M)$ for some $x_0 \in M$. Extend v to a neighborhood of x_0 by parallel translating along geodesics, and keep the resulting vector field constant in time. We then have $\frac{\partial}{\partial t} v = 0$, $\nabla v = 0$, and, choosing an orthonormal frame field e_i, we have $\triangle v = \sum_{i=1}^{n} (\nabla_{e_i} \nabla_{e_i} v - \nabla_{\nabla_{e_i} e_i} v) = \sum_{i=1}^{n} (\nabla_{e_i} 0 - \nabla_0 v) = 0$, using the fact that v and e_i are parallel translated along geodesics.

Using these computations, that $A_\epsilon(x_0,t) > 0$ for $t < t_0$, that $A_{\epsilon ij} v^i v^j$ is a local spatial minimum, and the null eigenvector condition on β, we have

$$\begin{aligned} 0 &\geq \frac{\partial}{\partial t}(A_{\epsilon ij} v^i v^j) \\ &= \frac{\partial}{\partial t}(A_{\epsilon ij}) v^i v^j \\ &> (\triangle A_{\epsilon ij} + \beta_{ij}) v^i v^j \\ &= \triangle (A_{\epsilon ij} v^i v^j) + \beta_{ij} v^i v^j \\ &\geq 0 \end{aligned}$$

giving the desired contradiction. □

Corollary 3. *If $Ric_g > 0$ at $t = 0$, then it remains so under the Ricci flow.*

Proof. Recall that the Ricci curvature evolves as

$$\begin{aligned} \frac{\partial}{\partial t} Ric(X,Y) &= \triangle Ric(X,Y) + 3R Ric(X,Y) \\ &\quad - 6 \sum_k Ric(X,e_k) Ric(Y,e_k) - g(X,Y)\left(R^2 - 2|Ric|^2\right). \end{aligned}$$

If V satisfies the null-eigenvector condition for Ric, i.e. $Ric(X,V) = 0$ for all X, then in particular we have $Ric(V,V) = 0$ and $\sum_k Ric(V,e_k) Ric(V,e_k) = 0$. Denote by λ, μ and ν the eigenvalues of Ric. By assumption, one of these, say ν, is 0. We then have

$$\begin{aligned} g(V,V)\left(2|Ric|^2 - R^2\right) &= g(V,V)\left(2(\lambda^2 + \mu^2) - (\lambda + \mu)^2\right) \\ &= g(V,V)(\lambda - \mu)^2 \\ &\geq 0. \end{aligned}$$

By the tensor maximum principle, positive Ricci curvature is preserved. □

We can in fact improve this estimate as follows:

Corollary 4. *If $Ric_g > \epsilon R g$ at $t = 0$, it remains so under the Ricci flow.*

We will apply the tensor maximum principle to the symmetric 2-tensor $\frac{1}{R} Ric - \epsilon g$, which requires deriving a reaction-diffusion equation. The following lemma will be useful for this computation.

Lemma 5. *Let T be an (r,s) tensor and let f be a positive smooth function on M. Then*

$$\triangle\left(\frac{1}{f}T\right)=\frac{f\triangle T-\triangle fT}{f^2}-\frac{2}{f}\sum_k e_k(f)D_{e_k}(\frac{1}{f}T).$$

Proof. We compute

$$\begin{aligned}\triangle\left(\frac{1}{f}T\right)&=\sum_k D^2_{e_k,e_k}\left(\frac{1}{f}T\right)\\&=\sum_k D_{e_k}D_{e_k}\left(\frac{1}{f}T\right)-\sum_k D_{D_{e_k}e_k}\left(\frac{1}{f}T\right)\\&=\sum_k D_{e_k}\left(\frac{1}{f}D_{e_k}T-\frac{e_k(f)}{f^2}T\right)-\sum_k\left(\frac{1}{f}D_{D_{e_k}e_k}T-\frac{D_{e_k}e_k(f)}{f^2}T\right)\\&=\frac{1}{f}\triangle T-\frac{1}{f^2}\triangle fT-2\sum_k\left(\frac{e_k(f)}{f^2}D_{e_k}T-\frac{e_k(f)^2}{f^3}T\right)\\&=\frac{f\triangle T-\triangle fT}{f^2}-\frac{2}{f}\sum_k e_k(f)D_{e_k}(\frac{1}{f}T)\end{aligned}$$

as desired. □

We now prove Corollary 4.

Proof. We first derive the evolution equation for $\frac{Ric(X,Y)}{R}-\epsilon g(X,Y)$ so that we can apply the tensor maximum principle. Let

$$Q(X,Y)=3RRic(X,Y)-6\sum_k Ric(X,e_k)Ric(Y,e_k)-g(X,Y)(R^2-2|Ric|^2)$$

so that $\frac{\partial}{\partial t}Ric(X,Y)-\triangle Ric(X,Y)+Q(X,Y)$. We then compute

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{Ric(X,Y)}{R}-\epsilon g\right)&=\frac{R\triangle Ric(X,Y)-\triangle RRic(X,Y)}{R^2}\\&\quad+\frac{RQ(X,Y)-2|Ric|^2Ric(X,Y)}{R^2}+2\epsilon Ric(X,Y),\end{aligned}$$

and using the previous lemma we have

$$\begin{aligned}\triangle\left(\frac{Ric(X,Y)}{R}-\epsilon g(X,Y)\right)&=\frac{R\triangle Ric(X,Y)-\triangle RRic(X,Y)}{R^2}\\&\quad-\frac{2}{R}\sum_k e_k(R)D_{e_k}\left(\frac{Ric(X,Y)}{R}-\epsilon g(X,Y)\right).\end{aligned}$$

Putting these computations together, we obtain the evolution equation

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{Ric(X,Y)}{R}-\epsilon g\right) &= \triangle\left(\frac{Ric(X,Y)}{R}-\epsilon g(X,Y)\right)\\ &\quad+\frac{2}{R}\sum_k e_k(R)D_{e_k}\left(\frac{Ric(X,Y)}{R}-\epsilon g(X,Y)\right)\\ &\quad+\frac{RQ(X,Y)-2|Ric|^2Ric(X,Y)}{R^2}+2\epsilon Ric(X,Y).\end{aligned}$$

To prove the claim, we will apply the tensor maximum principle with

$$\beta(X,Y)=\frac{RQ(X,Y)-2|Ric|^2Ric(X,Y)}{R^2}+2\epsilon Ric(X,Y).$$

The Ricci tensor is symmetric, so we may diagonalize it with an orthonormal set of eigenvectors $\{e_1,e_2,e_3\}$ associated to eigenvalues λ,μ and ν. Assume without loss of generality that e_1 is the null-eigenvector of $Ric-\epsilon Rg$, so that $\lambda=\epsilon R$. Then use this to eliminate ϵ in the expression for $\beta(e_1,e_1)$, obtaining

$$\begin{aligned}R^2\beta(e_1,e_1) &= R(3R\lambda-6\lambda^2-R^2+2|Ric|^2)+2\epsilon R^2\lambda-2\lambda|Ric|^2\\ &= R(3R\lambda-4\lambda^2-R^2+2|Ric|^2)-2\lambda|Ric|^2.\end{aligned}$$

Substituting $R=\lambda+\mu+\nu$ and $|Ric|^2=\lambda^2+\mu^2+\nu^2$ and simplifying, this expression becomes

$$R^2\beta(e_1,e_1)=\lambda^2(\mu+\nu-2\lambda)+(\mu+\nu)(\mu-\nu)^2.$$

The second term is nonnegative because positive Ricci curvature is preserved, so we focus our attention on the first. Observe that contracting the inequality $Ric\geq\epsilon Rg$ give $\epsilon\leq\frac{1}{3}$. Since $\epsilon=\frac{\lambda}{R}$, it follows that $\frac{R}{\lambda}=1+\frac{\mu+\nu}{\lambda}\geq 3$, which we may rearrange to obtain $\mu+\nu\geq 2\lambda$. This proves that the first term is positive, hence all the hypotheses of the tensor maximum principle are satisfied. □

9.2 Pointwise Pinching of Eigenvalues

In this section we prove a pointwise pinching estimate on the eigenvalues of the Ricci tensor. We first introduce a geometric quantity measuring the deviation of the eigenvalues from one another: note that

$$\begin{aligned}|Ric|^2-\frac{1}{3}R^2 &= \lambda^2+\mu^2+\nu^2-\frac{1}{3}(\lambda+\mu+\nu)^2\\ &= \frac{1}{3}\left(2\lambda^2+2\mu^2+2\nu^2-2\lambda\mu-2\lambda\nu-2\mu\nu\right)\\ &= \frac{1}{3}\left((\lambda-\mu)^2+(\lambda-\nu)^2+(\mu-\nu)^2\right).\end{aligned}$$

This implies in particular that $|Ric|^2\geq\frac{1}{3}R^2$, so that

$$\begin{aligned}\frac{\partial}{\partial t}R &= \triangle R+2|Ric|^2\\ &\geq \triangle R+\frac{2}{3}R^2.\end{aligned}$$

By assumption, $R \geq \rho$ for some $\rho > 0$ at time $t = 0$. The scalar maximum principle then implies that R is bounded below by the solution to the ODE $\frac{dh}{dt} = \frac{2}{3}h^2$ with $h(0) = \rho$. Solving, one obtains

$$R \geq \frac{3\rho}{3 - 2\rho t}$$

which tells us that the maximal time T for which a solution exists is bounded above by $\frac{3}{2\rho}$.

Remark 6. *It can be shown via interpolation inequalities (see [9]) that if $|Riem| < C < \infty$ on a time interval $[0, T)$, then the metric converges in the C^∞ sense to a limit metric $g(T)$. Using the short-time existence result, we can extend the solution past T, proving long-time existence. This implies in particular that $|Ric|$ blows up as t goes to T, and since $|Ric|^2 \leq R^2$ by the positivity of Ricci curvature, we must have R_{max} going to ∞.*

The main theorem of this section is the following.

Theorem 29. *There exist $\delta > 0$ and $C < \infty$ depending only on the initial metric such that for all times at which a solution exists,*

$$\frac{|Ric|^2 - \frac{1}{3}R^2}{R^2} \leq CR^{-\delta}.$$

This theorem, along with the previous remark, tells us that the quantity $\frac{|Ric|^2}{R^2} - \frac{1}{3}$ goes to 0 at the points where R blows up.

Proof. Let $\gamma = 2 - \delta$ and let

$$f = \frac{|Ric|^2 - \frac{1}{3}R^2}{R^\gamma} = \frac{|Ric|^2}{R^\gamma} - \frac{1}{3}R^{2-\delta}.$$

The idea is to derive an evolution equation for f of the form

$$\frac{\partial}{\partial t}f \leq \triangle f + \langle X, \nabla f \rangle$$

for some vector field X which may depend on time. The scalar maximum principle would then imply that if $f \leq C$ at $t = 0$, then it remains so, proving the claim. The derivation of the evolution equation is quite tedious, but it does not rely on any fancy identities or tricks. We refer the reader to [9] for a proof of the following:

Lemma 6. *The function f satisfies the evolution equation*

$$\begin{aligned}\frac{\partial}{\partial t}f = &\triangle f + \frac{2(\gamma - 1)}{R}\langle \nabla R, \nabla f \rangle \\ &- \frac{2}{R^{\gamma+2}}|R \cdot DRic - DR \cdot Ric|^2 \\ &- \frac{(2-\gamma)(\gamma-1)}{R^{\gamma+2}}\left(|Ric|^2 - \frac{1}{3}R^2\right)|DR|^2 \\ &+ \frac{2}{R^{\gamma+1}}\left(\delta|Ric|^2(|Ric|^2 - \frac{1}{3}R^2) - 2P\right)\end{aligned}$$

where

$$P = \lambda^2(\lambda - \mu)(\lambda - \nu) + \mu^2(\mu - \lambda)(\mu - \nu) + \nu^2(\nu - \lambda)(\nu - \mu).$$

Thus, the proof boils down producing a δ depending only on the initial metric such that $2P \geq \delta|Ric|^2 \left(|Ric|^2 - \frac{1}{3}R^2\right)$. This is the content of the next lemma:

Lemma 7. *For $R > 0$ and $Ric \geq \epsilon R g$, we have $P \geq \epsilon^2 |Ric|^2 \left(|Ric|^2 - \frac{1}{3}R^2\right)$.*

Let λ, μ and ν be the eigenvalues of the Ricci tensor. Without loss of generality, $\lambda \geq \mu \geq \nu$. The hypotheses give

$$\lambda \geq \mu \geq \nu \geq \epsilon R > 0.$$

We can rewrite the first two terms of P as follows:

$$\begin{aligned}
\lambda^2(\lambda-\mu)(\lambda-\nu) + \mu^2(\mu-\lambda)(\mu-\nu) &= \lambda^2(\lambda-\mu)((\lambda-\mu)+(\mu-\nu)) + \mu^2(\mu-\lambda)(\mu-\nu) \\
&= \lambda^2(\lambda-\mu)^2 + (\lambda-\mu)(\lambda^2(\mu-\nu) - \mu^2(\mu-\nu)) \\
&= (\lambda-\mu)^2\left(\lambda^2 + (\mu-\nu)(\lambda+\mu)\right) \\
&\geq \lambda^2(\lambda-\mu)^2.
\end{aligned}$$

Thus, we have the sequence of inequalities

$$\begin{aligned}
P &\geq \lambda^2(\lambda-\mu)^2 + \nu^2(\mu-\nu)^2 \\
&\geq \epsilon^2 R^2\left((\lambda-\mu)^2 + (\mu-\nu)^2\right) \\
&= \frac{\epsilon^2}{3}R^2\left((\lambda-\mu)^2 + (\mu-\nu)^2 + 2(\lambda-\mu)^2 + 2(\mu-\nu)^2\right) \\
&\geq \frac{\epsilon^2}{3}R^2\left((\lambda-\mu)^2 + (\mu-\nu)^2 + (\lambda-\nu)^2\right) \\
&\geq \frac{\epsilon^2}{3}|Ric|^2\left((\lambda-\mu)^2 + (\mu-\nu)^2 + (\lambda-\nu)^2\right) \\
&= \epsilon^2|Ric|^2\left(|Ric|^2 - \frac{1}{3}R^2\right),
\end{aligned}$$

proving the lemma. By taking $\delta < 2\epsilon^2$, we have the desired differential inequality for f, proving the claim. ☐

9.3 Bounding the Gradient of Scalar Curvature

Theorem 30. *For every $\eta > 0$ there exists $C < \infty$ depending only on η and the initial metric such that when a solution exists,*

$$|\nabla R|^2 \leq \eta R^3 + C(\eta).$$

The strategy is to find a positive function h and derive a differential inequality for the function $f = \frac{|\nabla R|^2}{R} - \eta R^2$ of the form

$$\frac{\partial}{\partial t}(f+h) \leq \triangle(f+h) + C_0(\eta)$$

for some constant C_0 depending only on η and the initial metric. We can then apply the maximum principle to conclude that

$$f \leq f + h \leq C_0(\eta)t,$$

which then implies

$$|\nabla R|^2 \leq \eta R^3 + C_0(\eta) R t.$$

Recall that the maximal time T for which a solution exists is finite and depends only on the initial minimum of R, so that

$$|\nabla R|^2 \leq \eta R^3 + C_1(\eta) R$$

for some $C_1(\eta)$ depending only on η and $g(0)$. Finally, if we choose $C(\eta)$ sufficiently large depending only on $C_1(\eta)$ and η we have

$$|\nabla R|^2 \leq \eta R^3 + C_1(\eta) R \leq 2\eta R^3 + C(\eta).$$

Since η is arbitrary, this would complete the proof.

We proceed with the derivation of the evolution equation, which is a bit more delicate than that of the last section. One easily computes (see [9])

$$\begin{aligned}\frac{\partial}{\partial t} f = \triangle f - \frac{2}{R^3} |R \mathrm{Hess} R - DR \cdot DR|^2 - 4\eta R |Ric|^2 \\ - 2\left(\frac{|Ric|^2}{R^2} - \eta\right) |\nabla R|^2 + \frac{4}{R} \left\langle \nabla R, \nabla |Ric|^2 \right\rangle .\end{aligned}$$

Recall from the previous section that $\frac{|Ric|^2}{R^2} \geq \frac{1}{3}$, so the third term is bounded above by $-\frac{4}{3}\eta R^3$, and by choosing $\eta < \frac{1}{3}$ the fourth term is nonpositive. The only remaining bad term is the last one. Working in normal coordinates, we obtain a bound for this term via Cauchy-Schwarz:

$$\begin{aligned}\left\langle \nabla R, \nabla |Ric|^2 \right\rangle &= g^{ij} \partial_i R \partial_j \left(g^{pq} g^{rs} R_{pr} R_{qs} \right) \\ &= 2 g^{ij} g^{pq} g^{rs} \partial_i R R_{pr} \partial_j R_{qs} \\ &= 2 \left\langle \partial_i R R_{jk}, \partial_i R_{jk} \right\rangle \\ &\leq 2 |\partial_i R| |R_{jk}| |\partial_i R_{jk}|.\end{aligned}$$

Since all the eigenvalues of the Ricci tensor are positive, we have $|Ric|^2 \leq R^2$. Furthermore, using normal coordinates again we have

$$\begin{aligned}|\partial_i R|^2 &= \sum_{i=1}^{3} \left(\sum_{j=1}^{3} \partial_i R_{jj} \right)^2 \\ &\leq 3 \sum_{i,j} (\partial_i R_{jj})^2 \\ &\leq 3 |\partial_i R_{jk}|^2\end{aligned}$$

so that $|\partial_i R| \leq 2 |\partial_i R_{jk}|$. This last term is thus bounded by $16 |\partial_i R_{jk}|^2$, giving us the differential inequality

$$\frac{\partial}{\partial t} f \leq \triangle f + 16 |\partial_i R_{jk}|^2 - \frac{4}{3} \eta R^3.$$

In order to eliminate the bad term on the right hand side of this inequality, we consider the function $h = |Ric|^2 - \frac{1}{3}R^2$. It is not too hard to show that h satisfies the differential inequality

$$\frac{\partial}{\partial t}h \leq \triangle h - 2\left(|\partial_i R_{jk}|^2 - \frac{1}{3}|\partial_i R|^2\right) + 4Rh.$$

It is interesting that the constant 3 in the estimate $|\partial_i R|^2 \leq 3|\partial_i R_{jk}|^2$ can be improved to $\frac{20}{7}$; then, by adding the appropriate multiple K_0 of h to f we can cancel the $|\partial_i R_{jk}|^2$ terms to obtain

$$\frac{\partial}{\partial t}(f + K_0 h) \leq \triangle(f + K_0 h) + K_1 Rh - \frac{4}{3}\eta R^3$$

for some constant K_1 depending only on K_0. Recall from the previous section that there are constants $K_2, \delta > 0$ depending only on $g(0)$ such that $h \leq K_3 R^{2-\delta}$, so that the last two terms are bounded above by

$$K_1 K_2 R^{3-\delta} - \frac{4}{3}\eta R^3.$$

This last expression is certainly less than some constant $C(\eta)$ depending only on K_2, δ and η, completing the proof.

To get the improved estimate $|\partial_i R|^2 \leq \frac{20}{7}|\partial_i R_{jk}|^2$, we call on the contracted second Bianchi identity $g^{ij}\partial_i R_{jk} = \frac{1}{2}\partial_k R$. We would like to write $\partial_i R_{jk}$ in its irreducible components. To that end, write

$$\partial_i R_{jk} = E_{ijk} + F_{ijk},$$

where $E_{ijk} = Ag_{ij}\partial_k R + Bg_{jk}\partial_i R + Cg_{ik}\partial_j R$ for some constants A, B and C to be determined by the condition that F is traceless. Taking the three traces of F and using the contracted second Bianchi identity, we obtain these conditions: $3A + B + C = A + B + 3C = \frac{1}{2}$ and $A + 3B + C = 1$. Solving for the coefficients, we get

$$E_{ijk} = \frac{1}{20}(g_{ij}\partial_k R + g_{ik}\partial_j R) + \frac{3}{10}g_{jk}\partial_i R.$$

One can easily compute $|E_{ijk}|^2 = \frac{7}{20}|\partial_i R|^2$, as desired.

9.4 Global Pinching of Eigenvalues

With a bound on the gradient of the scalar curvature, we can extend the pointwise pinching result to a global result by applying the Bonnet-Myers theorem, which relates the curvature of a manifold to its topology:

Theorem 31. *Let (M, g) be a complete Riemannian manifold such that $Ric_g \geq (n-1)\lambda$, where λ is a positive constant. Then $diam(M) \leq \frac{\pi}{\sqrt{\lambda}}$.*

This classical theorem is a consequence of the second variation formula for energy, and for brevity we omit the proof. Let T be the maximal time interval on which a solution exists.

Theorem 32. *The quantity $\frac{|Ric|^2}{R^2} - \frac{1}{3}$ goes to 0 as $t \to T$.*

This is a consequence of the following lemma:

Lemma 8. $\frac{R_{max}}{R_{min}} \to 1$ *as* $t \to T$.

Since we know that R_{max} blows up as t goes to T, this lemma tells us that R_{min} must blow up as well. The pointwise pinching estimate $\frac{|Ric|^2}{R^2} - \frac{1}{3} \leq CR^{-\delta} \leq CR_{min}^{-\delta}$ gives us the desired result. We proceed with a proof of the lemma:

Proof. The bound on the gradient of the scalar curvature tells us that

$$|\nabla R| \leq \frac{\eta^2}{2} R^{\frac{3}{2}} + C(\eta)$$

for an η to be chosen sufficiently small later. Since R_{max} blows up as $t \to T$, there exists $\tau(\eta)$ such that $C(\eta) \leq \frac{\eta^2}{2} R_{max}^{\frac{3}{2}}$ for all $t \in [\tau(\eta), T)$. It follows that

$$|\nabla R| \leq \eta^2 R_{max}^{\frac{3}{2}}$$

for all $t \in [\tau(\eta), T)$. Take any t in this time interval, and let $R(p,t) = R_{max}(t)$. Let $\rho = \frac{1}{\eta\sqrt{R_{max}}}$. Then by our gradient bound, for any $x \in B_\rho(p)$ (the ball of radius ρ with respect to the metric at time t) we have

$$\begin{aligned} R(x,t) &\geq R_{max}(t) - \left(\sup_M |\nabla R|\right) d(x,p) \\ &\geq R_{max}(t) - \eta^2 R_{max}(t)^{\frac{3}{2}} \frac{1}{\eta\sqrt{R_{max}(t)}} \\ &= R_{max}(t)(1-\eta). \end{aligned}$$

We claim that for η sufficiently small, this ball is all of M, and since we can choose η arbitrarily small, this would prove the lemma. By our tensor maximum principle estimate, we know that on this ball

$$Ric \geq \epsilon R g \geq \epsilon R_{max}(1-\eta) = 2\left(\frac{\epsilon R_{max}(1-\eta)}{2}\right) g.$$

It follows from the Bonnet-Myers theorem that geodesics of length exceeding $l = \frac{\pi\sqrt{2}}{\sqrt{\epsilon R_{max}(1-\eta)}}$ inside the ball $B_\rho(p)$ are unstable. For η small enough, notice that $l < \frac{\rho}{2}$. Thus, if there were any points q such that $d(p,q) \geq \rho$, we could join them by a minimizing geodesic $\gamma(s)$ parametrized by arclength. But in this case, $\gamma|_{[0,\frac{3\rho}{4}]}$ is unstable, contradicting the minimality of γ and proving the claim. □

We end this section with a few remarks about the remainder of the proof of Theorem 3. By dilating and reparametrizing, we get a solution $\tilde{g}$ to the normalized flow, and we can extend the above results to $\tilde{g}$ (see [9]). In particular, one can show that $\tilde{T} = \infty$, and using interpolation inequalities for tensors, one can show that $\tilde{g}$ approaches a limit metric $g(\infty)$ in the C^∞ sense. Finally, via a maximum principle argument, one shows that $\widetilde{Ric}(t)$ exponentially approaches $\frac{1}{3}\tilde{r}\tilde{g}$. Since the Ricci curvature determines the sectional curvatures in dimension 3, we see that $\tilde{g}(\infty)$ is a metric of constant curvature.

10 Acknowledgements

I am very grateful to Simon Brendle for his enthusiastic guidance and astute reading suggestions, both of which have inspired me to pursue geometry in my graduate studies; I could not have asked for a better thesis advisor. I would also like to thank Leon Simon for many instructive conversations about three-manifolds with positive Ricci curvature. Finally, I would like to thank my academic advisor András Vasy for suggesting that I work with Professor Brendle, and for all of his support during my undergraduate years.

References

[1] U. Abresch and J. Langer, *The normalized curve shortening flow and homothetic solutions*, J. Differential Geometry **23** (1986), 175–196.

[2] S. Altschuler, *Singularities for the curve-shortening flow for space curves*, J. Differential Geometry **34** (1991), 491–514.

[3] S. Brendle, *Ricci flow and the sphere theorem*, Graduate Studies in Mathematics, vol. 111, American Mathematical Society, 2010.

[4] B. Chow, *The ricci flow on the two-sphere*, J. Differential Geometry **33** (1991), 325–334.

[5] M. Gage, *Curve shortening makes convex curves circular*, Invent. Math.

[6] ———, *An isoperimetric inequality with applications to curve-shortening*, Duke Math. J. **50** (1983), 1225–1229.

[7] M. Gage and R. S. Hamilton, *The heat equation shrinking convex plane curves*, J. Differential Geometry **23** (1986), 69–96.

[8] M. A. Grayson, *The heat equation shrinks embedded plane curves to round points*, J. Differential Geometry **26** (1987), 285–314.

[9] R. S. Hamilton, *Three-manifolds with positive ricci curvature*, J. Differential Geometry **17** (1982), 255–306.

[10] ———, *The ricci flow on surfaces*, Math. and General Relativity, Contemporary Math., vol. 71, 1988, pp. 237–262.

[11] ———, *An isoperimetric estimate for the ricci flow on the two-sphere*, Modern Methods in Complex Analysis, Princeton Univ. Press, Princeton, NJ, 1995.

[12] G. Huisken, *Asymptotic behavior for singularities of the mean curvature flow*, J. Differential Geometry **31** (1990), 285–299.

[13] ———, *A distance comparison principle for evolving curves*, Asian J. Math **2** (1998), 127–133.

[14] J. Milnor, *Morse theory*, Princeton University Press, Princeton, NJ, 1963.

[15] B. O'Neill, *Semi-riemannian geometry with applications to relativity*, Academic Press, San Diego, CA, 1983.

[16] R. Osserman, *Bonneson-style isoperimetric inequalities*, Amer. Math. Monthly **86** (1979), 1–29.

[17] P. Lu X. Chen and G. Tian, *A note on uniformization of riemann surfaces by ricci flow*, Proc. Amer. Math. Soc. **134** (2006), 3391–3393.