Lectures on Differential Geometry

Bennett Chow

Yutze Chow

University of California San Diego

University of Wisconsin, Milwaukee (emeritus)

Chapter 14

Uniformization of Surfaces via Heat Flow

Chapter from a book in progress.

Recall that the differential geometric version of the uniformization theorem (Theorem 8.14) says that for any Riemannian metric g_0 on a closed surface M^2, there exists a positive function v such that the new metric vg_0 on M^2 has constant curvature. That is, by changing infinitesimal lengths but not infinitesimal angles associated to the metric, one can arrange so that the new metric is nice in the sense that it has constant curvature. In this chapter, we consider Hamilton's heat flow approach to the proof of this result. Namely, we start with a Riemannian metric on a closed surface and we deform the metric in its conformal class by a heat-type equation, called the *Ricci flow*, to a constant curvature metric. Figure 14.0.1 shows snapshots of a solution to the Ricci flow on a 2-sphere.

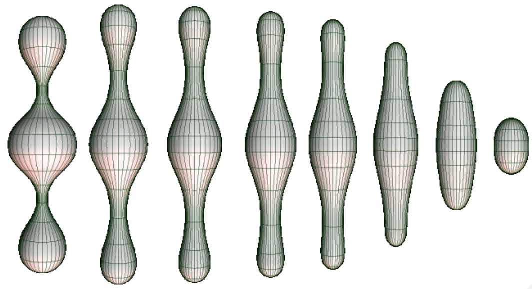

Figure 14.0.1. A rotationally symmetric solution to the Ricci flow on the 2-sphere. Metrics at later times are to the right. As the area decreases to zero, the metrics become rounder. Credit: Wikimedia Commons, Public Domain. Author: CBM

14.1. Families of conformally equivalent metrics on surfaces

Let M^2 be a closed oriented 2-dimensional manifold. Let g_0 be a Riemannian metric on M^2. Let $u(t) : M^2 \to \mathbb{R}$, $t \in I$, , where I is an interval, be a 1-parameter family of functions. Then

$$g(x,t) := \mathrm{e}^{2u(x,t)} g_0(x), \tag{14.1}$$

$t \in I$, is a 1-parameter family of metrics. By definition, each metric $g(t) = \mathrm{e}^{2u(t)} g_0$ is *conformal to* (or *conformally equivalent to*) g_0; that is, the infinitesimal angles defined by $g(t)$ are the same as those defined by g_0 (see §8.4). The function $\mathrm{e}^{2u(t)}$ is called the **conformal factor**. For simplicity, we will also call u the conformal factor.

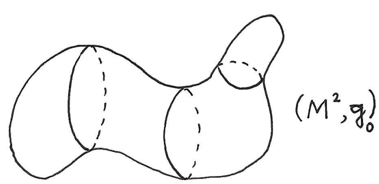

Figure 14.1.1. A Riemannian surface (M^2, g_0), where M^2 is diffeomorphic to S^2.

Not all metrics on S^2 can be isometrically embedded in $\mathbb{R}^3$, so the drawing of the Riemannian surface (M^2, g_0) in Figure 14.1.1 should not be viewed too literally. On the other hand, we can visualize the Riemannian metric g_0 on M^2 as follows. Let $\phi : S^2 \to M^2$ be a diffeomorphism, where S^2 is the unit sphere in $\mathbb{R}^3$. Consider the pulled back metric

$$h_0 := \phi^* g_0, \tag{14.2}$$

which is by definition isometric to g_0. So visualizing the metric h_0 on S^2 is the same as visualizing the metric g_0 on M^2.

The metric h_0 defines an inner product on each tangent space $T_{\mathbf{x}}S^2$, $\mathbf{x} \in S^2$. We visualize h_0 by drawing the set of unit vectors in $T_{\mathbf{x}}S^2$. Since $(h_0)_{\mathbf{x}}$ is an inner product on $T_{\mathbf{x}}S^2$, this set is an ellipse in a plane in $\mathbb{R}^3$; see Figure 14.1.2. A conformal metric $g = \mathrm{e}^{2u} g_0$ can now be visualized via its pullback metric

$$h := \phi^* g = \mathrm{e}^{2u\circ\phi} h_0. \tag{14.3}$$

Since the metric h is pointwise conformal to h_0, the set of unit vectors in $T_{\mathbf{x}}S^2$ with respect to $h_{\mathbf{x}}$ is an ellipse which is a *constant multiple* (scaling) of the ellipse for h_0.

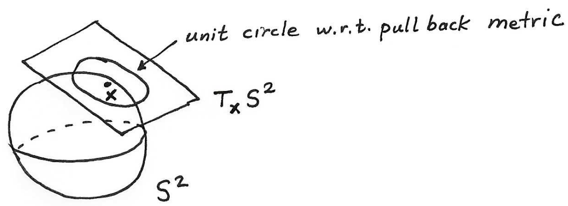

Figure 14.1.2. Visualizing a metric on a topological 2-sphere by pull-back: The unit 2-sphere, but with the pull-back metric $h_0 = \phi^* g_0$ defined by (14.2). The unit circle in $T_{\mathbf{x}}S^2$ with respect to h_0 is an ellipse.

We now consider the variation of a 1-parameter family of conformal metrics. Let

$$v(x,t) := 2\frac{\partial u}{\partial t}(x,t). \tag{14.4}$$

Differentiating (14.1) yields the equivalent formula

$$\left(\frac{\partial}{\partial t} g\right)(t) = 2\frac{\partial u}{\partial t}(t)\mathrm{e}^{2u(t)} g_0 = v(t)\, g(t). \tag{14.5}$$

Namely, it is easy to see that the conformal deformation of the metric $g(t)$ equation

$$\frac{\partial}{\partial t} g(t) = v(t)\, g(t) \tag{14.6}$$

holds if and only if the conformal factors $u(t)$ satisfy

$$2\frac{\partial u}{\partial t}(t) = v(t). \tag{14.7}$$

Even though $g(t)$ is just a 1-parameter family of conformally equivalent metrics, we say that $g(t)$ satisfying (14.6) is a **conformal deformation with velocity** $v(t)$.

Let $R(t) = 2K(t)$ denote the scalar curvature of $g(t)$, which is equal to twice the Gauss curvature of $g(t)$. By (8.46), we have that if $g(t) = \mathrm{e}^{2u(t)}g_0$, then

$$R(t) = \mathrm{e}^{-2u(t)}\big(R_0 - 2\Delta_0 u(t)\big) = \mathrm{e}^{-2u(t)}R_0 - 2\Delta_{g(t)}u(t), \tag{14.8}$$

where R_0 and Δ_0 denote the scalar curvature and Laplacian of g_0, respectively; the second equality follows from Lemma 11.2.

14.2. Variation of the curvature under a conformal variation of the metric

By differentiating (14.8), we calculate that if the metrics $g(t)$ satisfy (14.6), i.e., $\partial_t g = vg$, then

$$\begin{aligned}\frac{\partial R}{\partial t}(t) &= -2\frac{\partial u}{\partial t}(t)\mathrm{e}^{-2u(t)}\big(R_0 - 2\Delta_0 u(t)\big) - 2\mathrm{e}^{-2u(t)}\Delta_0\left(\frac{\partial u}{\partial t}(t)\right)\\ &= -v(t)R(t) - \mathrm{e}^{-2u(t)}\Delta_0 v(t)\\ &= -v(t)R(t) - \Delta_{g(t)}v(t).\end{aligned}$$

Summarizing, we have proved the following.

Lemma 14.1. *If a 1-parameter family of Riemannian metrics $g(t)$, $t \in I$, on a 2-dimensional smooth manifold M^2 satisfies $\frac{\partial}{\partial t}g(t) = v(t)g(t)$, where $v(t) : M^2 \to \mathbb{R}$ for each $t \in I$, then their scalar curvatures satisfy the equation*

$$\frac{\partial R}{\partial t}(t) = -\Delta_{g(t)}v(t) - v(t)R(t). \tag{14.9}$$

If we take $v(t) = -R(t)$, then we obtain the Ricci flow on surfaces.

Corollary 14.2. *If a 1-parameter family of Riemannian metrics $g(t)$ on a 2-dimensional manifold satisfies the equation $\frac{\partial}{\partial t}g(t) = -R(t)g(t)$, called the* ***Ricci flow on surfaces****, then their scalar curvatures satisfy the equation*

$$\frac{\partial R}{\partial t}(t) = \Delta_{g(t)}R(t) + R(t)^2. \tag{14.10}$$

Equation (14.10) is a nonlinear heat-type equation and also called a reaction-diffusion equation. On the right-hand side, the *diffusion term* is the Laplacian ΔR and the *reaction term* is the function of the solution R^2. Without the reaction term, from (14.10) we obtain the heat equation, which smooths out the solution. Without the diffusion term, from (14.10) we obtain an ODE, which in this case is $\frac{d\mathrm{R}}{dt} = \mathrm{R}^2$, where R is a function of t.

Example 14.3 (Shrinking 2-sphere)**.** Suppose that g_0 is the 2-sphere of radius ρ_0. Then its scalar curvature is $R_0 = \frac{2}{\rho_0^2}$. As we will see in Example 14.4, there exists a (unique) solution $g(t)$ to the Ricci flow satisfying the initial condition $g(0) = g_0$ which form round shrinking 2-spheres. Hence, for each t, $R(t)$ is a constant. Thus, $R(t)$ satisfies the ODE $\frac{dR}{dt}(t) = R(t)^2$. Solving this ODE, we obtain

$$R(t) = \frac{1}{R_0^{-1} - t} = \frac{1}{\frac{\rho_0^2}{2} - t}. \tag{14.11}$$

Observe that this solution exists on the maximal time interval $\left[0, \frac{\rho_0^2}{2}\right)$; in fact, it can be defined on the **ancient time interval** $\left(-\infty, \frac{\rho_0^2}{2}\right)$. As $t \to \frac{\rho_0^2}{2}$, we have that $R(t) \to \infty$ and the radius of the 2-sphere at time t tends to zero. See Figure 14.2.1

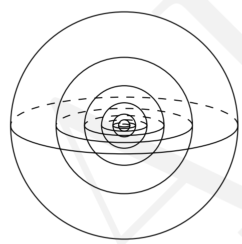

Figure 14.2.1. A constant curvature 2-sphere shrinking to a point under the Ricci flow.

14.3. The normalized Ricci flow equation on surfaces

As we have seen from the shrinking spheres in Example 14.3, the areas of the metrics is not preserved in general. The *normalized Ricci flow* rectifies this defect by scaling the metrics so that the area is constant in time.

Let $g(t)$ be a family of metrics on a closed oriented surface M^2. Let $r(t)$ denote the average scalar curvature of $g(t)$, that is,

$$r(t) := \frac{\int_{M^2} R(t)\, d\mu(t)}{\int_{M^2} d\mu(t)}, \tag{14.12}$$

where $d\mu(t)$ denotes the area form of $g(t)$. This is equal to the average of the function $R(t)$ on M^2 with respect to the area form $d\mu(t)$. Observe that

$$\int_{M^2} \big(R(t) - r(t)\big) d\mu(t) = 0. \tag{14.13}$$

Hamilton [**Ham88**] considered the following equation for $g(t)$:

$$\frac{\partial}{\partial t} g(t) = \big(r(t) - R(t)\big)\, g(t). \tag{14.14}$$

This equation, called the **normalized Ricci flow on surfaces**, is equivalent to the equation

$$2\frac{\partial u}{\partial t}(t) = r(t) - R(t) \tag{14.15}$$

for the conformal factor $u(t)$ defined by $g(t) = \mathrm{e}^{2u(t)} g_0$.

As for any equation, the main questions are: Do solutions exist and how do they behave?

Firstly, geometrically, we will see below that the metrics $\bar{g}(\bar{t})$ of a normalized Ricci flow are just metric rescalings and time reparametrizations[1] of the metrics $g(t)$ of a Ricci flow:

$$\frac{\partial}{\partial t} g(t) = -R(t)\, g(t). \tag{14.16}$$

Observe that the metric is (conformally) shrinking at points where the curvature is positive and the metric is expanding at points where the curvature is negative. See Figures 14.3.1 and 14.3.2.

In the next subsection (see (14.20) below), we will prove that the area of $\widetilde{g}(\widetilde{t})$ is constant under the normalized Ricci flow. On the other hand (see (14.24) below), under the Ricci flow the area of $g(t)$ is given by

$$\mathrm{Area}(g(t)) = \mathrm{Area}(g_0) - 4\pi\chi(M^2)t. \tag{14.17}$$

So:

(1) If $\chi(M^2) > 0$, then the area of $g(t)$ decreases at a constant rate.

(2) If $\chi(M^2) = 0$, then the area of $g(t)$ is constant.

(3) If $\chi(M^2) < 0$, then the area of $g(t)$ increases at a constant rate.

In particular, if M^2 is diffeomorphic to the 2-sphere S^2, then under the Ricci flow the area of $g(t)$ decreases at a constant rate until it limits to zero in a finite amount of time (provided one can show the solution exists as long as the area is positive).

Example 14.4 (Constant curvature solutions)**.** Suppose that (M^2, g_0) is a closed Riemannian surface with constant curvature $r_0 := R(g_0)$. Then:

(1) $g(t) \equiv g_0$, $t \in [0, \infty)$, is the unique maximal solution to the *normalized* Ricci flow with $g(0) = g_0$.

(2) $g(t) := (1 - r_0 t) g_0$, for all $t \geq 0$ satisfying $1 - r_0 t > 0$, is the unique maximal solution to the (*unnormalized*) Ricci flow with $g(0) = g_0$. Indeed, we check that

$$\partial_t g(t) = -r_0 g_0 = -R(0) g(0) = -R(t) g(t).$$

[1] This is why we denote the time parameter by $\bar{t}$ instead of t.

If $r_0 \le 0$, then this solution exists for all $t \in [0, \infty)$. On the other hand, if $r_0 > 0$, then this solution exists on the maximal time interval $[0, r_0^{-1})$; this agrees with Example 14.3.

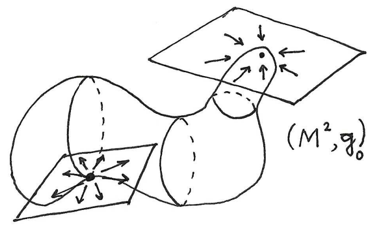

Figure 14.3.1. A Riemannian surface (M^2, g_0), where M^2 is diffeomorphic to the 2-sphere S^2. The arrows indicate that at points with positive curvature, the metric shrinks conformally under the Ricci flow.

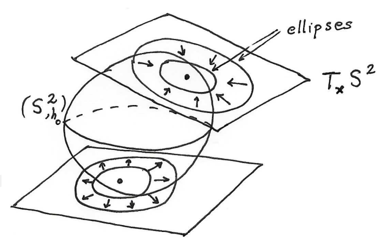

Figure 14.3.2. The unit 2-sphere, but with the pulled back metric $h_0 = \phi^* g_0$ defined by (14.2). At points with positive curvature, the ellipses shrink forward in time indicating that the metric is conformally shrinking at these points. At points with negative curvature, the ellipses expand.

14.4. Evolution of the area under the normalized and unnormalized Ricci flows

Now, suppose that we are given a solution $g(t)$ to the normalized Ricci flow on a closed oriented surface M^2. Suppose in addition that the time interval of existence is $I = [0, T)$, where $T \in (0, \infty]$, and that $g(0) = g_0$. (The last equality is equivalent to the conformal factor satisfying $u(0) = 0$.)

Let $\{\omega_0^1, \omega_0^2\}$ be a positively-oriented orthonormal coframe field for g_0 defined on an open subset $\mathcal{U}$ of M^2. Then $\{\omega^1(t), \omega^2(t)\} := \{e^{u(t)}\omega_0^1, e^{u(t)}\omega_0^2\}$ is a positively-oriented orthonormal coframe field for $g(t) = e^{2u} g_0$ on $\mathcal{U}$. Recall from (8.3) that the area form $d\mu(t) = d\mu_{g(t)}$ of $g(t)$ is given by

$$d\mu(t) = \omega^1(t) \wedge \omega^2(t) = e^{2u(t)}\omega_0^1 \wedge \omega_0^2 = e^{2u(t)} d\mu_{g_0}$$

on $\mathcal{U}$. Thus,

$$\frac{\partial}{\partial t} d\mu(t) = 2\frac{\partial u}{\partial t}(t)\,\mathrm{e}^{2u(t)} d\mu_{g_0} = \big(r(t) - R(t)\big) d\mu(t).$$

Hence, on all of M^2 we have under the normalized Ricci flow that the area form of $g(t)$ evolves by

$$\frac{\partial}{\partial t} d\mu(t) = \big(r(t) - R(t)\big) d\mu(t). \tag{14.18}$$

Since $r(t)$ is the average of $R(t)$, we have

$$\begin{aligned} \frac{d}{dt}\operatorname{Area}(g(t)) &= \frac{d}{dt}\int_{M^2} d\mu(t) = \int_{M^2} \frac{\partial}{\partial t} d\mu(t) \\ &= \int_{M^2} \big(r(t) - R(t)\big) d\mu(t) \\ &= 0. \end{aligned} \tag{14.19}$$

Thus, under the normalized Ricci flow,

$$\operatorname{Area}(g(t)) \equiv \operatorname{Area}(g_0) \tag{14.20}$$

for all $t \in [0, T)$. As a consequence, by the Gauss–Bonnet formula, we have

$$r(t) = \frac{\int_{M^2} R(t)\, d\mu(t)}{\int_{M^2} d\mu(t)} \equiv \frac{4\pi\chi(M^2)}{\operatorname{Area}(g_0)} \tag{14.21}$$

is a constant independent of t. So we denote $r := r(t)$.

On the other hand, under the (unnormalized) Ricci flow (14.14), we have similarly to (14.18) that

$$\frac{\partial}{\partial t} d\mu(t) = -R(t)\, d\mu(t). \tag{14.22}$$

Therefore, under the Ricci flow we have (cf. (14.19))

$$\frac{d}{dt}\operatorname{Area}(g(t)) = -\int_{M^2} R(t)\, d\mu(t) = -4\pi\chi(M^2) = -r_0 \operatorname{Area}(g_0), \tag{14.23}$$

where r_0 is the average scalar curvature at time zero. We conclude that under the Ricci flow,

$$\operatorname{Area}(g(t)) = \operatorname{Area}(g_0) - 4\pi\chi(M^2)\, t = \operatorname{Area}(g_0)(1 - r_0 t). \tag{14.24}$$

See Figure 14.4.1.

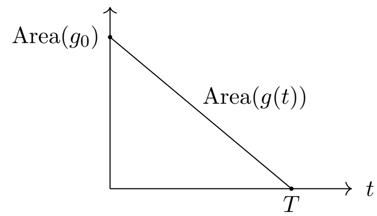

Figure 14.4.1. Area, as a function of time, of a closed surface with positive Euler characteristic under Ricci flow. The supremal time is $T = \frac{\text{Area}(g_0)}{4\pi\chi(M^2)}$.

14.5. The relation between the unnormalized and normalized Ricci flows

In this section we show that the unnormalized and normalized Ricci flows are related by a change in time parameter and by homothetic rescalings, depending on time, of the metrics. It is in this sense that solutions to the two flows with the same initial conditions are geometrically comparable: the shapes, but not the sizes, of the metrics are the same for the two flows.

Let $g(t)$ be a solution of the Ricci flow. Define space and time rescaled metrics by

$$\bar{g}(\bar{t}) := \frac{1}{1 - r_0 t} g(t), \tag{14.25}$$

where

$$\bar{t}(t) := \int_0^t \frac{1}{1 - r_0 \tau} d\tau = -\frac{1}{r_0} \ln(1 - r_0 t). \tag{14.26}$$

By (14.24), we have that

$$\text{Area}\left(\bar{g}(\bar{t})\right) \equiv \text{Area}(g_0). \tag{14.27}$$

We have

$$\frac{d\bar{t}}{dt}(t) = \frac{1}{1 - r_0 t}.$$

Using this, we compute that

$$\begin{aligned}
\frac{\partial}{\partial \bar{t}} \bar{g}(\bar{t}) &= \frac{1}{d\bar{t}/dt} \left(\frac{1}{1 - r_0 t} g(t) \right) \\
&= (1 - r_0 t) \frac{\partial}{\partial t} \left(\frac{1}{1 - r_0 t} g(t) \right) \\
&= \frac{\partial}{\partial t} g(t) + r_0 \frac{1}{1 - r_0 t} g(t) \\
&= -R(t) g(t) + r_0 \bar{g}(\bar{t}) \\
&= \left(r_0 - \bar{R}(\bar{t}) \right) \bar{g}(\bar{t}).
\end{aligned}$$

Thus, $\bar{g}(\bar{t})$ is a solution to the normalized Ricci flow with $\bar{g}(0) = g_0$.

Conversely, suppose that $\bar{g}(\bar{t})$ is a solution to the normalized Ricci flow with $\bar{g}(0) = g_0$. By reversing the discussion above, we have that if $t(\bar{t}) := \frac{1}{r_0}(1 - \mathrm{e}^{-r_0\bar{t}})$ and $g(t) := \mathrm{e}^{-r_0\bar{t}}\bar{g}(\bar{t})$, then $g(t)$ is a solution to the (unnormalized) Ricci flow with $g(0) = g_0$.

14.6. Short-time existence of the normalized Ricci flow

In order to use the Ricci flow, we need to first establish the short-time existence of solutions given an initial metric. By (14.15) and (14.8), we have that the function $u(x,t)$ satisfies

$$\frac{\partial u}{\partial t}(t) = \mathrm{e}^{-2u(t)}\Delta_0 u(t) - \mathrm{e}^{-2u(t)}\frac{R_0}{2} + \frac{r}{2}. \tag{14.28}$$

This is a heat-type equation in u. Technically, it has the fancy name of a *quasilinear second-order parabolic partial differential equation*. In any case, there is a well-developed theory of such equations and in particular we have the following well-known result. The proof of this result is beyond the scope of this book. See e.g. Friedman's book [**Fri64**] for the methods to prove such a result.

Lemma 14.5. *Given any function $u_0 : M^2 \to \mathbb{R}$, there exists $T \in (0,\infty]$ and a unique family of functions $u(t)$, $t \in [0,T)$, that satisfy the heat-type equation (14.28) with the initial condition $u(0) = u_0$.*

By taking $u_0 = 0$, i.e., the zero function, and by the equivalence of equations (14.28) and (14.14), we have the following.

Corollary 14.6 (Short-time existence and uniqueness)**.** *For any closed Riemannian surface (M^2, g_0), there exists $T \in (0,\infty]$ and a unique family of metrics $g(t)$, $t \in [0,T)$, that satisfy the normalized Ricci flow (14.14) with the initial condition $g(0) = g_0$.*

We take T to be the supremal time of existence. (In other words, $[0,T)$ is the maximal time interval of existence.) That is, by definition no continuation of the solution exists beyond time T. Later, we shall show that the supremal time of existence T of the normalized Ricci flow on surfaces is equal to ∞.

14.7. A lower bound for the curvature under the normalized Ricci flow

An important tool for studying heat-type equations is the parabolic maximum principle, which we introduce and apply in this section to study the behavior of the scalar curvatures of solutions to Ricci flow. We have seen the statement of the parabolic maximum principle for one-space and one-time dimensional heat-type equations in the previous chapter on the curve shortening flow. In this section we will give the statement and proof in more generality.

By Lemma 14.1, since $g(t)$ is a conformal deformation with velocity $v(t) = r - R(t)$, we have that scalar curvature satisfies the following evolution equation under the normalized Ricci flow:

$$\frac{\partial R}{\partial t}(t) = \Delta_{g(t)} R(t) + R(t)^2 - r\,R(t). \tag{14.29}$$

Using that r is constant in time, we may rewrite this formula as

$$\frac{\partial}{\partial t}(R(t) - r) = \Delta_{g(t)}(R(t) - r) + (R(t) - r)^2 + r\,(R(t) - r). \tag{14.30}$$

In particular, by dropping from the right-hand side the *square term*, which is non-negative, we obtain

$$\frac{\partial}{\partial t}(R(t) - r) \geq \Delta_{g(t)}(R(t) - r) + r\,(R(t) - r). \tag{14.31}$$

This, in turn, implies that

$$\frac{\partial}{\partial t}\big(\mathrm{e}^{-rt}(R(t) - r)\big) \geq \Delta_{g(t)}\big(\mathrm{e}^{-rt}(R(t) - r)\big). \tag{14.32}$$

14.7.1. The parabolic maximum principle on manifolds. In general, if $w(t) : M^n \to \mathbb{R}$, $t \in [0, T)$, are functions satisfying

$$\frac{\partial w}{\partial t}(x, t) \geq \Delta_{g(t)} w(x, t), \tag{14.33}$$

where $g(t)$, $t \in [0, T)$, is a 1-parameter family of Riemannian metrics on M^n, then we say that w is a **supersolution to the heat equation** (with normalized Ricci flow background). So $\mathrm{e}^{-rt}(R(t) - r)$ is a supersolution to the heat equation by (14.32).

The following is fundamentally important to estimating solutions to second-order parabolic partial differential equations. It has a wide range of applications and is "unreasonably effective".

Theorem 14.7 (Parabolic minimum principle for supersolutions to the heat equation)**.** *If $w : M^n \times [0, T) \to \mathbb{R}$, where M^n is compact, satisfies* (14.33) *and if $w(x, 0) \geq -C$ for all $x \in M^n$, where C is some constant, then*

$$w(x, t) \geq -C \quad \text{for all } x \in M^n,\ t \in [0, T). \tag{14.34}$$

Proof. The idea of the proof is simply the first and second derivative tests from calculus. The trick to implement this is to introduce a so-called fudge factor. To this end, let $\epsilon > 0$ and define

$$w_\epsilon(x, t) := w(x, t) + \epsilon t + \epsilon.$$

By (14.33), we have

$$\frac{\partial w_\epsilon}{\partial t}(x,t) \geq \Delta w_\epsilon(x,t) + \epsilon. \tag{14.35}$$

By hypothesis, $w_\epsilon(x,0) \geq -C + \epsilon$ for all $x \in M^n$.

Suppose for a contradiction that the function w_ϵ is less than $-C$ somewhere in $M^n \times [0,T)$. Then there exists a first time $t_0 \in (0,T)$ such that

$$w_\epsilon(x_0,t_0) = -C \quad \text{for some } \ x_0 \in M^n. \tag{14.36}$$

This is a rather intuitive result, true since w_ϵ is continuous and M^n is compact, which we will prove in the remark right after this proof.

By the choice of t_0, we have that $w_\epsilon(x,t) \geq -C$ for all $(x,t) \in M^n \times [0,t_0]$. By the first derivative test, since w_ϵ on $M^n \times [0,t_0]$ attains its minimum at (x_0,t_0), we have

$$\frac{\partial w_\epsilon}{\partial t}(x_0,t_0) \leq 0,$$
$$\nabla w_\epsilon(x_0,t_0) = \vec{0};$$

see Figure 14.7.1. By the second derivative test (11.5), we have that

$$(\nabla^2 w_\epsilon)_{(x_0,t_0)} \geq 0$$

is positive semi-definite. In particular, by tracing this, we obtain

$$(\Delta w_\epsilon)(x_0,t_0) \geq 0;$$

see Figure 14.7.2. By applying the first and second derivative tests to (14.35), we obtain

$$0 \geq \frac{\partial w_\epsilon}{\partial t}(x_0,t_0) \geq (\Delta w_\epsilon)(x,t) + \epsilon \geq \epsilon.$$

This is a contradiction since $\epsilon > 0$. Therefore, $w_\epsilon \geq -C$ on all of $M^n \times [0,T)$. By taking $\epsilon \to 0$, we conclude that $w \geq -C$ on all of $M^n \times [0,T)$. □

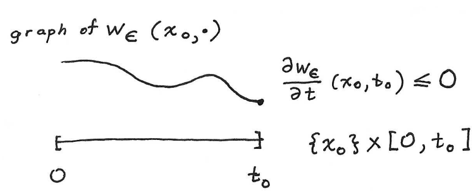

Figure 14.7.1. The first derivative test: At the minimum point (x_0,t_0) we have $\frac{\partial w_\epsilon}{\partial t} \leq 0$.

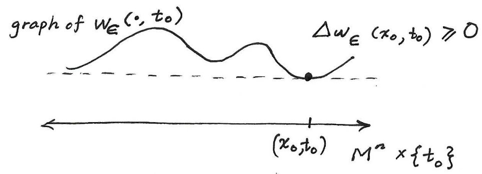

Figure 14.7.2. The second derivative test: At the minimum point (x_0, t_0) we have $\Delta w_\epsilon \geq 0$.

Remark 14.8. We give a proof of (14.36). Let

$$t_0 := \sup \left\{ \bar{t} \in [0, T) : w_\epsilon > -C \text{ on } M^n \times [0, \bar{t}] \right\}.$$

Firstly, since $w_\epsilon(\cdot, 0) \geq -C + \epsilon$ on M^n and since w_ϵ is continuous, we have that $t_0 > 0$. Secondly, since $w_\epsilon < -C$ somewhere in $M^n \times [0, T)$, we have $t_0 < T$. Thirdly, by the definition of t_0, we have $w_\epsilon(\cdot, t_0) \geq -C$ on M^n.

Suppose for a contradiction that $w_\epsilon(\cdot, t_0) > -C$ on all of M^n. Since M^n is compact, this implies that $w_\epsilon(\cdot, t_0) \geq -C + \delta$ on M^n for some constant $\delta > 0$. Since w_ϵ is continuous and since M^n is compact, there exists $\eta > 0$ such that $w_\epsilon \geq -C$ on $M^n \times [t_0 + \eta]$. This is a contradiction to the definition of t_0.[2] We conclude that $w_\epsilon(\cdot, t_0) = -C$ somewhere on M^n.

14.7.2. Applying the maximum principle to bound the scalar curvature from below. By applying the parabolic minimum principle (Theorem 14.7) to (14.32), we have that if $R_0 - r \geq -C$ (such a C always exists since M^2 is compact), then

$$\mathrm{e}^{-rt}(R(t) - r) \geq -C. \tag{14.37}$$

That is, under the normalized Ricci flow on surfaces, we have the estimate:

$$R(t) - r \geq -C\mathrm{e}^{rt}. \tag{14.38}$$

This estimate is particularly effective when $r < 0$. This is because in this case we have a lower bound for $\min_{x \in M^2}(R(x, t) - r)$ that is exponentially decaying in time. By the Gauss–Bonnet formula, the condition that $r < 0$ is equivalent to the topological condition that $\chi(M^2) < 0$, that is, the genus of M^2 is $\mathbf{g} := \mathbf{g}(M^2) > 1$.

Exercise 14.1 (Parabolic maximum principle for subsolutions of the heat equation)**.** *Prove that if $w : M^n \times [0, T) \to \mathbb{R}$, where M^n is compact, satisfies*

$$\frac{\partial w}{\partial t}(x, t) \leq \Delta w(x, t), \tag{14.39}$$

[2] A proof by contradiction of this: If no such η exists, then there exists a sequence (x_i, t_i) with $x_i \in M^n$ and $t_i \searrow t_0$ such that $w_\epsilon(x_i, t_i) \leq -C + \frac{1}{i}$. Since M^n is compact, we may pass to a subsequence so that $x_i \to x_\infty \in M^n$. By the continuity of w_ϵ, we have $w_\epsilon(x_\infty, t_0) = \lim_{i \to \infty} w_\epsilon(x_i, t_i) \leq -C$, which is a contradiction.

and if $w(x,0) \le C$ *for all* $x \in M^n$*, where* C *is some constant, then*

$$w(x,t) \le C \quad \text{for all } x \in M^n,\ t \in [0,T). \tag{14.40}$$

Exercise 14.2 (Parabolic maximum principles for linear heat-type equations)**.** (1) *Prove that if* $w : M^n \times [0,T) \to \mathbb{R}$*, where* M^n *is compact, satisfies*

$$\frac{\partial w}{\partial t}(x,t) \ge \Delta w(x,t) + cw(x,t), \tag{14.41}$$

and if $w(x,0) \ge -C$ *for all* $x \in M^n$*, where* c *and* C *are constants, then*

$$w(x,t) \ge -C\mathrm{e}^{ct} \quad \text{for all } x \in M^n,\ t \in [0,T). \tag{14.42}$$

(2) *Similarly, if*

$$\frac{\partial w}{\partial t}(x,t) \le \Delta w(x,t) + cw(x,t), \tag{14.43}$$

and if $w(\cdot,0) \le C$*, then*

$$w(x,t) \le C\mathrm{e}^{ct}. \tag{14.44}$$

14.8. Estimating the curvature from above under the normalized Ricci flow

14.8.1. The difficulty in obtaining an upper bound for the curvature. Unlike the case of a lower bound, an effective *upper* bound for $R(x,t) - r$ under the normalized Ricci flow on a 2-sphere is not as obvious. Indeed, let

$$\overline{R}(x,t) := R(x,t) - r$$

be the scalar curvature minus its average. Then (14.30) is the reaction-diffusion equation

$$\left(\frac{\partial}{\partial t} - \Delta\right)\overline{R} = \overline{R}^2 + r\overline{R}. \tag{14.45}$$

The associated ODE to the PDE (14.45) is obtained by dropping the Laplacian term; this yields the equation:

$$\frac{d}{dt}\mathrm{S} = \mathrm{S}^2 + r\mathrm{S}. \tag{14.46}$$

The solution to this ODE with initial data $\mathrm{S}(0) = \mathrm{S}_0 \neq 0$ is given by

$$\mathrm{S}(t) = \frac{r}{1 - (1 - r/\mathrm{S}_0)\mathrm{e}^{rt}}. \tag{14.47}$$

Observe that if $\mathrm{S}_0 > 0$, then

$$\mathrm{S}(t) \to \infty \quad \text{as } t \to T,$$

where $T := -\frac{1}{r}\ln(1 - r/\mathrm{S}_0)$. That is, we have finite-time blow up of the solution to the ODE.

The statement of the parabolic maximum principle for reaction-diffusion equations with nonlinear reaction terms is as follows.

Lemma 14.9. *Suppose that $g(t)$, $t \in [0, T)$, is a smooth 1-parameter family of Riemannian metrics on a closed differentiable manifold M^n. Let $u : M^n \times [0, T) \to \mathbb{R}$ be a supersolution to*

$$\frac{\partial u}{\partial t}(x,t) = \Delta_{g(t)} u(x,t) + F(u(x,t)), \tag{14.48}$$

where $F : \mathbb{R} \to \mathbb{R}$ is some smooth one-variable function. Let $U_0 \in \mathbb{R}$ satisfy $U_0 \geq \max_{M^n} u(\cdot, 0)$. Let $U(t)$, $t \in T'$, be the solution the associated ODE

$$\frac{dU}{dt}(t) = F(U(t)), \quad U(0) = U_0. \tag{14.49}$$

Then we have that

$$u(x,t) \leq U(t) \tag{14.50}$$

for all $x \in M^n$ and $t \in [0, \min\{T, T'\})$.

As a consequence of this parabolic maximum principle, by choosing $\mathrm{S}_0 := \max_{M^2} \overline{R}(\cdot, 0)$, we obtain the upper estimate for the scalar curvature:

$$R(x,t) - r \leq \mathrm{S}(t) \tag{14.51}$$

for all $x \in M^2$ and $t \in [0, \min\{T, T'\})$. See Figure 14.8.1. Unfortunately, $T' < \infty$ provided g_0 does not have constant curvature (which means $\mathrm{S}_0 > 0$), so we cannot get an upper bound for all time for R. We need another method.

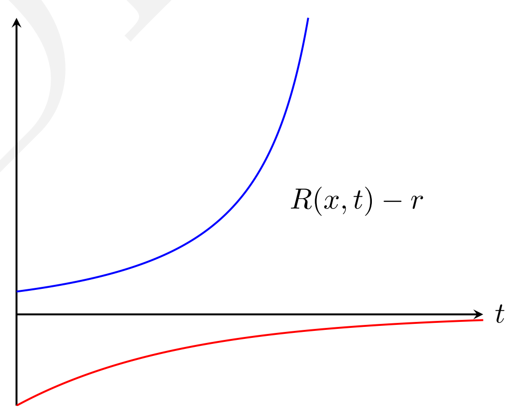

Figure 14.8.1. The lower bound (14.38) for $R(x,t) - r$ is represented by the red curve. The blue curve represents the upper bound given by the solution (14.47) to the associated ODE.

14.8.2. A key tool: The potential function. Necessity is the mother of invention. A simple, but not obvious, method to obtain an effective upper bound for $R(x,t)-r$ proceeds as follows. We carry this out in a few steps. Firstly, by definition,

$$\int_{M^2} \overline{R}(x,t)\, d\mu(x,t) = 0$$

for each time t. Because of this, by Corollary 11.19 there exists a function $f(t) : M^2 \to \mathbb{R}$ satisfying the Poisson-type equation:

$$\Delta_{g(t)} f(t) = \overline{R}(t) \tag{14.52}$$

on M^2. Note that each $f(t)$ is determined up to an additive constant. This is because any harmonic function on M^2 is a constant (see Lemma 11.13). We call $f(t)$ the **potential function**.

Recall that the curvature is defined in terms of the second derivatives of the metric. On the other hand, from (14.8) we saw that the scalar curvature of a conformally related metric may be expressed in terms of the Laplacian of the conformal factor. So, by analogy, the consideration of the potential function seems to be a reasonable thing to do. Let us now see if it helps.

14.8.3. Estimates for the potential function and its derivatives. Secondly, because we are in dimension 2, using Lemma 11.2 we calculate that

$$\begin{aligned} \frac{\partial}{\partial t}\big(\Delta_{g(t)} f(t)\big) &= \frac{\partial}{\partial t}\big(\mathrm{e}^{-2u(t)} \Delta_{g_0} f(t)\big) \\ &= -2\frac{\partial u}{\partial t}(t)\mathrm{e}^{-2u(t)} \Delta_{g_0} f(t) + \mathrm{e}^{-2u(t)} \Delta_{g_0} \left(\frac{\partial f}{\partial t}\right) \\ &= \overline{R}(t)\, \Delta_{g(t)} f(t) + \Delta_{g(t)} \left(\frac{\partial f}{\partial t}\right). \end{aligned} \tag{14.53}$$

Thus, by taking the time-derivative of (14.52), we obtain

$$\overline{R}(t)\, \Delta_{g(t)} f(t) + \Delta_{g(t)} \left(\frac{\partial f}{\partial t}\right) = \Delta \overline{R} + \overline{R}^2 + r\overline{R}.$$

In view of (14.52), we can rewrite this equation as

$$\Delta_{g(t)} \left(\frac{\partial f}{\partial t}\right) = \Delta\big(\Delta_{g(t)} f(t)\big) + r\Delta_{g(t)} f(t).$$

Again, since any harmonic function on M is a constant, this implies that there exist constants $C(t)$ such that

$$\frac{\partial f}{\partial t}(t) = \Delta_{g(t)} f(t) + r f(t) + C(t).$$

In the definition of $f(t)$ we can choose $f(t)$ so that these constants $C(t)$ are identically zero, that is, so that

$$\frac{\partial f}{\partial t}(t) = \Delta_{g(t)} f(t) + r f(t). \tag{14.54}$$

For simplicity, we write this equation as:

Lemma 14.10. *Under the normalized Ricci flow on a closed surface, the potential function f satisfies*

$$\left(\frac{\partial}{\partial t} - \Delta\right) f = r f. \tag{14.55}$$

If, given a family of metrics $g(t)$, we consider the equation

$$\left(\frac{\partial}{\partial t} - \Delta_{g(t)}\right) w(t) = r w(t), \tag{14.56}$$

then we have an equation which is linear in w. It is in this sense that (14.54) is a linear heat-type equation. On the other hand, $f(t)$ itself does not depend linearly on $g(t)$.

Thirdly, it is useful to consider the gradient of f. Let $g(t)^* = \langle \cdot, \cdot \rangle$ denote the inner product on T^*M dual to the metric $g(t)$. Then $\partial_t g(t)^* = \overline{R} g(t)^*$. So we compute that

$$\begin{aligned} \frac{\partial}{\partial t} \|\nabla f(t)\|^2_{g(t)} &= \frac{\partial}{\partial t} \big(g(t)^* \big(df(t), df(t)\big)\big) \\ &= \overline{R} g(t)^* \big(df(t), df(t)\big) + 2\big\langle \partial_t \big(df(t)\big), df(t)\big\rangle. \end{aligned} \tag{14.57}$$

Now,

$$\begin{aligned} \partial_t \big(df(t)\big) &= d\big(\partial_t f(t)\big) = d(\Delta f + r f) \\ &= \Delta df - \mathrm{Ric}(df) + r df \\ &= \Delta df - \frac{1}{2} R df + r df, \end{aligned} \tag{14.58}$$

where $\mathrm{Ric} : T^*M \to T^*M$ in the third line, where we used Lemma 11.5 to obtain the third equality, and where we used that $\mathrm{Ric} = \frac{1}{2} R g$ from $n = 2$ in the fourth line. Thus, by applying (14.58) to (14.57), we have that

$$\begin{aligned} \frac{\partial}{\partial t} \|\nabla f(t)\|^2_{g(t)} &= \overline{R} \|\nabla f(t)\|^2_{g(t)} + 2\langle \Delta df, df \rangle \\ &\quad - R \|\nabla f(t)\|^2_{g(t)} + 2r \|\nabla f(t)\|^2_{g(t)} \\ &= 2\langle \Delta df, df \rangle + r \|\nabla f(t)\|^2_{g(t)} \\ &= \Delta_{g(t)} \|\nabla f(t)\|^2_{g(t)} - 2\|\nabla^2 f(t)\|^2_{g(t)} + r \|\nabla f(t)\|^2_{g(t)}. \end{aligned} \tag{14.59}$$

For simplicity, we write this equation as:

Lemma 14.11. *Under the normalized Ricci flow on a closed surface, the norm squared of the gradient of the potential function satisfies*

$$\left(\frac{\partial}{\partial t} - \Delta\right) \|\nabla f\|^2 = -2\|\nabla^2 f\|^2 + r\|\nabla f\|^2. \tag{14.60}$$

We remark that the general formula for computing the heat operator applied to $\|\nabla v\|^2$ for a function $v = v(x,t)$ is given by (14.77) below. The "good" Hessian norm squared term is common to such calculations.

Fourthly, from the point of view of bounding the quantity by the parabolic maximum principle, the term $-2\|\nabla^2 f\|^2$ is a good term. In fact, we have

$$\|\nabla^2 f\|^2 \geq \frac{1}{2}\big(\operatorname{trace}_g(\nabla^2 f)\big)^2 = \frac{1}{2}(\Delta f)^2 = \overline{R}^2. \tag{14.61}$$

Because of this good term, the heat-type equation (14.60) for $\|\nabla f\|^2$ is useful for controlling the bad term $\overline{R}^2$ on the right-hand side of the equation (14.45) for $\overline{R}$. So we consider the sum

$$h := \overline{R} + \|\nabla f\|^2. \tag{14.62}$$

By (14.60) and (14.45), we have

$$\begin{aligned}\left(\frac{\partial}{\partial t} - \Delta\right) h &= \overline{R}^2 + r\overline{R} - 2\|\nabla^2 f\|^2 + r\|\nabla f\|^2 \\ &= -2\left\|-\frac{1}{2}\overline{R}g + \nabla^2 f\right\|^2 + rh.\end{aligned} \tag{14.63}$$

To see the last equality, we calculate using $\|g\|^2 = n = 2$ that

$$\left\|-\frac{1}{2}\overline{R}g + \nabla^2 f\right\|^2 = \frac{1}{2}\overline{R}^2 + \|\nabla^2 f\|^2 - \overline{R}\Delta f = \|\nabla^2 f\|^2 - \frac{1}{2}\overline{R}^2.$$

Consequently,

Lemma 14.12. *Under the normalized Ricci flow on a closed surface,*

$$\left(\frac{\partial}{\partial t} - \Delta\right) h \leq rh. \tag{14.64}$$

By applying the parabolic maximum principle (Exercise 14.2(2)), we have

$$h(x,t) \leq C\mathrm{e}^{rt}, \tag{14.65}$$

where $C := \max_{y \in M^2} h(y,0)$. In particular, since $\overline{R} \leq h$, we have

$$\overline{R}(x,t) \leq C\mathrm{e}^{rt}. \tag{14.66}$$

To wit, in order to estimate the curvature $\overline{R}$, we estimated the larger quantity h since it satisfies a better heat-type equation.

On the other hand, by (14.38) we have

$$\overline{R}(x,t) \geq -Ce^{rt} \tag{14.67}$$

for some constant C. Thus:

Lemma 14.13 (Curvature estimate under the normalized Ricci flow)**.** *Under the normalized Ricci flow on a closed surface, there exists a constant C depending only on the initial metric g_0 such that*

$$\left|\overline{R}\right|(x,t) \leq Ce^{rt} \tag{14.68}$$

for all $x \in M^2$ and $t \in [0,T)$. In particular, if the genus $\mathbf{g} > 1$, or equivalently $\chi(M^2) < 0$, so that $r < 0$, we have the exponential decay of $|\overline{R}|$.

14.9. Uniform convergence of the metric as $t \to T$

We now show that the exponential decay estimate in Lemma 14.13 is sufficient to prove the uniform convergence of $g(t)$ as $t \to T$. As in (14.1), define $u(t) : M^2 \to \mathbb{R}$, $t \in [0,T)$, by

$$g(t) =: e^{2u(t)} g_0.$$

Then, by (14.15), the conformal factor u satisfies

$$\frac{\partial u}{\partial t} = -\frac{1}{2}\overline{R}. \tag{14.69}$$

Integrating this, we see that for each $x \in M^2$ and $t_1 < t_2$,

$$u(x,t_1) - u(x,t_2) = \frac{1}{2}\int_{t_1}^{t_2} \overline{R}(x,t)dt.$$

Hence, using $r < 0$, we compute that

$$|u(x,t_1) - u(x,t_2)| \leq \frac{1}{2}\int_{t_1}^{t_2} \left|\overline{R}\right|(x,t)dt \leq C\int_{t_1}^{t_2} e^{rt}dt \leq \frac{C}{|r|}e^{rt_1}$$

for some constant C. Note that C is independent of $x \in M^2$ and $t_2 \in (t_1, T)$. As a consequence, we have:

(1) There exists a constant C such that

$$|u|(x,t) \leq C \tag{14.70}$$

for all $x \in M^2$ and $t \in [0,T)$.

(2) For each $x \in M^2$, the limit

$$\lim_{t \to T} u(x,t) =: u_T(x) \tag{14.71}$$

exists. This statement is true even if $T = \infty$. (We will prove later that $T = \infty$.)

The proof of (2) is as follows. Having seen the proof of (2), we leave the proof of (1) as an exercise. Choose any sequence $t_i \to T$. We have for any $i < j$ that

$$|u(x, t_i) - u(x, t_j)| \leq \frac{C}{|r|}(\mathrm{e}^{rt_i} - \mathrm{e}^{rt_j}) \leq \frac{C}{|r|}(\mathrm{e}^{rt_i} - \mathrm{e}^{rT}), \tag{14.72}$$

where $\mathrm{e}^{rT} := 0$ if $T = \infty$. This shows that $\{u(x, t_i)\}_{i=1}^{\infty}$ is a Cauchy sequence. Since every Cauchy sequence of real numbers converges, we have that

$$\lim_{i \to \infty} u(x, t_i) =: u_T(x)$$

exists for each $x \in M^2$. Now, for any $t \in (t_i, T)$ and $x \in M^2$, we have

$$|u(x, t_i) - u(x, t)| \leq \frac{C}{|r|}(\mathrm{e}^{rt_i} - \mathrm{e}^{rT}). \tag{14.73}$$

This implies that the convergence

$$\lim_{t \to T} u(x, t) =: u_T(x)$$

is *uniform*. By definition, this means that for any $\epsilon > 0$, there exists $t_\epsilon < T$ such that for all $x \in M^2$ and $t \in (t_\epsilon, T)$ we have

$$|u(x, t) - u_T(x)| < \epsilon.$$

Note that we have not yet established any regularity properties of u_T such as continuity or higher differentiability. This will be a goal of the following sections.

In any case, as a consequence of (14.70) in (1), we have

$$\mathrm{e}^{-2C} g_0 \leq g(t) \leq \mathrm{e}^{2C} g_0 \tag{14.74}$$

for all $t \in [0, T)$. In general, given two metrics g and g', we say that $g \leq g'$ if $g' - g$ is a positive semi-definite symmetric 2-tensor. Hence, if α is any k-tensor, then

$$\mathrm{e}^{-kC} \|\alpha\|_{g_0} \leq \|\alpha\|_{g(t)} \leq \mathrm{e}^{kC} \|\alpha\|_{g_0} \tag{14.75}$$

for all $t \in [0, T)$. As a consequence of (2) and (1), we have that

$$\lim_{t \to T} \|g(t) - g_T\|_{g_0} = 0, \tag{14.76}$$

where

$$g_T := \mathrm{e}^{2u_T} g_0.$$

14.10. Estimating the gradient of the curvature

Similarly to the previous chapter on the curve shortening flow, in view of the Arzelà–Ascoli Theorem, we need to estimate the derivatives of the curvature of our solution to the normalized Ricci flow.

14.10.1. Estimating the gradient of the curvature. In general, for a time-dependent function $v(t)$ and under the normalized Ricci flow on surfaces, using the same method as that to obtain (14.57) and (14.59), we compute that

$$\left(\frac{\partial}{\partial t}-\Delta\right)\|\nabla v(t)\|^2_{g(t)} = -2\|\nabla^2 v\|^2 - r\|\nabla v\|^2 + 2d\left(\left(\frac{\partial}{\partial t}-\Delta\right)v\right)\cdot dv. \tag{14.77}$$

By applying this formula to $v(t)=R(t)$, we obtain

$$\begin{aligned}\left(\frac{\partial}{\partial t}-\Delta\right)\|\nabla R\|^2 &= -2\|\nabla^2 R\|^2 - r\|\nabla R\|^2 + 2d(R^2)\cdot dR - 2r dR\cdot dR\\ &= -2\|\nabla^2 R\|^2 + 4R\|\nabla R\|^2 - 3r\|\nabla R\|^2.\end{aligned}$$

We rewrite this as

$$\left(\frac{\partial}{\partial t}-\Delta\right)\|\nabla R\|^2 = -2\|\nabla^2 R\|^2 + 4\overline{R}\|\nabla R\|^2 + r\|\nabla R\|^2. \tag{14.78}$$

Assume that $\chi(M^2)<0$. Since $\overline{R}\le Ce^{rt}$ and $r<0$, there exists $t_0<\infty$ such that $\overline{R}(x,t)\le -\frac{1}{8}r$ for all $t\ge t_0$ and $x\in M^2$. We then obtain for $t\ge t_0$ that

$$\left(\frac{\partial}{\partial t}-\Delta\right)\|\nabla R\|^2 \le -2\|\nabla^2 R\|^2 + \frac{r}{2}\|\nabla R\|^2 \le \frac{r}{2}\|\nabla R\|^2. \tag{14.79}$$

Hence, by Exercise 14.2(2) on the parabolic maximum principle, we have:

Lemma 14.14. *Under the normalized Ricci flow on a closed surface M^2 with $\chi(M^2)<0$, there exists a constant C depending only on the initial metric g_0 such that*

$$\|\nabla R\|^2(x,t)\le C\mathrm{e}^{\frac{r}{2}t}, \tag{14.80}$$

where the norm is with respect to $g(t)$.

14.10.2. Estimating the higher derivatives of curvature. For the higher-order derivatives of R, one can prove the following.

Lemma 14.15 (Higher derivatives of curvature estimate)**.** *Under the normalized Ricci flow on a closed surface M^2 with $\chi(M^2)<0$ and for each positive integer k, there exists a positive constants C_k depending only on the initial metric g_0 and k such that*

$$\|\nabla^k R\|^2(x,t)\le C_k\mathrm{e}^{\frac{r}{2}t} \tag{14.81}$$

for all $x\in M^2$ and $t\in[0,T)$.

As an example of how the proof of the higher derivative of curvature estimates proceed, we sketch the proof of the second derivative estimate; i.e., the case where $k = 2$. Details are given in Chapter 5 of [CK04]. By [CK04, Lemma 5.25], we have

$$\frac{\partial}{\partial t}\|\nabla^2 R\|^2 = \Delta\|\nabla^2 R\|^2 - 2\|\nabla^3 R\|^2 + (2R - 4r)\,\|\nabla^2 R\|^2 + 2R\,(\Delta R)^2 + 2\big\langle \nabla R, \nabla\,|\nabla R|^2\,\big\rangle.$$

Now let

$$\varphi := \|\nabla^2 R\|^2 - 3r\|\nabla R\|^2.$$

Then there exists a constant C depending only on $g(0)$ such that (see the proof of Corollary 5.26 in [CK04])

$$\frac{\partial\varphi}{\partial t} \leq \Delta\varphi + \frac{2r}{3}\varphi + C\mathrm{e}^{rt}.$$

In particular, for any (x, t) such that $\varphi(x,t) \geq -\frac{6C}{r}\mathrm{e}^{rt}$, we have

$$\frac{\partial\varphi}{\partial t}(x,t) \leq \Delta\varphi(x,t) + \frac{r}{2}\varphi(x,t).$$

By (a slight variant of) the parabolic maximum principle, we conclude that

$$\|\nabla^2 R\|^2 \leq \varphi \leq C\mathrm{e}^{\frac{r}{2}t} \tag{14.82}$$

for some constant C depending only on $g(0)$.

14.11. Long-time existence and convergence when the genus $\mathbf{g} > 1$

Given the curvature and its derivatives estimates of the previous section, we are now in position to prove the long-time existence and convergence to constant negative curvature of the normalized Ricci flow with any initial metric on a surface with genus $\mathbf{g} > 1$.

14.11.1. Arzelà–Ascoli Theorem and equicontinuous families of functions. Let (M, d) be a metric space. Recall that a family $\mathcal{F}$ of real-valued functions on M is **equicontinuous** if for any $\varepsilon > 0$ there exists $\delta > 0$ such that for all $\phi \in \mathcal{F}$ and all $x, y \in M$, if $d(x, y) < \delta$, then

$$|\phi(x) - \phi(y)| < \varepsilon.$$

Example 14.16. Let (M^n, g) be a Riemannian manifold. Suppose that $\mathcal{F}$ is a family of functions on M^n that are uniformly Lipschitz; that is, there exists a positive constant C such that for all $\phi \in \mathcal{F}$ and all $x, y \in M^n$,

$$|\phi(x) - \phi(y)| \leq Cd(x, y). \tag{14.83}$$

Then $\mathcal{F}$ is an equicontinuous family. Indeed, given $\varepsilon > 0$, we may let $\delta = \frac{\varepsilon}{C}$ for the definition of equicontinuity.

In particular, if $\mathcal{F}$ is a family of differentiable functions on M^n that satisfy a uniform derivative bound, i.e., for all $\phi \in \mathcal{F}$ and $x \in M^n$,

$$\|d\phi_x\| \leq C, \tag{14.84}$$

then $\mathcal{F}$ is an equicontinuous family.

Now suppose that $\phi : M^n \to \mathbb{R}$ is twice differentiable. We have that its derivative is a real-valued function on the tangent bundle: $d\phi : TM \to \mathbb{R}$. Observe that if $\|\nabla d\phi\| \leq C$ on M^n, then we have that the derivative of the restricted function $d\phi : SM \to \mathbb{R}$ is bounded by C, where SM denotes the unit tangent bundle. Thus, if $\mathcal{F}$ is a family of twice-differentiable functions on M^n such that $\|\nabla d\phi\| \leq C$ for all $\phi \in \mathcal{F}$ for some constant C, then the family

$$\mathcal{G} := \{d\phi : \phi \in \mathcal{F}\} \tag{14.85}$$

of functions on SM is equicontinuous.

We have the following fundamental result in analysis; see e.g. [**Rud76**, Theorem 7.25].

Theorem 14.17 (Arzelà and Ascoli)**.** *Suppose that (M, d) is a compact metric space. If $\{\phi_i\}$ is a uniformly bounded and equicontinuous sequence of real-valued functions on M, then there exists a subsequence $\{\phi_{i_j}\}$ that converges uniformly to a continuous function ϕ_∞ on M.*

We also have the following regarding the uniform convergence of derivatives; see e.g. [**Rud76**, Theorem 7.17] for the 1-dimensional case.

Theorem 14.18. *Let (M^n, g) be a Riemannian manifold and let $\{\phi_i\}$ be a sequence of real-valued functions on M^n. Suppose that $\{\phi_i\}$ converges uniformly to a function ϕ_∞ and that $\{d\phi_i\}$ converges uniformly to a 1-form ψ_∞. Then ϕ_∞ is differentiable and $d\phi_\infty = \psi_\infty$.*

By combining the preceding theorem with Theorem 14.17, we obtain:

Theorem 14.19. *Let (M^n, g) be a Riemannian manifold and let $\{\phi_i\}$ be a sequence of real-valued functions on M^n with the property that the functions and their first and second derivatives are uniformly bounded. Then there exists a subsequence $\{\phi_{i_j}\}$ such that $\{\phi_{i_j}\}$ converges uniformly to a continuously differentiable function ϕ_∞ on M^n and $\{d\phi_{i_j}\}$ converges uniformly to the function $d\phi_\infty$ on SM.*

We remark that the subsequence $\{d\phi_{i_j}\}$ converging uniformly to the function $d\phi_\infty$ on SM implies that $\{d\phi_{i_j}\}$ converges uniformly to $d\phi_\infty$ as sections of the cotangent bundle T^*M; i.e., as maps from M to T^*M whose composition with the projection map $T^*M \to M^n$ is the identity map of M^n.

We also recall the following result.

Lemma 14.20. *Let $\phi_t : X \to \mathbb{R}$, $t \in (0, T)$, where $T \in (0, \infty]$, be a family of functions in a set X with the property that*

$$|\partial_t \phi_t(x)| \leq \alpha(t) \tag{14.86}$$

for all $x \in X$ and $t \in [0, T)$, where $\alpha : [0, T) \to \mathbb{R}_+$ is a function satisfying

$$\int_0^T \alpha(t)\, dt < \infty. \tag{14.87}$$

Then there exists a function $\phi_T : X \to \mathbb{R}$ such that ϕ_t converges uniformly to ϕ_T as $t \to T$; that is, for any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in X$ and $t \in (T - \delta, T)$, we have

$$|\phi_t(x) - \phi_T(x)| < \varepsilon. \tag{14.88}$$

Proof. For any $0 \leq t_1 < t_2 < T$ and $x \in X$, we have

$$|\phi_{t_1}(x) - \phi_{t_2}(x)| \leq \int_{t_1}^{t_2} |\partial_t \phi_t(x)|\, dt \leq \int_{t_1}^{t_2} \alpha(t)\, dt.$$

Now let $\varepsilon > 0$. By hypothesis, there exists $\delta > 0$ such that $\int_{t_1}^{T} \alpha(t)\, dt < \varepsilon$ provided $t_1 \geq T - \delta$. Thus, for any $T - \delta \leq t_1 < t_2 < T$ and $x \in X$, we have

$$|\phi_{t_1}(x) - \phi_{t_2}(x)| < \varepsilon.$$

We leave it as an exercise to deduce the lemma. □

14.11.2. Convergence of the metrics $g(t)$ in each C^k-norm to a smooth metric g_T. We now proceed to prove that $g_T = \lim_{t \to T} g(t)$ is a C^∞ Riemannian metric on M^2. We start by estimating the first spatial derivative of u. We have

$$\frac{\partial}{\partial t} du(x, t) = -\frac{1}{2} dR(x, t) \tag{14.89}$$

as an equation for 1-forms. Thus,

$$du(x, t_1) - du(x, t_2) = \frac{1}{2} \int_{t_1}^{t_2} dR(x, t) dt \in T_x^* M.$$

For the right-hand side, the integration of the vector-valued function $t \mapsto dR(x, t)$ from $[0, T)$ to $T_x^* M$ is defined in the usual way. Taking norms, we obtain the estimate

$$\begin{aligned}\|du(x,t_1) - du(x,t_2)\|_{g_0} &\leq \frac{1}{2}\int_{t_1}^{t_2} \|dR(x,t)\|_{g_0} dt \\ &\leq C\int_{t_1}^{t_2} \mathrm{e}^{\frac{r}{4}t} dt \\ &\leq \frac{4C}{|r|}\left(\mathrm{e}^{\frac{r}{4}t_1} - \mathrm{e}^{\frac{r}{4}t_2}\right).\end{aligned}$$

From this it follows that the limit

$$\lim_{t\to T} du(x,t) =: v_T(x) \tag{14.90}$$

exists and that the convergence is uniform. By Theorem 14.18, u_T is differentiable and $du_T = v_T$. In fact, in a similar vein one can prove that for all $k \geq 1$, u_T is k-times differentiable and $\nabla^k u(t)$ converges uniformly to $\nabla^k u_T$ in the bundle of k-tensors $\otimes^k T^*M$. Hence, $g_T = \mathrm{e}^{2u_T} g_0$ is a C^∞ metric.

Now, if $T < \infty$, then we may continue the solution and there exists $\epsilon > 0$ and metrics $g(t)$, $t \in [T, T+\epsilon)$, solving the normalized Ricci flow (14.14) with $g(T) = g_T$. As such the two families of metrics $\{g(t)\}_{t\in[0,T)}$ and $\{g(t)\}_{t\in[T,T+\epsilon)}$ combine to form a solution to (14.14) on the time interval $[0, T+\epsilon)$ with $g(0) = g_0$. This contradicts T being the maximal time. Hence, we conclude that $T = \infty$.

Now that we know that $T = \infty$, we have shown above that g_∞ is a C^∞ metric on M^2. Furthermore, since $\nabla^k u(t)$ converges uniformly to $\nabla^k u_\infty$ as $t \to \infty$, we have that $R(t)$ converges uniformly to $R(g_\infty)$. By the estimate (14.68), we conclude that $R(g_\infty) \equiv r$. That is, g_∞ is a constant negative scalar curvature r metric. In summary, we have proved that for any initial metric on a surface of *genus greater than one* (i.e., *negative Euler characteristic*), the normalized Ricci flow exists for all positive time and converges to a constant negative curvature metric as time approaches infinity. This proves Theorem 14.21 below in the case where the genus $\mathbf{g} > 1$; i.e., the Euler characteristic of M^2 is negative.

Using similar techniques, one can prove that for any initial metric g_0 on a closed oriented surface with *zero Euler characteristic*, i.e., on a *torus*, a unique solution to the normalized Ricci flow exists for all $t \in [0, \infty)$ and that $g(t)$ converges to a C^∞ metric g_∞ as $t \to \infty$, where the curvature of g_∞ is identically zero. For details in this case, the reader may consult the original [**Ham88**] or Chapter 5 of the expository [**CK04**].

The statement of the global existence and convergence result for *all* closed surfaces is as follows.

Theorem 14.21 (Uniformization theorem by Ricci flow)**.** *Let (M^2, g_0) be a closed oriented Riemannian surface. Then there exists a solution $g(t)$ to* *the normalized Ricci flow for all time* $t \in [0,\infty)$ *with* $g(0) = g_0$. *As* $t \to \infty$, $g(t)$ *converges in each* C^k*-norm to a* C^∞ *metric* g_∞ *with constant scalar curvature equal to* $\frac{4\pi\chi(M^2)}{\mathrm{Area}(g_0)}$.

In the next section we consider the proof of this theorem in the special case where the Euler characteristic χ of M^2 is positive, having covered the case where $\chi(M^2) < 0$ above (references containing the $\chi(M^2) = 0$ are given above).

14.12. The Ricci flow on the 2-sphere

In this section and the next, we present the essential details of the proof of the convergence of the Ricci flow on closed surfaces with positive Euler characteristic. Since we assume that our surface is oriented, this means that our surface is diffeomorphic to the 2-sphere S^2.

14.12.1. Using monotone quantities to find more monotone quantities. Recall from (14.63) that, under the normalized Ricci flow on any closed surface M^2, the quantity $h = \overline{R} + \|\nabla f\|^2$, where $\overline{R} = R - r$, satisfies the evolution equation

$$\left(\frac{\partial}{\partial t} - \Delta\right) h = -2\left\|-\frac{1}{2}\overline{R}g + \nabla^2 f\right\|^2 + rh \leq rh. \tag{14.91}$$

This implies that

$$\left(\frac{\partial}{\partial t} - \Delta\right)(\mathrm{e}^{-rt}h) = -2\mathrm{e}^{-rt}\left\|-\frac{1}{2}\overline{R}g + \nabla^2 f\right\|^2 \leq 0.$$

So we have that $\mathrm{e}^{-rt}h$ is a monotone quantity in the sense that it is a subsolution to the heat equation and hence its spatial maximum is a nonincreasing function of time.

14.12.1.1. *The trace-free part* β *of the Hessian of the potential function* f. Motivated by Hamilton's idea that quantities that arise in the evolution equations of monotone quantities may also behave nicely under the normalized Ricci flow, one considers the symmetric 2-tensor

$$\beta := -\frac{1}{2}\overline{R}g + \nabla^2 f, \tag{14.92}$$

which by (14.91) has the property that

$$\left(\frac{\partial}{\partial t} - \Delta\right) h = -2\|\beta\|^2 + rh.$$

We also note that β is trace-free, that is:

$$\mathrm{trace}_g(\beta) = -\overline{R} + \Delta f = 0. \tag{14.93}$$

14.12.1.2. *Characterizing when β vanishes.* Observe that $\overline{R}$ vanishes if and only if $R \equiv r$, that is, g has constant curvature. Note also that if $\overline{R}$ vanishes, then f is constant, so that then β also vanishes.

Lemma 14.22. *Conversely, if β vanishes for some closed oriented Riemannian surface (M^2, g), then g has constant curvature.*

Proof. Suppose that $\beta = 0$. Then

$$\mathcal{L}_{\nabla f} g = 2\nabla^2 f = \overline{R} g. \tag{14.94}$$

Case 1: $\chi \leq 0$. Here we have that $r \leq 0$, which is a condition we will take advantage of. Taking the divergence of (14.94), we have

$$\begin{aligned} dR &= \operatorname{div}(\overline{R} g) \\ &= 2 \operatorname{div}(\nabla^2 f) \\ &= 2d(\Delta f) + 2 \operatorname{Ric}(df) \\ &= 2dR + R df. \end{aligned}$$

Therefore,

$$dR + R df = 0. \tag{14.95}$$

Taking a second divergence yields

$$\begin{aligned} 0 &= \Delta R + dR \cdot df + R \Delta f \\ &= \Delta R + dR \cdot df + R\overline{R} \\ &= \Delta \overline{R} + d\overline{R} \cdot df + \overline{R}^2 + r\overline{R}. \end{aligned}$$

Since $\overline{R}$ is a smooth function, and hence is continuous, and since M^2 is compact, there exists a point $x_0 \in M^2$ at which $\overline{R}$ attains its minimum: $\overline{R}(x_0) = \min_{x \in M^2} \overline{R}(x)$. We have

$$\Delta \overline{R}(x_0) \geq 0, \quad d\overline{R}(x_0) = \vec{0}.$$

Therefore,

$$\overline{R}(x_0)^2 + r\overline{R}(x_0) \leq 0.$$

Since $r \leq 0$, if $\overline{R}(x_0) < 0$, then $\overline{R}(x_0)^2 > 0$ and $r\overline{R}(x_0) \geq 0$ and thus we have a contradiction. Therefore, $\overline{R}(x_0) \geq 0$. Finally, since $\int_{M^2} \overline{R} d\mu = 0$, we conclude that $\overline{R} \equiv 0$ on all of M^2.

Case 2: $\chi > 0$. In this case, by the classification of surfaces (Theorem 8.11), we have that M^2 is diffeomorphic to the 2-sphere S^2.

By (14.94) and (12.26), we have that ∇f is a conformal vector field. Hence we may apply the Kazdan–Warner identity, i.e., Theorem 12.7, to obtain

$$0 = \int_{M^2} \langle \nabla_g R, \nabla f \rangle_g \, d\mu_g.$$

Integrating by parts, we obtain

$$0 = -\int_{M^2} R\Delta_g f \, d\mu_g = -\int_{M^2} \overline{R}\Delta_g f \, d\mu_g = -\int_{M^2} \overline{R}^2 d\mu_g.$$

We again conclude that $\overline{R} \equiv 0$ on M^2. □

14.12.1.3. *The evolution of β and its norm.* Since we know the evolution equations for $\overline{R}$, g, and f, we can compute the evolution of β. One catch is that we also have to calculate the evolution of the Hessian operator ∇^2 since it depends on $g(t)$. In any case, one arrives at the following formula:

$$\left(\frac{\partial}{\partial t} - \Delta\right)\beta = (r - 2R)\beta. \tag{14.96}$$

We refer to the read to [CK04] for an exposition of the details of this calculation of Hamilton.

In general, for any symmetric 2-tensor $\gamma(t)$, under the normalized Ricci flow on surfaces we have

$$\left(\frac{\partial}{\partial t} - \Delta\right)\|\gamma(t)\|^2_{g(t)} = -2\|\nabla\gamma\|^2 + 2\overline{R}\|\gamma\|^2 + 2\left(\frac{\partial}{\partial t} - \Delta\right)\gamma \cdot \gamma. \tag{14.97}$$

Hence, we obtain from (14.96) that

$$\left(\frac{\partial}{\partial t} - \Delta\right)\|\beta\|^2 = -2\|\nabla\beta\|^2 - 2R\|\beta\|^2. \tag{14.98}$$

14.12.1.4. *For any metric on S^2, a conformally equivalent metric has positive curvature.* Now assume that M^2 is diffeomorphic to the 2-sphere. Let g_0 be a Riemannian metric on M^2. Recall from (8.46) that if $g_1 = \mathrm{e}^{2u}g_0$, then

$$R_1 = \mathrm{e}^{-2u}\big(R_0 - 2\Delta_0 u\big).$$

Let r be the average scalar curvature of g_0. Recall by Corollary 11.19, which is a consequence of the Hodge theorem, that since $\int_M (R_0 - r) d\mu_0 = 0$, there exists a function $u : M^2 \to \mathbb{R}$ satisfying the Poisson equation

$$2\Delta_0 u = R_0 - r.$$

For this choice of u, we have

$$R_1 = \mathrm{e}^{-2u} r > 0.$$

14.12.1.5. *A uniform lower bound for the scalar curvature.* We consider the normalized Ricci flow $g(t)$ starting from the metric g_1. Using the techniques in §14.11 (which are for the case where $\chi(M^2) < 0$), one can show for the case where $\chi(M^2) > 0$ that $g(t)$ exists for all time $t \in [0, \infty)$. By the parabolic maximum principle applied to the equation (14.29), we have that

$$R(t) > 0$$

for all $t \geq 0$. Hamilton proved the following *apriori estimate*.

Proposition 14.23. Under the normalized Ricci flow on a closed surface with positive curvature, there exists a constant c such that

$$R(x,t) \geq c > 0 \tag{14.99}$$

for all $x \in M^2$ and $t \in [0,\infty)$.

We will finish the proof of this proposition in §14.13.6.4.

The point of the proposition is that the positive lower bound for the scalar curvature is *uniform*. That is, the proposition precludes the scalar curvature from decaying to zero as time tends to infinity. Another important significance of this estimate is that by (14.98) it implies that

$$\left(\frac{\partial}{\partial t} - \Delta\right) \|\beta\|^2 \leq -2c\|\beta\|^2.$$

Hence, by the parabolic maximum principle (Exercise 14.2(2)), there exists a constant C such that

$$\|\beta\|^2(x,t) \leq C\mathrm{e}^{-2ct} \tag{14.100}$$

for all $x \in M^2$ and $t \in [0,\infty)$, where $c > 0$.

14.12.2. The modified Ricci flow. In Riemannian geometry, isometric metrics are considered to be geometrically the same. So we first discuss the effect of pulling back by diffeomorphisms on a 1-parameter family of metrics. Let $\varphi_t : M^2 \to M^2$, $t \in [0,\infty)$, be a 1-parameter family of diffeomorphisms. Let V_t be the **1-parameter family of vector fields generated by** φ_t, that is, by definition,

$$\frac{\partial}{\partial t}\varphi_t(x) =: V_t(\varphi_t(x)) = (V_t \circ \varphi_t)(x).$$

Let $g(t)$, $t \in [0,\infty)$, be a solution to the normalized Ricci flow on M^2. The 1-parameter family of pullback metrics

$$\widetilde{g}(t) := \varphi_t^* g(t) \tag{14.101}$$

are by definition given by (see §6.6.1)

$$\widetilde{g}(t)(V,W) = g(t)\big(d\varphi_t(V), d\varphi_t(W)\big). \tag{14.102}$$

Also by definition, $\widetilde{g}(t)$ is isometric to $g(t)$. Thus, geometrically, the family $\widetilde{g}(t)$ is indistinguishable from $g(t)$.

Using the product rule and the consequence of the definition of the Lie derivative (6.91), we compute that

$$\begin{aligned}\frac{\partial}{\partial t}\widetilde{g}(t) &= \frac{\partial}{\partial t}\big(\varphi_t^* g(t)\big)\\ &= \varphi_t^*\left(\frac{\partial}{\partial t}g(t)\right) + \mathcal{L}_{d(\varphi_t^{-1})(\partial_t \varphi_t)}\widetilde{g}(t)\\ &= -\overline{R}_{\widetilde{g}(t)}\widetilde{g}(t) + \mathcal{L}_{d(\varphi_t^{-1})(V_t)}\widetilde{g}(t),\end{aligned} \tag{14.103}$$

where we used that $\varphi_t^*(\overline{R}_{g(t)}g(t)) = \overline{R}_{\widetilde{g}(t)}\widetilde{g}(t)$.

Define

$$\widetilde{f}(t) := f(t)\circ\varphi_t. \tag{14.104}$$

From now on, we choose the diffeomorphisms φ_t to be defined by

$$V_t = \nabla f(t) \quad \text{and} \quad \varphi_0 = \mathrm{id}_{M^2}, \tag{14.105}$$

so that

$$\frac{\partial}{\partial t}\varphi_t(x) = (\nabla f(t)\circ\varphi_t)(x). \tag{14.106}$$

Let $\widetilde{\nabla}$ denote the gradient with respect to $\widetilde{g}(t)$. By (14.103), we then have

$$\begin{aligned}\frac{\partial}{\partial t}\widetilde{g}(t) &= -\overline{R}_{\widetilde{g}(t)}\widetilde{g}(t) + \mathcal{L}_{d(\varphi_t^{-1})(\nabla f(t))}\widetilde{g}(t)\\ &= -\overline{R}_{\widetilde{g}(t)}\widetilde{g}(t) + \mathcal{L}_{\widetilde{\nabla}\widetilde{f}(t)}\widetilde{g}(t)\\ &= -\overline{R}_{\widetilde{g}(t)}\widetilde{g}(t) + 2\widetilde{\nabla}^2\widetilde{f}(t)\\ &=: 2\widetilde{\beta}(t),\end{aligned} \tag{14.107}$$

where to obtain the second equality we used that

$$\widetilde{\nabla}\widetilde{f}(t) = \nabla_{\varphi_t^* g(t)}(f(t)\circ\varphi_t) = \varphi_t^*(\nabla_{g(t)}f(t)) = d(\varphi_t^{-1})(\nabla f(t)). \tag{14.108}$$

We calculate that

$$\Delta_{\widetilde{g}(t)}\widetilde{f}(t) = \Delta_{\varphi_t^* g(t)}(f\circ\varphi_t) = (\Delta_{g(t)}f(t))\circ\varphi_t = (R_{g(t)} - r)\circ\varphi_t.$$

Therefore,

$$\Delta_{\widetilde{g}(t)}\widetilde{f}(t) = \overline{R}_{\widetilde{g}(t)} := R_{\widetilde{g}(t)} - r \tag{14.109}$$

since $r = r\circ\varphi_t$ follows from r being constant and since, by $\widetilde{g}(t) = \varphi_t^* g(t)$, we have $R_{\widetilde{g}(t)} = R_{g(t)}\circ\varphi_t$. Equation (14.109) is analogous to (14.52). Namely, $\widetilde{f}(t)$ is the potential function for $\widetilde{g}(t)$.

Observe that

$$\operatorname{trace}_{\widetilde{g}(t)}\big(\widetilde{\beta}(t)\big) = \operatorname{trace}_{\widetilde{g}(t)}\left(\frac{\partial}{\partial t}\widetilde{g}(t)\right) = -2\overline{R}_{\widetilde{g}(t)} + 2\widetilde{\Delta}\widetilde{f}(t) = 0. \tag{14.110}$$

Therefore, the area form of $\widetilde{g}(t)$ is independent of time under the modified Ricci flow:

$$\frac{\partial}{\partial t} d\mu_{\widetilde{g}(t)} = 0. \tag{14.111}$$

This implies that $\operatorname{Area}(\widetilde{g}(t))$ is constant, which we already know from the area of $g(t)$ being constant and $\widetilde{g}(t) = \varphi_t^* g(t)$.

We now calculate the evolution of the potential function for $\widetilde{g}(t)$:

$$\begin{aligned}
\frac{\partial \widetilde{f}}{\partial t}(t) &= \frac{\partial}{\partial t}(f(t) \circ \varphi_t) \\
&= \frac{\partial f}{\partial t}(t) \circ \varphi_t + df(t)\left(\frac{\partial}{\partial t}\varphi_t\right) \\
&= \left(\Delta_{g(t)} f(t) + r f(t)\right) \circ \varphi_t + df(t)\left(\nabla_{g(t)} f(t) \circ \varphi_t\right) \\
&= \Delta_{\varphi_t^* g(t)}(f(t) \circ \varphi_t) + r f(t) \circ \varphi_t + |\nabla_{g(t)} f(t)|^2 \circ \varphi_t.
\end{aligned}$$

That is,

$$\frac{\partial \widetilde{f}}{\partial t}(t) = \Delta_{\widetilde{g}(t)} \widetilde{f}(t) + \|\widetilde{\nabla} \widetilde{f}(t)\|^2_{\widetilde{g}(t)} + r \widetilde{f}(t) \tag{14.112}$$

This is analogous to (14.54), except that we have a gradient term. Observe that this gradient term may be rewritten as $\|\widetilde{\nabla} \widetilde{f}(t)\|^2_{\widetilde{g}(t)} = \mathcal{L}_{\widetilde{\nabla} \widetilde{f}(t)} \widetilde{f}$.

14.12.3. Convergence to constant curvature for the normalized Ricci flow on S^2. We can now begin to finish off the amazing proof of Hamilton. The long-time existence of the solution of the normalized Ricci flow on S^2 holds for the following reasons. Firstly, by (14.68), we have that $|R(x,t) - r| \leq C\mathrm{e}^{rt}$ for all $x \in M^2$ and $t \in [0,T)$ (Proposition 14.23 gives a much better lower bound for the scalar curvature). Secondly, by using this and similarly to Lemma 14.15, we can obtain time-dependent estimates for all derivatives of the curvature. Thirdly, similarly to §14.11, we can deduce from this that a unique solution $g(t)$ to the normalized Ricci flow on S^2 exists for all time $t \in [0, \infty)$.

Now recall from (14.100) and (14.92) that

$$\left\| -\frac{1}{2}\overline{R} g + \nabla^2 f \right\|^2_{g(t)} (x,t) \leq C\mathrm{e}^{-2ct}. \tag{14.113}$$

On the other hand, by (14.107),

$$2\widetilde{\beta}(t) = -\overline{R}_{\widetilde{g}(t)} \widetilde{g}(t) + 2\widetilde{\nabla}^2 \widetilde{f}(t) = \left(-\frac{1}{2}\overline{R} g + \nabla^2 f\right) \circ \varphi_t.$$

Therefore,

$$\left\| \frac{\partial}{\partial t}\widetilde{g}(t) \right\|^2_{\widetilde{g}(t)} = \left\| -\overline{R}_{\tilde{g}(t)}\widetilde{g}(t) + 2\widetilde{\nabla}^2\widetilde{f}(t) \right\|^2_{\widetilde{g}(t)}(x,t) \le C\mathrm{e}^{-2ct}. \tag{14.114}$$

One can show, analogously to Lemma 14.15, that for each positive integer k there exists a constant C_k such that

$$\|\widetilde{\nabla}^k \widetilde{\beta}(t)\|_{\widetilde{g}(t)} \le C_k. \tag{14.115}$$

Similarly to §14.11, we can deduce from this that the solution $\widetilde{g}(t)$ to the modified Ricci flow exists for all time $t \in [0, \infty)$ and that the metrics $\widetilde{g}(t)$ converge as $t \to \infty$ to a smooth Riemannian metric $\widetilde{g}_\infty$. Furthermore, this metric satisfies

$$\widetilde{\beta}_\infty := -\frac{1}{2}\overline{R}_{\tilde{g}_\infty}\widetilde{g}_\infty + \widetilde{\nabla}^2 \widetilde{f}_\infty = 0, \tag{14.116}$$

where $\widetilde{f}_\infty$ satisfies

$$\Delta_{\tilde{g}_\infty}\widetilde{f}_\infty = R_{\tilde{g}_\infty} - r. \tag{14.117}$$

Now, (14.116) implies that the vector field $\widetilde{\nabla}\widetilde{f}_\infty$ is a conformal vector field with respect to the metric $\widetilde{g}_\infty$. Thus we may apply the Kazdan–Warner identity (Theorem 12.7) to obtain

$$\begin{aligned} 0 &= \int_{M^2} \langle \widetilde{\nabla} R_{\tilde{g}_\infty}, \widetilde{\nabla}\widetilde{f}_\infty \rangle_{\tilde{g}_\infty} d\mu_{\tilde{g}_\infty} \\ &= \int_{M^2} R_{\tilde{g}_\infty} \Delta_{\tilde{g}_\infty}\widetilde{f}_\infty d\mu_{\tilde{g}_\infty} \\ &= \int_{M^2} R_{\tilde{g}_\infty}(R_{\tilde{g}_\infty} - r) d\mu_{\tilde{g}_\infty} \\ &= \int_{M^2} (R_{\tilde{g}_\infty} - r)^2 d\mu_{\tilde{g}_\infty}. \end{aligned}$$

We conclude that

$$R_{\tilde{g}_\infty} \equiv r. \tag{14.118}$$

Moreover, since the convergence of $\widetilde{g}(t)$ to $\tilde{g}_\infty$ is exponentially fast in each C^k norm, we have that $R_{\tilde{g}(t)}$ converges to r exponentially fast under the modified Ricci flow. We also have that $\|\widetilde{\nabla}^k R_{\tilde{g}(t)}\|_{\tilde{g}(t)}$ decays exponentially to 0 as $t \to \infty$ for each positive integer k. Since the solution $g(t)$ satisfies $R_{\tilde{g}(t)} = R_{g(t)} \circ$ and $\widetilde{\nabla}^k R_{\tilde{g}(t)} = \varphi_t^* \nabla^k_{g(t)} R_{g(t)}$, we have that $R_{g(t)}$ converges to r exponentially fast and each $\|\widetilde{\nabla}^k R_{\tilde{g}(t)}\|$ decays exponentially to 0 as $t \to \infty$. Therefore, the solution $g(t)$ to the normalized Ricci flow converges exponentially fast in each C^k norm to a smooth Riemannian metric g_∞. Since $R_{g(t)}$ converges to r, we conclude that $R_{g_\infty} \equiv r$.

14.13. The entropy and Harnack estimates

In this section we discuss the entropy and Harnack estimates that are used in the proof of the key estimate in Proposition 14.23, which says that the scalar curvature under the normalized Ricci flow is uniformly bounded from below by a positive constant.

14.13.1. The general idea of entropy.

The idea of *entropy* is important in thermodynamics, statistical mechanics, information theory, probability theory, and partial differential equations.

Let n be a positive integer and suppose that $\mathbf{p} := \{p_1, \ldots, p_n\}$ is a (discrete) probability distribution of a set of n elements; that is, $\sum_{i=1}^{n} p_i = 1$. Then the **entropy** of this probability distribution is defined to be equal to

$$N(\mathbf{p}) := -\sum_{i=1}^{n} p_i \ln(p_i). \tag{14.119}$$

14.13.2. Entropy for the heat equation.

Let (M^n, g) be a closed Riemannian manifold and let $f : M^n \to \mathbb{R}$ be a positive function with $\int_{M^n} f d\mu = 1$. The **relative entropy** of the probability distribution $f d\mu$ is defined as

$$N(f) := -\int_{M^n} f \ln(f) \, d\mu. \tag{14.120}$$

Now suppose that $f(t) : M^n \to \mathbb{R}$ is a solution to the **heat equation**

$$\frac{\partial f}{\partial t} = \Delta f. \tag{14.121}$$

We compute that

$$\begin{aligned} \frac{dN}{dt} &= -\int_{M^n} \left(\ln(f) \frac{\partial f}{\partial t} + f \frac{\partial}{\partial t} \ln(f) \right) d\mu \\ &= -\int_{M^n} \left(\ln(f) \Delta f + \Delta f \right) d\mu \\ &= \int_{M^n} \frac{\|\nabla f\|^2}{f} d\mu \\ &\geq 0, \end{aligned}$$

where we integrated by parts and used the divergence theorem. Thus, the entropy of a solution to the heat equation is a non-decreasing function of time.

14.13.3. Entropy in comparison to L^p-norms. For any real number $p > 1$, we have

$$\int_{M^n} f \ln f \, d\mu \leq 2 \left(\int_{M^n} |f-1|^p d\mu \right)^{1/p} + \frac{2}{p-1} \int_{M^n} |f-1|^p d\mu. \tag{14.122}$$

Now recall that the L^p-norm of a function $f : M^n \to \mathbb{R}$ is defined by

$$\|f\|_p := \left(\int_{M^n} |f|^p \, d\mu \right)^{1/p}. \tag{14.123}$$

Hölder's inequality says that for any $\alpha, \beta \in [1, \infty]$ with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$,[3] we have

$$\|fg\|_1 \leq \|f\|_\alpha \|g\|_\beta \tag{14.124}$$

for any functions f and g.

Now suppose that $\int_{M^n} d\mu = 1$ and let $q > p > 1$. By Hölder's inequality with $\alpha = \frac{q}{p}$, we have

$$\begin{aligned} (\|f\|_p)^p = \| |f|^p \|_1 = \| |f|^p \cdot 1 \|_1 &\leq \| |f|^p \|_\alpha \|1\|_\beta \\ = \| |f|^p \|_{\frac{q}{p}} &= \left(\int_{M^n} |f|^q d\mu \right)^{p/q} = \|f\|_q^p. \end{aligned}$$

Hence,

$$\|f\|_p \leq \|f\|_q. \tag{14.125}$$

So, for $q > p > 1$, the L^q-norm is "stronger" than the L^p-norm in the sense that $\|f\|_q \leq C$ tells you more than $\|f\|_p \leq C$.

Now, by (14.122), the entropy satisfies

$$\int_{M^n} f \ln f \, d\mu \leq 2\|f-1\|_p + \frac{2}{p-1} \|f-1\|_p^p. \tag{14.126}$$

So, the L^p-distance $\|f-1\|_p$ between f and the constant function 1 controls the entropy of f.

We will now see that the idea of entropy is useful in Ricci flow.

14.13.4. Hamilton's entropy estimate. Let (M^2, g) be a closed Riemannian surface. If g has positive curvature, then we can define **Hamilton's surface entropy** by

$$N(g) := \int_{M^2} R \ln R \, d\mu. \tag{14.127}$$

(This is the opposite of the usual sign convention for entropy, so we want to show that Hamilton's entropy decreases.)

[3] We use the convention that $\frac{1}{\infty} := 0$.

Let $(M^2, g(t))$ be a Ricci flow on a closed surface with positive curvature. The **surface entropy monotonicity formula** is:

(14.128)
$$\begin{aligned}\frac{d}{dt}N(g(t)) &= -2\int_{M^2}\left\|-\frac{1}{2}\overline{R}g + \nabla^2 f\right\|^2 d\mu - \int_{M^2}\frac{\|\nabla R + R\nabla f\|^2}{R}d\mu\\ &= -2\int_{M^2}\|\beta\|^2 d\mu - 4\int_{M^2}\frac{\|\operatorname{div}(\beta)\|^2}{R}d\mu\\ &\leq 0.\end{aligned}$$

This implies Hamilton's result that his surface entropy is monotonically non-increasing.

How did we obtain this monotonicity formula? The second equality in (14.128) follows from the definition of β and the calculations:

$$\operatorname{div}(\beta) = -\frac{1}{2}\nabla R + \operatorname{div}(\nabla^2 f)$$

and

$$\begin{aligned}\operatorname{div}(\nabla^2 f) &= \sum_{i=1}\nabla^3 f(e_i, \cdot, e_i)\\ &= \sum_{i=1}\nabla^3 f(\cdot, e_i, e_i) + \operatorname{Ric}(\nabla f)\\ &= \nabla(\Delta f) + \frac{1}{2}R\nabla f\\ &= \nabla R + \frac{1}{2}R\nabla f.\end{aligned}$$

The first equality in (14.128) follows from the formula

$$\frac{d}{dt}N(g(t)) = -\int_{M^2}\frac{\|\nabla R\|^2}{R}d\mu + \int_{M^2}\overline{R}^2 d\mu \qquad (14.129)$$

and an integration by parts. To see (14.129), we calculate as follows. Recall by (14.18) that

$$\frac{\partial}{\partial t}d\mu = -\overline{R}d\mu.$$

By combining this with (14.29), we obtain

$$\frac{\partial}{\partial t}(Rd\mu) = \frac{\partial R}{\partial t}d\mu + R\frac{\partial}{\partial t}d\mu = \Delta R d\mu. \qquad (14.130)$$

Note that a consistency check for this formula is that as a consequence we have

$$\frac{d}{dt}\int_{M^2}Rd\mu = \int_{M^2}\Delta R d\mu = 0,$$

where the last equality is by the divergence theorem. Indeed, we already know this from the Gauss–Bonnet formula.

Now, using (14.29) and (14.130), we calculate that

$$\begin{aligned}\frac{d}{dt}N &= \frac{d}{dt}\int_{M^2} \ln R\, R\, d\mu \\ &= \int_{M^2} \frac{\partial}{\partial t}(\ln R)\, R\, d\mu + \int_{M^2} \ln R \frac{\partial}{\partial t}(R\, d\mu) \\ &= \int_{M^2} \frac{1}{R}(\Delta R + R\overline{R})\, R\, d\mu + \int_{M^2} \ln R\, \Delta R\, d\mu.\end{aligned}$$

We can now integrate by parts to obtain

$$\frac{d}{dt}N = -\int_{M^2} \frac{\|\nabla R\|^2}{R} d\mu + \int_{M^2} \overline{R}^2 d\mu,$$

where we also used that $\int_{M^2} R\overline{R}d\mu = \int_{M^2} \overline{R}^2 d\mu$. See e.g. [CK04] for an exposition of the details of how to carry out the integration by parts to obtain (14.128) from (14.129).

14.13.5. Hamilton's Harnack estimate. In the study of the Ricci flow on surfaces, β is a natural quantity. Recall from the previous subsection that

$$2\operatorname{div}(\beta) = \nabla R + R\nabla f. \tag{14.131}$$

By simply taking a second divergence, we obtain

$$\begin{aligned}2\operatorname{div}^2(\beta) &= \operatorname{div}(\nabla R + R\nabla f) \\ &= \Delta R + \langle \nabla R, \nabla f\rangle + R\overline{R},\end{aligned} \tag{14.132}$$

where we used that $\Delta f = \overline{R}$. Now (14.131) implies that

$$-2\frac{\nabla R}{R}\cdot \operatorname{div}(\beta) = -\frac{\|\nabla R\|^2}{R} - \langle \nabla R, \nabla f\rangle.$$

Therefore,

$$\begin{aligned}Q &:= \frac{2}{R}\operatorname{div}^2(\beta) - 2\frac{\nabla R}{R^2}\cdot \operatorname{div}(\beta) \\ &= \frac{\Delta R}{R} - \frac{\|\nabla R\|^2}{R^2} + \overline{R} \\ &= \Delta \ln R + R - r.\end{aligned} \tag{14.133}$$

The quantity Q is called **Hamilton's Harnack quantity**. As we will see in the next section, Q vanishes on self-similar solutions to the Ricci flow, called *Ricci solitons* (as we will see, β vanishes on Ricci solitons). This is one motivation for considering Q as a natural quantity for which to compute the evolution equation.

One can show the estimate

$$Q(x,t) \geq -\frac{Cre^{rt}}{Ce^{rt}-1} =: q(t), \tag{14.134}$$

where $C > 1$ is a constant depending only on g_0. Note that the function $q(t)$ is increasing. In particular, we have that if $t \geq 1$, then

$$Q(x,t) \geq -\frac{Cr}{C-1} =: -C' \tag{14.135}$$

for all $x \in M^2$. This is called the **Harnack estimate** for the Ricci flow on surfaces.

The proof of (14.134) is simply to derive the following heat-type inequality:

$$\frac{\partial Q}{\partial t} \geq \Delta Q + 2\langle \nabla \ln R, \nabla Q \rangle + Q^2 + rQ. \tag{14.136}$$

By taking $C := \frac{q_0}{q_0 + r} > 1$, where $q_0 := \min Q(\cdot, 0)$, we have that $q(t)$ satisfies the ODE $\frac{dq}{dt} = q^2 + rq$ with $q(0) = \frac{Cr}{C-1} = q_0$. Now, applying the parabolic maximum principle to (14.136) yields the Harnack estimate (14.134).

Now, let us see why the Harnack estimate for the Ricci flow on surfaces is useful.

Using (14.29), we calculate that

$$\begin{aligned} \frac{\partial}{\partial t}\ln R &= \frac{1}{R}\frac{\partial R}{\partial t} = \frac{1}{R}(\Delta R + R^2 - rR) \\ &= \Delta \ln R + \|\nabla \ln R\|^2 + R - r. \end{aligned} \tag{14.137}$$

Therefore, the Harnack quantity Q defined by (14.133) may be re-expressed as the space-time gradient quantity

$$Q = \frac{\partial}{\partial t}\ln R - \|\nabla \ln R\|^2. \tag{14.138}$$

Thus, the Harnack estimate (14.135) says that

$$\frac{\partial}{\partial t}\ln R - \|\nabla \ln R\|^2 \geq -C' \tag{14.139}$$

for some constant C', provided $t \geq 1$.

In order to compare the curvatures of the solution at two different points (x_1, t_1) and (x_2, t_2) in space-time, we will integrate the differential expression Q along paths in space time. For this purpose, let

$$\gamma : [t_1, t_2] \to M^2$$

be a path with $\gamma(t_1) = x_1$ and $\gamma(t_2) = x_2$. Consider the associated space-time path

$$\widetilde{\gamma} : [t_1, t_2] \to M^2 \times [t_1, t_2]$$

defined by

$$\widetilde{\gamma}(t) := (\gamma(t), t). \tag{14.140}$$

Observe that $\widetilde{\gamma}(t_1) = (x_1, t_1)$ and $\widetilde{\gamma}(t_2) = (x_2, t_2)$.

We now apply the Fundamental Theorem of Calculus to the one-variable function $\ln R$ along $\widetilde{\gamma}$ to obtain

$$\begin{aligned}\ln R(x_2,t_2) - \ln R(x_1,t_1) \\ &= \int_{t_1}^{t_2} \frac{d}{dt}(\ln R(\gamma(t),t)\,dt \\ &= \int_{t_1}^{t_2} \left(\nabla \ln R(\gamma(t),t)\cdot\gamma'(t) + \frac{\partial \ln R}{\partial t}(\gamma(t),t)\right) dt,\end{aligned} \tag{14.141}$$

where the dot product $\cdot$ denotes the inner product with respect to the metric $g(t)$, also denoted by $\langle\cdot,\cdot\rangle$. By applying the Harnack estimate (14.139) to this, we obtain

$$\begin{aligned}\ln\frac{R(x_2,t_2)}{R(x_1,t_1)} &\geq \int_{t_1}^{t_2} \left(\nabla \ln R(\gamma(t),t)\cdot\gamma'(t) + \|\nabla \ln R\|^2(\gamma(t),t) - C'\right) dt \\ &\geq -\int_{t_1}^{t_2} \frac{1}{4}\|\gamma'(t)\|^2_{g(t)} dt - C'(t_2-t_1),\end{aligned}$$

where to obtain the last inequality we used the elementary (Peter–Paul) inequality $-ab+b^2 \geq -\frac{1}{4}a^2$ and that

$$\nabla \ln R(\gamma(t),t)\cdot\gamma'(t) \geq -\|\nabla \ln R(\gamma(t),t)\|_{g(t)}\,\|\gamma'(t)\|_{g(t)}.$$

We have proved the following:

Proposition 14.24. Let $(M^2, g(t))$ be a solution to the normalized Ricci flow on surfaces with positive curvature. Let $x_1, x_2 \in M^2$ and $t_1 < t_2$. Then for any path $\gamma : [t_1,t_2] \to M^2$ with $\gamma(t_1) = x_1$ and $\gamma(t_2) = x_2$, we have

$$\frac{R(x_2,t_2)}{R(x_1,t_1)} \geq \mathrm{e}^{-C'(t_2-t_1)} \exp\left(-\int_{t_1}^{t_2} \frac{1}{4}\|\gamma'(t)\|^2_{g(t)} dt\right). \tag{14.142}$$

To get the best estimate from (14.142), on the right-hand side we should take the supremum over all such paths γ. Since, in general we cannot compute the supremum, we will be satisfied with a rough lower estimate of the right-hand side which indeed will suffice for our purposes.

14.13.6. The uniform estimate for the scalar curvature.

14.13.6. The uniform estimate for the scalar curvature. We now proceed to obtain a uniform estimate for the scalar curvature R under the normalized Ricci flow.

14.13.6.1. *Uniform equivalence of the metrics on short time intervals.* Let t_1 be any positive time. Let $x_1 \in M^2$ be a point at which $R(\cdot,t_1)$ attains its maximum. Let $K_1 := \max_{M^2} R(\cdot,t_1) = R(x_1,t_1)$. We have

$$\frac{d}{dt}R_{\max}(t) \leq (R_{\max}(t))^2 - rR_{\max}(t) \leq (R_{\max}(t))^2.$$

The solution to the ODE $\frac{dk}{dt} = k^2$ with the initial condition $k(t_1) = K_1$ is given by

$$k(t) = \frac{1}{\frac{1}{K_1} + t_1 - t}$$

for $t \in [t_1, t_1 + \frac{1}{K_1}]$. Therefore, by the parabolic maximum principle, we have that

$$R(x,t) \leq 2K_1 \tag{14.143}$$

for all $x \in M^2$ and $t \in [t_1, t_1 + \frac{1}{2K_1}]$.

Let $t_2 = t_1 + \frac{1}{2K_1}$. Let $\bar{t} \in [t_1, t_2]$. We have for any $x \in M^2$,

$$\begin{aligned} g(x,t_2) &= \exp\left(\int_{\bar{t}}^{t_2} (r - R(x,t))\, dt\right) g(x,\bar{t}) \\ &\geq \exp\left(\int_{\bar{t}}^{t_2} (r - 2K_1)\, dt\right) g(x,\bar{t}) \\ &\geq \exp\left(\frac{r}{2K_1} - 1\right) g(x,\bar{t}) \\ &\geq \mathrm{e}^{-1} g(x,\bar{t}) \end{aligned}$$

for all $\bar{t} \in [t_1, t_2]$. That is, we have:

Lemma 14.25. *For any normalized Ricci flow, we have*

$$g(x,t) \leq \mathrm{e}\, g(x,t_2) \tag{14.144}$$

for all $x \in M^2$ and $t \in [t_1, t_2]$, where $t_2 = t_1 + \frac{1}{2K_1}$.

14.13.6.2. *Smoothing property of the curvature function.* Let x_2 be a point in M^2 and let $t_2 = t_1 + \frac{1}{2K_1}$, where (x_1, t_1) is as in the previous subsection so that $R(x_1, t_1) = \max_{M^2} R(\cdot, t_1) = K_1$. Let $\gamma : [t_1, t_2] \to M^2$ be a constant-speed minimal geodesic with respect to the metric $g(t_2)$ joining the point x_1 to the point x_2. Then

$$\|\gamma'(t)\|_{g(t_2)} = \frac{d_{g(t_2)}(x_1, x_2)}{t_2 - t_1}.$$

Further assume that $K \geq 1$. By Proposition 14.24, we have

$$\begin{aligned} \frac{R(x_2,t_2)}{R(x_1,t_1)} &\geq \mathrm{e}^{-C'(t_2 - t_1)} \exp\left(-\int_{t_1}^{t_2} \frac{1}{4} \|\gamma'(t)\|^2_{g(t)} dt\right) \\ &\geq \mathrm{e}^{-\frac{C'}{2K}} \exp\left(-\mathrm{e} \int_{t_1}^{t_2} \frac{1}{4} \|\gamma'(t)\|^2_{g(t_2)} dt\right) \\ &= \mathrm{e}^{-\frac{C'}{2K}} \exp\left(-\frac{\mathrm{e}}{4} \frac{d^2_{g(t_2)}(x_1, x_2)}{t_2 - t_1}\right). \end{aligned} \tag{14.145}$$

Recall that $t_2 - t_1 = \frac{1}{2K_1}$. Thus, if we assume that $d_{g(t_2)}(x_1, x_2) \leq \frac{1}{\sqrt{K_1}}$, then

$$\frac{R(x_2,t_2)}{R(x_1,t_1)} \geq \mathrm{e}^{-\frac{C'}{2K_1}}\mathrm{e}^{-\frac{\mathrm{e}}{2}} \geq \mathrm{e}^{-\frac{C'+\mathrm{e}}{2}} \tag{14.146}$$

where the last inequality is since $K_1 \geq 1$. Since $R(x_1, t_1) = K_1$, we obtain:

Lemma 14.26.

$$R(x_2, t_2) \geq \mathrm{e}^{-\frac{C'+\mathrm{e}}{2}} K_1 \tag{14.147}$$

for all $x_2 \in B^{g(t_2)}_{1/\sqrt{K_1}}(x_1)$.

This lemma reflects the smoothing property of the curvature function. Namely, if the curvature is large at a point (x_1, t_1), then the curvature is large in a small ball centered at that point at a slightly later time.

14.13.6.3. *Combining the entropy and differential Harnack estimates.* We are now in a position to combine the entropy and differential Harnack estimates to obtain the uniform bound for the scalar curvature.

Lemma 14.27. *There exists a universal constant $c > 0$ such that*

$$\mathrm{Area}(B^{g(t_2)}_{1/\sqrt{K_1}}(x_1)) \geq \frac{c}{K_1}. \tag{14.148}$$

That is, with respect to $g(t_2)$, the ball of radius $\rho := \frac{1}{\sqrt{K_1}}$ centered at x_1 has area at least $c\rho^2$.

Now recall that the monotonicity of the surface entropy says that there exists a constant (depending only on the initial metric g_0) such that

$$N(g(t)) = \int_{M^2} R \ln R \, d\mu \leq C$$

for all $t \in [0, \infty)$. On the other hand, recall the elementary inequality that for any $u \in (0, \infty)$, $u \ln u \geq -\frac{1}{\mathrm{e}}$. Thus we have that, where $B^{t_2} := B^{g(t_2)}$,

$$\begin{aligned}\int_{B^{t_2}_\rho(x_1)} R \ln R \, d\mu(t_2) &= \int_{M^2} R \ln R \, d\mu(t_2) - \int_{M^2 \setminus B^{t_2}_\rho(x_1)} R \ln R \, d\mu(t_2) \\ &\leq C + \frac{\mathrm{Area}(g(t_2))}{\mathrm{e}},\end{aligned}$$

where the right-hand side is constant depending only on g_0. By applying Lemmas 14.26 and 14.27, we obtain that

$$\begin{aligned}C + \frac{\mathrm{Area}(g_0)}{\mathrm{e}} &\geq \mathrm{Area}(B^{t_2}_\rho(x_1))\mathrm{e}^{-\frac{C'+\mathrm{e}}{2}} K_1 \ln\left(\mathrm{e}^{-\frac{C'+\mathrm{e}}{2}} K_1\right) \\ &\geq c\mathrm{e}^{-\frac{C'+\mathrm{e}}{2}} \ln\left(\mathrm{e}^{-\frac{C'+\mathrm{e}}{2}} K_1\right).\end{aligned}$$

This implies that for any time t_1, $\max_{t_1} R(\cdot, t_1) \leq K_1$ is bounded by a constant depending only on g_0. Since t_1 is arbitrary, this implies that the scalar curvature of the solution to the normalized Ricci flow is uniformly bounded.

14.13.6.4. *The uniform positive lower bound for the scalar curvature.* Since the metrics $g(t)$, $t \in [0, \infty)$, all have positive and uniformly bounded curvature and constant area, we have that the diameters of $g(t)$ are uniformly bounded from above (see e.g. Corollary 5.52 in [CK04]). We claim that we can thus use the Harnack estimate again to obtain a uniform positive lower bound for the scalar curvatures of $g(t)$. This will complete the proof of Proposition 14.23 and hence also of Theorem 14.21 in the $\chi > 0$ case.

Proof of the lower bound. Let C be such that $R(x,t) \leq C$ for all $x \in M^2$ and $t \in [0, \infty)$ and $\operatorname{diam}(g(t)) \leq C$ for all $t \in [0, \infty)$. Let (x_2, t_2) be a point with $t_2 \geq 1$. Let $t_1 := t_2 - 1$ and let x_1 be a point at which

$$R(x_1, t_1) = r;$$

such a point always exists since r is equal to the average of R at time t_1. By the same argument as to obtain (14.145), we have

$$\frac{R(x_2,t_2)}{R(x_1,t_1)} \geq \mathrm{e}^{-\frac{C'}{2C}} \exp\left(-\frac{\mathrm{e}}{4}\frac{d^2_{g(t_2)}(x_1,x_2)}{t_2 - t_1}\right).$$

By applying the uniform diameter bound to this inequality, we obtain

$$R(x_2,t_2) \geq r\mathrm{e}^{-\frac{C'}{2C}} \exp\left(-\frac{\mathrm{e}C^2}{4}\right). \tag{14.149}$$

This is the desired uniform positive lower bound for the scalar curvature.

14.14. Ricci solitons

One may consider the possibility of a normalized Ricci flow on a surface M^2 that just moves by diffeomorphisms. That is, the possibility that the solution is of the form $g(t) = \phi_t^* g_0$ for some 1-parameter family of diffeomorphisms of M^2. Recall that isometric metrics are geometrically the same. Thus, such a solution is geometrically a fixed point of the normalized Ricci flow. One can think about this abstractly. That is, let $\mathfrak{Met}$ denote the set of Riemannian metrics on M^2. Let $\mathfrak{Diff}$ denote the group of self-diffeomorphisms of M^2. The group $\mathfrak{Diff}$ acts on the set $\mathfrak{Met}$ by pull-back: We have

$$\sigma : \mathfrak{Diff} \times \mathfrak{Met} \to \mathfrak{Met} \tag{14.150}$$

defined by

$$\sigma(\phi, g) := \phi^*(g). \tag{14.151}$$

The quotient space $\mathfrak{Met}/\mathfrak{Diff}$ is the set of isometry classes of Riemannian metrics on M^2. A Ricci flow $g(t)$, $t \in I$, may equivalently be considered as the path $\gamma : I \to \mathfrak{Met}$ defined by $\gamma(t) := g(t)$. Let $\pi : \mathfrak{Met} \to \mathfrak{Met}/\mathfrak{Diff}$ be the canonical projection map. Then

$$\pi \circ \gamma : I \to \mathfrak{Met}/\mathfrak{Diff}$$

maps t to the isometry class of $g(t)$. We see that a Ricci flow $g(t)$ evolves by diffeomorphisms if and only if the associated path $\pi \circ \gamma$ is constant.

14.14.1. Shrinking and steady Ricci solitons. It has been long believed that constant curvature Riemannian metrics are the most natural metrics. Both the uniformization theorem and the Ricci flow version of its proof support this belief. The Ricci flow proof actually first proves convergence of the modified flow to what is called a shrinking Ricci soliton, which we now define.

Definition 14.28. A Riemannian surface (M^2, g) and a function f on M^2 is called a **shrinking Ricci soliton** if

$$\overline{R}g := (R - r)g = 2\nabla^2 f. \tag{14.152}$$

By (7.29), the shrinking Ricci soliton equation (14.152) says that

$$(R - r)g = \mathcal{L}_{\nabla f}\, g. \tag{14.153}$$

We claim that this equation is an infinitesimal version of the condition that a solution $g(t)$ to the normalized Ricci flow is of the form $g(t) = \phi_t^* g_0$ for some 1-parameter family of diffeomorphisms $\{\phi_t\}_{t \in \mathbb{R}}$. To see this, we compute that

$$(R_{g(t)} - r)g(t) = \frac{\partial}{\partial t} g(t) = \frac{\partial}{\partial t}(\phi_t^* g_0) = \mathcal{L}_{d(\phi_t^{-1})(\frac{\partial}{\partial t}\phi_t)}\, g(t).$$

Hence, if

$$d(\phi_t^{-1})\left(\frac{\partial}{\partial t}\phi_t\right) = \nabla_{g(t)} f(t) \tag{14.154}$$

for some function $f(t) : M^2 \to \mathbb{R}$, then we obtain

$$(R_{g(t)} - r)g(t) = \mathcal{L}_{\nabla f(t)}\, g(t). \tag{14.155}$$

Thus, in this case the Riemannian surface $(M^2, g(t))$ with $f(t)$ is a shrinking Ricci soliton. For such solutions to the *normalized* Ricci flow, the metric $g(t)$ is geometrically independent of time and moving only by the pull-back by diffeomorphisms. The reason we call it a *shrinking* Ricci soliton is as follows. Define

$$\tilde{g}(\tilde{t}) := \mathrm{e}^{-r_0 t} g(t) = \mathrm{e}^{-r_0 t} \phi_t^* g_0, \tag{14.156}$$

where $\tilde{t}(t) := \frac{1}{r_0}(1 - \mathrm{e}^{-r_0 t})$. By the discussion at the end of §14.5, we have that $\tilde{g}(\tilde{t})$ is a solution to the unnormalized Ricci flow. These rescaled metrics satisfy the Ricci flow and evolve by diffeomorphisms and scalings. Since $r_0 > 0$ and since $t(\tilde{t}) = -\frac{1}{r_0}\ln(1 - r_0\tilde{t})$ is an increasing function , we have the metrics $\tilde{g}(\tilde{t})$ are shrinking forward in time. This justifies the moniker "shrinking Ricci soliton".

In the previous section, we proved (albeit omitting some key details) that any solution $\widetilde{g}(t)$ to the modified Ricci flow on S^2 converges to a smooth metric $\widetilde{g}_\infty$ which satisfies the equation (14.116):

$$\overline{R}_{\widetilde{g}_\infty}\widetilde{g}_\infty = 2\nabla^2_{\widetilde{g}_\infty}\widetilde{f}_\infty.$$

Thus, we proved for that a flow that is geometrically the same as the normalized Ricci flow (i.e., the solutions of the two equations differ by the pull-back by diffeomorphisms), the solutions converge to shrinking gradient Ricci solitons. We then used the Kazdan–Warner identity to prove that any shrinking Ricci soliton on S^2 must have constant curvature. So we proved that the solution to the *modified* Ricci flow converges to a constant curvature metric. Moreover, we can conclude the same for the *normalized* Ricci flow because of the exponential rate of convergence to constant curvature, including the derivatives of curvature decaying exponentially to 0. That is, the solutions to the normalized Ricci flow converge to constant curvature metrics. This completes the proof of the differential geometric version of the uniformization theorem.

The discussion above begs the question: Are there Ricci solitons that are not constant curvature metrics (so that the potential functions are constants)?

We first consider **steady Ricci solitons**. These are Riemannian surfaces (M^2, g), together with functions $f : M^2 \to \mathbb{R}$, that satisfy the equation (cf. (14.152)):

$$Rg = 2\nabla^2 f. \tag{14.157}$$

14.14.2. Cigar soliton. An iconic example of a *steady* gradient Ricci soliton is the 2-dimensional **cigar soliton**. Its underlying manifold is the plane $\mathbb{R}^2$. Its Riemannian metric is defined by

$$g_\Sigma(x^1, x^2) := \frac{4g_{\mathrm{Euc}}}{1 + (x^1)^2 + (x^2)^2}, \tag{14.158}$$

where $g_{\mathrm{Euc}} = dx^1 \otimes dx^1 + dx^2 \otimes dx^2$ is the Euclidean metric, and its potential function is defined by

$$f_\Sigma\left(x^1, x^2\right) := -\ln\left(1 + (x^1)^2 + (x^2)^2\right). \tag{14.159}$$

See Figure 14.14.1. The reason for the factor of 4 in (14.158) is so that the maximum scalar curvature of g_Σ will be equal to 1.

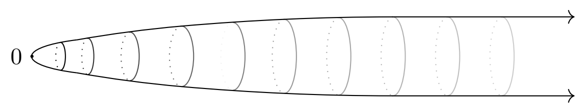

Figure 14.14.1. The cigar soliton metric g_Σ on $\mathbb{R}^2$.

The (exterior) derivative of the potential function f_Σ is given by

$$df_\Sigma = -\frac{1}{1+(x^1)^2+(x^2)^2}(2x^1\,dx^1 + 2x^2\,dx^2). \tag{14.160}$$

Thus, the gradient, with respect to g_Σ, of the potential function is given by

$$\nabla_{g_\Sigma} f_\Sigma\left(x^1, x^2\right) = \left(-\frac{x^1}{2}, -\frac{x^2}{2}\right). \tag{14.161}$$

Recall from (8.46) that if $\widetilde{g} = \mathrm{e}^{2u} g$, then

$$K_{\tilde{g}} = \mathrm{e}^{-2u}(K_g - \Delta_g u).$$

Using this, we compute that the Gauss curvature of g_Σ is equal to

$$\begin{aligned} K_\Sigma &= -\frac{1+(x^1)^2+(x^2)^2}{4}\Delta_{\mathrm{Euc}}\left(\frac{1}{2}\ln\frac{4}{1+(x^1)^2+(x^2)^2}\right) \\ &= \frac{1}{2(1+(x^1)^2+(x^2)^2)}. \end{aligned} \tag{14.162}$$

On the Riemannian surface $(\mathbb{R}^2, g_\Sigma)$ we have the global orthornormal frame field defined by

$$e_1 := \sqrt{1+(x^1)^2+(x^2)^2}\,\frac{\partial}{\partial x^1}, \qquad e_2 := \sqrt{1+(x^1)^2+(x^2)^2}\,\frac{\partial}{\partial x^2}.$$

The Hessian of f_Σ with respect to this frame field is given by

$$\nabla^2 f_\Sigma(e_i, e_j) = e_i(e_j(f_\Sigma)) - \sum_{k=1}^{2}\omega_j^k(e_i)\,e_k(f_\Sigma) \tag{14.163}$$

We have that $(e_{\mathrm{Euc}})_1 = \frac{\partial}{\partial x^1}$, $(e_{\mathrm{Euc}})_2 = \frac{\partial}{\partial x^2}$ is a global orthonormal frame field for the Euclidean metric g_{Euc}. Its dual orthonormal coframe field is given by $(\omega_{\mathrm{Euc}})^1 = dx^1$, $(\omega_{\mathrm{Euc}})^2 = dx^2$. By (8.43) and since $(\omega_{\mathrm{Euc}})_j^i = 0$, we have the connection 1-forms ω_j^i of g_Σ with respect to the orthonormal frame

e_1, e_2 are given by

$$\begin{aligned}\omega_j^k &= \frac{\partial}{\partial x^j}\left(\frac{1}{2}\ln\frac{4}{1+(x^1)^2+(x^2)^2}\right)dx^k \\ &\quad - \frac{\partial}{\partial x^k}\left(\frac{1}{2}\ln\frac{4}{1+(x^1)^2+(x^2)^2}\right)dx^j \\ &= \frac{x^k dx^j - x^j dx^k}{1+(x^1)^2+(x^2)^2}.\end{aligned}$$

Thus,

$$\omega_j^k(e_i) = \frac{x^k\delta_{ij} - x^j\delta_{ik}}{\sqrt{1+(x^1)^2+(x^2)^2}},$$

where δ_{ij} is the Kronecker delta symbol. We have

$$e_j(f_\Sigma) = -\frac{2x^j}{\sqrt{1+(x^1)^2+(x^2)^2}}$$

and

$$e_i(e_j(f_\Sigma)) = -2\delta_{ij} + \frac{2x^i x^j}{1+(x^1)^2+(x^2)^2}.$$

Moreover,

$$\begin{aligned}-\sum_{k=1}^{2}\omega_j^k(e_i)\,e_k(f_\Sigma) &= \sum_{k=1}^{2}\frac{x^k\delta_{ij} - x^j\delta_{ik}}{\sqrt{1+(x^1)^2+(x^2)^2}}\frac{2x^k}{\sqrt{1+(x^1)^2+(x^2)^2}} \\ &= \frac{2((x^1)^2+(x^2)^2)\delta_{ij} - 2x^i x^j}{1+(x^1)^2+(x^2)^2}.\end{aligned}$$

Hence, by (14.163) and by summing the last two displays, we obtain

$$\nabla^2 f_\Sigma(e_i, e_j) = -\frac{2\delta_{ij}}{\sqrt{1+(x^1)^2+(x^2)^2}} = -K_\Sigma\, g_\Sigma(e_i, e_j).$$

This proves that the cigar soliton $(\mathbb{R}^2, g_\Sigma, f_\Sigma)$ is a steady gradient Ricci soliton.

Exercise 14.3. *Show that the cigar soliton metric, defined by (14.158), may be expressed (except at the origin 0^2) by a change of coordinates as*

$$g_\Sigma = ds^2 + \tanh^2(s)\,d\theta^2, \tag{14.164}$$

for $s \in (0, \infty)$ and $\theta \in \mathbb{R}/2\pi\mathbb{Z}$.

Exercise 14.4. *Prove that the Gauss curvature of the cigar soliton metric is given by*

$$K_\Sigma = 2\,\mathrm{sech}^2(s). \tag{14.165}$$

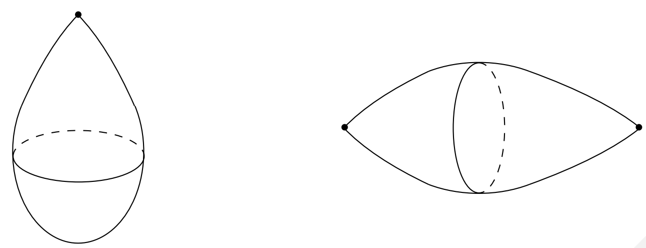

Figure 14.14.2. The teardrop (L) and football (R) shrinking Ricci solitons.

14.14.3. Shrinking Ricci solitons on orbifolds. Hamilton proved that any shrinking Ricci soliton on a bad orbifold must be rotationally symmetric. He also proved that the soliton is unique up to scaling and diffeomorphisms.

The proof (see case 2 in the proof of Lemma 14.22) of the fact that on S^2 the only shrinking Ricci solitons are the constant curvature metrics uses the Kazdan–Warner identity and hence uses the uniformization theorem. Chen, Lu, and Tian [**CLT06**] proved this result without using the uniformization theorem.

14.15. Uniformization of 2-dimensional orbifolds

Firstly, we remark that the Hodge Decomposition Theorem 11.18 extends to orbifolds. In particular, we have the following consequence in dimension 2 (cf. Corollary 11.19).

Proposition 14.29. Let (O^2, g) be a closed Riemannian orbifold with isolated singularities. If $\phi : O^2 \to \mathbb{R}$ is a function satisfying $\int_{O^2} \phi \, d\mu = 0$, then there exists a function $f : O^2 \to \mathbb{R}$ satisfying the Poisson equation

$$\Delta f = \phi. \tag{14.166}$$

Thus, for any closed Riemannian orbifold (O^2, g) with isolated singularities, there exists a function $f : O^2 \to \mathbb{R}$ satisfying

$$\Delta f = R - r, \tag{14.167}$$

where r is the average of the scalar curvature R.

By the works of Wu [**Wu91, CW91**], we have the following.

Theorem 14.30 (Uniformization of 2-dimensional orbifolds)**.** *Let (O^2, g_0) be a 2-dimensional closed oriented Riemannian orbifold. Then there exists a solution $g(t)$ to the modified Ricci flow for all time $t \in [0, \infty)$ with $g(0) = g_0$. As $t \to \infty$, $g(t)$ converges in each C^k-norm to a C^∞ metric g_∞. There exists a function f_∞ on O^2 such that (O^2, g_∞) together with f_∞ is a shrinking Ricci* *soliton. That is,*

$$(R_{g_\infty} - r)g_\infty = 2\nabla^2_{g_\infty} f_\infty. \tag{14.168}$$

(Note that $\Delta_{g_\infty} f_\infty = R_{g_\infty} - r$.) For a good orbifold, both the normalized and modified Ricci flows converge to constant curvature metrics.

For *good* orbifolds, this result is originally due to Hamilton. Any closed orbifold that admits a constant curvature metric must be a good orbifold. Therefore, any shrinking gradient Ricci soliton on a bad closed orbifold must be non-trivial; that is, its potential and curvature functions are not constant.