Mathematical
Surveys
and
Monographs

Volume 110

The Ricci Flow: An Introduction

Bennett Chow
Dan Knopf

American Mathematical Society

2000 *Mathematics Subject Classification.* Primary 53C44, 53C21, 58J35, 53C25, 35K55, 57M50, 35K57.

For additional information and updates on this book, visit
www.ams.org/bookpages/surv-110

Library of Congress Cataloging-in-Publication Data

Chow, Bennett.
The Ricci flow: an introduction/Bennett Chow, Dan Knopf.
p.cm.– (Mathematical surveys and monographs, ISSN 0076-5376; v. 110)
Includes bibliographical references and index.
ISBN 0-8218-3515-7 (alk. paper)
1. Global differential geometry. 2. Ricci flow. 3. Riemannian manifolds. I. Knopf, Dan, 1959– II. Title. III. Mathematical surveys and monographs; no. 110.

QA670.C46 2004
516.3′62–dc22 2004046148

10 9 8 7 6 5 4 3 2 1 09 08 07 06 05 04

Contents

Preface

This book and its planned sequel(s) are intended to compose an introduction to the Ricci flow in general and in particular to the program originated by Hamilton to apply the Ricci flow to approach Thurston's Geometrization Conjecture. The Ricci flow is the geometric evolution equation in which one starts with a smooth Riemannian manifold $(\mathcal{M}^n, g_0)$ and evolves its metric by the equation

$$\frac{\partial}{\partial t} g = -2 \operatorname{Rc},$$

where Rc denotes the Ricci tensor of the metric g. The Ricci flow was introduced in Hamilton's seminal 1982 paper, "Three-manifolds with positive Ricci curvature". In this paper, closed 3-manifolds of positive Ricci curvature are topologically classified as spherical space forms. Until that time, most results relating the curvature of a 3-manifold to its topology involved the influence of curvature on the fundamental group. For example, Gromov–Lawson and Schoen–Yau classified 3-manifolds with positive scalar curvature essentially up to homotopy. Among the many results relating curvature and topology that hold in arbitrary dimensions are Myers' Theorem and the Cartan–Hadamard Theorem.

A large number of innovations that originated in Hamilton's 1982 and subsequent papers have had a profound influence on modern geometric analysis. Here we mention just a few. Hamilton's introduction of a nonlinear heat-type equation for metrics, the Ricci flow, was motivated by the 1964 harmonic heat flow introduced by Eells and Sampson. This led to the renewed study by Huisken, Ecker, and many others of the mean curvature flow originally studied by Brakke in 1977. One of the techniques that has dominated Hamilton's work is the use of the maximum principle, both for functions and for tensors. This technique has been applied to control various geometric quantities associated to the metric under the Ricci flow. For example, the so-called pinching estimate for 3-manifolds with positive curvature shows that the eigenvalues of the Ricci tensor approach each other as the curvature becomes large. (This result is useful because the Ricci flow shrinks manifolds with positive Ricci curvature, which tends to make the curvature larger.) Another curvature estimate, due independently to Ivey and Hamilton, shows that the singularity models that form in dimension three necessarily have nonnegative sectional curvature. (A singularity model is a solution of the flow that arises as the limit of a sequence of dilations of an original solution approaching a singularity.) Since the underlying manifolds of such limit solutions are topologically simple, a detailed analysis of the singularities which arise in dimension three is therefore possible. For example, showing that certain singularity models are shrinking cylinders is one of the cornerstones for enabling geometric-topological surgeries to be performed on singular solutions to the Ricci flow on 3-manifolds. The Li–Yau–Hamilton-type differential Harnack quantity is another innovation; this yields an *a priori* estimate for a certain expression involving the curvature and its first and second derivatives in space. This estimate allows one to compare a solution at different points and times. In particular, it shows that in the presence of a nonnegative curvature operator, the scalar curvature does not decrease too fast. Another consequence of Hamilton's differential Harnack estimate is that slowly forming singularities (those in which the curvature of the original solution grows more quickly than the parabolically natural rate) lead to singularity models that are stationary solutions.

The convergence theory of Cheeger and Gromov has had fundamental consequences in Riemannian geometry. In the setting of the Ricci flow, Gromov's compactness theorem may be improved to obtain C^∞ convergence of a sequence of solutions to a smooth limit solution. For singular solutions, Perelman's recent *No Local Collapsing* Theorem allows one to dilate about sequences of points and times approaching the singularity time in such a way that one can obtain a limit solution that exists infinitely far back in time. The analysis of such limit solutions is important in Hamilton's singularity theory.

A guide for the reader

The reader of this book is assumed to have a basic knowledge of Riemannian geometry. Familiarity with algebraic topology and with nonlinear second-order partial differential equations would also be helpful but is not strictly necessary.

In Chapters 1 and 2, we begin the study of the Ricci flow by considering special solutions which exhibit typical properties of the Ricci flow and which guide our intuition. In the case of an initial homogeneous metric, the solution remains homogenous so that its analysis reduces to the study of a system of ODE. We study examples of such solutions in Chapter 1. Solutions which exist on long time intervals such as those which exist since time $-\infty$ or until time $+\infty$ are very special and can appear as singularity models. In Chapter 2, we both present such solutions explicitly and provide some intuition for how they arise. An important example of a stationary solution is the so-called "cigar soliton" on $\mathbb{R}^2$; intuitive solutions include the neckpinch and degenerate neckpinch. In Chapter 2, we discuss degenerate neckpinches heuristically. We also rigorously construct neckpinch solutions under certain symmetry assumptions.

The short-time existence theorem for the Ricci flow with an arbitrary smooth initial metric is proved in Chapter 3. This basic result allows one to use the Ricci flow as a practical tool. In particular, a number of smoothing results in Riemannian geometry can be proved using the short-time existence of the flow combined with the derivative estimates of Chapter 7. Since the Ricci flow system of equations is only weakly parabolic, the short-time existence of the flow does not follow directly from standard parabolic theory. Hamilton's original proof relied on the Nash–Moser inverse function theorem. Shortly thereafter, DeTurck gave a simplified proof by showing that the Ricci flow is equivalent to a strictly parabolic system.

In Chapter 4, the maximum principle for both functions and tensors is presented. This provides the technical foundations for the curvature, gradient of scalar curvature, and Li–Yau–Hamilton differential Harnack estimates proved later in the book.

In Chapter 5, we give a comprehensive treatment of the Ricci flow on surfaces. In this lowest nontrivial dimension, many of the techniques used in three and higher dimensions are exhibited. The concept of stationary solutions is introduced here and used to motivate the construction of various quantities, including Li–Yau–Hamilton (LYH) differential Harnack estimates. The entropy estimate is applied to solutions on the 2-sphere. In addition, derivative and injectivity radius estimates are first proved here. We also discuss the isoperimetric estimate and combine it with the Gromov-type compactness theorem for solutions of the Ricci flow to rule out the cigar soliton as a singularity model for solutions on surfaces. In higher dimensions, the cigar soliton is ruled out by Perelman's *No Local Collapsing* Theorem.

The original topological classification by Hamilton of closed 3-manifolds with positive Ricci curvature is proved in Chapter 6. Via the maximum principle for systems, the qualitative behavior of the curvature tensor may be reduced in this case to the study of a system of three ODE in three unknowns. Using this method, we prove the pinching estimate for the curvature mentioned above, which shows that the eigenvalues of the Ricci tensor are approaching each other at points where the scalar curvature is becoming large. This pinching estimate compares curvatures at the same point. Then we estimate the gradient of the scalar curvature in order to compare the curvatures at different points. The combination of these estimates shows that the Ricci curvatures are tending to constant. Using this fact together with estimates for the higher derivatives of curvature, we prove the convergence of the volume-normalized Ricci flow to a spherical space form.

The derivative estimates of Chapter 7 show that assuming an initial curvature bound allows one to bound all derivatives of the curvature for a short time, with the estimate deteriorating as the time tends to zero. (This deterioration of the estimate is necessary, because an initial bound on the curvature alone does not imply simultaneous bounds on its derivatives.) The derivative estimates established in this chapter enable one to prove the long-time existence theorem for the flow, which states that a unique solution to the Ricci flow exists as long as its curvature remains bounded.

In Chapter 8, we begin the analysis of singularity models by discussing the general procedures for obtaining limits of dilations about a singularity. This involves taking a sequence of points and times approaching the singularity, dilating in space and time, and translating in time to obtain a sequence of solutions. Depending on the type of singularity, slightly different methods must be employed to find suitable sequences of points and times so that one obtains a singularity model which can yield information about the geometry of the original solution near the singularity just prior to its formation.

In Chapter 9, we consider Type I singularities, those where the curvature is bounded proportionally to the inverse of the time remaining until the singularity time. This is the parabolically natural rate of singularity formation for the flow. In this case, one sees the $\mathbb{R} \times \mathcal{S}^2$ cylinder (or its quotients) as singularity models. The results in this chapter provide additional rigor to the intuition behind neckpinches.

In the two appendices, we provide elementary background for tensor calculus and some comparison geometry.

REMARK. At present (December 2003) there has been sustained excitement in the mathematical community over Perelman's recent groundbreaking progress on Hamilton's Ricci flow program intended to resolve Thurston's Geometrization Conjecture. Perelman's results allow one to view some of the material in Chapters 8 and 9 in a new light. We have retained that material in this volume for its independent interest, and hope to present Perelman's work in a subsequent volume.

A guide for the hurried reader

The reader wishing to develop a nontechnical appreciation of the Ricci flow program for 3-manifolds as efficiently as possible is advised to follow the fast track outlined below.

In Chapter 1, read Section 1 for a brief introduction to the Geometrization Conjecture.

In Chapter 2, the most important examples are the cigar, the neckpinch and the degenerate neckpinch. Read the discussion of the cigar soliton in Section 2. Read the statements of the main results on neckpinches derived in Section 5. And read the heuristic discussion of degenerate neckpinches in Section 6.

In Chapter 3, review the variation formulas derived in Section 1.

In Chapter 4, read the proof of the first scalar maximum principle (Theorem 4.2) and at least the statements of the maximum principles that follow.

In Chapter 5, review the entropy estimates derived in Section 8 and the differential Harnack estimates derived in Section 10. The entropy estimates will be used in Section 6 of Chapter 9. The Harnack estimates for surfaces are prototypes of those that apply in higher dimensions.

We suggest that Chapter 6 be read in its entirety.

In Chapter 7, read the statements of the main results. These give precise insight into the smoothing properties of the Ricci flow and its long-time behavior. The Compactness Theorem stated in Section 3 is an essential ingredient in important technique of analyzing singularity formation by taking limits of parabolic dilations.

In Chapter 8, read Sections 1–3.1. These present the classification of maximal-time solutions to the flow, and introduce the method of parabolic dilation.

In Chapter 9, read Section 1 for a heuristic description of singularity formation in 3-manifolds. Then read Sections 2 and 3 to gain an understanding of why positive curvature dominates near singularities in dimension three. Finally, review the results of Sections 4 and 6 to gain insight into the technique of dimension reduction.

Acknowledgments

During the preparation of this volume, Bennett Chow received support from NSF grant DMS-9971891, while Dan Knopf was partially supported by DMS-0328233.

B.C. would like to thank the following mathematicians from whom he has benefitted both mathematically and personally. He thanks Professor Shiing-Shen Chern whom he met in 1979, and who provided him with the inspiration to study differential geometry. From his graduate school days at Princeton and San Diego: Huai-Dong Cao, Craig Hodgson, Nikos Kapouleas, Igor Rivin, Fadi Twainy, and Cumrun Vafa. From his postdoc years at MIT, Rice, and NYU: Zhiyong Gao, Victor Guillemin, Richard Melrose, Louis Nirenberg, Dan Stroock, and Deane Yang. From the years at Minnesota: his colleagues Scot Adams, Bob Gulliver, Conan Leung, Wei-Ming Ni, and Jiaping Wang, as well as his former students Sun-Chin (Michael) Chu, Alexei Krioukov, Lii-Perng Liou, Dan O'Loughlin, Dong-Ho Tsai, and Yu Yuan. From San Diego: his colleagues Li-Tien Cheng, Zheng-Xu (John) He, Lei Ni, and Nolan Wallach, as well as his former student David Glickenstein. Special thanks to B.C.'s advisor, Professor Shing-Tung Yau, and teacher, Professor Richard Hamilton, for their invaluable teachings, support, and encouragement for more than 20 years, without which this book would not have been written. Very special thanks to Classic Dimension for encouragement, support and inspiration. B.C. is indebted to his parents for the love, support, and encouragement they have given him all of these years, and to his father for teaching him mathematics beginning in his youth. B.C. dedicates this book to his parents Yutze and Wanlin and to the memory of his sister Eleanor.

The authors would also like to thank Sigurd Angenent, Albert Chau, Hsiao-Bing Cheng, Christine Guenther, Joel Hass, Gerhard Huisken, Tom Ilmanen, Jim Isenberg, Tom Ivey, Peter Li, Chang-Shou Lin, Fang-Hua Lin, Peng Lu, Feng Luo, Rick Schoen, Miles Simon, Gang Tian, Mu-Tao Wang, Brian White, and Xi-Ping Zhu. They thank Helena Noronha for her encouragement, Ed Dunne and Sergei Gelfand at AMS for their help, and André Neves for proofreading and offering many useful suggestions.

In addition to the other mathematicians mentioned above, D.K. extends special thanks to his advisor Kevin McLeod. D.K. thanks his parents for all that they are, and especially his father for his loving patience and support. Thanks to Erik, Susie, Yoyi, Hilarie, and everyone at ERC. D.K. dedicates this book to his wife, Stephanie.

Bennett Chow and Dan Knopf

CHAPTER 1

The Ricci flow of special geometries

The **Ricci flow**

$$\frac{\partial}{\partial t}g = -2\,\mathrm{Rc}$$
$$g(0) = g_0$$

and its cousin the **normalized Ricci flow**

$$\frac{\partial}{\partial t}g = -2\,\mathrm{Rc} + \frac{2}{n}\frac{\int_{\mathcal{M}^n} R\,d\mu}{\int_{\mathcal{M}^n} d\mu}g$$
$$g(0) = g_0$$

are methods of evolving the metric of a Riemannian manifold $(\mathcal{M}^n, g_0)$ that were introduced by Hamilton in [**58**]. They differ only by a rescaling of space and time. Hamilton has crafted a well-developed program to use these flows to resolve Thurston's Geometrization Conjecture for closed 3-manifolds. The intent of this volume is to provide a comprehensive introduction to the foundations of Hamilton's program. Perelman's recent ground-breaking work [**105, 106, 107**] is aimed at completing that program.

Roughly speaking, Thurston's Geometrization Conjecture says that any closed 3-manifold can be canonically decomposed into pieces in such a way that each admits a unique homogeneous geometry. (See Section 1 below.) As we will learn in the chapters that follow, one cannot in general expect a solution $(\mathcal{M}^3, g(t))$ of the Ricci flow starting on an arbitrary closed 3-manifold to converge to a complete locally homogeneous metric. Instead, one must deduce topological and geometric properties of $\mathcal{M}^3$ from the behavior of $g(t)$. Hamilton's program outlines a highly promising strategy to do so.

By way of an intuitive introduction to this strategy, this chapter addresses the following natural question:

If g_0 is a complete locally homogeneous metric, how will $g(t)$ evolve?

The observations we collect in examining this question are intended to help the reader develop a sense and intuition for the properties of the flow in these special geometries. While knowledge of the Ricci flow's behavior in homogeneous geometries does not appear necessary for understanding its topological consequences, such knowledge is valuable for understanding analytic aspects of the flow, particularly those related to collapse.

REMARK 1.1. In dimension $n = 3$, the categories TOP, PL, and DIFF all coincide. In this volume, any manifold under consideration — regardless of its dimension — is assumed to be smooth (C^∞). Moreover, unless explicitly stated otherwise, every manifold is assumed to be without boundary, so that it is closed if and only if it is compact.

1. Geometrization of three-manifolds

To motivate our interest in locally homogeneous metrics on 3-manifolds, we begin with a heuristic discussion of the Geometrization Conjecture, which will be reviewed in somewhat more detail in the successor to this volume. (Good references are [**115**] and [**123**].) We begin that discussion with a brief review of some basic 3-manifold topology. One says an orientable closed manifold $\mathcal{M}^3$ is **prime** if $\mathcal{M}^3$ is not diffeomorphic to the 3-sphere $\mathcal{S}^3$ and if a connected sum decomposition $\mathcal{M}^3 = \mathcal{M}_1^3 \# \mathcal{M}_2^3$ is possible only if $\mathcal{M}_1^3$ or $\mathcal{M}_2^3$ is itself diffeomorphic to $\mathcal{S}^3$. One says an orientable closed manifold $\mathcal{M}^3$ is **irreducible** if every separating embedded 2-sphere bounds a 3-ball. It is well known that the only orientable 3-manifold that is prime but not irreducible is $\mathcal{S}^2 \times \mathcal{S}^1$.

A consequence of the **Prime Decomposition Theorem** [**85, 97**] is that an orientable closed manifold $\mathcal{M}^3$ can be decomposed into a finite connected sum of prime factors

$$\mathcal{M}^3 \approx (\#_j \mathcal{X}_j^3) \ \# \ (\#_k \mathcal{Y}_k^3) \ \# \ (\#_\ell (\mathcal{S}^2 \times \mathcal{S}^1)).$$

Each $\mathcal{X}_j^3$ is irreducible with finite fundamental group and universal cover a homotopy 3-sphere. Each $\mathcal{Y}_k^3$ is irreducible with infinite fundamental group and a contractible universal cover. The prime decomposition is unique up to re-ordering and orientation-preserving diffeomorphisms of the factors. From the standpoint of topology, the Prime Decomposition Theorem reduces the study of closed 3-manifolds to the study of irreducible 3-manifolds.

There is a further decomposition of irreducible manifolds, but we must recall more nomenclature before we can state it precisely. Let Σ^2 be a two-sided compact properly embedded surface in a manifold-with-boundary $\mathcal{N}^3$. Assume that Σ^2 has no components diffeomorphic to the 2-disc $\mathcal{D}^2$, and that Σ^2 either lies in $\partial\mathcal{N}^3$ or intersects $\partial\mathcal{N}^3$ only in $\partial\Sigma^2$. Under these conditions, one says Σ^2 is **incompressible** if for each $\mathcal{D}^2 \subset \mathcal{N}^3$ with $\mathcal{D}^2 \cap \Sigma^2 = \partial\mathcal{D}$, there exists a disc $\mathcal{D}_* \subset \Sigma^2$ with $\partial\mathcal{D}_* = \partial\mathcal{D}$.

Let $\mathcal{M}^3$ be irreducible. The **Torus Decomposition Theorem** [**79, 80**] says that there exists a finite (possibly empty) collection of disjoint incompressible 2-tori $\mathcal{T}_i^2$ such that each component $\mathcal{N}^3$ of $\mathcal{M}^3 \setminus \cup \mathcal{T}_i^2$ is either geometrically atoroidal or a Seifert fiber space, and that a minimal such collection $\{\mathcal{T}_i\}$ is unique up to homotopy. One says an irreducible manifold-with-boundary $\mathcal{N}^3$ is **geometrically atoroidal** if every incompressible torus $\mathcal{T}^2 \subset \mathcal{N}^3$ is isotopic to a component of $\partial\mathcal{N}$. One says a compact manifold $\mathcal{N}^3$ is a **Seifert fiber space** if it admits a foliation by $\mathcal{S}^1$ fibers.

REMARK 1.2. Seifert's original definition [**116**] of a fiber space required the existence of a fiber-preserving diffeomorphism of a tubular neighborhood of each fiber to a neighborhood of a fiber in some quotient of $\mathcal{S}^1 \times \mathcal{D}^2$ by a cyclic group action. But Epstein showed later [**39**] that if a compact 3-manifold is foliated by $\mathcal{S}^1$ fibers, then each fiber must possess a tubular neighborhood with the prescribed property. Hence the simpler definition we have given above is equivalent to the original.

One says $\mathcal{M}^3$ is **Haken** if it is prime and contains an incompressible surface other than $\mathcal{S}^2$. In [**122**], Thurston proved the important result that if $\mathcal{M}^3$ is Haken (in particular, if $\mathcal{M}^3$ admits a nontrivial torus decomposition) then $\mathcal{M}^3$ admits a canonical decomposition into finitely many pieces $\mathcal{N}_i^3$ such that each possesses a unique geometric structure. For now, the reader should regard a geometric structure as a complete locally homogeneous Riemannian metric g_i on $\mathcal{N}_i^3$. A more thorough discussion of geometric structures will be found in Sections 2 and 3 below.

Thurston's **Geometrization Conjecture** asserts that a geometric decomposition holds for non-Haken manifolds as well, namely that every closed 3-manifold can be canonically decomposed into pieces such that each admits a unique geometric structure. In light of the Torus Decomposition Theorem and more recent results from topology, one can summarize what is currently known about the Geometrization Conjecture as follows. Let $\mathcal{N}^3$ be an irreducible orientable closed 3-manifold. If its fundamental group $\pi_1(\mathcal{N}^3)$ contains a subgroup isomorphic to the fundamental group $\mathbb{Z} \oplus \mathbb{Z}$ of the torus, then either $\pi_1(\mathcal{N}^3)$ has a nontrivial center or else $\mathcal{N}^3$ contains an incompressible torus. (See [**114**] and also [**124**].) If $\pi_1(\mathcal{N}^3)$ has a nontrivial center, then results of Casson–Jungreis [**22**] and Gabai [**42**] imply that $\mathcal{N}^3$ is a Seifert fiber space, all of which are known to be geometrizable. On the other hand, if $\mathcal{N}^3$ contains an incompressible torus, then it is Haken, hence geometrizable by Thurston's result. Thus only the following two cases of Thurston's Conjecture are open today:

CONJECTURE 1.3 (Elliptization). *Let $\mathcal{N}^3$ be an irreducible orientable closed 3-manifold of finite fundamental group $\pi_1(\mathcal{N}^3)$. Then $\mathcal{N}^3$ is diffeomorphic to a quotient $\mathcal{S}^3/\Gamma$ of the 3-sphere by a finite subgroup Γ of $\mathrm{O}(4)$. In particular, $\mathcal{N}^3$ admits a Riemannian metric of constant positive curvature.*

CONJECTURE 1.4 (Hyperbolization). *Let $\mathcal{N}^3$ be an irreducible orientable closed 3-manifold of infinite fundamental group $\pi_1(\mathcal{N}^3)$ such that $\pi_1(\mathcal{N}^3)$ contains no subgroup isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$. Then $\mathcal{N}^3$ admits a complete hyperbolic metric of finite volume.*

REMARK 1.5. A complete proof of the Elliptization Conjecture would imply the Poincaré Conjecture, which asserts that any homotopy 3-sphere is actually a topological 3-sphere. (See Chapter 6.)

REMARK 1.6. Noteworthy progress toward the Hyperbolization Conjecture has come from topology. For example, Gabai has proved [**43**] that if

$\mathcal{N}^3$ is closed, irreducible, and homotopic to a hyperbolic 3-manifold, then $\mathcal{N}^3$ is homeomorphic to a hyperbolic 3-manifold.

2. Model geometries

A **geometric structure** on an n-dimensional manifold $\mathcal{M}^n$ may be regarded as a complete locally homogeneous Riemannian metric g. A metric is locally homogeneous if it looks the same at every point. More precisely, one says that g is **locally homogeneous** if for all $x, y \in \mathcal{M}^n$ there exist neighborhoods $\mathcal{U}_x \subseteq \mathcal{M}^n$ of x and $\mathcal{U}_y \subseteq \mathcal{M}^n$ of y and a g-isometry $\gamma_{xy} : \mathcal{U}_x \to \mathcal{U}_y$ such that $\gamma_{xy}(x) = y$. In this section, we shall develop a more abstract way to think about geometric structures.

One says that a Riemannian metric g on $\mathcal{M}^n$ is **homogeneous** (to wit, *globally* homogeneous) if for every $x, y \in \mathcal{M}^n$ there exists a g-isometry $\gamma_{xy} : \mathcal{M}^n \to \mathcal{M}^n$ such that $\gamma_{xy}(x) = y$. These concepts are equivalent when $\mathcal{M}^n$ is simply connected.

PROPOSITION 1.7 (Singer [**120**]). *Any locally homogeneous metric g on a simply-connected manifold is globally homogeneous.*

We also have the following well-known fact.

PROPOSITION 1.8. *If $(\mathcal{M}^n, g)$ is homogeneous, then it is complete.*

Thus if $(\mathcal{M}^n, g)$ is locally homogeneous, we say that its geometry is given by the **homogeneous model** $(\widetilde{\mathcal{M}^n}, \tilde{g})$, where $\widetilde{\mathcal{M}^n}$ is the universal cover of $\mathcal{M}^n$ and $\tilde{g}$ is the complete metric obtained by lifting the metric g. Thus to study geometric structures, it suffices to study homogeneous models.

A **model geometry** in the sense of Klein is a tuple $(\mathcal{M}^n, \mathcal{G}, \mathcal{G}_*)$, where $\mathcal{M}^n$ is a simply-connected smooth manifold and $\mathcal{G}$ is a group of diffeomorphisms that acts smoothly and transitively on $\mathcal{M}^n$ such that for each $x \in \mathcal{M}^n$, the **isotropy group** (point stabilizer)

$$\mathcal{G}_x \doteqdot \{\gamma \in \mathcal{G} : \gamma(x) = x\}$$

is isomorphic to $\mathcal{G}_*$. We say a model geometry $(\mathcal{M}^n, \mathcal{G}, \mathcal{G}_*)$ is a **maximal model geometry** if $\mathcal{G}$ is maximal among all subgroups of the diffeomorphism group $\mathfrak{D}(\mathcal{M}^n)$ that have compact isotropy groups. (An example of particular relevance occurs when $\mathcal{M}^n = \mathcal{G}$ is a 3-dimensional unimodular Lie group. These maximal models were classified by Milnor in [**98**] and are briefly reviewed in Section 4 of this chapter.)

The concepts of model geometry and homogeneous model are essentially equivalent (up to a subtle ambiguity that will be addressed in Remark 1.13). One direction is easy to establish.

LEMMA 1.9. *Every model geometry $(\mathcal{M}^n, \mathcal{G}, \mathcal{G}_*)$ may be regarded as a complete homogeneous space $(\mathcal{M}^n, g)$.*

PROOF. One may obtain a complete homogeneous $\mathcal{G}$-invariant Riemannian metric on $\mathcal{M}^n$ as follows. Choose any $x \in \mathcal{M}^n$. If $\hat{g}_x$ is any scalar product on $T_x\mathcal{M}^n$, one obtains a $\mathcal{G}_x$-invariant scalar product g_x by averaging $\hat{g}_x$ under the action of $\mathcal{G}_x \cong \mathcal{G}_*$. Then since $\mathcal{G}$ is transitive, there exists for any $y \in \mathcal{M}^n$ an element $\gamma_{yx} \in \mathcal{G}$ such that $\gamma_{yx}(y) = x$. So one may define a scalar product g_y for all $V, W \in T_y\mathcal{M}^n$ by

$$g_y(V,W) \doteqdot \left(\gamma_{yx}^* g_x\right)(V,W) = g_x\left(\left(\gamma_{yx}\right)_* V,\ \left(\gamma_{yx}\right)_* W\right).$$

To see that g_y is well defined, suppose that γ'_{yx} is another element of $\mathcal{G}$ such that $\gamma'_{yx}(y) = x$. Then $\gamma'_{yx} \circ \gamma_{yx}^{-1} \in \mathcal{G}_x$, so by $\mathcal{G}_x$-invariance of g_x, we have

$$\begin{aligned}\left(\gamma_{yx}^* g_x\right)(V,W) &= g_x\left(\left(\gamma'_{yx} \circ \gamma_{yx}^{-1}\right)_* \left(\gamma_{yx}\right)_* V,\ \left(\gamma'_{yx} \circ \gamma_{yx}^{-1}\right)_* \left(\gamma_{yx}\right)_* W\right) \\ &= g_x\left(\left(\gamma'_{yx}\right)_* V,\ \left(\gamma'_{yx}\right)_* W\right) = \left(\left(\gamma'_{yx}\right)^* g_x\right)(V,W).\end{aligned}$$

Because the action of $\mathcal{G}$ is smooth, one obtains in this way a smooth Riemannian metric g on $\mathcal{M}^n$. Because the action of $\mathcal{G}$ is transitive, g is complete by Proposition 1.8. □

To complete the connection between homogeneous models and model geometries, we want to establish the converse of the construction above. Namely, we want to show that every homogeneous model $(\widetilde{\mathcal{M}}^n, \tilde{g})$ may be regarded as a model geometry $(\mathcal{M}^n, \mathcal{G}, \mathcal{G}_*)$. We begin by recalling some classical facts about the set $\mathrm{Isom}(\mathcal{M}^n, g)$ of isometries of a Riemannian manifold. $\mathrm{Isom}(\mathcal{M}^n, g)$ is clearly a subgroup of the diffeomorphism group $\mathfrak{D}(\mathcal{M}^n)$. Moreover, it is clear that $(\mathcal{M}^n, g)$ is homogeneous if and only if $\mathrm{Isom}(\mathcal{M}^n, g)$ acts transitively on $\mathcal{M}^n$. The following facts are classical. (See [**100**] and [**87**].)

PROPOSITION 1.10. *Let $(\mathcal{M}^n, g)$ be a smooth Riemannian manifold.*

(1) $\mathrm{Isom}(\mathcal{M}^n, g)$ *is a Lie group and acts smoothly on* $\mathcal{M}^n$.
(2) *For each $x \in \mathcal{M}^n$, the isotropy group*

$$I_x(\mathcal{M}^n, g) \doteqdot \{\gamma \in \mathrm{Isom}(\mathcal{M}^n, g) : \gamma(x) = x\}$$

is a compact subgroup of $\mathrm{Isom}(\mathcal{M}^n, g)$.
(3) *For each $x \in \mathcal{M}^n$, the linear isotropy representation*

$$\lambda_x : I_x(\mathcal{M}^n, g) \to \mathrm{O}(T_x\mathcal{M}^n, g(x))$$

defined by

$$\lambda_x(\gamma) = \gamma_* : T_x\mathcal{M}^n \to T_x\mathcal{M}^n$$

is an injective group homomorphism onto a closed subgroup of the orthogonal group $\mathrm{O}(T_x\mathcal{M}^n, g(x))$. *In particular,*

$$I_x(\mathcal{M}^n, g) \cong \lambda_x(I_x(\mathcal{M}^n, g)) \subseteq \mathrm{O}(T_x\mathcal{M}^n, g(x))$$

is compact.

COROLLARY 1.11.

(1) *Given* $x, y \in \mathcal{M}^n$ *and a linear isometry*

$$f : (T_x\mathcal{M}^n, g(x)) \to (T_y\mathcal{M}^n, g(y)),$$

there exists at most one isometry $\gamma \in \operatorname{Isom}(\mathcal{M}^n, g)$ *with* $\gamma(x) = y$ *and* $\gamma_* = f$.

(2) *If* $\mathcal{M}^n$ *is compact, then* $\operatorname{Isom}(\mathcal{M}^n, g)$ *is compact.*

When $(\mathcal{M}^n, g)$ is homogeneous, all isotropy groups look the same.

LEMMA 1.12. *If* $\operatorname{Isom}(\mathcal{M}^n, g)$ *acts transitively on* $\mathcal{M}^n$, *then all isotropy groups* $I_x(\mathcal{M}^n, g)$ *are isomorphic.*

PROOF. Given any $x, y \in \mathcal{M}^n$, there exists an isometry $\gamma : \mathcal{M}^n \to \mathcal{M}^n$ such that $\gamma(x) = y$. Define

$$\gamma_{\#} : I_x(\mathcal{M}^n, g) \to I_y(\mathcal{M}^n, g)$$

by

$$\gamma_{\#}(\beta) \doteqdot \gamma \circ \beta \circ \gamma^{-1}.$$

Clearly, $\gamma_{\#}$ is a group isomorphism with inverse $(\gamma^{-1})_{\#}$. □

Thus one may regard a complete homogeneous space $(\mathcal{M}^n, g)$ as a model geometry $(\mathcal{M}^n, \operatorname{Isom}(\mathcal{M}^n, g), I_x(\mathcal{M}^n, g))$.

REMARK 1.13. There is an ambiguity in the equivalence of model geometries and homogeneous models that we now address. The variety of possible Riemannian metrics g compatible with a given model geometry $(\mathcal{M}^n, \mathcal{G}, \mathcal{G}_*)$ depends strongly on the size of the isotropy group $\mathcal{G}_*$. If the isotropy group $\mathcal{G}_*$ is large enough, for example if $\mathcal{G}_* \approx \mathrm{O}(n)$, then one may show easily that a unique $\mathcal{G}_*$-invariant metric g, hence a unique homogenous model $(\mathcal{M}^n, g)$ is determined up to a scalar multiple. For smaller isotropy groups $\mathcal{G}_*$, there may be more choices for g, and these may endow $\mathcal{M}^n$ with somewhat different geometric properties. On the other hand, there may be closed proper subgroups of $\operatorname{Isom}(\mathcal{M}^n, g)$ acting transitively on $(\mathcal{M}^n, g)$, so that a homogeneous model $(\mathcal{M}^n, g)$ may be regarded as a model geometry $(\mathcal{M}^n, \mathcal{G}, \mathcal{G}_*)$ under different groups $(\mathcal{M}^n, \mathcal{G}', \mathcal{G}'_*)$. For the purpose of giving a standard description of the geometric structures that occur in the Geometrization Conjecture, this ambiguity is addressed by choosing the isotropy groups to be as large as possible, in other words by considering only maximal model geometries. In fact, Thurston proved that there are exactly eight maximal model geometries that have compact 3-manifold representatives.

3. Classifying three-dimensional maximal model geometries

In dimension three, a surprisingly rich collection of model geometries is obtained by taking $\mathcal{M}^3 = \mathcal{G}^3$ to be a Lie group acting transitively on itself by multiplication. Since we are interested only in closed manifolds modeled on $\mathcal{G}^3$, we can make a further restriction. One says a Lie group

Signature	Lie group	Description
$(-1,-1,-1)$	$\mathrm{SU}(2)$	compact, simple
$(-1,-1,0)$	$\widetilde{\mathrm{Isom}}(\mathbb{R}^2)$	solvable
$(-1,-1,+1)$	$\widetilde{\mathrm{SL}}(2,\mathbb{R})$	noncompact, simple
$(-1,0,0)$	nil	nilpotent
$(-1,0,+1)$	sol	solvable
$(0,0,0)$	$\mathbb{R}\oplus\mathbb{R}\oplus\mathbb{R}$	commutative

TABLE 1. 3-dimensional unimodular simply-connected Lie groups

$\mathcal{G}^n$ is **unimodular** if its volume form is bi-invariant. This is equivalent to the statement that the 1-parameter family of diffeomorphisms generated by any left-invariant vector field preserves volume. Only unimodular Lie groups admit compact quotients, because only those groups admit quotients of finite volume. Thus one can make the following observation.

LEMMA 1.14. *If $(\mathcal{M}^n, g)$ is a homogeneous model such that* $\mathrm{Isom}(\mathcal{M}^n, g)$ *is diffeomorphic to $\mathcal{M}^n$, then $\mathcal{M}^n$ is a Lie group. If moreover there exists a discrete subgroup Γ of* $\mathrm{Isom}(\mathcal{M}^n, g)$ *such that $\mathcal{M}^n/\Gamma$ is compact, then $\mathcal{M}^n$ is unimodular.*

In [**98**], Milnor classified all 3-dimensional unimodular Lie groups. (His method is sketched in Section 4 below.) Any simply-connected 3-dimensional unimodular Lie group $\mathcal{M}^3$ must be isomorphic to one of the six groups listed in Table 1. Remarkably, the groups in this list make up five of the eight maximal models that appear in the Geometrization Conjecture. (Thurston discards $\widetilde{\mathrm{Isom}}(\mathbb{R}^2)$ because it is not maximal; see Remark 1.13.)

The key to classifying all maximal model geometries $(\mathcal{M}^3, \mathcal{G}, \mathcal{H})$ with compact representatives is to look more closely at the isotropy group $\mathcal{G}_*$. Let $(\mathcal{M}^3, \mathcal{G}, \mathcal{H})$ be a maximal model geometry, and let $\mathcal{G}'$ denote the connected component of the identity in $\mathcal{G}$. Then $\mathcal{G}'$ acts transitively on $\mathcal{M}^3$, so by Lemma 1.12, all isotropy groups $\mathcal{G}'_x$ of the $\mathcal{G}'$ action are isomorphic. If $\mathcal{G}''_x$ denotes the connected component of the identity in $\mathcal{G}_x$, then the collection $\mathcal{G}'_x/\mathcal{G}''_x$ forms a covering space of $\mathcal{M}^n$. Since $\mathcal{M}^n$ is simply connected, this covering must be trivial. It follows that all $\mathcal{G}'_x$ are connected, hence isomorphic to a connected closed subgroup $\mathcal{G}'_*$ of $\mathrm{SO}(3)$. Since the only proper nontrivial Lie subalgebra of $\mathfrak{so}(3)$ is $\mathfrak{so}(2)$, one concludes that $\mathcal{G}'_*$ is either $\mathrm{SO}(3)$ itself or $\mathrm{SO}(2)$ or else the trivial group.

Following Thurston's argument in Chapter 3.8 of [**123**], one is thus able to classify all 3-dimensional maximal model geometries which admit compact representatives:

PROPOSITION 1.15. *Let $(\mathcal{M}^3, \mathcal{G}, \mathcal{G}_*)$ be a maximal model geometry represented by at least one compact 3-manifold. Then exactly one of the following is true:*

(1) $\mathcal{G}_* \cong \mathrm{SO}(3)$, *in which case either*

   (a) $\mathcal{M}^3$ *is a Lie group isomorphic to* SU (2)*;*
   (b) $\mathcal{M}^3$ *is a Lie group isomorphic to* $\mathbb{R}^3$*; or*
   (c) $\mathcal{M}^3$ *is diffeomorphic to hyperbolic 3-space* $\mathcal{H}^3$.
(2) $\mathcal{G}_* \cong \mathrm{SO}(2)$, *in which case either*
   (a) $\mathcal{M}^3$ *is a trivial bundle over a 2-dimensional maximal model, so that*
      (i) $\mathcal{M}^3$ *is diffeomorphic to* $\mathcal{S}^2 \times \mathbb{R}$*; or*
      (ii) $\mathcal{M}^3$ *is diffeomorphic to* $\mathcal{H}^2 \times \mathbb{R}$*;*
   (b) $\mathcal{M}^3$ *constitutes a nontrivial bundle over a 2-dimensional maximal model, so that*
      (i) $\mathcal{M}^3$ *is a Lie group isomorphic to* nil*; or*
      (ii) $\mathcal{M}^3$ *is a Lie group isomorphic to* $\widetilde{\mathrm{SL}(2, \mathbb{R})}$.
(3) $\mathcal{G}_*$ *is trivial, in which case* $\mathcal{M}^3$ *is a Lie group isomorphic to* sol.

4. Analyzing the Ricci flow of homogeneous geometries

In a homogeneous geometry, every point looks the same as any other. As a consequence, the Ricci flow is reduced from a system of partial differential equations to a system of ordinary differential equations. In this section, we shall outline how this reduction may be done for the five model geometries that can be realized as a pair $(\mathcal{G}, \mathcal{G})$, where $\mathcal{G}$ is a simply-connected unimodular Lie group. The theory in these cases is most elegant, but the remaining three model geometries can be treated in a similar fashion.

Let $\mathcal{G}^n$ be any Lie group, and let $\mathfrak{g}$ be the Lie algebra of all left-invariant vector fields on $\mathcal{G}$. Since a left-invariant metric on $\mathcal{G}$ is equivalent to a scalar product on $\mathfrak{g}$, the set of all such metrics can be identified with the set $\mathcal{S}_n^+$ of symmetric positive-definite $n \times n$ matrices. $\mathcal{S}_n^+$ is an open convex subset of $\mathbb{R}^{n(n+1)/2}$. For each left-invariant metric g on $\mathcal{G}$, the Ricci flow may thus be regarded as a path $t \mapsto g(t) \in \mathcal{S}_n^+$. One would thus expect the Ricci flow to reduce to a system of $n(n+1)/2$ coupled ordinary differential equations. In fact, we can do better.

Let us equip $\mathcal{G}$ with a left-invariant moving frame $\{F_i\}$. Then the structure constants c_{ij}^k for the Lie algebra $\mathfrak{g}$ of left-invariant vector fields on $\mathcal{G}$ are defined by

$$[F_i, F_j] = c_{ij}^k F_k,$$

and the adjoint representation of $\mathfrak{g}$ is the map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \cong \mathfrak{gl}(n, \mathbb{R})$ given by

$$(\mathrm{ad}\, V)\, W \doteqdot [V, W] = V^i W^j c_{ij}^k F_k.$$

If g is a left-invariant metric on $\mathcal{G}$, then the adjoint $\mathrm{ad}^* : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ with respect to g of the map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is given by

$$\langle (\mathrm{ad}\, X)^* Y, Z \rangle = \langle Y, (\mathrm{ad}\, X)\, Z \rangle = \langle Y, [X, Z] \rangle .$$

Now suppose $\mathcal{G}$ is a 3-dimensional Lie group. Let $\{F_i\}$ be a left-invariant moving frame chosen to be orthonormal with respect to a given left-invariant metric g.. Then one may define a vector space isomorphism $\mathfrak{g} \to \wedge^2\mathfrak{g}$ by

$$F_i \mapsto F_{i+1 \bmod 3} \wedge F_{i+2 \bmod 3}.$$

Composing this with the commutator, regarded as the linear map $\wedge^2\mathfrak{g} \to \mathfrak{g}$ sending $V \wedge W \mapsto [V, W]$, yields a vector space endomorphism $\Gamma : \mathfrak{g} \to \mathfrak{g}$ whose matrix with respect to the ordered basis $\beta = (F_1, F_2, F_3)$ is

$$\Gamma_\beta = \begin{pmatrix} c_{23}^1 & c_{31}^1 & c_{12}^1 \\ c_{23}^2 & c_{31}^2 & c_{12}^2 \\ c_{23}^3 & c_{31}^3 & c_{12}^3 \end{pmatrix}.$$

Observe that the matrix Γ_β displays all the structure constants of $\mathfrak{g}$.

Suppose further that $\mathcal{G}$ is unimodular. In this case, one has $\operatorname{tr}\operatorname{ad} V = 0$ for every $V \in \mathfrak{g}$, which implies in particular that Γ is self adjoint with respect to g. So by an orthogonal change of basis $\beta \mapsto \alpha = (\bar{F}_1, \bar{F}_2, \bar{F}_3)$, the matrix Γ_β of Γ can be transformed into

$$\Gamma_\alpha = \begin{pmatrix} \bar{c}_{23}^1 & & \\ & \bar{c}_{31}^2 & \\ & & \bar{c}_{12}^3 \end{pmatrix} \equiv \begin{pmatrix} 2\lambda & & \\ & 2\mu & \\ & & 2\nu \end{pmatrix}, \tag{1.1}$$

where the $\bar{c}_{ij}^k$ are the Lie algebra structure constants for $\{\bar{F}_i\}$. We can arrange that $\lambda, \mu, \nu \in \{-1, 0, 1\}$ without changing the metric g if we are willing to let the frame $\{\tilde{F}_i\}$ be merely orthogonal. For instance, if λ, μ, ν are all nonzero, we get $\tilde{\lambda}, \tilde{\mu}, \tilde{\nu} = \pm 1$ by putting

$$\tilde{F}_1 = \frac{1}{\sqrt{|\mu\nu|}}\bar{F}_1, \quad \tilde{F}_2 = \frac{1}{\sqrt{|\lambda\nu|}}\bar{F}_2, \quad \tilde{F}_3 = \frac{1}{\sqrt{|\lambda\mu|}}\bar{F}_3.$$

The general case is just as easy. We can also arrange that $\lambda \leq \mu \leq \nu$. For instance, if $\tilde{F}_1 = -\bar{F}_2$, $\tilde{F}_2 = \bar{F}_1$, $\tilde{F}_3 = \bar{F}_3$, the orientation of the manifold is unchanged, but $(\tilde{\lambda}, \tilde{\mu}, \tilde{\nu}) = (\mu, \lambda, \nu)$. Finally if we are willing to reverse the orientation of $\mathcal{G}$, we can arrange that there are at least as many negative structure constants as positive, thereby displaying exactly six distinct signatures. This construction and classification was first described in [**98**], so we call an orthogonal frame field for which $\lambda \leq \mu \leq \nu \in \{-1, 0, 1\}$ a **Milnor frame** for the left-invariant metric g. Table 1 (above) displays the possible simply-connected unimodular Lie groups.

Let $\{F_i\}$ be a Milnor frame for some left-invariant metric g on a 3-dimensional unimodular Lie group $\mathcal{G}$. Then there are positive constants A, B, and C so that g may be written with respect to the set $\{\omega^i\}$ of 1-forms dual to $\{F_i\}$ as

$$g = A\,\omega^1 \otimes \omega^1 + B\,\omega^2 \otimes \omega^2 + C\,\omega^3 \otimes \omega^3. \tag{1.2}$$

If we want to evolve g by the Ricci flow, we must study its curvature. For brevity, we display the map ad^* in the form

$$\begin{pmatrix} (\mathrm{ad}\, F_1)^* F_1 & (\mathrm{ad}\, F_2)^* F_1 & (\mathrm{ad}\, F_3)^* F_1 \\ (\mathrm{ad}\, F_1)^* F_2 & (\mathrm{ad}\, F_2)^* F_2 & (\mathrm{ad}\, F_3)^* F_2 \\ (\mathrm{ad}\, F_1)^* F_3 & (\mathrm{ad}\, F_2)^* F_3 & (\mathrm{ad}\, F_3)^* F_3 \end{pmatrix}.$$

If $\{F_i\}$ is a Milnor frame, it is easy to compute that ad^* is determined by

$$\begin{pmatrix} 0 & 2\lambda \frac{A}{C} F_3 & -2\lambda \frac{A}{B} F_2 \\ -2\mu \frac{B}{C} F_3 & 0 & 2\mu \frac{B}{A} F_1 \\ 2\nu \frac{C}{B} F_2 & -2\nu \frac{C}{A} F_1 & 0 \end{pmatrix}. \tag{1.3}$$

The Levi-Civita connection of g is determined by ad and ad^* via the formula

$$\nabla_X Y = \frac{1}{2}\left\{[X,Y] - (\mathrm{ad}\, X)^* Y - (\mathrm{ad}\, Y)^* X\right\}. \tag{1.4}$$

Then once the connection is known, it is straightforward to compute that the curvature tensor of g is given by

$$\begin{aligned} \langle R(X,Y)Y,X\rangle &= \frac{1}{4}\left|(\mathrm{ad}\, X)' Y + (\mathrm{ad}\, Y)' X\right|^2 - \left\langle (\mathrm{ad}\, X)' X, (\mathrm{ad}\, Y)' Y\right\rangle \\ &\quad - \frac{3}{4}\left|[X,Y]\right|^2 - \frac{1}{2}\langle [[X,Y],Y],X\rangle - \frac{1}{2}\langle [[Y,X],X],Y\rangle. \end{aligned}$$

LEMMA 1.16. *Suppose that $\{F_i\}$ is a Milnor frame for a left-invariant metric on a 3-dimensional unimodular Lie group $\mathcal{G}$. Then*

$$\langle R(F_k,F_i)F_j,F_k\rangle = 0$$

for all k and any $i \neq j$.

PROOF. Without loss of generality, we may assume that $i=1$, $j=2$, and $k=3$. Using (1.1), (1.3), and (1.4), and noting in particular that $\nabla_{F_\ell} F_\ell = -(\mathrm{ad}\, F_\ell)^* F_\ell = 0$ for any ℓ, we compute that

$$\begin{aligned} \langle R(F_3,F_1)F_2,F_3\rangle &\doteqdot \left\langle \nabla_{F_3}(\nabla_{F_1}F_2) - \nabla_{F_1}(\nabla_{F_3}F_2) - \nabla_{[F_3,F_1]}F_2,\ F_3\right\rangle \\ &= \langle \nabla_{F_3}F_2,\ \nabla_{F_1}F_3\rangle - \langle \nabla_{F_1}F_2,\ \nabla_{F_3}F_3\rangle \\ &\quad - \left\langle \nabla_{[F_3,F_1]}F_2,\ F_3\right\rangle \\ &= \langle \nabla_{F_3}F_2,\ \nabla_{F_1}F_3\rangle \\ &= \frac{1}{2}\left\langle \left(-\lambda - \mu\frac{B}{A} + \nu\frac{C}{A}\right)F_1,\ \left(-\mu - \nu\frac{C}{B} + \lambda\frac{A}{B}\right)F_2\right\rangle \\ &= 0. \end{aligned}$$

□

By the lemma, any choice of Milnor frame for a left-invariant metric g on $\mathcal{G}$ lets us globally identify both g and $\operatorname{Rc}(g)$ with diagonal matrices. Hence, we may regard the Ricci flow as a coupled system of three (rather than six) ordinary differential equations for the positive quantities A, B, and C appearing in formula (1.2), now regarded as functions of time. (Compare with Section 4 of Chapter 6.) We shall see in particular that the behavior of solutions to this system depends on the structure constants of the Lie algebra $\mathfrak{g}$. Heuristically, the evolution of the metric will in some sense try to expand the non-commutative directions. Put another way, one expects the Ricci flow to try to increase the isotropy of the initial metric.

5. The Ricci flow of a geometry with maximal isotropy SO (3)

In the remaining sections of this chapter, we will consider one example of the behavior of the Ricci flow on each isotropy class of model geometry in dimension three. An excellent study of the behavior of the normalized Ricci flow on all the homogeneous models is [**76**]. (The behavior of the unnormalized Ricci flow on model geometries was also studied from another perspective in [**86**].)

The fully isotropic geometries are $\mathcal{S}^3$, $\mathbb{R}^3$, and $\mathcal{H}^3$. We shall take $\mathcal{S}^3$ as our example, identifying it with the Lie group SU (2). Algebraically

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} w & -z \\ \bar{z} & \bar{w} \end{pmatrix} : w, z \in \mathbb{C}, \quad |w|^2 + |z|^2 = 1 \right\}.$$

Clearly, SU (2) may be identified topologically with the standard 3-sphere of radius 1 embedded in $\mathbb{R}^4$. The signature of a Milnor frame on SU (2) is given by $\lambda = \mu = \nu = -1$.

Some of the most significant differences in the behavior of the Ricci flow on the various 3-dimensional model geometries involve the issue of collapse. Recall that a Riemannian manifold $(\mathcal{M}^n, g)$ is said to be **ε-collapsed** if its injectivity radius satisfies $\operatorname{inj}(x) \leq \varepsilon$ for all $x \in \mathcal{M}^n$. Intuitively, such a manifold appears to be of lower dimension when viewed at length scales much larger than ε. A manifold $\mathcal{M}^n$ is said to **collapse** with bounded curvature if it admits a family $\{g_\varepsilon : \varepsilon > 0\}$ of Riemannian metrics such that $\sup_{x \in \mathcal{M}^n} |\operatorname{Rm}[g_\varepsilon](x)|_{g_\varepsilon}$ is bounded independently of ε, but

$$\lim_{\varepsilon \searrow 0} \left(\sup_{x \in \mathcal{M}^n} \operatorname{inj}_{g_\varepsilon}(x) \right) = 0.$$

To introduce this issue, we shall consider the behavior of the Ricci flow on SU (2) with respect to a 1-parameter family of initial data exhibiting collapse to a lower-dimensional manifold. Recall that the Hopf fibration

$$\mathcal{S}^1 \hookrightarrow \mathcal{S}^3 \to \mathcal{S}^2$$

is induced by the projection $\pi : \mathcal{S}^3 \approx \mathrm{SU}(2) \longrightarrow \mathbb{CP}^1 \approx \mathcal{S}^2$ defined by $\pi(w, z) = [w, z]$. In [**14**], Berger showed how to collapse $\mathcal{S}^3$ with bounded curvature by shrinking the circles of the Hopf fibration. The Gromov-Hausdorff limit of the collapse is (S^2, h), where h is a metric of constant sectional curvature 4. Motivated by his example, we will consider a family of left-invariant initial metrics $\{g_\varepsilon : 0 < \varepsilon \leq 1\}$ written with respect to a fixed Milnor frame $\{F_i\}$ on SU(2) in the form

$$g_\varepsilon = \varepsilon A\,\omega^1 \otimes \omega^1 + B\,\omega^2 \otimes \omega^2 + C\,\omega^3 \otimes \omega^3.$$

The sectional curvatures of g_ε are then

$$\begin{aligned}
K(F_2 \wedge F_3) &= \frac{1}{\varepsilon ABC}(B-C)^2 - 3\frac{\varepsilon A}{BC} + \frac{2}{B} + \frac{2}{C} \\
K(F_3 \wedge F_1) &= \frac{1}{\varepsilon ABC}(\varepsilon A-C)^2 - 3\frac{B}{\varepsilon AC} + \frac{2}{\varepsilon A} + \frac{2}{C} \\
K(F_1 \wedge F_2) &= \frac{1}{\varepsilon ABC}(\varepsilon A-B)^2 - 3\frac{C}{\varepsilon AB} + \frac{2}{\varepsilon A} + \frac{2}{B}.
\end{aligned}$$

Notice that the special case $A = B = C = 1$ recovers the classic ε-collapsed **Berger collapsed sphere** metric whose sectional curvatures are

$$K(F_2 \wedge F_3) = 4 - 3\varepsilon, \quad K(F_3 \wedge F_1) = \varepsilon, \quad K(F_1 \wedge F_2) = \varepsilon.$$

Note too that all geometric quantities are bounded as long as $B - C = O(\varepsilon)$ as $\varepsilon \to 0$. In particular, the Ricci tensor of g_ε is determined by the following:

$$\begin{aligned}
\operatorname{Rc}(F_1, F_1) &= \frac{2}{BC}\left[(\varepsilon A)^2 - (B-C)^2\right] \\
\operatorname{Rc}(F_2, F_2) &= \frac{2}{\varepsilon AC}\left[B^2 - (\varepsilon A - C)^2\right] = 4 - 2\frac{\varepsilon A}{C} + 2\frac{B^2 - C^2}{\varepsilon AC} \\
\operatorname{Rc}(F_3, F_3) &= \frac{2}{\varepsilon AB}\left[C^2 - (\varepsilon A - B)^2\right] = 4 - 2\frac{\varepsilon A}{B} + 2\frac{C^2 - B^2}{\varepsilon AB}
\end{aligned}$$

Our first observation is that regardless of how collapsed an initial metric g_ε may be, the Ricci flow starting at g_ε — that is, the Ricci flow of any homogeneous metric on SU(2) — shrinks to a point in finite time and becomes asymptotically round. This is equivalent to the convergence result for the normalized flow that was obtained in [**76**].

PROPOSITION 1.17. *For any $\varepsilon \in (0,1]$ and any choice of initial data $\varepsilon A_0, B_0, C_0 > 0$, the unique solution $g_\varepsilon(t)$ to (1.5) exists for a maximal finite time interval $0 \leq t < T < \infty$. The metric $g_\varepsilon(t)$ becomes asymptotically round as $t \nearrow T$.*

PROOF. Define $D \doteqdot \varepsilon A$. Then B, C, D satisfy the system

$$\begin{aligned}
\frac{d}{dt}B &= -8 + 4\frac{C^2 + D^2 - B^2}{CD} \\
\frac{d}{dt}C &= -8 + 4\frac{B^2 + D^2 - C^2}{BD} \\
\frac{d}{dt}D &= -8 + 4\frac{B^2 + C^2 - D^2}{BC}.
\end{aligned}$$

From the symmetry in these equations, we may assume without loss of generality that $D_0 \leq C_0 \leq B_0$. The computation

$$\frac{d}{dt}(B-D) = 4(B-D)\frac{C^2-(B+D)^2}{BCD}$$

shows that the inequalities $D \leq C \leq B$ persist for as long as the solution exists. Thus we may assume that $B_0 - D_0 > 0$. The estimate

$$\frac{d}{dt}B \leq -8 + 4\frac{D}{C} \leq -4,$$

implies that the solution exists only on a finite time interval; in particular, there exists $T < \infty$ such that $D \searrow 0$ as $t \nearrow T$. The estimate

$$\frac{d}{dt}\left(\frac{B-D}{D}\right) = 8\left(\frac{B-D}{D}\right)\frac{C-D-B^2}{CD} \leq 0$$

shows that $\frac{B-D}{D} \leq \delta$, where

$$\delta \doteqdot \frac{B_0 - D_0}{D_0} > 0.$$

Thus we have $0 < B - D \leq \delta D$ for $0 \leq t < T$, whence the result follows. $\square$

REMARK 1.18. When $B = C = D$ at $t = 0$, this symmetry persists. It follows easily that SU (2) remains round and shrinks at the rate

$$D(t) = D_0 - 4t = D_0 - 2(n-1)t.$$

This is the 3-dimensional version of the 'shrinking round sphere' discussed in Subsection 3.1 of Chapter 2.

Now we consider the 2-parameter family of metrics $g_\varepsilon(t)$ obtained by letting the an initial metric g_ε evolve by the Ricci flow on a maximal time interval $0 \leq t < T_\varepsilon$. We will see that the Ricci flow on SU (2) (in contrast to the homogeneous geometries we shall consider below) strongly avoids collapse. Instead of studying A, B, and C directly, it will be more convenient to introduce quantities E, F defined by

$$E \doteqdot B + C, \qquad F \doteqdot \frac{B-C}{\varepsilon}.$$

Then the Ricci flow of $g_\varepsilon(t)$ is equivalent to the following system:

$$\frac{d}{dt}A = \frac{4\varepsilon}{BC}\left(F^2 - A^2\right) \tag{1.5a}$$

$$\frac{d}{dt}E = -16 + \frac{4\varepsilon}{ABC}\left(A^2 - F^2\right)E \tag{1.5b}$$

$$\frac{d}{dt}F = \frac{4}{ABC}\left(\varepsilon A^2 - \frac{E^2}{\varepsilon}\right)F. \tag{1.5c}$$

As initial conditions, we posit $A(0,\varepsilon) = A_0 > 0$, $E(0,\varepsilon) = B_0 + C_0 > 0$, and without loss of generality $F(0,\varepsilon) = (B_0 - C_0)/\varepsilon \geq 0$. A solution of this system exists as long as $ABC \in (0,\infty)$. Our next observation is that solutions exist on a common time interval, regardless of ε.

LEMMA 1.19. *There exists a time $T_0 > 0$ depending only on the initial data (A_0, B_0, C_0) such that the solution $g_\varepsilon(t)$ of (1.5) exists for all $0 \le t < T_0$ and all $0 < \varepsilon \le 1$.*

PROOF. Since $\frac{d}{dt}(AE) = -16A < 0$, there is k_0 such that $ABC \le k_0/A$ for as long as a solution exists. So it will suffice to prove that $ABC > 0$. The calculation

$$\frac{d}{dt}\left(\frac{E}{BC}\right) = \frac{4}{AB^2C^2}\left[2AE^2 - 4ABC - \varepsilon E\left(A^2 + F^2\right)\right] \le 8\left(\frac{E}{BC}\right)^2$$

shows that there is $k_1 > 0$ depending only on B_0 and C_0 such that

$$E \le \frac{BC}{k_1 - 8t}$$

for as long as a solution exists. Then since

$$\frac{d}{dt}(BC) = 8(\varepsilon A - E) \ge -8E \ge -\frac{8}{k_1 - 8t}(BC),$$

there is $k_2 > 0$ depending only on B_0 and C_0 such that

$$BC \ge k_2(k_1 - 8t) \tag{1.6}$$

for as long as a solution exists. This implies that

$$\frac{d}{dt}A \ge -\frac{4}{BC}A^2 \ge -\frac{4}{k_2(k_1 - 8t)}A^2,$$

whence the conclusion follows readily. □

Our final observation is that there may be a jump discontinuity in the collapsing structure at $t = 0$. This reflects the fact that the Ricci flow on SU(2) assiduously avoids collapse.

LEMMA 1.20. *If $(A_0, B_0, C_0)_\varepsilon$ is a family of initial data parameterized by $\varepsilon \in (0,1)$ such that*

$$\lim_{\varepsilon \to 0} F(0, \varepsilon) > 0,$$

then for all $t \in (0, T_0)$,

$$\lim_{\varepsilon \to 0} F(t, \varepsilon) = 0.$$

PROOF. It will suffice to prove the result for $t \in [0, \tau]$, where $\tau \in (0, T_0)$ is arbitrary. Define

$$P \doteqdot \frac{4A}{BC} \qquad \text{and} \qquad Q \doteqdot \frac{4E^2}{ABC},$$

recalling that

$$\frac{d}{dt}F = \left(\varepsilon P - \frac{Q}{\varepsilon}\right)F.$$

Since

$$\frac{d}{dt}(AF) = -\frac{16F}{\varepsilon} < 0,$$

there is by (1.6) some $k_3 < \infty$ such that

$$\frac{d}{dt}F \leq PF = \frac{4}{BC}AF \leq \frac{k_3}{k_2\left(k_1 - 8t\right)}.$$

Hence there is a positive function $f_1(t)$ bounded on $[0,\tau]$ such that

$$F(t,\cdot) \leq f_1(t)$$

for all $t \in [0,\tau]$. Then since

$$\frac{d}{dt}A \leq \frac{4F^2}{BC} \leq \frac{4\left(f_1(t)\right)^2}{k_2\left(k_1 - 8t\right)},$$

there are positive functions $f_2(t)$ and $f_3(t)$ bounded on $[0,\tau]$ such that

$$A(t,\cdot) \leq f_2(t)$$

and thus

$$P(t,\cdot) \leq \frac{4f_2(t)}{k_2\left(k_1 - 8t\right)} \leq f_3(t)$$

for all $t \in [0,\tau]$. Finally, since $E^2/(BC) \geq 4$, we have

$$Q(t,\cdot) \geq \frac{16}{A} \geq \frac{16}{f_2(t)}$$

for all $t \in [0,\tau]$. Hence as $\varepsilon \searrow 0$, we have

$$\varepsilon P - \frac{Q}{\varepsilon} \to -\infty$$

uniformly on $[0,\tau]$, which proves the result. □

6. The Ricci flow of a geometry with isotropy SO (2)

Under the Ricci flow, the $\mathcal{S}^2$ factor of $\mathcal{S}^2 \times \mathbb{R}$ shrinks homothetically while the $\mathbb{R}$ factor remains unchanged. As a consequence, the solution becomes singular in finite time, and the manifold converges in the pointed Gromov–Hausdorff sense to $\mathbb{R}$. In Section 5 of Chapter 2, we will discuss Ricci flow singularities on portions of a manifold that are geometrically close to $\mathcal{S}^{n-1} \times (-\ell,\ell)$.

On the other hand, the product $\mathcal{H}^2 \times \mathbb{R}$ and the Lie groups nil and $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ all share the property that certain directions selected by the Lie algebra expand under the Ricci flow while the remaining directions converge to a finite (possibly zero) value. As a result, solutions of the normalized Ricci flow on compact representatives of the latter three geometries all collapse with bounded curvature, exhibiting Gromov–Hausdorff convergence to 1- or 2-dimensional manifolds. (See [**76**].)

We shall use nil as an example of a geometry with isotropy SO (2), because its behavior can be computed explicitly. Let $\mathcal{G}^3$ denote the 3-dimensional Heisenberg group. Algebraically, $\mathcal{G}^3$ is isomorphic to the set of upper-triangular 3×3 matrices

$$\mathcal{G}^3 \cong \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{R} \right\}$$

endowed with the usual matrix multiplication. Topologically, $\mathcal{G}^3$ is diffeomorphic to $\mathbb{R}^3$ under the map

$$\mathcal{G}^3 \ni \gamma = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mapsto (x, y, z) \in \mathbb{R}^3.$$

Under this identification, left multiplication by γ corresponds to the map

$$L_\gamma (a, b, c) = (a + x,\ b + y,\ c + xb + z).$$

The signature of a Milnor frame on $\mathcal{G}^3$ is $\lambda = -1$ and $\mu = \nu = 0$. Any left-invariant metric g on $\mathcal{G}^3$ may be written in a Milnor frame $\{F_i\}$ for g as

$$g = A\,\omega^1 \otimes \omega^1 + B\,\omega^2 \otimes \omega^2 + C\,\omega^3 \otimes \omega^3,$$

where A, B, C are positive. Then the sectional curvatures of $(\mathcal{G}^3, g)$ are

$$K (F_2 \wedge F_3) = -3\frac{A}{BC}, \quad K (F_3 \wedge F_1) = \frac{A}{BC}, \quad K (F_1 \wedge F_2) = \frac{A}{BC},$$

and its Ricci tensor is

$$\mathrm{Rc} = 2\frac{A^2}{BC}\,\omega^1 \otimes \omega^1 - 2\frac{A}{C}\,\omega^2 \otimes \omega^2 - 2\frac{A}{B}\,\omega^3 \otimes \omega^3.$$

Hence we get a solution $(\mathcal{G}^3, g(t))$ of the Ricci flow for given initial data $A_0, B_0, C_0 > 0$ if A, B, and C evolve by

$$\frac{d}{dt}A = -4\frac{A^2}{BC} \tag{1.7a}$$

$$\frac{d}{dt}B = 4\frac{A}{C} \tag{1.7b}$$

$$\frac{d}{dt}C = 4\frac{A}{B}. \tag{1.7c}$$

System (1.7) can be solved explicitly. Observing that B/C is a conserved quantity, we compute that

$$\frac{d}{dt}\left(\frac{A}{B^2}\right) = -12\left(\frac{B}{C}\right)\left(\frac{A}{B^2}\right)^2 = -12\left(\frac{B_0}{C_0}\right)\left(\frac{A}{B^2}\right)^2.$$

Solving this gives

$$\frac{A}{B^2} = \frac{C_0/B_0}{12t + B_0C_0/A_0}$$

and shows that

$$\frac{d}{dt}A = -4\left(\frac{B}{C}\right)\left(\frac{A}{B^2}\right)A = -\frac{4}{12t + B_0C_0/A_0}A.$$

Observing that AB is another conserved quantity, we can thus obtain the full solution

$$A = A_0^{2/3} B_0^{1/3} C_0^{1/3} \left(12t + B_0 C_0 / A_0\right)^{-1/3} \tag{1.8a}$$

$$B = A_0^{1/3} B_0^{2/3} C_0^{-1/3} \left(12t + B_0 C_0 / A_0\right)^{1/3} \tag{1.8b}$$

$$C = A_0^{1/3} B_0^{-1/3} C_0^{2/3} \left(12t + B_0 C_0 / A_0\right)^{1/3}. \tag{1.8c}$$

Thus we have proved the following.

PROPOSITION 1.21. *For any choice of initial data $A_0, B_0, C_0 > 0$, the unique solution of (1.7) is given by (1.8). There exist constants $0 < c_1 \leq c_2 < \infty$ depending only on the initial data such that each sectional curvature K is bounded for all $t \geq 0$ by*

$$\frac{c_1}{t} \leq K \leq \frac{c_2}{t},$$

and such that the diameter of any compact quotient $\mathcal{N}^3$ of $\mathcal{G}^3$ is bounded for all $t \geq 0$ by

$$c_1 t^{1/6} \leq \operatorname{diam} \mathcal{N}^3 \leq c_2 t^{1/6}.$$

In particular, $\mathcal{N}^3$ is almost flat.

Recall that a Riemannian manifold $(\mathcal{M}^n, g)$ is said to be ε-flat if its curvature is bounded in terms of its diameter by

$$|\mathrm{Rm}| \leq \frac{\varepsilon}{\operatorname{diam}^2 (\mathcal{M}^n, g)}. \tag{1.9}$$

One says $\mathcal{M}^n$ is **almost flat** if it is ε-flat for all $\varepsilon > 0$.

COROLLARY 1.22. *On any compact* nil-*geometry manifold, the normalized Ricci flow undergoes collapse, exhibiting Gromov–Hausdorff convergence to $\left(\mathbb{R}^2, g_{\mathrm{can}}\right)$, where g_{can} is the standard flat metric.*

7. The Ricci flow of a geometry with trivial isotropy

The only geometry with trivial isotropy is sol, which may also be regarded as the group $\operatorname{Isom}\left(\mathbb{E}_1^1\right)$ of rigid motions of Minkowski 2-space.

The signature of a Milnor frame on sol is $\lambda = -1$, $\mu = 0$, and $\nu = 1$. Any left-invariant metric g may be written in a Milnor frame $\{F_i\}$ for g as

$$g = A\,\omega^1 \otimes \omega^1 + B\,\omega^2 \otimes \omega^2 + C\,\omega^3 \otimes \omega^3.$$

The sectional curvatures are

$$K(F_2 \wedge F_3) = \frac{(A - C)^2 - 4A^2}{ABC}$$

$$K(F_3 \wedge F_1) = \frac{(A + C)^2}{ABC}$$

$$K(F_1 \wedge F_2) = \frac{(A - C)^2 - 4C^2}{ABC},$$

so the Ricci flow is equivalent to the system

$$\frac{d}{dt}A = 4\frac{C^2 - A^2}{BC} \tag{1.10a}$$

$$\frac{d}{dt}B = 4\frac{(A+C)^2}{AC} \tag{1.10b}$$

$$\frac{d}{dt}C = 4\frac{A^2 - C^2}{AB}. \tag{1.10c}$$

Observing that $AC \equiv A_0 C_0$ and $B\,(C - A) \equiv B_0\,(C_0 - A_0)$ are conserved quantities, we define $G \doteqdot A/C$ and consider the simplified system

$$\frac{d}{dt}B = 8 + 4\frac{1+G^2}{G} \tag{1.11a}$$

$$\frac{d}{dt}G = 8\frac{1-G^2}{B}, \tag{1.11b}$$

which has a solution for all positive time. Observing that $\frac{d}{dt}B \geq 16$, we conclude in particular that $B \to \infty$ as $t \to \infty$. If $A_0 = C_0$, then $G \equiv 1$ and B is a linear function of time. If not, then it is straightforward to show that G is strictly monotone and approaches a limit G_∞ satisfying either $G_0 < G_\infty \leq 1$ or $1 \leq G_\infty < G_0$. Since

$$\frac{d}{dG}\log B = \frac{1}{2G}\frac{1+G}{1-G},$$

the condition $G_\infty \neq 1$ is incompatible with the observation that $B \to \infty$. It follows easily that $A, C \to \sqrt{A_0 C_0}$ and that B grows linearly.

PROPOSITION 1.23. *For any choice of initial data $A_0, B_0, C_0 > 0$, the unique solution of (1.10) exists for all positive time. For any $\varepsilon > 0$, there exists $T_\varepsilon \geq 0$ such that*

$$\left|A - \sqrt{A_0 C_0}\right| \leq \varepsilon, \quad \left|C - \sqrt{A_0 C_0}\right| \leq \varepsilon, \quad \text{and} \quad \left|\frac{d}{dt}B - 16\right| \leq \varepsilon$$

for all $t \geq T_\varepsilon$. Moreover, there exist constants $0 < c_1 \leq c_2 < \infty$ depending only on the initial data such that each sectional curvature K is bounded for all $t \geq 0$ by

$$\frac{c_1}{t} \leq K \leq \frac{c_2}{t}.$$

COROLLARY 1.24. *On any compact* sol*-geometry manifold, the normalized Ricci flow undergoes collapse, exhibiting Gromov–Hausdorff convergence to $\mathbb{R}$.*

To gain insight into the behavior of sol metrics on compact manifolds, we recall a construction from [**67**]. One chooses $\Lambda \in \mathrm{SL}(2, \mathbb{Z})$ with eigenvalues $\lambda_+ > 1 > \lambda_-$. Then in coordinates θ, x, y on $\mathbb{R}^3$, chosen so that the x, y axes coincide with the eigenvectors of Λ, one defines an initial metric by

$$g \doteqdot e^{2\alpha}\, d\theta \otimes d\theta + e^{\beta+\gamma}\, dx \otimes dx + e^{\beta-\gamma}\, dy \otimes dy, \tag{1.12}$$

where α and β are constant and γ is a linear function of θ such that $\zeta \doteqdot e^{-\alpha}(d\gamma/d\theta) > 0$. It is not hard to check that a Milnor frame for g is given by

$$\begin{aligned} F_1 &= e^{-\gamma/2}\frac{\partial}{\partial x} - e^{\gamma/2}\frac{\partial}{\partial y} \\ F_2 &= \frac{4}{\zeta}e^{-\alpha}\frac{\partial}{\partial \theta} \\ F_3 &= e^{-\gamma/2}\frac{\partial}{\partial x} + e^{\gamma/2}\frac{\partial}{\partial y}. \end{aligned}$$

Indeed, one has $[F_1, F_2] = 2F_3$, $[F_2, F_3] = -2F_1$, $[F_3, F_1] = 0$, and can write the metric (1.12) in the form

$$g = 2e^{\beta}\,\omega^1 \otimes \omega^1 + \frac{16}{\zeta^2}\,\omega^2 \otimes \omega^2 + 2e^{\beta}\,\omega^3 \otimes \omega^3.$$

Clearly, g descends to a metric on the product of the line and the torus $\mathcal{T}^2$, and Λ acts on $\mathbb{R} \times \mathcal{T}^2$ by $(\theta, x, y) \mapsto (\theta + 2\pi, \lambda_- x, \lambda_+ y)$. If

$$W(\theta + 2\pi) = W(\theta) + 2\log\lambda_+,$$

then Λ is an isometry, whence g becomes a well defined metric on the compact mapping torus $\mathcal{N}^3_\Lambda$, regarded as a twisted $\mathcal{T}^2$ bundle over $\mathcal{S}^1$. By (1.12), it is easy to see that α governs the length of the base circle, while β and γ describe the scale and skew of the fibers, respectively. Under the Ricci flow, the metric (1.12) evolves as follows: β and γ remain fixed, while one has

$$\alpha(t) = \alpha(0) + \ln\sqrt{1 + \zeta^2 t}.$$

Intuitively, one may interpret this by saying that the Ricci flow attempts to 'untwist' the bundle $\mathcal{N}^3_\Lambda$.

Notes and commentary

Some good sources for basic 3-manifold topology are [**71**], [**78**], and [**70**]. Homogeneous Riemannian spaces are reviewed in Chapter 3 of [**27**] and Chapter 7 of [**20**]. The curvatures of left-invariant metrics on Lie groups are studied in [**98**].

Detailed analyses of the behavior of the normalized and unnormalized Ricci flows on the remaining homogeneous geometries may be found in [**76**] and [**86**], respectively. Qualitatively, one sees a variety of behaviors. For example, metrics in the $\widetilde{\mathrm{Isom}}(\mathbb{R}^2)$ family converge. Metrics in the $\mathcal{H}^3$ family converge under the normalized Ricci flow, and give rise to immortal solutions of the unnormalized flow. (See Section 4 of Chapter 2.) Under the normalized flow, compact manifolds in the $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ and $\mathcal{H}^2 \times \mathbb{R}$ families collapse to 2-dimensional limits, while compact manifolds in the $\mathcal{S}^2 \times \mathbb{R}$ and sol families collapse to 1-dimensional limits.

CHAPTER 2

Special and limit solutions

In this chapter, we continue our study of special and intuitive solutions to the Ricci flow. We introduce self-similar solutions, often called Ricci solitons, which may be regarded as generalized fixed points of the flow. We then give examples of solutions that exist for infinite time: eternal solutions (those existing for all times $-\infty < t < \infty$), ancient solutions (those which exist for times $-\infty < t < \omega$), and immortal solutions (those which exist for $\alpha < t < \infty$). Because these solutions have infinite time to diffuse, they should have very special properties. Our interest in those properties arises for the following reason. A singularity model for the Ricci flow is a complete nonflat solution obtained as a limit of dilations about a singularity. (In Chapters 8 and 9, we shall study singularity models and their formation.) Every singularity model exists for infinite time, and the special properties one expects to find in a singularity model should yield valuable information about the geometry of the original solution near the singularity just prior to its formation. In this chapter, we also study two important singularities directly: we present a rigorous analysis of neckpinch singularities (under certain symmetry hypotheses) and a heuristic analysis of a degenerate neckpinch.

1. Generalized fixed points

There is only a small class of genuine fixed points of the Ricci flow. A Riemannian manifold $(\mathcal{M}^n, g)$ is a fixed point of the unnormalized Ricci flow

$$\frac{\partial}{\partial t} g = -2 \operatorname{Rc} \tag{2.1}$$

if and only if the metric g is Ricci flat. A compact manifold $(\mathcal{M}^n, g)$ is a fixed point of the normalized Ricci flow

$$\frac{\partial}{\partial t} g = -2 \operatorname{Rc} + \frac{2}{n} \left(\frac{\int_{\mathcal{M}^n} R \, d\mu}{\int_{\mathcal{M}^n} d\mu} \right) g$$

if and only if its average scalar curvature is constant $\left(\int_{\mathcal{M}^n} R \, d\mu / \int_{\mathcal{M}^n} d\mu \right) \equiv \rho$ and

$$\operatorname{Rc} = \frac{\rho}{n} g,$$

hence if and only if g is an Einstein metric.

There is however a larger class of solutions which may be regarded as generalized fixed points. These are called **self-similar solutions** or **Ricci**

solitons. We will study Ricci solitons in detail in a chapter of the successor to this volume. But because these solutions often arise as limits of dilations about singularities, we shall encounter some examples in this chapter. Hence we provide a brief introduction to them here.

Suppose that $(\mathcal{M}^n, g(t))$ is a solution of the unnormalized Ricci flow on a time interval (α, ω) containing 0, and set $g_0 = g(0)$. One says $g(t)$ is a **self-similar solution** of the Ricci flow if there exist scalars $\sigma(t)$ and diffeomorphisms ψ_t of $\mathcal{M}^n$ such that

$$g(t) = \sigma(t)\ \psi_t^*(g_0) \tag{2.2}$$

for all $t \in (\alpha, \omega)$. Thus a self-similar solution is a solution of the evolution equation (2.1) having the special form (2.2). A metric having this form changes only by diffeomorphism and rescaling. To see the significance of this fact, let us denote by $S_2\mathcal{M}^n$ the bundle of symmetric $(2,0)$-tensors on $\mathcal{M}^n$ and by $S_2^+\mathcal{M}^n$ the sub-bundle of positive-definite tensors. Then the space of Riemannian metrics on $\mathcal{M}^n$ may be written as $\mathfrak{M}(\mathcal{M}^n) \doteqdot C^\infty(\mathcal{M}^n, S_2^+\mathcal{M}^n)$. Let $\mathfrak{D}(\mathcal{M}^n)$ denote the diffeomorphism group of $\mathcal{M}^n$. Because the Ricci flow is invariant under diffeomorphism, it may be regarded as a dynamical system on the moduli space

$$\mathfrak{M}(\mathcal{M}^n)/\mathfrak{D}(\mathcal{M}^n)$$

of Riemannian metrics modulo diffeomorphisms. In this context, Ricci solitons correspond to generalized fixed points.

On the other hand, suppose that $(\mathcal{M}^n, g_0)$ is a fixed Riemannian manifold such that the identity

$$-2\operatorname{Rc}(g_0) = \mathcal{L}_X g_0 + 2\lambda g_0 \tag{2.3}$$

holds for some constant λ and some complete vector field X on $\mathcal{M}^n$. In this case, we say g_0 is a **Ricci soliton**. If X vanishes identically, a Ricci soliton is simply an Einstein metric. Consequently, any solution of (2.3) (which is actually a coupled elliptic system for g_0 and X) may be regarded as a generalization of an Einstein metric.

REMARK 2.1. By rescaling, one may assume that $\lambda \in \{-1, 0, 1\}$ in (2.3). These three cases correspond to **shrinking**, **steady**, and **expanding** solitons, respectively. And as we shall see below, these yield examples of ancient, eternal, and immortal solutions, respectively.

REMARK 2.2. In case the vector field X appearing in (2.3) is the gradient field of a potential function $-f$, one has

$$\nabla\nabla f = \operatorname{Rc} + \lambda g$$

and says g_0 is a **gradient Ricci soliton**.

A vector field X that makes a Riemannian metric into a Ricci soliton may not be unique. The next example illustrates this.

EXAMPLE 2.3. Let $(\mathbb{R}^n, g_{\text{can}})$ denote Euclidean space with its standard metric. Since the metric is flat, the trivial solution of (2.3) obtained by choosing $X \equiv 0$ lets us regard this as a steady Ricci soliton. But it may also be regarded as an expanding gradient Ricci soliton, called the **Gaussian soliton**, by taking $\lambda = 1$ and choosing the potential function

$$f(x) = \frac{1}{2}|x|^2.$$

In this section, we will prove the following observation.

LEMMA 2.4. *If $(\mathcal{M}^n, g(t))$ is a solution of the Ricci flow (2.1) having the special form (2.2), then there exists a vector field X on $\mathcal{M}^n$ such that $(\mathcal{M}^n, g_0, X)$ solves (2.3). Conversely, given any solution $(\mathcal{M}^n, g_0, X)$ of (2.3), there exist 1-parameter families of scalars $\sigma(t)$ and diffeomorphisms ψ_t of $\mathcal{M}^n$ such that $(\mathcal{M}^n, g(t))$ becomes a solution of the Ricci flow (2.1) when $g(t)$ is defined by (2.2).*

In words, this result says that there is a bijection between the families of self-similar solutions and Ricci solitons which allows us to regard the concepts as equivalent.

PROOF OF LEMMA 2.4. First suppose that $(\mathcal{M}^n, g(t))$ is a solution of the Ricci flow having the form (2.2). We may assume without loss of generality that $\sigma(0) = 1$ and $\psi_0 = \text{id}$. Then we have

$$-2\operatorname{Rc}(g_0) = \frac{\partial}{\partial t}g(t)\Big|_{t=0} = \sigma'(0)\, g_0 + \mathcal{L}_{Y(0)}g_0,$$

where $Y(t)$ is the family of vector fields generating the diffeomorphisms ψ_t. This implies that g_0 satisfies (2.3) with $\lambda = \frac{1}{2}\sigma'(0)$ and $X = Y(0)$.

Conversely, suppose that g_0 satisfies (2.3). Define

$$\sigma(t) \doteqdot 1 + 2\lambda t,$$

and define a 1-parameter family of vector fields Y_t on $\mathcal{M}^n$ by

$$Y_t(x) \doteqdot \frac{1}{\sigma(t)}X(x).$$

Let ψ_t denote the diffeomorphisms generated by the family Y_t, where $\psi_0 = \text{id}_{\mathcal{M}^n}$, and define a smooth 1-parameter family of metrics on $\mathcal{M}^n$ by

$$g(t) \doteqdot \sigma(t) \cdot \psi_t^*(g_0).$$

Then $g(t)$ has the special form (2.2). The computation

$$\frac{\partial}{\partial t}g = \frac{d\sigma}{dt} \cdot \psi_t^*(g_0) + \sigma(t) \cdot \psi_t^*(\mathcal{L}_{Y_t}g_0) = \psi_t^*(2\lambda g_0 + \mathcal{L}_X g_0).$$

implies by (2.3) that

$$\frac{\partial}{\partial t}g = \psi_t^*(-2\operatorname{Rc}[g_0]) = -2\operatorname{Rc}[g],$$

hence that $g(t)$ is a solution of the Ricci flow. □

2. Eternal solutions

An **eternal solution** of the Ricci flow is one that exists for all time. Such solutions should have very special properties. Intuitively, one may arrive at this expectation as follows. We shall see in Chapters 5 and 6 that the curvatures of a solution to the Ricci flow evolve by reaction-diffusion equations. Equations of this type involve a competition between the diffusion term (which seeks to disperse concentrations of curvature uniformly over the manifold as time moves forward) and the reaction term (which tends to create concentrations of curvature as time moves forward). On an eternal solution, one can look arbitrarily far backwards or forwards in time without encountering any concentration phenomena. This means that such solutions must be very stable, with no concentrations of curvature at any finite time in either the past or the future.

2.1. The cigar soliton. Hamilton's **cigar soliton** is the complete Riemannian surface $(\mathbb{R}^2, g_\Sigma)$, where

$$g_\Sigma = \frac{dx \otimes dx + dy \otimes dy}{1 + x^2 + y^2}. \tag{2.4}$$

This manifold is also known in the physics literature as **Witten's black hole**. (See [**125**].) As we shall see below, it is a steady soliton, hence corresponds to an eternal self-similar solution. Recalling that the Christoffel symbols Γ_{ij}^k of a metric g are given by

$$\Gamma_{ij}^k = \frac{1}{2} g^{k\ell} \left(\frac{\partial}{\partial x^i} g_{j\ell} + \frac{\partial}{\partial x^j} g_{i\ell} - \frac{\partial}{\partial x^\ell} g_{ij} \right),$$

it is easy to compute that the Christoffel symbols of g_Σ with respect to the coordinate system $(z^1 = x,\ z^2 = y)$ on $\mathbb{R}^2$are

$$\begin{array}{lll} \Gamma_{11}^1 = -x/\left(1+r^2\right) & \Gamma_{12}^1 = -y/\left(1+r^2\right) & \Gamma_{22}^1 = x/\left(1+r^2\right) \\ \Gamma_{11}^2 = y/\left(1+r^2\right) & \Gamma_{12}^2 = -x/\left(1+r^2\right) & \Gamma_{22}^2 = -y/\left(1+r^2\right) \end{array},$$

where $r \doteqdot \sqrt{x^2 + y^2}$. Because g_Σ is rotationally symmetric, it is natural to write it in polar coordinates as

$$g_\Sigma = \frac{dr^2 + r^2\, d\theta^2}{1 + r^2}. \tag{2.5}$$

The scalar curvature of g_Σ is

$$R_\Sigma = \frac{4}{1 + r^2}, \tag{2.6}$$

and its area form is

$$d\mu_\Sigma = \frac{1}{1 + r^2}\, dx \wedge dy.$$

Since $r^2/\left(1 + r^2\right) \to 1$ as $r \to \infty$, equations (2.5) and (2.6) show that the metric is asymptotic at infinity to a cylinder of radius 1. In fact, we shall see below that this convergence happens exponentially fast in the distance scale induced by the metric. One can also see the cigar's asymptotic approach to the cylinder in another way: the geodesics in the cigar metric are exactly those curves $\gamma(t) = (r(t), \theta(t))$ given in polar coordinates by solutions to the ODE system

$$\theta'' + \frac{2}{r(1+r^2)}\theta' r' = 0$$
$$r'' - \frac{r}{1+r^2}\left[(r')^2 + (\theta')^2\right] = 0.$$

From this, it is not hard to see that $\theta(t)$ can be written in the form

$$\theta(t) = a + bt + b\int^t \frac{d\tau}{r^2(\tau)}$$

where $a, b \in \mathbb{R}$ are arbitrary constants. In particular, when $r_0 = r(0)$ is large, a Euclidean circle of radius r_0 in $\mathbb{R}^2$ is close to being geodesic for a short time.

REMARK 2.5. The cigar soliton is actually a Kähler metric

$$g_\Sigma = \frac{dz\, d\bar{z}}{1+|z|^2}$$

on $\mathbb{C} \approx \mathbb{R}^2$, and hence a Kähler-Ricci soliton. In fact, the cigar is the simplest representative of an entire family of Kähler-Ricci solitons that exist on $\mathbb{C}^{2m}$ for all $m \in \mathbb{N}$ and on certain other complex manifolds. Various other metrics in this family were introduced in the papers [**88**], [**23, 24**], and [**40**]. See also [**77**].

REMARK 2.6. The cigar metric on the topological cylinder

$$\mathcal{C} = \{(x,y) : x \in \mathbb{R} \text{ and } y \in \mathbb{R}/\mathbb{Z}\}$$

is

$$\frac{dx^2 + dy^2}{1+e^{-2x}}.$$

The details are left as an easy exercise for the reader.

To write the cigar metric in its natural coordinates on $\mathbb{R}^2$, define

$$s \doteqdot \operatorname{arcsinh} r = \log\left(r + \sqrt{1+r^2}\right).$$

It is easy to see that s represents the metric distance from the origin. Indeed, since

$$dr = \cosh s\, ds = \sqrt{1+r^2}\, ds,$$

formula (2.5) becomes

$$g_\Sigma = ds^2 + \tanh^2 s\, d\theta^2. \tag{2.7}$$

Then formula (2.6) becomes

$$R_\Sigma = \frac{4}{1+r^2} = \frac{4}{\cosh^2 s} = \frac{16}{(e^s + e^{-s})^2}.$$

In particular, $R_\Sigma = O\left(e^{-2s}\right)$ as $s \to \infty$, which illustrates the exponential decay claimed above.

We will now show that the cigar is essentially unique. The argument will also construct the potential function, hence the gradient vector field that makes g_Σ into a soliton.

LEMMA 2.7. *Up to homothety, the cigar is the unique rotationally symmetric gradient Ricci soliton of positive curvature on $\mathbb{R}^2$.*

PROOF. Without loss of generality, one may write any rotationally symmetric metric g on $\mathbb{R}^2$ in the form

$$g = ds^2 + \varphi(s)^2\, d\theta^2, \tag{2.8}$$

where $\varphi(s)$ is a strictly positive function. (Compare with equation (2.7) above.) We shall calculate using moving frames, as will be reviewed in Section 1 of Chapter 5. It is natural to write

$$g = \delta_{ij}\,\omega^i \otimes \omega^j$$

in terms of the orthonormal coframe field $\left\{\omega^1, \omega^2\right\}$ given by

$$\omega^1 = ds, \qquad \omega^2 = \varphi(s)\, d\theta.$$

The orthonormal frame field $\{e_1, e_2\}$ dual to $\left\{\omega^1, \omega^2\right\}$ is then given by

$$e_1 = \frac{\partial}{\partial s}, \qquad e_2 = \frac{1}{\varphi(s)} \cdot \frac{\partial}{\partial \theta}.$$

It is straightforward to check that for any vector field

$$X = X^1 \frac{\partial}{\partial s} + X^2 \frac{\partial}{\partial \theta}$$

on $\mathbb{R}^2$, one has

$$\begin{aligned} \nabla_X e_1 &= \frac{\varphi'(s)}{\varphi(s)} X^2 \frac{\partial}{\partial \theta} = \varphi'(s)\, d\theta(X) \cdot e_2 \\ \nabla_X e_2 &= -\varphi'(s)\, X^2 \frac{\partial}{\partial s} = -\varphi'(s)\, d\theta(X) \cdot e_1. \end{aligned}$$

It follows that the Levi-Civita connection 1-forms $\{\omega_i^j\}$ defined with respect to $\{e_1, e_2\}$ by

$$\nabla_X e_i = \omega_i^j(X) \cdot e_j$$

have only two nonzero components:

$$\omega_1^2 = -\omega_2^1 = \varphi'(s)\, d\theta. \tag{2.9}$$

In general, Cartan's first and second structural equations are

$$d\omega^i = \omega^j \wedge \omega_j^i \qquad \text{and} \qquad \Omega_i^j = d\omega_i^j - \omega_i^k \wedge \omega_k^j.$$

In the case at hand, we have $d\omega^1 = 0$, $d\omega^2 = \frac{\varphi'(s)}{\varphi(s)}\,\omega^1 \wedge \omega^2$, and

$$\Omega_1^2 = d\omega_1^2 = \varphi''(s)\, ds \wedge d\theta = \frac{\varphi''(s)}{\varphi(s)}\,\omega^1 \wedge \omega^2.$$

So the Gauss curvature K is given by

$$K = \Omega_1^2 (e_2, e_1) = -\frac{\varphi''(s)}{\varphi(s)}. \tag{2.10}$$

Recalling equation (2.3), Remark 2.1, and Remark 2.2, we see that g constitutes a steady gradient Ricci soliton if and only if there exists a function f such that

$$Kg = \nabla\nabla f. \tag{2.11}$$

Our assumptions are that $K > 0$ and that f is a radial function $f = f(s)$. This implies in particular that f is a convex function. Recalling (2.9), we observe that

$$\nabla_{e_1} e_1 = \omega_1^2 (e_1) \cdot e_2 = 0$$

and

$$\nabla_{e_2} e_2 = \omega_2^1 (e_2) \cdot e_1 = -\frac{\varphi'(s)}{\varphi(s)} e_1.$$

Since $g = \delta$ with respect to the frame $\{e_1, e_2\}$, it follows that (2.11) holds if and only if both

$$K = e_1 (e_1 f) - (\nabla_{e_1} e_1) f = f''(s)$$

and

$$K = e_2 (e_2 f) - (\nabla_{e_2} e_2) f = \frac{\varphi'(s) f'(s)}{\varphi(s)}$$

hold, hence by (2.10) if and only if

$$-\frac{\varphi''(s)}{\varphi(s)} = f''(s) = \frac{\varphi'(s) f'(s)}{\varphi(s)}. \tag{2.12}$$

Solving the separable ODE $f''/f' = \varphi'/\varphi$ implied by the second equality in (2.12) implies that

$$f'(s) = \alpha\varphi(s) \tag{2.13}$$

for some constant α. We cannot have $\alpha = 0$ or else g would be flat. Since $\varphi > 0$, this forces f' to have a sign. Since f is convex, this sign must be positive; so we may set $\alpha = 2a^2$ for some $a \neq 0$. Then substituting $f' = 2a^2\varphi$ into the first equality in (2.12) yields the ODE

$$\varphi'' + 2a^2\varphi\varphi' = 0.$$

Integrating this gives

$$\varphi' + a^2\varphi^2 = b$$

for some constant b. In order to g to extend to a smooth metric at $s = 0$, we must have

$$\varphi(0) = 0 \qquad \text{and} \qquad \varphi'(0) = 1$$

by Lemma 2.10 (below). This is possible if and only if $b = 1$. So we solve the linear ODE $\varphi' + a^2\varphi^2 = 1$ to obtain

$$\varphi(s) = \frac{1}{a} \tanh(as).$$

Hence by (2.8), g must have the form

$$g = ds^2 + \frac{1}{a^2}\tanh^2(as)\, d\theta^2$$

for some $a > 0$. We claim that g is nothing other than a constant multiple of the cigar soliton. To see this, make the substitution $\sigma = as$ to get

$$g = \frac{1}{a^2}\left(d\sigma^2 + \tanh^2\sigma\, d\theta^2\right) = \frac{1}{a^2} g_\Sigma.$$

□

REMARK 2.8. By (2.13), we have $f' = a\varphi$ and hence

$$f(s) = \frac{1}{a}\log(\cosh(as)).$$

In particular

$$X = -\operatorname{grad} f = -a\varphi \cdot e_1 = -\tanh(as)\cdot\frac{\partial}{\partial s}.$$

EXERCISE 2.9. Find the analog of the cigar metric with curvature $K < 0$.

3. Ancient solutions

An **ancient solution** of the Ricci flow is one which exists on a maximal time interval $-\infty < t < \omega$, where $\omega < \infty$.

3.1. The round sphere. The shrinking round sphere is in a sense the canonical ancient solution of the Ricci flow. Let g_{can} denote the standard round metric on $\mathcal{S}^n$ of radius 1, and consider the 1-parameter family of conformally equivalent metrics

$$g(t) \doteqdot r(t)^2\, g_{\mathrm{can}},$$

where $r(t)$ is to be determined. Observe that $g(t)$ is a solution of the Ricci flow if and only if

$$2r\frac{dr}{dt}\cdot g_{\mathrm{can}} = \frac{\partial}{\partial t}g = -2\operatorname{Rc}[g] = -2\operatorname{Rc}[g_{\mathrm{can}}] = -2(n-1)\, g_{\mathrm{can}},$$

hence if and only if $r(t)$ is a solution of the ODE

$$\frac{dr}{dt} = -\frac{n-1}{r}.$$

Evidently, setting

$$r(t) = \sqrt{r_0^2 - 2(n-1)t} = \sqrt{2(n-1)}\cdot\sqrt{T-t} \tag{2.14}$$

yields an ancient solution $(\mathcal{S}^n, g(t))$ of the Ricci flow that exists for the time interval $-\infty < t < T$, where $T < \infty$ is the singularity time defined by

$$T \doteqdot \frac{r_0^2}{2(n-1)}.$$

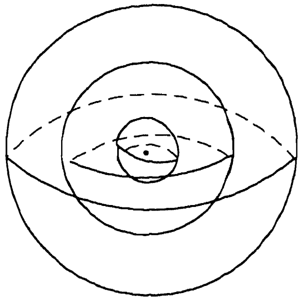

FIGURE 1. Shrinking round n-sphere

3.2. The cylinder-to-sphere rule. Before considering our next example, it will be helpful to recall the following result from Riemannian geometry.

LEMMA 2.10. *Let $0 < \omega \leq \infty$, and let g be a metric on the topological cylinder $(-\omega, \omega) \times S^n$ of the form*

$$g = \varphi(z)^2 \, dz^2 + \psi(z)^2 \, g_{\mathrm{can}},$$

where $\varphi, \psi : (-\omega, \omega) \to \mathbb{R}_+$ and g_{can} is the canonical round metric of radius 1 on S^n. Then g extends to a smooth metric on S^{n+1} if and only if

$$\int_{-\omega}^{\omega} \varphi(\zeta) \, d\zeta < \infty, \tag{2.15}$$

$$\lim_{z \to \pm\omega} \psi(z) = 0, \tag{2.16}$$

$$\lim_{z \to \pm\omega} \frac{\psi'(z)}{\varphi(z)} = \mp 1, \tag{2.17}$$

and

$$\lim_{z \to \pm\omega} \frac{d^{2k}\psi}{ds^{2k}}(z) = 0 \tag{2.18}$$

for all $k \in \mathbb{N}$, where ds is the element of arc length induced by φ.

PROOF. Since the result is standard, we will only prove sufficiency. The argument is essentially two-dimensional, so to simplify the notation we shall assume that $n = 1$ and

$$g = \varphi(z)^2 \, dz^2 + \psi(z)^2 \, d\theta^2,$$

where $\theta \in \mathbb{R}/2\pi\mathbb{Z}$. It will suffice to consider the 'north pole' $z = \omega$. Let

$$s(z) \doteqdot \int_0^z \varphi(\zeta) \, d\zeta,$$

and set $\bar{s} \doteqdot \lim_{z \to \omega} s(z) < \infty$. Let

$$r(z) \doteqdot \bar{s} - s,$$

and define

$$\rho(r) \equiv \rho(r(z)) \doteqdot \frac{\psi(z)}{r}.$$

Then the metric may be written as

$$g = dr^2 + r^2\rho^2\, d\theta^2.$$

Introduce 'Cartesian coordinates'

$$x \doteqdot r\cos\theta,$$
$$y \doteqdot r\sin\theta.$$

Then $r^2 = x^2 + y^2$ and we have

$$dx = \frac{x}{r}\, dr - y\, d\theta$$
$$dy = \frac{y}{r}\, dr + x\, d\theta,$$

whence

$$dr = \frac{x}{r}\, dx + \frac{y}{r}\, dy$$
$$d\theta = -\frac{y}{r^2}\, dx + \frac{x}{r^2}\, dy.$$

A simple calculation shows that

$$g = \left(1 + \frac{\rho^2 - 1}{r^2} y^2\right) dx^2 - \left(2xy\frac{\rho^2 - 1}{r^2}\right) dx\, dy + \left(1 + \frac{\rho^2 - 1}{r^2} x^2\right) dy^2.$$

So g extends to a smooth metric if

$$f(r) \doteqdot \frac{\rho(r)^2 - 1}{r^2}$$

extends to an even function which is smooth at $r = 0$. By Taylor's theorem,

$$f(r) = \frac{\rho(0) - 1}{r^2} + \frac{2\rho(0)\rho'(0)}{r} + \sum_{k=2}^{K} \left.\frac{d^k(\rho^2)}{dr^k}\right|_{r=0} \frac{r^{k-2}}{k!} + O\left(r^{K-1}\right).$$

Hence f yields a smooth even function at zero if and only if $\rho(0) = 1$, $\rho'(0) = 0$, and all odd derivatives of ρ^2 vanish at the origin. Noting that

$$\frac{d^k(\rho^2)}{dr^k} = 2\rho\rho^{(k)} + \sum_{j=1}^{k-1} c_{jk}\rho^{(j)}\rho^{(k-j)}$$

for appropriate integers c_{jk}, we conclude by induction that f yields a smooth even function at zero if and only if $\rho(0) = 1$ and $\rho^{(2k+1)}(0) = 0$ for all $k \geq 0$. To translate back to $\psi(z) = r\rho(r)$, we simply observe that

$$\frac{d^k\psi}{dr^k} = k\frac{d^{k-1}\rho}{dr^{k-1}} + r\frac{d^k\rho}{dr^k}.$$

In particular, the observation

$$-\frac{1}{\varphi}\frac{d\psi}{dz} = \frac{d\psi}{dr} = \rho + r\frac{d\rho}{dr},$$

shows $\rho(0) = 1$ and $\rho'(0) = 0$ if and only if

$$\lim_{z \to \omega} \psi(z) = 0 \qquad \text{and} \qquad \lim_{z \to \omega} \frac{\psi'(z)}{\varphi(z)} = -1.$$

$\square$

3.3. The Rosenau solution. In Lemma 5.3, we will show that if g and h are metrics on a surface $\mathcal{M}^2$ conformally related by $g = e^v h$, then the scalar curvature R_g of g is related to the scalar curvature R_h of h by $R_g = e^{-v}(-\Delta_h v + R_h)$. In the case that

$$g = u \cdot h$$

for $u : \mathcal{M}^2 \to \mathbb{R}_+$, this formula becomes

$$R_g = \frac{-\Delta_h(\log u) + R_h}{u}. \tag{2.19}$$

Now if u is allowed to depend on time but h is regarded as fixed, g will satisfy the Ricci flow equation for a surface,

$$\frac{\partial}{\partial t} g = -R \cdot g,$$

if and only if

$$\frac{\partial u}{\partial t} \cdot h = \left(\frac{\Delta_h \log u - R_h}{u} \right) g,$$

hence if and only if u evolves by

$$\frac{\partial u}{\partial t} = \Delta_h \log u - R_h.$$

In the special case that $(\mathcal{M}^2, h)$ is flat, this reduces to

$$\frac{\partial u}{\partial t} = \Delta_h \log u. \tag{2.20}$$

REMARK 2.11. There is an interesting connection between equation (2.20) and the **porous media flow**, which is the PDE

$$\frac{\partial}{\partial \tau} u = \Delta u^m. \tag{2.21}$$

When $m = 1$, this is just the ordinary linear heat equation. For all other $m > 0$, it is nonlinear. When $0 < m < 1$, equation (2.21) appears in various models in plasma physics. When $m > 1$, it models ionized gases at high temperatures; and when $m > 2$, it models ideal gases flowing isentropically in a homogeneous porous medium. (See [**8**].) If $t \doteqdot m\tau$, then equation (2.21) becomes

$$\frac{\partial}{\partial t} u = \frac{1}{m} \Delta u^m.$$

In the formal limit $m = 0$, this equation may be applied to model the thickness $u > 0$ of a thin lubricating film if one neglects certain fourth-order effects. (See [**19**].) Since

$$\lim_{m \searrow 0} \frac{u^m - 1}{m} = \log u$$

for $u > 0$, the calculation

$$\begin{aligned} \lim_{m \searrow 0} \Delta \left(\frac{u^m - 1}{m} \right) &= \lim_{m \searrow 0} \left[u^{m-1} \Delta u + (m-1)\, u^{m-2} \left| \nabla u \right|^2 \right] \\ &= \frac{\Delta u}{u} - \frac{\left| \nabla u \right|^2}{u^2} = \Delta \log u \end{aligned}$$

shows that (2.20) is exactly the limit obtained for $m = 0$. This connection between the porous media flow and the Ricci flow in dimension $n = 2$ was first made by Sigurd Angenent. (See [**128, 129**].)

Now let h be the flat metric on the manifold $\mathcal{M}^2 = \mathbb{R} \times \mathcal{S}^1_1$, where $\mathcal{S}^1_1$ is the circle of radius 1. Give $\mathcal{M}^2$ coordinates $x \in \mathbb{R}$ and $\theta \in \mathcal{S}^1_1 = \mathbb{R}/2\pi\mathbb{Z}$. The **Rosenau solution** [**111**] of the Ricci flow is the metric $g = u \cdot h$ defined for $t < 0$ by

$$u(x,t) = \frac{2\beta \sinh(-\alpha\lambda t)}{\cosh \alpha x + \cosh \alpha \lambda t} \tag{2.22}$$

for parameters where $\alpha, \beta, \lambda > 0$ to be determined. Because u is independent of θ, we have

$$\Delta_h \log u = \frac{\partial^2}{\partial x^2} \log u.$$

Hence the computations

$$\frac{\partial}{\partial t} u(x,t) = -2\alpha\beta\lambda \frac{\cosh \alpha\lambda t \cdot \cosh \alpha x + 1}{(\cosh \alpha x + \cosh \alpha \lambda t)^2} \tag{2.23}$$

and

$$\frac{\partial^2}{\partial x^2} \log u(x,t) = -\alpha^2 \frac{\cosh \alpha\lambda t \cdot \cosh \alpha x + 1}{(\cosh \alpha x + \cosh \alpha \lambda t)^2} \tag{2.24}$$

show that u satisfies (2.20) (hence that $g = uh$ solves the Ricci flow on $\mathcal{M}^2$) if and only if

$$2\beta\lambda = \alpha. \tag{2.25}$$

Note that the Rosenau solution is ancient but not eternal, since by equation (2.22),

$$\lim_{t \nearrow 0} u(x,t) = 0.$$

By (2.19), the scalar curvature of g is given by

$$R[g(t)] = -\frac{\Delta_h \log u}{u} = \frac{\alpha^2}{2\beta} \frac{\cosh \alpha\lambda t \cdot \cosh \alpha x + 1}{\sinh(-\alpha\lambda t)(\cosh \alpha x + \cosh \alpha \lambda t)}. \tag{2.26}$$

In particular, the solution has positive curvature for as long as it exists.

Now regard $\mathcal{M}^2$ as the 2-sphere $\mathcal{S}^2$ punctured at the poles:

$$\mathbb{R}\times\mathcal{S}^1 \approx \mathcal{S}^2 \setminus \{ \text{ north and south poles } \}.$$

Since $g(t)$ has positive scalar curvature for all $t < 0$, it is reasonable to investigate under what conditions $g(t)$ extends to a smooth solution of the Ricci flow on $\mathcal{S}^2$. We want to apply Lemma 2.10 with $z = x$ and $\varphi = \psi = \sqrt{u}$. Letting $s(x,t)$ denote the distance from the 'equator' $x = 0$ to the point $x \in \mathbb{R}$ at time t, we estimate

$$\begin{aligned} s(x,t) &= \sqrt{2\beta \sinh(-\alpha\lambda t)} \int_0^x \frac{1}{\sqrt{\cosh \alpha x + \cosh \alpha \lambda t}}\, dx \\ &\leq 2\sqrt{\beta \sinh(-\alpha\lambda t)} \int_0^{|x|} e^{-\frac{\alpha}{2}|x|}\, dx. \end{aligned}$$

Hence the distance to the 'poles' $x = \pm\infty$ is bounded at all times $t < 0$ by $\frac{4}{\alpha}\sqrt{\beta \sinh(-\alpha\lambda t)} < \infty$. This shows that (2.15) is satisfied. It is easy to check that for all $t < 0$, we have

$$\lim_{|x|\to\infty} u(x,t) = 0,$$

hence that (2.16) is satisfied. Since

$$\lim_{x\to\pm\infty} \left(\frac{\frac{\partial}{\partial x}\sqrt{u(x,t)}}{\sqrt{u(x,t)}} \right) = -\alpha \lim_{x\to\pm\infty} \left(\frac{\sinh \alpha x}{\cosh \alpha x + \cosh \alpha\lambda t} \right) = \mp\alpha,$$

we see that (2.17) is satisfied if and only if $\alpha = 1$. Finally, we compute

$$\frac{\partial}{\partial s}\sqrt{u} = \frac{1}{\sqrt{u}}\frac{\partial}{\partial x}\sqrt{u} = \frac{1}{2}\frac{\partial}{\partial x}\log u = -\frac{\alpha \sinh \alpha x}{2(\cosh \alpha x + \cosh \alpha\lambda t)}$$

and recall (2.24) to see that

$$\frac{\partial^2}{\partial s^2}\sqrt{u} = \frac{1}{2\sqrt{u}}\frac{\partial^2}{\partial x^2}\log u = -\frac{\alpha^2}{2\sqrt{2\beta \sinh(-\alpha\lambda t)}} \frac{\cosh \alpha\lambda t \cdot \cosh \alpha x + 1}{(\cosh \alpha x + \cosh \alpha\lambda t)^{3/2}}.$$

It follows that (2.18) is satisfied for $k = 1$; higher derivatives may be estimated similarly. Applying Lemma 2.10, we get the following result.

LEMMA 2.12. *The metric defined by (2.22) for $t < 0$ extends to an ancient solution of the Ricci flow on $\mathcal{S}^2$ if and only if*

$$\alpha = 1 \qquad \textit{and} \qquad \beta = \frac{1}{2\lambda}. \tag{2.27}$$

From now on, we assume (2.27) holds. Then it follows from (2.26) that the scalar curvature $R_{\pm\infty}(t)$ at the 'poles' $x = \pm\infty$ is strictly positive for all times $t < 0$:

$$\begin{aligned} R_{\pm\infty}(t) &\doteqdot \lim_{|x|\to\infty} R = \lim_{|x|\to\infty} \frac{\lambda(\cosh \lambda t \cdot \cosh x + 1)}{\sinh(-\lambda t)(\cosh x + \cosh \lambda t)} \\ &= \lambda \coth(-\lambda t) > 0. \end{aligned}$$

Moreover, the curvature $R_{\pm\infty}(t)$ at the poles is actually the maximum curvature of $(\mathcal{S}^2, g(t))$ for all $t < 0$, since for all $x > 0$ we have

$$\begin{aligned} \frac{\partial}{\partial x} R &= \frac{\partial}{\partial x} \left(\frac{\lambda (\cosh \lambda t \cdot \cosh x + 1)}{\sinh(-\lambda t)(\cosh x + \cosh \lambda t)} \right) \\ &= \frac{\sinh x \cdot \sinh(-\lambda t)}{(\cosh x + \cosh \lambda t)^2} > 0. \end{aligned}$$

If we compare the scalar curvature at an arbitrary point with its maximum, we find that

$$\lim_{t \nearrow 0} \frac{R(t)}{R_{\pm\infty}(t)} = \lim_{t \nearrow 0} \frac{(\cosh \lambda t \cdot \cosh x + 1)}{\cosh(\lambda t)(\cosh x + \cosh \lambda t)} = 1.$$

This confirms a result equivalent to what we shall prove in Chapter 5: any solution of the unnormalized Ricci flow on a topological $\mathcal{S}^2$ shrinks to a round point in finite time.

EXERCISE 2.13. The Rosenau solution provides an example of some relevance to our study of singularity models in Chapter 8: it is a Type II ancient solution which gives rise to an eternal solution if we take a limit looking infinitely far back in time (as described in Section 6 of that chapter). In particular, there is a theorem which says that if one takes a limit of the Rosenau solution at either pole $x = \pm\infty$ as $t \to -\infty$, one gets a copy of the cigar soliton studied above. Demonstrate this explicitly.

REMARK 2.14. The Rosenau solution is of interest here in part because it could potentially occur as a dimension-reduction limit of a 3-manifold singularity. (Techniques of forming limits at singularities by parabolic dilation are presented in Section 3 of Chapter 8. The method of dimension reduction is introduced in Section 4 of Chapter 9 and will be discussed further in the successor to this volume.) Recent work of Perelman [**105**] eliminates this possibility for finite-time singularities. Indeed, if the Rosenau solution occurred as a dimension-reduction limit, one could by Exercise 2.13 adjust the sequence of points and times about which one dilated in order to get a cigar limit. Perelman's *No Local Collapsing* Theorem excludes this possibility.

4. Immortal solutions

An **immortal solution** of the Ricci flow is one which exists on a maximal time interval $\alpha < t < \infty$, where $\alpha > -\infty$. The example below appears in Appendix A of [**55**]. It yields a self-similar solution on $\mathbb{R}^2$ which expands from a cone with cone angle $2\pi\zeta$, where $0 < \zeta < 1$. For $t > 0$, it forms a smooth complete metric on $\mathbb{R}^2$ with positive curvature that decays exponentially with respect to the distance from the origin. The reader is invited to compare this solution with the expanding Kähler–Ricci solitons on $\mathbb{C}^n$ described in [**24**] that converge to a cone as $t \searrow 0$. (See also [**40**].)

Consider the one-parameter family of expanding metrics $g(t)$ defined for $t > 0$ in the polar coordinate system (r, θ) on $\mathbb{R}^2 \backslash \{$ any ray $\}$ by

$$g(t) = t\left(f(r)^2\, dr^2 + r^2\, d\theta^2\right), \tag{2.28}$$

where f is a positive function to be determined. We will study the metrics $g(t)$ in local coordinates $\left(z^1 = r,\ z^2 = \theta\right)$, instead of using moving frames as we did for the cigar soliton in Lemma 2.7 of Section 2. Again recalling the formula

$$\Gamma_{ij}^k = \frac{1}{2} g^{k\ell}\left(\frac{\partial}{\partial x^i} g_{j\ell} + \frac{\partial}{\partial x^j} g_{i\ell} - \frac{\partial}{\partial x^\ell} g_{ij}\right),$$

we compute that the Christoffel symbols of g are

$$\begin{array}{lll} \Gamma_{11}^1 = f'/f & \Gamma_{12}^1 = 0 & \Gamma_{22}^1 = -r/f^2 \\ & & \\ \Gamma_{11}^2 = 0 & \Gamma_{12}^2 = 1/r & \Gamma_{22}^2 = 0 \end{array}.$$

Then the standard formula

$$R_{ijk}^\ell = \frac{\partial}{\partial x^i}\Gamma_{jk}^\ell - \frac{\partial}{\partial x^j}\Gamma_{ik}^\ell + \Gamma_{im}^\ell \Gamma_{jk}^m - \Gamma_{jm}^\ell \Gamma_{ik}^m$$

shows that the only nonzero components of the Ricci tensor $R_{jk} = R_{ijk}^i$ of g are

$$R_{11} = R_{211}^2 = \frac{f'(r)}{rf(r)} \qquad \text{and} \qquad R_{22} = R_{122}^1 = \frac{rf'(r)}{f(r)^3}. \tag{2.29}$$

Because $\mathrm{Rc} = Kg$, where K is the Gauss curvature of g, this shows that

$$K(t) \equiv K[g(t)] = \frac{1}{t} \cdot \frac{f'(r)}{rf(r)^3}. \tag{2.30}$$

Now let X be the time-dependent vector field defined on $\mathbb{R}^2 \backslash \{0\}$ by

$$X(t) \doteqdot \frac{r}{tf(r)} \cdot \frac{\partial}{\partial r}. \tag{2.31}$$

Notice that the 1-form ξ metrically dual to X is time-independent:

$$\xi = rf(r)\, dr. \tag{2.32}$$

The components of the Lie derivative $\mathcal{L}_X g$ of g with respect to X are

$$(\mathcal{L}_X g)_{11} = \nabla_1 \xi_1 = \frac{\partial}{\partial r}(rf(r)) - \Gamma_{11}^1 \xi_1 = f(r) \tag{2.33a}$$

$$(\mathcal{L}_X g)_{12} = \nabla_1 \xi_2 = 0 \tag{2.33b}$$

$$(\mathcal{L}_X g)_{21} = \nabla_2 \xi_1 = 0 \tag{2.33c}$$

$$(\mathcal{L}_X g)_{22} = \nabla_2 \xi_2 = -\Gamma_{22}^1 \xi_1 = \frac{r^2}{f(r)}. \tag{2.33d}$$

We want to determine f so that $g(t)$ evolves by the modified Ricci flow

$$\frac{\partial}{\partial t} g(t) = -2\,\mathrm{Rc}[g(t)] + \mathcal{L}_{X(t)} g(t). \tag{2.34}$$

(Compare with the Ricci–DeTurck flow analyzed in Section 3 of Chapter 3.) If we can construct a solution $g(t)$ of the modified Ricci flow (2.34), we will obtain an expanding self-similar solution $\bar{g}(t)$ of the Ricci flow as follows. Observe that there exists a one-parameter family of diffeomorphisms $\varphi_t : \mathbb{R}^2 \to \mathbb{R}^2$ defined for all $t > 0$ by

$$\frac{\partial}{\partial t}\varphi_t(p) = -X(\varphi_t(p), t).$$

(Because $\mathbb{R}^2$ is not compact, long-time existence of φ_t must be checked explicitly, in contrast with Lemma 3.15.) Define an expanding self-similar family of metrics $\bar{g}(t)$ by

$$\bar{g}(t) \doteqdot \varphi_t^*(g(t)) = t \cdot \varphi_t^*\left(f(r)^2\, dr^2 + r^2\, d\theta^2\right). \tag{2.35}$$

Using the identity

$$\left.\frac{\partial}{\partial s}\right|_{s=0}\left(\varphi_t^{-1} \circ \varphi_{t+s}\right) = \left(\varphi_t^{-1}\right)_*\left(\left.\frac{\partial}{\partial s}\right|_{s=0}\varphi_{t+s}\right) = \left(\varphi_t^{-1}\right)_* X(t),$$

we compute that

$$\begin{aligned}
\frac{\partial}{\partial t}\bar{g}(t) &= \frac{\partial}{\partial t}\left(\varphi_t^* g(t)\right) = \left.\frac{\partial}{\partial s}\right|_{s=0}\left(\varphi_{t+s}^* g(t+s)\right) \\
&= \varphi_t^*\left(\frac{\partial}{\partial t}g(t)\right) + \left.\frac{\partial}{\partial s}\right|_{s=0}\left(\varphi_{t+s}^* g(t)\right) \\
&= \varphi_t^*\left\{-2\,\mathrm{Rc}[g(t)] + \mathcal{L}_{X(t)}g(t)\right\} + \left.\frac{\partial}{\partial s}\right|_{s=0}\left[\left(\varphi_t^{-1}\circ\varphi_{t+s}\right)^*\varphi_t^* g(t)\right] \\
&= -2\,\mathrm{Rc}[\bar{g}(t)] + \varphi_t^*\left(\mathcal{L}_{X(t)}g(t)\right) - \mathcal{L}_{\left[\left(\varphi_t^{-1}\right)_* X(t)\right]}\left(\varphi_t^* g(t)\right) \\
&= -2\,\mathrm{Rc}[\bar{g}(t)].
\end{aligned}$$

Hence the self-similar metrics $\bar{g}(t)$ defined by (2.35) solve the Ricci flow.

To construct a solution $g(t)$ of the modified Ricci flow, we observe by recalling (2.29) and (2.33) that $g(t)$ will solve (2.34) if and only if the following system of equations is satisfied:

$$\begin{aligned}
f(r)^2 &= \frac{\partial}{\partial t}g_{11} = -2R_{11} + 2\nabla_1\xi_1 = -2\frac{f'(r)}{rf(r)} + f(r) \\
r^2 &= \frac{\partial}{\partial t}g_{22} = -2R_{22} + 2\nabla_2\xi_2 = -2\frac{rf'(r)}{f(r)^3} + \frac{r^2}{f(r)}.
\end{aligned}$$

Remarkably, multiplying the second equation by f^2/r^2 gives the first, so that this system is equivalent to the single first-order ODE

$$f' = \frac{r}{2}\left(f^2 - f^3\right). \tag{2.36}$$

Integrating this separable equation using partial fractions gives the relation

$$\frac{r^2}{4} = C - \frac{1}{f(r)} + \log\frac{f(r)}{|f(r) - 1|}. \tag{2.37}$$

Notice that $f(r) \to 1$ is possible only if $r \to \infty$, so that we either have $0 < f(r) < 1$ or $1 < f(r)$ for all r. But combining equations (2.30) and (2.36) shows that the Gauss curvature of $g(t)$ is

$$K(t) = \frac{1}{2t} \cdot \frac{1-f}{f}. \tag{2.38}$$

So in order to have a metric of positive curvature, we must find a solution satisfying $0 < f(r) < 1$. Such a solution does in fact exist, and can be given explicitly. Recall that the function $w : (0,\infty) \to (0,\infty)$ defined by $w(x) = xe^x$ is invertible; its inverse is the Lambert-W (product log) function $W : (0,\infty) \to (0,\infty)$. One can verify directly that

$$f(r) = \frac{1}{1 + W\left(\left(\frac{1}{\zeta} - 1\right) \exp\left(\left(\frac{1}{\zeta} - 1\right) - \frac{r^2}{4}\right)\right)} \tag{2.39}$$

is the solution of (2.37) satisfying $f(0) = \zeta \in (0,1)$.

Following [55], we shall now demonstrate how one might arrive at the representation (2.39) for f under the *Ansatz* that $0 < f(r) < 1$. Define

$$F(r) \doteqdot \frac{1}{f(r)} - 1,$$

noting that the *Ansatz* forces $F(r) > 0$ for all $r \geq 0$. Then equation (2.37) is equivalent to

$$\frac{r^2}{4} = C' - F(r) - \log F(r),$$

where $C' = C + 1$, which we write as

$$C' - \frac{r^2}{4} = F(r) + \log F(r) = \log\left(F(r)\, e^{F(r)}\right).$$

Observing that $C' = \log\left(F(0)\, e^{F(0)}\right)$, we can put this into the form

$$F(r)\, e^{F(r)} = e^{C' - r^2/4} = F(0) \exp\left(F(0) - \frac{r^2}{4}\right). \tag{2.40}$$

By the *Ansatz*, we have $F(r) > 0$ for all $r \geq 0$. Thus we may apply W to both sides of equation (2.40), obtaining

$$F(r) = W\left(F(0) \exp\left(F(0) - \frac{r^2}{4}\right)\right),$$

whence equation (2.39) follows immediately.

Note that $\lim_{r\to\infty} F(r) = 0$, so that $\lim_{r\to\infty} f(r) = 1$. Note too that equation (2.38) implies that $\lim_{r\to\infty} K(t) = 0$ for all $t > 0$. This reflects the fact that the metrics $g(t)$ are asymptotic at infinity to a flat cone.

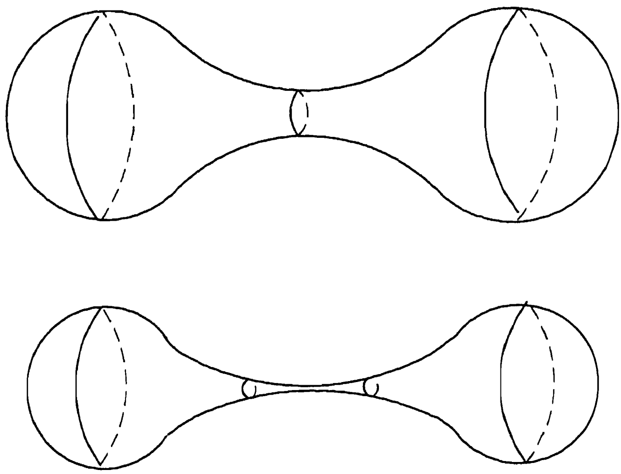

FIGURE 2. A neckpinch forming

5. The neckpinch

The shrinking sphere we considered in Subsection 3.1 is the simplest example of a finite-time singularity of the Ricci flow. In this section, we consider what is its next simplest and arguably its most important singularity, the 'neckpinch'.

One says a solution $(\mathcal{M}^n, g(t))$ of the Ricci flow encounters a **local singularity** at $T < \infty$ if there exists a proper compact subset $K \subset \mathcal{M}^n$ such that

$$\sup_{K \times [0,T)} |\mathrm{Rm}| = \infty$$

but

$$\sup_{(\mathcal{M}^n \backslash K) \times [0,T)} |\mathrm{Rm}| < \infty.$$

This type of singularity formation is also called **pinching behavior** in the literature. The first rigorous examples of pinching behavior for the Ricci flow were constructed by Miles Simon [**119**] on noncompact warped products $\mathbb{R} \times_f S^n$. In these examples, a supersolution of the Ricci flow PDE is used as an upper barrier to force a singularity to occur on a compact subset in finite time. Another family of examples was constructed in [**40**]. Here, the metric is a complete $\mathrm{U}(n)$-invariant shrinking gradient Kähler–Ricci soliton on the holomorphic line bundle L^{-k} over $\mathbb{CP}^{n-1}$ with twisting number $k \in \{1, \ldots, n-1\}$. As $t \nearrow T$, the $\mathbb{CP}^{n-1}$ which constitutes the zero-section of the bundle pinches off, while the metric remains nonsingular and indeed converges to a Kähler cone on the set $(\mathbb{C}^n \backslash \{0\})/\mathbb{Z}_k$ which constitutes the rest of the bundle. Both of these families of examples live on noncompact manifolds.

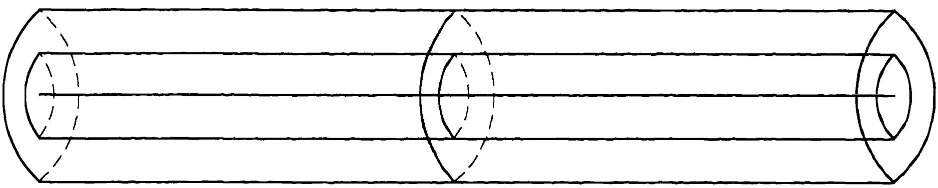

FIGURE 3. The shrinking cylinder soliton

A neckpinch is a special type of local singularity. There exist quantitative measures of necklike behavior for a solution of the Ricci flow. (See [**65**], as well as Section 4 of Chapter 9.) However, the following qualitative characterization will suffice for our present purposes. One says a solution $\left(\mathcal{M}^{n+1}, g(t)\right)$ of the Ricci flow encounters a **neckpinch singularity** at time $T < \infty$ if there exists a time-dependent open subset $N_t \subseteq \mathcal{M}^{n+1}$ such that N_t is diffeomorphic to a quotient of $\mathbb{R} \times \mathcal{S}^n$ by a finite group acting freely, and such that the pullback of the metric $g(t)\,|N_t$ to $\mathbb{R} \times \mathcal{S}^n$ asymptotically approaches the 'shrinking cylinder' soliton

$$ds^2 + 2(n-1)(T-t)\,g_{\text{can}} \tag{2.41}$$

in a suitable sense as $t \nearrow T$. (Here g_{can} denotes the round metric of radius 1 on $\mathcal{S}^n$.)

From the perspective of topology, the neckpinch is perhaps the most important and intensively studied singularity which the Ricci flow can encounter, especially in dimensions three and four. In Chapter 9, we shall see some reasons why this is so. For example, it is expected that one can perform a geometric-topological surgery on the underlying manifold just prior to a neckpinch in such a way that the maximum curvature of the solution is reduced by an amount large enough to permit the flow to be continued on the piece or pieces that remain after the surgery. (We plan to discuss such surgeries in a successor to this volume.)

From the perspective of asymptotic analysis, on the other hand, remarkably little is known about singularity formation in the Ricci flow. For example, one cannot find in the literature examples of either formal or rigorous asymptotic analysis for Ricci flow singularities comparable to what has been done for the mean curvature flow [**7**]. (Also see [**1**].) By contrast, excellent examples of singularity analysis for related reaction-diffusion equations like $u_t = \Delta u + u^p$ can be found in [**46, 47, 48**] and [**41**], among other sources.

Because neckpinches are such important singularities, we shall study them in considerable detail in this section. The first rigorous examples of neckpinches for the Ricci flow were constructed by Sigurd Angenent and the second author in [**6**]. In fact, these are the first examples of any sort of pinching behavior of the Ricci flow on compact manifolds. This section closely follows that paper. The proofs are considerably more involved than anything we have presented thus far, but illustrate many important ideas that will be developed further in the chapters to come.

The main results of [**6**] may be summarized as follows.

THEOREM 2.15. *If $n \geq 2$, there exists an open subset of the family of* $\mathrm{SO}(n+1)$*-invariant metrics on $\mathcal{S}^{n+1}$ such that any solution of the Ricci flow starting at a metric in this set will develop a neckpinch at some time $T < \infty$. The singularity is Type-I (rapidly-forming). Any sequence of parabolic dilations formed at the developing singularity converges to a shrinking cylinder soliton (2.41) uniformly in any ball of radius $o\left(\sqrt{-(T-t)\log(T-t)}\right)$ centered at the neck.*

Any $\mathrm{SO}(n+1)$-invariant metric on $\mathcal{S}^{n+1}$ can be written as

$$g = \varphi^2\, dx^2 + \psi^2\, g_{\mathrm{can}} \tag{2.42}$$

on the set $(-1,1) \times \mathcal{S}^n$, which may be identified in the natural way with the sphere $\mathcal{S}^{n+1}$ with its north and south poles removed. The quantity $\psi(x,t) > 0$ may thus be regarded as the 'radius' of the totally geodesic hypersurface $\{x\} \times \mathcal{S}^n$ at time t. It is natural to write geometric quantities related to g in terms of the distance

$$s\,(x) = \int_0^x \varphi\,(x)\, dx$$

from the equator. Then writing $\frac{\partial}{\partial s} = \frac{1}{\varphi}\frac{\partial}{\partial x}$ and $ds = \varphi\, dx$, one can write the metric (2.42) in the nicer form of a warped product

$$g = ds^2 + \psi^2\, g_{\mathrm{can}}. \tag{2.43}$$

Armed with this notation, one can make a precise statement about the asymptotics of the developing singularity.

THEOREM 2.16. *Let $\bar{s}\,(t)$ denote the location of the smallest neck. Then there are constants $\delta > 0$ and $C < \infty$ such that for t sufficiently close to T one has the estimate*

$$1 + o(1) \leq \frac{\psi(x,t)}{\sqrt{2(n-1)(T-t)}} \leq 1 + \frac{C}{-(T-t)\log(T-t)}(s-\bar{s})^2 \tag{2.44}$$

in the inner layer $|s-\bar{s}| \leq 2\sqrt{-(T-t)\log(T-t)}$, and the estimate

$$\frac{\psi(x,t)}{\sqrt{T-t}} \leq C \frac{s-\bar{s}}{\sqrt{-(T-t)\log(T-t)}} \log \frac{s-\bar{s}}{\sqrt{-(T-t)\log(T-t)}} \tag{2.45}$$

in the intermediate layer $2\sqrt{-(T-t)\log(T-t)} \leq s - \bar{s} \leq (T-t)^{(1-\delta)/2}$.

The estimates in the theorem are exactly those one gets when one writes the evolution equation (2.47) satisfied by ψ with respect to the self-similar space coordinate $\sigma = (s-\bar{s})\,/\sqrt{T-t}$ and time coordinate $\tau = -\log\,(T-t)$ and derives formal matched asymptotics. For this reason, it is expected that these estimates are sharp.

REMARK 2.17. In the special case of a reflection-symmetric metric on $\mathcal{S}^{n+1}$ with a single neck at $x = 0$ and two bumps, the solution exhibits what we call 'single point pinching'; to wit, the neckpinch occurs only on the totally-geodesic hypersurface $\{0\} \times \mathcal{S}^n$, unless the diameter of the solution $(\mathcal{S}^{n+1}, g(t))$ becomes infinite as the singularity time is approached. (The latter alternative is not expected to occur.)

We shall justify this remark in Subsection 5.6.

5.1. How the solution evolves. The first task in studying neckpinches is to compute basic geometric quantities related to the metric (2.42). One begins by observing that g will solve the Ricci flow if and only if φ and ψ evolve by

$$\varphi_t = n \frac{\psi_{ss}}{\psi} \varphi \tag{2.46}$$

and

$$\psi_t = \psi_{ss} - (n-1) \frac{1 - \psi_s^2}{\psi} \tag{2.47}$$

respectively. In order that $g(t)$ extend to a smooth solution of the Ricci flow on S^{n+1}, it suffices to impose the boundary conditions

$$\lim_{x \to \pm 1} \psi_s = \mp 1. \tag{2.48}$$

Indeed, if (2.48) is satisfied by the initial data, then the resulting solution will satisfy the hypotheses of Lemma 2.10 for all times $t > 0$ that it exists.

REMARK 2.18. The partial derivatives ∂_s and ∂_t do not commute, but instead satisfy

$$[\partial_t, \partial_s] = -n \frac{\psi_{ss}}{\psi} \partial_s.$$

REMARK 2.19. Equation (2.46) will effectively disappear in what follows, because the evolution of φ is controlled by the quantity ψ_{ss}/ψ.

The Riemann curvature tensor of (2.43) is determined by the sectional curvatures

$$K_0 = -\frac{\psi_{ss}}{\psi} \tag{2.49}$$

of the n 2-planes perpendicular to the spheres $\{x\} \times \mathcal{S}^n$, and the sectional curvatures

$$K_1 = \frac{1 - \psi_s^2}{\psi^2} \tag{2.50}$$

of the $\binom{n}{2} = n(n-1)/2$ 2-planes tangential to these spheres. The Ricci tensor of g is thus

$$\mathrm{Rc} = (nK_0)\, ds^2 + (K_0 + (n-1) K_1)\, g_{\mathrm{can}}, \tag{2.51}$$

and its scalar curvature is

$$R = 2nK_0 + n(n-1)K_1. \tag{2.52}$$

5.2. Bounds on curvature and other derivatives. The first part of our analysis of a neckpinch singularity is to obtain bounds for the first and second derivatives of ψ. To do so, it will be useful to consider the scale-invariant measure of the difference between the two sectional curvatures defined by

$$a \doteqdot \psi^2 (K_1 - K_0). \tag{2.53}$$

We collect the key results of this part of the analysis in our first proposition.

PROPOSITION 2.20. *Let $g(t)$ be a solution of the Ricci flow having the form (2.43) and satisfying the bounds $|\psi_s| \leq 1$ and $R > 0$ initially.*

(1) *For as long as the solution exists, $|\psi_s| \leq 1$.*
(2) *For as long as the solution exists, one has the bound*
$$-\frac{\alpha}{\psi^2} \leq K_1 - K_0 \leq \frac{\alpha}{\psi^2},$$
where $\alpha \doteqdot \sup |a(\cdot, 0)|$.
(3) *There exists $C = C(n, g(0))$ such that for as long as the solution exists,*
$$|\mathrm{Rm}| \leq \frac{C}{\psi^2}.$$
(4) *The quantity $(K_1)_{\min}$ is nondecreasing.*
(5) *For as long as the solution exists, one has $R > 0$ and $\psi_t < 0$.*
(6) *ψ^2 is a uniformly Lipschitz-continuous function of time; in fact, one has*
$$\left|\left(\psi^2\right)_t\right| \leq 2(\alpha + n).$$
(7) *If $g(t)$ exists for $0 \leq t < T$, then the limit*
$$\psi(x, T) \doteqdot \lim_{t \nearrow T} \psi(x, t)$$
exists for each $x \in [-1, 1]$.

We establish the claims above in the remainder of this subsection. Denoting the first derivative by $v \doteqdot \psi_s$, one first computes that

$$v_t = v_{ss} + \frac{n-2}{\psi} v v_s + \frac{n-1}{\psi^2}\left(1 - v^2\right) v. \tag{2.54}$$

This equation has a happy consequence.

LEMMA 2.21. *Assume $g(t)$ is a solution to the Ricci flow having the form (2.43) and satisfying (2.48) for $0 \leq t < T$. Then*

$$1 \leq \sup |v(\cdot, t)| \leq \sup |v(\cdot, 0)|.$$

PROOF. Applying the maximum principle to (2.54), one concludes that at any maximum of v which exceeds 1, one has

$$v_t \leq \frac{n-1}{\psi^2}\left(1-v^2\right)v < 0.$$

Similarly, at any minimum of v with $v < -1$, one has $v_t > 0$. The result follows once we observe that (2.48) implies $\sup|v(\cdot,0)| \geq 1$. □

Denoting the second derivative of ψ by $w \doteqdot \psi_{ss}$, one calculates that

$$w_t = w_{ss} + (n-2)\frac{v}{\psi}w_s - 2\frac{w^2}{\psi} - (4n-5)\frac{v^2}{\psi^2}w + \frac{n-1}{\psi^2}w \tag{2.55a}$$

$$-2(n-1)\frac{v^2\left(1-v^2\right)}{\psi^3}. \tag{2.55b}$$

Together, the evolution equations for v and w help us obtain a good estimate for a.

LEMMA 2.22. *Under the Ricci flow, the quantity a evolves by*

$$a_t = a_{ss} + (n-4)\frac{\psi_s}{\psi}a_s - 4(n-1)\frac{\psi_s^2}{\psi^2}a.$$

PROOF. Noting that

$$a = \psi w - v^2 + 1,$$

one computes that

$$a_s = \psi_s w + \psi w_s - 2vv_s = \psi w_s - vw$$

and

$$a_{ss} = \psi_s w_s + \psi w_{ss} - v_s w - vw_s = \psi w_{ss} - w^2.$$

Then recalling equations (2.47), (2.54), and (2.55), one derives the equation

$$\begin{aligned}
a_t &= w\psi_t + \psi w_t - 2vv_t \\
&= w\left\{w - (n-1)\frac{1-v^2}{\psi}\right\} \\
&\quad + \psi\left\{\begin{array}{c} w_{ss} + (n-2)\frac{vw_s}{\psi} - 2\frac{w^2}{\psi} - (4n-5)\frac{v^2w}{\psi^2} \\ +(n-1)\frac{w}{\psi^2} - 2(n-1)\frac{v^2(1-v^2)}{\psi^3} \end{array}\right\} \\
&\quad - 2v\left\{w_s + (n-2)\frac{vw}{\psi} + (n-1)\frac{v(1-v^2)}{\psi^2}\right\} \\
&= \psi w_{ss} - w^2 + (n-4)\,vw_s - (5n-8)\frac{v^2w}{\psi} - 4(n-1)\frac{v^2\left(1-v^2\right)}{\psi^2} \\
&= a_{ss} + (n-4)\frac{v}{\psi}a_s - 4(n-1)\frac{v^2}{\psi^2}a.
\end{aligned}$$

□

Applying the maximum principle proves that a is uniformly bounded.

COROLLARY 2.23. $\sup |a(\cdot,t)| \leq \alpha \doteqdot \sup |a(\cdot,0)|$.

We can also show that the curvature is controlled by the radius.

COROLLARY 2.24. *If $g(t)$ is a solution of the Ricci flow of the form (2.43), then there exists C depending only on n and $g(0)$ such that*

$$|\mathrm{Rm}| \leq \frac{C}{\psi^2}.$$

PROOF. Since v is bounded by Lemma 2.21, so is $\psi^2 K_1 = 1 - v^2$. Then since a is bounded, it follows that $\psi^2 K_0 = \psi^2 K_1 - a$ is bounded as well. Since

$$|\mathrm{Rm}|^2 = 2nK_0^2 + n(n-1) K_1^2,$$

the result follows. □

Now to simplify the notation, we define

$$K \doteqdot -K_0 = \frac{\psi_{ss}}{\psi} \qquad \text{and} \qquad L \doteqdot K_1 = \frac{1-\psi_s^2}{\psi^2},$$

recalling that K_0 and K_1 are the sectional curvatures defined in (2.49) and (2.50), respectively. Noting that the Laplacian of a radially symmetric function f is given by

$$\Delta f = \frac{\partial^2 f}{\partial s^2} + n \frac{\psi_s}{\psi}\frac{\partial f}{\partial s},$$

we compute the evolution of the sectional curvatures. One finds that K evolves by

$$K_t = \Delta K + 2(n-1) KL - 2K^2 - 2(n-1)\frac{\psi_s^2}{\psi^2}(K+L).$$

The evolution equation for L is derived as follows.

LEMMA 2.25. *Under the Ricci flow, the quantity L evolves by*

$$\begin{aligned} L_t &= \Delta L + 2\frac{\psi_s}{\psi} L_s + 2\left[K^2 + (n-1)L^2\right] \\ &= \Delta L - 4\frac{\psi_s^2}{\psi^2}(K+L) + 2\left[K^2 + (n-1)L^2\right]. \end{aligned}$$

PROOF. Using equations (2.47) and (2.54), one computes that

$$\begin{aligned} L_t &= -2\frac{\psi_s}{\psi^2}(\psi_s)_t - 2\frac{L}{\psi}\psi_t \\ &= -2\frac{\psi\psi_{sss}}{\psi^2} - 2(n-1)\frac{\psi_s^2}{\psi^2}(K+L) + 2\left(\frac{\psi_s^2}{\psi^2} - L\right)K + 2(n-1)L^2. \end{aligned}$$

Then observing that

$$L_s = -2\frac{\psi_s}{\psi}(K+L),$$

one calculates

$$L_{ss} = -2\frac{\psi\psi_{sss}}{\psi^2} + 6\frac{\psi_s^2}{\psi^2}(K+L) + 2\left(\frac{\psi_s^2}{\psi^2} - L\right)K - 2K^2.$$

Combining these equations yields

$$L_t = L_{ss} - 2(n+2)\frac{\psi_s^2}{\psi^2}(K+L) + 2\left[K^2 + (n-1)L^2\right],$$

whence the result follows. □

COROLLARY 2.26. *$L_{\min}(t)$ is nondecreasing; moreover, if $L_{\min}(0) \neq 0$ then*

$$L_{\min}(t) \geq \frac{1}{L_{\min}^{-1}(0) - 2(n-1)t}.$$

We can now show that the radius ψ is strictly decreasing in time. We will make use of the observation that ψ_t, a, and the sectional curvatures satisfy the relations

$$\psi_t = -\psi\left[K_0 + (n-1)K_1\right] = \psi\left(K_0 - \frac{1}{n}R\right) = \frac{a}{\psi} - n\psi K_1.$$

LEMMA 2.27. *Let $g(t)$ be a solution of the Ricci flow having the form (2.43) and satisfying $|\psi_s| \leq 1$ initially. If the scalar curvature is positive initially, then it remains so, and one has*

$$\psi_t < 0$$

for as long as the solution exists.

PROOF. That $R \geq R_{\min}(0)$ is a general fact we shall prove in Lemma 6.8. To show that $\psi_t < 0$, first note that the bound $|\psi_s| \leq 1$ forces $K_1 = \left(1 - \psi_s^2\right)/\psi^2$ to be nonnegative everywhere. There are now two cases. If $\psi_{ss} < 0$, then one has

$$\psi_t = \psi_{ss} - (n-1)\psi K_1 < 0.$$

On the other hand, if $\psi_{ss} \geq 0$, then $K_0 = -\psi_{ss}/\psi \leq 0$ and hence

$$-\psi_t = \psi\left[K_0 + (n-1)K_1\right] = \psi\left(\frac{1}{n}R - K_0\right) > 0.$$

□

We conclude this part of the analysis by showing that ψ^2 is a uniformly Lipschitz continuous function of time.

LEMMA 2.28. *Let $g(t)$ be a solution of the Ricci flow having the form (2.43) and satisfying $|\psi_s| \leq 1$ initially. Then*

$$\left|\left(\psi^2\right)_t\right| \leq 2(\alpha + n).$$

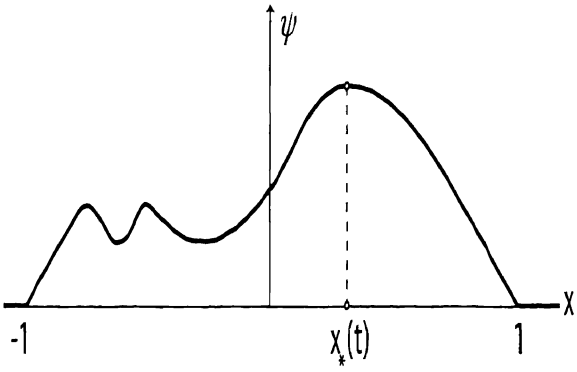

FIGURE 4. A typical profile

PROOF. Noting that

$$\psi_{ss} = \frac{a + \psi_s^2 - 1}{\psi},$$

one has

$$\psi\psi_t = \psi\psi_{ss} - (n-1)\left(1 - \psi_s^2\right) = a - n\left(1 - \psi_s^2\right).$$

□

COROLLARY 2.29. *If $g(t)$ exists for $0 \le t < T$, then $\lim_{t \nearrow T} \psi(x,t)$ exists for each $x \in [-1,1]$.*

5.3. The profile of the solution. We call local minima of $x \mapsto \psi(x,t)$ 'necks' and local maxima 'bumps'. We are interested in solutions whose initial data has at least one neck. The second part of our analysis is to establish the sense in which the profile of the initial data persists, in particular to show that the solution will become singular at its smallest neck and nowhere else.

Let $g(t)$ be a solution of the Ricci flow having the form (2.43). We denote the radius of the smallest neck by

$$r_{\min}(t) \doteqdot \min\{\psi(x,t) : \psi_x(x,t) = 0\}$$

for each t such that the solution exists; if the solution has no necks, then $r_{\min}$ will be undefined. We let $x_*(t)$ denote the right-most bump, and call the region right of $x_*(t)$ the 'polar cap'.

The main results of this part of the analysis are as follows.

PROPOSITION 2.30. *Let $g(t)$ be a solution to the Ricci flow of the form (2.43) such that $|\psi_s| \le 1$ and $R > 0$ initially. Assume that the solution keeps at least one neck.*

(1) *At any time, the derivative ψ_s has finitely many zeroes. The number of zeroes is nonincreasing in time, and if ψ ever has a degenerate critical point (to wit, a point such that $\psi_s = \psi_{ss} = 0$ simultaneously) the number of zeroes of ψ_s drops.*

(2) *There exists a time T bounded above by $r_{\min}(0)^2/(n-1)$ such that the radius $r_{\min}(t)$ of the smallest neck satisfies*

$$(n-1)(T-t) \leq r_{\min}(t)^2 \leq 2(n-1)(T-t).$$

(3) *The solution is concave ($\psi_{ss} < 0$) on the polar cap.*

(4) *If the right-most bump persists, then $D \doteqdot \lim_{t \nearrow T} \psi(x_*(t), t)$ exists. If $D > 0$, then no singularity occurs on the polar cap.*

We begin with the observation that the number of necks cannot increase with time, and further that all bumps/necks will be nondegenerate maxima/minima unless one or more necks and bumps come together and annihilate each other.

LEMMA 2.31. *Let $g(t) : 0 \leq t < T$ be a solution of the Ricci flow of the form (2.43). At any time $t \in (0, T)$, the derivative $v = \psi_s$ has a finite number of zeroes when regarded as a function of $x \in (-1, 1)$. This number of zeroes is nonincreasing in time. Moreover, if ψ ever has a degenerate critical point, the number of zeroes of ψ_s drops.*

PROOF. The derivative $v = \psi_s$ satisfies (2.54), which can be written as a liner parabolic equation $v_t = v_{ss} + Qv$, with

$$Q = (n-2)\,\frac{w}{\psi} + (n-1)\,\frac{1-v^2}{\psi}.$$

Since $\frac{\partial}{\partial s} = \frac{1}{\varphi}\frac{\partial}{\partial x}$, one can in turn write (2.54) as

$$v_t = \varphi^{-1}\left(\varphi^{-1} v_x\right)_x + Q\,(x,t)\,v = A\,(x,t)\,v_{xx} + B\,(x,t)\,xv + C\,(x,t)\,v$$

for suitable functions A, B, C. Since $v \to \mp 1$ as $x \to \pm 1$, the Sturmian theorem [**5**] applies. □

We next derive upper and lower bounds for the rate at which a neck shrinks. These estimates show that a singularity will develop at the smallest neck in finite time, unless the solution loses all its necks first.

LEMMA 2.32. *Let $g(t)$ be a solution to the Ricci flow of the form (2.43) such that $|\psi_s| \leq 1$ and $R > 0$ initially. Then*

$$(n-1)\,(T-t) \leq r_{\min}\,(t)^2 \leq 2\,(n-1)\,(T-t)\,.$$

In particular, the solution must either lose all its necks or else become singular at or before

$$T = \frac{r_{\min}\,(0)^2}{n-1}.$$

PROOF. Note that $\psi\,(\cdot, t)$ is a Morse function except perhaps for finitely many times, and that $r_{\min}\,(t)$ is a Lipschitz continuous function. We claim that

$$-\frac{n-1}{r_{\min}\,(t)} \leq \frac{d}{dt} r_{\min}\,(t) \leq -\frac{n-1}{2r_{\min(t)}}$$

holds for almost all times. The lemma follows from the claim by integration.

To prove the claim, fix some t_0 such that $\psi(\cdot, t_0)$ is a Morse function, and let its smallest critical value be attained at x_0. Then the Implicit Function Theorem implies that there exists a smooth function $x(\cdot)$ defined for t near t_0 such that $x(t_0) = x_0$ and $\psi_x(x(t), t) = 0$. Taking the total derivative, one obtains

$$\begin{aligned}\frac{d}{dt}\psi(x(t), t) &= \psi_t(x(t), t) + \psi_x(x(t), t)\frac{dx}{dt}(t) \\ &= \psi_t(x(t), t) \\ &= \psi_{ss}(x(t), t) - \frac{n-1}{r_{\min}(t)}.\end{aligned}$$

The first inequality in the claim follows from this when one recalls that $\psi_{ss} \geq 0$ at a neck. To prove the other inequality, recall that $R = 2nK_0 + n(n-1)K_1$, where $K_0 = -\psi_{ss}/\psi$ and $K_1 = \left(1 - \psi_s^2\right)/\psi^2$. So at a neck,

$$\psi_{ss} = -\psi K_0 = \psi\left[\frac{n-1}{2}K_1 - \frac{R}{2n}\right] = \frac{n-1}{2\psi} - \psi\frac{R}{2n}.$$

Since the inequality $R > 0$ is preserved by the flow, the second inequality follows. □

The final steps in this part of the proof show that no singularity occurs on the polar cap: the region between the last bump and the pole. To do this, we first use the tensor maximum principle to show that the Ricci curvature is positive there. We shall invoke the following slight modification of Theorem 4.6.

LEMMA 2.33. *Let $(\mathcal{N}_t, \partial\mathcal{N}_t, g(t) : 0 \leq t < T)$ be a smooth 1-parameter family of compact Riemannian manifolds with boundary. Let S and U be symmetric $(2,0)$-tensor fields on $\mathcal{N}_t$ such that S evolves by*

$$\frac{\partial}{\partial t}S = \Delta S + U * S,$$

where $U * S$ denotes the symmetrized product $(U * S)_{ij} = U_i^k S_{kj} + S_i^k U_{kj}$. Suppose that $\inf_{p\in\mathcal{N}_t} S(p, 0) > 0$ and that $S(q, t) > 0$ for all points $q \in \partial\mathcal{N}_t$ and times $t \in [0, T)$. If $(U * S)(V, \cdot) \geq 0$ whenever $S(V, \cdot) = 0$, then

$$\inf_{p\in\mathcal{N}_t,\ t\in[0,T)} S(p, t) \geq 0.$$

Once one observes that the hypotheses imply that S can first attain a zero eigenvalue only at an interior point of $\mathcal{N}_t$, the proof is almost identical to what we shall present in Chapter 4. We will use the proposition to prove that ψ is strictly concave on the polar caps.

LEMMA 2.34. *If $\psi_{ss}(x, 0) \leq 0$ for $x_*(0) < x < 1$, then $\psi_{ss}(x, t) \leq 0$ for $x_*(t) < x < 1$ and all $0 < t < T$.*

PROOF. First we show that the Ricci tensor satisfies

$$\frac{\partial}{\partial t}\operatorname{Rc} = \Delta\operatorname{Rc} + U * \operatorname{Rc}, \tag{2.56}$$

where U is the (2,0) tensor given by

$$U = (K_1 - K_0)\left[(n-1)\, ds^2 + \psi^2 g_{\mathrm{can}}\right].$$

To verify (2.56) at any given point $(x, P) \in (-1,1) \times \mathcal{S}^n$, we adopt the convention that Roman indices belong to $\{0, \dots n\}$ while Greek indices belong to $\{1, \dots, n\}$.

We choose coordinates $\{y^1, \dots, y^n\}$ near P on $\mathcal{S}^n$ such that $\hat{g} = g_{\mathrm{can}}$ has components $\hat{g}_{\alpha\beta} = \delta_{\alpha\beta}$ at P. Then we set $y^0 = s$ so that $\{y^0, y^1, \dots, y^n\}$ is a coordinate system near (x, P) on $(-1,1) \times \mathcal{S}^n$. The only nonvanishing components of the metric g in these coordinates are $g_{00} = 1$ and $g_{\alpha\alpha} = \psi^2 \hat{g}_{\alpha\alpha}$. Moreover, all components of the Riemann tensor vanish in these coordinates except $R_{\alpha 00\alpha} = \psi^2 K_0$ and $R_{\alpha\beta\beta\alpha} = \psi^4 K_1$ $(\alpha \neq \beta)$. Similarly, all components of the Ricci tensor vanish except $R_{00} = nK_0$ and $R_{\alpha\alpha} = \psi^2\left[K_0 + (n-1) K_1\right]$. The evolution equation (2.56) now follows from formula (6.7). We have established (2.56) on the punctured sphere $(-1,1) \times \mathcal{S}^n$; by continuity it remains valid at the poles $\{\pm 1\} \times \mathcal{S}^n$.

To apply Proposition 2.33, let $\mathcal{N}_t$ denote the topological $(n+1)$-ball

$$\mathcal{N}_t = \{(x, P) : x \geq x_*(t),\ P \in \mathcal{S}^n\}$$

endowed with the metric $g(t)$. Observe that the sectional curvatures K_0 and K_1 are strictly positive on

$$\partial \mathcal{N}_t = \{x_*(t)\} \times \mathcal{S}^n,$$

because ψ has a local maximum at $x_*(t)$ and ψ_s has a simple zero there by Lemma 2.31. So $\mathrm{Rc} > 0$ on $\partial\mathcal{N}_t$. If $\psi(\cdot, 0)$ is strictly convex for all $x \geq x_*(0)$, then $\mathrm{Rc}(\cdot, 0) > 0$ on $\mathcal{N}_t$. So if Rc ever acquires a zero eigenvalue, it must do so at some point $p \in \operatorname{int}\mathcal{N}_t$ and time $t \in (0, T)$. If $\mathrm{Rc}(V,V)|_{(p,t)} = 0$ for some vector $V \in T_p\mathcal{N}_t$, then $(U * \mathrm{Rc})(V,V) = 0$, because U and Rc commute. Hence Lemma 2.33 implies that $\mathrm{Rc} \geq 0$ on $\mathcal{N}_t$ for as long as $g(t)$ exists. The lemma follows immediately. □

Finally, we prove that when $\eta > 0$ is chosen sufficiently small, the quantity

$$b \doteqdot \frac{|a|}{\psi^\eta} = \psi^{2-\eta} |K_1 - K_0|$$

remains bounded in a neighborhood $\mathcal{B}$ of the pole. Because the exponent η breaks scale invariance, b may be regarded as a pinching inequality for the curvatures on the polar cap. The bound on b thus lets us apply a parabolic dilation argument that shows that singularity formation on a polar cap under our hypotheses would lead to a contradiction. (We shall study parabolic dilations in greater generality in Chapter 8.)

LEMMA 2.35. *Let $g(t)$ be as in Proposition 2.30. If $D > 0$, no singularity occurs on the polar cap*

$$((x_*(t), 1) \times S^n(1)) \cup \{P_+\}.$$

PROOF. By Lemma 2.28, we may let C_0 be an upper bound for $|(\psi^2)_t|$. Choose $t_1 \in (0,T)$ so that $C_0(T-t_1) < D^2/8$. Then one has

$$\psi(x_*(t),t)^2 \geq D^2 - C_0(T-t) > \frac{7}{8}D^2$$

for all $t \in [t_1, T)$. Now let x_1 be the unique solution of $\psi(x_1,t_1)^2 = \frac{3}{4}D^2$ in the interval $[x_*(t_1), 1]$. Lemma 2.28 implies for all $t \in [t_1, T)$ that

$$\psi(x_1,t)^2 \leq \frac{3}{4}D^2 + C_0(t-t_1) < \frac{7}{8}D^2 < \psi(x_*(t),t)^2.$$

Thus one has $x_*(t) < x_1 < 1$ for all $t \in [t_1,T)$, and hence $\psi_s < 0$ and $\psi_{ss} < 0$ on the interval $[x_1, 1]$ for all $t \in [t_1, T)$. It follows that the metric distance

$$d_1(t) \doteqdot s(1,t) - s(x_1,t)$$

from (x_1,t) to the pole P_+ is decreasing in time. Indeed,

$$\frac{d}{dt}d_1(t) = \int_{x_1}^{1} \frac{n\psi_{ss}}{\psi} ds < 0.$$

Next let $x_2 \in (x_1, 1)$ be defined by $\psi(x_2,t_1)^2 = \frac{3}{8}D^2$. Then for $t \in [t_1, T)$ one has

$$\psi(x_2,t)^2 \leq \frac{3}{8}D^2 + C_0(T-t_1) < \frac{1}{2}D^2$$

and

$$\psi(x_1,t)^2 \geq \frac{3}{4}D^2 - C_0(T-t_1) > \frac{5}{8}D^2.$$

Thus $\psi(x_1,t)^2 - \psi(x_2,t)^2 \geq D^2/8$. Hence by crudely estimating the quantity $\psi(x_1,t) + \psi(x_2,t)$ from below by

$$\psi(x_2,t) \geq \sqrt{\frac{3}{8}D^2 - C_0(T-t_1)} \geq \sqrt{D^2/4} = D/2,$$

we obtain

$$\psi(x_1,t) - \psi(x_2,t) \geq \frac{D^2/8}{D/2} = D/4.$$

Lemma 2.34 implies that for $x \in [x_2, 1)$, one has

$$-\psi_s > \frac{\psi(x_1,t) - \psi(x,t)}{s(x_1,t) - s(x,t)} > \frac{D/4}{s(x_1,t_1) - s(1,t_1)} \doteqdot \delta.$$

At this point we once again consider the quantity a defined in (2.53). We found that $L(a) = 0$, where L is the differential operator

$$L = \partial_t - \partial_s^2 - (n-4)\frac{\psi_s}{\psi}\partial_s + 4(n-1)\frac{\psi_s^2}{\psi^2}.$$

One can compute that the quantity $u \doteqdot \psi^\eta$ satisfies

$$\begin{aligned} L(u) &= (4-\eta)\frac{\psi_s}{\psi}u_s + \frac{n-1}{\psi^2}(4\psi_s^2 - \eta)u \\ &= (4-\eta)\psi^{\eta-2}\psi_s^2 + \frac{n-1}{\psi^2}(4\psi_s^2 - \eta)u. \end{aligned}$$

We have $|\psi_s| \geq \delta$ in the region $\mathcal{Q}_2 = [x_2, 1) \times [t_1, T)$. If we take $\eta < 4\delta^2 < 4$, then we have $L(u) > 0$ in $\mathcal{Q}_2$. Then by the maximum principle, there exists $C < \infty$ such that $|a| \leq Cu$ in $\mathcal{Q}_2$. Indeed, b must attain its maximum C on the parabolic boundary of $\mathcal{Q}_2$. At the left end (that is at $x = x_2$) we have $u = \psi^\eta \geq (D/64)^\eta$, while Corollary 2.23 implies that $|a|$ is bounded by $\sup |a(\cdot, 0)|$. At the other vertical side of $\mathcal{Q}_2$ (that is at $x = 1$) we have $a = 0$. Since a is smooth, this implies $a = O(s(1,t) - s(x,t))$. On the other hand, because $u = \psi^\eta$ and $\psi_s = -1$ at $x = 1$, we have $\lim_{x \nearrow 1} |a/u| = 0$ for all $t < T$. Finally, at $t = t_1$, the quantity b is continuous for $x_2 \leq x < 1$; while we have just verified that $b \to 0$ as $x \nearrow 1$. Thus b is bounded on the parabolic boundary of $\mathcal{Q}_2$, and hence bounded on $\mathcal{Q}_2$.

We now perform a parabolic dilation. Let $\mathcal{B}_2$ denote the portion of $\mathcal{S}^{n+1}$ where $x \geq x_2$. Then we have a solution $g(t)$ to the Ricci flow on $\mathcal{B}_2$ defined for $t \in [t_1, T)$. By Lemma 2.34, this solution has $\mathrm{Rc} \geq 0$. Because of spherical symmetry, $\mathcal{B}_2$ is a geodesic ball in $(\mathcal{S}^{n+1}, g(t))$ whose radius is bounded from above by $d_1(t)$. Since $\psi = \psi(x_2, t) \geq D/2$ on the boundary of $\mathcal{B}_2$, it follows from the estimate $|\psi_s| \leq 1$ that the radius of $\mathcal{B}_2$ is bounded from below by $D/2$.

Assume that the sectional curvatures of the metrics $g(t)$ on $\mathcal{B}_2$ are not bounded as $t \nearrow T$. Then there is a sequence of points $P_k \in \mathcal{B}_2$ and time $t_k \in [t_2, T)$ such that $|\mathrm{Rm}(P_k, t_k)| \to \infty$ as $k \to \infty$. We may choose this sequence so that

$$\sup_{Q \in \mathcal{B}_2} |\mathrm{Rm}\,(Q, t)| = |\mathrm{Rm}\,(P_k, t_k)|$$

holds for $t \leq t_k$. Writing x_k for the x coordinate of P_k, we note that $\psi(x_k, t_k) \to 0$; and from $|\psi_s| \geq \delta$ we conclude that

$$\lim_{k \to \infty} d_{t_k}(P_k, P_+) = 0,$$

where d_t is the distance measured with the metric $g(t)$.

Define $\epsilon_k = |\mathrm{Rm}(P_k, t_k)|^{-1/2}$, and introduce rescaled metrics

$$g_k(t) = \frac{1}{\epsilon_k^2}\, g(t_k + \epsilon_k^2 t).$$

Let C_1 be the constant from Corollary 2.24 for which $|\mathrm{Rm}| \leq C_1 \psi^{-2}$ holds. Then we have

$$\psi(x_k, t_k) \leq \sqrt{C}_1 \epsilon_k;$$

and because $|\psi_s| \geq \delta$, we also have

$$d_{t_k}(P_k, P_+) \leq \frac{C_1}{\delta} \epsilon_k.$$

The distance from P_k to the pole P_+ measured in the rescaled metric $g_k(0) = \epsilon_k^{-2} g(t_k)$ is therefore at most C_1/δ. In particular, this distance is uniformly bounded.

Translating to the rescaled metric, we find that $g_k(t)$ is a solution of the Ricci flow defined for $t \in (-\epsilon_k^{-2} t_k, 0]$ on the region $\mathcal{B}_2$. The Riemann curvature of g_k is uniformly bounded by $|\mathrm{Rm}| \leq 1$, with equality attained at P_k at $t = 0$. One may then extract a convergent subsequence whose limit is an ancient solution $g_\infty(t)$ of the Ricci flow on $\mathbb{R}^{n+1} \times (-\infty, 0]$ with uniformly bounded sectional curvatures. The limit solution has nonzero sectional curvature at $t = 0$ at some point P_* whose distance to the origin is at most C_1/δ.

We introduce the radial coordinate $r = r(x, t_k) = \epsilon_k^{-1}(s(1, t_k) - s(x, t_k))$. Then the metric $g_k(0) = \epsilon_k^{-2} g(t_k)$ seen through the exponential map at the pole P_+ is given by

$$g_k(0) = (dr)^2 + \Psi_k(r)^2 \hat{g}, \text{ with } \Psi_k(r) = \epsilon_k^{-1}\psi(s(1, t_k) - \epsilon_k r, t_k).$$

The metrics g_k converge in C^∞ on regions $r \leq R$ for any finite R, and the functions Ψ_k hence also converge in C^∞ on any interval $[0, R]$. The scale invariant quantity a is given by $a = \psi\psi_{ss} - \psi_s^2 + 1 = \Psi\Psi_{rr} - \Psi_r^2 + 1$, and it satisfies $|a| \leq C\psi^\eta \leq C\epsilon_k^\eta \Psi_k^\eta$. Thus we find that the limit $\Psi_\infty = \lim \Psi_k$ satisfies

$$\Psi\Psi_{rr} - \Psi_r^2 + 1 = 0.$$

Hence for some λ, μ, one has

$$\Psi(r; \lambda, \mu) = \begin{cases} \frac{1}{\lambda}\sin\lambda(r - \mu), & \lambda < \infty \\ r - \mu & \lambda = \infty \end{cases}.$$

Since $-\Psi_r = -\psi_s \in [\delta, 1]$ cannot vanish, the only valid solution is the one with $\lambda = \infty$ and $\mu = 0$, that is $\Psi_\infty(r) = r$. But then the limiting metric $g_\infty = \lim g_k(0)$ is $(dr)^2 + r^2\hat{g}$; to wit, g_∞ is the flat Euclidean metric. Because this is impossible, we conclude that the sectional curvatures of the metrics $g(t)$ are in fact bounded from above for all $x \in (x_2(t), 1)$ and $t < T$. Since Corollary 2.24 bounds the sectional curvatures for $x \in (x_*(t), x_2(t))$, we have proved Lemma 2.35. □

5.4. Convergence to a shrinking cylinder. In the third part of our analysis, we derive estimates which indicate that a neckpinch asymptotically approaches the shrinking cylinder soliton (2.41). By the construction in Subsection 5.5 below, we may assume that the solution keeps at least one neck. Then letting $x_-(t)$ and $x_+(t)$ denote the left-most and right-most bumps, respectively, we define the 'waist'

$$\mathcal{W}(t) = [x_-(t), x_+(t)]$$

for all times t such that the solution exists. The key to this part of our analysis is the quantity

$$F \doteqdot -\frac{K_0}{K_1}\log K_1 = \frac{K}{L}\log L.$$

Notice that F is positive in a neighborhood of a neck. An application of the maximum principle will let us bound F from above when it is positive and $K_1 = L$ is sufficiently large. The value of this estimate is that the factor

$\log K_1$ breaks scale invariance. Our bound on F will thus show that K_0/K_1 becomes small whenever K_1 is large, in particular near a forming neckpinch.

The main results obtained in this part of the proof show that the singularity is Type I (rapidly forming) and estimate its asymptotics. (We shall discuss the classification of singularities in Section 1 of Chapter 8.)

PROPOSITION 2.36. *Let $g(t) : 0 \leq t < T$ be a maximal solution of the Ricci flow having the form (2.43) such that $|\psi_s| \leq 1$ and $R \geq 0$. Assume that the solution has at least one neck.*

(1) *A singularity occurs at the smallest neck at some time $T < \infty$. This singularity is of Type I; in particular, there exists $C = C(n, g_0)$ such that*
$$|\mathrm{Rm}| \leq \frac{C}{T-t}.$$
(2) *There exists $C = C(n, g_0)$ such that*
$$\frac{K}{L}\left[\log L + 2 - \log L_{\min}(0)\right] \leq C.$$
(3) *Let $\bar{s}(t)$ denote the location of the smallest neck. Then there are constants $\delta > 0$ and $C < \infty$ such that for t sufficiently close to T, one has the estimate*
$$1 \leq \frac{\psi}{r_{\min}} \leq 1 + \frac{C}{-\log r_{\min}}\left(\frac{s-\bar{s}}{r_{\min}}\right)^2$$
in the inner layer $|s-\bar{s}| \leq 2r_{\min}\sqrt{-\log r_{\min}}$, and the estimate
$$\frac{\psi}{r_{\min}} \leq C\frac{s-\bar{s}}{r_{\min}\sqrt{-\log r_{\min}}}\log\frac{s-\bar{s}}{r_{\min}\sqrt{-\log r_{\min}}}$$
in the intermediate layer $2r_{\min}\sqrt{-\log r_{\min}} \leq s-\bar{s} \leq r_{\min}^{1-\delta}$, where
$$r_{\min}(t) = [1+o(1)]\sqrt{2(n-1)(T-t)}.$$

LEMMA 2.37. *Let $g(t)$ be a solution to the Ricci flow of the form (2.43) such that $|\psi_s| \leq 1$ and $R > 0$ initially. Then there exists $C = C(n, g_0)$ such that*
$$|\mathrm{Rm}| \leq \frac{C}{T-t}.$$

PROOF. By Lemma 2.34, the curvature remains bounded on the polar caps. On the waist $\mathcal{W}(t)$, we obtain the stated bound by combining Corollary 2.24 with Lemma 2.32 □

LEMMA 2.38. *The quantity $F = \frac{K}{L}\log L$ evolves by*

$$F_t = \Delta F + 2\left(\frac{\log L - 1}{L\log L}\right)L_sF_s + \left(\frac{2-\log L}{\log L}\right)\frac{KL_s^2}{L^3} \tag{2.57a}$$
$$-2P\left(\frac{\psi_s}{\psi}\right)^2\frac{K+L}{L} + 2QK, \tag{2.57b}$$

where

$$P = (n-1)\log L - 2\frac{K}{L}(\log L - 1) \tag{2.58}$$

and

$$Q = n - 1 - \frac{K^2}{L^2}(\log L - 1) - F. \tag{2.59}$$

PROOF. Straightforward computation. □

Rather than work with F directly, we shall consider a related quantity $\hat{F}$, which is invariant under simultaneous rescaling of the metric $g \mapsto \lambda g$ and time $t \mapsto \lambda t$.

LEMMA 2.39. *Let $g(t) : 0 \leq t < T$ be a maximal solution of the Ricci flow having the form (2.43) such that $|\psi_s| \leq 1$ and $R \geq 0$. For all $t \in [0,T)$ and $x \in \mathcal{W}(t)$, the scaling-invariant quantity*

$$\hat{F} \doteqdot \frac{K}{L}\left[\log L + 2 - \log\left(L_{\min}(0)\right)\right]$$

satisfies

$$\sup_{\mathcal{W}(t)} \hat{F}(\cdot, t) \leq \max\left\{n-1, \sup_{\mathcal{W}(0)} \hat{F}(\cdot, 0)\right\}. \tag{2.60}$$

PROOF. We first consider the case that $L_{\min}(0) \geq e^2$. Then by Corollary 2.26, we have $L_{\min}(t) \geq e^2$ for as long as the solution exists. We will apply the maximum principle to equation (2.57) in the region where $F \geq n-1$, noting that $K > 0$ in this region as well. Our assumption that $0 \leq R = n[-2K + (n-1)L]$ implies that

$$K \leq \frac{n-1}{2}L.$$

Using $L > e^2 > e$, we conclude that the coefficient P in (2.58) satisfies

$$P \geq (n-1)\log L - (n-1)(\log L - 1) = n - 1.$$

Thus when $K > 0$ and $L > e^2$, equation (2.57) implies the differential inequality

$$F_t \leq \Delta F + 2\left(\frac{\log L - 1}{L \log L}\right) L_s F_s + 2QK.$$

But if $F \geq n-1$, then the coefficient Q in (2.59) satisfies

$$Q \leq -\frac{K^2}{L^2}(\log L - 1) \leq 0.$$

Hence

$$F_t \leq \Delta F + 2\left(\frac{\log L - 1}{L \log L}\right) L_s F_s.$$

We conclude that if $L_{\min}(0) \geq e^2$, then $\sup F(\cdot, t)$ cannot increase whenever it exceeds $n-1$.

To complete the proof we must deal with possibility that $L_{\min}(0) < e^2$. In this case, we consider a rescaled solution

$$\tilde{g}(t) = \lambda g\left(\frac{t}{\lambda}\right)$$

of the Ricci flow. Denoting its sectional curvatures by $\tilde{K}_0 = -\tilde{K}$ and $\tilde{K}_1 = \tilde{L}$, we have $\tilde{L}_{\min} = \lambda^{-1} L_{\min}$ and $\tilde{K} = \lambda^{-1} K$. Thus the choice $\lambda = e^{-2} L_{\min}(0) > 0$ implies that $\tilde{L}(0) \geq e^2$, whence the preceding arguments apply to the metric $\tilde{g}$. We conclude therefore that $\max\{n-1, \tilde{F}\}$ is nonincreasing, where $\tilde{F} = (\tilde{K}/\tilde{L}) \log \tilde{L}$. Translated back to the original metric $g(t)$, this implies that

$$\max\left\{n-1, \frac{K}{L} \log\left(\frac{e^2}{L_{\min}(0)} L\right)\right\}$$

does not increase with time, as claimed. $\square$

The bound for $\hat{F}$ we obtained above is an example of what is called a 'pinching estimate' for the sectional curvatures, and is a foreshadowing of similar results we will see in Chapters 6 and 9.

Our next result will let us compare the actual radius $\psi(x,t)$ near a neck with the radius $\sqrt{2(n-1)(T-t)}$ of the cylinder soliton which is its singularity model.

LEMMA 2.40. *Let $g(t)$ be as in Proposition 2.36. For $t \in [0,T)$, choose $x_0(t) \in \mathcal{W}(t)$ so that*

$$\psi(x_0(t), t) = r_{\min}(t).$$

Define

$$\sigma = s(x,t) - s(x_0(t), t),$$

Then there are constants $\delta > 0$ and $C < \infty$ such that for t sufficiently close to T, one has

$$1 \leq \frac{\psi(x,t)}{r_{\min}(t)} \leq 1 + \frac{C}{-\log r_{\min}(t)} \left(\frac{\sigma}{r_{\min}(t)}\right)^2 \tag{2.61}$$

for $|\sigma| \leq 2 r_{\min} \sqrt{-\log r_{\min}}$, and

$$\frac{\psi(x,t)}{r_{\min}(t)} \leq C \frac{\sigma}{r_{\min}(t)\sqrt{-\log r_{\min}(t)}} \log \frac{\sigma}{r_{\min}(t)\sqrt{-\log r_{\min}(t)}} \tag{2.62}$$

for $2 r_{\min}(t)\sqrt{-\log r_{\min}(t)} \leq \sigma \leq r_{\min}(t)^{1-\delta}$.

PROOF. Let ε denote a small positive number to be chosen below. We regard t as fixed, and consider the neighborhood of the neck $x_0(t)$ in which $\psi \leq \varepsilon$ and $|\psi_s| < \varepsilon$. In this region, one always has $L \geq (1-\varepsilon^2)/\varepsilon^2$, so that there exists a constant $C < \infty$ such that

$$|2 - \log L_{\min}(0)| \leq CL.$$

Taking C larger if necessary, we may assume by Lemma 5.4 that $\frac{K}{L}\log L \leq C$ as well. This implies that

$$\frac{\psi_{ss}}{\psi} \leq \frac{1-\psi_s^2}{\psi^2} \cdot \frac{C}{\log \frac{1-\psi_s^2}{\psi^2}}$$

and thus

$$\psi\psi_{ss} \leq C \frac{1-\psi_s^2}{\log(1-\psi_s^2) - 2\log\psi}.$$

Hence

$$\frac{\psi\psi_{ss}}{1-\psi_s^2} \leq \frac{C}{-2\log\psi} \cdot \frac{1}{1 - \frac{1}{2}\frac{\log(1-\psi_s^2)}{\log\psi}}.$$

Since we have restricted attention to the region where $\psi \leq \varepsilon$ and $|\psi_s| \leq \varepsilon$, the denominator of the second factor on the right-hand side obeys the bound

$$1 - \frac{1}{2}\frac{\log(1-\psi_s^2)}{\log\psi} \geq 1 - \frac{1}{2}\frac{\log(1-\varepsilon^2)}{\log\varepsilon}.$$

So by choosing ε small enough, we can ensure that

$$\frac{\psi\psi_{ss}}{1-\psi_s^2} \leq \frac{C}{-\log\psi}.$$

We now further restrict our attention to the region to the right of the neck. There $\psi_s > 0$, which allows us to choose the radius ψ as a coordinate and thereby regard all quantities as functions of ψ. Then we have

$$-\frac{d}{d\log\psi}\left(\log(1-\psi_s^2)\right) = \frac{2\psi_s\psi_{ss}}{1-\psi_s^2} \cdot \frac{\psi}{\psi_s} \leq \frac{C}{-\log\psi}.$$

Integrating this differential inequality from the center of the neck (where $\psi = r_{\min}$ and $\psi_s = 0$) to an arbitrary point, one gets

$$-\log(1-\psi_s^2) \leq \int_{u=r_{\min}}^{\psi} \frac{C}{-\log u}\, d(\log u) = C \log \frac{\log r_{\min}}{\log\psi}.$$

Using the calculus inequalities $x \leq -\log(1-x)$ and $\log x \leq x - 1$, one then obtains

$$\psi_s^2 \leq C \log \frac{\log r_{\min}}{\log\psi} \leq C\left(\frac{\log r_{\min}}{\log\psi} - 1\right). \tag{2.63}$$

Since we are assuming $|\psi_s| \leq \varepsilon$, this last inequality will only be useful if the right-hand side is no more than ε^2. Henceforth we assume that

$$r_{\min} \leq \psi \leq \left(\frac{1}{r_{\min}}\right)^{\varepsilon^2/2C} r_{\min}. \tag{2.64}$$

Then using the fact that $e^{-\varepsilon^2/C} \leq 1 - \varepsilon^2/2C$ for small ε, one finds that (2.64) and (2.63) imply that $|\psi_s| \leq \varepsilon$ and $\psi < \varepsilon$, as required.

Now we integrate once again to get

$$\sqrt{C}\sigma \geq \int_{r_{\min}}^{\psi} \frac{du}{\sqrt{\frac{\log r_{\min}}{\log u} - 1}}.$$

Substitute $u = r_{\min} v$. Then it follows from the inequalities $r_{\min} \leq u \leq \psi$ that $1 \leq v \leq \psi / r_{\min}$. By (2.64), this implies that $0 \leq \log v \leq -\frac{\varepsilon^2}{2C} \log r_{\min}$. Thus we have

$$\frac{\sqrt{C}\sigma}{r_{\min}} \geq \int_1^{\psi/r_{\min}} \frac{\sqrt{-\log r_{\min} - \log v}}{\sqrt{\log v}}\, dv \tag{2.65a}$$

$$\geq \frac{1}{2}\sqrt{-\log r_{\min}} \int_1^{\psi/r_{\min}} \frac{dv}{\sqrt{\log v}}. \tag{2.65b}$$

To put this inequality into a more useful form, we shall use the following fact, which is left to the reader to verify.

CLAIM 2.41. *The function* $Z : [0, \infty) \to [1, \infty)$ *defined by*

$$\xi = \int_1^{Z(\xi)} \frac{dv}{\sqrt{\log v}}$$

is monotone increasing and has the asymptotic behavior that

$$Z(\xi) = 1 + \frac{1}{2}\xi^2 + o\left(\xi^2\right) \qquad \text{as} \qquad \xi \searrow 0$$

and

$$Z(\xi) = [1 + o(1)]\, \xi \sqrt{\log \xi} \qquad \text{as} \qquad \xi \nearrow \infty.$$

In light of the claim, estimate (2.65) can be recast as

$$\frac{\psi}{r} \leq Z\left(\frac{2\sqrt{C}\sigma}{r_{\min}\sqrt{-\log r_{\min}}} \right).$$

Estimate (2.61) follows from this and the expansion of $Z(\xi)$ for small ξ.

For larger σ we get

$$\frac{\psi}{r_{\min}} \leq C \frac{\sigma}{r_{\min}\sqrt{-\log r_{\min}}} \log \frac{\sigma}{r_{\min}\sqrt{-\log r_{\min}}},$$

which is exactly (2.62). This estimate will be valid in the region where (2.64) is satisfied and $\sigma \geq \delta r_{\min}\sqrt{-\log r_{\min}}$. Using $C\sqrt{-\log r_{\min}} = r_{\min}^{o(1)}$, we conclude that (2.62) will hold if $\psi / r_{\min} \leq (1/r_{\min})^{\varepsilon^2/2C + o(1)}$. □

By Lemma 2.37, the neckpinch forms a Type I (rapidly forming) singularity. One can therefore construct a sequence of parabolic dilations near the developing neck which converge to a singularity model that is an ancient solution of the Ricci flow. (See Subsection 4.1 of Chapter 8.) The lemma above shows that the singularity model must in fact be the cylinder solution (2.41). It follows that

$$r_{\min}(t) = [1 + o(1)] \sqrt{2(n-1)(T-t)}.$$

In this way, one obtains part (3) of Proposition 2.36 as an immediate corollary of Lemma 2.40.

5.5. Neckpinches happen. In this part of the analysis, we show that there exist initial data $\Psi = \psi(0)$ meeting our hypotheses. In particular, we construct simple examples obtained by removing a neighborhood of the equator of a standard sphere and replacing it with a long thin neck. These examples satisfy $\Psi = \sqrt{A + Bs^2}$ near the equator (for appropriate constants A and B) and blend smoothly into the standard sphere metric on the polar caps. Our construction justifies

PROPOSITION 2.42. *There exist initial metrics*

$$g = ds^2 + \Psi^2 g_{\mathrm{can}}$$

for the Ricci flow on S^{n+1} which satisfy $|\Psi_s| \leq 1$, have positive scalar curvature, and possess a neck sufficiently small and a bump sufficiently large so that under the flow, the neck must disappear before the bump can vanish. Hence these solutions exhibit a neckpinch singularity in finite time.

Our method will be to remove a sufficiently large neighborhood of the equator of the round metric $ds^2 + (\cos s)^2 g_{\mathrm{can}}$ on $\mathcal{S}^{n+1}$ and replace with a sufficiently narrow neck. Let $A > 0$, and let B satisfy

$$\begin{array}{ll} 0 < B < 1/2 & \text{if } n = 2 \\ 0 < B < 1 & \text{if } n \geq 3 \end{array}.$$

Define the function

$$W(s) \doteqdot \sqrt{A + Bs^2}.$$

LEMMA 2.43. *The metric*

$$ds^2 + W(s)^2 g_{\mathrm{can}}$$

on $\mathbb{R} \times \mathcal{S}^n$ has positive scalar curvature, and the scale-invariant measure of its curvature pinching

$$a = W(s)W''(s) - W'(s)^2 + 1$$

obeys the bounds

$$1 - B < a < 1 + B.$$

PROOF. One computes that

$$\begin{aligned} \frac{W(s)^2}{n} \cdot R &= -2W(s)W''(s) + (n-1)\{1 - W'(s)^2\} \\ &= n - 1 - 2B + (3 - n)\frac{B^2 s^2}{W}(s)^2. \end{aligned}$$

Since $Bs^2 < A + Bs^2 = W(s)^2$, we find that when $n \geq 3$,

$$\frac{W(s)^2}{n} \cdot R \geq n - 1 - 2B + 3 - n = 2(1 - B).$$

For $n = 2$ we get

$$\frac{W(s)^2}{n} \cdot R \geq n - 1 - 2B = 1 - 2B.$$

To estimate a, we write

$$a = WW_{ss} - W_s^2 + 1 = \frac{1}{2}(W^2)_{ss} - 2W_s^2 + 1,$$

which implies that

$$a = B - 2\frac{B^2 s^2}{A + Bs^2} + 1,$$

hence that $1 - B < a < 1 + B$. □

Now for A and B chosen as above, we define

$$\hat{\psi}_{A,B}(s) = \min\{W(s)\cos s\} = \begin{cases} W(s) & \text{if } |s| \leq S_{A,B} \\ \cos s & \text{if } S_{A,B} < |s| \leq \pi/2 \end{cases},$$

where $S_{A,B}$ is the unique positive solution of $\cos s = \sqrt{A + Bs^2}$. The function $\hat{\psi}_{A,B}$ is piecewise smooth an satisfies $\left|\frac{d}{ds}\hat{\psi}_{A,B}\right| \leq 1$ for all $s \in [-\pi/2, \pi/2] \setminus \{S_{A,B}\}$. Moreover, it is easy to check that the metric $ds^2 + \hat{\psi}_{A,B}^2 g_{\text{can}}$ has positive scalar curvature.

We now smooth out the corner that $\hat{\psi}_{A,B}$ has at $S_{A,B}$. First we construct a new function $\breve{\psi}_{A,B}$ which coincides with $\hat{\psi}_{A,B}$ outside a small interval $I_\epsilon \doteqdot (S_{A,B} - \epsilon, S_{A,B} + \epsilon)$ and has $\frac{d^2}{ds^2}\breve{\psi}_{A,B}$ constant in I_ϵ. This constant may be chosen so that $\breve{\psi}_{A,B}$ is C^1. Since $\hat{\psi}_{A,B}$ is increasing for $0 \leq s < S_{A,B}$ and decreasing thereafter, we have $\frac{d}{ds}\breve{\psi}_{A,B} < 0$ in I_ϵ. Moreover, we also have $|\frac{d}{ds}\breve{\psi}_{A,B}| < 1$ in I_ϵ. Hence the metric $ds^2 + \breve{\psi}_{A,B}^2 g_{\text{can}}$ will have $R > 0$ everywhere.

Now the function $\breve{\psi}_{A,B}^2$ is C^1, and its second derivative $\frac{d^2}{ds^2}\breve{\psi}_{A,B}$ has simple jump discontinuities at $S_{A,B} \pm \epsilon$. So we may smooth it in arbitrarily small neighborhoods of the two points $S_{A,B} \pm \epsilon$ in such a way that the smoothed function $\psi_{A,B}$ is C^∞, satisfies $R > 0$ everywhere it is defined, and coincides with $\hat{\psi}_{A,B}$ outside of $I_{2\epsilon}$.

Finally, we observe that is possible to perform the construction above in such a way that the initial neck is small enough that the solution will encounter a neckpinch singularity before its diameter shrinks to zero. Indeed, it is not hard to check that one may choose $\bar{A} > 0$ sufficiently small so that for all $A \in (0, \bar{A})$, the smoothed function $\psi_{A,B}$ will coincide with $\hat{\psi}_{A,B}$ in the interval $|s| \leq a$. Moreover, its derivatives $\frac{d^2}{ds^2}\psi_{A,B}$ will be bounded for all $|s| \geq a$ uniformly in $A \in (0, \bar{A})$.

We have proved the following.

LEMMA 2.44. *There exists a family of initial metrics*

$$g_{A,B} = ds^2 + \psi_{A,B}(s)\, g_{\text{can}}$$

such that each metric in the family has positive scalar curvature, satisfies $\left|\frac{d}{ds}\psi_{A,B}\right| \leq 1$*, has a neck of radius* $r_{\min}(0) = \sqrt{A}$*, and has a bump of height no less that* $\bar{A}\sqrt{B}$*. Moreover, the scale-invariant measure* a *of its curvature pinching is bounded uniformly for all* $A \in (0, \bar{A})$.

Lemma 2.32 implies that the solution of the Ricci flow starting from $g_{A,B}$ must lose its neck before time

$$T_A \doteqdot \frac{r_{\min}(0)^2}{n-1} = \frac{A}{n-1}.$$

On the other hand, the solution will have a bump at some $x_*(t)$; and since $|a|$ is uniformly bounded for all solutions under consideration, the height of this bump will be bounded from below by

$$\psi(x_*(t), t)^2 \geq \bar{A}\sqrt{B} - Ct \geq \bar{A}\sqrt{B} - \frac{CA}{n-1}.$$

If $A \in (0, \bar{A})$ is small enough, the neck must disappear before the bump can vanish. This concludes our proof of Proposition 2.42

5.6. Single point pinching. Finally, we justify Remark 2.17. We consider the special case of a reflection-invariant metric on S^{n+1} with a single symmetric neck at $x = 0$ and two bumps. (See Figure 2.) For solutions of this type whose diameter remains bounded as the singularity time is approached, the neckpinch singularity will occur only on the totally-geodesic hypersurface $\{0\} \times S^n$. We shall prove this claim by constructing a family of subsolutions $\underline{v}$ for $v = \psi_s$.

Recall that $x_*(t)$ denotes the location of the right-hand bump. (For reflection symmetric data, this is the unique point in $(0,1)$ where $\psi_{\max}(t)$ is attained). Recall too that $D = \lim_{t \nearrow T} \psi(x_*(t), t)$ denotes the final height of that bump. By Proposition 2.42, we may assume that $D > 0$. Define the function

$$\rho(t) = n \int_0^t \int_0^D \left(\frac{\psi_s}{\psi}\right)^2 ds\, dt,$$

noting that ρ is monotone increasing in time, so that $\rho(T) = \lim_{t \nearrow T} \rho(t)$ exists.

Now let $s_*(t) = s(x_*(t), t)$ denote the distance from the equator to the right-hand bump. By Lemma 2.21, one has $|\psi_s| \leq 1$; and because $\partial_t \psi_{\max} \leq -(n-1)/\psi_{\max}$, one also has $\psi(s_*(t), t) > D$ for all $t \in [0, T)$. Together, these results imply that $s_*(t) > D$ for all $t \in [0, T)$. Let $x_D(t)$ denote the unique point in $(0, x_*(t))$ such that $s(x_D(t), t) = D$. Notice that for any s, the hypothesis of reflection symmetry about $x = 0$ lets one integrate by parts to obtain the identity

$$\frac{\partial s}{\partial t} = n \int_0^x \frac{\psi_{ss}}{\psi} \frac{\partial s}{\partial x}\, dx = n \left\{ \frac{\psi_s}{\psi} + \int_0^{s(x)} \left(\frac{\psi_s}{\psi}\right)^2 ds \right\}.$$

Since Lemma 2.31 implies that $\psi_s \geq 0$ when $0 \leq s \leq s_*(t)$, one may then estimate at any $\hat{x} \in [x_D(t), x_*(t)]$ that

$$s_*(t) \geq s(\hat{x}, t) \geq n \int_0^t \int_0^{s(\hat{x})} \left(\frac{\psi_s}{\psi}\right)^2 ds\, dt \geq \rho(t). \tag{2.66}$$

In particular, $\rho(t)$ is bounded above by the distance from the equator to the bump.

We are now ready to prove that the singularity occurs only on the equatorial hypersurface.

PROPOSITION 2.45. *If the diameter of the solution $g(t)$ remains bounded as $t \nearrow T$, then $\psi(s, T) > 0$ for all $0 < s < D/2$.*

To establish this result, let $t_0 \in (0, T)$ and $\delta > 0$ be given. For $\varepsilon > 0$ to be chosen below, define

$$\underline{v}(s,t) = \varepsilon\{s - [\rho(t) - \rho(t_0 - \delta)]\}.$$

By (2.66), the finite-diameter assumption implies that $\rho(T) < \infty$, hence that

$$\sup_{t_0 - \delta < t < T} [\rho(t) - \rho(t_0 - \delta)]$$

becomes arbitrarily small when $t_0 - \delta$ is sufficiently close to T. Proposition 2.45 is thus an immediate consequence of

LEMMA 2.46. *If $\rho(T) < \infty$, then for any $t_0 \in (0, T)$ and $\delta \in (0, t_0)$, there exists $\varepsilon > 0$ such that $v = \psi_s$ satisfies*

$$v(s,t) \geq \underline{v}(s,t)$$

for all points $0 < s < D/2$ and times $t_0 \leq t < T$.

PROOF. Since $\underline{v}(0, t_0) < 0$ and $v(s, t_0) > 0$ for all $s \in (0, s_*(t_0))$, one may choose ε_1 such that if $0 < \varepsilon < \varepsilon_1$, then $v(s, t_0) > \underline{v}(s, t_0)$ whenever $0 \leq s \leq D/2$. So if the lemma is false, there will be a first time $\bar{t} \in (t_0, T)$ and a point $\bar{s} \in (0, D/2)$ such that $v(\bar{s}, \bar{t}) = \underline{v}(\bar{s}, \bar{t})$. At $(\bar{s}, \bar{t})$, one then has

$$v_t \leq \underline{v}_t = \varepsilon \left(\frac{\partial s}{\partial t} - \rho'\right) \tag{2.67}$$

as well as $v = \underline{v}$, $v_s = \underline{v}_s = \varepsilon$, and $v_{ss} \geq \underline{v}_{ss} = 0$. Hence

$$v_t = v_{ss} + \frac{(n-2)}{\psi} v v_s + \frac{(n-1)}{\psi^2}(1 - v^2) v \geq \varepsilon \frac{(n-2)}{\psi} v + \frac{(n-1)(1 - \underline{v}^2)}{\psi^2} v.$$

Whenever $0 < \varepsilon < \varepsilon_2 = \sqrt{2}/D$, one has $\underline{v} \leq \varepsilon s < 1/\sqrt{2}$ for all $0 \leq s \leq D/2$ and $t \in [t_0, T)$. Then because $v \geq 0$ for $s \in (0, D/2) \subset [0, s_*(t)]$, one estimates at $(\bar{s}, \bar{t})$ that

$$\begin{aligned}
\underline{v}_t - v_t &\leq \varepsilon\left(\frac{\partial s}{\partial t} - \rho' - \frac{n-2}{\psi} v\right) - (n-1)(1-\underline{v}^2)\frac{v}{\psi^2} \\
&\leq \varepsilon\left(\frac{\partial s}{\partial t} - \rho'\right) - \frac{n-1}{2}\frac{v}{\psi^2} \\
&= \varepsilon n\left[\frac{v}{\psi} - \int_{\bar{s}}^{D}\left(\frac{v}{\psi}\right)^2 ds\right] - \frac{n-1}{2}\frac{v}{\psi^2} \\
&< \frac{v}{\psi}\left[\varepsilon n - \frac{n-1}{2\psi}\right].
\end{aligned}$$

Choose $\varepsilon_3 < (n-1)/[2n\psi_{\max}(0)]$. Then if $0 < \varepsilon < \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$, the consequence $\psi_t < 0$ of Lemma 2.27 implies the inequality $\underline{v}_t - v_t < 0$. This contradicts (2.67), hence proves the result. □

6. The degenerate neckpinch

In contrast to the contents of Section 5, the discussion in this section is heuristic and not fully rigorous. Nonetheless, degenerate neckpinches have been rigorously demonstrated [7] for the mean curvature flow of a surface in $\mathbb{R}^3$. Although an analogous result is still lacking for solutions of the Ricci flow, the mean curvature flow is similar in so many respects that it is strongly conjectured that degenerate neckpinches exist for the Ricci flow.

6.1. An intuitive picture of degenerate neckpinches. Let us consider how a degenerate neckpinch should arise. Imagine a family of rotationally symmetric solutions

$$\{(\mathcal{S}^n, g_\alpha(t)) : \alpha \in [0,1]\}$$

of the Ricci flow parameterized by the unit interval . When $\alpha = 0$, let the initial metric have the profile described in Subsection 5.5. This is a symmetric dumbbell with two equally-sized hemispherical regions joined by a thin neck. (See Figure 2.) We proved in Section 5 that there exist such initial conditions which lead to a neckpinch singularity of the Ricci flow at some time $T_0 < \infty$. On the other hand, when $\alpha = 1$, let the initial metric be a round metric on the n-sphere. As we saw in Subsection 3.1, such a metric remains round and shrinks to a point at some time $T_1 < \infty$.

Now suppose that for α close to 1, the initial metric $g_\alpha(0)$ is a lopsided dumbbell where the neck is fat and short and one of the hemispheres is smaller than the other. In this case, if α is close enough to 1, the smoothing effects of the Ricci flow are conjectured to be strong enough to allow the smaller hemisphere to pull through before the neck shrinks. (See Figure 5.) In this case, the metric should eventually acquire positive curvature everywhere and shrink to a round point. (This is of course known if $g_\alpha(0)$ is a metric of positive Ricci curvature on the 3-sphere; see Chapter 6.)

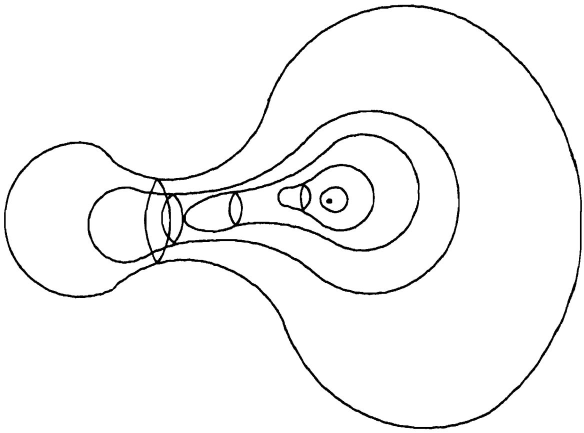

FIGURE 5. A lopsided dumbell shrinking to a round point

As α increases from 0 to 1, we may thus imagine that the initial metrics $g_\alpha(0)$ are dumbbells in which the necks become shorter and fatter while one of the hemispheres becomes progressively smaller. Such metrics can in fact be constructed having positive scalar curvature. Then for each $\alpha \in [0,1]$, the solution will exist up to a time $T_\alpha < \infty$ when a singularity forms. (By Theorem 6.45, a finite-time singularity occurs at some time T if and only if the curvature becomes unbounded as $t \nearrow T$. And by Lemma 6.53, a finite-time singularity is inevitable if the scalar curvature ever becomes everywhere positive.) By the continuous dependence of a well-posed PDE on its initial conditions, one expects that there is some parameter $\hat{\alpha} \in (0,1)$ such that the neck pinches off for all $\alpha \in [0, \hat{\alpha})$. On the other hand, one expect for all $\alpha \in (\hat{\alpha}, 1]$ that the necks will not pinch off but that the metrics will approach constant curvature while shrinking to a point. (Depending on the initial family of metrics $g_\alpha(0)$, there may of course be more than one such bifurcation point in $(0,1)$ at which the solutions change qualitatively.) For the solution $g_{\hat{\alpha}}(t)$, we expect that the neck pinches off at $T_{\hat{\alpha}} < \infty$ exactly at the same time that the smaller hemisphere shrinks to a point. As illustrated in Figure 6, this should result in a cusplike singularity known at a **degenerate neckpinch**. (Another discussion of the intuition behind this example may be found in Section 3 of [**63**].)

The degenerate neckpinch is the standard conjectural model for what is known as a *slowly forming singularity*. As we shall see in Chapter 8, a finite-time singularity of a solution $(\mathcal{M}^n, g(t))$ to the Ricci flow is classified as Type I and called 'rapidly forming' if it occurs at the natural parabolic growth rate suggested by the example of the round sphere in Subsection 3.1, hence if there exists $C < \infty$ such that

$$\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}| \cdot (T-t) \leq C < \infty.$$

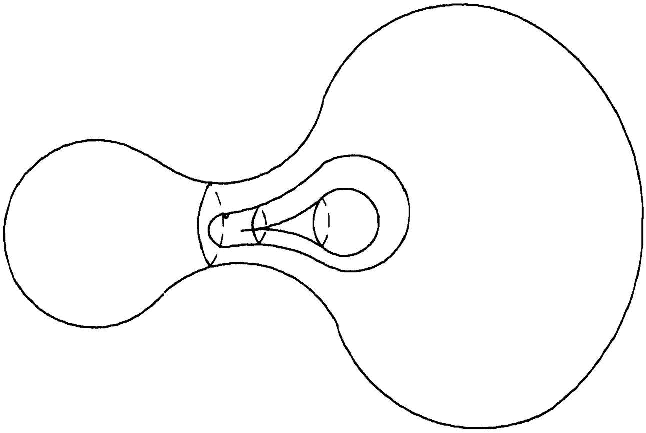

FIGURE 6. A cusp forming

On the other hand, the singularity is classified as Type IIa if

$$\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}| \cdot (T-t) = \infty.$$

Type IIa singularities are called 'slowly forming'. (We will explain this terminology below.) The heuristic picture we have described motivates the following question about the formation of slowly forming singularities.

PROBLEM 2.47. Given a compact smooth 3-manifold $\mathcal{M}^3$ and a one-parameter family of initial metrics $\{g_\beta : \beta \in [0,1]\}$ such that the Ricci flow starting at g_1 forms a Type I singularity model which is a quotient $\mathcal{S}^3/\Gamma_1$ of the 3-sphere, while the Ricci flow starting at g_0 forms a Type I singularity model which is a quotient $(\mathcal{S}^2 \times \mathbb{R})/\Gamma_0$ of the cylinder, does there exist $\beta \in (0,1)$ such that the flow starting at g_β forms a Type IIa singularity?

It is expected that at any $\beta \in (0,1)$ where there is a transition in the singularity model, a Type IIa singularity should form.

6.2. An intuitive picture of slowly forming singularities. Because a Type IIa singularity has the property that

$$\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}| \cdot (T-t) = \infty,$$

the terminology 'slowly forming' may initially seem counterintuitive. On the contrary, it is in fact perfectly logical, as we now explain.

Let us first explore why $\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}| \cdot (T-t)$ is a meaningful quantity for finite-time singularities. Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow defined on a maximal time interval $\alpha \leq t < T$, and define

$$K(t) \doteqdot \sup_{x \in \mathcal{M}^n} |\mathrm{Rm}(x,t)|.$$

In Theorem 6.45, we shall prove that $T < \infty$ only if

$$\lim_{t \nearrow T} K(t) = \infty.$$

Because the metric g scales like $(\text{length})^2$ and $\text{Rm} = (R_{ijk}^{\ell})$ is dimensionless, the quantity $|\text{Rm}|$ scales like $(\text{length})^{-2}$. We will see in Section 3 of Chapter 3 that the Ricci flow is equivalent to a parabolic PDE. Since parabolic PDE equate time with $(\text{length})^2$, the quantity $(T-t)K(t)$ is naturally dimensionless.

We shall discuss short-time and long-time existence theory for the Ricci flow in Chapters 3 and 7, respectively. (See also Sections 7 and 8 of Chapter 6.) In Corollary 7.7, we will prove a short-time existence result which implies that

$$T - t \geq \frac{c}{K(t)}, \tag{2.68}$$

where $c > 0$ depends only on the dimension n. In other words, the natural quantity $(T-t)K(t)$ is always bounded from below. One classifies finite-time singularities, therefore, by whether or not it is bounded from above.

One says the singularity at time $T < \infty$ is **rapidly forming** if the estimate (2.68) is sharp in the sense that there exists $C \in [c, \infty)$ such that for all $t < T$ one has

$$\frac{c}{K(t)} \leq T - t \leq \frac{C}{K(t)}.$$

A rapidly forming singularity is evidently Type I, because $(T-t)K(t) \leq C$. On the other hand, if for every $C \in [c, \infty)$ there exists a time $t_C < T$ such that

$$T - t_C > \frac{C}{K(t_C)}, \tag{2.69}$$

one says the singularity is **slowly forming**. Equation (2.69) reveals why this terminology makes sense: there is more time remaining until extinction than the maximum curvature predicts.

The conjectural picture sketched earlier in this section is very useful for developing intuition for understanding slowly-forming singularities. Regardless of their role in Hamilton's program for obtaining a geometric and topological classification of 3-manifolds via the Ricci flow, degenerate neckpinches and indeed any slowly forming singularities are interesting in their own right. Indeed, a rigorous proof of the existence of degenerate neckpinches for the Ricci flow and an asymptotic analysis of their behaviors would be highly valuable in forming productive conjectures concerning certain analytic (as opposed to topological) properties of singularity formation.

Notes and commentary

Although no examples of slowly forming (Type IIa) singularities on compact manifolds are known at the time of this writing, a rigorous example [**35**] has recently been constructed on $\mathbb{R}^2$.

CHAPTER 3

Short time existence

A foundational step in the study of any system of evolutionary partial differential equations is to show that it enjoys short-time existence and uniqueness. In this chapter, we prove short-time existence of the Ricci flow. Because the flow is quasilinear and only weakly parabolic, short-time existence does not follow from standard parabolic theory. Hamilton originally used the sophisticated machinery of the Nash–Moser implicit function theorem to establish short-time existence in [**58**]. Our proof follows the elegant method suggested by Dennis DeTurck in [**36**].

1. Variation formulas

Given any smooth family of metrics $g(t)$ on a smooth manifold $\mathcal{M}^n$, one may compute the variations of the Levi-Civita connection and its associated curvature tensors. These calculations will be applied later to derive the evolution equations for these quantities under the Ricci flow. (See Chapter 6.) We compute the variation formulas now because the variation of the Ricci tensor determines the symbol of the Ricci flow equation when we regard the Ricci tensor $\mathrm{Rc}(g)$ as a second order partial differential operator on the metric g. (These variation formulas may also be found in [**20**].)

We first derive the variation formula for the Levi-Civita connection. Although a connection is not a tensor, the difference of two connections is a $(2,1)$-tensor. In particular, the time-derivative of a connection is a $(2,1)$-tensor. In this section, we shall assume that

$$\frac{\partial}{\partial t} g_{ij} = h_{ij}\,, \tag{3.1}$$

where h is some symmetric 2-tensor.

LEMMA 3.1. *The metric inverse g^{-1} evolves by*

$$\frac{\partial}{\partial t} g^{ij} = -g^{ik} g^{j\ell} h_{k\ell}\,. \tag{3.2}$$

PROOF. Since the Kronecker delta satisfies

$$\delta^i_\ell = g^{ik} g_{k\ell},$$

equation (3.1) implies that

$$0 = \left(\frac{\partial}{\partial t} g^{ik}\right) g_{k\ell} + g^{ik} h_{k\ell}.$$

Hence

$$\frac{\partial}{\partial t}g^{ij} = \left(\frac{\partial}{\partial t}g^{ik}\right) g_{k\ell}g^{j\ell} = -g^{ik}g^{j\ell}h_{k\ell}\,.$$

□

LEMMA 3.2. *The variation of the Levi-Civita connection* Γ *is given by*

$$\frac{\partial}{\partial t}\Gamma_{ij}^{k} = \frac{1}{2}g^{k\ell}\left(\nabla_i h_{j\ell} + \nabla_j h_{i\ell} - \nabla_\ell h_{ij}\right). \tag{3.3}$$

PROOF. Recall that

$$\Gamma_{ij}^{k} = \frac{1}{2}g^{k\ell}\left(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}\right)$$

in local coordinates $\{x^i\}$, where $\partial_i \doteqdot \frac{\partial}{\partial x^i}$. Hence

$$\begin{aligned}\frac{\partial}{\partial t}\Gamma_{ij}^{k} &= \frac{1}{2}\frac{\partial}{\partial t}g^{k\ell}\cdot\left(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}\right)\\ &\quad + \frac{1}{2}g^{k\ell}\left(\partial_i\left(\frac{\partial}{\partial t}g_{j\ell}\right) + \partial_j\left(\frac{\partial}{\partial t}g_{i\ell}\right) - \partial_\ell\left(\frac{\partial}{\partial t}g_{ij}\right)\right).\end{aligned}$$

In geodesic coordinates centered at $p \in \mathcal{M}^n$, one has $\Gamma_{ij}^k(p) = 0$. It follows that $\partial_i A_{jk} = \nabla_i A_{jk}$ at p for any tensor A; in particular, $\partial_i g_{jk}(p) = 0$ for all i, j, k. Thus we obtain

$$\frac{\partial}{\partial t}\Gamma_{ij}^{k}(p) = \frac{1}{2}g^{k\ell}\left(\nabla_i\frac{\partial}{\partial t}g_{j\ell} + \nabla_j\frac{\partial}{\partial t}g_{i\ell} - \nabla_\ell\frac{\partial}{\partial t}g_{ij}\right)(p)\,.$$

Since both sides of this equation are components of tensors, the result holds in any coordinate system and at any point. □

Since the Riemann curvature tensor is defined solely in terms of the Levi-Civita connection, we can readily compute its evolution.

LEMMA 3.3. *The evolution of the Riemann curvature tensor* Rm *is given by*

$$\frac{\partial}{\partial t}R_{ijk}^{\ell} = \frac{1}{2}g^{\ell p}\left\{\begin{array}{c}\nabla_i\nabla_j h_{kp} + \nabla_i\nabla_k h_{jp} - \nabla_i\nabla_p h_{jk}\\ -\nabla_j\nabla_i h_{kp} - \nabla_j\nabla_k h_{ip} + \nabla_j\nabla_p h_{ik}\end{array}\right\}. \tag{3.4}$$

PROOF. In local coordinates $\{x^i\}$, we have the standard formula

$$R_{ijk}^{\ell} = \partial_i\Gamma_{jk}^{\ell} - \partial_j\Gamma_{ik}^{\ell} + \Gamma_{jk}^{p}\Gamma_{ip}^{\ell} - \Gamma_{ik}^{p}\Gamma_{jp}^{\ell}.$$

Thus we compute

$$\begin{aligned}\frac{\partial}{\partial t}R_{ijk}^{\ell} &= \partial_i\left(\frac{\partial}{\partial t}\Gamma_{jk}^{\ell}\right) - \partial_j\left(\frac{\partial}{\partial t}\Gamma_{ik}^{\ell}\right)\\ &\quad + \frac{\partial}{\partial t}\left(\Gamma_{jk}^{p}\right)\cdot\Gamma_{ip}^{\ell} + \Gamma_{jk}^{p}\cdot\left(\frac{\partial}{\partial t}\Gamma_{ip}^{\ell}\right) - \left(\frac{\partial}{\partial t}\Gamma_{ik}^{p}\right)\cdot\Gamma_{jp}^{\ell} - \Gamma_{ik}^{p}\cdot\left(\frac{\partial}{\partial t}\Gamma_{jp}^{\ell}\right).\end{aligned}$$

As in Lemma 3.2, we use geodesic coordinates centered at $p \in \mathcal{M}^n$ to calculate that

$$\frac{\partial}{\partial t} R_{ijk}^{\ell}(p) = \nabla_i \left(\frac{\partial}{\partial t} \Gamma_{jk}^{\ell} \right)(p) - \nabla_j \left(\frac{\partial}{\partial t} \Gamma_{ik}^{\ell} \right)(p),$$

and then observe that this formula holds everywhere. The present lemma follows from substituting (3.3) into this equation. □

REMARK 3.4. By commuting derivatives, one can also write the evolution of the Riemann tensor in the form

$$\frac{\partial}{\partial t} R_{ijk}^{\ell} = \frac{1}{2} g^{\ell p} \left\{ \begin{array}{c} \nabla_i \nabla_k h_{jp} + \nabla_j \nabla_p h_{ik} - \nabla_i \nabla_p h_{jk} - \nabla_j \nabla_k h_{ip} \\ -R_{ijk}^{q} h_{qp} - R_{ijp}^{q} h_{kq} \end{array} \right\}.$$

LEMMA 3.5. *The evolution of the Ricci tensor* Rc *is given by:*

$$\frac{\partial}{\partial t} R_{jk} = \frac{1}{2} g^{pq} \left(\nabla_q \nabla_j h_{kp} + \nabla_q \nabla_k h_{jp} - \nabla_q \nabla_p h_{jk} - \nabla_j \nabla_k h_{qp} \right). \tag{3.5}$$

PROOF. This follows from contracting on $i = \ell$ in Lemma 3.3. □

REMARK 3.6. Recall that the divergence (3.19) of a $(2,0)$-tensor is given by

$$(\delta h)_k = -(\operatorname{div} h)_k = -g^{ij} \nabla_i h_{jk},$$

and denote the **Lichnerowicz Laplacian** [**93**] of a $(2,0)$-tensor by

$$(\Delta_L h)_{jk} \doteqdot \Delta h_{jk} + 2 g^{qp} R_{qjk}^{r} h_{rp} - g^{qp} R_{jp} h_{qk} - g^{qp} R_{kp} h_{jq}. \tag{3.6}$$

(The Lichnerowicz Laplacian is discussed in Section 4 of Appendix A). Denoting the trace of h by

$$H \doteqdot \operatorname{tr}_g h = g^{pq} h_{pq},$$

one can write the evolution of the Ricci tensor in the form

$$\frac{\partial}{\partial t} R_{jk} = -\frac{1}{2} \left[\Delta_L h_{jk} + \nabla_j \nabla_k H + \nabla_j (\delta h)_k + \nabla_k (\delta h)_j \right].$$

LEMMA 3.7. *The evolution of the scalar curvature function R is given by:*

$$\frac{\partial}{\partial t} R = -\Delta H + \nabla^p \nabla^q h_{pq} - \langle h, \mathrm{Rc} \rangle. \tag{3.7}$$

PROOF. Using Lemmas 3.1 and 3.5, we compute that

$$\begin{aligned} \frac{\partial}{\partial t} R &= \left(\frac{\partial}{\partial t} g^{jk} \right) R_{jk} + g^{jk} \left(\frac{\partial}{\partial t} R_{jk} \right) \\ &= -g^{ij} g^{k\ell} \left(\nabla_i \nabla_j h_{k\ell} - \nabla_i \nabla_k h_{j\ell} + h_{ik} R_{j\ell} \right). \end{aligned}$$

□

REMARK 3.8. One can also write the evolution of the scalar curvature in the invariant form

$$\frac{\partial}{\partial t} R = -\Delta H + \operatorname{div}(\operatorname{div} h) - \langle h, \mathrm{Rc} \rangle.$$

The general formula for the evolution of volume is given by the following observation.

LEMMA 3.9. *The volume element $d\mu$ evolves by*

$$\frac{\partial}{\partial t} d\mu = \frac{1}{2}\left(g^{ij}\frac{\partial}{\partial t} g_{ij}\right) d\mu = \frac{H}{2}\, d\mu.$$

PROOF. If $\left\{x^i\right\}_{i=1}^n$ are oriented local coordinates, then

$$d\mu = \sqrt{\det g}\, dx^1 \wedge \cdots \wedge dx^n.$$

Hence

$$\frac{\partial}{\partial t} d\mu = \frac{1}{2}\left(\frac{\partial}{\partial t}\log\det g\right)\sqrt{\det g}\, dx^1 \wedge \cdots \wedge dx^n.$$

□

COROLLARY 3.10. *The total scalar curvature $\int_{\mathcal{M}^n} R\, d\mu$ evolves by*

$$\frac{d}{dt}\left(\int_{\mathcal{M}^n} R\, d\mu\right) = \int_{\mathcal{M}^n}\left(\frac{1}{2}RH - \langle \mathrm{Rc}, h\rangle\right) d\mu.$$

We will also need to understand the evolution of length.

LEMMA 3.11. *Let γ_t be a time-dependent family of curves with fixed endpoints in $(\mathcal{M}^n, g(t))$, and let $L_t(\gamma_t)$ denote the length of γ_t with respect to the metric $g(t)$. If*

$$\frac{\partial}{\partial t} g = h,$$

then

$$\frac{d}{dt} L_t(\gamma_t) = \frac{1}{2}\int_{\gamma_t} h(S,S)\, ds - \int_{\gamma_t}\langle \nabla_S S, V\rangle\, ds, \tag{3.8}$$

where S is the unit tangent to γ_t, V is the variation vector field, and ds is the element of arc length.

PROOF. Consider the variation

$$\gamma : [a,b] \times (-\varepsilon, \varepsilon) \to \mathcal{M}^n,$$

regarded as the family of curves $\gamma_t(u) = \gamma(u,t)$. Assume $\gamma(a,\cdot) \equiv x \in \mathcal{M}^n$ and $\gamma(b,\cdot) \equiv y \in \mathcal{M}^n$. Define the vector fields

$$U = \gamma_*(\partial/\partial u) \quad \text{and} \quad V = \gamma_*(\partial/\partial t)$$

along γ, and set

$$s(u) \doteqdot \int_a^u \langle U, U\rangle^{1/2}\, du.$$

Then the unit tangent to γ_t is

$$S \doteqdot \langle U, U\rangle^{-1/2}\, U.$$

Since $[U, V] = 0$, we have

$$\begin{aligned}
\frac{d}{dt} L_t(\gamma_t) &= \frac{d}{dt} \int_a^b g(U,U)^{1/2}\, du \\
&= \frac{1}{2} \int_a^b g(U,U)^{-1/2} \frac{\partial g}{\partial t}(U,U)\, du \\
&+ \int_a^b g(U,U)^{-1/2} g(U, \nabla_U V)\, du \\
&= \frac{1}{2} \int_a^b \frac{\partial g}{\partial t}(S,S)\, ds + \int_a^b \langle S, \nabla_S V\rangle\, ds \\
&= \frac{1}{2} \int_a^b h(S,S)\, ds + \left(\langle S, V\rangle|_a^b - \int_a^b \langle \nabla_S S, V\rangle\, ds \right).
\end{aligned}$$

□

2. The linearization of the Ricci tensor and its principal symbol

2.1. The symbol of a nonlinear differential operator. Let $\mathcal{V}$ and $\mathcal{W}$ be vector bundles over $\mathcal{M}^n$, and let

$$L : C^\infty(\mathcal{V}) \to C^\infty(\mathcal{W})$$

be a linear differential operator of order k, written as

$$L(V) = \sum_{|\alpha| \le k} L_\alpha \partial^\alpha V,$$

where $L_\alpha \in \hom(\mathcal{V}, \mathcal{W})$ is a bundle homomorphism (namely, a linear map when restricted to each fiber) for each multi-index α. If $\zeta \in C^\infty(T^*\mathcal{M}^n)$, then the **total symbol** of L in the direction ζ is the bundle homomorphism

$$\sigma[L](\zeta) \doteqdot \sum_{|\alpha| \le k} L_a (\Pi_j \zeta^{\alpha_j}).$$

Note that $\sigma[L](\zeta)$ is a linear map of degree at most k in ζ. The **principal symbol** of L in the direction ζ is the bundle homomorphism

$$\hat{\sigma}[L](\zeta) \doteqdot \sum_{|\alpha| = k} L_a (\Pi_j \zeta^{\alpha_j}).$$

The principal symbol of a differential operator L captures algebraically those analytic properties of L that depend only on its highest derivatives. The following basic property of the total symbol will be useful: if $\mathcal{X}$ is another vector bundle over $\mathcal{M}^n$ and

$$M : C^\infty(\mathcal{W}) \to C^\infty(\mathcal{X})$$

is a linear differential operator of order ℓ, then the symbol of $M \circ L$ in the direction ζ is the bundle homomorphism

$$\sigma[M \circ L](\zeta) = \sigma[M](\zeta) \circ \sigma[L](\zeta) : \mathcal{V} \to \mathcal{X} \tag{3.9}$$

of degree at most $k+\ell$ in ζ.

We now regard the Ricci tensor $\operatorname{Rc}(g)$ as a nonlinear partial differential operator on the metric g,

$$\operatorname{Rc}_g \equiv \operatorname{Rc}(g) : C^\infty\left(S_2^+ T^*\mathcal{M}^n\right) \to C^\infty\left(S_2 T^*\mathcal{M}^n\right).$$

Here $C^\infty\left(S_2^+ T^*\mathcal{M}^n\right)$ denotes the space of positive definite symmetric $(2,0)$-tensors (in other words, Riemannian metrics) and $C^\infty\left(S_2 T^*\mathcal{M}^n\right)$ denotes the space of symmetric $(2,0)$-tensors. By Lemma 3.5, the linearization

$$D\left(\operatorname{Rc}_g\right) : C^\infty\left(S_2 T^*\mathcal{M}^n\right) \to C^\infty\left(S_2 T^*\mathcal{M}^n\right)$$

of the Ricci tensor is given by

$$\left[D\left(\operatorname{Rc}_g\right)(h)\right]_{jk} = \frac{1}{2} g^{pq}\left(\nabla_q \nabla_j h_{kp} + \nabla_q \nabla_k h_{jp} - \nabla_q \nabla_p h_{jk} - \nabla_j \nabla_k h_{qp}\right).$$

Let $\zeta \in C^\infty\left(T^*\mathcal{M}^n\right)$ be a covector. The **principal symbol** in the direction ζ of the linear partial differential operator $D\left(\operatorname{Rc}_g\right)$ is the bundle homomorphism

$$\hat{\sigma}\left[D\left(\operatorname{Rc}_g\right)\right](\zeta) : S_2 T^*\mathcal{M}^n \to S_2 T^*\mathcal{M}^n$$

obtained by replacing the covariant derivative ∇_i by the covector ζ_i, namely

$$\left[\hat{\sigma}\left[D\left(\operatorname{Rc}_g\right)\right](\zeta)(h)\right]_{jk} = \frac{1}{2} g^{qp}\left\{\begin{array}{c} \zeta_q \zeta_j h_{kp} + \zeta_q \zeta_k h_{jp} \\ -\zeta_q \zeta_p h_{jk} - \zeta_j \zeta_k h_{qp} \end{array}\right\}. \tag{3.10}$$

A linear partial differential operator L is said to be **elliptic** if its principal symbol $\hat{\sigma}[L](\zeta)$ is an isomorphism whenever $\zeta \neq 0$. A nonlinear operator N is said to be elliptic if its linearization DN is elliptic. There is a rich existence and regularity theory for linear elliptic operators. However, as we shall see below, the fact that the Ricci tensor is invariant under diffeomorphism,

$$\operatorname{Rc}\left(\varphi^* g\right) = \varphi^*\left(\operatorname{Rc}(g)\right), \tag{3.11}$$

implies that the principal symbol $\hat{\sigma}\left[D\left(\operatorname{Rc}_g\right)\right](\zeta)$ of the nonlinear partial differential operator Rc_g has a nontrivial kernel.

2.2. The Bianchi identities. To increase our understanding of the consequences of the diffeomorphism invariance of the Riemann curvature tensor, we shall show that it implies the first and second Bianchi identities. We follow a proof of Kazdan [**81**], which is implicit in work of Hilbert. We will see that the Bianchi identities are a consequence of the Jacobi identity for the Lie bracket of vector fields, which in turn follows from the diffeomorphism invariance of the bracket.

Let X, Y, and Z be arbitrary vector fields on a manifold $\mathcal{M}^n$. Let φ_t be the one-parameter group of diffeomorphisms generated by X such that $\varphi_0 = \mathrm{id}_{\mathcal{M}^n}$. Diffeomorphism invariance of the Lie bracket implies that

$$\varphi_t^*[Y, Z] = \left[\varphi_t^* X, \varphi_t^* Y\right].$$

Differentiating at $t = 0$, we obtain

$$\begin{aligned}[X,[Y,Z]] &= \frac{d}{dt}\varphi_t^*[Y,Z] = \left[\frac{d}{dt}(\varphi_t^* Y), \varphi_t^* Z\right] + \left[\varphi_t^* Y, \frac{d}{dt}(\varphi_t^* Z)\right] \\ &= [[X,Y],Z] + [Y,[X,Z]],\end{aligned}$$

which is equivalent to the Jacobi identity.

To illustrate the basic idea, we first consider the diffeomorphism invariance

$$R[\varphi_t^* g] = \varphi_t^*(R[g])$$

of the scalar curvature function R. Linearizing this identity, we get

$$DR_g(\mathcal{L}_X g) = \mathcal{L}_X R = \nabla_X R. \tag{3.12}$$

If the variation of g is h, then formula 3.7 of Lemma 3.7 may be written as

$$DR_g(h) = -g^{ij}g^{k\ell}(\nabla_i\nabla_j h_{k\ell} - \nabla_i\nabla_k h_{j\ell} + R_{ik}h_{j\ell}).$$

Substituting

$$h_{ij} = (\mathcal{L}_X g)_{ij} = \nabla_i X_j + \nabla_j X_i$$

and commuting covariant derivatives yields

$$\begin{aligned}DR_g(\mathcal{L}_X g) &= -2\Delta\nabla_i X^i - 2R_{ij}\nabla^i X^j + \nabla^i\nabla_j\nabla_i X^j + \nabla_i\nabla_j\nabla^j X^i \\ &= 2X^i\nabla^j R_{ij}.\end{aligned}$$

Combining this result with equation (3.12), we obtain

$$2X^i\nabla^j R_{ij} = X^i\nabla_i R.$$

Since X is arbitrary, this proves that the diffeomorphism invariance of R implies the contracted second Bianchi identity:

$$\nabla^j R_{ij} = \frac{1}{2}\nabla_i R. \tag{3.13}$$

Similarly, recalling formula (A.2) from Section 2 of Appendix A, we observe that the diffeomorphism invariance (3.11) of the Ricci tensor implies that

$$[D(\mathrm{Rc}_g)(\mathcal{L}_X g)]_{jk} = (\mathcal{L}_X \mathrm{Rc})_{jk} = \langle X, \nabla \mathrm{Rc}\rangle + R_{ik}\nabla_j X^i + R_{ji}\nabla_k X^i. \tag{3.14}$$

On the other hand, Lemma 3.5 implies

$$[D(\mathrm{Rc}_g)(h)]_{jk} = \frac{1}{2}\left(\nabla^i\nabla_j h_{ki} + \nabla^i\nabla_k h_{ji} - \Delta h_{jk} - \nabla_j\nabla_k h_i^i\right).$$

When $h = \mathcal{L}_X g$, this becomes

$$\begin{aligned}[D(\mathrm{Rc}_g)(\mathcal{L}_X g)]_{jk} = R_{ij}\nabla_k X^i + R_{ik}\nabla_j X^i + \frac{1}{2}\left(X^i\nabla_j R_{ik} + X^i\nabla_k R_{ij}\right) \\ + \frac{1}{2}\left(\nabla^\ell R_{\ell jki}X^i + \nabla^\ell R_{\ell kji}X^i\right).\end{aligned}$$

Combining this equation with (3.14) and again using the fact that X is arbitrary, we obtain the following consequence of the second Bianchi identity

$$\nabla_i R_{jk} = \frac{1}{2}\left(\nabla^\ell R_{\ell kji} + \nabla^\ell R_{\ell jki} + \nabla_j R_{ik} + \nabla_k R_{ij}\right). \tag{3.15}$$

Notice that this yields the contracted second Bianchi identity when one contracts on the indices j, k.

Finally, we observe that the diffeomorphism invariance of the Riemann curvature tensor implies that

$$\left[D\left(\operatorname{Rm}_g\right)\left(\mathcal{L}_X g\right)\right]_{ijk}^{\ell} = \left(\mathcal{L}_X \operatorname{Rm}\right)_{ijk}^{\ell},$$

where by formula (A.2), the components of the Lie derivative of the Riemann tensor are

$$\left(\mathcal{L}_X \operatorname{Rm}\right)_{ijk}^{\ell} = X^p \nabla_p R_{ijk}^{\ell} + R_{pjk}^{\ell} \nabla_i X^p + R_{ipk}^{\ell} \nabla_j X^p \tag{3.16a}$$

$$+ R_{ijp}^{\ell} \nabla_k X^p - R_{ijk}^{p} \nabla_p X^\ell. \tag{3.16b}$$

On the other hand, Lemma 3.3 implies that

$$\begin{aligned} 2\left[D\left(\operatorname{Rm}_g\right)(h)\right]_{ijk}^{\ell} = {} & \nabla_i \nabla_j h_k^\ell + \nabla_i \nabla_k h_j^\ell - \nabla_i \nabla^\ell h_{jk} \\ & - \nabla_j \nabla_i h_k^\ell - \nabla_j \nabla_k h_i^\ell + \nabla_j \nabla^\ell h_{ik}. \end{aligned}$$

Substituting $h = \mathcal{L}_X g$ and commuting derivatives, we find that

$$\begin{aligned} 2\left[D\left(\operatorname{Rm}_g\right)\left(\mathcal{L}_X g\right)\right]_{ijk}^{\ell} = {} & \nabla_i \nabla_j \nabla^\ell X_k - \nabla_i \nabla^\ell \nabla_j X_k - \nabla_j \nabla_i \nabla^\ell X_k \\ & - \nabla_j \nabla^\ell \nabla_i X_k + \nabla_i \nabla_k \nabla^\ell X_j - \nabla_i \nabla^\ell \nabla_k X_j \\ & - \nabla_j \nabla_k \nabla^\ell X_i + \nabla_j \nabla^\ell \nabla_k X_i - \nabla_i \nabla_j \nabla_k X^\ell \\ & + \nabla_i \nabla_k \nabla_j X^\ell + \nabla_j \nabla_i \nabla_k X^\ell - \nabla_j \nabla_k \nabla_i X^\ell \\ & + 2\nabla_i \nabla_j \nabla_k X^\ell - 2\nabla_j \nabla_i \nabla_k X^\ell. \end{aligned}$$

Thus if we rewrite (3.16) as

$$\begin{aligned} \left(\mathcal{L}_X \operatorname{Rm}\right)_{ijk}^{\ell} &= g^{\ell p}\left\{\begin{array}{c} X^q \nabla_q R_{ijkp} - \nabla_i X_q R_{kpj}^q - \nabla_j X_q R_{pki}^q \\ -\nabla_k X_q R_{ijp}^q - \nabla_q X_p R_{ijk}^q \end{array}\right\} \\ &= g^{\ell p}\left\{\begin{array}{c} -\nabla_i\left(R_{kpj}^q X_q\right) - \nabla_j\left(R_{pki}^q X_q\right) - \nabla_k X_q R_{ijp}^q \\ -\nabla_q X_p R_{ijk}^q + X_q\left(\nabla^q R_{ijkp} + \nabla_i R_{kpj}^q + \nabla_j R_{pki}^q\right) \end{array}\right\} \end{aligned}$$

and compare terms, we get

$$0 = [D(\mathrm{Rm}_g)(\mathcal{L}_X g)]^{\ell}_{ijk} - (\mathcal{L}_X \mathrm{Rm})^{\ell}_{ijk}$$
$$= \frac{1}{2} g^{\ell p} \left\{ \begin{array}{c} -\nabla_i \left[\left(R^q_{jpk} - R^q_{kpj} - R^q_{jkp} \right) X^q \right] \\ \\ +\nabla_j \left[\left(R^q_{ipk} - R^q_{kpi} - R^q_{ikp} \right) X_q \right] \\ \\ -2X^q \left(\nabla_q R_{ijkp} + \nabla_i R_{kpjq} + \nabla_j R_{pkiq} \right) \end{array} \right\}.$$

Now let $p \in \mathcal{M}^n$ be arbitrary. Since X is an arbitrary vector field, we may first prescribe $X(p) = 0$ and $\nabla_i X_j(p) = g_{ij}(p)$, obtaining

$$0 = -\left(R_{j\ell ki} - R_{k\ell ji} - R_{jk\ell i} \right) + \left(R_{i\ell kj} - R_{k\ell ij} - R_{ik\ell j} \right).$$

Recalling the symmetries $R_{ijk\ell} = -R_{jik\ell} = -R_{ij\ell k} = R_{k\ell ij}$, we see that this implies the first Bianchi identity

$$0 = R_{ijk\ell} + R_{ik\ell j} + R_{i\ell jk}. \tag{3.17}$$

Then if we choose X to be an element of a local orthonormal frame field and use the first Bianchi identity to cancel terms, we obtain the second Bianchi identity

$$0 = \nabla_q R_{ijk\ell} + \nabla_i R_{jqk\ell} + \nabla_j R_{qik\ell}. \tag{3.18}$$

2.3. The principal symbol of the differential operator $\mathrm{Rc}(g)$**.** We shall now explore the principal symbol $\hat{\sigma}[D(\mathrm{Rc}_g)]$ of the Ricci tensor, regarded as a nonlinear partial differential operator on the metric g.

Recall that the formal adjoint of the **divergence**

$$\delta_g : C^\infty(S_2 T^* \mathcal{M}^n) \to C^\infty(T^* \mathcal{M}^n) \tag{3.19a}$$

$$(\delta_g h)_k \doteqdot -g^{ij} \nabla_i h_{jk} \tag{3.19b}$$

with respect to the L^2 inner product

$$(V, W) = \int_{\mathcal{M}^n} \langle V, W \rangle \, d\mu_g \tag{3.20}$$

is the linear differential operator

$$\delta_g^* : C^\infty(T^* \mathcal{M}^n) \to C^\infty(S_2 T^* \mathcal{M}^n) \tag{3.21a}$$

$$\left(\delta_g^* X \right)_{jk} \doteqdot \frac{1}{2} \left(\nabla_j X_k + \nabla_k X_j \right) = \frac{1}{2} \left(\mathcal{L}_{X^\sharp} g \right)_{jk}. \tag{3.21b}$$

In other words, $\delta_g^* X$ is just a scalar multiple of the **Lie derivative** $\mathcal{L}_{X^\sharp} g$ of g with respect to the vector field $X^\sharp$ which is g-dual to X. The total symbol of δ_g^* in the direction ζ is the bundle homomorphism

$$\sigma\left[\delta_g^*\right](\zeta) : T^* \mathcal{M}^n \to S_2 T^* \mathcal{M}^n$$

that acts by

$$\left(\sigma\left[\delta_g^*\right](\zeta)(X) \right)_{jk} = \frac{1}{2} \left(\zeta_j X_k + \zeta_k X_j \right). \tag{3.22}$$

Consider the composition

$$D(\mathrm{Rc}_g) \circ \delta_g^* : C^\infty(T^*\mathcal{M}^n) \to C^\infty(S_2 T^*\mathcal{M}^n).$$

This is *a priori* a third-order differential operator, so its principal symbol

$$\hat{\sigma}\left[\left(D(\mathrm{Rc}_g) \circ \delta_g^*\right)\right](\zeta) : T^*\mathcal{M}^n \to S_2 T^*\mathcal{M}^n$$

is the degree 3 part of its total symbol. But as we observed in Section 2.2, the invariance (3.11) of the Ricci tensor under diffeomorphism implies that

$$\left[\left(D(\mathrm{Rc}_g) \circ \delta_g^*\right)(X)\right]_{jk} = \frac{1}{2}\left[\mathcal{L}_{X^\sharp}(\mathrm{Rc}_g)\right]_{jk}.$$

Since the right-hand-side involves only one derivative of X, we see that its total symbol is at most of degree 1 in ζ. In other words, the principal (degree 3) symbol $\hat{\sigma}\left[\left(D(\mathrm{Rc}_g) \circ \delta_g^*\right)\right](\zeta)$ is in fact the zero map. Since by (3.9), one has

$$0 = \hat{\sigma}\left[\left(D(\mathrm{Rc}_g) \circ \delta_g^*\right)\right](\zeta) = \hat{\sigma}\left[D(\mathrm{Rc}_g)\right](\zeta) \circ \hat{\sigma}\left[\delta_g^*\right](\zeta),$$

it follows that

$$\mathrm{im}\left(\hat{\sigma}\left[\delta_g^*\right](\zeta)\right) \subseteq \ker\left(\hat{\sigma}\left[D(\mathrm{Rc}_g)\right](\zeta)\right),$$

hence that $\hat{\sigma}\left[D(\mathrm{Rc}_g)\right](\zeta)$ has at least an n-dimensional kernel in each $(n(n+1)/2)$-dimensional fiber:

$$\dim\left(\ker\left(\hat{\sigma}\left[D(\mathrm{Rc}_g)\right](\zeta)\right)\right) \geq n. \tag{3.23}$$

REMARK 3.12. In Section 2.2, we proved that the diffeomorphism invariance of the Ricci tensor implies the contracted second Bianchi identity. They are actually equivalent, since commuting derivatives and applying the contracted second Bianchi identity reveals that

$$\begin{aligned}
\left[\left(D(\mathrm{Rc}_g) \circ \delta_g^*\right)(X)\right]_{jk} &= \frac{1}{4}\nabla^p\nabla_j\left(\nabla_k X_p + \nabla_p X_k\right) \\
&\quad + \frac{1}{4}\nabla^p\nabla_k\left(\nabla_j X_p + \nabla_p X_j\right) \\
&\quad - \frac{1}{4}\Delta\left(\nabla_j X_k + \nabla_k X_j\right) + \frac{1}{2}\nabla_j\nabla_k\left(\delta_g X\right) \\
&= \frac{1}{4}X^p\left(\nabla_j R_{kp} + \nabla_k R_{jp} + \nabla^q R_{qjkp} + \nabla^q R_{qkjp}\right) \\
&\quad + \frac{1}{2}\left(R_j^p\nabla_k X_p + R_k^p\nabla_j X_p\right) \\
&= \frac{1}{2}\left(X^p\nabla_p R_{jk} + R_j^p\nabla_k X_p + R_k^p\nabla_j X_p\right) \\
&= \frac{1}{2}\left[\mathcal{L}_{X^\sharp}(\mathrm{Rc}_g)\right]_{jk}.
\end{aligned}$$

Now consider the linear operator

$$B_g : C^\infty(S_2 T^*\mathcal{M}^n) \to C^\infty(T^*\mathcal{M}^n)$$

defined by

$$[B_g(h)]_k \doteqdot g^{ij}\left(\nabla_i h_{jk} - \frac{1}{2}\nabla_k h_{ij}\right). \tag{3.24}$$

The total symbol of B_g in the direction ζ is the bundle homomorphism

$$S_2T^*\mathcal{M}^n \to T^*\mathcal{M}^n$$

given by

$$(\sigma[B_g](\zeta)(h))_k = g^{ij}\left(\zeta_i h_{jk} - \frac{1}{2}\zeta_k h_{ij}\right).$$

Notice that $\sigma[B_g](\zeta)$ is of degree 1 in ζ. Writing the contracted second Bianchi identity (3.13) in the form

$$B_g(\mathrm{Rc}_g) = 0$$

shows that Rc_g belongs to the kernel of B_g. Linearizing, we obtain

$$(D(B_g))[\mathrm{Rc}(g+h)] + B_g[((D\,\mathrm{Rc}_g)(h))] = 0.$$

Since the differential operator $D(B_g)$ is of order 1, its degree 3 symbol is zero. But it is easy to check that the linear operator $B_g \circ D(\mathrm{Rc}_g)$ is of order 3. It follows that the principal (degree 3) symbol

$$\hat{\sigma}[B_g \circ D(\mathrm{Rc}_g)](\zeta) = \hat{\sigma}[B_g](\zeta) \circ \hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)$$

must be the zero map, hence that

$$\mathrm{im}(\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)) \subseteq \ker(\hat{\sigma}[B_g](\zeta)) \subseteq S_2T^*\mathcal{M}^n.$$

The considerations above have shown that for all ζ, the following short sequence constitutes an algebraic chain complex:

$$0 \longrightarrow T^*\mathcal{M}^n \xrightarrow{\hat{\sigma}[\delta_g^*](\zeta)} S_2T^*\mathcal{M}^n \xrightarrow{\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)} S_2T^*\mathcal{M}^n \xrightarrow{\hat{\sigma}[B_g](\zeta)} T^*\mathcal{M}^n \longrightarrow 0. \tag{3.25}$$

We shall now show that it is in fact a short exact sequence by proving that the nontrivial kernel of $\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)$ (hence the failure of $D(\mathrm{Rc}_g)$ to be elliptic) is due only to the invariance (3.11) of the Ricci tensor under diffeomorphism. We begin by studying the kernel of $\sigma[B_g](\zeta)$. Given a nonzero 1-form ζ, define

$$A_\zeta \doteqdot \{\zeta \otimes X + X \otimes \zeta - \langle \zeta, X\rangle g : X \in T^*\mathcal{M}^n\} \subseteq S_2T^*\mathcal{M}^n$$

and

$$K_\zeta \doteqdot \ker(\sigma[B_g](\zeta)) \subseteq S_2T^*\mathcal{M}^n.$$

Clearly, $\dim A_\zeta = n$ in each fiber where $\zeta \neq 0$. In any such fiber, the calculation

$$\langle \sigma[B_g](\zeta)(h), X\rangle = \frac{1}{2}\langle \zeta \otimes X + X \otimes \zeta - \langle \zeta, X\rangle g, h\rangle$$

proves that $K_\zeta = A_\zeta^\perp$, hence that $\dim K_\zeta = n(n-1)/2$. Now we further consider the kernel of $\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)$. For all $h \in K_\zeta$, it follows easily from (3.10) that

$$[\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)(h)]_{jk} = -\frac{1}{2}|\zeta|^2 h_{jk},$$

hence that

$$\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta) = -\frac{1}{2}|\zeta|^2 \mathrm{id}_{K_\zeta}.$$

In particular, $\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)|_{K_\zeta} : K_\zeta \to K_\zeta$ is an automorphism in each fiber where $\zeta \neq 0$, which proves that

$$\dim(\mathrm{im}(\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta))) \geq \dim K_\zeta = \frac{n(n-1)}{2}.$$

By (3.23), this implies that

$$\dim(\ker(\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta))) = n. \tag{3.26}$$

Before we conclude this discussion, it is worth comparing the global and infinitesimal viewpoints it affords. The facts that

$$\mathrm{im}\,\delta_g^* \perp \ker \delta_g$$

under the global inner product (3.20), and that

$$C^\infty(S_2T^*\mathcal{M}^n) = \mathrm{im}\,\delta_g^* \oplus \ker \delta_g$$

are classical. (See [**38**] and [**15**].) On the other hand, define

$$I_\zeta \doteqdot \mathrm{im}\left(\hat{\sigma}\left[\delta_g^*\right](\zeta)\right) = \ker(\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)).$$

Then in each fiber where $\zeta \neq 0$, we have

$$S_2T^*\mathcal{M}^n = I_\zeta + K_\zeta.$$

However, this decomposition is not orthogonal with respect to the inner product $\langle \cdot, \cdot \rangle$ induced by g on each fiber. This is because the operator $\hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)$ is not self-adjoint: indeed, if $h, k \in S_2T^*\mathcal{M}^n$, we have

$$\begin{aligned}&\langle \hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)(h), k\rangle - \langle h, \hat{\sigma}[D(\mathrm{Rc}_g)](\zeta)(k)\rangle \\ &= \frac{1}{2}\langle (\mathrm{tr}_g k) h - (\mathrm{tr}_g h) k, \zeta \otimes \zeta\rangle.\end{aligned}$$

3. The Ricci–DeTurck flow and its parabolicity

The short exact sequence (3.25) show that the nonlinear differential operator Rc_g is not an elliptic operator on the metric g. Because of this, we cannot immediately apply standard theory to conclude that a unique solution of the Ricci flow exists for a short time. In spite of this fact, the Ricci flow does enjoy short-time existence and uniqueness:

THEOREM 3.13 (Hamilton). *If $(\mathcal{M}^n, g_0)$ is a closed Riemannian manifold, there exists a unique solution $g(t)$ to the Ricci flow defined on some positive time interval $[0, \varepsilon)$ such that $g(0) = g_0$.*

REMARK 3.14. We shall see in Corollary 7.7 that the lifetime of a maximal solution is bounded below by

$$\frac{c}{\max_{\mathcal{M}^n} |\mathrm{Rm}\,[g_0]|_{g_0}},$$

where c is a universal constant depending only on n.

Hamilton's original proof [58] of this result relied on the Nash–Moser inverse function theorem and was lengthy and technically difficult. However, he showed that the degeneracy of the equation is due only to its diffeomorphism invariance. Soon after, DeTurck found a much simpler proof of Theorem 3.13. DeTurck showed that it is possible to modify the Ricci flow and thereby obtain a *parabolic* PDE by a clever trick: one modifies the right-hand side of the equation by adding a term which is a Lie derivative of the metric with respect to a certain vector field which in turn depends on the metric. Remarkably, one then can obtain a solution to the original Ricci flow equation by pulling back the solution of the modified flow by appropriately chosen diffeomorphisms.

To motivate how the **DeTurck trick** is done, we rewrite the linearization of the Ricci tensor as follows:

(3.27)
$$\begin{aligned} -2\left[D\left(\mathrm{Rc}_g\right)(h)\right]_{jk} = {} & \Delta\, h_{jk} \\ & - \nabla_j(g^{pq}\nabla_q h_{pk}) - \nabla_k(g^{pq}\nabla_q h_{pj}) + \nabla_j\nabla_k(g^{pq}h_{qp}) \\ & + 2g^{qp}R^r_{qjk}\,h_{rp} - g^{qp}R_{jp}\,h_{kq} - g^{qp}R_{kp}\,h_{jq}. \end{aligned}$$

Define a 1-form $V = V(g,h)$ by $V = B_g(h)$ where B_g is given by (3.24), so that

$$V_k \doteqdot g^{pq}\nabla_p h_{qk} - \frac{1}{2}\nabla_k(g^{pq}h_{pq}).$$

We may then express the linearization of the Ricci tensor as

$$-2\left[D\left(\mathrm{Rc}_g\right)(h)\right]_{jk} = \Delta\, h_{jk} - \nabla_j V_k - \nabla_k V_j + S_{jk}, \tag{3.28}$$

where the symmetric 2-tensor $S = S(g,h)$ is defined by

$$S_{jk} \doteqdot 2g^{qp}R^r_{qjk}\,h_{rp} - g^{qp}R_{jp}\,h_{kq} - g^{qp}R_{kp}\,h_{jq}\,.$$

Note that S involves no derivatives of h. The 1-form V may be rewritten as

$$V_k = \frac{1}{2}g^{pq}\left(\nabla_p h_{qk} + \nabla_q h_{pk} - \nabla_k h_{pq}\right) = g^{pq}g_{kr}\left[D\left(\Gamma_g\right)(h)\right]^r_{pq},$$

where

$$D\left(\Gamma_g\right) : C^\infty\left(S_2T^*\mathcal{M}^n\right) \to C^\infty\left(S_2T^*\mathcal{M}^n \otimes T\mathcal{M}^n\right)$$

denotes the linearization of the Levi-Civita connection and is given by

$$\left[D\left(\Gamma_g\right)(h)\right]^k_{ij} = \left.\frac{\partial}{\partial s}\right|_{s=0} \Gamma^k_{ij}$$

when

$$\left.\frac{\partial}{\partial s}\right|_{s=0} g = h.$$

Now let $\tilde{\Gamma}$ be a fixed torsion-free connection. (For instance, we could take $\tilde{\Gamma}$ to be the Levi-Civita connection of a fixed background metric $\tilde{g}$.) The considerations above lead us to define a vector field $W = W(g, \tilde{\Gamma})$ by

$$W^k = g^{pq}\left(\Gamma_{pq}^k - \tilde{\Gamma}_{pq}^k\right). \tag{3.29}$$

Since the difference of two connections is a tensor, W is a globally well-defined vector field (independent of the coordinates used to describe it locally). Because W involves one derivative of the metric g, the map

$$P \equiv P(\tilde{\Gamma}) : C^\infty\left(S_2T^*\mathcal{M}^n\right) \to C^\infty\left(S_2T^*\mathcal{M}^n\right)$$

that corresponds to taking the Lie derivative of g with respect to W, namely

$$P(g) \doteqdot \mathcal{L}_W g,$$

is a second-order partial differential operator. The linearization of P is

$$[DP(h)]_{jk} = \nabla_j V_k + \nabla_k V_j + T_{jk}, \tag{3.30}$$

where T_{jk} is a linear first-order expression in h. Comparing the equation above with equation (3.28) leads us to consider a modified Ricci operator, namely

$$Q \doteqdot -2\operatorname{Rc} + P : C^\infty\left(S_2T^*\mathcal{M}^n\right) \to C^\infty\left(S_2T^*\mathcal{M}^n\right).$$

From (3.28) and (3.30), the linearization of Q satisfies

$$DQ(h) = \Delta\, h + U,$$

where $U_{jk} = -2S_{jk} + T_{jk}$ is a linear first-order expression in h. Hence the principal symbol of DQ is given by

$$\sigma\left[DQ\right](\zeta)(h) = |\zeta|^2\, h, \tag{3.31}$$

which implies in particular that Q is elliptic.

DeTurck's strategy for constructing a unique short-time solution $\bar{g}\,(t)$ of the Ricci flow

$$\frac{\partial}{\partial t}\bar{g} = -2\operatorname{Rc}(\bar{g}), \tag{3.32a}$$

$$\bar{g}\,(0) = g_0 \tag{3.32b}$$

on a closed manifold $\mathcal{M}^n$ proceeds in four steps:

STEP 1. One defines the **Ricci–DeTurck flow** by

$$\frac{\partial}{\partial t}g_{ij} = -2R_{ij} + \nabla_i W_j + \nabla_j W_i, \tag{3.33a}$$

$$g\,(0) = g_0, \tag{3.33b}$$

where the time-dependent 1-form W is g-dual to the vector field (3.29). In particular,

$$W_j = g_{jk}W^k \doteqdot g_{jk}g^{pq}\left(\Gamma_{pq}^k - \tilde{\Gamma}_{pq}^k\right) \tag{3.34}$$

depends on $g\,(t)$, its Levi-Civita connection $\Gamma\,(t)$, and the fixed background connection $\tilde{\Gamma}$. It follows from (3.31) that the Ricci–DeTurck flow is a strictly parabolic system of partial differential equations. It is a standard result that for any smooth initial metric g_0, there exists $\varepsilon > 0$ depending on g_0 such that a unique smooth solution $g(t)$ to (3.33) will exist for a short time $0 \leq t < \varepsilon$.

STEP 2. One observes that the one-parameter family of vector fields $W(t)$ defined by (3.29) exists as long as the solution $g(t)$ of (3.33) exists. Then one defines a 1-parameter family of maps $\varphi_t : \mathcal{M}^n \to \mathcal{M}^n$ by

$$\frac{\partial}{\partial t}\varphi_t(p) = -W(\varphi_t(p), t) \tag{3.35}$$
$$\varphi_0 = \mathrm{id}_{\mathcal{M}^n}.$$

Notice that $\varphi_t(p)$ is constructed by solving a non-autonomous ODE at each $p \in \mathcal{M}^n$. Because $\mathcal{M}^n$ is compact, it follows from Lemma 3.15 (below) that all $\varphi_t(p)$ exist and remain diffeomorphisms for as long as the solution $g(t)$ exists, namely for $t \in [0, \varepsilon)$.

STEP 3. One observes that the family of metrics

$$\bar{g}(t) \doteqdot \varphi_t^* g(t) \qquad (0 \leq t < \varepsilon) \tag{3.36}$$

is a solution to the Ricci flow (3.32). Indeed, we have $\bar{g}(0) = g(0) = g_0$, because $\varphi_0 = \mathrm{id}_{\mathcal{M}^n}$. Then we compute that

$$\begin{aligned}
\frac{\partial}{\partial t}(\varphi_t^* g(t)) &= \left.\frac{\partial}{\partial s}\right|_{s=0} (\varphi_{t+s}^* g(t+s)) \\
&= \varphi_t^*\left(\frac{\partial}{\partial t} g(t)\right) + \left.\frac{\partial}{\partial s}\right|_{s=0} (\varphi_{t+s}^* g(t)) \\
&= \varphi_t^*\left(-2\,\mathrm{Rc}(g(t)) + \mathcal{L}_{W(t)} g(t)\right) \\
&\quad + \left.\frac{\partial}{\partial s}\right|_{s=0} \left[(\varphi_t^{-1}\circ\varphi_{t+s})^* \varphi_t^* g(t)\right] \\
&= -2\,\mathrm{Rc}(\varphi_t^* g(t)) + \varphi_t^*\left(\mathcal{L}_{W(t)} g(t)\right) - \mathcal{L}_{[(\varphi_t^{-1})_* W(t)]}(\varphi_t^* g(t)) \\
&= -2\,\mathrm{Rc}(\varphi_t^* g(t)).
\end{aligned}$$

The equality on the penultimate line follows from the identity

$$\left.\frac{\partial}{\partial s}\right|_{s=0} (\varphi_t^{-1}\circ\varphi_{t+s}) = (\varphi_t^{-1})_*\left(\left.\frac{\partial}{\partial s}\right|_{s=0}\varphi_{t+s}\right) = (\varphi_t^{-1})_* W(t).$$

Hence $\bar{g}(t) \doteqdot \varphi_t^* g(t)$ is a solution of the Ricci flow for $t \in [0, \varepsilon)$. This completes the proof of the existence claim in Theorem 3.13.

STEP 4. All that remains is to show that $\bar{g}(t) \doteqdot \varphi_t^* g(t)$ is the *unique* solution of the Ricci flow with initial data $\bar{g}(0) = g_0$. This will be easier to do after we have the machinery of the harmonic map heat flow at our disposal. Consequently, we shall postpone the proof of uniqueness: see Step 4 of the existence and uniqueness proof outlined in Section 4.4.

3.1. Existence of the DeTurck diffeomorphisms. In this subsection, we establish the existence of a family $\{\varphi_t\}$ of diffeomorphisms solving the ODE (3.35).

LEMMA 3.15. *If $\{X_t : 0 \leq t < T \leq \infty\}$ is a continuous time-dependent family of vector fields on a compact manifold $\mathcal{M}^n$, then there exists a one-parameter family of diffeomorphisms $\{\varphi_t : \mathcal{M}^n \to \mathcal{M}^n : 0 \leq t < T \leq \infty\}$ defined on the same time interval such that*

$$\frac{\partial \varphi_t}{\partial t}(x) = X_t[\varphi_t(x)] \tag{3.37a}$$

$$\varphi_0(x) = x \tag{3.37b}$$

for all $x \in \mathcal{M}^n$ and $t \in [0, T)$.

PROOF. We may assume there is $t_0 \in [0, T)$ such that $\varphi_s(y)$ exists for all $0 \leq s \leq t_0$ and $y \in \mathcal{M}^n$. Let $t_1 \in (t_0, T)$ be given. We shall show that φ_t exists for all $t \in [t_0, t_1]$. Since t_1 is arbitrary, this implies the lemma. Given any $x_0 \in \mathcal{M}^n$, choose local coordinate systems $(\mathcal{U}, \mathbf{x})$ and $(\mathcal{V}, \mathbf{y})$ such that $x_0 \in \mathcal{U}$ and $\varphi_{t_0}(x_0) \in \mathcal{V}$. As long as $x \in \mathcal{U}$ and $\varphi_t(x) \in \mathcal{V}$, the equation (3.37a) is equivalent to

$$\begin{aligned}\frac{\partial}{\partial t}\left[\mathbf{y} \circ \varphi_t \circ \mathbf{x}^{-1}(p)\right] &= \mathbf{y}_*\left[\frac{\partial \varphi_t}{\partial t}\left[\mathbf{x}^{-1}(p)\right]\right] \\ &= \left(\mathbf{y}_* X_t \circ \mathbf{y}^{-1}\right)\left(\mathbf{y} \circ \varphi_t \circ \mathbf{x}^{-1}(p)\right)\end{aligned}$$

for $p \in \mathbf{x}(\mathcal{U})$ such that $\varphi_t \circ \mathbf{x}^{-1}(p) \in \mathcal{V}$. Setting $\mathbf{z}_t = \mathbf{y} \circ \varphi_t \circ \mathbf{x}^{-1}$ and $\mathbf{F}_t = \mathbf{y}_* X_t \circ \mathbf{y}^{-1}$, we get

$$\frac{\partial}{\partial t}\mathbf{z}_t = \mathbf{F}_t(\mathbf{z}_t)$$

where $\mathbf{z}_t$ and $\mathbf{F}_t$ are time-dependent maps between subsets of $\mathbb{R}^n$. Thus we see that (3.37a) is locally equivalent to a nonlinear ODE in $\mathbb{R}^n$. Hence for all $x \in \mathcal{U}$ such that $\varphi_{t_0}(x) \in \mathcal{V}$, a unique solution to (3.37a) exists for a short time $t \in [t_0, t_0 + \varepsilon)$. Since the vector fields X_t are uniformly bounded on the compact set $\mathcal{M}^n \times [t_0, t_1]$, there exists an $\bar{\varepsilon} > 0$ independent of $x \in \mathcal{M}^n$ and $t \in [t_0, t_1]$ such that a unique solution $\varphi_t(x)$ exists for $t \in [t_0, t_0 + \bar{\varepsilon}]$. Since the same claim holds for the flow starting at $\varphi_{t+\bar{\varepsilon}}(x)$, a simple iteration finishes the argument. □

REMARK 3.16. The lemma may fail to be true if $\mathcal{M}^n$ is not compact. For example, if $\mathcal{M}^1 = \mathbb{R}$ and $X_t(u) = u^2$, then we have the ODE

$$\begin{aligned}\frac{\partial}{\partial t}\varphi_t(x) &= [\varphi_t(x)]^2 \\ \varphi_0(x) &= x\end{aligned}$$

whose solution is $\varphi_t(0) \equiv 0$ if $x = 0$ and

$$\varphi_t(x) = \frac{1}{x^{-1} - t}$$

if $x \neq 0$. If $x < 0$, the maximal solution exists for $t \in (x^{-1}, \infty)$. If $x > 0$, the maximal solution exists for $t \in (-\infty, x^{-1})$. In this case, we note that only φ_0 is a diffeomorphism of $\mathbb{R}$, whereas for $t > 0$ we have the diffeomorphism

$$\varphi_t : \left(-\infty, t^{-1}\right) \to \left(-t^{-1}, \infty\right).$$

(To see this, note that if $x > 0$, then $t < x^{-1}$ implies that $x < t^{-1}$.)

3.2. DeTurck's notations. This subsection (which is not needed for the sequel) describes DeTurck's original formulation of the Ricci–DeTurck flow.

Define the algebraic **Einstein operator**

$$G : C^{\infty}\left(S_2 T^* \mathcal{M}^n\right) \to C^{\infty}\left(S_2 T^* \mathcal{M}^n\right)$$

by

$$G\left(v\right)_{ij} \doteqdot v_{ij} - \frac{1}{2}\left(\operatorname{tr}_g v\right) g_{ij},$$

noting that G takes the Ricci tensor to the Einstein tensor:

$$G\left(\mathrm{Rc}\right) = \mathrm{Rc} - \frac{1}{2} R g.$$

Recall the divergence δ and its formal L^2 adjoint δ^* introduced in (3.19) and (3.21), respectively, noting that δ^* of a 1-form X is just a scalar multiple of the Lie derivative of g with respect to the vector field metrically dual to X. Using these operators, we can rewrite the linearization of the Ricci tensor (3.27) as

$$\left[D\left(\mathrm{Rc}_g\right)\left(h\right)\right]_{jk} = -\frac{1}{2}\left[\Delta h_{jk} + 2R_{pjkq}h^{pq} - R_j^\ell h_{\ell k} - R_k^\ell h_{j\ell}\right] + Y_{jk},$$

where

$$\begin{aligned}
Y_{jk} &= \frac{1}{2}\nabla_j \nabla^\ell h_{\ell k} + \frac{1}{2}\nabla_k \nabla^\ell h_{\ell j} - \frac{1}{2}\nabla_k \nabla_j \left(\operatorname{tr}_g h\right) \\
&= \frac{1}{2}\nabla_j\left(\nabla^\ell h_{\ell k} - \frac{1}{2}\nabla_k\left(\operatorname{tr}_g h\right)\right) + \frac{1}{2}\nabla_k\left(\nabla^\ell h_{\ell j} - \frac{1}{2}\nabla_j\left(\operatorname{tr}_g h\right)\right) \\
&= \frac{1}{2}\nabla_j \nabla^\ell\left(h_{\ell k} - \frac{1}{2}\left(\operatorname{tr}_g h\right) g_{\ell k}\right) + \frac{1}{2}\nabla_k \nabla^\ell\left(h_{\ell j} - \frac{1}{2}\left(\operatorname{tr}_g h\right) g_{\ell j}\right) \\
&= -\left[\delta^*\left(\delta\left[G\left(h\right)\right]\right)\right]_{jk}.
\end{aligned}$$

Thus by recalling the Lichnerowicz Laplacian defined in (3.6), we obtain the following result.

LEMMA 3.17. *The variation of the Ricci tensor has the form*

$$D\left(\mathrm{Rc}_g\right)\left(h\right) = -\frac{1}{2}\left(\Delta_L h\right) - \left[\delta^*\left(\delta\left[G\left(h\right)\right]\right)\right].$$

We now reconsider the Ricci–DeTurck flow (3.33) and rewrite the 1-forms $W\left(t\right)$ by computing at any point $p \in \mathcal{M}^n$ in a coordinate system chosen so that $\left(\Gamma_{g(t)}\right)_{ij}^k\left(p\right) = 0$. Since we have

$$\left(\nabla_{g(t)}\right)_i Q = \partial_i Q$$

at p for any tensor Q, it follows from (3.34) that the identity

$$\begin{aligned}
W_j(t) &= g_{jk}(t)\, g^{pq}(t) \left(\left(\Gamma_{g(t)}\right)_{ij}^{k} - \left(\Gamma_{\tilde{g}}\right)_{pq}^{k} \right) \\
&= -g_{jk}(t)\, g^{pq}(t) \left(\Gamma_{\tilde{g}}\right)_{pq}^{k} \\
&= -\frac{1}{2} g_{jk}(t)\, g^{pq}(t)\, \tilde{g}^{k\ell} \left(\left(\nabla_{g(t)}\right)_p \tilde{g}_{\ell q} + \left(\nabla_{g(t)}\right)_q \tilde{g}_{p\ell} - \left(\nabla_{g(t)}\right)_\ell \tilde{g}_{pq} \right) \\
&= g_{jk}(t)\, \tilde{g}^{k\ell} \left(\delta \left[G(\tilde{g})\right]\right)_\ell
\end{aligned}$$

holds at p. Thus if one defines

$$\tilde{g}^{-1} : C^\infty (T^* \mathcal{M}^n) \to C^\infty (T^* \mathcal{M}^n)$$

by

$$\left(\tilde{g}^{-1} Q\right)_j \doteqdot g_{jk}(t)\, \tilde{g}^{k\ell} Q_\ell,$$

one can write

$$W(t) = \tilde{g}^{-1} \left(\delta \left[G(\tilde{g})\right]\right).$$

In particular,

$$\nabla_i W_j + \nabla_j W_i = 2 \left(\delta^* W\right)_{ij} = 2 \left(\delta^* \left[\tilde{g}^{-1} \left(\delta \left[G(\tilde{g})\right]\right)\right]\right)_{ij}.$$

Hence the Ricci–DeTurck flow (3.33) can be written in the alternate form

$$\frac{\partial}{\partial t} g = -2 \left\{ \operatorname{Rc} - \delta^* \left[\tilde{g}^{-1} \left(\delta \left[G(\tilde{g})\right]\right)\right] \right\}, \tag{3.38a}$$

$$g(0) = g_0. \tag{3.38b}$$

4. The Ricci–DeTurck flow in relation to the harmonic map flow

It is a surprising and useful observation that equation (3.35) for the diffeomorphisms φ_t may be interpreted in terms of the harmonic map heat flow.

4.1. The harmonic map heat flow. Let $(\mathcal{M}^n, g)$ and $(\mathcal{N}^m, h)$ be two Riemannian manifolds, and let $f : \mathcal{M}^n \to \mathcal{N}^m$ be a smooth map. The derivative of f is

$$df \equiv f_* \in C^\infty \left(T^* \mathcal{M}^n \otimes f^* T \mathcal{N}^m\right),$$

where $f^* T\mathcal{N}^m$ is the pullback bundle over $\mathcal{M}^n$. Using local coordinates $\{x^i\}$ on $\mathcal{M}^n$ and $\{y^\alpha\}$ on $\mathcal{N}^m$, we denote the Levi-Civita connection of g by $\Gamma(g)_{ij}^k$ and that of h by $\Gamma(h)_{\alpha\beta}^\gamma$. Then

$$df = (df)_j^\alpha \left(dx^j \otimes \frac{\partial}{\partial y^\alpha}\right) \equiv \frac{\partial f^\alpha}{\partial x^j} \left(dx^j \otimes \frac{\partial}{\partial y^\alpha}\right).$$

The induced connection

$$\nabla : C^\infty \left(T^* \mathcal{M}^n \otimes f^* T \mathcal{N}^m\right) \to C^\infty \left(T^* \mathcal{M}^n \otimes T^* \mathcal{M}^n \otimes f^* T \mathcal{N}^m\right)$$

is determined by

$$\left(f^* \Gamma\right)_{i\beta}^\gamma = \frac{\partial f^\alpha}{\partial x^i} \left(\Gamma_h \circ f\right)_{\alpha\beta}^\gamma.$$

Hence

$$\nabla(df) = \sum_{i,j,\alpha} (\nabla df)_{ij}^{\alpha}\, dx^i \otimes dx^j \otimes \frac{\partial}{\partial y^{\alpha}}$$

where

$$\begin{aligned}(\nabla df)_{ij}^{\alpha} &\equiv \nabla_i d_j f^{\alpha} \\ &= \frac{\partial}{\partial x^i}\left(\frac{\partial f^{\alpha}}{\partial x^j}\right) - (\Gamma_g)_{ij}^{k}\left(\frac{\partial f^{\alpha}}{\partial x^k}\right) + \frac{\partial f^{\beta}}{\partial x^i}(\Gamma_h \circ f)_{\beta\gamma}^{\alpha}\left(\frac{\partial f^{\gamma}}{\partial x^j}\right).\end{aligned}$$

The **harmonic map Laplacian** with respect to the domain metric g and codomain metric h is defined to be the trace

$$\Delta_{g,h} f \doteqdot \operatorname{tr}_g(\nabla(df)) \in C^{\infty}(f^* T\mathcal{N}^m),$$

namely

$$\Delta_{g,h} f = \left(g^{ij}\nabla_i d_j f^{\gamma}\right)\frac{\partial}{\partial y^{\gamma}}.$$

In components,

$$(\Delta_{g,h} f)^{\gamma} = g^{ij}\left[\frac{\partial^2 f^{\gamma}}{\partial x^i \partial x^j} - (\Gamma_g)_{ij}^{k}\frac{\partial f^{\gamma}}{\partial x^k} + \left((\Gamma_h)_{\alpha\beta}^{\gamma} \circ f\right)\frac{\partial f^{\alpha}}{\partial x^i}\frac{\partial f^{\beta}}{\partial x^j}\right]. \tag{3.39}$$

Given $f_0 : \mathcal{M}^n \to \mathcal{N}^m$, the **harmonic map flow** introduced by Eells and Sampson is

$$\frac{\partial f}{\partial t} = \Delta_{g,h} f, \tag{3.40a}$$

$$f(0) = f_0. \tag{3.40b}$$

This is a parabolic equation, so a unique solution exists for a short time.

When f is a diffeomorphism, the Laplacian $\Delta_{g,h} f$ may be rewritten in a useful way:

LEMMA 3.18. *Let $f : (\mathcal{M}^n, g) \to (\mathcal{N}^n, h)$ be a diffeomorphism of Riemannian manifolds. Then*

$$(\Delta_{g,h} f)^{\gamma}(x) = \left[\left(f^{-1}\right)^* g\right]^{\alpha\beta}\left(-\Gamma\left(\left(f^{-1}\right)^* g\right)_{\alpha\beta}^{\gamma} + \Gamma(h)_{\alpha\beta}^{\gamma}\right)(f(x)).$$

PROOF. We first compute the Christoffel symbols of the pull-back of a metric in local coordinates $\{x^i\}$ on $\mathcal{M}^n$ and $\{y^{\alpha}\}$ on $\mathcal{N}^n$. Let $\varphi : \mathcal{M}^n \to \mathcal{N}^n$

be a diffeomorphism. Then if κ is a metric on $\mathcal{N}^n$, we have

$$\begin{aligned}
\Gamma_{ij}^k(\varphi^*\kappa)\frac{\partial}{\partial x^k} &\doteqdot \nabla(\varphi^*\kappa)_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} \\
&= \left(\varphi^{-1}\right)_*\left(\nabla(\kappa)_{\varphi_*\left(\frac{\partial}{\partial x^i}\right)}\varphi_*\left(\frac{\partial}{\partial x^j}\right)\right) \\
&= \left(\varphi^{-1}\right)_*\left(\nabla(\kappa)_{\frac{\partial\varphi^\alpha}{\partial x^i}\frac{\partial}{\partial y^\alpha}}\left(\frac{\partial\varphi^\beta}{\partial x^j}\frac{\partial}{\partial y^\beta}\right)\right) \\
&= \left(\varphi^{-1}\right)_*\left(\frac{\partial^2\varphi^\beta}{\partial x^i\partial x^j}\frac{\partial}{\partial y^\beta}+\Gamma(\kappa)_{\alpha\beta}^\gamma\frac{\partial\varphi^\alpha}{\partial x^i}\frac{\partial\varphi^\beta}{\partial x^j}\frac{\partial}{\partial y^\gamma}\right) \\
&= \left(\frac{\partial^2\varphi^\beta}{\partial x^i\partial x^j}\frac{\partial\left(\varphi^{-1}\right)^k}{\partial y^\beta}+\Gamma(\kappa)_{\alpha\beta}^\gamma\frac{\partial\varphi^\alpha}{\partial x^i}\frac{\partial\varphi^\beta}{\partial x^j}\frac{\partial\left(\varphi^{-1}\right)^k}{\partial y^\gamma}\right)\frac{\partial}{\partial x^k}.
\end{aligned}$$

Hence

$$\Gamma_{ij}^k(\varphi^*\kappa)\frac{\partial\varphi^\gamma}{\partial x^k} = \frac{\partial^2\varphi^\gamma}{\partial x^i\partial x^j}+\Gamma(\kappa)_{\alpha\beta}^\gamma\frac{\partial\varphi^\alpha}{\partial x^i}\frac{\partial\varphi^\beta}{\partial x^j}.$$

Since

$$\kappa^{\alpha\beta} = (\varphi^*\kappa)^{ij}\frac{\partial\varphi^\alpha}{\partial x^i}\frac{\partial\varphi^\beta}{\partial x^j},$$

we multiply by $(\varphi^*\kappa)^{ij}$ to obtain

$$-\kappa^{\alpha\beta}\Gamma(\kappa)_{\alpha\beta}^\gamma = (\varphi^*\kappa)^{ij}\left(\frac{\partial^2\varphi^\gamma}{\partial x^i\partial x^j}-\Gamma_{ij}^k(\varphi^*\kappa)\frac{\partial\varphi^\gamma}{\partial x^k}\right). \tag{3.41}$$

Putting $\varphi = f$ and $\kappa = \left(f^{-1}\right)^* g$ in (3.41) and substituting into (3.39), we get

$$(\Delta f)^\gamma = \left(\left(f^{-1}\right)^* g\right)^{\alpha\beta}\left[-\Gamma\left(\left(f^{-1}\right)^* g\right)_{\alpha\beta}^\gamma+\Gamma(h)_{\alpha\beta}^\gamma\right], \tag{3.42}$$

whence the lemma follows. □

COROLLARY 3.19. *Taking $\mathcal{M}^n = \mathcal{N}^n$ and f to be the identity, we have*

$$(\Delta_{g,h}\,\mathrm{id})^\gamma = g^{\alpha\beta}\left(-\Gamma(g)_{\alpha\beta}^\gamma+\Gamma(h)_{\alpha\beta}^\gamma\right).$$

A particular case of this result is of some independent interest:

COROLLARY 3.20. *If $(\mathcal{M}^n, g)$ is a manifold of strictly positive Ricci curvature, then*

$$\mathrm{id} : (\mathcal{M}^n, g) \to (\mathcal{M}^n, \mathrm{Rc}(g))$$

is a harmonic map.

PROOF. Note that $h \doteqdot \mathrm{Rc}(g)$ is a metric, and let A denote the globally-defined tensor field

$$A = \Gamma(h) - \Gamma(g).$$

At the origin of a normal coordinate system for g, we have

$$A_{ij}^k = \frac{1}{2}\left(\mathrm{Rc}^{-1}\right)^{k\ell}\left(\nabla_i R_{j\ell} + \nabla_j R_{i\ell} - \nabla_\ell R_{ij}\right).$$

Since both sides of this identity are components of tensors, it holds everywhere. Applying the contracted second Bianchi identity to the result of the previous corollary, we get

$$(\Delta_{g,h}\,\mathrm{id})^k = g^{ij}A_{ij}^k = \frac{1}{2}\left(\mathrm{Rc}^{-1}\right)^{k\ell}\left[g^{ij}\left(\nabla_i R_{j\ell} + \nabla_j R_{i\ell} - \nabla_\ell R_{ij}\right)\right] = 0.$$

□

REMARK 3.21. If $\mathrm{Rc} < 0$, then the result holds for the target manifold $(\mathcal{M}^n, -\mathrm{Rc}(g))$.

4.2. The cross curvature tensor. Corollary 3.20 raises some interesting albeit tangential questions. We shall introduce these here only briefly, because they are not needed for the remainder of this volume. However, we plan to provide a more complete discussion in the next volume. (Also see [**32**] for more details.)

PROBLEM 3.22. Given a Riemannian manifold $(\mathcal{M}^n, g)$, does there exist a metric $\gamma \neq g$ on $\mathcal{M}^n$ such that $\mathrm{id} : (\mathcal{M}^n, \gamma) \to (\mathcal{M}^n, g)$ is a harmonic map?

Assuming the sectional curvatures of g have a definite sign, the answer in dimension three is given by the cross curvature tensor. If we choose a local orthonormal frame in which the sectional curvatures are $\kappa_1 \doteqdot R_{2332}$, $\kappa_2 \doteqdot R_{1331}$, and $\kappa_3 \doteqdot R_{1221}$, then the Ricci tensor corresponds to the matrix

$$\mathrm{Rc} = \begin{pmatrix} \kappa_2+\kappa_3 & & \\ & \kappa_1+\kappa_3 & \\ & & \kappa_1+\kappa_2 \end{pmatrix}$$

and the **cross curvature tensor** $c = c(g)$ of $(\mathcal{M}^3, g)$ corresponds to the matrix

$$c = \begin{pmatrix} \kappa_2\kappa_3 & & \\ & \kappa_1\kappa_3 & \\ & & \kappa_1\kappa_2 \end{pmatrix}.$$

More explicitly, one can write c in terms of the Einstein tensor

$$E = \mathrm{Rc} - \frac{1}{2}Rg,$$

which has eigenvalues $-\kappa_1$, $-\kappa_2$, and $-\kappa_3$. We have

$$\begin{aligned} c_{ij} &= \left(\frac{\det E}{\det g}\right)\left(E^{-1}\right)_{ij} = \frac{1}{2}\mu^{ipq}\mu^{jrs}E_{pr}E_{qs} \\ &= \frac{1}{8}\mu^{pqk}\mu^{rs\ell}R_{i\ell pq}R_{kjrs} \end{aligned}$$

where μ^{ijk} are the components of the volume form with indices raised and normalized so that $\mu^{123} = 1$ in an oriented orthonormal frame.

EXERCISE 3.23. Verify that

$$\left(c^{-1}\right)^{ij} \nabla_i c_{jk} = \frac{1}{2} \left(c^{-1}\right)^{ij} \nabla_k c_{ij}.$$

Use this to prove that if the sectional curvatures of $\left(\mathcal{M}^3, g\right)$ are negative everywhere (or positive everywhere), then c_g is a metric and

$$\mathrm{id} : \left(\mathcal{M}^3, c_g\right) \to \left(\mathcal{M}^3, g\right)$$

is a harmonic map.

Since we are interested in curvature flows, one may consider the cross-curvature flow. The **cross curvature flow**, is defined on a 3-manifold of negative sectional curvature by

$$\frac{\partial}{\partial t} g = c(g),$$

and on a 3-manifold of positive sectional curvature by

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

When $\mathcal{M}^3$ is closed, it can be proven that a solution exists for a short time. The main open problem is to determine the long-time existence and convergence properties of the cross curvature flow. When the initial metric has negative sectional curvature, the following monotonicity formulas have been proven by Hamilton and the first author.

PROPOSITION 3.24. *As long as a solution to the cross curvature flow exists, one has*

$$\frac{\partial}{\partial t} \mathrm{Vol}(E) \geq 0$$

and

$$\frac{d}{dt} \int_{\mathcal{M}^3} \left[\frac{1}{3} \left(g^{ij} E_{ij}\right) - \left(\frac{\det E}{\det g} \right)^{1/3} \right] d\mu \leq 0.$$

Note that $\mathrm{Vol}(E)$ is scale-invariant in g. The integrand

$$\frac{1}{3} \left(g^{ij} E_{ij}\right) - \left(\det E / \det g\right)^{1/3}$$

is the difference between the arithmetic and geometric means of $-\kappa_1$, $-\kappa_2$, and $-\kappa_3$. It vanishes identically if and only if the sectional curvature is constant. The monotonicity of the integral is surprising, because the metric g is expanding, and the integral scales like $g^{1/2}$.

CONJECTURE 3.25. *If $\left(\mathcal{M}^3, g_0\right)$ is a manifold of negative sectional curvature, the cross curvature flow with initial condition g_0 should exist for all positive time and converge.*

REMARK 3.26. Ben Andrews [4] has observed that when the universal cover of the initial Riemannian manifold may be isometrically embedded as a hypersurface in Euclidean or Minkowski 4-space, the Gauss curvature flow of that hypersurface effects the cross curvature flow of the induced metric; in this case, he can prove convergence results.

4.3. The DeTurck diffeomorphisms and the harmonic map flow. Recall that the diffeomorphisms φ_t defined by (3.29) and (3.35) satisfy

$$\frac{\partial}{\partial t}\varphi_t = -W \circ \varphi = g^{pq}\left(-\Gamma_{pq}^k + \tilde{\Gamma}_{pq}^k\right).$$

If $\tilde{\Gamma}$ is the Levi-Civita connection of a metric $\tilde{g}$, then it follows from Lemma 3.18 and equation (3.36) that

$$g^{pq}\left(-\Gamma_{pq}^k + \tilde{\Gamma}_{pq}^k\right) = \left(\left(\varphi_t^{-1}\right)^* \bar{g}\right)^{pq}\left(-\Gamma_{pq}^k + \tilde{\Gamma}_{pq}^k\right) = \Delta_{\bar{g}(t),\tilde{g}}\ \varphi_t$$

Therefore

$$\frac{\partial}{\partial t}\varphi_t = \Delta_{\bar{g}(t),\tilde{g}}\ \varphi_t\,.$$

Thus we come to the interesting observation that the DeTurck diffeomorphisms satisfy the harmonic map heat flow:

LEMMA 3.27. *The diffeomorphisms φ_t used to obtain the solution $\bar{g}(t)$ of the Ricci flow (3.32) from the solution $g(t)$ of the Ricci–DeTurck flow (3.33) satisfy the harmonic map flow with domain metric $\bar{g}(t)$ and codomain metric $\tilde{g}$.*

4.4. An alternate approach to existence and uniqueness. In [**63**], Hamilton introduced a variant of DeTurck's argument for short-time existence and uniqueness of the Ricci flow on a closed manifold $\mathcal{M}^n$. This method uses both the Ricci–DeTurck flow and the harmonic map flow, and provides a particularly elegant proof of uniqueness. Hamilton's strategy is as follows:

STEP 1. Fix a background metric $\tilde{g}$ on $\mathcal{M}^n$.

STEP 2. If a solution to the Ricci flow (3.32) exists, denote it by $\bar{g}(t)$. In this case, let φ_t denote the solution of the harmonic map heat flow

$$\frac{\partial}{\partial t}\varphi_t = \Delta_{\bar{g}(t),\tilde{g}}\ \varphi_t$$

with respect to $\bar{g}(t)$ and $\tilde{g}$. Note that, *a priori,* we only know that the maps φ_t exist and remain diffeomorphisms for a short time if $\bar{g}(t)$ exists.

STEP 3. Observe that if $\bar{g}(t)$ exists, then

$$g(t) \doteqdot (\varphi_t)_* \bar{g}(t)$$

is a solution of the Ricci–DeTurck flow (3.33). Indeed, using (3.35) we compute that

$$\begin{aligned}\frac{\partial}{\partial t} g(t) &= \frac{\partial}{\partial t}\left(\left(\varphi_t^{-1}\right)^* \bar{g}(t)\right) \\ &= \left(\varphi_t^{-1}\right)^* \left(\frac{\partial}{\partial t}\bar{g}(t)\right) + \mathcal{L}_{(\varphi_t)_*\left(\frac{\partial}{\partial t}\varphi_t\right)}\left[\left(\varphi_t^{-1}\right)^* \bar{g}(t)\right] \\ &= \left(\varphi_t^{-1}\right)^* \left(-2\operatorname{Rc}\left[\bar{g}(t)\right]\right) + \mathcal{L}_{(\varphi_t)_*(\varphi_t^*[W(t)])}\left[g(t)\right] \\ &= -2\operatorname{Rc}\left[g(t)\right] + \mathcal{L}_{W(t)}\left[g(t)\right].\end{aligned}$$

But since (3.33) is parabolic, we know that a unique solution $g(t)$ does exist. And once we have $g(t)$, we can obtain the diffeomorphisms φ_t by solving the non-autonomous ODE (3.35) at each point, as in Step 2 of DeTurck's method. So

$$\bar{g}(t) = \varphi_t^* g(t)$$

does exist after all, and solves the Ricci flow.

STEP 4. We now prove that a solution of the Ricci flow is uniquely determined by its initial data. Suppose $\bar{g}_1(t)$ and $\bar{g}_2(t)$ are two solutions of the Ricci flow (3.32) on a common time interval. Let $(\varphi_1)_t$ denote the solution of the harmonic map flow with respect to $\bar{g}_1(t)$ and $\tilde{g}$. Let $(\varphi_2)_t$ denote the solution of the harmonic map flow with respect to $\bar{g}_2(t)$ and $\tilde{g}$. Then

$$g_1(t) \doteqdot ((\varphi_1)_t)_* \bar{g}_1(t) \quad \text{and} \quad g_2(t) \doteqdot ((\varphi_2)_t)_* \bar{g}_2(t)$$

are both solutions of the Ricci–DeTurck flow (3.33). Because $g_1(0) = g_2(0)$ and (3.33) enjoys unique solutions, we have $g_1(t) = g_2(t)$ for as long as both exist. But then both $(\varphi_1)_t$ and $(\varphi_2)_t$ are solutions of the ODE

$$\frac{\partial}{\partial t}(\varphi_i)_t(p) = -W((\varphi_i)_t(p), t) \qquad (i = 1, 2)$$

generated by the same vector field

$$W^k = g^{pq}\left(\Gamma_{pq}^k - \tilde{\Gamma}_{pq}^k\right).$$

Hence $(\varphi_1)_t = (\varphi_2)_t$ for as long as they are both defined, which implies in particular that

$$\bar{g}_1(t) = (\varphi_1)_t^* g_1(t) = (\varphi_2)_t^* g_2(t) = \bar{g}_2(t).$$

5. The Ricci flow regarded as a heat equation

Heuristically, the Ricci flow may be interpreted as a nonlinear heat equation for a Riemannian metric. In this section, we develop this point of view, which is related to DeTurck's trick. The first step is to consider the Ricci tensor in local harmonic coordinates.

DEFINITION 3.28. Local coordinates $\{x^i\}$ are called **harmonic coordinates** if each coordinate function x^i is harmonic:

$$0 = \Delta x^i = g^{jk}\left(\partial_j\partial_k - \Gamma_{jk}^\ell \partial_\ell\right)x^i = -g^{jk}\Gamma_{jk}^i.$$

Existence of harmonic coordinates is a straightforward consequence of standard existence theory for elliptic partial differential equations.

THEOREM 3.29 (Local solvability of elliptic PDE). *Let*

$$F(x, u, \partial u, \ldots, \partial^k u)$$

be a C^∞ function, and let $x_0 \in \mathbb{R}^n$. If there exists a function $v \in C^k(U)$ defined in a neighborhood U of x_0 such that

$$F(x, v, \partial v, \ldots, \partial^k v)(x_0) = 0$$

and if F is elliptic at v (that is, if the linearization of F at v is elliptic), then there exists a neighborhood V of x_0 and a function u such that

$$F(x, u, \partial u, \ldots, \partial^k u)(x) = 0$$

for all $x \in V$ and

$$\partial^j u(x_0) = \partial^j v(x_0)$$

for all multi-indices j such that $|j| < k$.

COROLLARY 3.30. *Given a point $p \in M$, there exist harmonic coordinates defined in some neighborhood of p.*

PROOF. Let $\{x^i\}$ denote geodesic coordinates centered at p. Because $\Gamma_{ij}^k(p) = 0$, we have

$$\Delta x^i(p) = \left[g^{ij}(\partial_i \partial_j - \Gamma_{ij}^k \partial_k) x^i \right](p) = 0.$$

By the theorem above, there exist a neighborhood V of p and functions $\{u^i\}$ such that $\Delta u^i = 0$ holds in V, with $u^i(p) = 0$ and $\partial_j u^i(p) = \partial_j x^i(p) = \delta_j^i$. As a consequence, $\{u^i\}$ are harmonic coordinates in some neighborhood $W \subset V$ of p. □

Standard elliptic regularity theory implies that if the metric g is in a Hölder class $C^{k,\alpha}$ (respectively C^ω) in the original coordinates, then the harmonic coordinates themselves are in the class $C^{k+1,\alpha}$ (respectively C^ω). In particular, the map from the original coordinates to the harmonic coordinates is of class $C^{k+1,\alpha}$ (C^ω). Since transforming a tensor involves at most first derivatives of the map, it follows readily that g itself has optimal regularity in harmonic coordinates.

LEMMA 3.31. *If a metric g is of class $C^{k,\alpha}$ (C^ω) in a given coordinate chart, then any tensor of class $C^{k,\alpha}$ (C^ω) in that chart belongs to the same Hölder class in harmonic coordinates. In particular, g itself is of class $C^{k,\alpha}$ (C^ω).*

The insight into the Ricci flow which harmonic coordinates provide stems from the following observation.

LEMMA 3.32. *In harmonic coordinates, the Ricci tensor is given by*

$$-2R_{ij} = \Delta(g_{ij}) + Q_{ij}\left(g^{-1}, \partial g\right),$$

where $\Delta(g_{ij})$ denotes the Laplacian of the component g_{ij} of the metric regarded as a scalar function, and Q denotes a sum of terms which are quadratic in the metric inverse g^{-1} and its first derivatives ∂g.

PROOF. The Ricci tensor is given in local coordinates by

$$\begin{aligned}-2R_{jk} = -2R^q_{qjk} &= -2\left(\partial_q\Gamma^q_{jk} - \partial_j\Gamma^q_{qk} + \Gamma^p_{jk}\Gamma^q_{qp} - \Gamma^p_{qk}\Gamma^q_{jp}\right)\\ &= -\partial_q\left[g^{qr}\left(\partial_j g_{kr} + \partial_k g_{jr} - \partial_r g_{jk}\right)\right]\\ &\quad + \partial_j\left[g^{qr}\left(\partial_q g_{kr} + \partial_k g_{qr} - \partial_r g_{qk}\right)\right] + \Gamma^p_{jk}\Gamma^q_{qp} - \Gamma^p_{qk}\Gamma^q_{jp}.\end{aligned}$$

Using Lemma 3.1, we write this as

$$\begin{aligned}-2R_{jk} &= g^{qr}\left(\partial_q\partial_r g_{jk} - \partial_q\partial_k g_{jr} + \partial_j\partial_k g_{qr} - \partial_j\partial_r g_{qk}\right) + g^{-1} * g^{-1} * \partial g * \partial g\\ &= \Delta(g_{jk}) - g^{qr}\partial_k\left(\Gamma^s_{qr}g_{sj}\right) - g^{qr}\partial_j\left(\Gamma^s_{qr}g_{sk}\right) + g^{-1} * g^{-1} * \partial g * \partial g,\end{aligned}$$

where $g^{-1} * g^{-1} * \partial g * \partial g$ denotes a sum of contractions whose exact formula is irrelevant. In harmonic coordinates, the second and third terms of the last line vanish, and we obtain

$$-2R_{jk} = \Delta(g_{jk}) + g^{-1} * g^{-1} * \partial g * \partial g.$$

□

COROLLARY 3.33. *In harmonic coordinates, the Ricci flow takes the form of a system of nonlinear heat equations for the components of the metric tensor:*

$$\frac{\partial}{\partial t}g_{ij} = \Delta(g_{ij}) + Q_{ij}\left(g^{-1}, \partial g\right). \tag{3.43}$$

REMARK 3.34. The reader is cautioned that equation (3.43) is not tensorial. Since the metric is parallel, it is of course clear that $\Delta g \equiv 0$. Moreover, one should not in general expect coordinates which are harmonic at time $t = 0$ to be harmonic at times $t > 0$.

Notes and commentary

Hamilton's original proof of short-time existence [**58**] used the Nash–Moser implicit function theorem. [**57**] DeTurck's proof [**36, 37**] is a substantial simplification. Other presentations of DeTurck's proof may be found in §6 of [**63**] and in [**112**].

The Lichnerowicz Laplacian acting on symmetric 2-tensors (which as we saw above is essentially the linearization of the Ricci flow operator) is formally the same as the Hodge–de Rham Laplacian acting on 2-forms. See Remark A.3 in Appendix A.

CHAPTER 4

Maximum principles

Maximum principles are among the most important tools in the study of second-order parabolic differential equations. Moreover, they are robust enough to be effective on manifolds. For this reason, they are particularly useful for the study of the Ricci flow. In this chapter, we review some basic theorems of this type. For pedagogical reasons, the chapter is organized as follows: we begin with the most elementary results and then introduce progressively more general ones. We conclude by stating some powerful and advanced theorems and then presenting a brief discussion of strong maximum principles.

1. Weak maximum principles for scalar equations

1.1. The heat equation with a gradient term. The heat equation is the prototype for parabolic equations. One of the most important properties it satisfies is the **maximum principle**. On a compact manifold, the maximum principle says that whatever pointwise bounds hold for a smooth solution to the heat equation at the initial time $t = 0$ persist for times $t > 0$.

PROPOSITION 4.1 (scalar maximum principle, first version: pointwise bounds are preserved). *Let $u : \mathcal{M}^n \times [0,T) \to \mathbb{R}$ be a C^2 solution to the heat equation*

$$\frac{\partial u}{\partial t} = \Delta_g u$$

on a closed Riemannian manifold, where Δ_g denotes the Laplacian with respect to the metric g. If there are constants $C_1 \leq C_2 \in \mathbb{R}$ such that $C_1 \leq u(x,0) \leq C_2$ for all $x \in \mathcal{M}^n$, then $C_1 \leq u(x,t) \leq C_2$ for all $x \in \mathcal{M}^n$ and $t \in [0,T)$.

The proposition follows immediately from a more general result in which one allows a gradient term on the right-hand side. Let $g(t) : t \in [0,T)$ be a 1-parameter family of Riemannian metrics and $X(t) : t \in [0,T)$ a 1-parameter family of vector fields on a closed manifold $\mathcal{M}^n$. We say a C^2 function $u : \mathcal{M}^n \times [0,T) \to \mathbb{R}$ is a **supersolution** to the heat-type equation

$$\frac{\partial v}{\partial t} = \Delta_{g(t)} v + \langle X, \nabla v \rangle$$

at $(x,t) \in \mathcal{M}^n \times [0,T)$ if

$$\frac{\partial u}{\partial t}(x,t) \geq \left(\Delta_{g(t)} u\right)(x,t) + \langle X, \nabla u \rangle (x,t). \tag{4.1}$$

THEOREM 4.2 (scalar maximum principle, second version: lower bounds are preserved for supersolutions). *Let $g(t) : t \in [0,T)$ be a 1-parameter family of Riemannian metrics and $X(t) : t \in [0,T)$ a 1-parameter family of vector fields on a closed manifold $\mathcal{M}^n$. Let $u : \mathcal{M}^n \times [0,T) \to \mathbb{R}$ be a C^2 function. Suppose that there exists $\alpha \in \mathbb{R}$ such that $u(x,0) \geq \alpha$ for all $x \in \mathcal{M}^n$ and that u is a supersolution of the heat equation at any $(x,t) \in \mathcal{M}^n \times [0,T)$ such that $u(x,t) < \alpha$. Then $u(x,t) \geq \alpha$ for all $x \in \mathcal{M}^n$ and $t \in [0,T)$.*

The idea of the proof is eminently simple: in essence, one just applies the first and second derivative tests in calculus.

PROOF. If $H : \mathcal{M}^n \times [0,T) \to \mathbb{R}$ is a C^2 function and (x_0, t_0) is a point and time where H attains its minimum among all points and earlier times, namely

$$H(x_0, t_0) = \min_{\mathcal{M}^n \times [0,t_0]} H,$$

then

$$\frac{\partial H}{\partial t}(x_0, t_0) \leq 0, \tag{4.2a}$$

$$\nabla H(x_0, t_0) = 0, \tag{4.2b}$$

$$\Delta H(x_0, t_0) \geq 0. \tag{4.2c}$$

Consider the function H defined by

$$H(x,t) \doteqdot [u(x,t) - \alpha] + \varepsilon t + \varepsilon,$$

where ε is any positive number. Note that $H \geq \varepsilon > 0$ at $t = 0$. Using (4.1), we find that H satisfies

$$\frac{\partial H}{\partial t} \geq \Delta H + \langle X, \nabla H \rangle + \varepsilon \tag{4.3}$$

at any point where $u < \alpha$. To prove the theorem, it will suffice to prove the claim that $H > 0$ for all $t \in [0,T)$. To prove that claim, suppose that $H \leq 0$ at some $(x_1, t_1) \in \mathcal{M}^n \times [0,T)$. Then since $\mathcal{M}^n$ is compact and $H > 0$ at $t = 0$, there is a first time $t_0 \in (0, t_1]$ such that there exists a point $x_0 \in \mathcal{M}^n$ such that $H(x_0, t_0) = 0$. Then since

$$u(x_0, t_0) = \alpha - \varepsilon t_0 - \varepsilon < \alpha,$$

combining (4.2) with (4.3) implies that

$$0 \geq \frac{\partial H}{\partial t}(x_0, t_0) \geq \Delta H(x_0, t_0) + \langle X, \nabla H \rangle (x_0, t_0) + \varepsilon \geq \varepsilon > 0.$$

This contradiction proves the claim and hence the theorem. $\square$

1.2. The heat equation with a linear reaction term. More generally, one may add in a reaction term. We first consider the case where the reaction term is linear. In particular, if $g(t)$ is a 1-parameter family of metrics, $X(t)$ is a 1-parameter family of vector fields, and $\beta : \mathcal{M}^n \times [0,T) \to \mathbb{R}$ is a given function, we say u is a **supersolution** to the linear heat equation

$$\frac{\partial v}{\partial t} = \Delta_{g(t)} v + \langle X, \nabla v \rangle + \beta v$$

at any points and times where

$$\frac{\partial u}{\partial t} \geq \Delta_{g(t)} u + \langle X, \nabla u \rangle + \beta u. \tag{4.4}$$

PROPOSITION 4.3 (scalar maximum principle, third version: linear reaction terms preserve lower bounds). *Let $u : \mathcal{M}^n \times [0,T) \to \mathbb{R}$ be a C^2 supersolution to (4.4) on a closed manifold. Suppose that for each $\tau \in [0,T)$, there exists a constant $C_\tau < \infty$ such that $\beta(x,t) \leq C_\tau$ for all $x \in \mathcal{M}^n$ and $t \in [0,\tau]$. If $u(x,0) \geq 0$ for all $x \in \mathcal{M}^n$, then $u(x,t) \geq 0$ for all $x \in \mathcal{M}^n$ and $t \in [0,T)$.*

PROOF. Given $\tau \in (0,T)$, define

$$J(x,t) \doteqdot e^{-C_\tau t} u(x,t),$$

where C_τ is as in the hypothesis. One computes that

$$\frac{\partial J}{\partial t} \geq \Delta_{g(t)} J + \langle X, \nabla J \rangle + (\beta - C_\tau) J.$$

Since $\beta - C_\tau \leq 0$ on $\mathcal{M}^n \times [0,\tau]$, one has

$$\frac{\partial J}{\partial t}(x,t) \geq \left(\Delta_{g(t)} J\right)(x,t) + \langle X, \nabla J \rangle (x,t)$$

for all $(x,t) \in \mathcal{M}^n \times [0,\tau]$ such that $J(x,t) \leq 0$. By Theorem 4.2, one concludes that $J \geq 0$ on $\mathcal{M}^n \times [0,\tau)$. Hence $u \geq 0$ on $\mathcal{M}^n \times [0,\tau)$. Since $\tau \in (0,T)$ was arbitrary, the proposition follows. □

1.3. The heat equation with a nonlinear reaction term. Now we treat the case where the reaction term is nonlinear. In particular, we consider the semilinear heat equation

$$\frac{\partial v}{\partial t} = \Delta_{g(t)} v + \langle X, \nabla v \rangle + F(v) \tag{4.5}$$

where $g(t)$ is a 1-parameter family of metrics, $X(t)$ is a 1-parameter family of vector fields, and $F : \mathbb{R} \to \mathbb{R}$ is a locally Lipschitz function. We say u is a **supersolution** of (4.5) if

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

and a **subsolution** if

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

THEOREM 4.4 (scalar maximum principle, fourth version: ODE gives pointwise bounds for PDE). *Let $u : \mathcal{M}^n \times [0,T) \to \mathbb{R}$ be a C^2 supersolution to (4.5) on a closed manifold. Suppose there exists $C_1 \in \mathbb{R}$ such that that $u(x,0) \geq C_1$ for all $x \in \mathcal{M}^n$, and let φ_1 be the solution to the associated ordinary differential equation*

$$\frac{d\varphi_1}{dt} = F(\varphi_1)$$

satisfying

$$\varphi_1(0) = C_1.$$

Then

$$u(x,t) \geq \varphi_1(t)$$

for all $x \in \mathcal{M}^n$ and $t \in [0,T)$ such that $\varphi_1(t)$ exists.
Similarly, suppose that u is a subsolution to (4.4) and $u(x,0) \leq C_2$ for all $x \in M$. Let $\varphi_2(t)$ be the solution to the initial value problem

$$\frac{d\varphi_2}{dt} = F(\varphi_2)$$
$$\varphi_2(0) = C_2.$$

Then

$$u(x,t) \leq \varphi_2(t)$$

for all $x \in \mathcal{M}^n$ and $t \in [0,T)$ such that $\varphi_2(t)$ exists.

PROOF. We will just prove the lower bound, since the upper bound is similar. We compute that

$$\frac{\partial}{\partial t}(u - \varphi_1) \geq \Delta(u - \varphi_1) + \langle X, \nabla(u - \varphi_1)\rangle + F(u) - F(\varphi_1).$$

The assumptions on the initial data imply that $u - \varphi_1 \geq 0$ at $t = 0$. We claim that $u - \varphi_1 \geq 0$ for all $t \in [0,T)$. To prove the claim, let $\tau \in (0,T)$ be given. Since $\mathcal{M}^n$ is compact, there exists a constant $C_\tau < \infty$ such that $|u(x,t)| \leq C_\tau$ and $|\varphi_1(t)| \leq C_\tau$ for all $(x,t) \in \mathcal{M}^n \times [0,\tau]$. Since F is locally Lipschitz, there exists a constant $L_\tau < \infty$ such that

$$|F(v) - F(w)| \leq L_\tau |v - w|$$

for all $v, w \in [-C_\tau, C_\tau]$. Hence we have

$$\frac{\partial}{\partial t}(u - \varphi_1) \geq \Delta(u - \varphi_1) + \langle X, \nabla(u - \varphi_1)\rangle - L_\tau \operatorname{sgn}(u - \varphi_1) \cdot (u - \varphi_1)$$

on $\mathcal{M}^n \times [0,\tau]$, where $\operatorname{sgn}(\cdot) \in \{-1,0,1\}$ denotes the signum function. Applying Proposition 4.3 with $\beta \doteqdot -L_\tau \operatorname{sgn}(u - \varphi_1)$, we obtain

$$u - \varphi_1 \geq 0$$

on $\mathcal{M}^n \times [0,\tau]$. This proves the claim. The theorem follows, since $\tau \in (0,T)$ was arbitrary. □

REMARK 4.5. In what follows, we shall freely apply Theorem 4.4 without explicit reference by simply invoking *the (parabolic) maximum principle.*

2. Weak maximum principles for tensor equations

The maximum principle is extremely robust: it applies to general classes of second-order parabolic equations and even to some systems, such as the Ricci flow. The following result is a simple example of the maximum principle for systems. Recall that one writes $A \geq 0$ for a symmetric 2-tensor A if the quadratic form induced by A is positive semidefinite.

THEOREM 4.6 (tensor maximum principle, first version: non-negativity is preserved). *Let $g(t)$ be a smooth 1-parameter family of Riemannian metrics on a closed manifold $\mathcal{M}^n$. Let $\alpha(t) \in C^\infty(T^*\mathcal{M}^n \otimes_S T^*\mathcal{M}^n)$ be a symmetric $(2,0)$-tensor satisfying the semilinear heat equation*

$$\frac{\partial}{\partial t}\alpha \geq \Delta_{g(t)}\alpha + \beta,$$

where $\beta(\alpha, g, t)$ is a symmetric $(2,0)$-tensor which is locally Lipschitz in all its arguments and satisfies the **null eigenvector assumption** *that*

$$\beta(V,V)(x,t) = \left(\beta_{ij}V^iV^j\right)(x,t) \geq 0$$

whenever $V(x,t)$ is a null eigenvector of $\alpha(t)$, that is whenever

$$\left(\alpha_{ij}V^j\right)(x,t) = 0.$$

If $\alpha(0) \geq 0$ (that is, if $\alpha(0)$ is positive semidefinite), then $\alpha(t) \geq 0$ for all $t \geq 0$ such that the solution exists.

This result is in a sense a prototype for more advanced tensor maximum principles that we shall encounter later. So before giving the full proof, we will describe the strategy and key concepts behind it.

IDEA OF THE PROOF. Recall that one proves the scalar maximum principle (for example, Theorem 4.2) by a purely local argument at a point and the first time when the solution becomes zero. The tensor maximum principle essentially follows from the scalar maximum principle by applying the tensor to a fixed vector field. To illustrate this, suppose that $\alpha > 0$ for all $0 \leq t < t_0$, but that (x_0, t_0) is a point and time and $v \in T_{x_0}\mathcal{M}^n$ is a vector such that

$$\alpha_{ij}v^j(x_0,t_0) = 0.$$

Then $\alpha_{ij}W^iW^j(x,t) \geq 0$ for all $x \in \mathcal{M}^n$, $t \in [0,t_0]$, and tangent vectors $W \in T_x\mathcal{M}^n$. One wants to extend v to a vector field V defined in a space-time neighborhood of (x_0,t_0) such that $V(x_0,t_0) = v$ and

$$\frac{\partial V}{\partial t}(x_0,t_0) = 0, \tag{4.6a}$$

$$\nabla V(x_0,t_0) = 0, \tag{4.6b}$$

$$\Delta V(x_0,t_0) = 0. \tag{4.6c}$$

This may be accomplished by parallel translation (with respect to $g(t_0)$) of v in space along geodesic rays (with respect to $g(t_0)$) emanating from x_0, and then taking V to be independent of time. To see that the Laplacian of V vanishes at (x_0, t_0), choose any frame $\{e_i \in T_{x_0}\mathcal{M}^n\}_{i=1}^n$ which is orthonormal with respect to $g(t_0)$ and parallel translate it in a spatial neighborhood along geodesic rays emanating from x_0. With respect to this local orthonormal frame, the Laplacian of V is

$$\Delta V(x_0, t_0) = \sum_{i=1}^{n} \left[\nabla_{e_i}(\nabla_{e_i} V) - \nabla_{\nabla_{e_i} e_i} V\right](x_0, t_0)$$
$$= \sum_{i=1}^{n} \left[\nabla_{e_i}\vec{0} - \nabla_{\vec{0}} V\right](x_0, t_0) = \vec{0}.$$

Then at any point in the space-time neighborhood of (x_0, t_0), one has

$$\frac{\partial}{\partial t}\left(\alpha_{ij} V^i V^j\right) = \left(\frac{\partial}{\partial t}\alpha_{ij}\right) V^i V^j = \left(\Delta \alpha_{ij} + \beta_{ij}\right) V^i V^j.$$

Now since $\left(\alpha_{ij} V^i V^j\right)(x_0, t_0) = 0$ and $\left(\alpha_{ij} V^i V^j\right)(x, t_0) \geq 0$ for all x in a spatial neighborhood of x_0, one has

$$\Delta\left(\alpha_{ij} V^i V^j\right) \geq 0.$$

But equations (4.6) imply that at (x_0, t_0),

$$\Delta\left(\alpha_{ij} V^i V^j\right) = \left(\Delta \alpha_{ij}\right) V^i V^j.$$

Combining these observations with the assumption

$$\left(\beta_{ij} V^i V^j\right)(x_0, t_0) \geq 0$$

shows that

$$\frac{\partial}{\partial t}\left(\alpha_{ij} V^i V^j\right) = \Delta\left(\alpha_{ij} V^i V^j\right) + \beta_{ij} V^i V^j \geq 0$$

at (x_0, t_0). Hence, if $\alpha_{ij} V^i V^j$ ever becomes zero, it cannot decrease further.

Now we give the rest of the argument.

Proof of Theorem 4.6. Given any $\tau \in (0, T)$, we shall show that there exists $\delta \in (0, \tau]$ such that for all $t_0 \in [0, \tau - \delta]$, if $\alpha \geq 0$ at $t = t_0$, then $\alpha \geq 0$ on $M \times [t_0, t_0 + \delta]$. The theorem follows easily from this statement.

Fix any $t_0 \in [0, \tau - \delta]$. For $0 < \varepsilon \leq 1$, consider the modified $(2,0)$-tensor A_ε defined for all $x \in \mathcal{M}^n$ and $t \in [t_0, t_0 + \delta]$ by

$$A_\varepsilon(x, t) \doteqdot \alpha(x, t) + \varepsilon\left[\delta + (t - t_0)\right] \cdot g(x, t),$$

where $\delta > 0$ will be chosen below. As in the proof of the scalar maximum principle, the term $\varepsilon\delta g$ makes A_ε strictly positive definite at $t = t_0$ because

$$A_\varepsilon(x, t_0) = \alpha(x, t_0) + \varepsilon\delta g(x, t_0) > 0,$$

and the term $\varepsilon(t - t_0) g$ will make

$$\frac{\partial}{\partial t} A_\varepsilon > \frac{\partial}{\partial t}\alpha$$

for $t \in [t_0, t_0 + \delta]$ when we choose $\delta > 0$ sufficiently small, depending on

$$\max_{\mathcal{M}^n \times [0,\tau]} \left| \frac{\partial}{\partial t} g \right| .$$

The evolution of A_ε is given by

$$\frac{\partial}{\partial t} A_\varepsilon = \frac{\partial}{\partial t} \alpha + \varepsilon g + \varepsilon \left[\delta + (t - t_0) \right] \frac{\partial g}{\partial t}.$$

Since $\Delta A_\varepsilon = \Delta \alpha$, we have

$$\frac{\partial}{\partial t} A_\varepsilon \geq \Delta A_\varepsilon + \beta + \varepsilon g + \varepsilon \left[\delta + (t - t_0) \right] \frac{\partial g}{\partial t},$$

which we rewrite as

$$\frac{\partial}{\partial t} A_\varepsilon \geq \Delta A_\varepsilon + \beta \left(A_\varepsilon, g, t \right) + \left[\beta \left(\alpha, g, t \right) - \beta \left(A_\varepsilon, g, t \right) \right] \tag{4.7a}$$

$$+ \varepsilon g + \varepsilon \left[\delta + (t - t_0) \right] \frac{\partial g}{\partial t}. \tag{4.7b}$$

We first choose $\delta_0 > 0$ depending on $g(t)$ for $t \in [0, \tau]$ to be small enough so that on $\mathcal{M}^n \times [t_0, t_0 + \delta_0]$, we have

$$\frac{\partial}{\partial t} g \geq - \frac{1}{4\delta_0} g.$$

This implies in particular that

$$\varepsilon g + \varepsilon \left[\delta_0 + (t - t_0) \right] \frac{\partial g}{\partial t} \geq \frac{1}{2} \varepsilon g \tag{4.8}$$

on $\mathcal{M}^n \times [t_0, t_0 + \delta_0]$. Since β is locally Lipschitz, there exists a constant K depending on the bounds for α and g on $\mathcal{M}^n \times [0, \tau]$ (but not on ε) which is large enough that

$$\beta \left(\alpha, g, t \right) - \beta \left(A_\varepsilon, g, t \right) \geq -K \varepsilon \left[\delta_0 + (t - t_0) \right] g \geq -2K \varepsilon \delta_0 g$$

on $\mathcal{M}^n \times [t_0, t_0 + \delta_0]$. Then if we choose $\delta \in (0, \delta_0)$ so small that

$$\delta < \frac{1}{4K},$$

we have

$$\beta \left(\alpha, g, t \right) - \beta \left(A_\varepsilon, g, t \right) > - \frac{1}{2} \varepsilon g. \tag{4.9}$$

Hence combining (4.8) and (4.9) to (4.7) shows that

$$\frac{\partial}{\partial t} A_\varepsilon > \Delta A_\varepsilon + \beta \left(A_\varepsilon, g, t \right) \tag{4.10}$$

on $\mathcal{M}^n \times [t_0, t_0 + \delta]$. We claim that $A_\varepsilon > 0$ on $\mathcal{M}^n \times [t_0, t_0 + \delta]$. Suppose the claim is false. Then there exists a point and time $(x_1, t_1) \in \mathcal{M}^n \times (t_0, t_0 + \delta]$ and a nonzero vector $v \in T_{x_1} \mathcal{M}^n$ such that $A_\varepsilon > 0$ for all times $t_0 \leq t < t_1$, but

$$\left(\left(A_\varepsilon \right)_{ij} V^j \right) (x_1, t_1) = 0.$$

Extend v to a vector field V defined in a space-time neighborhood of (x_1, t_1) by the method described above. Then (4.10) and the null-eigenvector assumption imply that at (x_1, t_1), we have

$$\begin{aligned} 0 &\geq \frac{\partial}{\partial t}\left((A_\varepsilon)_{ij} V^i V^j\right) = \left(\frac{\partial}{\partial t} A_\varepsilon\right)_{ij} V^i V^j \\ &> \left[(\Delta A_\varepsilon)_{ij} + \beta_{ij}(A_\varepsilon, g, t)\right] V^i V^j \\ &= \Delta\left((A_\varepsilon)_{ij} V^j V^j\right) + \beta_{ij}(A_\varepsilon, g, t) V^i V^j \geq 0. \end{aligned}$$

This contradiction proves the claim. Then since $\delta > 0$ depends only on

$$\max_{\mathcal{M}^n \times [0,T]} \left|\frac{\partial}{\partial t} g\right|$$

and K, and is in particular independent of ε, we can let $\varepsilon \searrow 0$. Theorem 4.6 follows. $\square$

REMARK 4.7. The proof above corrects a minor oversight in the original argument of Section 9 of [**58**], which failed to consider the $\frac{\partial g}{\partial t}$ term.

3. Advanced weak maximum principles for systems

Theorem 4.6 was proved in [**58**]; it may be regarded as the tensor analogue of Theorem 4.2. There is a tensor analogue of Theorem 4.4 which shows how a tensor evolving by a nonlinear PDE may be controlled by a system of ODE; it was proved in [**59**]. We shall state the result here but postpone our discussion of its proof.

The set-up is as follows. Let $\mathcal{M}^n$ be a closed oriented manifold equipped with a smooth 1-parameter family of metrics $g(t) : t \in [0, T]$ and their Levi-Civita connections $\nabla(t)$. Let $\pi : \mathcal{E} \to \mathcal{M}^n$ be a vector bundle over $\mathcal{M}^n$ with a fixed bundle metric h. Let

$$\bar{\nabla}(t) : C^\infty(\mathcal{E}) \to C^\infty(\mathcal{E} \otimes T^*\mathcal{M}^n)$$

be a smooth family of connections compatible with h in the sense that for all vector fields $X \in T\mathcal{M}^n$, sections $\varphi, \psi \in C^\infty(\mathcal{E})$, and times $t \in [0, T]$ one has

$$X(h(\varphi, \psi)) = h\left(\bar{\nabla}_X \varphi, \psi\right) + h\left(\varphi, \bar{\nabla}_X \psi\right).$$

Define

$$\hat{\nabla}(t) : C^\infty(\mathcal{E} \otimes T^*\mathcal{M}^n) \to C^\infty(\mathcal{E} \otimes T^*\mathcal{M}^n \otimes T^*\mathcal{M}^n)$$

for all $X \in T\mathcal{M}^n$, $\xi \in T^*\mathcal{M}^n$, and $\varphi \in C^\infty(\mathcal{E})$ by

$$\hat{\nabla}_X(\varphi \otimes \xi) \doteqdot \bar{\nabla}_X \varphi \otimes \xi + \varphi \otimes \nabla_X \xi.$$

Then the time-dependent bundle Laplacian $\hat{\Delta}(t)$ is defined for all $\varphi \in C^\infty(\mathcal{E})$ as the metric trace

$$\hat{\Delta}\varphi \doteqdot \operatorname{tr}_g \hat{\nabla}\left(\bar{\nabla}\varphi\right).$$

Let $F : \mathcal{E} \times [0,T] \to \mathcal{E}$ be a continuous map such that $F(\cdot,\cdot,t) : \mathcal{E} \to \mathcal{E}$ is fiber-preserving for each $t \in [0,T]$, and $F(\cdot,x,t) : \mathcal{E}_x \to \mathcal{E}_x$ is Lipschitz for each $x \in \mathcal{M}^n$ and $t \in [0,T]$. Let $\mathcal{K}$ be a closed subset of $\mathcal{E}$. Then we can state the following result from [**59**]. Although it is quite general, we still call it a 'tensor maximum principle' because we shall apply it in the case that $g(t)$ is a solution of the Ricci flow and $\mathcal{E}$ is a tensor bundle.

THEOREM 4.8 (tensor maximum principle, second version: ODE gives pointwise bounds for PDE). *Under the assumptions above, let $\alpha(t) : 0 \le t \le T$ be a solution of the nonlinear* PDE

$$\frac{\partial}{\partial t}\alpha = \hat{\Delta}\alpha + F(\alpha)$$

such that $\alpha(0) \in \mathcal{K}$. Assume further that:

- *$\mathcal{K}$ is invariant under parallel translation by $\bar{\nabla}(t)$ for all $t \in [0,T]$; and*
- *$\mathcal{K}_x \doteqdot \mathcal{K} \cap \pi^{-1}(x)$ is a closed convex subset of $\mathcal{E}_x \doteqdot \pi^{-1}(x)$ for all $x \in \mathcal{M}^n$.*

Then if every solution of the ODE

$$\frac{d}{dt}a = F(a)$$
$$a(0) \in \mathcal{K}_x$$

defined in each fiber $\mathcal{E}_x$ remains in $\mathcal{K}_x$, the solution $\alpha(t)$ of the PDE *remains in $\mathcal{K}$.*

There are two important generalizations of this result which will be useful for us. Both are proved in [**33**] (though some special cases appear in Hamilton's work). In the first, one allows the set $\mathcal{K}$ to depend on time, obtaining a time-dependent maximum principle for systems.

THEOREM 4.9 (tensor maximum principle, third version: ODE controls PDE in time-dependent subsets). *Adopt the assumptions above, allowing $\mathcal{K}(t)$ to be a closed subset of $\mathcal{E}$ for all $t \in [0,T]$. Let $\alpha(t) : 0 \le t \le T$ be a solution of the nonlinear* PDE

$$\frac{\partial}{\partial t}\alpha = \hat{\Delta}\alpha + F(\alpha)$$

such that $\alpha(0) \in \mathcal{K}(0)$. Assume further that:

- *the space-time track $\cup_{t\in[0,T]}(\mathcal{K}(t) \times \{t\})$ is a closed subset of $\mathcal{E} \times [0,T]$;*
- *$\mathcal{K}(t)$ is invariant under parallel translation by $\bar{\nabla}(t)$ for all $t \in [0,T)$; and*
- *$\mathcal{K}_x(t) \doteqdot \mathcal{K}(t) \cap \pi^{-1}(x)$ is a closed convex subset of $\mathcal{E}_x$ for all $x \in \mathcal{M}^n$ and $t \in [0,T]$.*

Then if every solution of the ODE

$$\frac{d}{dt}a = F(a)$$
$$a(0) \in \mathcal{K}_x(0)$$

defined in each fiber $\mathcal{E}_x$ remains in $\mathcal{K}_x(t)$, the solution $\alpha(t)$ of the PDE *remains in $\mathcal{K}(t)$.*

The final generalization we will state is an enhanced version of the result above. It is motivated by applications in which the ODE solution might escape from a certain subset $\mathcal{A} \subset \mathcal{K}$, but where we may be able to *assume* that the solution of the PDE avoids $\mathcal{A}$. Accordingly, we call $\mathcal{A}(t) \subset \mathcal{K}(t)$ the **avoidance set**.

THEOREM 4.10 (tensor maximum principle, fourth version: ODE controls PDE outside avoidance sets). *Adopt the assumptions above, where $\mathcal{K}(t)$ is a closed subset of $\mathcal{E}$ for all $t \in [0,T]$. Let $\alpha(t) : 0 \leq t \leq T$ be a solution of the nonlinear* PDE

$$\frac{\partial}{\partial t}\alpha = \hat{\Delta}\alpha + F(\alpha)$$

such that $\alpha(0) \in \mathcal{K}(0)$. Assume further that:

- *the space-time tracks $\cup_{t\in[0,T]} (\mathcal{K}(t) \times \{t\})$ and $\cup_{t\in[0,T]} (\mathcal{A}(t) \times \{t\})$ are closed subsets of $\mathcal{E} \times [0,T]$;*
- *$\mathcal{K}(t)$ is invariant under parallel translation by $\bar{\nabla}(t)$ for all $t \in [0,T)$;*
- *$\mathcal{K}_x(t) \doteqdot \mathcal{K}(t) \cap \pi^{-1}(x)$ is a closed convex subset of $\mathcal{E}_x$ for all $x \in \mathcal{M}^n$ and $t \in [0,T]$; and*
- *$\alpha(t)$ avoids $\mathcal{A}(t)$, meaning that $\alpha(t) \notin \mathcal{A}(t)$ for any $t \in [0,T]$.*

Then if every solution of the ODE

$$\frac{d}{dt}a = F(a)$$
$$a(0) \in \mathcal{K}_x(0) \setminus \mathcal{A}_x(0)$$

defined in each fiber $\mathcal{E}_x$ either remains in $\mathcal{K}_x(t)$ for all $t \in [0,T]$ or else enters $\mathcal{A}(t_0)$ at some time $t_0 \in (0,T]$, the solution $\alpha(t)$ of the PDE *remains in $\mathcal{K}(t)$.*

In the successor to this volume, we will present detailed proofs of all the results stated in this section.

4. Strong maximum principles

It is well known that any nonnegative solution u of the heat equation

$$u_t = \Delta u$$

has $u > 0$ for all $t > 0$ unless $u \equiv 0$. (See [**110**].) In particular, if one knows at $t = 0$ that $u \geq 0$ everywhere and that $u > 0$ at one point, one can conclude that u is strictly positive for all positive time. This is an example of a *strong maximum principle.* The availability of strong maximum principles for parabolic equations is one of the principal advantages of using evolution equations in geometry. Heuristically, strong maximum principles show that a pointwise bound which holds at some time t can propagate to a global bound at any time $t + \varepsilon$.

Strong maximum principles apply to both scalar and tensor quantities evolving under the Ricci flow. We will study their general formulation in the next volume. But we shall benefit from them in subsequent chapters of this monograph. For example, if one has a solution of the Ricci flow with initially nonnegative scalar curvature, the strong maximum principle for scalar equations implies that $R > 0$ for all times $t > 0$ that the solution exists — unless all $g(t)$ are scalar flat metrics. (See for instance Section 7 of Chapter 5.) The strong maximum principle for tensors is also highly useful. We shall sketch a key example here; rigorous arguments will appear in the chapters which follow. If $\left(\mathcal{M}^3, g(t)\right)$ is a solution of the Ricci flow with nonnegative sectional curvature at $t = 0$, then the eigenvalues $\nu \leq \mu \leq \lambda$ of the curvature operator satisfy exactly one of the following patterns

$$
\begin{aligned}
0 &= \nu = \mu = \lambda \\
0 &= \nu = \mu < \lambda \\
0 &< \nu \leq \mu \leq \lambda
\end{aligned}
$$

for all points and times $t > 0$ such that the solution exists. (The Lie algebra structure of $\mathfrak{so}(3)$ rules out the case $0 = \nu < \mu \leq \lambda$.) In the first case, the manifold is flat; in the second case, it splits as the product of a one-dimensional factor with a positively curved surface; and in the third case, the manifold has strictly positive sectional curvature everywhere. Because (as we shall see in Corollary 9.7) any solution of the Ricci flow constructed as a limit of parabolic dilations about a finite-time singularity in dimension $n = 3$ has nonnegative sectional curvature, this rigidity — a consequence of the strong maximum principle for tensors — is very important for the analysis of singularities.

Notes and commentary

Maximum principles (weak and strong) are among the most important technical tools used to study the Ricci flow. They have numerous applications. For example, the scalar maximum principle implies that the bound $R \geq R_{\min}(0)$ for the scalar curvature is preserved (Lemma 6.8). The tensor maximum principle implies that nonnegative Ricci curvature and hence nonnegative sectional curvature is preserved in dimension $n = 3$ (Corollary 6.11) and that nonnegativity of the curvature operator is preserved in all dimensions (Corollary 6.27).

A nice presentation of the scalar maximum principle for generalized heat equations may be found on pages 101–102 of [**56**]. The maximum principle for tensor equations (Theorem 4.6) first appeared in [**58**]. The tensor maximum principle for systems (Theorem 4.8) appeared in [**59**]. Maximum principles for systems with time-dependent convex sets (Theorem 4.9) and avoidance sets (Theorem 4.10) are studied in [**33**]. Certain maximum principles of these types are used implicitly in some of Hamilton's work.

A reader interested in discussions of maximum principles for more general classes of equations is encouraged to consult [**110**] and [**121**].

CHAPTER 5

The Ricci flow on surfaces

One of the triumphs of nineteenth-century mathematics was the **Uniformization Theorem**, which implies that every smooth surface admits an essentially unique conformal metric of constant curvature. This result may be interpreted as the statement that every 2-dimensional manifold admits a canonical geometry. Looked at another way, it may be regarded as a classification of such manifolds into three families — those of constant positive, zero, or negative curvature. In this sense, the Uniformization Theorem may be regarded as a forerunner, at least in spirit, of current efforts to classify the geometries of closed 3-manifolds.

In this chapter, we study the Ricci flow on closed 2-dimensional manifolds. We shall show that a solution of the normalized Ricci flow exists for all time and converges to a constant-curvature metric conformal to the initial metric. The existence of a constant curvature metric in each conformal class is a classical fact which is equivalent to the Uniformization Theorem. So in this sense, the Ricci flow may be regarded as a natural homotopy between a given Riemannian metric and the canonical metric in its conformal class whose existence is guaranteed by the Uniformization Theorem.

If $(\mathcal{M}^2, g)$ is a compact Riemannian surface (to wit, a 2-dimensional Riemannian manifold) we denote its average scalar curvature by

$$r \doteqdot \left(\int_{\mathcal{M}^2} R\, d\mu\right) \Big/ \left(\int_{\mathcal{M}^2} d\mu\right).$$

By the Gauss-Bonnet theorem, r is determined by the Euler characteristic $\chi(\mathcal{M}^2)$ of the surface, hence is independent of the metric g. The objective of this chapter is to prove the following result.

THEOREM 5.1. *If $(\mathcal{M}^2, g_0)$ is a closed Riemannian surface, there exists a unique solution $g(t)$ of the normalized Ricci flow*

$$\frac{\partial}{\partial t} g = (r - R)\, g$$
$$g(0) = g_0.$$

The solution exists for all time. As $t \to \infty$, the metrics $g(t)$ converge uniformly in any C^k-norm to a smooth metric g_∞ of constant curvature.

One of the objectives of this chapter is to introduce important techniques. Therefore, in the interest of pedagogical clarity, we shall not always attempt to present the most efficient proofs known. The organization of the chapter is as follows. In Sections 1 through 4, we study the evolution of a metric on an arbitrary surface and derive certain *a priori* estimates that apply to all 2-dimensional solutions $(\mathcal{M}^2, g(t))$ of the Ricci flow. Then we establish Theorem 5.1 by considering the three natural cases. In Section 5, we prove exponential convergence to a metric of constant negative curvature in the case that $\chi(\mathcal{M}^2) < 0$. In Section 6, we prove convergence to a flat metric on surfaces satisfying $\chi(\mathcal{M}^2) = 0$. In the remainder of the chapter, we treat the far more difficult case that $\chi(\mathcal{M}^2) > 0$, following the strategy outlined in Section 7.

REMARK 5.2. The Ricci flow still has not provided a complete proof of the Uniformization Theorem if $\chi(\mathcal{M}^2) > 0$. As we shall see in Section 7, the proof of Theorem 5.1 for this case uses the Kazdan–Warner identity to show that the only Ricci solitons on $\mathcal{S}^2$ are the metrics of constant curvature. The proof of the Kazdan–Warner identity requires the Uniformization Theorem. (In [**60**], Hamilton gives another proof that the only solitons on $\mathcal{S}^2$ are trivial. But this proof uses the fact that $\mathcal{S}^2 \backslash \{ \text{ point } \}$ is conformal to $\mathbb{R}^2$, hence also requires Uniformization.)

1. The effect of a conformal change of metric

Let us recall how the curvature, Laplacian, and volume element of a Riemannian manifold $(\mathcal{M}^n, g)$ transform under a conformal change of metric.

To study the curvature, we shall use moving frames. Let $\{e_i\}_{i=1}^n$ be a local orthonormal frame field: a local orthonormal basis for $T\mathcal{M}^n$ in an open set $\mathcal{U} \subset \mathcal{M}^n$. The dual orthonormal basis of $T^*\mathcal{M}^n$ (induced coframe field) is denoted by $\{\omega^i\}_{i=1}^n$ and defined by $\omega^i(e_j) = \delta_j^i$ for all $i, j = 1, \ldots, n$. The metric is then given by

$$g = \sum_{i=1}^n \omega^i \otimes \omega^i.$$

The connection 1-forms $\omega_i^j \in \Omega^1(\mathcal{U})$ are the components of the Levi-Civita connection with respect to $\{e_i\}_{i=1}^n$ and are defined by

$$\nabla_X e_i = \omega_i^j(X)\, e_j,$$

for all $i, j = 1, \ldots, n$ and all vector fields $X \in C^\infty(T\mathcal{M}^n|_{\mathcal{U}})$. They are antisymmetric: $\omega_i^j = -\omega_j^i$. We write the first and second structure equations of Cartan in the form

$$d\omega^i = \omega^j \wedge \omega_j^i$$
$$\mathrm{Rm}_i^j \equiv \Omega_i^j = d\omega_i^j - \omega_i^k \wedge \omega_k^j.$$

(Note the Einstein summation convention.)

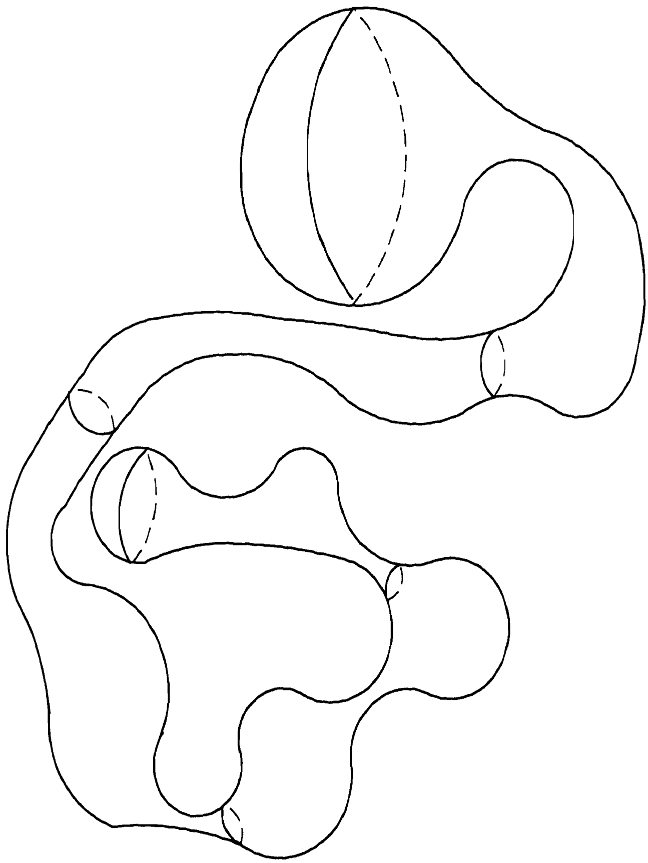

FIGURE 1. A 2-sphere whose curvature is of mixed sign

In the case of a surface $\mathcal{M}^2$, we have $d\omega^1 = \omega^2 \wedge \omega_2^1$, $d\omega^2 = \omega^1 \wedge \omega_1^2$, and

$$\mathrm{Rm}_2^1 = d\omega_2^1.$$

In particular, the Gauss curvature is given by

$$K \doteqdot \langle R(e_1, e_2) e_2, e_1 \rangle = \mathrm{Rm}_2^1(e_1, e_2).$$

LEMMA 5.3. *If g and h are metrics on a surface $\mathcal{M}^2$ conformally related by*

$$g = e^{2u} h,$$

then their scalar curvatures are related by

$$R_g = e^{-2u}(-2\Delta_h u + R_h).$$

PROOF. Let $\{f_1, f_2\}$ be an orthonormal frame field for the metric h. Let $\{\eta^1, \eta^2\}$ be its dual coframe field, and denote the corresponding connection 1-form by η_2^1. Then

$$e_1 \doteqdot e^{-u} f_1 \qquad e_2 \doteqdot e^{-u} f_2$$

is an orthonormal frame field for g and

$$\omega^1 \doteqdot e^u \eta^1 \qquad \omega^2 \doteqdot e^u \eta^2$$

is an orthonormal coframe field for g. To compute the connection 1-form ω_2^1 for $\{e_i\}$ with respect to g, we begin by calculating

$$\begin{aligned} d\omega^1 &= e^u \left(d\eta^1 + du \wedge \eta^1\right) = e^u \left(\eta^2 \wedge \eta_2^1 + f_2(u)\, \eta^2 \wedge \eta^1\right) \\ d\omega^2 &= e^u \left(d\eta^2 + du \wedge \eta^2\right) = e^u \left(\eta^1 \wedge \eta_1^2 + f_1(u)\, \eta^1 \wedge \eta^2\right). \end{aligned}$$

Combining these equations with the general formula

$$\omega_2^1 = d\omega^1(e_2, e_1)\, \omega^1 + d\omega^2(e_2, e_1)\, \omega^2, \tag{5.1}$$

we obtain

$$\begin{aligned} \omega_2^1 &= \left[\eta_2^1(e_1) + e^{-u} f_2(u)\right] \omega^1 - \left[\eta_1^2(e_2) + e^{-u} f_1(u)\right] \omega^2 \\ &= \eta_2^1 + f_2(u)\, \eta^1 - f_1(u)\, \eta^2. \end{aligned}$$

Then by applying this result and the second structure equation, we can write the curvature 2-form of g as

$$\begin{aligned} \operatorname{Rm}[g]_2^1 = d\omega_2^1 &= d\eta_2^1 + d\left[f_2(u)\, \eta^1 - f_1(u)\, \eta^2\right] \\ &= \operatorname{Rm}[h]_2^1 + d\left[f_2(u)\right] \wedge \eta^1 - d\left[f_1(u)\right] \wedge \eta^2 + f_2(u)\, d\eta^1 - f_1(u)\, d\eta^2 \\ &= \operatorname{Rm}[h]_2^1 + f_2 f_2(u)\, \eta^2 \wedge \eta^1 - f_1 f_1(u)\, \eta^1 \wedge \eta^2 \\ &+ f_2(u)\, \eta_2^1(f_1)\, \eta^2 \wedge \eta^1 - f_1(u)\, \eta_1^2(f_2)\, \eta^1 \wedge \eta^2, \end{aligned}$$

where we again used the identity $d\eta^k = \eta_1^2(f_k)\, \eta^1 \wedge \eta^2$ implied by (5.1). Because $\Delta_h u = \nabla^2_{f_1, f_1} u + \nabla^2_{f_2, f_2} u$, this proves

$$\operatorname{Rm}[g]_2^1 = \operatorname{Rm}[h]_2^1 - (\Delta_h u)\, \eta^1 \wedge \eta^2$$

and hence

$$R_g = 2K_g = 2 \operatorname{Rm}[g]_2^1(e_1, e_2) = 2e^{-2u}(K_h - \Delta_h u) = e^{-2u}(R_h - 2\Delta_h u).$$

□

To describe how the Laplacian changes, we begin with a general observation.

LEMMA 5.4. *Let $g(t)$ be a smooth 1-parameter family of metrics on $\mathcal{M}^n$. Then*

$$\frac{\partial}{\partial t} \Delta_{g(t)} = -\left(\frac{\partial}{\partial t} g_{ij}\right) \nabla^i \nabla^j - g^{k\ell}\left(g^{ij} \nabla_i \left(\frac{\partial}{\partial t} g_{j\ell}\right) - \frac{1}{2} \nabla_\ell \left(g^{ij} \frac{\partial}{\partial t} g_{ij}\right)\right) \nabla_k.$$

PROOF. If u is an arbitrary smooth function, we have

$$\begin{aligned} \frac{\partial}{\partial t}(\Delta u) &= \frac{\partial}{\partial t}\left[g^{ij}\left(\partial_i \partial_j - \Gamma_{ij}^k \partial_k\right) u\right] \\ &= \left(\frac{\partial}{\partial t} g^{ij}\right) \nabla_i \nabla_j u - g^{ij}\left(\frac{\partial}{\partial t} \Gamma_{ij}^k\right) \nabla_k u + \Delta\left(\frac{\partial}{\partial t} u\right), \end{aligned}$$

where

$$g^{ij}\frac{\partial}{\partial t}\Gamma_{ij}^{k} = \frac{1}{2}g^{ij}g^{k\ell}\left(\nabla_i\left(\frac{\partial}{\partial t}g_{j\ell}\right) + \nabla_j\left(\frac{\partial}{\partial t}g_{j\ell}\right) - \nabla_\ell\left(\frac{\partial}{\partial t}g_{ij}\right)\right)$$

$$= g^{k\ell}\left(g^{ij}\nabla_i\left(\frac{\partial}{\partial t}g_{j\ell}\right) - \frac{1}{2}\nabla_\ell\left(g^{ij}\frac{\partial}{\partial t}g_{ij}\right)\right).$$

□

This formula simplifies nicely in the special case of a conformal deformation.

COROLLARY 5.5. *If $g(t)$ is a smooth 1-parameter family of metrics on $\mathcal{M}^n$ such that*

$$\frac{\partial}{\partial t}g = fg$$

for a scalar function $f : \mathcal{M}^n \to \mathbb{R}$, then

$$\frac{\partial}{\partial t}\Delta = -f\Delta + \left(\frac{n}{2} - 1\right)\nabla f \cdot \nabla.$$

In particular, if $n = 2$, then

$$\frac{\partial}{\partial t}\Delta = -f\Delta.$$

PROOF. It suffices to calculate

$$g^{k\ell}\left(g^{ij}\nabla_i\left(\frac{\partial}{\partial t}g_{j\ell}\right) - \frac{1}{2}\nabla_\ell\left(g^{ij}\frac{\partial}{\partial t}g_{ij}\right)\right) = \left(1 - \frac{n}{2}\right)\nabla^k f.$$

□

For later use, we also record the following immediate consequence of Lemma 6.5.

COROLLARY 5.6. *If $(\mathcal{M}^n, g(t))$ is a solution of the normalized Ricci flow*

$$\frac{\partial}{\partial t}g = -2\,\mathrm{Rc} + \frac{2r}{n}g,$$

where $r(t)$ denotes the average scalar curvature, then

$$\frac{\partial}{\partial t}d\mu = (r - R)\,d\mu. \tag{5.2}$$

2. Evolution of the curvature

In this section, we compute the PDE for the evolution of the scalar curvature R and study the corresponding ODE obtained formally by ignoring the Laplacian term. This analysis will let us apply the maximum principle to obtain a lower bound for R.

LEMMA 5.7. *If $g(t)$ is a smooth 1-parameter family of metrics on a Riemannian surface $\mathcal{M}^2$ such that $\partial g/\partial t = fg$ for a scalar function f, then*

$$\frac{\partial}{\partial t}R = -\Delta f - Rf.$$

PROOF. Let h be a fixed metric on $\mathcal{M}^2$ and let g be conformally related to h by $g = e^u h$. By Lemma 5.3, the scalar curvatures of g and h are related by

$$R_g = e^{-u}\left(-\Delta_h u + R_h\right).$$

If $\partial g/\partial t = fg$, then $\partial u/\partial t = f$, and differentiating the equation above with respect to time yields

$$\frac{\partial}{\partial t}R_g = -\left(\frac{\partial}{\partial t}u\right)e^{-u}\left(-\Delta_h u + R_h\right) - e^{-u}\Delta_h\left(\frac{\partial}{\partial t}u\right) = -fR_g - \Delta_g f.$$

□

COROLLARY 5.8. *Under the normalized Ricci flow on a surface, we have*

$$\frac{\partial}{\partial t}R = \Delta R + R\left(R - r\right), \tag{5.3}$$

where the average scalar curvature r is constant in time.

This type of evolution is known as a **reaction-diffusion equation**. The Laplacian term promotes diffusion of R, whereas the quadratic reaction terms promote concentration of R. If the right-hand side contained only the Laplacian term, the equation would be the heat equation, albeit with respect to a time-dependent metric; and one would expect R to tend to a constant as $t \to \infty$. On the other hand, if the right-hand side contained only the $R\,(R-r)$ term, the equation would be an ODE; and the solution would blow up in finite time for any initial data satisfying $R(0) > \max\{r, 0\}$. The answer to the question of how the scalar curvature evolves under the normalized Ricci flow depends on which term dominates. We shall later see that it is the diffusion term which determines the qualitative behavior of the equation.

As we discussed in Chapter 4, one may ignoring the Laplacian term in the PDE (5.3) in order to compare its solutions with those of the ODE

$$\frac{d}{dt}s = s\left(s - r\right), \qquad s(0) = s_0, \tag{5.4}$$

where the function $s = s(t)$ plays the role of R. When $r \neq 0$ and $s_0 \neq 0$, the solution of this ODE is

$$s(t) = \frac{r}{1 - \left(1 - \frac{r}{s_0}\right)e^{rt}};$$

when $r = 0$, we have $s(t) = s_0/\left(1 - s_0 t\right)$; and when $s_0 = 0$, the solution is $s(t) \equiv 0$. It follows that whenever $s_0 > \max\{r, 0\}$, there is $T < \infty$ given by

$$T = \begin{cases} -\frac{1}{r}\log\left(1 - r/s_0\right) > 0 & \text{if } r \neq 0 \\ 1/s_0 > 0 & \text{if } r = 0 \end{cases}$$

such that

$$\lim_{t \to T} s(t) = \infty.$$

On the other hand, the ODE behaves much better when $s_0 < \min\{r, 0\}$, in which case we have

$$s(t) - r \geq s_0 - r.$$

The brief analysis above shows that we cannot obtain a good upper bound for the curvature under the normalized Ricci flow simply by applying the maximum principle to equation (5.3). Nevertheless, our formulas for $s(t)$ do allow us to derive useful lower bounds. Set $R_{\min}(t) \doteqdot \inf_{x \in \mathcal{M}^2} R(x,t)$. Then we have the following estimates.

LEMMA 5.9. *Let $g(t)$ be a complete solution with bounded curvature of the normalized Ricci flow on a compact surface.*

- *If $r < 0$, then*

$$R - r \geq \frac{r}{1 - \left(1 - \frac{r}{R_{\min}(0)}\right) e^{rt}} - r \geq \left(R_{\min}(0) - r\right) e^{rt}.$$

- *If $r = 0$, then*

$$R \geq \frac{R_{\min}(0)}{1 - R_{\min}(0)\, t} > -\frac{1}{t}.$$

- *If $r > 0$ and $R_{\min}(0) < 0$, then*

$$R \geq \frac{r}{1 - \left(1 - \frac{r}{R_{\min}(0)}\right) e^{rt}} \geq R_{\min}(0)\, e^{-rt}.$$

Notice that in each case, the right-hand side tends to 0 as $t \to \infty$.

In summary, we have uniform lower bounds for the curvature under the normalized Ricci flow. If the average scalar curvature r is negative, then $R_{\min}(t)$ approaches its average r exponentially fast. If r is positive and $R_{\min}(t)$ ever becomes nonnegative, it remains so for all time. If r is positive but $R_{\min}(t)$ is negative, then $R_{\min}(t)$ approaches zero exponentially fast.

Since the upper bounds for the curvature which one can derive directly from the maximum principle blow up in finite time, we shall use other techniques in the next section to obtain a uniform upper bound for the curvature when $r \leq 0$ and an exponential upper bound when $r > 0$.

3. How Ricci solitons help us estimate the curvature from above

Let $\mathfrak{M}_V(\mathcal{M}^n)$ denote the space of metrics of volume V on a manifold $\mathcal{M}^n$. There is a natural right action of the group $\mathfrak{D}_0(\mathcal{M}^n)$ of volume-preserving diffeomorphisms of $\mathcal{M}^n$ on $\mathfrak{M}_V$ given by $(g, \varphi) \mapsto \varphi^* g$. If $\mathcal{M}^2$ is a surface, then $\mathfrak{M}_A/\mathfrak{D}_0$ is the space of geometric structures of area A in a given conformal class.

Ricci solitons (equivalently, self-similar solutions of the Ricci flow) may be regarded as fixed points of the normalized Ricci flow acting $\mathfrak{M}_A/\mathfrak{D}_0$. These special solutions motivate us to consider certain quantities that may guide us in developing estimates for general solutions. It turns out to be particularly useful to study functions which are constant in space on Ricci solitons. In particular, we shall obtain upper bounds for the scalar curvature by estimating such a quantity. These upper bounds are uniform except when $r > 0$, in which case they are exponential. (Ricci solitons and self-similar solutions were introduced briefly in Section 1 of Chapter 2. In a chapter of the planned successor to this volume, we will make a more systematic study of Ricci solitons in all dimensions.)

DEFINITION 5.10. A solution $g(t)$ of the normalized Ricci flow on a surface $\mathcal{M}^2$ is called a **self-similar solution** of the normalized Ricci flow if there exists a one-parameter family of conformal diffeomorphisms $\varphi(t)$ such that

$$g(t) = \varphi(t)^* g(0).$$

Differentiating this equation with respect to time implies that

$$\frac{\partial}{\partial t} g = \mathcal{L}_X g, \tag{5.5}$$

where $X(t)$ is the one-parameter family of vector fields generated by $\varphi(t)$. By definition of the normalized Ricci flow, equation (5.5) is equivalent to

$$(r - R) g_{ij} = \nabla_i X_j + \nabla_j X_i. \tag{5.6}$$

If $X = -\nabla f$ for some function $f(x,t)$, we obtain

$$(R - r) g_{ij} = 2 \nabla_i \nabla_j f. \tag{5.7}$$

DEFINITION 5.11. We say a metric $g(t)$ on a surface $\mathcal{M}^2$ is a **Ricci soliton** if it satisfies equation (5.6). We say $g(t)$ is a **gradient Ricci soliton** if it satisfies equation (5.7).

Tracing (5.7) yields the equation

$$\Delta f = R - r. \tag{5.8}$$

This equation is solvable even on non-soliton solutions, because

$$\int_{\mathcal{M}^2} (R - r)\, dA = 0.$$

Thus for *any* solution $g(t)$ of the normalized Ricci flow, the solution f of (5.8) is defined to be the **potential of the curvature** . On a compact (closed) surface $\mathcal{M}^2$, the potential is unique up to addition of function $c(t)$ of time alone, because the only harmonic functions there are constants.

Denoting the trace-free part of the Hessian of the potential f of the curvature by

$$M \doteqdot \nabla\nabla f - \frac{1}{2} \Delta f \cdot g, \tag{5.9}$$

we observe that the gradient Ricci soliton equation (5.7) is equivalent to $M \equiv 0$. Recalling that $R_{ik} = \frac{R}{2} g_{ik}$ and taking the divergence of M, we obtain

$$\begin{aligned}(\operatorname{div} M)_i &\doteqdot \nabla^j M_{ji} = \nabla_j \nabla_i \nabla^j f - \frac{1}{2} \nabla_i \nabla_j \nabla^j f \\ &= R_{ik} \nabla^k f + \frac{1}{2} \nabla_i \Delta f = \frac{1}{2} \left(R \nabla_i f + \nabla_i R \right).\end{aligned}$$

Thus when $R > 0$, the gradient Ricci soliton equation (5.7) implies that $\nabla(\log R + f) \equiv 0$, hence that $\log R + f$ is constant in space. More generally, regardless of the sign of R, we have

$$\begin{aligned}0 = 2(\operatorname{div} M)_i &= \nabla_i R + (R - r) \nabla_i f + r \nabla_i f \\ &= \nabla_i R + 2 \nabla_i \nabla^j f \nabla_j f + r \nabla_i f = \nabla_i \left(R + |\nabla f|^2 + r f \right).\end{aligned}$$

Hence on any Ricci gradient soliton, there is a function $C(t)$ of time alone such that

$$F \doteqdot R + |\nabla f|^2 + rf = C.$$

Since F is constant in space on Ricci gradient solitons, we expect quantities related to it to satisfy nice evolution equations. In fact, we may always assume that the potential function f itself satisfies a nice equation.

LEMMA 5.12. *Let $f_0(x,t)$ be a potential of the curvature for a solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow on a compact surface. Then there is a function $c(t)$ of time alone such that the potential function $f \doteqdot f_0 + c$ satisfies the evolution equation*

$$\frac{\partial}{\partial t} f = \Delta f + rf. \tag{5.10}$$

PROOF. By Corollary 5.5, we have $\frac{\partial}{\partial t}\Delta = (R - r)\Delta$. Thus when we differentiate (5.8) with respect to time and use equation (5.3), we obtain

$$(R - r)^2 + \Delta \left(\frac{\partial}{\partial t} f_0 \right) = \frac{\partial}{\partial t} R = \Delta \Delta f_0 + R(R - r),$$

which implies that

$$\Delta \left(\frac{\partial}{\partial t} f_0 \right) = \Delta \left(\Delta f_0 + r f_0 \right).$$

Since the only harmonic functions on a closed manifold are constants, there is a function $\gamma(t)$ of time alone such that

$$\frac{\partial}{\partial t} f_0 = \Delta f_0 + r f_0 + \gamma.$$

The lemma follows by choosing $c(t) \doteqdot -e^{rt} \int_0^t e^{-r\tau} \gamma(\tau)\, d\tau$. □

REMARK 5.13. Our normalization for f differs from that introduced in [**60**]. Compare the result above with Definition 4.1 and the evolution equation derived in Section 4.2 of that paper.

Applying the maximum principle to equation (5.10) yields a useful estimate.

COROLLARY 5.14. *Under the normalized Ricci flow on a compact surface, there exists a constant C such that*

$$|f| \leq Ce^{rt}.$$

A consequence of this when $r \leq 0$ is that the metrics $g(t)$ are uniformly equivalent for as long as a solution exists.

PROPOSITION 5.15. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a compact surface with $r \leq 0$. Then there exists a constant $C \geq 1$ depending only on the initial metric $g(0)$ such that for as long as the solution exists,*

$$\frac{1}{C} g(0) \leq g(t) \leq C\, g(0).$$

PROOF. It follows from (5.8) and (5.10) that

$$\frac{\partial}{\partial t} g = (r - R)\, g = (\Delta f)\, g = \left(\frac{\partial}{\partial t} f - rf\right) g.$$

Integrating this equation with respect to time implies

$$g(x,t) = \exp\left[f(x,t) - f(x,0) - r\int_0^t f(x,\tau)\, d\tau\right] g(x,0).$$

Applying the corollary, we conclude that for any fixed vector V,

$$\left|\log\left(\frac{g(V,V)|_{(x,t)}}{g(V,V)|_{(x,0)}}\right)\right| \leq C\left(e^{rt} + 1\right).$$

□

Now define

$$H \doteqdot R - r + |\nabla f|^2. \tag{5.11}$$

We expect H to satisfy a nice evolution equation, both because f itself does and because $H + rf = F - r$ is constant in space on Ricci gradient solitons.

PROPOSITION 5.16. *On a solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow on a compact surface, the quantity H defined in (5.11) evolves by*

$$\frac{\partial}{\partial t} H = \Delta H - 2|M|^2 + rH, \tag{5.12}$$

where M is the tensor defined in (5.9).

PROOF. Using (5.8), we rewrite equation (5.3) in the form

$$\frac{\partial}{\partial t}(R - r) = \Delta R + R(R - r) = \Delta\,(R - r) + (\Delta f)^2 + r\,(R - r). \tag{5.13}$$

To compute the evolution equation for $|\nabla f|^2$, we use equation (5.10) and recall that $Rg = 2\,\mathrm{Rc}$ on a surface, obtaining

$$\begin{aligned}
\frac{\partial}{\partial t}|\nabla f|^2 &= \frac{\partial}{\partial t}\left(g^{ij}\nabla_i\nabla_j f\right) = \left(\frac{\partial}{\partial t}g^{ij}\right)\nabla_i f\nabla_j f + 2\left(\frac{\partial}{\partial t}\nabla_i f\right)\nabla^i f \\
&= (R-r)|\nabla f|^2 + 2\langle\nabla\Delta f + r\nabla f, \nabla f\rangle \\
&= r|\nabla f|^2 + 2\langle\Delta\nabla f, \nabla f\rangle \\
&= \Delta|\nabla f|^2 - 2|\nabla\nabla f|^2 + r|\nabla f|^2 . \qquad (5.14)
\end{aligned}$$

Combining (5.13) and (5.14) gives the result, because

$$|M|^2 = |\nabla\nabla f|^2 - \frac{1}{2}(\Delta f)^2 .$$

□

Applying the maximum principle to equation (5.12) yields a useful estimate.

COROLLARY 5.17. *On a solution of the normalized Ricci flow on a compact surface, there exists a constant C depending only on the initial metric such that*

$$R - r \le H \le Ce^{rt}.$$

Combining this result with Lemma 5.9 lets us estimate R both from above and below.

PROPOSITION 5.18. *For any solution $\left(\mathcal{M}^2, g(t)\right)$ of the normalized Ricci flow on a compact surface, there exists a constant $C > 0$ depending only on the initial metric such that:*

- *If $r < 0$, then*
$$r - Ce^{rt} \le R \le r + Ce^{rt}.$$
- *If $r = 0$, then*
$$-\frac{C}{1+Ct} \le R \le C.$$
- *If $r > 0$, then*
$$-Ce^{-rt} \le R \le r + Ce^{rt}.$$

In summary, we now have time-dependent upper and lower bounds for the scalar curvature that are valid for as long as a solution exists. As we shall see later in this volume, the long-time existence of solutions is a consequence of these estimates. For now, we shall assume Corollary 7.2, which tells us that long-time existence of solutions follows from appropriate *a priori* bounds on their curvature. To guide the reader's understanding, we summarize the relevant arguments here.

STEP 1. The smoothing properties of the Ricci flow (Theorem 7.1) reveal that bounds on the curvature imply *a priori* bounds on all its derivatives for a short time. We call these **Bernstein–Bando–Shi (BBS) derivative estimates**. They originate in Bernstein's ideas [**16, 17, 18**] for proving gradient bounds via the maximum principle, and were derived for the Ricci flow in [**9**] and [**117, 118**]. (We begin our analysis of BBS estimates in Sections 5, 6, and 9. We discuss their general n-dimensional formulation and provide further references in Chapter 7.)

STEP 2. Short-time existence results (Theorem 3.13 and Corollary 7.7) tell us that the lifetime of a maximal solution $(\mathcal{M}^n, g(t))$ is bounded below by

$$\frac{c}{\max_{\mathcal{M}^n} |\mathrm{Rm}[g_0]|_{g_0}},$$

where c is a universal constant depending only on n. (See also the *doubling-time estimate* of Lemma 5.45.)

STEP 3. Long-time existence results (Theorem 6.45) then imply that the flow cannot be extended past $T < \infty$ only if

$$\lim_{t \nearrow T} \left(\sup_{x \in \mathcal{M}^n} |\mathrm{Rm}(x,t)| \right) = \infty.$$

Because of these facts, the bounds in Proposition 5.18 allow us immediately to apply Corollary 7.2 to get the following long-time existence result for solutions on surfaces.

PROPOSITION 5.19. *If $(\mathcal{M}^2, g_0)$ is a closed Riemannian surface, a unique solution $g(t)$ of the normalized Ricci flow exists for all time and satisfies $g(0) = g_0$.*

In Sections 5 and 6, we will derive bounds for all derivatives of the curvature on surfaces of nonpositive Euler characteristic, thus enabling a direct proof of long-time existence in those cases. In Section 9, we apply the BBS method to bound the gradient of the scalar curvature on a surface of positive Euler characteristic. Finally, in Chapters 6 and 7, we will prove more general results that yield bounds on all higher derivatives of the curvature and establish maximal-time existence (under appropriate hypotheses) on compact manifolds in all dimensions.

4. Uniqueness of Ricci solitons

The main objective of this section is to prove that the only gradient Ricci solitons on surfaces are the metrics of constant curvature. This fact will be important when we consider the long-time behavior of the normalized Ricci flow on a surface of positive Euler characteristic. We start with a result that holds in all dimensions.

4.1. Uniqueness of steady and expanding solitons. If $(\mathcal{M}^n, g)$ is a compact Riemannian manifold, let us denote its volume by

$$V \doteqdot \int_{\mathcal{M}^n} d\mu$$

and its average scalar curvature by

$$\rho \doteqdot \frac{\int_{\mathcal{M}^n} R \, d\mu}{V}.$$

If g is a Ricci soliton, then (2.3) implies that

$$-2\,\mathrm{Rc} = \mathcal{L}_X g + 2\lambda g,$$

whence we get the divergence identity

$$n\lambda + R = \delta X$$

by tracing. Integrating this identity shows that

$$(n\lambda + \rho) V = 0,$$

hence that $\rho = -n\lambda$. It follows that the average scalar curvature of a soliton that is respectively expanding, steady, or shrinking must be respectively negative, zero, or positive. This is not surprising, because applying Lemma 3.9 to the corresponding self-similar solution implies that

$$\frac{d}{dt} \log V(t) = -\rho(t).$$

PROPOSITION 5.20. *Any expanding or steady self-similar solution of the Ricci flow on a compact n-dimensional manifold is Einstein.*

PROOF. We consider the corresponding solution $(\mathcal{M}^n, g(t))$ of the normalized Ricci flow. (See Section 9 of Chapter 6.) Because $g(t)$ flows by diffeomorphisms, there exists a one-parameter family of 1-forms $X(t)$ such that

$$-\frac{1}{2}\frac{\partial}{\partial t} g = \mathrm{Rc} - \frac{\rho}{n} g = \mathcal{L}_{X^\natural} g.$$

Thus by Corollary 6.64, we have

$$-\langle \nabla R, X\rangle = -\mathcal{L}_X R = \frac{\partial}{\partial t} R = \Delta R + 2|\mathrm{Rc}|^2 - \frac{2\rho}{n} R,$$

which implies that the identity

$$\Delta R + \langle \nabla R, X\rangle + 2|\mathrm{Rc}|^2 - \frac{2\rho}{n} R = 0 \tag{5.15}$$

holds identically in time.

Now assume $g(t)$ is an expanding or steady self-similar solution. Then $\rho(t) \leq 0$ as we observed above. Replacing Rc by $\mathrm{Rc} - \frac{\rho}{n} g$, one finds that equation (5.15) is equivalent to

$$\Delta(R-\rho) + \langle \nabla(R-\rho), X\rangle + 2\left|\mathrm{Rc} - \frac{\rho}{n} g\right|^2 + \frac{2\rho}{n}(R-\rho) = 0. \tag{5.16}$$

By (5.16), at any point $x \in \mathcal{M}^n$ such that $R(x,t) = R_{\min}(t)$, we have

$$2\left|\mathrm{Rc} - \frac{\rho}{n} g\right|^2 + \frac{2\rho}{n}(R-\rho) \leq 0.$$

Since $R(x,t) \le \rho(t) \le 0$, this is possible only if

$$\left|\mathrm{Rc} - \frac{\rho}{n} g\right|^2 = 0$$

at (x,t). Tracing then implies that

$$R_{\min}(t) = R(x,t) = \rho(t),$$

hence that $R \equiv r$ is constant in space. Substituting back into (5.16), we conclude that

$$\mathrm{Rc}(\cdot,t) \equiv \frac{\rho(t)}{n} g(\cdot,t)$$

holds identically, hence that $g(t)$ is Einstein. $\square$

4.2. Surface solitons and the Kazdan–Warner identity. Now we prove the main assertion of this section. Since the argument uses a particular formulation of the Kazdan–Warner identity [**82**], we follow it with a discussion of that result.

PROPOSITION 5.21. *If $(\mathcal{M}^2, g(t))$ is a self-similar solution of the normalized Ricci flow on a Riemannian surface, then $g(t) \equiv g(0)$ is a metric of constant curvature.*

PROOF. By Proposition 5.20, we may assume that $r > 0$. By passing to a cover space if necessary, we may thus assume that $\mathcal{M}^2$ is diffeomorphic to $\mathcal{S}^2$. Contracting the Ricci soliton equation (5.6) by Rg^{-1} yields

$$2R(r-R) = 2R \operatorname{div} X,$$

and hence

$$-\int_{\mathcal{S}^2} (R-r)^2 \, dA = \int_{\mathcal{S}^2} R(r-R) \, dA = \int_{\mathcal{S}^2} R \operatorname{div} X \, dA.$$

Since X is a conformal Killing vector field, integrating by parts and applying the Kazdan–Warner identity (5.18) implies that

$$\int_{\mathcal{S}^2} (R-r)^2 \, dA = \int_{\mathcal{S}^2} \langle \nabla R, X \rangle \, dA = 0.$$

Hence $R \equiv r$, and the proposition is proved. $\square$

We begin our discussion of the Kazdan–Warner identity by stating its usual formulation. (See for example page 195 of [**113**].) Let g be an arbitrary metric on a topological $\mathcal{S}^2$. We shall denote the standard round metric on $\mathcal{S}^2$ by $\bar{g}$ and will use bars to indicate geometric quantities associated to $\bar{g}$. Let φ be a first eigenfunction of the rough Laplacian $\bar{\Delta}$ on $\mathcal{S}^2$. (In other words, φ is a spherical harmonic.) By the Uniformization Theorem, one may write $g = e^{2u}\bar{g}$. Then one can write the standard Kazdan–Warner identity in the form

$$\int_{\mathcal{S}^2} \bar{\nabla}\varphi\,(K)\, e^{2u} \, d\bar{A} = 0, \tag{5.17}$$

where $K = R/2$ denotes the Gauss curvature of g.

We now derive the form of (5.17) that we used in proving Proposition 5.21. The real vector space of conformal Killing vector fields of $\left(\mathcal{S}^2, \bar{g}\right)$ is 6-dimensional. Three of the dimensions arise from vector fields of the form $\bar{\nabla}\varphi$, where φ is a spherical harmonic. The other three dimensions come from the Killing vector fields for $\bar{g}$. We claim that if Y is a Killing vector field for $\bar{g}$, then

$$\int_{\mathcal{S}^2} Y(K)\, e^{2u}\, d\bar{A} = 0.$$

By the dimension count $6 = 3 + 3$, the claim implies that for any conformal vector field X, one has

$$\int_{\mathcal{S}^2} \langle X, \nabla K\rangle\, dA = \int_{\mathcal{S}^2} X(K)\, dA = 0, \tag{5.18}$$

since $dA = e^{2u}\, d\bar{A}$. Note that conformality is well defined, because X is conformal with respect to g if and only if X is conformal with respect to $\bar{g}$.

The claim is derived from an integration by parts. Since $\bar{K} = 1$, Lemma 5.3 implies that

$$K = e^{-2u}\left(1 - \bar{\Delta} u\right).$$

Recalling that $\overline{\operatorname{div}} Y = 0$ whenever Y is a Killing vector field for $\bar{g}$, we compute that

$$\begin{aligned}\int_{\mathcal{S}^2} Y(K)\, e^{2u}\, d\bar{A} &= \int_{\mathcal{S}^2} \langle Y, \bar{\nabla} K_g\rangle_{\bar{g}}\, e^{2u}\, d\bar{A} \\ &= -2\int_{\mathcal{S}^2} K\left\langle Y, \bar{\nabla} u\right\rangle_{\bar{g}} e^{2u}\, d\bar{A}.\end{aligned}$$

Now

$$\begin{aligned}\int_{\mathcal{S}^2} K\left\langle Y, \bar{\nabla} u\right\rangle_{\bar{g}} e^{2u}\, d\bar{A} &= \int_{\mathcal{S}^2} \left(1 - \bar{\Delta} u\right)\left\langle Y, \bar{\nabla} u\right\rangle_{\bar{g}} d\bar{A} \\ &= -\int_{\mathcal{S}^2} \left(\bar{\Delta} u\right)\left\langle Y, \bar{\nabla} u\right\rangle_{\bar{g}} d\bar{A},\end{aligned}$$

because

$$\int_{\mathcal{S}^2} \left\langle Y, \bar{\nabla} u\right\rangle_{\bar{g}} d\bar{A} = -\int_{\mathcal{S}^2} u \cdot \overline{\operatorname{div}} Y\, d\bar{A} = 0.$$

Integrating by parts again and using the fact that $\bar{\nabla} Y$ is antisymmetric, we obtain

$$\begin{aligned}\int_{\mathcal{S}^2} \left(\bar{\Delta} u\right)\left\langle Y, \bar{\nabla} u\right\rangle_{\bar{g}} d\bar{A} &= -\int_{\mathcal{S}^2} \bar{\nabla}_i u \bar{\nabla}^i \bar{\nabla}_j u Y^j\, d\bar{A} - \int_{\mathcal{S}^2} \bar{\nabla}_i u \bar{\nabla}_j u \bar{\nabla}^i Y^j\, d\bar{A} \\ &= -\frac{1}{2}\int_{\mathcal{S}^2} \left\langle Y, \bar{\nabla}\left|\bar{\nabla} u\right|_{\bar{g}}^2\right\rangle_{\bar{g}} d\bar{A} \\ &= \frac{1}{2}\int_{\mathcal{S}^2} \left|\bar{\nabla} u\right|_{\bar{g}}^2 \left(\overline{\operatorname{div}} Y\right) d\bar{A} = 0.\end{aligned}$$

5. Convergence when $\chi(\mathcal{M}^2) < 0$

In this section, we will prove the following case of Theorem 5.1.

THEOREM 5.22. *Let $(\mathcal{M}^2, g_0)$ be a closed Riemannian surface with average scalar curvature $r < 0$. Then the unique solution $g(t)$ of the normalized Ricci flow with $g(0) = g_0$ converges exponentially in any C^k-norm to a smooth constant-curvature metric g_∞ as $t \to \infty$.*

By Proposition 5.19, the solution $g(t)$ exists for $0 < t < \infty$. By Proposition 5.15, the metrics $g(t)$ are all uniformly equivalent. And by Proposition 5.18, there exists a constant $C > 0$ depending only on g_0 such that R is exponentially approaching its average in the sense that

$$|R - r| \le Ce^{rt}. \tag{5.19}$$

So to prove the theorem, it will suffice to show that all derivatives of R are dying exponentially.

LEMMA 5.23. *On any solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow, $|\nabla R|^2$ evolves by*

$$\frac{\partial}{\partial t}|\nabla R|^2 = \Delta|\nabla R|^2 - 2|\nabla\nabla R|^2 + (4R - 3r)|\nabla R|^2. \tag{5.20}$$

PROOF. Using (5.3) and the Ricci identity for a surface

$$\nabla\Delta = \Delta\nabla - \frac{1}{2}R\nabla,$$

we get

$$\frac{\partial}{\partial t}(\nabla R) = \nabla(\Delta R + R(R - r)) = \Delta\nabla R + \frac{3}{2}R\nabla R - r\nabla R.$$

Hence

$$\begin{aligned}\frac{\partial}{\partial t}|\nabla R|^2 &= \frac{\partial}{\partial t}\left(g^{ij}\nabla_i R\nabla_j R\right)\\ &= (R - r)|\nabla R|^2 + 2\left\langle \Delta\nabla R + \frac{3}{2}R\nabla R - r\nabla R, \nabla R\right\rangle.\end{aligned}$$

Now the lemma follows from the fact that

$$\Delta|\nabla R|^2 = 2\langle\Delta\nabla R, \nabla R\rangle + 2|\nabla\nabla R|^2.$$

□

COROLLARY 5.24. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow such that $r < 0$, then there exists $C_1 > 0$ such that*

$$|\nabla R|^2 \le C_1 e^{rt/2}. \tag{5.21}$$

PROOF. By (5.19), one has

$$\frac{\partial}{\partial t}|\nabla R|^2 \le \Delta|\nabla R|^2 - 2|\nabla\nabla R|^2 + \left(r + 4Ce^{rt}\right)|\nabla R|^2.$$

So for all $t > 0$ large enough, one gets

$$\frac{\partial}{\partial t}|\nabla R|^2 \leq \Delta|\nabla R|^2 + \frac{r}{2}|\nabla R|^2,$$

whence the result follows from the maximum principle. □

Before treating the general case, we provide another example, in order to illustrate the role played by the evolution of the Levi-Civita connection.

LEMMA 5.25. *On any solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow, $|\nabla\nabla R|^2$ evolves by*

$$\begin{aligned}\frac{\partial}{\partial t}|\nabla\nabla R|^2 &= \Delta|\nabla\nabla R|^2 - 2|\nabla\nabla\nabla R|^2 + (2R - 4r)|\nabla\nabla R|^2 \\ &\quad + 2R(\Delta R)^2 + 2\left\langle \nabla R, \nabla|\nabla R|^2\right\rangle.\end{aligned}$$

PROOF. In dimension $n = 2$, the standard variational formula (3.3) becomes

$$\begin{aligned}\frac{\partial}{\partial t}\Gamma_{ij}^k &= \frac{1}{2}g^{k\ell}\left[\nabla_i\left(\frac{\partial}{\partial t}g_{jl}\right) + \nabla_j\left(\frac{\partial}{\partial t}g_{i\ell}\right) - \nabla_\ell\left(\frac{\partial}{\partial t}g_{ij}\right)\right] \\ &= -\frac{1}{2}\left(\nabla_i R\delta_j^k + \nabla_j R\delta_i^k - \nabla^k R g_{ij}\right).\end{aligned}$$

We find by commuting derivatives that

$$\nabla_i\nabla_j\Delta R = \Delta\nabla_i\nabla_j R - 2R\nabla_i\nabla_j R + \left(R\Delta R + \frac{1}{2}|\nabla R|^2\right)g_{ij} - \nabla_i R\nabla_j R.$$

Thus we get

$$\begin{aligned}\frac{\partial}{\partial t}\nabla_i\nabla_j R &= \nabla_i\nabla_j\left(\frac{\partial}{\partial t}R\right) - \left(\frac{\partial}{\partial t}\Gamma_{ij}^k\right)\nabla_k R \\ &= \Delta\nabla_i\nabla_j R + R\Delta R g_{ij} + 2\nabla_i R\nabla_j R - r\nabla_i\nabla_j R\end{aligned}$$

and hence

$$\begin{aligned}\frac{\partial}{\partial t}|\nabla\nabla R|^2 &= \frac{\partial}{\partial t}\left(g^{ij}g^{k\ell}\nabla_i\nabla_j R\nabla_k\nabla_\ell R\right) \\ &= \Delta|\nabla\nabla R|^2 - 2|\nabla\nabla\nabla R|^2 + (2R - 4r)|\nabla\nabla R|^2 \\ &\quad + 2R(\Delta R)^2 + 2\left\langle \nabla R, \nabla|\nabla R|^2\right\rangle.\end{aligned}$$

□

COROLLARY 5.26. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow such that $r < 0$, then there exists $C_2 > 0$ such that*

$$|\nabla\nabla R|^2 \leq C_2 e^{rt/2}.$$

PROOF. By (5.19), there exists $t_0 \geq 0$ such that $R \leq 0$ for all $t \geq t_0$. Applying (5.21) at such times, we find that there exists $C_1' > 0$ such that

$$\frac{\partial}{\partial t}|\nabla\nabla R|^2 \leq \Delta|\nabla\nabla R|^2 - 4r|\nabla\nabla R|^2 + C_1' e^{rt/2}|\nabla\nabla R|.$$

Define $\varphi \geq |\nabla\nabla R|^2$ by

$$\varphi \doteqdot |\nabla\nabla R|^2 - 3r\,|\nabla R|^2 .$$

Then there exists $t_1 \geq t_0$ large enough such that for all $t \geq t_1$, one has $4R - 3r \leq \frac{3}{4}r$ and hence

$$\begin{aligned}\frac{\partial}{\partial t}\varphi &\leq \Delta\varphi + 2r\,|\nabla\nabla R|^2 + (4R-3r)\left(-3r\,|\nabla R|^2\right) + C_1' e^{rt/2}\,|\nabla\nabla R| \\ &\leq \Delta\varphi + \frac{3}{4}r\varphi + C_1' e^{rt/2}\,|\nabla\nabla R| \\ &\leq \Delta\varphi + \frac{3}{4}r\varphi + C_1' e^{rt/2}\sqrt{\varphi} \\ &\leq \Delta\varphi + \frac{2}{3}r\varphi + C_2' e^{rt},\end{aligned}$$

where $C_2' \geq 3\left(C_1'\right)^2/|r|$. By the maximum principle, there exists $C_2'' > 0$ such that $\varphi \leq C_2'' e^{rt/2}$, which implies the result. $\square$

We are now ready for the general case.

PROPOSITION 5.27. *Let $\left(\mathcal{M}^2, g(t)\right)$ be a solution of the normalized Ricci flow on a closed surface with $r < 0$. Then for each positive integer k, there exists a constant $C_k < \infty$ such that for all $t \in [0,\infty)$,*

$$\sup_{x\in\mathcal{M}^2} \left|\nabla^k R(x,t)\right|^2 \leq C_k e^{rt/2}.$$

PROOF. The proof is by complete induction on k; we may suppose the result is known for $0 \leq j \leq k-1$. There is a commutator of the sort

$$\nabla^k \Delta R - \Delta\nabla^k R = \sum_{j=0}^{\lfloor k/2\rfloor} \left(\nabla^j R\right) \otimes_g \left(\nabla^{k-j} R\right),$$

where $\lfloor\cdot\rfloor$ is the greatest-integer function, and $X \otimes_g Y$ denotes a finite linear combination of contractions of the tensors X and Y taken with respect to the metric $g(t)$. Similarly, there are formulas

$$\nabla^k\left(R^2\right) = \sum_{j=0}^{\lfloor k/2\rfloor} \left(\nabla^j R\right) \otimes_g \left(\nabla^{k-j} R\right)$$

and

$$\left(\frac{\partial}{\partial t}\Gamma\right) \otimes_g \left(\nabla^j R\right) = (\nabla R) \otimes_g \left(\nabla^j R\right).$$

Thus by recursive application of the identity

$$\frac{\partial}{\partial t}\left(\nabla_i \nabla_j \cdots \nabla_q R\right) = \nabla_i\left(\frac{\partial}{\partial t}\nabla_j \cdots \nabla_q R\right) - \left(\frac{\partial}{\partial t}\Gamma_{ij}^p\right)\nabla_p \cdots \nabla_q R,$$

we obtain

$$\frac{\partial}{\partial t}\left(\nabla^k R\right) = \Delta \nabla^k R + \sum_{j=0}^{\lfloor k/2 \rfloor} \left(\nabla^j R\right) \otimes_g \left(\nabla^{k-j} R\right) - r\left(\nabla^k R\right)$$

and hence

$$\begin{aligned}
\frac{\partial}{\partial t}\left|\nabla^k R\right|^2 &= \frac{\partial}{\partial t}\left(g^{ip}\cdots g^{jq}\,\nabla_i \cdots \nabla_p R\,\nabla_j \cdots \nabla_q R\right) \\
&= \Delta\left|\nabla^k R\right|^2 - 2\left|\nabla^{k+1} R\right|^2 - (k+2)\,r\left|\nabla^k R\right|^2 \\
&\quad + \left(\nabla^k R\right) \otimes_g \left[\textstyle\sum_{j=0}^{\lfloor k/2 \rfloor} \left(\nabla^j R\right) \otimes_g \left(\nabla^{k-j} R\right)\right].
\end{aligned}$$

By induction, there exists a time $t_0 \geq 0$ and a constant C such that for $t \geq t_0$ one has

$$\frac{\partial}{\partial t}\left|\nabla^k R\right|^2 \leq \Delta\left|\nabla^k R\right|^2 - (k+2)\,r\left|\nabla^k R\right|^2 + Ce^{rt/2}\left(1 + \left|\nabla^k R\right|\right).$$

The induction hypothesis also implies that

$$\begin{aligned}
\frac{\partial}{\partial t}\left|\nabla^{k-1} R\right|^2 &= \Delta\left|\nabla^{k-1} R\right|^2 - 2\left|\nabla^k R\right|^2 - (k+1)\,r\left|\nabla^{k-1} R\right|^2 \\
&\quad + \left(\nabla^{k-1} R\right) \otimes_g \left[\textstyle\sum_{j=0}^{\lfloor (k-1)/2 \rfloor} \left(\nabla^j R\right) \otimes_g \left(\nabla^{k-1-j} R\right)\right] \\
&\leq \Delta\left|\nabla^{k-1} R\right|^2 - 2\left|\nabla^k R\right|^2 + Ce^{rt/2}.
\end{aligned}$$

Define $\Phi \geq \left|\nabla^k R\right|^2$ by

$$\Phi \doteqdot \left|\nabla^k R\right|^2 - (k+1)\,r\left|\nabla^{k-1} R\right|^2.$$

Then there exist C' and C'' such that

$$\begin{aligned}
\frac{\partial}{\partial t}\Phi &\leq \Delta\Phi + kr\left|\nabla^k R\right|^2 + C'e^{rt/2}\left(1 + \left|\nabla^k R\right|\right) \\
&\leq \Delta\Phi + \frac{kr}{2}\Phi + C''e^{rt}.
\end{aligned}$$

As above, the conclusion now follows readily from the maximum principle. □

Our proof of Theorem 5.22 is now complete.

6. Convergence when $\chi\left(\mathcal{M}^2\right) = 0$

In this section, we will prove the following case of Theorem 5.1.

THEOREM 5.28. *Let $\left(\mathcal{M}^2, g_0\right)$ be a closed Riemannian surface with average scalar curvature $r = 0$. Then the unique solution $g(t)$ of the normalized Ricci flow with $g(0) = g_0$ converges uniformly in any C^k-norm to a smooth constant-curvature metric g_∞ as $t \to \infty$.*

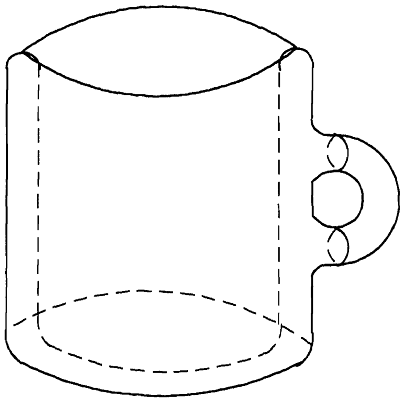

FIGURE 2. A surface of Euler characteristic zero

By Proposition 5.19, the solution $g(t)$ exists for $0 < t < \infty$. By Proposition 5.15, the metrics $g(t)$ are all uniformly equivalent. To prove the theorem, we shall show that the scalar curvature and all its derivatives vanish in the limit as $t \to \infty$.

We begin by deriving an estimate for the decay of the scalar curvature. Recall from Section 3 that the potential of the curvature f satisfies

$$\Delta f = R$$

and is normalized by a function of time alone so that

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

LEMMA 5.29. *Let $(\mathcal{M}^2, g(t))$ be a solution of the Ricci flow on a closed surface with $r = 0$. Then there exists $C < \infty$ depending only on $g(0)$ such that for all $t \in [0, \infty)$, the potential of the curvature satisfies*

$$\sup_{x \in \mathcal{M}^2} |\nabla f(x,t)|^2 \leq \frac{C}{1+t}.$$

PROOF. It follows from Proposition 5.16 that

$$\frac{\partial}{\partial t} |\nabla f|^2 = \Delta |\nabla f|^2 - 2 |\nabla \nabla f|^2 .$$

By the maximum principle, there is $C_0 = C_0(g(0))$ such that $|\nabla f|^2 \leq C_0$ for all time. To improve this estimate, we apply another BBS technique. (The method of obtaining gradient bounds via maximum-principle arguments will be discussed further in Chapter 7. Please see the references cited therein.) Noticing that

$$\frac{\partial}{\partial t} \left(t |\nabla f|^2 \right) \leq \Delta \left(t |\nabla f|^2 \right) + |\nabla f|^2$$

and

$$\frac{\partial}{\partial t} f^2 = \Delta (f^2) - 2 |\nabla f|^2 ,$$

we obtain the inequality

$$\frac{\partial}{\partial t}\left(t|\nabla f|^2 + f^2\right) \le \Delta\left(t|\nabla f|^2 + f^2\right).$$

Hence there is $C_1 = C_1(g(0))$ such that $t|\nabla f|^2 + f^2 \le C_1$; in particular, we have $|\nabla f|^2 \le C_1/t$. The lemma follows from combining these estimates. □

PROPOSITION 5.30. *Let $(\mathcal{M}^2, g(t))$ be a solution of the Ricci flow on a closed surface with $r = 0$. Then there exists $C < \infty$ depending only on $g(0)$ such that for all $t \in [0,\infty)$,*

$$\sup_{x\in\mathcal{M}^2}\left(|R(x,t)| + |\nabla f(x,t)|^2\right) \le \frac{C}{1+t}.$$

The method of proof is a hybrid of BBS techniques and certain innovations of Hamilton. (See for instance Propositions 5.16 and 5.50.)

PROOF. Recalling that $R^2 = (\Delta f)^2 \le 2|\nabla\nabla f|^2$, we compute

$$\begin{aligned}\frac{\partial}{\partial t}\left(R + 2|\nabla f|^2\right) &= \Delta\left(R + 2|\nabla f|^2\right) + R^2 - 4|\nabla\nabla f|^2\\ &\le \Delta\left(R + 2|\nabla f|^2\right) - R^2.\end{aligned}$$

The motivation to consider this quantity is the favorable $-R^2$ term on the right-hand side. Not only do we have $R + 2|\nabla f|^2 \le C_0(g(0))$ by the maximum principle, but we can also apply the BBS method to obtain

$$\begin{aligned}\frac{\partial}{\partial t}\left[t\left(R + 2|\nabla f|^2\right)\right] &\le \Delta\left[t\left(R + 2|\nabla f|^2\right)\right] - tR^2 + R + 2|\nabla f|^2\\ &\le \Delta\left[t\left(R + 2|\nabla f|^2\right)\right] - \frac{t}{2}R^2 + R + 2|\nabla f|^2\\ &\quad - \frac{t}{2}\left(R + 2|\nabla f|^2\right)^2 + 2t|\nabla f|^2\left(R + |\nabla f|^2\right).\end{aligned}$$

Since $t|\nabla f|^2 \le C_1$ by Lemma 5.29, there is $C' < \infty$ so large that at any point and time where $t\left(R + 2|\nabla f|^2\right) \ge C'$, we have $R \ge 0$ and hence

$$\begin{aligned}\frac{\partial}{\partial t}\left[t\left(R + 2|\nabla f|^2\right)\right] &\le \Delta\left[t\left(R + 2|\nabla f|^2\right)\right] - \frac{t}{2}R^2 - \frac{t}{2}\left(R + 2|\nabla f|^2\right)^2\\ &\quad + (1 + 2C_1)R + 2(1 + C_1)|\nabla f|^2\\ &\le \Delta\left[t\left(R + 2|\nabla f|^2\right)\right] - \frac{1}{2t}\left[t\left(R + 2|\nabla f|^2\right)\right]^2\\ &\quad - \left[\sqrt{\frac{t}{2}}R - \sqrt{\frac{1}{2t}}(1 + 2C_1)\right]^2 + \frac{C_2}{t}.\end{aligned}$$

Thus there is $C \ge C'$ large enough that $t\left(R + 2|\nabla f|^2\right) \ge C$ implies

$$\frac{\partial}{\partial t}\left[t\left(R + 2|\nabla f|^2\right)\right] \le \Delta\left[t\left(R + 2|\nabla f|^2\right)\right].$$

By the maximum principle, $R + 2|\nabla f|^2 \leq C/t$ for all positive times. Because Proposition 5.18 and Lemma 5.29 uniformly bound $R + |\nabla f|^2$, the proposition follows easily. □

We now proceed to derive suitable bounds on the derivatives of the curvature. For pedagogical reasons, our method will again emphasize clarity of exposition rather than efficiency. The reader interested only in seeing the general case may proceed directly to Proposition 5.33.

LEMMA 5.31. *Let $(\mathcal{M}^2, g(t))$ be a solution of the Ricci flow on a closed surface with $r = 0$. Then there exists $C < \infty$ depending only on $g(0)$ such that for all $t \in [0, \infty)$,*

$$\sup_{x \in \mathcal{M}^2} |\nabla R(x,t)|^2 \leq \frac{C}{(1+t)^3}.$$

PROOF. When $r = 0$, equation (5.20) takes the form

$$\frac{\partial}{\partial t} |\nabla R|^2 = \Delta |\nabla R|^2 - 2|\nabla\nabla R|^2 + 4R|\nabla R|^2 \leq \Delta |\nabla R|^2 + 4R|\nabla R|^2.$$

Let α be a constant to be chosen later, and consider the quantity

$$\varphi \doteqdot t^4 |\nabla R|^2 + \alpha t^3 R^2,$$

which satisfies $\varphi(\cdot, 0) \equiv 0$. Observing that

$$\frac{\partial}{\partial t}\left(t^4 |\nabla R|^2\right) \leq \Delta\left(t^4 |\nabla R|^2\right) + 4t^3 (tR + 1)|\nabla R|^2$$

and

$$\frac{\partial}{\partial t}\left(t^3 R^2\right) = \Delta\left(t^3 R^2\right) - 2t^3 |\nabla R|^2 + 2t^3 R^3 + 3t^2 R^2,$$

we compute

$$\frac{\partial}{\partial t}\varphi \leq \Delta\varphi + 4t^3 (tR + 1 - \alpha/2)|\nabla R|^2 + \alpha\left(2t^3 R^3 + 3t^2 R^2\right).$$

By Proposition 5.30, one may choose $\alpha < \infty$ such that $tR + 1 \leq \alpha/2$. Therefore

$$\frac{\partial}{\partial t}\varphi \leq \Delta\varphi + C,$$

where $C = \alpha\left(\frac{1}{4}\alpha^3 + \frac{3}{4}\alpha^2\right)$. Hence by the maximum principle,

$$t^4 |\nabla R|^2 + \alpha t^3 R^2 \doteqdot \varphi \leq Ct.$$

□

This method can clearly be generalized to higher derivatives of the curvature. To motivate the proof for the general case, we will next estimate $|\nabla\nabla R|$.

LEMMA 5.32. *Let $(\mathcal{M}^2, g(t))$ be a solution of the Ricci flow on a closed surface with $r = 0$. Then there exists $C < \infty$ depending only on $g(0)$ such that for all $t \in [0, \infty)$,*

$$\sup_{x \in \mathcal{M}^2} |\nabla\nabla R(x,t)|^2 \leq \frac{C}{(1+t)^4}.$$

PROOF. Recall from Lemma 5.25 that

$$\frac{\partial}{\partial t} |\nabla\nabla R|^2 = \Delta |\nabla\nabla R|^2 - 2 |\nabla\nabla\nabla R|^2 + 2R |\nabla\nabla R|^2 + 2R(\Delta R)^2 + 2\left\langle \nabla R, \nabla |\nabla R|^2 \right\rangle.$$

Let $\beta > 0$ be a constant to be chosen later, and define

$$\psi \doteqdot t^5 |\nabla\nabla R|^2 + \beta t^4 |\nabla R|^2,$$

noticing that $\psi(\cdot, 0) \equiv 0$. Using the facts that $(\Delta R)^2 \leq 2 |\nabla\nabla R|^2$ and

$$\left\langle \nabla R, \nabla |\nabla R|^2 \right\rangle = 2 (\nabla\nabla R)(\nabla R, \nabla R),$$

it is easy to estimate

$$\frac{\partial}{\partial t}\psi \leq \Delta\psi + (6tR + 5 - 2\beta) t^4 |\nabla\nabla R|^2 + 4\left(t^3 |\nabla R|^2\right) \sqrt{t^4 |\nabla\nabla R|^2} + 4\beta (tR + 1)\left(t^3 |\nabla R|^2\right).$$

By Proposition 5.30, there is $C_1 < \infty$ such that $tR \leq C_1$. By Lemma 5.31, there is $C_2 < \infty$ such that $t^3 |\nabla R|^2 \leq C_2$. So one can choose $\beta = \beta(C_1)$ large enough and then $C = C(C_1, C_2)$ such that

$$\frac{\partial}{\partial t}\psi \leq \Delta\psi + C.$$

Hence we have $\psi \leq Ct$ by the maximum principle. □

We now treat the general case.

PROPOSITION 5.33. *Let $(\mathcal{M}^2, g(t))$ be a solution of the Ricci flow on a closed surface with $r = 0$. Then for each positive integer k, there exists a constant $C_k < \infty$ depending only on $g(0)$ such that for all $t \in [0, \infty)$,*

$$\sup_{x \in \mathcal{M}^2} \left|\nabla^k R(x,t)\right|^2 \leq \frac{C_k}{(1+t)^{k+2}}.$$

PROOF. The proof is by complete induction on k; we may suppose the result is known for $0 \leq j \leq k-1$. As in Proposition 5.27, we compute that

$$\begin{aligned}\frac{\partial}{\partial t}\left|\nabla^k R\right|^2 &= \frac{\partial}{\partial t}\left(g^{ip} \cdots g^{jq}\, \nabla_i \cdots \nabla_p R\, \nabla_j \cdots \nabla_q R\right) \\ &= \Delta\left|\nabla^k R\right|^2 - 2\left|\nabla^{k+1} R\right|^2 \\ &\quad + \left(\nabla^k R\right) \otimes_g \left[\textstyle\sum_{j=0}^{\lfloor k/2 \rfloor} \left(\nabla^j R\right) \otimes_g \left(\nabla^{k-j} R\right)\right].\end{aligned}$$

Let N be a constant to be determined, and set

$$\Phi \doteqdot t^{k+3}\left|\nabla^k R\right|^2 + N t^{k+2}\left|\nabla^{k-1} R\right|^2.$$

Then there are constants a, b_j, and c_j such that we can estimate

$$\begin{aligned}\frac{\partial}{\partial t}\Phi &\leq \Delta\Phi + t^{k+2}\left|\nabla^k R\right|^2\left[atR + (k+3-2N)\right] \\ &+ \sqrt{t^{k+2}\left|\nabla^k R\right|^2}\cdot\sum_{j=1}^{\lfloor k/2\rfloor} b_j\sqrt{t^{j+2}\left|\nabla^j R\right|^2\cdot t^{k-j+2}\left|\nabla^{k-j}R\right|^2} \\ &+ N\sum_{j=0}^{\lfloor (k-1)/2\rfloor} c_j\sqrt{t^{k+1}\left|\nabla^{k-1}R\right|^2\cdot t^{j+2}\left|\nabla^j R\right|^2\cdot t^{k-j+1}\left|\nabla^{k-j-1}R\right|^2} \\ &+ N(k+2)\,t^{k+1}\left|\nabla^{k-1}R\right|^2.\end{aligned}$$

By the inductive hypothesis, there is a sufficiently large constant N depending only on $g(0)$, and positive constants A, B, C, D depending on $g(0)$ and N such that

$$\frac{\partial}{\partial t}\Phi \leq \Delta\Phi - At^{k+2}\left|\nabla^k R\right|^2 + B\sqrt{t^{k+2}\left|\nabla^k R\right|^2} + C \leq \Delta\Phi + D.$$

As in Lemma 5.32, this suffices to yield the result. □

Now that our proof of Proposition 5.33 is complete, Theorem 5.28 follows readily.

7. Strategy for the case that $\chi\left(\mathcal{M}^2 > 0\right)$

We have established Theorem 5.1 for surfaces such that $r \leq 0$. In the remainder of this chapter, we tackle the far more difficult case that $r > 0$. For the convenience of the reader, we now summarize the path we will take to prove that the normalized Ricci flow on a surface of positive Euler characteristic converges exponentially to a metric of constant positive curvature.

7.1. The case that $R(\cdot, 0) \geq 0$. We first consider the special case that the scalar curvature is initially nonnegative. By the strong maximum principle, one then has $R_{\min}(t) > 0$ for any $t > 0$ unless $R \equiv 0$ everywhere, which is possible only if the initial manifold was a flat torus. Hence (by restarting the flow after some fixed time $\varepsilon > 0$ has elapsed) we may assume that $R(\cdot, 0) > 0$.

Recall that the trace-free part of the Hessian of the potential f of the curvature is the tensor M defined in (5.9) as

$$M = \nabla\nabla f - \frac{1}{2}\Delta f \cdot g,$$

where by (5.8), one has $\Delta f = R - r$. In Section 3, we observed that the tensor M vanishes identically on a Ricci soliton. And in Section 4, we saw that the only self-similar solutions of the Ricci flow on a compact surface are the metrics of constant curvature. Taken together, these observations suggest one might be able to show that $g(t)$ converges to a metric of constant positive curvature by proving that M decays sufficiently rapidly. To explore this idea, we compute the evolution equation satisfied by M.

LEMMA 5.34. *On a solution* $(\mathcal{M}^2, g(t))$ *of the normalized Ricci flow, the tensor* M *evolves by*

$$\frac{\partial}{\partial t}M = \Delta M + (r - 2R)\,M.$$

PROOF. By Lemma 3.2, the Levi-Civita connection of g evolves by

$$\begin{aligned}
\frac{\partial}{\partial t}\Gamma_{ij}^k &= \frac{1}{2}g^{k\ell}\left[\nabla_i\left(\frac{\partial}{\partial t}g_{j\ell}\right) + \nabla_j\left(\frac{\partial}{\partial t}g_{i\ell}\right) - \nabla_\ell\left(\frac{\partial}{\partial t}g_{ij}\right)\right] \\
&= \frac{1}{2}g^{k\ell}\left(-\nabla_i R \cdot g_{j\ell} - \nabla_j R \cdot g_{i\ell} + \nabla_\ell R \cdot g_{ij}\right) \\
&= \frac{1}{2}\left(-\nabla_i R \cdot \delta_j^k - \nabla_j R \cdot \delta_i^k + \nabla^k R \cdot g_{ij}\right).
\end{aligned}$$

Using this formula and recalling equations (5.3) and (5.10), we calculate that

$$\begin{aligned}
\frac{\partial}{\partial t}M_{ij} &= \frac{\partial}{\partial t}\left(\nabla_i\nabla_j f - \frac{1}{2}(R - r)\,g_{ij}\right) \\
&= \nabla_i\nabla_j\left(\frac{\partial f}{\partial t}\right) - \left(\frac{\partial}{\partial t}\Gamma_{ij}^k\right)\nabla_k f - \frac{1}{2}\left(\frac{\partial}{\partial t}R\right)g_{ij} - \frac{1}{2}(R - r)\frac{\partial}{\partial t}g_{ij} \\
&= \nabla_i\nabla_j(\Delta f + rf) + \frac{1}{2}\left(\nabla_i R \cdot \delta_j^k + \nabla_j R \cdot \delta_i^k - \nabla^k R \cdot g_{ij}\right)\nabla_k f \\
&\quad - \frac{1}{2}\left[\Delta R + R(R - r)\right]g_{ij} + \frac{1}{2}(R - r)^2 g_{ij} \\
&= \nabla_i\nabla_j\Delta f + \frac{1}{2}\left(\nabla_i R \nabla_j f + \nabla_i f \nabla_j R - \langle\nabla R, \nabla f\rangle g_{ij}\right) \\
&\quad - \frac{1}{2}(\Delta R)\,g_{ij} + rM_{ij}.
\end{aligned}$$

Next we use the fact that

$$R_{ijk\ell} = \frac{R}{2}\left(g_{i\ell}g_{jk} - g_{ik}g_{j\ell}\right)$$

on a surface to compute the commutator $[\nabla\nabla, \Delta]$, obtaining

$$\begin{aligned}
\nabla_i \nabla_j \Delta f &= \nabla_i \nabla_j \nabla_k \nabla^k f \\
&= \nabla_i \nabla_k \nabla_j \nabla^k f - \nabla_i \left(R_{j\ell} \nabla^\ell f \right) \\
&= \nabla_k \nabla_i \nabla_j \nabla^k f - R_{ikj}^\ell \nabla_\ell \nabla^k f - R_{i\ell} \nabla_j \nabla^\ell f \\
&- R_{j\ell} \nabla_i \nabla^\ell f - \nabla_i R_{j\ell} \nabla^\ell f \\
&= \Delta \nabla_i \nabla_j f - \nabla^k \left(R_{ikj}^\ell \nabla_\ell f \right) - R_{ikj}^\ell \nabla_\ell \nabla^k f \\
&- R_{i\ell} \nabla_j \nabla^\ell f - R_{j\ell} \nabla_i \nabla^\ell f - \nabla_i R_{j\ell} \nabla^\ell f \\
&= \Delta \nabla_i \nabla_j f - \frac{1}{2} \left(\nabla_i R \nabla_j f + \nabla_i f \nabla_j R - \langle \nabla R, \nabla f \rangle g_{ij} \right) \\
&- 2R \left(\nabla_i \nabla_j f - \frac{1}{2} (\Delta f) g_{ij} \right).
\end{aligned}$$

Combining these results, we get

$$\begin{aligned}
\frac{\partial}{\partial t} M_{ij} &= \Delta \nabla_i \nabla_j f - \frac{1}{2} (\Delta R) g_{ij} + (r - 2R) M_{ij} \\
&= \Delta \left(\nabla_i \nabla_j f - \frac{1}{2} (R - r) g_{ij} \right) + (r - 2R) M_{ij}.
\end{aligned}$$

□

COROLLARY 5.35. *On a solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow, the norm squared of the tensor M evolves by*

$$\frac{\partial}{\partial t} |M|^2 = \Delta |M|^2 - 2 |\nabla M|^2 - 2R |M|^2. \tag{5.22}$$

PROOF. Recalling Lemma 3.1 and using the result above, we obtain

$$\begin{aligned}
\frac{\partial}{\partial t} |M|^2 &= \frac{\partial}{\partial t} \left(g^{ik} g^{j\ell} M_{ij} M_{k\ell} \right) \\
&= 2 \langle M, \Delta M + (r - 2R) M \rangle + 2 (R - r) |M|^2 \\
&= \Delta |M|^2 - 2 |\nabla M|^2 - 2R |M|^2.
\end{aligned}$$

□

Equation (5.22) is the key result that motivates the following strategy. If we can prove that $R \geq c$ for some constant $c > 0$ independent of t, we will obtain an estimate of the sort

$$|M| \leq Ce^{-ct}.$$

Then we can consider the modified Ricci flow

$$\frac{\partial}{\partial t} g = 2M = 2\nabla\nabla f - (R - r) g = (r - R) g + \mathcal{L}_{\nabla f} g. \tag{5.23}$$

As we saw in Section 4 of Chapter 2, the solution to (5.23) differs from the solution to the normalized Ricci flow

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

only by by the one-parameter family of diffeomorphisms φ_t generated by the time-dependent vector fields $\nabla f(t)$. Since $\mathcal{M}^2$ is compact, Lemma 3.15 implies that these diffeomorphisms exist as long as the potential function $f(t)$ does. And since the quantity $|M|^2$ is invariant under diffeomorphism, the estimate $|M|^2\leq Ce^{-ct}$ will hold on the solution to the modified Ricci flow (5.23). Moreover, as we did for the curvature in Sections 5 and 6, one can obtain estimates for all derivatives of M, which prove that the solution $g(t)$ to the modified flow converges exponentially fast in all C^k to a metric g_∞ such that M_∞ vanishes identically. By equation (5.7), this will imply that g_∞ is a gradient soliton. Then Proposition 5.21 (which uses the Kazdan–Warner identity, hence the Uniformization Theorem) will imply that g_∞ is a metric of constant positive curvature. It will then follow that there exist positive constants c_k, C_k for each $k\in\mathbb{N}$ such that the solution $g(t)$ of the modified flow satisfies

$$\left|\nabla^k R\right|\leq C_k e^{-c_k t}.$$

But by diffeomorphism invariance, the same estimates must hold for the solution of the unmodified flow. In this way, we will be able to conclude that the normalized Ricci flow starting at a metric of strictly positive scalar curvature converges exponentially fast to a constant curvature metric.

In order to obtain uniform positive bounds for R, we proceed as follows. We begin by developing two important technical tools.

STEP 1. The *surface entropy* of a compact 2-manifold $(\mathcal{M}^2, g(t))$ of positive curvature evolving by the normalized Ricci flow is defined as

$$N(g(t))\doteqdot\int_{\mathcal{M}^2}R\log R\,dA.$$

In Section 8, we will prove that $N(g(t))$ is strictly decreasing unless $g(t)$ is Einstein; we shall later use this result to obtain uniform bounds on the scalar curvature of a solution to the normalized Ricci flow on a surface of positive Euler characteristic. (If $R(0)$ changes sign, there is a modified entropy formula; in this case, in this case, we will prove that $N(g(t))$ remains bounded.)

STEP 2. In Section 9, we obtain uniform bounds for the metric, the scalar curvature, the gradient of the scalar curvature, and the diameter of a solution on a surface of positive Euler characteristic. In particular, we show that if $|R|\leq\kappa$ on a time interval $[t_0, t_0+1/(4\kappa)]$, then $|\nabla R|\leq 4\kappa^{3/2}$ on the same interval. Together with the entropy result obtained in Step 1, we use this to argue that $|R|$ is uniformly bounded independently of time, and that the diameter of the solution is bounded. In this step, the positivity of the curvature allows us to apply Klingenberg's Theorem to obtain an injectivity radius bound.

STEP 3. In Section 10, we derive a *differential Harnack estimate of Li-Yau-Hamilton type* for the normalized Ricci flow. This is a lower bound for the time derivative of the curvature

$$\frac{\partial}{\partial t} \log R - |\nabla \log R|^2 .$$

(Again, there is a modified formula for the case that $R(0)$ changes sign.) Integrating this along a space-time path from (x_1, t_1) to (x_2, t_2) such that $0 < t_1 < t_2$ yields a classical Harnack inequality, which is an *a priori* comparison of the sort

$$\frac{R(x_2, t_2)}{R(x_1, t_1)} \geq \frac{t_1}{t_2} \exp\left[-\frac{d(x_1, x_2, t_1)^2}{c(t_2 - t_1)}\right].$$

STEP 4. In Section 11, we combine the differential Harnack estimate with the bounds on $|R|$ and the diameter derived in Step 2 to obtain a positive lower bound for R. As outlined above, this allows us to prove the desired convergence result for an initial metric of strictly positive curvature.

REMARK 5.36. In terms of getting a positive lower bound for R, the BBS estimates are useful locally, whereas the differential Harnack estimate is useful globally. Indeed, if

$$R(x, t) = \kappa \doteqdot R_{\max}(t) > 0,$$

then the BBS estimates show that $R(\cdot, t) > \kappa/2$ in a $g(t)$-metric ball of radius $c/\sqrt{\kappa}$, where $c > 0$ is a universal constant. On the other hand, if the diameter is bounded, then the Harnack estimate yields a uniform lower bound for R.

7.2. The case that $R(\cdot, 0)$ changes sign. To obtain convergence for an arbitrary initial metric having $r > 0$, we prove that the scalar curvature of any such solution eventually becomes everywhere positive. Once this happens, the argument outlined above goes through.

In order to obtain suitable bounds on the curvature, we need the modified entropy formula and the modified Harnack estimate that apply when the curvature is of mixed sign. We also need an *ad hoc* injectivity radius estimate, since Klingenberg's Theorem may not apply if the curvature is negative somewhere. We obtain an adequate estimate for the injectivity radius in Section 12, and bound the curvature in Section 13.

In Sections 14 and 15, we illustrate alternative methods for bounding the injectivity radius and curvature of solutions whose curvature is initially of mixed sign.

8. Surface entropy

Since we have not been able to employ the maximum principle to obtain a uniform upper bound for the curvature in the case that $r > 0$, we shall consider integral quantities. Perhaps the most important such quantity for the Ricci flow on surfaces is the **surface entropy**.

8.1. The case that $R(\cdot, 0) > 0$. We first consider the case that $R > 0$. The surface entropy N is defined for a metric of strictly positive curvature on a closed manifold $\mathcal{M}^n$ by

$$N(g) \doteqdot \int_{\mathcal{M}^n} R \log R \, d\mu.$$

The reason this quantity is called an *entropy* is because it formally resembles certain classical entropies, each of which is the integral of a positive function times its logarithm. Our sign convention is opposite the usual one, and so we shall show that the entropy is decreasing (instead of increasing) under the normalized Ricci flow. We first compute the time derivative of the entropy.

LEMMA 5.37. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow on a compact surface, then*

$$\frac{\partial}{\partial t}(R dA) = \Delta R \, dA. \tag{5.24}$$

PROOF. By (5.2) and (5.3), we have

$$\begin{aligned} \frac{\partial}{\partial t}(R \, dA) &= \left(\frac{\partial}{\partial t} R\right) dA + R\left(\frac{\partial}{\partial t} dA\right) \\ &= [\Delta R + R(R-r)] \, dA + R(r-R) \, dA. \end{aligned}$$

□

LEMMA 5.38. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow on a compact surface such that $R(\cdot, 0) > 0$, then the entropy evolves by*

$$\frac{dN}{dt} = -\int_{\mathcal{M}^2} \frac{|\nabla R|^2}{R} dA + \int_{\mathcal{M}^2} (R-r)^2 \, dA \tag{5.25}$$

PROOF. Using (5.24) and integrating by parts, we calculate

$$\begin{aligned} \frac{dN}{dt} &= \int_{\mathcal{M}^2} \left(\frac{\partial}{\partial t} \log R\right) R \, dA + \int_{\mathcal{M}^2} \log R \cdot \frac{\partial}{\partial t}(R \, dA) \\ &= \int_{\mathcal{M}^2} [\Delta R + R(R-r)] \, dA \\ &\quad + \int_{\mathcal{M}^2} \log R \cdot \Delta R \, dA \\ &= \int_{\mathcal{M}^2} R(R-r) \, dA - \int_{\mathcal{M}^2} \frac{|\nabla R|^2}{R} dA. \end{aligned}$$

□

PROPOSITION 5.39. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow on a compact surface with $R(\cdot, 0) > 0$, then*

$$\frac{dN}{dt} = -\int_{\mathcal{M}^2} \frac{|\nabla R + R\nabla f|^2}{R} \, dA - 2\int_{\mathcal{M}^2} |M|^2 \, dA \le 0,$$

where M is the trace-free part of the Hessian of f defined in (5.9). In particular, the entropy is strictly decreasing unless $(\mathcal{M}^2, g(t))$ is a gradient Ricci soliton.

PROOF. Recall that $\Delta f = R - r$. Expanding the first term on the right-hand side and integrating by parts, we obtain

(5.26)
$$\int_{\mathcal{M}^2} \frac{|\nabla R + R\nabla f|^2}{R} \, dA = \int_{\mathcal{M}^2} \left(\frac{|\nabla R|^2}{R} - 2R(R-r) + R|\nabla f|^2 \right) dA.$$

On the other hand, commuting covariant derivatives and integrating by parts shows that

$$\begin{aligned}
\int_{\mathcal{M}^2} (R-r)^2 \, dA &= \int_{\mathcal{M}^2} (\Delta f)^2 \, dA \\
&= -\int_{\mathcal{M}^2} \langle \nabla f, \nabla \Delta f \rangle \, dA \\
&= -\int_{\mathcal{M}^2} \left(\langle \nabla f, \Delta \nabla f \rangle - \mathrm{Rc}(\nabla f, \nabla f) \right) dA \\
&= \int_{\mathcal{M}^2} \left(|\nabla\nabla f|^2 + \frac{1}{2} R |\nabla f|^2 \right) dA,
\end{aligned}$$

hence that

$$\begin{aligned}
-2\int_{\mathcal{M}^2} |M|^2 \, dA &= \int_{\mathcal{M}^2} \left((\Delta f)^2 - 2|\nabla\nabla f|^2 \right) dA \\
&= \int_{\mathcal{M}^2} \left(\frac{1}{2} R |\nabla f|^2 - |\nabla\nabla f|^2 \right) dA \\
&= \int_{\mathcal{M}^2} \left(R|\nabla f|^2 - (R-r)^2 \right) dA. \qquad (5.27)
\end{aligned}$$

Subtracting (5.26) from (5.27) yields

$$-2\int_{\mathcal{M}^2} |M|^2 \, dA - \int_{\mathcal{M}^2} \frac{|\nabla R + R\nabla f|^2}{R} \, dA = \int_{\mathcal{M}^2} \left((R-r)^2 - \frac{|\nabla R|^2}{R} \right) dA,$$

whence the result follows by Lemma 5.38. □

COROLLARY 5.40. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a closed surface with $R(\cdot, 0) > 0$. Then the entropy is a strictly-decreasing function of time unless $R(\cdot, 0) \equiv r$, in which case it is constant in time.*

PROOF. If $dN/dt = 0$ at some time $t_0 \in [0, \infty)$, then $M(\cdot, t_0) \equiv 0$. By equation (5.7), $g(t_0)$ is a gradient Ricci soliton. Hence $R(\cdot, t_0) \equiv r$ is constant by Proposition 5.21, and thus $g(t) \equiv g(t_0)$. □

8.2. The case that $R(\cdot, 0)$ changes sign. We want to extend our previous entropy estimates to initial metrics on $\mathcal{S}^2$ or $\mathbb{RP}^2$ whose curvature changes sign. We first need to define a suitable modification of the notion of entropy. By passing to the double cover if necessary, we may assume $\mathcal{M}^2$ is diffeomorphic to $\mathcal{S}^2$. Recall that the ODE corresponding to the PDE satisfied by R is

$$\frac{d}{dt}s = s(s-r).$$

Recall that $r > 0$ on $\mathcal{M}^2 \approx \mathcal{S}^2$, and let $s(t)$ be the ODE solution with initial condition $s(0) = s_0 < R_{\min}(0) < 0$, namely

$$s(t) = \frac{r}{1 - \left(1 - \frac{r}{s_0}\right) e^{rt}}. \qquad (s_0 < R_{\min}(0) < 0) \tag{5.28}$$

Then the difference of R and s evolves by

$$\frac{\partial}{\partial t}(R-s) = \Delta\,(R-s) + (R-r+s)(R-s).$$

Since $R_{\min}(0) - s_0 > 0$, the maximum principle implies that $R - s > 0$ for as long as the solution of the normalized Ricci flow exists. This suggests that we define a **modified entropy** by

$$\hat{N}(g(t), s(t)) \doteqdot \int_{\mathcal{M}^2} (R-s)\log(R-s)\, dA. \tag{5.29}$$

LEMMA 5.41. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a closed surface with an arbitrary initial metric satisfying $r > 0$. Then*

$$\frac{d}{dt}\hat{N} = \int_{\mathcal{M}^2}\left(-\frac{|\nabla R|^2}{R-s} + (R-s)(R-r+s+s\log(R-s))\right) dA.$$

PROOF. Put $P \doteqdot R - s$. Noting that

$$\frac{\partial}{\partial t}(P\,dA) = (\Delta P + (R-r+s)P)\,dA + P(r-R)\,dA = (\Delta P + sP)\,dA,$$

we integrate by parts to obtain

$$\begin{aligned}\frac{d}{dt}\hat{N} &= \int_{\mathcal{M}^2} [(\Delta P + (R-r+s)P) + (\log P)(\Delta P + sP)]\,dA \\ &= \int_{\mathcal{M}^2}\left[-\frac{1}{P}|\nabla P|^2 + (R-r+s+s\log P)P\right] dA.\end{aligned}$$

□

When the initial metric has strictly positive curvature, Corollary 5.40 shows that the entropy is decreasing. Not surprisingly, one can no longer prove this result for initial metrics whose curvature changes sign. However, one can show that the entropy is uniformly bounded from above, which is what we need for our applications.

PROPOSITION 5.42. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow corresponding to an arbitrary initial metric with $r > 0$. Then the modified entropy evolves by*

$$\begin{aligned} \frac{d}{dt}\hat{N} &= -\int_{\mathcal{M}^2} \frac{|\nabla R + (R-s)\nabla f|^2}{R-s}\, dA - 2\int_{\mathcal{M}^2} |M|^2\, dA \\ &\quad - s\int_{\mathcal{M}^2} \left(|\nabla f|^2 + s - r - (R-s)\log(R-s)\right) dA, \end{aligned} \tag{5.30}$$

where M is the trace-free part of $\nabla\nabla f$.

PROOF. Expand the first term on the right-hand side and integrate by parts to get

$$\begin{aligned} &\int_{\mathcal{M}^2} \frac{|\nabla R + (R-s)\nabla f|^2}{R-s}\, dA \\ &\quad = \int_{\mathcal{M}^2} \left(\frac{|\nabla R|^2}{R-s} - 2R(R-r) + (R-s)|\nabla f|^2\right) dA. \end{aligned}$$

But as we observed in (5.27),

$$-2\int_{\mathcal{M}^2} |M|^2\, dA = \int_{\mathcal{M}^2} \left(R|\nabla f|^2 - R(R-r)\right) dA.$$

Subtracting these identities yields

$$\begin{aligned} &-2\int_{\mathcal{M}^2} |M|^2\, dA - \int_{\mathcal{M}^2} \frac{|\nabla R + (R-s)\nabla f|^2}{R}\, dA \\ &\quad = \int_{\mathcal{M}^2} \left(R(R-r) + s|\nabla f|^2 - \frac{|\nabla R|^2}{R-s}\right) dA, \end{aligned}$$

whence the proposition follows by Lemma 5.38. □

Since $s < 0$, we expect the last line of equation (5.30) to be positive. The most difficult term to control is $\int_{\mathcal{M}^2} |\nabla f|^2\, dA$. To obtain an upper bound for the modified entropy, therefore, we shall estimate its integral with respect to time.

LEMMA 5.43. *There exists a constant C depending only on $g(0)$ such that if the solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow with $r > 0$ exists for $0 \le t < T$, then*

$$\int_0^T e^{-rt} \int_{\mathcal{M}^2} |\nabla f|^2\, dA\, dt \le C.$$

PROOF. By (5.10) and (5.2),

$$\begin{aligned}\frac{d}{dt}\int_{\mathcal{M}^2} f\,dA &= \int_{\mathcal{M}^2} \left((\Delta f + rf) + f\,(r-R)\right) dA \\ &= \int_{\mathcal{M}^2} \left((\Delta f + rf) - f\Delta f\right) dA \\ &= \int_{\mathcal{M}^2} \left(rf + |\nabla f|^2\right) dA.\end{aligned}$$

Hence

$$\frac{d}{dt}\left(e^{-rt}\int_{\mathcal{M}^2} f\,dA\right) = e^{-rt}\int_{\mathcal{M}^2} |\nabla f|^2\,dA.$$

But by Corollary 5.14, we have $\int_{\mathcal{M}^2} |f|\,dA \leq Ce^{rt}$ for some $C<\infty$, so that integrating the equation above with respect to time yields

$$\int_0^T e^{-rt}\int_{\mathcal{M}^2} |\nabla f|^2\,dA\,dt = \left[e^{-rt}\int_{\mathcal{M}^2} f\,dA\right]_{t=0}^{t=T} \leq 2C.$$

□

PROPOSITION 5.44. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow corresponding to an arbitrary initial metric with $r>0$. Then there is a constant C depending only on $g(0)$ such that*

$$\hat{N}\left(g(t), s(t)\right) \leq C.$$

PROOF. By (5.28), there is $C>0$ depending only on $g(0)$ such that $-Ce^{-rt} \leq s \leq 0$. Hence (5.30) implies that

$$\begin{aligned}\frac{d}{dt}\hat{N} &\leq -s\int_{\mathcal{M}^2}\left(|\nabla f|^2 + s - r - (R-s)\log(R-s)\right) dA \\ &\leq Ce^{-rt}\int_{\mathcal{M}^2}|\nabla f|^2\,dA + Ce^{-rt}\left|\hat{N}\right|.\end{aligned}$$

The proposition follows by integrating this inequality with respect to time and applying Lemma 5.43. □

9. Uniform upper bounds for R and $|\nabla R|$

In this section, we derive further estimates for a solution $g(t)$ on a surface of positive Euler characteristic. In the first part, we obtain a key *doubling-time estimate,* and derive upper and lower bounds for $g(t)$. In the second part, we derive a BBS estimate for the gradient of the scalar curvature. In the final part, we apply the entropy estimate to obtain uniform bounds for the scalar curvature itself under the additional hypothesis that $R(\cdot, 0) > 0$.

9.1. Bounds for the metric on a surface with $\chi\left(\mathcal{M}^2\right) > 0$. Define $R_{\min}(t) \doteqdot \min_{x\in\mathcal{M}^2} R(x,t)$ and $R_{\max}(t) \doteqdot \max_{x\in\mathcal{M}^2} R(x,t)$.

LEMMA 5.45 (Doubling-time estimate). *Let $\left(\mathcal{M}^2, g(t)\right)$ be any solution of the normalized Ricci flow on a closed surface with $r > 0$. Then for any $t_0 \in [0,\infty)$, the estimate*

$$R_{\max}(t) \leq 2R_{\max}(t_0)$$

holds for all $x \in \mathcal{M}^2$ and

$$t \in \left[t_0, t_0 + \frac{1}{2R_{\max}(t_0)}\right].$$

PROOF. By (5.3), the scalar curvature evolves by

$$\frac{\partial}{\partial t}R = \Delta R + R^2 - rR.$$

Since $r > 0$, we have $R_{\max}(t) > 0$ for all time. So at a maximum in space,

$$\frac{\partial}{\partial t}R \leq \Delta R + R^2.$$

The solution of the initial value problem

$$\frac{d\rho}{dt} = \rho^2, \qquad \rho(t_0) = R_{\max}(t_0)$$

is

$$\rho(t) = \frac{1}{R_{\max}^{-1}(t_0) + t_0 - t}.$$

Hence the maximum principle implies that for $t_0 \leq t \leq t_0 + 1/2R_{\max}(t_0)$,

$$R_{\max}(t) \leq 2R_{\max}(t_0).$$

□

REMARK 5.46. This is a prototype of the general **Doubling-time estimate** derived in Corollary 7.5.

The lemma implies that the metrics are uniformly equivalent in the same time interval. In particular, a lower bound for the metric is readily obtained.

COROLLARY 5.47. *Let $\left(\mathcal{M}^2, g(t)\right)$ be any solution of the normalized Ricci flow on a closed surface with $r > 0$. Given any $t_0 \in [0,\infty)$, the estimate*

$$g(x,t) \geq \frac{1}{e} g(x,t_0)$$

holds for all $x \in \mathcal{M}^2$ and $t \in \left[t_0, t_0 + \frac{1}{2R_{\max}(t_0)}\right]$.

PROOF. At any $x \in \mathcal{M}^2$, we may write the metric as

$$g(x,t) = e^{\int_{t_0}^{t}(r - R(x,\tau))\,d\tau} g(x,t_0). \tag{5.31}$$

If $t_0 \le t \le t_0 + 1/2R_{\max}(t_0)$, then Lemma 5.45 implies that

$$\begin{aligned}\int_{t_0}^{t} (r - R(x,\tau))\, d\tau &\ge -\int_{t_0}^{t} R(x,\tau)\, d\tau \\ &\ge -2\int_{t_0}^{t_0+\frac{1}{2R_{\max}(t_0)}} R_{\max}(t_0)\, d\tau = -1.\end{aligned}$$

□

It takes only slightly more work to derive an upper bound for the metric.

LEMMA 5.48. *Let $(\mathcal{M}^2, g(t))$ be any solution of the normalized Ricci flow on a closed surface with $r > 0$.*

- *If $R(\cdot,0) \ge 0$, then for any times $0 \le t_0 \le t < \infty$,*
$$g(x,t) \le e^{r(t-t_0)} g(x,t_0).$$
- *If $R(\cdot,0)$ changes sign, then for any times $0 \le t_0 \le t < \infty$,*
$$g(x,t) \le \left[e^{r(t-t_0)} \frac{\left(1 - \frac{r}{s_0}\right) - e^{-rt}}{\left(1 - \frac{r}{s_0}\right) - e^{-rt_0}} \right] g(x,t_0).$$

PROOF. Recall from Section 2 that $R_{\min}(t)$ is bounded below by the solution of the ODE

$$\frac{d}{dt}s = s(s-r), \qquad s(0) = \begin{cases} 0 & \text{if } R_{\min}(0) \ge 0 \\ R_{\min}(0) & \text{if } R_{\min}(0) < 0, \end{cases}$$

namely

$$s(t) = \begin{cases} 0 & \text{if } R_{\min}(0) \ge 0 \\ \frac{r}{1-(1-r/R_{\min}(0))e^{rt}} & \text{if } R_{\min}(0) < 0. \end{cases}$$

Hence for any $x \in \mathcal{M}^2$, we have

$$\int_{t_0}^{t} (r - R(x,\tau))\, d\tau \le \int_{t_0}^{t} (r - s(\tau))\, d\tau.$$

In the case that $R_{\min}(0) \ge 0$, it follows therefore from (5.31) that

$$g(x,t) \le e^{\int_{t_0}^{t}(r-s(\tau))\,d\tau} g(x,t_0) \le e^{r(t-t_0)} g(x,t_0).$$

In the case that $R(\cdot,0)$ changes sign, we compute

$$-\int_{t_0}^{t} s(\tau)\, d\tau = \int_{t_0}^{t} \frac{-re^{-r\tau}}{e^{-r\tau} - \left(1 - \frac{r}{s_0}\right)}\, d\tau = \log\left[e^{-r\tau} - \left(1 - \frac{r}{s_0}\right)\right]_{\tau=t_0}^{\tau=t}$$

in order to conclude that

$$g(x,t) \le e^{\int_{t_0}^{t}(r-s(\tau))\,d\tau} g(x,t_0) \le \left[e^{r(t-t_0)} \frac{\left(1 - \frac{r}{s_0}\right) - e^{-rt}}{\left(1 - \frac{r}{s_0}\right) - e^{-rt_0}} \right] g(x,t_0).$$

□

When the initial scalar curvature is nonnegative, these two lemmas lead to the following observation.

COROLLARY 5.49. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a closed surface with $R(\cdot,0) \geq 0$. Then for any $t_0 \in [0,\infty)$, the estimate*

$$\frac{1}{e} g(x,t_0) \leq g(x,t) \leq \sqrt{e}\, g(x,t_0) \tag{5.32}$$

holds for all $x \in \mathcal{M}^2$ and $t \in \left[t_0, t_0 + \frac{1}{2R_{\max}(t_0)}\right]$.

9.2. Estimating the gradient of the scalar curvature. If the scalar curvature R is bounded above by some constant κ on a time interval $[0,\tau)$, then applying the maximum principle to (5.20) proves that

$$|\nabla R|^2 \leq e^{4\kappa t - 3\int_0^t r(\bar{t})\,d\bar{t}} \sup_{x\in\mathcal{M}^2} |\nabla R|^2 (x,0)\,.$$

However, we prefer a bound for $|\nabla R|^2$ which depends only on the initial scalar curvature $R(\cdot,0)$ and not on its derivatives. Of course, any such bound must blow up as $t \searrow 0$.

PROPOSITION 5.50. *There exists a universal constant $C < \infty$ such that for any solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow on a compact surface that satisfies $r \geq 0$ and $|R(\cdot,0)| \leq \kappa$ for some $\kappa > 0$, the estimate*

$$\sup_{x\in\mathcal{M}^2} |\nabla R|(x,t) \leq \frac{C\kappa}{\sqrt{t}}$$

holds for all times $0 < t \leq 1/(C\kappa)$.

PROOF. The idea is to compute the evolution equation of

$$G \doteqdot t\,|\nabla R|^2 + R^2,$$

and apply the maximum principle. Notice that $G^2 \leq R^2 \leq \kappa^2$ at $t=0$. It follows from (5.20) that

$$\frac{\partial}{\partial t}\left(t\,|\nabla R|^2\right) = \Delta\left(t\,|\nabla R|^2\right) - 2t\,|\nabla\nabla R|^2 + [t(4R-3r)+1]\,|\nabla R|^2\,.$$

Adding this equation to

$$\frac{\partial}{\partial t}(R^2) = \Delta(R^2) - 2\,|\nabla R|^2 + 2R^2(R-r)$$

shows that

$$\frac{\partial}{\partial t} G \leq \Delta G + (4tR-1)\,|\nabla R|^2 + 2R^3. \tag{5.33}$$

In order to apply the maximum principle to (5.33), we need to estimate $|R|$ on a suitable time interval. Since

$$\frac{\partial}{\partial t} R \geq R(R-r)$$

at a minimum of R, we have $R_{\min}(t) \geq \min\{0, R_{\min}(0)\} \geq -\kappa$. And since $r \geq 0$, we know $R_{\max} \geq 0$, which implies that

$$\frac{\partial}{\partial t} R \leq R^2$$

at a maximum of R. Combining these observations proves

$$|R| \leq \frac{\kappa}{1 - \kappa t}$$

for $0 \leq t < 1/\kappa$. This implies that $|R| \leq 2\kappa$ if $0 \leq t \leq 1/(2\kappa)$, and hence that $4t|R| \leq 1$ if $0 \leq t \leq 1/(8\kappa)$. Thus equation (5.33) yields the estimate

$$\frac{\partial}{\partial t} G \leq \Delta G + 16\kappa^3$$

for times $0 \leq t \leq 1/(8\kappa)$, whence it follows from the maximum principle that on that interval,

$$G \leq \kappa^2 + 16\kappa^3 t \leq \kappa^2 + \frac{16\kappa^3}{8\kappa} = 3\kappa^2.$$

□

9.3. Estimates for solutions with $R(\cdot, 0) > 0$. In the remainder of this section, we assume that $R(\cdot, 0) \geq 0$. By the strong maximum principle, one then has $R_{\min}(t) > 0$ for any $t > 0$ unless $R \equiv 0$ everywhere, which is possible only if the initial manifold was a flat torus. Hence (by restarting the flow after some fixed time $\varepsilon > 0$ has elapsed) we may assume that $R(\cdot, 0) > 0$. Then it follows from Lemma 5.9 and Proposition 5.18 that

$$ce^{-rt} \leq R(x,t) \leq Ce^{rt}$$

for some constants $c, C \in (0, \infty)$ depending only on g_0. We can now obtain a uniform upper bound for R by using the entropy estimate.

PROPOSITION 5.51. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a closed surface of positive Euler characteristic. If $R(\cdot, 0) > 0$, then there exists a constant $C \in [1, \infty)$ depending only on g_0 such that*

$$\sup_{\mathcal{M}^2 \times [0,\infty)} R \leq C.$$

PROOF. Since we already have time-dependent bounds for R, we may set

$$\kappa_1 \doteqdot \max_{\mathcal{M}^2 \times [0,1]} R.$$

Given any time $T \in [1, \infty)$, define

$$\kappa(T) \doteqdot \max_{\mathcal{M}^2 \times [0,T]} R \geq \kappa_1.$$

We want to show that κ is bounded independently of T. Assume that $\kappa(T) > \max\{\kappa_1, 1/4\}$, so that $T > 1$. Let $(x_1, t_1) \in \mathcal{M}^2 \times (1, T]$ be a point such that

$$R(x_1, t_1) = \max_{\mathcal{M}^2 \times [0,T]} R = \kappa.$$

Let $t_0 \doteqdot t_1 - \frac{1}{4\kappa} > 0$. Examining the proof of Proposition 5.50, we see that if $|R| \leq \kappa$ on a time interval $[t_0, t_0 + 1/(4\kappa)]$, then

$$|\nabla R(x,t)| \leq \frac{2\kappa}{\sqrt{t - t_0}}$$

for all $x \in \mathcal{M}^2$ and $t \in (t_0, t_0 + 1/(4\kappa)]$. In particular, we have

$$|\nabla R(x, t_1)| \leq 4\kappa^{3/2}$$

for all $x \in \mathcal{M}^2$. Denote by $B_{g(t_1)}(x,\rho)$ the ball with center $x \in \mathcal{M}^2$ and radius ρ, measured with respect to the metric $g(t_1)$. Let

$$y \in B_{g(t_1)}\left(x_1, 1/\sqrt{64\kappa}\right),$$

and choose a unit speed minimal geodesic $\gamma(s)$ joining y to x_1 at time t_1. Then we have

$$\begin{aligned} R(x_1, t_1) - R(y, t_1) &= \int_\gamma \frac{\partial}{\partial s}[R(\gamma(s), t_1)]\, ds \\ &\leq \int_\gamma |\nabla R(\gamma(s), t_1)|\, ds \leq \frac{4\kappa^{3/2}}{8\kappa^{1/2}} = \frac{\kappa}{2}. \end{aligned}$$

Hence we get

$$R(y, t_1) \geq \frac{\kappa}{2}$$

for all $y \in B_{g(t_1)}\left(x_1, 1/\sqrt{64\kappa}\right)$.

Now by Klingenberg's Theorem, the injectivity radius of any closed orientable surface whose sectional curvatures k are bounded by $0 < k \leq k^*$ satisfies $\operatorname{inj}\left(\mathcal{M}^2, g\right) \geq \pi/\sqrt{k^*}$. (See Theorem 5.9 of [**27**].) Hence

$$\operatorname{inj}\left(\mathcal{M}^2, g(t_1)\right) \geq \frac{\pi}{\sqrt{\frac{R_{\max}(t_1)}{2}}} > \frac{\pi}{\sqrt{\kappa}}.$$

By the maximum principle, we have $R(\cdot, t_1) > 0$. So the entropy $N(g(t_1))$ is well defined. Then since $R(\cdot, 0) \geq 0$, we can let $s_0 \nearrow 0$ in Proposition 5.44 to obtain $C(g_0)$ such that $N(g(t_1)) \leq C$. Since $\inf_{y>0}(y \log y) = -1/e$, we have

$$\begin{aligned} C \geq N(g(t_1)) &\doteqdot \int_{\mathcal{M}^2} R(\log R)\, dA \\ &\geq \int_{B_{g(t_1)}\left(x_1, 1/\sqrt{64\kappa}\right)} R(\log R)\, dA - \frac{1}{e}\operatorname{Area}\left(\mathcal{M}^2, g(t_1)\right). \end{aligned}$$

But the area comparison theorem (§3.4 of [**25**]) implies that there is a universal constant $c > 0$ such that

$$\begin{aligned} \int_{B_{g(t_1)}\left(x_1, 1/\sqrt{64\kappa}\right)} R(\log R)\, dA &\geq \frac{\kappa}{2}\left(\log\frac{\kappa}{2}\right) \cdot \operatorname{Area}\left(B_{g(t_1)}\left(x_1, 1/\sqrt{64\kappa}\right)\right) \\ &\geq c \log\frac{\kappa}{2}. \end{aligned}$$

We conclude that $\kappa(T)$ has a uniform upper bound. $\square$

The uniform upper bound for the scalar curvature implies a uniform upper bound for the diameter.

COROLLARY 5.52. *For any solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow on a closed surface with $R(\cdot, 0) > 0$, there exists a constant $C > 0$ depending only on g_0 such that*

$$\operatorname{diam}\left(\mathcal{M}^2, g(t)\right) \leq C.$$

PROOF. Suppose there are N points $p_1, \ldots, p_N \in \mathcal{M}^2$ such that

$$\operatorname{dist}_{g(t)}(p_i, p_j) \geq \frac{2\pi}{\sqrt{R_{\max}(t)}}$$

for all $1 \leq i \neq j \leq N$. It follows from Klingenberg's Lemma that

$$\operatorname{inj}\left(\mathcal{M}^2, g(t)\right) \geq \frac{\pi}{\sqrt{R_{\max}(t)/2}},$$

hence that the balls $B_{g(t)}\left(p_i, \pi/\sqrt{R_{\max}(t)/2}\right)$ are embedded and pairwise disjoint. The area comparison estimate then implies that there exists $\varepsilon > 0$ such that

$$\operatorname{Area}\left(\mathcal{M}^2, g(t)\right) \geq \sum_{i=1}^{N} \operatorname{Area} B_{g(t)}\left(p_i, \frac{\pi}{\sqrt{R_{\max}(t)/2}}\right) \geq N \frac{\varepsilon}{R_{\max}(t)}.$$

Hence by Proposition 5.51, there exists $C > 0$ such that

$$N \leq \frac{R_{\max}(t)}{\varepsilon} \operatorname{Area}\left(\mathcal{M}^2, g(t)\right) \leq C \cdot \operatorname{Area}\left(\mathcal{M}^2, g_0\right).$$

□

10. Differential Harnack estimates of LYH type

In the context of parabolic differential equations, a classical Harnack inequality is an *a priori* lower bound for a positive solution of a parabolic equation at some point and time in terms of that solution at another point and an earlier time. In their important paper [**92**], Li and Yau obtained new space-time gradient estimates for solutions of parabolic equations. Their estimates are differential inequalities which substantially generalize classical Harnack estimates. Moreover, their method relies primarily on the maximum principle, hence is robust enough to work in the context of Riemannian geometry. Because of later advances made by Hamilton in the ideas Li and Yau pioneered, such estimates are today called **differential Harnack estimates of LYH type**.

We plan to discuss differential Harnack estimates for the Ricci flow in n dimensions in a chapter of the successor to this volume. Here, we derive certain differential Harnack inequalities that let us estimate the scalar curvature function on a surface evolving by the normalized Ricci flow.

10.1. The case that $R(\cdot, 0) > 0$. Recall that the gradient Ricci soliton equation (5.7) implies that

$$\nabla R + R\nabla f = 0. \tag{5.34}$$

Let $L \doteqdot \log R$ and define

$$Q \doteqdot \Delta L + R - r. \tag{5.35}$$

Q is known as the **differential Harnack quantity** for a surface of positive curvature. On a gradient soliton of positive curvature, we can divide equation (5.34) by R and take the divergence to find that $Q \equiv 0$. As was mentioned in Section 3, quantities which are constant in space on soliton solutions often yield useful estimates on general solutions. For such solutions, we shall obtain a lower bound for Q depending only on the initial metric g_0 by applying the maximum principle to its evolution equation.

LEMMA 5.53. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a compact surface with strictly positive scalar curvature. Then*

$$\frac{\partial}{\partial t} L = \Delta L + |\nabla L|^2 + R - r. \tag{5.36}$$

PROOF. By (5.3), we have

$$\frac{\partial}{\partial t} L = \frac{1}{R}\frac{\partial}{\partial t} R = \frac{1}{R}\left(\Delta R + R(R-r)\right) = \Delta L + |\nabla L|^2 + R - r.$$

□

COROLLARY 5.54. *The differential Harnack quantity Q satisfies the identity*

$$Q = \frac{\partial}{\partial t} L - |\nabla L|^2 .$$

LEMMA 5.55. *On any solution of the normalized Ricci flow on a compact surface with strictly positive scalar curvature, one has*

$$\frac{\partial}{\partial t} Q = \Delta Q + 2\langle \nabla Q, \nabla L\rangle + 2\left|\nabla\nabla L + \frac{1}{2}(R-r)g\right|^2 + rQ. \tag{5.37}$$

PROOF. Using (5.5) with (5.36) and recalling the Bochner identity, we calculate

$$\begin{aligned}
\frac{\partial}{\partial t} Q &= (R-r)\,\Delta L + \Delta\left(\frac{\partial}{\partial t} L\right) + R\left(\frac{\partial}{\partial t} L\right) \\
&= \Delta Q + \Delta|\nabla L|^2 + (R-r)\,\Delta L + R\left(\Delta L + |\nabla L|^2 + R - r\right) \\
&= \Delta Q + 2\left(\langle \nabla \Delta L, \nabla L\rangle + |\nabla\nabla L|^2\right) \\
&\quad + 2R|\nabla L|^2 + (2R-r)\,\Delta L + R\,(R-r) \\
&= \Delta Q + 2\langle \nabla Q, \nabla L\rangle + 2|\nabla\nabla L|^2 + 2(R-r)\,\Delta L + (R-r)^2 + rQ.
\end{aligned}$$

The result follows. □

COROLLARY 5.56. *Q satisfies the evolutionary inequality*

$$\frac{\partial}{\partial t}Q \geq \Delta Q + 2\langle \nabla Q, \nabla L\rangle + Q^2 + rQ. \tag{5.38}$$

PROOF. Use (5.37) and the standard fact that $n|S|^2 \geq (\operatorname{tr}_g S)^2$ for any symmetric 2-tensor S on an n-dimensional Riemannian manifold. □

To obtain an estimate for Q from the maximum principle, we need to consider the ODE corresponding to (5.38), namely

$$\frac{d}{dt}q = q^2 + rq.$$

The solution with initial data $q(0) = q_o < -r < 0$ is

$$q(t) = -\frac{rq_o e^{rt}}{q_o e^{rt} - q_o - r} = -\frac{Cre^{rt}}{Ce^{rt} - 1},$$

where $C \doteqdot q_o/(q_o + r) > 1$. Therefore, applying the maximum principle to equation (5.38) implies the following.

PROPOSITION 5.57. *On a complete solution of the normalized Ricci flow with bounded positive scalar curvature, there exists a constant $C > 1$ depending only on g_0 such that*

$$\frac{\partial}{\partial t}\log R - |\nabla \log R|^2 = Q \geq -\frac{Cre^{rt}}{Ce^{rt} - 1}.$$

This estimate for Q is known as a **differential Harnack inequality.** Integrating it along paths in space-time yields a classical Harnack inequality: a lower bound for the curvature at some point and time (x_2, t_2) in terms of the curvature at some (x_1, t_1), where $0 \leq t_1 < t_2$. In particular, let $\gamma : [t_1, t_2] \to \mathcal{M}^2$ be a C^1-path joining x_1 and x_2. Then by the fundamental theorem of calculus,

$$\begin{aligned}
\log\frac{R(x_2,t_2)}{R(x_1,t_1)} &= \int_{t_1}^{t_2} \frac{d}{dt}(\log R)(\gamma(t),t)\,dt \\
&= \int_{t_1}^{t_2} \left(\frac{\partial}{\partial t}(\log R)(\gamma(t),t) + \left\langle \nabla \log R, \frac{d\gamma}{dt}\right\rangle\right) dt \\
&\geq \int_{t_1}^{t_2} \left(|\nabla \log R|^2 - \frac{Cre^{rt}}{Ce^{rt}-1} - |\nabla \log R| \cdot \left|\frac{d\gamma}{dt}\right|\right) dt \\
&\geq \int_{t_1}^{t_2} \left(-\frac{Cre^{rt}}{Ce^{rt}-1} - \frac{1}{4}\left|\frac{d\gamma}{dt}\right|^2\right) dt \\
&= -\frac{1}{4}\int_{t_1}^{t_2} \left|\frac{d\gamma}{dt}\right|^2 dt - \log\left(Ce^{rt_2} - 1\right) + \log\left(Ce^{rt_1} - 1\right).
\end{aligned} \tag{5.39}$$

Now define

$$A = A(x_1, t_1, x_2, t_2) \doteqdot \inf_{\gamma} \int_{t_1}^{t_2} \left|\frac{d\gamma}{dt}\right|^2 dt,$$

where the infimum is taken over all C^1-paths $\gamma : [t_1, t_2] \to \mathcal{M}^2$ joining x_1 and x_2. Then exponentiation of equation (5.39) yields a classical Harnack inequality.

PROPOSITION 5.58. *Let $(\mathcal{M}^2, g(t))$ be a complete solution of the normalized Ricci flow with bounded positive scalar curvature.* *Then there exist constants $C_1 > 1$ and $C > 0$ depending only on g_0 such that for all $x_1, x_2 \in \mathcal{M}^2$ and $0 \leq t_1 < t_2$,*

$$\frac{R(x_2, t_2)}{R(x_1, t_1)} \geq e^{-\frac{A}{4}} \frac{C_1 e^{rt_1} - 1}{C_1 e^{rt_2} - 1} \geq e^{-\frac{A}{4} - C(t_2 - t_1)}. \tag{5.40}$$

10.2. The case that $R(\cdot, 0)$ changes sign. We now extend the differential Harnack inequality to solutions of the normalized Ricci flow whose curvature changes sign. Assume that $r > 0$. As in Section 8.2, we consider the solution

$$s(t) = \frac{r}{1 - \left(1 - \frac{r}{s_0}\right) e^{rt}}$$

of the ODE

$$\frac{d}{dt} s = s(s - r)$$

with initial condition

$$s(0) = s_0 < R_{\min}(0) < 0.$$

We define

$$\hat{L} = \hat{L}(g, s) \doteqdot \log(R - s).$$

A computation similar to that in Lemma 5.53 shows that $\hat{L}$ evolves by

$$\frac{\partial}{\partial t} \hat{L} = \Delta \hat{L} + \left| \nabla \hat{L} \right|^2 + R - r + s. \tag{5.41}$$

Then we define

$$\hat{Q} = \hat{Q}(g, s) \doteqdot \Delta \hat{L} + R - r = \frac{\partial}{\partial t} \hat{L} - \left| \nabla \hat{L} \right|^2 - s.$$

LEMMA 5.59. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a compact surface with any initial metric such that $r > 0$. Then*

$$\begin{aligned} \frac{\partial}{\partial t} \hat{Q} &= \Delta \hat{Q} + 2 \left\langle \nabla \hat{Q}, \nabla \hat{L} \right\rangle + 2 \left| \nabla \nabla \hat{L} + \frac{1}{2}(R - r) g \right|^2 \\ &\quad + s \left| \nabla \hat{L} \right|^2 + (r - s) \hat{Q} + s(R - r). \end{aligned} \tag{5.42}$$

PROOF. As in Lemma 5.55, we compute

$$\begin{aligned}\frac{\partial}{\partial t}\hat{Q} &= (R-r)\,\Delta\hat{L} + \Delta\left(\Delta\hat{L} + \left|\nabla\hat{L}\right|^2 + R - r + s\right) \\ &\quad + (R-s)\left(\Delta\hat{L} + \left|\nabla\hat{L}\right|^2 + R - r + s\right) + s\,(s-r) \\ &= \Delta\hat{Q} + 2\left\langle\nabla\hat{Q},\nabla\hat{L}\right\rangle + 2\left|\nabla\nabla\hat{L}\right|^2 + 2\,(R-r)\,\Delta\hat{L} + (R-r)^2 \\ &\quad + s\left|\nabla\hat{L}\right|^2 + (r-s)\left(\Delta\hat{L} + R - r\right) + s\,(R-r).\end{aligned}$$

□

Our aim is to obtain a lower bound for $\hat{Q}$. This is accomplished in the following estimate.

PROPOSITION 5.60. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a closed surface. Let g_0 be any initial metric such that $r > 0$. Then there exists a constant $C > 0$ depending only on g_0 such that*

$$Q \geq -C.$$

PROOF. By (5.28) and Proposition 5.18, there is $C > 0$ depending only on g_0 such that $-Ce^{-rt} \leq s \leq 0$ and $R \leq Ce^{rt}$. Hence the only possibly troublesome negative term on the right-hand side of (5.42) is $s|\nabla\hat{L}|^2$, over which we have no control. However, we can avoid this bad term by considering the quantity

$$P \doteqdot \hat{Q} + s\hat{L}.$$

Indeed, it follows from (5.4) and (5.41) that

$$\frac{\partial}{\partial t}\left(s\hat{L}\right) = \Delta\left(s\hat{L}\right) + s\left|\nabla\hat{L}\right|^2 + s\,(R-r+s) + (s-r)\left(s\hat{L}\right).$$

Rewriting the gradient term and using the fact that $\hat{L} = \log(R-s) \geq -c\,(1+t)$ for some $c > 0$, we see that there is $C_1 > 0$ such that

$$\frac{\partial}{\partial t}\left(s\hat{L}\right) \geq \Delta\left(s\hat{L}\right) + 2\langle\nabla\left(s\hat{L}\right),\nabla\hat{L}\rangle - s\left|\nabla\hat{L}\right|^2 - C_1.$$

By adding the evolution equations for $\hat{Q}$ and $s\hat{L}$ and recalling the standard fact that $n\,|S|^2 \geq (\operatorname{tr}_g S)^2$ for any symmetric 2-tensor S, we obtain

$$\frac{\partial}{\partial t}P \geq \Delta P + 2\left\langle\nabla P,\nabla\hat{L}\right\rangle + \hat{Q}^2 + (r-s)\hat{Q} - C_2$$

for some $C_2 > 0$. Since $s\hat{L} = P - \hat{Q}$ is bounded, it follows that there exists a constant $C > 0$ such that

$$\frac{\partial}{\partial t}P \geq \Delta P + 2\left\langle\nabla P,\nabla\hat{L}\right\rangle + \frac{1}{2}\left(P^2 - C^2\right).$$

Applying the maximum principle, we get

$$\hat{Q} + s\hat{L} = P \geq \min\left\{\min_{x\in\mathcal{M}^2} P(x,0), -C\right\},$$

and hence reach the conclusion. □

This is our differential Harnack inequality in the case that the curvature changes sign. As in the case of strictly positive curvature, we can integrate to obtain a classical Harnack inequality.

PROPOSITION 5.61. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a compact surface. Let g_0 be any initial metric such that $r > 0$. Then there exists a constant $C > 0$ depending only on g_0 such that for any $x_1, x_2 \in \mathcal{M}^2$ and $0 \leq t_1 < t_2$,*

$$\frac{R(x_2, t_2) - s(t_2)}{R(x_1, t_1) - s(t_1)} \geq e^{-\frac{A}{4} - C(t_2 - t_1)}, \tag{5.43}$$

where $A \doteqdot \inf_\gamma \int_{t_1}^{t_2} \left|\frac{d\gamma}{dt}\right|^2 dt$.

11. Convergence when $R(\cdot, 0) > 0$

We are now ready to prove Theorem 5.1 in the case the $R(\cdot, 0) > 0$. The first step is to apply the LYH differential Harnack inequality to obtain a uniform positive lower bound for R.

PROPOSITION 5.62. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a compact surface such that $R[g_0] > 0$. Then there exists a constant $c > 0$ depending only on g_0 such that*

$$R(x, t) \geq c > 0$$

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

PROOF. Observe that the differential Harnack inequality implies that for any points $x_1, x_2 \in \mathcal{M}^2$ and times $0 \leq t_1 < t_2$, one has

$$\frac{R(x_2, t_2)}{R(x_1, t_1)} \geq e^{-\frac{A}{4} - C(t_2 - t_1)},$$

where

$$A = A(x_1, t_1, x_2, t_2) \doteqdot \inf_\gamma \int_{t_1}^{t_2} \left|\frac{d\gamma}{dt}\right|^2 dt.$$

Notice that $R \geq ce^{-r}$ for times $0 \leq t \leq 1$, and consider any point and time (x, t) with $t \geq 1$. Choose any $x_1 \in \mathcal{M}^2$ such that $r \leq R(x_1, t-1) \leq R_{\max}(t-1)$. Then we have

$$R(x, t) \geq e^{-\frac{A}{4} - C} R(x_1, t-1) \geq re^{-C} \cdot e^{-A/4}.$$

Thus to obtain a uniform lower bound for R, it will suffice to obtain a uniform upper bound for $A(x_1, t-1, x, t)$.

It follows from Proposition 5.51 and formula (5.31) that

$$e^{-C} g(t-1) \leq g(\tau) \leq e^{r} g(t-1).$$

for $\tau \in [t-1, t]$. Let γ be a geodesic joining x and x_1 of constant speed with respect to $g(t)$, so that

$$\left|\gamma'(\tau)\right|_{g(t)} = \operatorname{dist}_{g(t)}(x, x_1).$$

Then we have

$$\begin{aligned} A &\leq \int_{t-1}^{t} \left|\gamma'(\tau)\right|_{g(\tau)}^{2} d\tau \\ &\leq e^{r+C} \int_{t-1}^{t} \left|\gamma'(\tau)\right|_{g(t)}^{2} d\tau = e^{r+C} \left[\operatorname{dist}_{g(t)}(x, x_1)\right]^2 . \end{aligned}$$

Hence the desired uniform bound for A is implied by the diameter bound of Corollary 5.52. □

Applying Corollary 5.35, we conclude that there are constants $0 < c \leq C < \infty$ such that

$$|M| \leq Ce^{-ct}.$$

Then arguing as we did for the curvature in Sections 5 and 6, it is straightforward to obtain estimates for all derivatives of M on a solution of the modified Ricci flow (5.23).

COROLLARY 5.63. *If $(\mathcal{M}^2, g(t))$ is a solution of the modified Ricci flow (5.23) on a surface $\mathcal{M}^2$ of positive Euler characteristic, there exist for every $k \in \mathbb{N}$ constants c_k and C_k depending only on g_0 such that*

$$\left|\nabla^k M\right| \leq C_k e^{-c_k t}.$$

Following the argument detailed in Section 7, we thus obtain the following result.

THEOREM 5.64. *Let $(\mathcal{M}^2, g_0)$ be a closed Riemannian surface with average scalar curvature $r > 0$. If $R(g_0) > 0$, then the unique solution $g(t)$ of the normalized Ricci flow with $g(0) = g_0$ converges exponentially in any C^k-norm to a smooth constant-curvature metric g_∞ as $t \to \infty$.*

12. A lower bound for the injectivity radius

In the next section, we shall derive uniform upper and lower bounds for the scalar curvature of a solution $(\mathcal{M}^2, g(t))$ to the normalized Ricci flow on a closed Riemannian surface $\mathcal{M}^2$ of positive Euler characteristic whose curvature is initially of mixed sign. In order to do this, we will require a lower bound for the injectivity radius of such a solution. This section is devoted to obtaining a suitable estimate.

PROPOSITION 5.65. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a closed surface with Euler characteristic $\chi(\mathcal{M}^2) > 0$, and let $K_{\max}(t)$ denote its maximum Gaussian curvature. Then for all $0 \le t < \infty$,*

$$\operatorname{inj}\left(\mathcal{M}^2, g(t)\right) \ge \min\left\{\operatorname{inj}\left(\mathcal{M}^2, g_0\right), \min_{\tau\in[0,t]} \frac{\pi}{\sqrt{K_{\max}(\tau)}}\right\}.$$

We begin with a collection of standard technical tools that are valid in any dimension. We denote the open ball of radius ρ centered $\xi \in T_x\mathcal{M}^n$ by $B_x(\xi, \rho)$ and the length of a path γ by $L(\gamma)$.

LEMMA 5.66. *Let $(\mathcal{M}^n, g)$ be a complete Riemannian manifold with sectional curvatures bounded above by $K_{\max} > 0$. Let β be a geodesic path joining points p and q in $\mathcal{M}^n$ with $L(\beta) < \pi/\sqrt{K_{\max}}$. Then for any points p^* and q^* sufficiently near p and q respectively, there is a unique geodesic path β^* joining p^* and q^* which is close to β.*

PROOF. By the Rauch comparison theorem (Theorem B.20 of Appendix B), the map

$$\exp_p\big|\, B_p\left(0, \pi/\sqrt{K_{\max}}\right) : B_p\left(0, \pi/\sqrt{K_{\max}}\right) \to \mathcal{M}^n$$

is a local diffeomorphism. We may assume that the geodesic $\beta : [0,1] \to \mathcal{M}^n$ is parameterized so that $\beta(s) = \exp_p(sV)$ for some $V \in B_p\left(0, \pi/\sqrt{K_{\max}}\right)$ with $\exp_p(V) = q$. Since $\exp_p\big|\, B_p\left(0, \pi/\sqrt{K_{\max}}\right)$ is a local diffeomorphism, there is for each p^* and q^* sufficiently close to p and q respectively a unique vector $V^* \in T_{p^*}\mathcal{M}^n$ close to $V \in T_p\mathcal{M}^n$ such that $\exp_{p^*}(V^*) = q^*$. (Here we have identified $T_p\mathcal{M}^n \cong \mathbb{R}^n \cong T_{p^*}\mathcal{M}^n$.) The path $\beta^* : [0,1] \to \mathcal{M}^n$ defined by $\beta^*(s) \doteqdot \exp_{p^*}(sV^*)$ is the unique geodesic near β joining p^* and q^*. $\square$

We now anticipate Definitions B.66 and B.67 (found in Appendix B): if $k \in \mathbb{N}$, a **proper geodesic k-gon** is a collection

$$\Gamma = \{\gamma_i : [0, \ell_i] \to \mathcal{M}^n : i = 1, \ldots, k\}$$

of unit-speed geodesic paths between k pairwise-distinct vertices $p_i \in \mathcal{M}^n$ such that $p_i = \gamma_i(0) = \gamma_{i-1}(\ell_{i-1})$ for each i, where all indices are interpreted modulo k. We say Γ is a **nondegenerate proper geodesic k-gon** if $\measuredangle_{p_i}(-\dot\gamma_{i-1}, \dot\gamma_i) \neq 0$ for each $i = 1, \ldots, k$; if $k = 1$, we interpret this to mean $L(\Gamma) > 0$. A **(nondegenerate) geodesic k-gon** is a (nondegenerate) proper geodesic j-gon for some $j = 1, \ldots, k$.

LEMMA 5.67. *Let $(\mathcal{M}^n, g)$ be a closed Riemannian manifold with sectional curvatures bounded above by $K_{\max} > 0$. Let β be a nondegenerate geodesic 2-gon between $p \neq q \in \mathcal{M}^n$ of edge lengths $L(\beta_i) < \pi/\sqrt{K_{\max}}$ for $i = 1, 2$. Then there is a smooth nontrivial geodesic loop γ on $\mathcal{M}^n$ with*

$$L(\gamma) \le L(\beta_1) + L(\beta_2).$$

PROOF. Choose $d < \pi/\sqrt{K_{\max}}$ so that $\max\{L(\beta_i)\} < d$. Let Γ be the space of all nondegenerate geodesic 2-gons γ in $\mathcal{M}^n$ with $L(\gamma_i) < d$ for $i = 1, 2$. Note that $\Gamma \ni \beta$ is a nonempty open $2n$-manifold locally parameterized by points in $\mathcal{M}^n \times \mathcal{M}^n$. Indeed, by Lemma 5.66 there is for every pair of points p^* and q^* sufficiently close to p and q a unique nondegenerate geodesic 2-gon β^* with $L(\beta_i^*) < d$ for $i = 1, 2$. Define

$$m \doteqdot \inf_{\gamma \in \Gamma} \{L(\gamma_1) + L(\gamma_2)\}.$$

Note that $m \geq \operatorname{inj} \mathcal{M}^n > 0$, since the map $\exp_x | B_x(0, \rho) : B_x(0, \rho) \to \mathcal{M}^n$ is an embedding at any $x \in \mathcal{M}^n$ whenever $\rho < \operatorname{inj} \mathcal{M}^n$. Let $\beta(i) \in \Gamma$ be a sequence of nondegenerate geodesic 2-gons with

$$\lim_{i \to \infty} \{L(\beta(i)_1) + L(\beta(i)_2)\} = m.$$

Let $p(i) \neq q(i)$ denote the endpoints of the geodesic paths $\beta(i)_1$ and $\beta(i)_2$. By compactness of $\mathcal{M}^n$, we may pass to a subsequence such that the limits $p_\infty \doteqdot \lim_{j \to \infty} p(j)$ and $q_\infty \doteqdot \lim_{j \to \infty} q(j)$ exist and satisfy $\operatorname{dist}(p_\infty, q_\infty) \leq d < \pi/\sqrt{K_{\max}}$. Hence geodesics

$$\beta(\infty)_1 \doteqdot \lim_{j \to \infty} \beta(j)_1 \qquad \text{and} \qquad \beta(\infty)_2 \doteqdot \lim_{j \to \infty} \beta(j)_2$$

exist and satisfy $L(\beta(\infty)_1) + L(\beta(\infty)_2) = m$ by continuity.

There are now two cases. If $p_\infty = q_\infty$, then some path, say $\beta(\infty)_1$, is a nondegenerate geodesic 1-gon at p_∞ satisfying $L(\beta(\infty)_1) = m$. The loop $\beta(\infty)_1$ must be smooth at p_∞, or else the first variation formula (see 1.3 of [**27**]) would let us shorten $\beta(\infty)_1$ inside Γ.

In the second case, $p_\infty \neq q_\infty$. If $\beta(\infty) \doteqdot \{\beta(\infty)_1, \beta(\infty)_2\}$ is degenerate, then $\beta(\infty)_2$ must be the same path as $\beta(\infty)_1$ but with the opposite orientation. By the uniqueness statement in Lemma 5.66, this is possible only if $\beta(i)_1 = \beta(i)_2$ for sufficiently large i, which contradicts the nondegeneracy of $\beta(i)$. Hence $\beta(\infty)$ is nondegenerate. As above, $\beta(\infty)$ must be smooth, or else the first variation formula would let us shorten $\beta(\infty)$ inside Γ. $\square$

DEFINITION 5.68. We say a geodesic loop γ is **stable** if every nearby loop γ^* satisfies $L(\gamma^*) \geq L(\gamma)$. We say a geodesic loop γ is **weakly stable** if the second variation of arc length of γ is nonnegative.

Clearly, stable implies weakly stable. The following result is well known.

LEMMA 5.69. *Let $(\mathcal{M}^n, g)$ be a closed Riemannian manifold with sectional curvatures bounded above by $K_{\max} > 0$. If α is the shortest geodesic loop in $\mathcal{M}^n$ and $L(\alpha) < 2\pi/\sqrt{K_{\max}}$, then α is stable.*

PROOF. If not, there exists a loop α^* near α with $L(\alpha^*) < L(\alpha)$. Pick distinct points p and q on α that divide α into two geodesic paths α_1 and α_2 of equal length. It is always possible to choose distinct points p^* and q^* on α^* near p and q respectively that divide α^* into paths α_1^* and α_2^* of length less than $\pi/\sqrt{K_{\max}}$. Then since $L(\alpha_1) = L(\alpha_2) < \pi/\sqrt{K_{\max}}$, there are by Lemma 5.66 unique geodesic paths β_1^* and β_2^* between p^* and q^* that are close to α_1 and α_2 respectively, hence close to α_1^* and α_2^*. Now as a consequence of the Gauss Lemma (Lemma B.1), we have $L(\beta_i^*) \leq L(\alpha_i^*)$ for $i = 1, 2$. But then by Lemma 5.67, there is a smooth geodesic loop γ with

$$L(\gamma) \leq L(\beta_1^*) + L(\beta_2^*) \leq L(\alpha^*) < L(\alpha),$$

which is a contradiction. □

LEMMA 5.70. *If γ is a weakly stable geodesic loop in an orientable Riemannian surface $(\mathcal{M}^2, g)$, then*

$$\int_\gamma R\, ds \leq 0.$$

PROOF. Let T denote the unit tangent vector to γ. Since $\mathcal{M}^2$ is orientable, there is a well-defined unit normal N along γ. Since γ is geodesic and $n = 2$, we have $\nabla_T N = 0$, because

$$\langle \nabla_T N, T \rangle = T \langle N, T \rangle = 0 \qquad \text{and} \qquad \langle \nabla_T N, N \rangle = \frac{1}{2} T \langle N, N \rangle = 0.$$

Let γ_θ be a 1-parameter family of loops with $\gamma_0 = \gamma$ and $\partial \gamma_\theta / \partial \theta|_{\theta=0} = N$. Then by the second variation formula (see 1.14 of [**27**])

$$\left. \frac{\partial^2}{\partial \theta^2} L(\gamma_\theta) \right|_{\theta=0} = \int_\gamma \langle R(N,T) N, T \rangle\, ds = -\int_\gamma \frac{R}{2}\, ds.$$

Since γ is weakly stable, the result follows. □

Now let γ_t be a smooth 1-parameter family of loops in a time-dependent Riemannian manifold $(\mathcal{M}^n, g(t))$. That is, each γ_t is a smooth loop in $\mathcal{M}^n$ depending smoothly on t in some open interval $\mathcal{I}$. If $L_t(\gamma)$ denotes the length of a loop γ with respect to the metric $g(t)$, then by Lemma 3.11, we have

$$\frac{d}{dt} L_t(\gamma_t) = \frac{1}{2} \int_{\gamma_t} \frac{\partial g}{\partial t}(T,T)\, ds - \int_{\gamma_t} \langle \nabla_T T, V \rangle\, ds, \tag{5.44}$$

where T is the unit tangent to γ_t, V is the variation vector field $\gamma_*(\partial/\partial t)$, and ds is the element of arc length. This implies the following observation.

LEMMA 5.71. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow and γ_t is a smooth 1-parameter family of geodesic loops, then*

$$\left. \frac{d}{dt} L_t(\gamma_t) \right|_{t=\tau} = \frac{1}{2} \int_{\gamma_\tau} (r - R)\, ds.$$

Combining this result with Lemma 5.70 yields the following result.

COROLLARY 5.72. *Let γ_t be a smooth 1-parameter family of loops in a solution $(\mathcal{M}^2, g(t))$ of the normalized Ricci flow. If γ_τ is a weakly stable geodesic loop, then*

$$\frac{d}{dt} L_t(\gamma_t)\bigg|_{t=\tau} \geq \frac{r}{2} L_\tau(\gamma_\tau).$$

We are now prepared to show that the length of the shortest closed geodesic is strictly increasing whenever it is small enough.

LEMMA 5.73. *Let $(\mathcal{M}^2, g(t))$ be a solution of the normalized Ricci flow on a compact surface of positive Euler characteristic, and let $K_{\max}(t) > 0$ denote its maximum Gaussian curvature. Suppose γ_τ is the shortest closed $g(\tau)$-geodesic at some time $\tau > 0$. If $L_\tau(\gamma_\tau) < 2\pi/\sqrt{K_{\max}(\tau)}$, then there is $\varepsilon > 0$ small enough so that for all $t \in (\tau - \varepsilon, \tau)$, there is a $g(t)$-geodesic γ_t such that*

$$L_t(\gamma_t) < L_\tau(\gamma_\tau).$$

Note, however, that we do not claim the geodesics γ_t depend smoothly on $t \in (\tau - \varepsilon, \tau)$.

PROOF. By Corollary 5.72, there is $\varepsilon_0 > 0$ sufficiently small so that $L_t(\gamma_\tau) < L_\tau(\gamma_\tau)$ for $t \in (\tau - \varepsilon_0, \tau)$. Choose points p and q on γ_τ that divide γ_τ into two segments γ_τ^1 and γ_τ^2 of equal length with respect to the metric $g(t)$. These may not be $g(t)$-geodesic segments, but for some $\varepsilon \in (0, \varepsilon_0)$, we have $L_t(\gamma_\tau^i) < \pi/\sqrt{K_{\max}(t)}$ for $i = 1, 2$ and all $t \in (\tau - \varepsilon, \tau)$. Thus by Lemma 5.66, there exist unique $g(t)$-geodesics β_1 and β_2 between p and q which are near γ_τ^1 and γ_τ^2 respectively. Then by the Gauss lemma, we have

$$L_t(\beta_1) + L_t(\beta_2) \leq L_t(\gamma_\tau^1) + L_t(\gamma_\tau^2) < L_\tau(\gamma_\tau).$$

Hence by Lemma 5.67, there is a smooth $g(t)$-geodesic loop γ with

$$L_t(\gamma) \leq L_t(\beta_1) + L_t(\beta_2) < L_\tau(\gamma_\tau).$$

□

We are now ready to prove the main result of this section.

PROOF OF PROPOSITION 5.65. Lemma B.34 in Appendix B (Klingenberg's Lemma) says that inj $(\mathcal{M}^2, g(t))$ is no smaller than the minimum of $\pi/\sqrt{K_{\max}(t)}$ and half the length of the shortest closed $g(t)$-geodesic. But by Lemma 5.73, the length of the shortest closed geodesic is strictly increasing as long as it is less than $2\pi/\sqrt{K_{\max}(t)}$. □

13. The case that $R(\cdot, 0)$ changes sign

In this section, we complete our proof of Theorem 5.1 for the case of a compact Riemannian surface $(\mathcal{M}^2, g_0)$ such that $\chi(\mathcal{M}^2) > 0$ and $R_{\min}(g_0) < 0$. The road map for obtaining convergence in this case is as follows. Armed with the entropy estimate of Proposition 5.44 and the injectivity radius estimate of Proposition 5.65, we proceed as in Proposition 5.51 to obtain a uniform upper bound for R. As in Corollary 5.52, this implies an upper bound for the diameter. Then we use the differential Harnack estimate of Proposition 5.61 to argue that R eventually becomes strictly positive within finite time. Once this happens, the proof of Theorem 5.64 goes through exactly as written.

Our first result is an analog of Proposition 5.51.

LEMMA 5.74. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow on a closed surface of positive Euler characteristic, then there exists a constant $C > 0$ depending only on g_0 such that*

$$\sup_{\mathcal{M}^2 \times [0,\infty)} R \leq C.$$

PROOF. By Proposition 5.18, there exists $C_1 > 0$ depending only on g_0 such that $\sup_{\mathcal{M}^2 \times [0,1]} R \leq r + C_1 e^r$. Given $T \geq 1$, define

$$\kappa(T) \doteqdot \max_{\mathcal{M}^2 \times [0,T]} R.$$

We intend to prove that κ is bounded independently of T. We may assume that $\kappa(T) > \max\{\kappa_1, 1/4\}$, so that $T > 1$. Choose $(x_1, t_1) \in \mathcal{M}^2 \times (1, T]$ such that $R(x_1, t_1) = \max_{\mathcal{M}^2 \times [0,T]} R = \kappa(T)$. Following the proof of Proposition 5.51, one shows that

$$R(y, t_1) \geq \frac{\kappa}{2} \tag{5.45}$$

for all $y \in B_{g(t_1)}(x_1, 1/\sqrt{64\kappa})$. Define $\rho \doteqdot \operatorname{inj}(\mathcal{M}^2, g_0)$. Then by Proposition 5.65, one has

$$\operatorname{inj}(\mathcal{M}^2, g(t)) \geq \min\left\{\rho, \frac{\pi}{\sqrt{\kappa}}\right\}.$$

As long as $\rho \leq \pi/\sqrt{\kappa}$, we have the uniform upper bound $\kappa \leq (\pi/\rho)^2$. So we may assume $\rho > \pi/\sqrt{\kappa}$. By Proposition 5.44, there exists $C_2 > 0$ such that

$$\begin{aligned} C_2 \geq \hat{N}(g(t_1)) &\doteqdot \int_{\mathcal{M}^2} (R - s)\log(R - s)\, dA \\ &\geq \int_{B_{g(t_1)}\left(x_1, \frac{1}{\sqrt{64\kappa}}\right)} (R - s)\log(R - s)\, dA \\ &\quad - \frac{1}{e}\operatorname{Area}(\mathcal{M}^2, g(t_1)), \end{aligned}$$

where $s = s(t_1) \leq 0$ is given by formula (5.28). By the area comparison theorem (§3.4 of [**25**]) and estimate (5.45), there exists $\varepsilon > 0$ such that

$$\begin{aligned} \int_{B_{g(t_1)}\left(x_1, \frac{1}{\sqrt{64\kappa}}\right)} (R - s)\log(R - s)\, dA &\geq \frac{\varepsilon}{\kappa}\left(\frac{\kappa}{2} - s\right)\log\left(\frac{\kappa}{2} - s\right) \\ &\geq \frac{\varepsilon}{2}\log\frac{\kappa}{2}. \end{aligned}$$

Since Area $(\mathcal{M}^2, g(t_1)) = $ Area $(\mathcal{M}^2, g_0)$, this implies the desired uniform upper bound. $\square$

We next bound the diameter of the solution as we did in Corollary 5.52.

COROLLARY 5.75. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow on a closed surface of positive Euler characteristic, then there exists a constant $C > 0$ depending only on g_0 such that*

$$\operatorname{diam}\left(\mathcal{M}^2, g(t)\right) \leq C.$$

PROOF. Suppose there exist points $p_1, \ldots, p_N \in \mathcal{M}^2$ such that

$$\operatorname{dist}_{g(t)}(p_i, p_j) \geq \frac{2\pi}{\sqrt{R_{\max}(t)}}$$

for all $1 \leq i \neq j \leq N$. Define

$$\delta(t) \doteqdot \min\left\{\operatorname{inj}\left(\mathcal{M}^2, g_0\right), \frac{\pi}{\sqrt{R_{\max}(t)/2}}\right\}.$$

By Proposition 5.65, the balls $B_{g(t)}(p_i, \delta(t))$ are embedded and pairwise disjoint. Arguing as we did in Corollary 5.52, one estimates that

$$N \leq \frac{\operatorname{Area}\left(\mathcal{M}^2, g_0\right)}{\delta(t)}.$$

But by Lemma 5.74, δ is bounded from below by a constant depending only on g_0. The result follows. $\square$

We can now show that R becomes strictly positive within finite time.

LEMMA 5.76. *If $(\mathcal{M}^2, g(t))$ is a solution of the normalized Ricci flow on a closed surface of positive Euler characteristic, then there exists $T < \infty$ such that*

$$\inf_{\mathcal{M}^2 \times [T,\infty)} R > 0.$$

PROOF. We begin by deriving a positive lower bound for $R - s$, where s is given by formula (5.28). Given any (x, t) with $t \geq 1$, chose $x_1 \in \mathcal{M}^2$ such that

$$0 < r \leq R(x_1, t-1) \leq R_{\max}(t-1).$$

Then the differential Harnack estimate (5.43) derived in Proposition 5.61 implies that

$$R(x,t) - s(t) \geq e^{-C}\left[R(x_1, t-1) - s(t-1)\right] e^{-A/4} \geq e^{-C} r e^{-A/4}, \tag{5.46}$$

where C is a constant depending only on the initial metric, and A is defined by

$$A(x_1, t-1, x, t) \doteqdot \inf_\gamma \int_{t-1}^t |\dot\gamma|^2 \, d\tau.$$

As in Proposition .5.62, the desired positive lower bound for $R-s$ will follow from (5.46) once we bound A from above. Recall the integral formula (5.31) for the metric, which implies that for any $\tau \geq t-1$, one can write

$$g(x,\tau) = g(x,t-1)\exp\int_{t-1}^{\tau}(r-R)\,d\hat{\tau}.$$

Using the consequence of Lemma 5.9 that $R \geq R_{\min}(0)$ and the uniform upper bound for R obtained in Lemma 5.74, it follows from this formula that there exist $c_1 < C_1$ independent of $x \in \mathcal{M}^2$ such that for all $\tau \in [t-1,t]$, one has

$$c_1 \cdot g(x,t-1) \leq g(x,\tau) \leq C_2 \cdot g(x,t-1).$$

As in the proof of Proposition .5.62, this estimate implies that there exists $C_2 > 0$ such that

$$A \leq C_2\left[\operatorname{dist}_{g(t)}(x,x_1)\right]^2,$$

whence the desired upper bound for A follows from the diameter bound of Lemma 5.75.

Now by substituting the uniform upper bound for A into estimate (5.46) and using the fact that $x \in \mathcal{M}^2$ is arbitrary, one concludes that there exists $\varepsilon > 0$ such that

$$R_{\min}(t) = \inf_{x\in\mathcal{M}^2} R(x,t) \geq \varepsilon + s(t).$$

Since $s(t) \nearrow 0$ as $t \to \infty$, this proves the lemma. □

Once the curvature becomes strictly positive, we may proceed exactly as in Section 11. This completes our proof of the following result.

THEOREM 5.77. *Let $(\mathcal{M}^2, g_0)$ be a closed Riemannian surface with average scalar curvature $r > 0$. The unique solution $g(t)$ of the normalized Ricci flow with $g(0) = g_0$ converges exponentially in any C^k-norm to a smooth constant-curvature metric g_∞ as $t \to \infty$.*

14. Monotonicity of the isoperimetric constant

In the remainder of this chapter, we illustrate an alternative strategy for proving convergence of the Ricci flow starting from an arbitrary initial metric on a surface of positive Euler characteristic. The basic idea is as follows. A solution of the unnormalized Ricci flow with $r > 0$ will become singular in finite time. If one can estimate the injectivity radius of such a solution on a scale appropriate to the curvature and if one can rule out slowly-forming singularities, one can then show that a sequence of parabolic dilations formed at the developing singularity will converge to a round sphere metric. This is called a *blow-up strategy,* because one 'blows up' the solution at the developing singularity. Blow-up strategies are extremely important in understanding singularity formation of the Ricci flow in higher dimensions.

REMARK 5.78. The result that a sequence of blow-ups $(\mathcal{M}^2, g_j(t))$ pulled back by diffeomorphisms converges smoothly to a shrinking round sphere metric $(\mathcal{S}^2, g_\infty(t))$ is weaker than the pointwise C^∞ convergence obtained above. Moreover, if one did not already know the Uniformization Theorem, one could not conclude that g_∞ is in the same conformal class as the original solution. An analogous problem occurs when studying the Kähler–Ricci flow in complex dimensions greater than one: one does not know in general whether or not a limit g_∞ is Kähler with respect to the same complex structure as the original solution.

The strategy of proving sequential convergence $g_j(t) \to g_\infty(t)$ uses several techniques that will be developed later in this volume. We provide the details in Section 15. In this section, we develop the background necessary to obtain a suitable injectivity radius estimate.

DEFINITION 5.79. Let $(\mathcal{M}^n, g)$ be a closed orientable Riemannian manifold. We say a smooth embedded closed (but possibly disconnected) hypersurface $\Sigma^{n-1} \subset \mathcal{M}^n$ **separates** $\mathcal{M}^n$ if $\mathcal{M}^n \backslash \Sigma^{n-1}$ has two connected components $\mathcal{M}_1^n$ and $\mathcal{M}_2^n$ such that

$$\partial \mathcal{M}_1^n = \partial \mathcal{M}_2^n = \Sigma^{n-1}.$$

DEFINITION 5.80. If Σ^{n-1} separates $\mathcal{M}^n$, the **isoperimetric ratio** of Σ^{n-1} in $(\mathcal{M}^n, g)$ is

$$C_S\left(\Sigma^{n-1}\right) \doteqdot \left(\operatorname{Area}\left(\Sigma^{n-1}\right)\right)^n \left(\frac{1}{\operatorname{Vol}\left(\mathcal{M}_1^n\right)} + \frac{1}{\operatorname{Vol}\left(\mathcal{M}_2^n\right)}\right)^{n-1}.$$

The **isoperimetric constant** of $(\mathcal{M}^n, g)$ is

$$C_S(\mathcal{M}^n, g) \doteqdot \inf_{\Sigma^{n-1}} C_S\left(\Sigma^{n-1}\right),$$

where the infimum is taken over all smooth embedded Σ^{n-1} that separate $\mathcal{M}^n$.

REMARK 5.81. $C_S\left(\Sigma^{n-1}\right)$ is invariant under homothetic rescaling of the metric $g \mapsto e^\lambda g$ and is equivalent to the standard isoperimetric ratio

$$C_I\left(\Sigma^{n-1}\right) \doteqdot \frac{\left(\operatorname{Area}\left(\Sigma^{n-1}\right)\right)^n}{\min\left\{\operatorname{Vol}\left(\mathcal{M}_1^n\right), \operatorname{Vol}\left(\mathcal{M}_2^n\right)\right\}^{n-1}}.$$

Indeed, one has

$$C_I\left(\Sigma^{n-1}\right) \leq C_S\left(\Sigma^{n-1}\right) \leq 2^{n-1} C_I\left(\Sigma^{n-1}\right).$$

Now we restrict ourselves to the case of a closed orientable Riemannian surface $(\mathcal{M}^2, g)$. If γ separates $\mathcal{M}^2$ into two open surfaces $\mathcal{M}_1^2$ and $\mathcal{M}_2^2$, we have

$$C_S(\gamma) = L_g(\gamma)^2 \left(\frac{1}{A_g(\mathcal{M}_1^2)} + \frac{1}{A_g(\mathcal{M}_2^2)} \right)$$

$$(5.47) \qquad = L_g(\gamma)^2 \frac{A_g(\mathcal{M}^2)}{A_g(\mathcal{M}_1^2) \cdot A_g(\mathcal{M}_2^2)},$$

where L_g denotes length and A_g denotes area measured with respect to g. Moreover, the infimum in Definition 5.80 is attained by a smooth embedded curve.

LEMMA 5.82. *If $(\mathcal{M}^2, g)$ is a closed orientable Riemannian surface, there exists a smooth embedded (possibly disconnected) closed curve γ such that*

$$C_S(\gamma) = C_S(\mathcal{M}^2, g).$$

For a proof, see Theorem 3.4 of [**69**]. Analogous results in general dimensions can be found in Theorems 5-5 and 8-6 of [**99**].

REMARK 5.83. The minimizing curve γ might be disconnected, as can easily be seen by considering a rotationally symmetric 2-dimensional torus with two very thin necks. In this case, γ is the union of two disjoint embedded loops. More generally, no connected embedded homotopically nontrivial curve will separate a torus.

Now we restrict ourselves further to the case that $(\mathcal{M}^2, g)$ is diffeomorphic to $\mathcal{S}^2$. Define a *loop* in $\mathcal{M}^2$ to be a connected closed curve. It follows from the Schoenflies theorem that any smooth embedded loop γ separates $\mathcal{M}^2$ into smooth discs $\mathcal{M}_1^2$ and $\mathcal{M}_2^2$. We denote the isoperimetric ratio of a smooth loop on a topological sphere by

$$C_H(\gamma) \doteqdot C_S(\gamma).$$

Then we define

$$C_H(\mathcal{M}^2, g) \doteqdot \inf_{\gamma} C_H(\gamma),$$

where the infimum is taken over all smooth embedded loops γ. Clearly,

$$C_S(\mathcal{M}^2, g) \le C_H(\mathcal{M}^2, g).$$

If $(\mathcal{S}^2, g_{\mathrm{can}})$ is a constant-curvature 2-sphere, then $C_H(\mathcal{S}^2, g_{\mathrm{can}}) = 4\pi$. Moreover, this is an upper bound.

LEMMA 5.84. *If $(\mathcal{M}^2, g)$ is a closed orientable Riemannian surface, then*

$$C_H(\mathcal{M}^2, g) \le 4\pi.$$

PROOF. Let $x \in \mathcal{M}^2$ be given and define

$$\gamma_\varepsilon \doteqdot \{ y \in \mathcal{M}^2 : \operatorname{dist}_g(x, y) = \varepsilon \}.$$

For $\varepsilon > 0$ small enough, γ_ε is a smooth embedded loop. Since g is Euclidean up to first order at x, we see from (5.47) that as $\varepsilon \searrow 0$,

$$C_H(\gamma_\varepsilon) = [2\pi(\varepsilon + o(\varepsilon))]^2 \frac{A_g(\mathcal{M}^2)}{\left[\pi(\varepsilon + o(\varepsilon))^2\right][A_g(\mathcal{M}^2) - O(\varepsilon^2)]} \to 4\pi.$$

□

We now sketch a proof of the following existence result; further details may be found in §3 of [**62**].

LEMMA 5.85. *If $(\mathcal{M}^2, g)$ is a Riemannian surface diffeomorphic to S^2 such that $C_H(\mathcal{M}^2, g) < 4\pi$, then there exists a smooth embedded loop β such that*

$$C_H(\beta) = C_H(\mathcal{M}^2, g).$$

Our proof of Lemma 5.85 will use the following algebraic fact.

LEMMA 5.86. *If L_1, L_2 and A_1, A_2, A_3 are any positive numbers, then*

$$\begin{aligned}&(L_1 + L_2)^2 \left(\frac{1}{A_1 + A_2} + \frac{1}{A_3}\right)\\ &> \min\left\{L_1^2\left(\frac{1}{A_1} + \frac{1}{A_2 + A_3}\right),\ L_2^2\left(\frac{1}{A_2} + \frac{1}{A_1 + A_3}\right)\right\}.\end{aligned}$$

PROOF. To obtain a contradiction, suppose the result is false. Then there exist positive numbers L_1, L_2 and A_1, A_2, A_3 such that

$$(L_1 + L_2)^2 \left(\frac{1}{A_1 + A_2} + \frac{1}{A_3}\right) \leq L_1^2\left(\frac{1}{A_1} + \frac{1}{A_2 + A_3}\right)$$

and

$$(L_1 + L_2)^2 \left(\frac{1}{A_1 + A_2} + \frac{1}{A_3}\right) \leq L_2^2\left(\frac{1}{A_2} + \frac{1}{A_1 + A_3}\right).$$

We rewrite these inequalities as

$$\frac{A_1(A_2 + A_3)}{A_3(A_1 + A_2)} \leq \frac{L_1^2}{(L_1 + L_2)^2} \tag{5.48}$$

and

$$\frac{A_2(A_1 + A_3)}{A_3(A_1 + A_2)} \leq \frac{L_2^2}{(L_1 + L_2)^2}. \tag{5.49}$$

Summing (5.48) and (5.49) yields

$$\frac{2A_1A_2}{A_3(A_1 + A_2)} \leq -\frac{2L_1L_2}{(L_1 + L_2)^2} < 0,$$

which is impossible. □

SKETCH OF PROOF OF LEMMA 5.85. Let L and A denote length and area, respectively, measured with respect to the metric g. Define $\eta > 0$ by

$$C_H\left(\mathcal{M}^2, g\right) = 4\pi - 2\eta.$$

There exists $c_0 > 0$ depending only on g and η such that $L(\gamma) \geq c_0$ whenever γ is a smooth embedded loop such that $C_H(\gamma) \leq 4\pi - \eta$. By (5.47), we have

$$L(\gamma)^2 \frac{A\left(\mathcal{M}^2\right)}{A\left(\mathcal{M}_1^2\right) \cdot A\left(\mathcal{M}_2^2\right)} = C_H(\gamma) < 4\pi,$$

whence it follows that

$$A\left(\mathcal{M}_a^2\right) > \frac{c_0^2}{4\pi}$$

for $a = 1, 2$. We also have

$$L(\gamma)^2 \leq 4\pi \frac{A\left(\mathcal{M}_1^2\right) \cdot A\left(\mathcal{M}_2^2\right)}{A\left(\mathcal{M}^2\right)} \leq 2\pi A\left(\mathcal{M}^2\right).$$

This gives us some control on the curves γ. Now let $\bar{\gamma}_i$ be a sequence of smooth embedded loops such that

$$\lim_{i\to\infty} C_H\left(\bar{\gamma}_i\right) = C_H\left(\mathcal{M}^2, g\right).$$

By applying the curve shortening flow (in particular, Grayson's theorem [**49**] that any embedded loop either converges to a geodesic loop or shrinks to a round point) one may show that there exist a new sequence γ_i of smooth embedded loops such that

$$\lim_{i\to\infty} C_H\left(\gamma_i\right) = C_H\left(\mathcal{M}^2, g\right)$$

and

$$\int_{\gamma_i} k_i^2\, ds \leq C,$$

where k_i denotes the Gaussian curvature of the curve γ_i, and the constant $C < \infty$ depends only on g and η. From this, one shows that the γ_i are locally uniformly bounded in $L^{1,2}$ and $C^{1,1/2}$. One then concludes that for any $p < 2$ and $\alpha < 1/2$, there exists a subsequence that converges in $L^{1,p}$ and $C^{1,\alpha}$ to an immersed curve γ_∞. Since γ_∞ minimizes length among all nearby curves bounding regions of fixed areas, the formula for first variation of arc length shows that γ_∞ has constant curvature. Hence γ_∞ is smooth.

We now show that γ_∞ is actually embedded. Since it is a smooth limit of smooth embedded curves, it can at worst have points of self tangency. If this is so, one can regard γ_∞ as the union of two curves γ_1 and γ_2 of positive lengths L_1 and L_2, respectively, bounding positive areas A_1 and A_2, respectively. Then the 'inside' of γ_∞ has area $A_1 + A_2$. Let A_3 be the area 'outside' of γ_∞. We then have

$$C_H(\gamma_\infty) = (L_1 + L_2)^2 \left(\frac{1}{A_3} + \frac{1}{A_1 + A_2} \right)$$
$$C_H(\gamma_1) = L_1^2 \left(\frac{1}{A_1} + \frac{1}{A_2 + A_3} \right)$$
$$C_H(\gamma_2) = L_2^2 \left(\frac{1}{A_2} + \frac{1}{A_1 + A_3} \right).$$

But Lemma 5.86 implies that

$$C_H(\gamma_\infty) > \min \{C_H(\gamma_1), C_H(\gamma_2)\},$$

which contradicts the minimality of γ_∞. □

It is interesting to note that $C_H(\gamma)$ may be represented in another way.

LEMMA 5.87. *Let $(\mathcal{M}^2, g)$ be diffeomorphic to $\mathcal{S}^2$. Let γ be a smooth loop in $(\mathcal{M}^2, g)$, and let $(\mathcal{S}^2, \bar{g})$ be a 2-sphere of constant curvature chosen so that*

$$A_{\bar{g}}(\mathcal{S}^2) = A_g(\mathcal{M}^2).$$

Let $\bar{\gamma}$ be a shortest loop (namely, a round circle) that separates $(\mathcal{S}^2, \bar{g})$ into two discs $\mathcal{D}_1^2$ and $\mathcal{D}_2^2$ with $A_{\bar{g}}(\mathcal{D}_1^2) = A_g(\mathcal{M}_1^2)$ and $A_{\bar{g}}(\mathcal{D}_2^2) = A_g(\mathcal{M}_2^2)$. Then

$$C_H(\gamma) = 4\pi \frac{L_g(\gamma)^2}{L_{\bar{g}}(\bar{\gamma})^2}.$$

PROOF. Identify $(\mathcal{S}^2, \bar{g})$ with the 2-sphere of radius $\rho \doteqdot \sqrt{A_g(\mathcal{M}^2)/4\pi}$ centered at $0 \in \mathbb{R}^3$. We may assume without loss of generality that

$$0 < A_1 \doteqdot A_g(\mathcal{M}_1^2) \le A_2 \doteqdot A_g(\mathcal{M}_2^2).$$

Choose $h \in [0, \rho)$ such that

$$A_{\bar{g}}\{(x, y, z) \in \mathcal{S}^2 : z > h\} = A_1$$
$$A_{\bar{g}}\{(x, y, z) \in \mathcal{S}^2 : z < h\} = A_2.$$

Then

$$\bar{\gamma} \doteqdot \{(x, y, z) \in \mathcal{S}^2 : z = h\}$$

is a round circle that divides $(\mathcal{S}^2, \bar{g})$ into two discs with areas A_1 and A_2. Because

$$\frac{d}{dh} A_{\bar{g}}\{(x, y, z) \in \mathcal{S}^2 : z < h\} = \frac{d}{dh} \int_{-\rho}^{h} 2\pi\rho \, dz = 2\pi\rho,$$

we have $A_2 = 2\pi\rho(\rho + h)$ and $A_1 = 2\pi\rho(\rho - h)$. Hence

$$4\pi \frac{A_1 A_2}{A_1 + A_2} = \frac{[2\pi\rho(\rho - h)][2\pi\rho(\rho + h)]}{\rho^2} = 4\pi^2(\rho^2 - h^2) = (L_{\bar{g}}(\bar{\gamma}))^2.$$

The lemma follows. □

We are now ready to state the main assertion of this section. Notice that it does not matter whether we consider the normalized or unnormalized Ricci flow, because the isoperimetric constant is scale invariant:

$$C_S\left(\mathcal{M}^2, g\right) = C_S\left(\mathcal{M}^2, e^\lambda g\right).$$

THEOREM 5.88. *Let $\left(\mathcal{M}^2, g(t)\right)$ be a solution of the Ricci flow on a topological 2-sphere. At any time t such that $C_H\left(\mathcal{M}^2, g(t)\right) < 4\pi$, one has*

$$\frac{d}{dt} C_H\left(\mathcal{M}^2, g(t)\right) \geq 0$$

in the sense of the lim inf *of forward difference quotients.*

REMARK 5.89. Because $C_H\left(\mathcal{M}^2, g(t)\right)$ is a continuous function of time satisfying the upper bound $C_H\left(\mathcal{M}^2, g(t)\right) \leq 4\pi$, the theorem implies in particular that $C_H\left(\mathcal{M}^2, g(t)\right)$ is nondecreasing in time.

Before proving the theorem, we introduce some notation. Assume in the remainder of this section that $\left(\mathcal{M}^2, g(t)\right)$ is a solution of the Ricci flow on a topological 2-sphere for $t \in [0, \omega)$. Let

$$A \equiv A(t) \doteqdot A_{g(t)}\left(\mathcal{M}^2\right).$$

Observe that it follows from Lemma 6.5 and the **Gauss–Bonnet Theorem** for a closed surface, that

$$\frac{d}{dt} A = -\int_{\mathcal{M}^2} R \, dA = -4\pi\chi\left(\mathcal{M}^2\right) = -8\pi. \tag{5.50}$$

Given any $t_0 \in [0, \omega)$ and any smooth embedded loop γ_0 separating $\mathcal{M}^2$ into $\mathcal{M}_0^+$ and $\mathcal{M}_0^-$, let γ_ρ denote the parallel curve of signed distance ρ from γ_0 measured with respect to the metric $g(t_0)$, where $\rho < 0$ for curves in $\mathcal{M}_0^+$ and $\rho > 0$ for curves in $\mathcal{M}_0^-$. For $|\rho|$ sufficiently small, γ_ρ is a smooth embedded loop that separates $\mathcal{M}^2$ into $\mathcal{M}_\rho^+$ and $\mathcal{M}_\rho^-$. We shall thus consider the following functions of both ρ and t:

$$\begin{aligned} L &\equiv L(\rho, t) \doteqdot L_{g(t)}(\gamma_\rho) \\ A_\pm &\equiv A_\pm(\rho, t) \doteqdot A_{g(t)}\left(\mathcal{M}_\rho^\pm\right) \\ C_H &\equiv C_H(\rho, t) \doteqdot (L(\rho, t))^2 \frac{A(t)}{A_+(\rho, t) \cdot A_-(\rho, t)}. \end{aligned}$$

Notice that by (5.47), C_H is the isoperimetric ratio of the curve γ_ρ taken with respect to the metric $g(t)$. The next three lemmas accomplish the calculations necessary to prove the theorem.

LEMMA 5.90. *Let $\left(\mathcal{M}^2, g(t)\right)$ be a solution of the Ricci flow on a topological 2-sphere. Then*

$$\frac{\partial}{\partial \rho} L = \int_{\gamma_\rho} k \, ds \qquad \frac{\partial}{\partial t} L = -\int_{\gamma_r} K \, ds$$

$$\frac{\partial}{\partial \rho} A_\pm = \pm L \qquad \frac{\partial}{\partial t} A_\pm = -4\pi \pm 2\int_{\gamma_r} k \, ds,$$

where $K = R/2$ is the Gaussian curvature and $k = \langle T, \nabla_T N\rangle$ is the geodesic curvature of the curve γ_ρ with unit tangent vector T and unit normal vector N oriented outward to $\mathcal{M}_\rho^+$.

PROOF. The derivatives with respect to ρ are easily verified. By Lemma 3.11,

$$\frac{\partial}{\partial t}L = -\frac{1}{2}\int_{\gamma_\rho} R\, ds = -\int_{\gamma_r} K\, ds.$$

By the Gauss–Bonnet formula for a disc with outward unit normal N, one has

$$\begin{aligned} 2\pi = 2\pi\chi\left(\mathcal{M}_\rho^\pm\right) &= \int_{\mathcal{M}_\rho^\pm} K\, dA \mp \int_{\partial\mathcal{M}_\rho^\pm} \langle \nabla_T T, N\rangle\, ds \\ &= \int_{\mathcal{M}_\rho^\pm} K\, dA \pm \int_{\gamma_\rho} \langle T, \nabla_T N\rangle\, ds, \end{aligned} \tag{5.51}$$

and hence

$$\frac{\partial}{\partial t}A_\pm = -\int_{\mathcal{M}_\rho^\pm} R\, dA = -4\pi \pm 2\int_{\gamma_\rho} k\, ds.$$

□

LEMMA 5.91. *Under the same hypothesis, L and $A_\pm$ evolve by*

$$\frac{\partial}{\partial t}L = \frac{\partial^2}{\partial\rho^2}L$$

and

$$\frac{\partial}{\partial t}A_\pm = \frac{\partial^2}{\partial\rho^2}A_\pm - 4\pi \pm \int_{\gamma_\rho} k\, ds,$$

respectively.

Notice that these resemble heat equations.

PROOF. Differentiating (5.51) with respect to ρ yields

$$0 = \int_{\gamma_\rho} K\, ds + \frac{\partial}{\partial\rho}\int_{\gamma_\rho} k\, ds.$$

Thus by Lemma 5.90,

$$\frac{\partial}{\partial t}L = \frac{\partial}{\partial\rho}\int_{\gamma_\rho} k\, ds = \frac{\partial^2}{\partial\rho^2}L.$$

The formulas for $A_\pm$ follow from the observation that

$$\frac{\partial^2}{\partial\rho^2}A_\pm = \pm\frac{\partial}{\partial\rho}L = \pm\int_{\gamma_\rho} k\, ds = \frac{\partial}{\partial t}A_\pm + 4\pi \mp \int_{\gamma_\rho} k\, ds.$$

□

LEMMA 5.92. *If $(\mathcal{M}^2, g(t))$ is a solution of the Ricci flow on a topological 2-sphere, the isoperimetric ratios $C_H(\rho, t)$ of the parallel loops γ_ρ measured with respect to the metric $g(t)$ satisfy the heat-type equation*

$$\frac{\partial}{\partial t}(\log C_H) = \frac{\partial^2}{\partial \rho^2}(\log C_H) + \frac{\Gamma}{L} \cdot \frac{\partial}{\partial \rho}(\log C_H) + \frac{4\pi - C_H}{A}\left(\frac{A_+}{A_-} + \frac{A_-}{A_+}\right),$$

where

$$\Gamma \doteqdot \int_{\gamma_\rho} k\, ds.$$

PROOF. Notice that

$$\log C_H = 2\log L - \log A_+ - \log A_- + \log A.$$

By Lemmas 5.90 and 5.91, we have

$$\begin{aligned} \frac{\partial}{\partial t}(\log L) = L^{-1}\frac{\partial^2}{\partial \rho^2}L &= \frac{\partial^2}{\partial \rho^2}(\log L) + \left(\frac{\partial}{\partial \rho}(\log L)\right)^2 \\ &= \frac{\partial^2}{\partial \rho^2}(\log L) + \frac{\Gamma}{L} \cdot \frac{\partial}{\partial \rho}(\log L). \end{aligned}$$

Similarly, we obtain

$$\begin{aligned} \frac{\partial}{\partial t}(\log A_\pm) &= A_\pm^{-1}\left[\frac{\partial^2}{\partial \rho^2}A_\pm - 4\pi \pm \Gamma\right] \\ &= \frac{\partial^2}{\partial \rho^2}(\log A_\pm) + \left(\frac{\partial}{\partial \rho}(\log A_\pm)\right)^2 - \frac{4\pi}{A_\pm} \pm \frac{\Gamma}{A_\pm} \\ &= \frac{\partial^2}{\partial \rho^2}(\log A_\pm) + \left(\frac{L}{A_\pm}\right)^2 - \frac{4\pi}{A_\pm} + \frac{\Gamma}{L} \cdot \frac{\partial}{\partial \rho}(\log A_\pm), \end{aligned}$$

because $L = \pm(\partial A_\pm / \partial \rho)$. Finally. we recall (5.50) to get

$$\frac{d}{dt}(\log A) = -\frac{8\pi}{A}.$$

The lemma follows when we combine these results with the identities

$$\frac{4\pi}{A}\left(\frac{A_+}{A_-} + \frac{A_-}{A_+}\right) = \frac{4\pi}{A_+ + A_-}\frac{(A_+ + A_-)^2 - 2A_+A_-}{A_+A_-} = \frac{4\pi}{A_+} + \frac{4\pi}{A_-} - \frac{8\pi}{A}$$

and

$$\frac{C_H}{A}\left(\frac{A_+}{A_-} + \frac{A_-}{A_+}\right) = \frac{L^2}{A_+A_-}\left(\frac{A_+}{A_-} + \frac{A_-}{A_+}\right) = \frac{L^2}{A_-^2} + \frac{L^2}{A_+^2}.$$

□

If $C_H < 4\pi$, then Lemma 5.92 shows that $\log C_H$ is a supersolution of a type of heat equation. This leads one to expect that it should be nondecreasing. We now prove this rigorously.

PROOF OF THEOREM 5.88. We shall show that for any $\varepsilon > 0$, we have

$$\left.\frac{d}{dt} C_H\left(\mathcal{M}^2, g(t)\right)\right|_{t=t_0} \geq -\varepsilon$$

at all times t_0 such that $C_H\left(\mathcal{M}^2, g(t_0)\right) < 4\pi$. Given such a time t_0, there is by Lemma 5.85 a smooth embedded loop γ_0 whose isoperimetric ratio measured with respect to $g(t_0)$ satisfies the identity

$$C_H(0, t_0) = C_H\left(\mathcal{M}^2, g(t)\right).$$

Since γ_0 is a minimizer of C_H at time t_0, we have the relations

$$\left(\frac{\partial}{\partial \rho} \log C_H\right)(0, t_0) = 0$$
$$\left(\frac{\partial^2}{\partial \rho^2} \log C_H\right)(0, t_0) \geq 0.$$

Using these and the inequality $4\pi - C_H \geq 0$, we conclude from Lemma 5.92 that

$$\left(\frac{\partial}{\partial t} \log C_H\right)(0, t_0) \geq 0.$$

This is possible only if for each $\varepsilon > 0$, there exists $\delta > 0$ depending on ε such that for all $t \in (t_0 - \delta, t_0)$, one has

$$\log C_H(0, t) \leq \log C_H(0, t_0) + \varepsilon(t_0 - t) = \log C_H\left(\mathcal{M}^2, g(t_0)\right) + \varepsilon(t_0 - t).$$

Hence for all $t \in (t_0 - \delta, t_0)$, we have

$$\log C_H\left(\mathcal{M}^2, g(t)\right) + \varepsilon t \leq \log C_H\left(\mathcal{M}^2, g(t_0)\right) + \varepsilon t_0.$$

□

15. An alternative strategy for the case $\chi\left(\mathcal{M}^2 > 0\right)$

We shall now complete the alternative strategy for proving convergence of the Ricci flow on a surface of positive Euler characteristic such that $R(\cdot, 0)$ might change sign. This approach uses the the isoperimetric estimates derived in the previous section. It also uses the entropy estimate of Section 8 and the LYH differential Harnack estimate of Section 10 for metrics of positive curvature, but (perhaps surprisingly) not their extensions to the case of variable curvature. It also relies on other important methods that will be developed later in this volume. In particular, the approach uses techniques of dilating about a singularity. (We will study such techniques in Chapter 8.)

REMARK 5.93. The reader will find related results in Section 6 of Chapter 9.

For the remainder of this section, we shall adopt notation different from that used in the rest of this chapter. Namely, we will denote by $(\mathcal{M}^2, g(t))$ the solution of the Ricci flow

$$\frac{\partial}{\partial t} g = -Rg, \qquad g(0) = g_0$$

and by $(\mathcal{M}^2, \bar{g}(\bar{t}))$ the corresponding solution of the normalized Ricci flow

$$\frac{\partial}{\partial \bar{t}} \bar{g} = -\bar{R}\bar{g}, \qquad \bar{g}(0) = g_0,$$

which exists for all $\bar{t} \in [0, \infty)$ by Proposition 5.19. As we shall see in Section 9 of Chapter 6, these are related by rescaling time

$$t = \frac{1}{\bar{r}}\left(1 - e^{-\bar{r}\bar{t}}\right), \qquad \bar{t} = \frac{1}{\bar{r}} \log \frac{1}{1 - \bar{r}t}$$

and dilating space

$$g(t) = e^{-\bar{r}\bar{t}}\, \bar{g}(\bar{t}) = (1 - \bar{r}t)\; \bar{g}(\bar{t}). \tag{5.52}$$

Here $\bar{r}(\bar{t}) \equiv \bar{r}(0)$ denotes the average scalar curvature of the metric g_0. Notice that $g(t)$ exists for $0 \leq t < T$, where

$$T \doteqdot \frac{1}{\bar{r}}. \tag{5.53}$$

If $A(t) \doteqdot \operatorname{Area}(\mathcal{M}^2, g(t))$, the Gauss-Bonnet formula

$$8\pi = 4\pi\chi\left(\mathcal{M}^2\right) \equiv \int_{\mathcal{M}^2} R[g(t)]\, d\mu[g(t)]$$

implies in particular that $dA/dt = -8\pi$ and $\bar{r}A(0) = 8\pi$. It follows that

$$A(t) = A(0) - 8\pi t = A(0)(1 - \bar{r}t), \tag{5.54}$$

hence that $A(t) \searrow 0$ and $R_{\max}(t) \nearrow \infty$ as $t \nearrow T$, where $R_{\max}(t)$ denotes the maximum scalar curvature at time t.

The first step in the current strategy is to show that one can bound the injectivity radius by the isoperimetric constant. This fact is a consequence of Klingenberg's result (Lemma B.34) that $\operatorname{inj}(\mathcal{M}^n, g)$ is no smaller than the minimum of $\pi/\sqrt{K_{\max}}$ and half the length of the shortest closed geodesic, where $K_{\max}$ denotes the maximum sectional curvature.

LEMMA 5.94. *If $(\mathcal{M}^2, g)$ is a topological 2-sphere with isoperimetric constant $C_H(\mathcal{M}^2, g)$ and maximum Gaussian curvature $K_{\max}$, then*

$$\left[\operatorname{inj}\left(\mathcal{M}^2, g\right)\right]^2 \geq \frac{\pi}{4K_{\max}} C_H\left(\mathcal{M}^2, g\right).$$

PROOF. Let γ be a geodesic loop of length $L(\gamma)$ dividing $\mathcal{M}^2$ into two topological discs $\mathcal{M}_1^2$ and $\mathcal{M}_2^2$ with $\partial\mathcal{M}_1^2 = \partial\mathcal{M}_2^2 = \gamma$. For $i = 1, 2$, let A_i

denote the area of $\mathcal{M}_i^2$. Since γ is a geodesic, the Gauss-Bonnet formula (5.51) reduces to

$$2\pi = \chi\left(\mathcal{M}_i^2\right) = \int_{\mathcal{M}_i^2} K\, dA$$

for $i = 1, 2$. This implies that $K_{\max} \cdot \min\{A_1, A_2\} \geq 2\pi$, equivalently that

$$\max\left\{\frac{1}{A_1}, \frac{1}{A_2}\right\} \leq \frac{K_{\max}}{2\pi}.$$

Thus by (5.47), we get the estimate

$$C_H\left(\mathcal{M}^2, g\right) \leq C_H(\gamma) = L(\gamma)^2\left(\frac{1}{A_1} + \frac{1}{A_2}\right) \leq \frac{K_{\max}}{\pi} L(\gamma)^2,$$

which we write in the form

$$L(\gamma)^2 \geq \frac{\pi}{K_{\max}} C_H\left(\mathcal{M}^2, g\right).$$

If $L_{\min}$ denotes the length of the shortest geodesic loop in $\left(\mathcal{M}^2, g\right)$, then Klingenberg's result says

$$\left[\operatorname{inj}\left(\mathcal{M}^2, g\right)\right]^2 \geq \min\left\{\frac{\pi^2}{K_{\max}}, \frac{L_{\min}^2}{4}\right\}.$$

Since we have $C_H\left(\mathcal{M}^2, g\right) \leq 4\pi$ by Lemma 5.84, we conclude that

$$\left[\operatorname{inj}\left(\mathcal{M}^2, g\right)\right]^2 \geq \min\left\{\frac{\pi^2}{K_{\max}} \cdot \frac{C_H\left(\mathcal{M}^2, g\right)}{4\pi}, \frac{\pi}{4K_{\max}} C_H\left(\mathcal{M}^2, g\right)\right\}.$$

□

Combining this result with Theorem 5.88, one obtains the following:

COROLLARY 5.95. *If $\left(\mathcal{M}^2, g(t)\right)$ is a solution of the Ricci flow on a topological $\mathcal{S}^2$ with Gaussian curvature bounded above by $K_{\max}(t)$, then*

$$\left[\operatorname{inj}\left(\mathcal{M}^2, g(t)\right)\right]^2 \geq \frac{\pi}{4K_{\max}(t)} C_H\left(\mathcal{M}^2, g(0)\right).$$

Our next task is to extract a limit of dilations about the singularity. Since the singularity time is $T < \infty$, the solution $\left(\mathcal{M}^2, g(t)\right)$ exhibits either a Type I singularity

$$\sup_{\mathcal{M}^2 \times [0,T)} |R(\cdot, t)|(T - t) < \infty$$

or Type IIa singularity

$$\sup_{\mathcal{M}^2 \times [0,T)} |R(\cdot, t)|(T - t) = \infty.$$

(See Section 1 of Chapter 8.) In the first case, we take a sequence of points $x_j \in \mathcal{M}^2$ and times $t_j \nearrow T$ chosen to give a Type I limit, as in Section 4.1 of Chapter 8; in the second case, we chose $x_j \in \mathcal{M}^2$ and $t_j \nearrow T$ to give a Type II limit, as in Section 4.2 of Chapter 8. In either case, we ensure that

$|R(x_i, t_i)| = \max_{\mathcal{M}^2} |R(\cdot, t_i)|$ and consider the sequence of parabolically dilated solutions

$$g_j(t) \doteqdot R(x_j, t_j) \cdot g\left(t_j + \frac{t}{R(x_j, t_j)}\right) \tag{5.55}$$

defined for $-t_j R(x_j, t_j) \leq t < (T - t_j) R(x_j, t_j)$. The dilated solutions $(\mathcal{M}^2, g_j(t))$ satisfy a curvature bound which depends on t but is uniform in j. Indeed, if we are dilating about a Type I singularity, there exists

$$\Omega \doteqdot \limsup_{t \nearrow T} (T - t) R_{\max}(t) < \infty$$

such that

$$R[g_j(t)] \leq \frac{\Omega}{\Omega - t};$$

whereas if the original singularity is Type IIa, there is a sequence $\Omega_j \nearrow \infty$ such that

$$R[g_j(t)] \leq \frac{1}{1 - t/\Omega_j}.$$

Since by Corollary 5.95, there is a uniform positive lower bound for all $\operatorname{inj}(\mathcal{M}^2, g_j(0))$, we can apply the Compactness Theorem from Section 3 of Chapter 7 to obtain a pointed subsequence $(\mathcal{M}^2, g_j(t), x_j)$ that converges uniformly on any compact time interval and in every C^k-norm to a complete pointed solution

$$(\mathcal{M}^2_\infty, g_\infty(t), x_\infty)$$

of the Ricci flow with scalar curvature R_∞. In the case of a Type I limit, $g_\infty(t)$ is an ancient solution that exists for $-\infty < t < \Omega$ and satisfies the curvature bound

$$R_\infty \leq \frac{\Omega}{\Omega - t}.$$

In the case of a Type IIa limit, $g_\infty(t)$ is an eternal solution that exists for $-\infty < t < \infty$ and satisfies

$$R_\infty(\cdot, \cdot) \leq 1 = R_\infty(x_\infty, 0).$$

The key step toward our goal will be to show that a Type IIa limit cannot occur. This is a consequence of the fact that an eternal solution with strictly positive curvature that attains the space-time maximum of its scalar curvature must be a steady Ricci soliton. In the particular case of a surface, the result is the following result.

LEMMA 5.96. *The only eternal solution $(\mathcal{N}^2, g(t))$ of the Ricci flow on a surface of strictly positive curvature that attains its maximum curvature in space and time is the cigar $(\mathbb{R}^2, g_\Sigma(t))$.*

PROOF. We claim that $(\mathcal{N}^2, g(t))$ is a gradient Ricci soliton. The result follows from this claim and Lemma 2.7, which proves that the only gradient soliton of positive curvature in dimension $n = 2$ is the cigar.

To prove the claim we consider the trace Harnack quantity for the unnormalized Ricci flow on a surface with bounded positive scalar curvature, namely

$$Q \doteqdot \frac{\partial}{\partial t}\log R - |\nabla \log R|^2 = \Delta \log R + R.$$

This is the quantity defined in (5.35) without the $-r$ term. Following the proof of Lemma 5.55, one computes that it evolves by

$$\frac{\partial}{\partial t}Q = \Delta Q + 2\langle \nabla \log R, \nabla Q\rangle + 2\left|\nabla\nabla \log R + \frac{1}{2}Rg\right|^2, \tag{5.56}$$

hence estimates that

$$\frac{\partial}{\partial t}Q \geq \Delta Q + 2\langle \nabla \log R, \nabla Q\rangle + Q^2. \tag{5.57}$$

(The reader should compare the equality and inequality above to (5.37) and (5.38), respectively.) If Q ever becomes nonnegative, then it remains so by the maximum principle. On the other hand, if $Q_{\min}(t_1) < 0$ for some t_1, then $Q_{\min}(t_0) < 0$ for all $t_0 \leq t_1$, and one has

$$Q_{\min}(t) \geq \frac{1}{(Q(t_0))^{-1} - (t - t_0)} \geq -\frac{1}{t - t_0}$$

for all $t > t_0$. Since $(\mathcal{N}^2, g(t))$ is an eternal solution, one may let $t_0 \searrow -\infty$, obtaining

$$Q_{\min}(t) \geq \lim_{t_0 \to -\infty} \frac{1}{t_0 - t} = 0.$$

Hence on $(\mathcal{N}^2, g(t))$, we have

$$\Delta \log R + R = Q \geq 0.$$

This is the ancient version of Hamilton's trace Harnack inequality. (Note the similarity between this argument and the proof of Lemma 9.15.)

Now by hypothesis, there exists $(p_0, t_0) \in \mathcal{N}^2 \times \mathbb{R}$ such that

$$R(p_0, t_0) = \sup_{\mathcal{N}^2 \times \mathbb{R}} R.$$

Thus we have $\partial R/\partial t = 0$ and $|\nabla R| = 0$ at (p_0, t_0), and hence $Q(p_0, t_0) = 0$. Since $Q \geq 0$, applying the strong maximum principle to equation (5.56) then shows that Q must vanish identically on $\mathcal{N}^2 \times \mathbb{R}$. By examining its evolution equation (5.56), we see that this is possible only if the tensor

$$\nabla\nabla \log R + \frac{1}{2}Rg$$

vanishes on $\mathcal{N}^2 \times \mathbb{R}$. In other words

$$\frac{\partial}{\partial t}g = -Rg = \mathcal{L}_{(\nabla \log R)^\sharp}\, g.$$

Hence $g(t)$ is a gradient Ricci soliton flowing along $\operatorname{grad}(\log R)$. $\square$

PROPOSITION 5.97. *Let $(\mathcal{M}^2, g(t))$ be a solution of the Ricci flow on a topological 2-sphere with $g(0) = g_0$. Then $g(t)$ exists for $0 \leq t < T \doteqdot 1/\bar{r}$, where $\bar{r}$ denotes the average scalar curvature of the metric g_0, and the singularity at time T is of Type I.*

PROOF. Suppose the singularity is of Type IIa. Since

$$A(t) = A(0) - 8\pi t = 8\pi\left(\frac{1}{\bar{r}} - t\right) = 8\pi(T - t)$$

by (5.53) and (5.54), this happens if and only if

$$\sup_{t\in[0,T)} R_{\max}(t) \cdot A(t) = \infty.$$

Since

$$\text{Area}\left(\mathcal{M}^2, g_j(0)\right) = R(x_j, t_j) \cdot A(t_j) = 8\pi R(x_j, t_j) \cdot (T - t_j) \to \infty$$

by (5.55), the limit $(\mathcal{M}^2_\infty, g_\infty(0))$ has infinite area. In particular, $\mathcal{M}^2_\infty$ is noncompact, and $g_\infty(t)$ has infinite area for all times $t \in (-\infty, \infty)$. Since we have

$$\lim_{\bar{t}\to\infty} \min_{\mathcal{M}^2} \bar{R}(\cdot, \bar{t}) \nearrow 0,$$

hence

$$\lim_{t\to T} \min_{\mathcal{M}^2} R(\cdot, t) \nearrow 0$$

by Lemma 5.9, the scalar curvature of $(\mathcal{M}^2_\infty, g_\infty(0))$ is nonnegative, hence strictly positive by the strong maximum principle. Moreover, $R^2_\infty \in (0, 1]$ attains its maximum at $(x_\infty, 0)$. By Lemma 5.96, $(\mathcal{M}^2_\infty, g_\infty(t))$ is the cigar soliton.

As we observed in Section 2 of Chapter 2, the cigar is the metric given in polar coordinates (ρ, θ) on $\mathbb{R}^2$ by

$$g_\Sigma = \frac{1}{1+\rho^2} d\rho^2 + \frac{\rho^2}{1+\rho^2} d\theta^2.$$

As we saw in Section 2 of Chapter 2, it attains its maximum curvature at the origin and is asymptotic to a cylinder as $\rho \to \infty$. In particular, for any $\varepsilon > 0$ there is $C \gg 1$ large enough so that the circle $\rho = C$ has length $2\pi - \varepsilon < L < 2\pi$ and the open disc $\rho < C$ has area $A > 1/\varepsilon$. Hence if $(\mathcal{M}^2, g_j(t), x_j)$ is a sequence of closed pointed manifolds converging to the cigar $(\mathbb{R}^2, g_\Sigma, 0)$, it is easy to see that there is $j \in \mathbb{N}$ so large that $C_H(\mathcal{M}^2, g_j) < \varepsilon$. Since this contradicts Theorem 5.88, we conclude that the Type IIa singularity is impossible. $\square$

COROLLARY 5.98. *The scalar curvature of any solution $(\mathcal{M}^2, \bar{g}(\bar{t}))$ of the normalized Ricci flow on a topological 2-sphere is uniformly bounded.*

PROOF. Since

$$g(t) = (1 - \bar{r}t)\, \bar{g}(\bar{t}) = \bar{r}(T - t)\, \bar{g}(\bar{t})$$

by (5.52) and (5.53), we have

$$\bar{R}(\cdot, \bar{t}) = \bar{r}(T-t)\, R(\cdot, t).$$

But since $g(t)$ has a Type I singularity, the right-hand side is uniformly bounded:

$$\limsup_{t \nearrow T} (T-t)\, R_{\max}(t) \doteqdot \Omega < \infty.$$

□

Now we finish the argument.

PROPOSITION 5.99. *If $(\mathcal{M}^2, g_0)$ is a closed Riemannian surface with average scalar curvature $r > 0$, then the unique solution $g(t)$ of the unnormalized Ricci flow with $g(0) = g_0$ encounters a singularity at some finite time T. A sequence $(\mathcal{M}^2, g_j(t))$ of parabolic dilations formed as $t_j \nearrow T$ converges locally smoothly to a shrinking round sphere $(\mathcal{S}^2, g_\infty(t))$.*

PROOF. We know that $(\mathcal{M}^2, g_j(t), x_j)$ converges locally smoothly in the pointed category to a limit $(\mathcal{M}^2_\infty, g_\infty(t), x_\infty)$ that must be an ancient solution with bounded curvature. By Lemma 9.15, its scalar curvature is positive. Let $\tau \doteqdot -t$. Then by Proposition 5.39, the entropy $N(g(\tau))$ is increasing in τ and is bounded above by the entropy of a soliton. Thus (going backwards in time) we can take a limit as $\tau \to \infty$ to get a solution $(\mathcal{M}^2_{2\infty}, g_{2\infty}(t), x_{2\infty})$ which has constant entropy, hence must be a soliton. By Proposition 5.21, $(\mathcal{M}^2_{2\infty}, g_{2\infty}(t), x_{2\infty})$ is a shrinking round sphere. But its entropy is minimal. Since $N(g(\tau))$ cannot decrease as τ increases, this implies that $N(g(\tau))$ must be constant, hence that the original limit $(\mathcal{M}^2_\infty, g_\infty(t), x_\infty)$ was itself the shrinking round sphere. □

REMARK 5.100. The assertion that the limit $(\mathcal{M}^2_\infty, g_\infty(t), x_\infty)$ must be the shrinking round 2-sphere also follows from Proposition 9.23.

Notes and commentary

The main reference for this chapter is Hamilton's paper on the Ricci flow on surfaces [**60**]. Unless otherwise noted, the results in this chapter may found there. In particular, the methods of entropy and Harnack estimates, Ricci solitons, and potential of the curvature were first introduced to the Ricci flow in [**60**]. Curvature derivative estimates were established by Shi in [**117, 118**]. Originally, curvature derivative estimates were obtained in Corollary 8.2 of [**60**] by bounding the L^2-norms of all derivatives and applying the Sobolev inequality. (See also Theorem 13.4 in [**58**].) The method of proof we give for the upper bound of R in Proposition 5.30 follows [**30**]. The direct proof of the entropy estimate of Proposition 5.39 was given in [**29**]; the results in subsection 8.1 are also from there. The method of proving the differential Harnack inequality in section 10 was first used by Li and Yau [**92**] for solutions of the heat equation. (See Section 6 of [**60**].) The extension to the case where R changes sign in Section 10.2 was established in [**28**]. A new proof without the use of the potential function was given by Hamilton and Yau [**68**]. The lower bound for the injectivity radius on $\mathcal{S}^2$ was proved by Hamilton. (See Section 5 of [**28**] or Section 12 of [**63**].) The isoperimetric estimate in Section 14 was given by Hamilton [**62**]; the application of this to give a new proof that R is uniformly bounded uses the Gromov-type compactness theorem established in [**64**] and reviewed in Section 3 of Chapter 7.

Results on the Ricci flow on noncompact surfaces are established in [**128**], [**129**], [**72**], [**73**], and [**35**]. Convergence theorems on 2-orbifolds are given in [**127**] and [**34**]. A new proof of the convergence on $\mathcal{S}^2$ using the Aleksandrov method was obtained by Bartz, Struwe, and Ye [**10**]. A related flow on $\mathcal{S}^2$ was studied by Leviton and Rubinstein [**90**].

The methods established for the Ricci flow are related to methods for other geometric evolution equations. Both the entropy and differential Harnack estimates were established by Hamilton for the curve shortening flow [**44**]. The entropy estimate was extended to higher dimensional hypersurface flows in [**31**] and [**3**]. References for differential Harnack inequalities of LYH type in higher dimensions and for other geometric evolution equations will be given in a chapter of the planned successor to this volume.

CHAPTER 6

Three-manifolds of positive Ricci curvature

The topic of this chapter is Hamilton's application of the Ricci flow to the classification of closed 3-manifolds of positive Ricci curvature, in particular his landmark result that any such manifold is diffeomorphic to a spherical space form. Let $(\mathcal{M}^n, g)$ be a closed Riemannian n-manifold with positive Ricci curvature. By Myers' Theorem, $\mathcal{M}^n$ is compact. Since its universal cover $\widetilde{\mathcal{M}^n}$ also has positive Ricci curvature, it too is compact; and hence the fundamental group of $\mathcal{M}^n$ is finite. When $n = 3$, we have the well-known

CONJECTURE 6.1 (Poincaré Conjecture). *Any simply connected closed smooth 3-manifold is diffeomorphic to $\mathcal{S}^3$.*

A successful resolution of the Poincaré Conjecture implies that $\widetilde{\mathcal{M}^3}$ is diffeomorphic to $\mathcal{S}^3$. Furthermore, there is also the following

CONJECTURE 6.2 (Spherical Space Form Conjecture). *Any finite group of diffeomorphisms acting freely on $\mathcal{S}^3$ is conjugate to a group of isometries of the standard sphere.*

In the presence of a complete proof of the Spherical Space Form Conjecture, the diffeomorphism $\widetilde{\mathcal{M}^3} \approx \mathcal{S}^3$ would imply that $\mathcal{M}^3$ is diffeomorphic to $\mathcal{S}^3/\Gamma$, where Γ is a finite subgroup of $\mathrm{O}(4)$. In other words, one could conclude that $\mathcal{M}^3$ is diffeomorphic to a space form, all of which are classified. [**126**] As a consequence, $\mathcal{M}^3$ would admit a metric of constant positive sectional curvature. It is this last statement — the existence of a constant positive sectional curvature metric on any closed 3-manifold with positive Ricci curvature — that Hamilton proved in 1982 using the then newly-introduced Ricci flow. Our goal in this chapter is to present his result:

THEOREM 6.3 (Hamilton). *Let $(\mathcal{M}^3, g_0)$ be a closed Riemannian 3-manifold of positive Ricci curvature. Then a unique solution $g(t)$ of the normalized Ricci flow with $g(0) = g_0$ exists for all positive time; and as $t \to \infty$, the metrics $g(t)$ converge exponentially fast in every C^m-norm to a metric g_∞ of constant positive sectional curvature.*

Although the result is stated for the normalized flow, it will be more convenient to work with the unnormalized flow for as long as possible. The two equations differ only by rescaling space and time. (See Subsection 9.1.) In fact, the theorem may be interpreted as a statement of the convergence of the unnormalized flow modulo rescaling; but the advantage of finally converting to the normalized flow is that it will allow us more naturally to demonstrate exponential convergence. Our strategy for proving Theorem 6.3 will be to learn enough about how curvature evolves under the unnormalized flow to obtain certain *a priori* estimates that will suffice to prove long-time existence and convergence of its normalized cousin.

1. The evolution of curvature under the Ricci flow

In this section, we compute the evolution equations for the Levi-Civita connection and the Riemann, Ricci, and scalar curvatures of a solution to the Ricci flow. We start by recalling certain general evolution equations for a one-parameter family of metrics which were derived in Section 1 of Chapter 3.

REMARK 6.4. Unless explicitly stated otherwise, all results in this section hold for any dimension $n \geq 2$.

LEMMA 6.5. *Suppose that $g(t)$ is a smooth one-parameter family of metrics on a manifold $\mathcal{M}^n$ such that*

$$\frac{\partial}{\partial t} g = h.$$

(1) *The Levi-Civita connection Γ of g evolves by*

$$\frac{\partial}{\partial t}\Gamma_{ij}^k = \frac{1}{2} g^{k\ell} \left(\nabla_i h_{j\ell} + \nabla_j h_{i\ell} - \nabla_\ell h_{ij}\right).$$

(2) *The $(3,1)$-Riemann curvature tensor* Rm *of g evolves by*

$$\begin{aligned}\frac{\partial}{\partial t} R_{ijk}^{\ell} &= \frac{1}{2} g^{\ell p} \left\{ \begin{array}{c} \nabla_i \nabla_j h_{kp} + \nabla_i \nabla_k h_{jp} - \nabla_i \nabla_p h_{jk} \\ -\nabla_j \nabla_i h_{kp} - \nabla_j \nabla_k h_{ip} + \nabla_j \nabla_p h_{ik} \end{array} \right\} \\ &= \frac{1}{2} g^{\ell p} \left\{ \begin{array}{c} \nabla_i \nabla_k h_{jp} + \nabla_j \nabla_p h_{ik} - \nabla_i \nabla_p h_{jk} - \nabla_j \nabla_k h_{ip} \\ -R_{ijk}^q h_{qp} - R_{ijp}^q h_{kq} \end{array} \right\}.\end{aligned}$$

(3) *The Ricci tensor* Rc *of g evolves by*

$$\begin{aligned}\frac{\partial}{\partial t} R_{jk} &= \frac{1}{2} g^{pq} \left(\nabla_q \nabla_j h_{kp} + \nabla_q \nabla_k h_{jp} - \nabla_q \nabla_p h_{jk} - \nabla_j \nabla_k h_{qp}\right) \\ &= -\frac{1}{2}\left[\Delta_L h_{jk} + \nabla_j \nabla_k \left(\operatorname{tr}_g h\right) + \nabla_j \left(\delta h\right)_k + \nabla_k \left(\delta h\right)_j\right],\end{aligned}$$

where Δ_L denotes the Lichnerowicz Laplacian defined in formula (3.6).

(4) *The scalar curvature R of g evolves by*

$$\begin{aligned}\frac{\partial}{\partial t} R &= g^{ij} g^{k\ell} \left(-\nabla_i \nabla_j h_{kl} + \nabla_i \nabla_k h_{j\ell} - h_{ik} R_{j\ell}\right) \\ &= -\Delta \left(\operatorname{tr}_g h\right) + \operatorname{div}\left(\operatorname{div} h\right) - \langle h, \mathrm{Rc}\rangle\end{aligned}$$

(5) *The volume form $d\mu$ of g evolves by*

$$\frac{\partial}{\partial t} d\mu = \frac{1}{2} (\operatorname{tr}_g h)\, d\mu.$$

Substituting $h = -2\,\mathrm{Rc}$ into the lemma yields the following result.

COROLLARY 6.6. *Suppose $g(t)$ is a solution of the Ricci flow.*

(1) The Levi-Civita connection Γ of g evolves by

$$\frac{\partial}{\partial t}\Gamma_{ij}^{k} = -g^{k\ell} \left(\nabla_i R_{j\ell} + \nabla_j R_{i\ell} - \nabla_\ell R_{ij}\right). \tag{6.1}$$

(2) The $(3,1)$-Riemann curvature tensor Rm of g evolves by

$$\frac{\partial}{\partial t} R_{ijk}^{\ell} = g^{\ell p} \left\{ \begin{array}{c} -\nabla_i \nabla_j R_{kp} - \nabla_i \nabla_k R_{jp} + \nabla_i \nabla_p R_{jk} \\ \\ +\nabla_j \nabla_i R_{kp} + \nabla_j \nabla_k R_{ip} - \nabla_j \nabla_p R_{ik} \end{array} \right\} \tag{6.2}$$

(3) The Ricci tensor Rc of g evolves by

$$\frac{\partial}{\partial t} R_{jk} = \Delta R_{jk} + \nabla_j \nabla_k R - g^{pq} \left(\nabla_q \nabla_j R_{kp} + \nabla_q \nabla_k R_{jp}\right) \tag{6.3}$$

(4) The scalar curvature R of g evolves by

$$\frac{\partial}{\partial t} R = 2\Delta R - 2 g^{jk} g^{pq} \nabla_q \nabla_j R_{kp} + 2 \left|\mathrm{Rc}\right|^2 \tag{6.4}$$

(5) The volume form $d\mu$ of g evolves by

$$\frac{\partial}{\partial t} d\mu = -R\, d\mu. \tag{6.5}$$

We shall soon obtain more useful forms of many of the equations above by applying various curvature identities. To motivate the type of equation we are looking for, recall formula (5.3) for the evolution of the scalar curvature of a solution of the normalized Ricci flow on a surface:

$$\frac{\partial}{\partial t} R = \Delta R + R\,(R - r).$$

This is an example of a **reaction-diffusion equation** with a quadratic nonlinearity involving the scalar curvature. Equations of this form are also called **heat-type equations**. We shall see below that in any dimension $n \geq 2$, the Riemann curvature Rm, the Ricci curvature Rc, and the scalar curvature R all satisfy heat-type equations whose nonlinear terms are various quadratic contractions of the Riemann curvature.

1.1. The scalar curvature function. We first consider the scalar curvature function, since its evolution equation is the simplest. After two contractions of the second Bianchi identity

$$\nabla_m R_{ijk\ell} + \nabla_k R_{ij\ell m} + \nabla_\ell R_{ijmk} = 0,$$

one gets $g^{jk} \nabla_j R_{k\ell} = \frac{1}{2} \nabla_\ell R$. Applying this to (6.4) gives a nicer equation.

LEMMA 6.7. *The scalar curvature of a solution to the Ricci flow evolves by*

$$\frac{\partial}{\partial t} R = \Delta R + 2\left|\mathrm{Rc}\right|^2 . \tag{6.6}$$

This is the general (n-dimensional) version of the evolution equation for the curvature of a surface derived in Chapter 5.

Although the evolution equation (6.6) for the scalar curvature R involves the Ricci tensor, the latter quantity enters only in a nonnegative term. Thus the weak maximum principle for scalar equations implies that positive scalar curvature is preserved under the Ricci flow. More generally, we have the following result.

LEMMA 6.8. *Let $g(t)$ be a solution of the Ricci flow with $g(0) = g_0$. If the scalar curvature of g_0 is bounded below by some constant ρ, then $g(t)$ has scalar curvature $R \geq \rho$ for as long as the solution exists.*

1.2. The Ricci curvature tensor. Applying the contracted second Bianchi identity to the second equation for the evolution of Rc in Lemma 6.5 puts equation (6.3) in a nicer form.

LEMMA 6.9. *Under the Ricci flow, the Ricci tensor evolves by*

$$\frac{\partial}{\partial t} R_{jk} = \Delta_L R_{jk} = \Delta R_{jk} + 2g^{pq}g^{rs}R_{pjkr}R_{qs} - 2g^{pq}R_{jp}R_{qk}. \tag{6.7}$$

PROOF. We have

$$\begin{aligned}\frac{\partial}{\partial t} R_{jk} &= \Delta_L R_{jk} + \nabla_j\nabla_k R - g^{pq}\left(\nabla_j\nabla_p R_{qk} + \nabla_k\nabla_p R_{jq}\right) \\ &= \Delta_L R_{jk} + \nabla_j\nabla_k R - \frac{1}{2}\left(\nabla_j\nabla_k R + \nabla_k\nabla_j R\right).\end{aligned}$$

□

The presence of the Riemann tensor in the evolution equation for the Ricci tensor is an obstacle to showing that nonnegative Ricci curvature is preserved in arbitrary dimensions. (In fact, there exist examples [**101**] where positive Ricci is *not* preserved.) One wants, therefore, to attain a better understanding of the contribution of the Riemann tensor to (6.7). Recall that in any dimension, the Riemann tensor admits the orthogonal decomposition

$$\mathrm{Rm} = \frac{R}{2n(n-1)}\left(g \odot g\right) + \frac{1}{n-2}\left(\overset{\circ}{\mathrm{Rc}} \odot g\right) + W, \tag{6.8}$$

where $\odot$ denotes the Kulkarni–Nomizu product of symmetric tensors,

$$\left(P \odot Q\right)_{ijk\ell} \doteqdot P_{i\ell}Q_{jk} + P_{jk}Q_{i\ell} - P_{ik}Q_{j\ell} - P_{j\ell}Q_{ik},$$

$\overset{\circ}{\mathrm{Rc}}$ denotes the trace-free part of the Ricci tensor, and W is the Weyl tensor. Because the Weyl tensor vanishes in dimension $n = 3$, equation (6.8) reduces to

$$R_{ijk\ell} = R_{i\ell}g_{jk} + R_{jk}g_{i\ell} - R_{ik}g_{j\ell} - R_{j\ell}g_{ik} - \frac{R}{2}\left(g_{i\ell}g_{jk} - g_{ik}g_{j\ell}\right). \tag{6.9}$$

This identity shows in particular that the Ricci tensor determines the Riemann tensor when $n = 3$. Because of this, the reaction terms in (6.7) depend only on the Ricci tensor.

LEMMA 6.10. *In dimension $n = 3$, the Ricci tensor of a solution to the Ricci flow evolves by*

$$\frac{\partial}{\partial t} R_{jk} = \Delta R_{jk} + 3RR_{jk} - 6g^{pq} R_{jp} R_{qk} + \left(2\left|\mathrm{Rc}\right|^2 - R^2\right) g_{jk}. \tag{6.10}$$

PROOF. Substituting (6.9) into the second term on the right-hand side of (6.7) yields

$$2g^{pq} g^{rs} R_{pjkr} R_{qs} = \left(2\left|\mathrm{Rc}\right|^2 - R^2\right) g_{jk} - 4g^{pq} R_{jp} R_{qk} + 3RR_{jk}.$$

□

The weak maximum principle for tensors (discussed in Chapter 4) implies the following important consequence of formula (6.10).

COROLLARY 6.11. *Let $g(t)$ be a solution of the Ricci flow with on a 3-manifold with $g(0) = g_0$. If g_0 has positive (nonnegative) Ricci curvature, then $g(t)$ has positive (nonnegative) Ricci curvature for as long as the solution exists.*

PROOF. At any point and time where the Ricci tensor has a null eigenvector, it has at most two nonzero eigenvalues. Then one has $\left|\mathrm{Rc}\right|^2 \geq R^2/2$, whence the result follows immediately from Theorem 4.6. □

REMARK 6.12. More generally, in any dimension $n \geq 3$, the decomposition (6.8) lets one write the evolution equation (6.7) satisfied by the Ricci tensor in the form

$$\begin{aligned}\frac{\partial}{\partial t} R_{jk} &= \Delta R_{jk} - \frac{2n}{n-2} R_j^\ell R_{\ell k} + \frac{2n}{(n-1)(n-2)} RR_{jk} \\ &\quad + \frac{2}{n-2}\left(\left|\mathrm{Rc}\right|^2 - \frac{1}{n-1} R^2\right) g_{jk} + 2R^{pq} W_{jpqk}.\end{aligned}$$

Since $\left|\mathrm{Rc}\right|^2 \geq R^2/(n-1)$ whenever Rc has a zero eigenvalue, one concludes in particular that the Weyl tensor is the sole obstacle to preserving the condition $\mathrm{Rc} \geq 0$.

1.3. The Riemann curvature tensor. The observations made in the previous two subsections suggest that the Riemann curvature tensor also satisfies a heat-type (reaction-diffusion) equation.

LEMMA 6.13. *Under the Ricci flow, the $(3,1)$-Riemann curvature tensor evolves by*

$$\frac{\partial}{\partial t} R_{ijk}^\ell = \Delta R_{ijk}^\ell + g^{pq}\left(R_{ijp}^r R_{rqk}^\ell - 2R_{pik}^r R_{jqr}^\ell + 2R_{pir}^\ell R_{jqk}^r\right) \tag{6.11a}$$

$$- R_i^p R_{pjk}^\ell - R_j^p R_{ipk}^\ell - R_k^p R_{ijp}^\ell + R_p^\ell R_{ijk}^p. \tag{6.11b}$$

PROOF. By applying the second Bianchi identity and commuting covariant derivatives, we compute

$$
\begin{aligned}
\Delta R_{ijk}^{\ell} &= g^{pq}\nabla_p\nabla_q R_{ijk}^{\ell} = g^{pq}\nabla_p\left(-\nabla_i R_{jqk}^{\ell} - \nabla_j R_{qik}^{\ell}\right) \\
&= g^{pq}\left\{\begin{array}{l} -\nabla_i\nabla_p R_{jqk}^{\ell} + R_{pij}^{r}R_{rqk}^{\ell} + R_{piq}^{r}R_{jrk}^{\ell} + R_{pik}^{r}R_{jqr}^{\ell} - R_{pir}^{\ell}R_{jqk}^{r} \\ +\nabla_j\nabla_p R_{qik}^{\ell} - R_{pji}^{r}R_{rqk}^{\ell} - R_{pjq}^{r}R_{irk}^{\ell} - R_{pjk}^{r}R_{iqr}^{\ell} + R_{pjr}^{\ell}R_{iqk}^{r} \end{array}\right\}
\end{aligned}
$$

and then apply the second Bianchi identity again to obtain

$$
g^{pq}\nabla_p R_{jqk}^{\ell} = g^{pq}g^{\ell m}\left(-\nabla_k R_{jqmp} - \nabla_m R_{jqpk}\right) = \nabla_k R_j^{\ell} - \nabla^{\ell} R_{jk}.
$$

This lets us rewrite the formula for ΔR_{ijk}^{ℓ} given above as

$$
\begin{aligned}
\Delta R_{ijk}^{\ell} &= -\nabla_i\nabla_k R_j^{\ell} + \nabla_i\nabla^{\ell} R_{jk} + \nabla_j\nabla_k R_i^{\ell} - \nabla_j\nabla^{\ell} R_{ik} \\
&\quad + g^{pq}\left\{\begin{array}{l} R_{pij}^{r}R_{rqk}^{\ell} + R_{piq}^{r}R_{jrk}^{\ell} + R_{pik}^{r}R_{jqr}^{\ell} - R_{pir}^{\ell}R_{jqk}^{r} \\ -R_{pji}^{r}R_{rqk}^{\ell} - R_{pjq}^{r}R_{irk}^{\ell} - R_{pjk}^{r}R_{iqr}^{\ell} + R_{pjr}^{\ell}R_{iqk}^{r} \end{array}\right\}.
\end{aligned}
$$

Now the first Bianchi identity shows that

$$
R_{pij}^{r}R_{rqk}^{\ell} - R_{pji}^{r}R_{rqk}^{\ell} = -R_{ijp}^{r}R_{rqk}^{\ell},
$$

which implies that ΔR_{ijk}^{ℓ} may be written as

$$
\Delta R_{ijk}^{\ell} = -\nabla_i\nabla_k R_j^{\ell} + \nabla_i\nabla^{\ell} R_{jk} + \nabla_j\nabla_k R_i^{\ell} \tag{6.12a}
$$

$$
- \nabla_j\nabla^{\ell} R_{ik} - R_i^{r} R_{jrk}^{\ell} + R_j^{r} R_{irk}^{\ell} \tag{6.12b}
$$

$$
+ g^{pq}\left\{\begin{array}{c} -R_{ijp}^{r}R_{rqk}^{\ell} + R_{pik}^{r}R_{jqr}^{\ell} + R_{pjr}^{\ell}R_{iqk}^{r} \\ -R_{pir}^{\ell}R_{jqk}^{r} - R_{pjk}^{r}R_{iqr}^{\ell} \end{array}\right\}. \tag{6.12c}
$$

On the other hand, rewriting the commutator term $g^{\ell p}\left(\nabla_j\nabla_i - \nabla_i\nabla_j\right)R_{kp}$ in formula (6.2) yields

$$
\frac{\partial}{\partial t}R_{ijk}^{\ell} = -\nabla_i\nabla_k R_j^{\ell} + \nabla_i\nabla^{\ell} R_{jk} + \nabla_j\nabla_k R_i^{\ell} - \nabla_j\nabla^{\ell} R_{ik} \tag{6.13a}
$$

$$
+ g^{\ell p}\left(R_{ijk}^{q}R_{qp} + R_{ijp}^{q}R_{kq}\right). \tag{6.13b}
$$

The reaction-diffusion equation for the Riemann curvature tensor now follows from comparing terms in formulas (6.12) and (6.13). ☐

We define the $(4,0)$-Riemann curvature tensor by $R_{ijk\ell} \doteqdot g_{\ell m}R_{ijk}^{m}$, so that $R_{1221} > 0$ on the round sphere. The following observation then follows immediately from the lemma above.

COROLLARY 6.14. *Under the Ricci flow, the* $(4,0)$*-Riemann curvature tensor satisfies the following reaction-diffusion equation:*

$$\frac{\partial}{\partial t} R_{ijk\ell} = \Delta R_{ijk\ell} + g^{pq}\left(R_{ijp}^{r} R_{rqk\ell} - 2R_{pik}^{r} R_{jqr\ell} + 2R_{pir\ell} R_{jqk}^{r}\right) \tag{6.14a}$$

$$- \left(R_i^p R_{pjk\ell} + R_j^p R_{ipk\ell} + R_k^p R_{ijp\ell} + R_\ell^p R_{ijkp}\right). \tag{6.14b}$$

There is a nice way to rewrite the evolution equation (6.14) in terms of a $(4,0)$-tensor B defined by

$$B_{ijk\ell} \doteqdot -g^{pr} g^{qs} R_{ipjq} R_{kr\ell s} = -R_{pij}^{q} R_{q\ell k}^{p}. \tag{6.15}$$

(The minus sign here, which does not appear in Hamilton's original paper [**58**], occurs because we are using the opposite sign convention for $R_{ijk\ell}$.) Note that B is quadratic in the Riemann curvature tensor and satisfies the following algebraic identity

$$B_{ijk\ell} = B_{ji\ell k} = B_{k\ell ij}. \tag{6.16}$$

LEMMA 6.15. *Under the Ricci flow, the* $(4,0)$*-Riemann curvature tensor evolves by*

$$\frac{\partial}{\partial t} R_{ijk\ell} = \Delta R_{ijk\ell} + 2\left(B_{ijk\ell} - B_{ij\ell k} + B_{ikj\ell} - B_{i\ell jk}\right) \tag{6.17a}$$

$$- \left(R_i^p R_{pjk\ell} + R_j^p R_{ipk\ell} + R_k^p R_{ijp\ell} + R_\ell^p R_{ijkp}\right). \tag{6.17b}$$

PROOF. First observe that the last two terms on the right-hand side of (6.14a) may be directly expressed in terms of B:

$$\begin{aligned} -2g^{pq} R_{pik}^{r} R_{jqr\ell} &= -2g^{pq} R_{ipkr} R_{jq\ell r} = 2B_{ikj\ell}, \\ 2g^{pq} R_{pir\ell} R_{jqk}^{r} &= 2g^{pq} R_{ip\ell r} R_{jqkr} = -2B_{i\ell jk}. \end{aligned}$$

This leaves us with $g^{pq} R_{ijp}^{r} R_{rqk\ell}$. Applying the first Bianchi identity gives

$$\begin{aligned} g^{pq} R_{ijp}^{r} R_{rqk\ell} &= g^{pq} g^{rs} R_{rpji} R_{sqk\ell} \\ &= g^{pq} g^{rs} \left(-R_{rjip} - R_{ripj}\right)\left(-R_{sk\ell q} - R_{s\ell qk}\right) \\ &= -B_{jik\ell} + B_{jilk} + B_{ijk\ell} - B_{ij\ell k}, \end{aligned}$$

whence identity (6.16) implies that

$$g^{pq} R_{ijp}^{r} R_{rqk\ell} = 2\left(B_{ijk\ell} - B_{ij\ell k}\right).$$

□

REMARK 6.16. Observe that although B does not in general satisfy the first Bianchi identity, the tensor C does, where

$$C_{ijk\ell} \doteqdot B_{ijk\ell} - B_{ij\ell k} + B_{ikj\ell} - B_{i\ell jk}.$$

Similarly, although $B_{jik\ell} \neq -B_{ijk\ell}$ in general, one has $C_{jik\ell} = -C_{ijk\ell}$.

2. Uhlenbeck's trick

Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow, and let $\{e_a^0\}_{a=1}^n$ be a local frame field defined in an open set $\mathcal{U} \subset \mathcal{M}^n$. Suppose that $\{e_a^0\}$ is orthonormal with respect to the initial metric g_0. We would like to evolve this frame field so that it stays orthonormal with respect to $g(t)$. This may easily be done by considering the following ODE system defined in $T_x\mathcal{M}^n$ for each $x \in \mathcal{U}$:

$$\frac{d}{dt} e_a(x,t) = \operatorname{Rc}(e_a(x,t)), \tag{6.18a}$$

$$e_a(x,0) = e_a^0(x). \tag{6.18b}$$

Here we regard the Ricci tensor $\operatorname{Rc} \equiv \operatorname{Rc}[g(t)]$ as a $(1,1)$-tensor, that is, an endomorphism $\operatorname{Rc} : T\mathcal{M}^n \to T\mathcal{M}^n$. Since (6.18) is a linear system of n ordinary differential equations, a unique solution exists as long as the solution $g(t)$ of the Ricci flow exists.

LEMMA 6.17. *The inner product $g(e_a, e_b)$ is constant for any solution $g(t)$ of the Ricci flow and any frame field $\{e_a(t)\}$ satisfying (6.18).*

PROOF. Noting that g, Rc, and each e_a all depend on time, we compute that

$$\begin{aligned}
\frac{\partial}{\partial t}[g(e_a, e_b)] &= \left(\frac{\partial}{\partial t} g\right)(e_a, e_b) + g\left(\frac{\partial}{\partial t} e_a, e_b\right) + g\left(e_a, \frac{\partial}{\partial t} e_b\right) \\
&= -2\operatorname{Rc}(e_a, e_b) + g(\operatorname{Rc}(e_a), e_b) + g(e_a, \operatorname{Rc}(e_b)) \\
&= 0.
\end{aligned}$$

□

COROLLARY 6.18. *If $\{e_a^0\}$ is orthonormal with respect to g_0, then $\{e_a(t)\}$ remains orthonormal with respect to $g(t)$.*

REMARK 6.19. If $\mathcal{M}^n$ is parallelizable (in other words, if $T\mathcal{M}^n$ is a trivial bundle) there exists a *global* frame field that is orthonormal with respect to any given Riemannian metric g. Indeed, one may take any global frame field and apply the Gram–Schmidt process (which preserves smoothness of the frame field).

REMARK 6.20. Every closed 3-manifold is parallelizable, hence may be assigned a global orthonormal frame field.

The idea of evolving a frame field to compensate for the evolution of a Riemannian metric has a more abstract formulation, due to Karen Uhlenbeck, which we now present. The **Uhlenbeck trick** allows us to put the evolution equation (6.17) satisfied by Rm into a particularly nice form. We begin as follows. Let $(\mathcal{M}^n, g(t) : t \in [0,T))$ be a solution to the Ricci flow with $g(0) = g_0$. Let V be a vector bundle over $\mathcal{M}^n$ isomorphic to $T\mathcal{M}^n$, and let $\iota_0 : V \to T\mathcal{M}^n$ be a bundle isomorphism. (In other words, the restrictions $(\iota_0)_x : V_x \to T_x\mathcal{M}^n$ are vector space isomorphisms depending smoothly on $x \in \mathcal{M}^n$.) Then if we define a metric h_0 on V by

$$h_0 \doteqdot \iota_0^* (g_0),$$

we automatically obtain a bundle isometry

$$\iota_0 : (V, h_0) \to (T\mathcal{M}^n, g_0).$$

Corresponding to the evolution of the metric $g(t)$ by the Ricci flow, we evolve the isometry $\iota(t)$ by

$$\frac{\partial}{\partial t}\iota = \mathrm{Rc} \circ \iota, \tag{6.19a}$$

$$\iota(0) = \iota_0. \tag{6.19b}$$

Here as in (6.18), we regard the Ricci tensor $\mathrm{Rc} \equiv \mathrm{Rc}[g(t)]$ as a $(1,1)$-tensor. For each $x \in \mathcal{M}^n$, equation (6.19) represents a system of linear ordinary differential equations. Hence a unique solution exists for $t \in [0,T)$ (namely, for as long as the solution $g(t)$ of the Ricci flow exists). Clearly, $\iota(t)$ remains a smooth bundle isomorphism for all $t \in [0,T)$. But more is true.

CLAIM 6.21. *Define $h(t) \doteqdot (\iota(t))^* (g(t))$. Then the bundle maps*

$$\iota(t) : (V, h(t)) \to (T\mathcal{M}^n, g(t))$$

remain isometries.

PROOF. Let $x \in \mathcal{M}^n$ and $X, Y \in V_x$ be arbitrary. Then recalling that g, Rc, and ι all depend on time, we compute that

$$\begin{aligned}
\frac{\partial}{\partial t} h(X,Y) &= \frac{\partial}{\partial t}\left[(\iota^* g)(X,Y)\right] \\
&= \frac{\partial}{\partial t}\left[g(\iota(X), \iota(Y))\right] \\
&= -2\,\mathrm{Rc}(\iota(X), \iota(Y)) \\
&\quad + g(\mathrm{Rc}(\iota(X)), \iota(Y)) + g(\iota(X), \mathrm{Rc}(\iota(Y))) \\
&= 0.
\end{aligned}$$

□

The Levi-Civita connections

$$\nabla(t) : C^\infty(T\mathcal{M}^n) \times C^\infty(T\mathcal{M}^n) \to C^\infty(T\mathcal{M}^n)$$

of the metrics $g(t)$ pull back to connections $D(t)$ on V defined by

$$D(t) \doteqdot \iota(t)^* \nabla(t) : C^\infty(T\mathcal{M}^n) \times C^\infty(V) \to C^\infty(V),$$

where for $X \in C^\infty(T\mathcal{M}^n)$ and $\xi \in C^\infty(V)$, we have

$$(\iota^*\nabla)(X,\xi) \doteqdot (\iota^*\nabla)_X(\xi) \doteqdot \nabla_X(\iota(\xi)).$$

Using the usual product rule, $\nabla(t)$ and $D(t)$ define connections on tensor product bundles of $T\mathcal{M}^n$, V, and their dual bundles $T^*\mathcal{M}^n$ and V^*. We denote these connections by ∇_D.

The next step in Uhlenbeck's trick is to pull back the Riemann curvature tensor to V. For $x \in \mathcal{M}^n$ and $X, Y, Z, W \in V_x$, the tensor

$$\iota^* \operatorname{Rm} \in C^\infty\left(\wedge^2 V \otimes_S \wedge^2 V\right)$$

is defined by

$$(6.20) \qquad (\iota^* \operatorname{Rm})(X, Y, Z, W) \doteqdot \operatorname{Rm}(\iota(X), \iota(Y), \iota(Z), \iota(W)).$$

Let $\left\{x^k\right\}_{k=1}^n$ be local coordinates defined on an open set $\mathcal{U} \subset \mathcal{M}^n$, and let $\{e_a\}_{a=1}^n$ be a basis of sections of V restricted to $\mathcal{U}$. The components $\left(\iota_a^k\right)$ of the bundle isomorphism $\iota : (V, h) \to (T\mathcal{M}^n, g(t))$ are defined by

$$\iota(e_a) \doteqdot \sum_{k=1}^n \iota_a^k \frac{\partial}{\partial x^k}.$$

Then the components (R_{abcd}) of $\iota^* \operatorname{Rm}$ are defined by

$$R_{abcd} \doteqdot (\iota^* \operatorname{Rm})(e_a, e_b, e_c, e_d) = \sum_{i,j,k,\ell=1}^n \iota_a^i \iota_b^j \iota_c^k \iota_d^\ell R_{ijk\ell}.$$

How is the evolution equation for $\iota^*(\operatorname{Rm})$ related to that for Rm? In order to answer this question, we define the Laplacian acting on tensor bundles of $T\mathcal{M}^n$ and V by

$$\Delta_D \doteqdot \operatorname{tr}_g(\nabla_D \circ \nabla_D) = \sum_{i,j=1}^n g^{ij} (\nabla_D)_i (\nabla_D)_j,$$

where $(\nabla_D)_j(\xi) = \nabla_j(\iota(\xi))$. We then get the following formula.

LEMMA 6.22. *If $g(t)$ is a solution of the Ricci flow and $\iota(t)$ is a solution of (6.19), then the curvature $\iota^* \operatorname{Rm}$ defined in (6.20) evolves by*

$$(6.21) \qquad \frac{\partial}{\partial t} R_{abcd} = \Delta_D R_{abcd} + 2\left(B_{abcd} - B_{abdc} + B_{acbd} - B_{adbc}\right),$$

where

$$B_{abcd} \doteqdot h^{eg} h^{fh} R_{aebf} R_{cgdh}.$$

PROOF. Observing that

$$\frac{\partial}{\partial t} \iota_a^k = R_\ell^k \iota_a^\ell,$$

we compute

$$\begin{aligned}
\frac{\partial}{\partial t} R_{abcd} &= \sum_{i,j,k,\ell=1}^{n} \frac{\partial}{\partial t}\left(\iota_a^i \iota_b^j \iota_c^k \iota_d^\ell R_{ijk\ell}\right) \\
&= \left(\frac{\partial}{\partial t}\iota_a^i\right) \iota_b^j \iota_c^k \iota_d^\ell R_{ijk\ell} + \iota_a^i \left(\frac{\partial}{\partial t}\iota_b^j\right) \iota_c^k \iota_d^\ell R_{ijk\ell} + \iota_a^i \iota_b^j \left(\frac{\partial}{\partial t}\iota_c^k\right) \iota_d^\ell R_{ijk\ell} \\
&\quad + \iota_a^i \iota_b^j \iota_c^k \left(\frac{\partial}{\partial t}\iota_d^\ell\right) R_{ijk\ell} + \iota_a^i \iota_b^j \iota_c^k \iota_d^\ell \left(\frac{\partial}{\partial t} R_{ijk\ell}\right) \\
&= R_m^i \iota_a^m \iota_b^j \iota_c^k \iota_d^\ell R_{ijk\ell} + \iota_a^i R_m^j \iota_b^m \iota_c^k \iota_d^\ell R_{ijk\ell} \\
&\quad + \iota_a^i \iota_b^j R_m^k \iota_c^m \iota_d^\ell R_{ijk\ell} + \iota_a^i \iota_b^j \iota_c^k R_m^\ell \iota_d^m R_{ijk\ell} \\
&\quad + \iota_a^i \iota_b^j \iota_c^k \iota_d^\ell \left\{ \begin{array}{c} \Delta R_{ijk\ell} + 2\left(B_{ijk\ell} - B_{ij\ell k} + B_{ikj\ell} - B_{i\ell jk}\right) \\ \\ -R_i^p R_{pjk\ell} - R_j^p R_{ipk\ell} - R_k^p R_{ijp\ell} - R_\ell^p R_{ijkp} \end{array} \right\} \\
&= \iota_a^i \iota_b^j \iota_c^k \iota_d^\ell \left[\Delta R_{ijk\ell} + 2\left(B_{ijk\ell} - B_{ij\ell k} + B_{ikj\ell} - B_{i\ell jk}\right)\right].
\end{aligned}$$

Since $\nabla_D(\iota) = 0$, we have $\left[\iota^*(\Delta Rm)\right]_{abcd} = \Delta_D R_{abcd}$. Also, one may easily check that $B_{abcd} = (\iota^* B)(e_a, e_b, e_c, e_d)$. The lemma follows. $\square$

3. The structure of the curvature evolution equation

By looking more closely at the structure of the Riemann curvature operator, we can obtain valuable insight into the structure of its evolution equation (6.21).

We begin with the observation that although B is quadratic in Rm, it is not the square of Rm. A natural way of defining the square is first to regard Rm as the operator on 2-forms

$$\mathrm{Rm} : \wedge^2 T^* \mathcal{M}^n \to \wedge^2 T^* \mathcal{M}^n$$

defined for all $U \in \wedge^2 T^* \mathcal{M}^n$ by

$$(\mathrm{Rm}(U))_{ij} \doteqdot -g^{kp} g^{\ell q} R_{ijk\ell} U_{pq}.$$

The operator Rm is then self adjoint with respect to the inner product defined for all $U, V \in \wedge^2 T^* \mathcal{M}^n$ by

$$\langle U, V \rangle \doteqdot g^{ik} g^{j\ell} U_{ij} V_{k\ell}; \tag{6.22}$$

indeed, it follows easily from symmetries of the Riemann curvature tensor that

$$\langle \mathrm{Rm}(U), V \rangle = \langle U, \mathrm{Rm}(V) \rangle.$$

Note that

$$\mathrm{Rm} = 2\kappa \, \mathrm{id}_{\wedge^2 T^* \mathcal{M}^n}$$

in case $(\mathcal{M}^n, g)$ has constant sectional curvature κ, because then

$$R_{ijk\ell} = \kappa \left(g_{i\ell} g_{jk} - g_{ik} g_{j\ell}\right),$$

whence

$$(\mathrm{Rm}(U))_{ij} = \kappa (U_{ij} - U_{ji}) = 2\kappa U_{ij}.$$

Now we can square the operator Rm to obtain the operator

$$\mathrm{Rm}^2 \equiv \mathrm{Rm} \circ \mathrm{Rm} : \wedge^2 T^* \mathcal{M}^n \to \wedge^2 T^* \mathcal{M}^n$$

given in local coordinates by

$$(\mathrm{Rm}^2)_{ijk\ell} = g^{pq} g^{rs} R_{ijps} R_{rqk\ell}. \tag{6.23}$$

REMARK 6.23. An equivalent theory results if one regards Rm and Rm^2 as symmetric bilinear forms on $\wedge^2 T\mathcal{M}^n$, hence as smooth sections of $\wedge^2 T^* \mathcal{M}^n \otimes_S \wedge^2 T^* \mathcal{M}^n$. This point of view will be useful below.

Although Rm^2 is the most natural definition of the square of the Riemann curvature operator, there is another concept of square which will be useful. This can be defined whenever one has a Lie algebra $\mathfrak{g}$ endowed with an inner product $\langle \cdot, \cdot \rangle$. Choose a basis $\{\varphi^\alpha\}$ of $\mathfrak{g}$ and let $C_\gamma^{\alpha\beta}$ denote the structure constants defined by

$$\left[\varphi^\alpha, \varphi^\beta\right] \doteqdot \sum_\gamma C_\gamma^{\alpha\beta} \varphi^\gamma,$$

where $[\cdot, \cdot]$ is the Lie bracket of $\mathfrak{g}$. Let $\{\varphi_\alpha^*\}$ denote the basis algebraically dual to $\{\varphi^\alpha\}$, so that $\varphi_\alpha^* \left(\varphi^\beta\right) = \delta_\alpha^\beta$. Given a symmetric bilinear form L on $\mathfrak{g}^*$, we may regard L as the element of $\mathfrak{g} \otimes_S \mathfrak{g}$ whose components are given by

$$L_{\alpha\beta} \doteqdot L\left(\varphi_\alpha^*, \varphi_\beta^*\right).$$

Then there is a commutative bilinear operation $\#$ defined for all $L, M \in \mathfrak{g} \otimes_S \mathfrak{g}$ by

$$(L \# M)_{\alpha\beta} \doteqdot C_\alpha^{\gamma\varepsilon} C_\beta^{\delta\zeta} L_{\gamma\delta} M_{\varepsilon\zeta}.$$

Abusing notation slightly, we define the **Lie algebra square** $L^\# \in \mathfrak{g} \otimes_S \mathfrak{g}$ of L by

$$(L^\#)_{\alpha\beta} \doteqdot (L \# L)_{\alpha\beta} = C_\alpha^{\gamma\delta} C_\beta^{\varepsilon\zeta} L_{\gamma\varepsilon} L_{\delta\zeta}. \tag{6.24}$$

The following property will be needed in showing that positivity of the Riemann curvature operator is preserved by the Ricci flow.

LEMMA 6.24. *If $L \geq 0$, then $L^\# \geq 0$.*

PROOF. Without loss of generality, we may choose a basis $\{\varphi^\alpha\}$ that diagonalizes L, so that $L_{\alpha\beta} = \delta_{\alpha\beta} L_{\alpha\alpha}$. Then for any $v = v^\alpha \varphi_\alpha^* \in \mathfrak{g}^*$, we have

$$L^\# (v, v) = \left(v^\alpha C_\alpha^{\gamma\delta}\right) \left(v^\beta C_\beta^{\varepsilon\zeta}\right) L_{\gamma\varepsilon} L_{\delta\zeta} = \left(v^\alpha C_\alpha^{\gamma\delta}\right)^2 L_{\gamma\gamma} L_{\delta\delta}.$$

□

REMARK 6.25. The inner product $\langle \cdot, \cdot \rangle$ of $\mathfrak{g}$ defines an metric isomorphism $\mathfrak{g} \to \mathfrak{g}^*$ by $v \mapsto \langle v, \cdot \rangle$. This isomorphism allows us to regard $L : \mathfrak{g} \to \mathfrak{g}$ as a self-adjoint endomorphism.

We now adopt this general definition to introduce another square of the operator Rm. Observe that for each $x \in \mathcal{M}^n$, the vector space $\wedge^2 T_x\mathcal{M}^n$ can be given the structure of a Lie algebra $\mathfrak{g}$ isomorphic to $\mathfrak{so}(n)$. Indeed, given $U, V \in \wedge^2 T^*\mathcal{M}^n$, we define their Lie bracket by

$$[U,V]_{ij} \doteqdot g^{k\ell}\left(U_{ik}V_{\ell j} - V_{ik}U_{\ell j}\right). \tag{6.25}$$

In a local orthonormal frame field $\{e_i\}$, any 2-form U may be naturally identified with an antisymmetric matrix (U_{ij}) so that formula (6.25) corresponds to

$$[U,V]_{ij} = (UV - VU)_{ij}.$$

This yields an evident Lie algebra isomorphism between $\mathfrak{g} \doteqdot \left(\wedge^2 T_x^*\mathcal{M}^n, [\cdot,\cdot]\right)$ and $\mathfrak{so}(n)$ for each $x \in \mathcal{M}^n$. We endow $\mathfrak{g}$ with the inner product defined by (6.22). Then the construction outlined above applies. In particular, in local coordinates $\left\{x^i\right\}$, there is a basis $\left\{\varphi^{(ij)} : 1 \le i < j \le n\right\}$ for $\mathfrak{g}$ defined by

$$\varphi^{(ij)} \doteqdot dx^i \wedge dx^j.$$

The corresponding structure constants $C_{(ij)}^{(pq),(rs)}$ are defined by

$$[dx^p \wedge dx^q, dx^r \wedge dx^s] = \sum_{(ij)} C_{(ij)}^{(pq),(rs)} dx^i \wedge dx^j.$$

Since

$$\begin{aligned} dx^p \wedge dx^q &= \frac{1}{2}\left(dx^p \otimes dx^q - dx^q \otimes dx^p\right) \\ &= \frac{1}{2}\left(\delta_k^p \delta_\ell^q - \delta_k^q \delta_\ell^p\right) dx^k \otimes dx^\ell, \end{aligned}$$

formula (6.25) implies that

$$\begin{aligned} C_{(ij)}^{(pq),(rs)} &= [dx^p \wedge dx^q, dx^r \wedge dx^s]_{ij} \\ &= \frac{1}{4} g^{k\ell} \left\{ \begin{array}{c} \left(\delta_i^p \delta_k^q - \delta_i^q \delta_k^p\right)\left(\delta_\ell^r \delta_j^s - \delta_\ell^s \delta_j^r\right) \\ -\left(\delta_i^r \delta_k^s - \delta_i^s \delta_k^r\right)\left(\delta_\ell^p \delta_j^q - \delta_\ell^q \delta_j^p\right) \end{array} \right\} \\ &= \frac{1}{4} \left\{ \begin{array}{c} g^{qr}\delta_i^p \delta_j^s - g^{qs}\delta_i^p \delta_j^r - g^{pr}\delta_i^q \delta_j^s + g^{ps}\delta_i^q \delta_j^r \\ -g^{sp}\delta_i^r \delta_j^q + g^{sq}\delta_i^r \delta_j^p + g^{rp}\delta_i^s \delta_j^q - g^{rq}\delta_i^s \delta_j^p \end{array} \right\} \\ &= \frac{1}{4} \left\{ \begin{array}{c} g^{qr}\left(\delta_i^p \delta_j^s - \delta_i^s \delta_j^p\right) + g^{qs}\left(\delta_i^r \delta_j^p - \delta_i^p \delta_j^r\right) \\ +g^{pr}\left(\delta_i^s \delta_j^q - \delta_i^q \delta_j^s\right) + g^{ps}\left(\delta_i^q \delta_j^r - \delta_i^r \delta_j^q\right) \end{array} \right\}. \end{aligned}$$

Then formula (6.24) yields

$$\left(\mathrm{Rm}^{\#}\right)_{ijk\ell} = R_{pquv}R_{rswx}C_{(ij)}^{(pq),(rs)}C_{(\ell k)}^{(uv),(wx)}. \tag{6.26}$$

Having completed these constructions, we can now write the evolution of the Riemann curvature operator in a way conducive to applying the tensor maximum principle.

THEOREM 6.26. *If $g(t)$ is a solution of the Ricci flow, the curvature $\iota^* \operatorname{Rm}$ defined in (6.20) evolves by*

$$\frac{\partial}{\partial t}(\iota^* \operatorname{Rm}) = \Delta_D \operatorname{Rm} + \operatorname{Rm}^2 + \operatorname{Rm}^\#. \tag{6.27}$$

PROOF. By exploiting appropriate antisymmetries, we get

$$R_{pquv} C_{(ij)}^{(pq),(rs)} = \frac{1}{2} R_{pquv} \left(g^{qr} \left(\delta_i^p \delta_j^s - \delta_i^s \delta_j^p \right) + g^{ps} \left(\delta_i^q \delta_j^r - \delta_i^r \delta_j^q \right) \right)$$

and

$$R_{pquv} R_{rswx} C_{(ij)}^{(pq),(rs)} R_{rswx} = R_{pquv} R_{rswx} g^{qr} \left(\delta_i^p \delta_j^s - \delta_i^s \delta_j^p \right).$$

So by (6.26), we have

$$\begin{aligned} \left(\operatorname{Rm}^\# \right)_{ijk\ell} &= R_{pquv} R_{rswx} C_{(ij)}^{(pq),(rs)} C_{(\ell k)}^{(uv),(wx)} \\ &= R_{pquv} R_{rswx} g^{qr} \left(\delta_i^p \delta_j^s - \delta_i^s \delta_j^p \right) g^{vw} \left(\delta_\ell^u \delta_k^x - \delta_\ell^x \delta_k^u \right) \\ &= R_{uvp}^q R_{sqx}^v \left(\delta_i^p \delta_j^s - \delta_i^s \delta_j^p \right) \left(\delta_\ell^u \delta_k^x - \delta_\ell^x \delta_k^u \right) \\ &= R_{\ell vi}^q R_{jqk}^v - R_{kvi}^q R_{jq\ell}^v - R_{\ell vj}^q R_{iqk}^v + R_{kvj}^q R_{iq\ell}^v. \end{aligned}$$

Now recalling that $B_{ijk\ell} = -R_{pij}^q R_{q\ell k}^p$ by definition (6.15), we invoke (6.16) in order to write $\operatorname{Rm}^\#$ as

$$\left(\operatorname{Rm}^\# \right)_{ijk\ell} = -B_{\ell ikj} + B_{ki\ell j} + B_{\ell jki} - B_{kj\ell i} = 2 \left(B_{ikj\ell} - B_{i\ell jk} \right).$$

On the other hand, applying the first Bianchi identity to formula (6.23) gives

$$\begin{aligned} \left(\operatorname{Rm}^2 \right)_{ijk\ell} &= g^{pq} g^{rs} R_{ijps} R_{qr\ell k} = g^{pq} g^{rs} \left(R_{ipsj} + R_{isjp} \right) \left(R_{krq\ell} + R_{kq\ell r} \right) \\ &= \left(R_{pij}^r - R_{pji}^r \right) \left(R_{rk\ell}^p - R_{r\ell k}^p \right) = -B_{ij\ell k} + B_{ijk\ell} + B_{ji\ell k} - B_{jik\ell} \\ &= 2 \left(B_{ijk\ell} - B_{ij\ell k} \right). \end{aligned}$$

So pulling back by the bundle isomorphism ι defined in (6.19), we have

$$\left(\operatorname{Rm}^2 \right)_{abcd} = 2 \left(B_{abcd} - B_{abdc} \right)$$

and

$$\left(\operatorname{Rm}^\# \right)_{abcd} = 2 \left(B_{acbd} - B_{adbc} \right).$$

Hence by Lemma (6.22), we conclude that

$$\begin{aligned} \left(\frac{\partial}{\partial t} - \Delta_D \right) R_{abcd} &= 2 \left(B_{abcd} - B_{abdc} - B_{adbc} + B_{acbd} \right) \\ &= \left(\operatorname{Rm}^2 \right)_{abcd} + \left(\operatorname{Rm}^\# \right)_{abcd}. \end{aligned}$$

□

Having written the evolution of the curvature operator in this form, we can immediately apply the tensor maximum principle for systems (discussed in Chapter 4) to get the following important result.

COROLLARY 6.27. *If $(\mathcal{M}^n, g(t))$ is a solution of the Ricci flow whose curvature operator is positive (negative) initially, then that condition is preserved for as long as the solution exists.*

4. Reduction to the associated ODE system

The maximum principle for systems introduced in Section 3 of Chapter 4 allows us to obtain qualitative information about the evolution of the Riemann curvature operator under the PDE (6.27) by studying associated systems (6.31) and (6.32) of ODE. This point of view originated in [**59**] and will be useful below in proving pinching estimates for the Ricci curvature of a solution to the Ricci flow on a closed 3-manifold with positive Ricci curvature.

Let $(\mathcal{M}^n, g)$ be a Riemannian manifold, and let $\{e_i\}$ be an orthonormal moving frame defined on an open set $\mathcal{U} \subseteq \mathcal{M}^n$. The frame $\{e_i\}$ induces an orthonormal basis $\left\{\theta^k = \theta^k_{ij}\, e_i \wedge e_j\right\}$ of $\wedge^2 T\mathcal{U}$. Hence for any $x \in \mathcal{U}$, there is a well defined Lie algebra isomorphism $\varphi_x : \wedge^2 T_x\mathcal{M}^n \longrightarrow \wedge^2\mathbb{R}^n$ taking the ordered basis $(\theta^1, \ldots, \theta^N)$ to an ordered basis $\beta = (\beta_1, \ldots, \beta_N)$ of $\wedge^2\mathbb{R}^n$, where $N = \binom{n}{2}$.

Let $\mathfrak{K}$ denote the space of those self-adjoint linear transformations $\mathbb{M} \in \mathrm{Sym}^2(\wedge^2\mathbb{R}^n)$ that obey the first Bianchi identity. The ODE corresponding to the reaction-diffusion PDE (6.27) satisfied by the curvature operator of a solution to the Ricci flow may then be written as

$$\frac{d}{dt}\mathbb{M} = \mathbb{M}^2 + \mathbb{M}^\#, \tag{6.28}$$

where $\mathbb{M}(0) = \mathbb{M}_0 \in \mathfrak{K}$ and $\mathbb{M}^\#$ is the Lie algebra square introduced in Section 3. The basis β allows one to represent any $\mathbb{M} \in \mathfrak{K}$ by an $N \times N$ symmetric matrix $\mathbb{M}_\beta$ defined by

$$\mathbb{M}(\beta_j) = \beta_i (\mathbb{M}_\beta)_{ij},$$

with the obvious summation in effect. The structure constants C_β for $\wedge^2 T_x\mathcal{M}^n \cong \wedge^2\mathbb{R}^n \cong \mathrm{so}(n)$ with respect to the basis β are defined by

$$[\beta_j, \beta_k] = \beta_i (C_\beta)_{ijk}.$$

In terms of the basis β, the transformation $\mathbb{M} \mapsto \mathbb{Q} \doteqdot \mathbb{M}^2 + \mathbb{M}^\#$ is then given by

$$(\mathbb{Q}_\beta)_{ij} = (\mathbb{M}_\beta)_{ik} (\mathbb{M}_\beta)_{kj} + (C_\beta)_{ipq} (C_\beta)_{jrs} (\mathbb{M}_\beta)_{pr} (\mathbb{M}_\beta)_{qs}.$$

In what follows, we shall suppress dependence on the basis β.

Let us now specialize to the case that $\mathcal{M}^3$ is a closed 3-manifold. Since $\mathcal{M}^3$ is parallelizable (as we noted in Remark 6.20) we may assume there exists a globally defined orthonormal moving frame $\{e_i\}$. We fix an orthonormal basis $\left\{\theta^k = \theta^k_{ij}\, e_i \wedge e_j\right\}$ of $\wedge^2 T\mathcal{M}^3$. For example, such a basis is given by taking

$$\theta^1 = \frac{1}{\sqrt{2}} * (e_1) = \frac{1}{\sqrt{2}} (e_2 \wedge e_3) \sim \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 0 \end{pmatrix},$$

$$\theta^2 = \frac{1}{\sqrt{2}} * (e_2) = \frac{1}{\sqrt{2}} (e_3 \wedge e_1) \sim \begin{pmatrix} 0 & 0 & -1/\sqrt{2} \\ 0 & 0 & 0 \\ 1/\sqrt{2} & 0 & 0 \end{pmatrix},$$

$$\theta^3 = \frac{1}{\sqrt{2}} * (e_3) = \frac{1}{\sqrt{2}} (e_1 \wedge e_2) \sim \begin{pmatrix} 0 & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where $* : \wedge^1 T\mathcal{M}^3 \to \wedge^2 T\mathcal{M}^3$ corresponds to the Hodge star operator. In dimension $n = 3$, the Lie algebra square is easily computed. Indeed, it is readily verified that $\left\langle \left[\theta^i, \theta^j\right], \theta^k \right\rangle$ is fully alternating in (i, j, k), hence that

$$\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}^{\#} = \begin{pmatrix} df - e^2 & ce - bf & be - cd \\ ce - bf & af - c^2 & bc - ae \\ be - cd & bc - ae & ad - b^2 \end{pmatrix}. \tag{6.29}$$

Now we identify Rm with the quadratic form $\mathbb{M}$ defined on $\wedge^2 T\mathcal{M}^3$ by

$$\mathbb{M}\left(e_i \wedge e_j, e_\ell \wedge e_k\right) = \left\langle R\left(e_i, e_j\right) e_k, e_\ell \right\rangle.$$

Using the basis $\left\{\theta^1, \theta^2, \theta^3\right\}$ of $\wedge^2 T\mathcal{U}$, we further identify $\mathbb{M}$ with the matrix $(\mathbb{M}_{pq})$ defined on each fiber $\wedge^2 T_x\mathcal{M}^3$ of the bundle $\wedge^2 T\mathcal{M}^3$ by

$$\left\langle R\left(e_i, e_j\right) e_k, e_\ell \right\rangle = \mathbb{M}_{pq} \theta^p_{ij} \theta^q_{\ell k}. \tag{6.30}$$

If $\{e_i\}$ evolves to remain orthonormal, the PDE (6.27) governing the behavior of Rm corresponds to the ODE

$$\frac{d}{dt}\mathbb{M} = \mathbb{M}^2 + \mathbb{M}^{\#} \tag{6.31}$$

satisfied by $\mathbb{M}$ in each fiber. If $\{e_i\}$ is chosen so that $\mathbb{M}_0$ is diagonal at $x \in \mathcal{M}^3$ with eigenvectors $\lambda(0) \geq \mu(0) \geq \nu(0)$, then (6.29) and (6.31) combine to yield the system

$$\frac{d}{dt}\begin{pmatrix} \lambda & & \\ & \mu & \\ & & \nu \end{pmatrix} = \begin{pmatrix} \lambda^2 & & \\ & \mu^2 & \\ & & \nu^2 \end{pmatrix} + \begin{pmatrix} \mu\nu & & \\ & \lambda\nu & \\ & & \lambda\mu \end{pmatrix} \tag{6.32}$$

posed on $\mathbb{R}^3$. In particular, $\mathbb{M}(t)$ remains diagonal (which is not in general the case in higher dimensions). Moreover, $\lambda(t) \geq \mu(t) \geq \nu(t)$ for all $t \geq 0$

such that a solution exists, because

$$\frac{d}{dt}(\lambda-\mu)=(\lambda-\mu)(\lambda+\mu-\nu)$$
$$\frac{d}{dt}(\mu-\nu)=(\mu-\nu)(-\lambda+\mu+\nu).$$

By (6.30), the eigenvalues λ, μ, ν of $\mathbb{M}$ are twice the sectional curvatures. Hence at $x\in\mathcal{M}^3$, we may regard the Ricci tensor as the matrix

$$\mathrm{Rc}=\frac{1}{2}\begin{pmatrix}\mu+\nu & & \\ & \lambda+\nu & \\ & & \lambda+\mu\end{pmatrix}. \tag{6.33}$$

For later reference, we note that the trace-free parts of Rm and Rc are related at the same point by

$$\overset{\circ}{\mathrm{Rm}}=\frac{1}{3}\begin{pmatrix}2\lambda-\mu-\nu & & \\ & -\lambda+2\mu-\nu & \\ & & -\lambda-\mu+2\nu\end{pmatrix}=-2\overset{\circ}{\mathrm{Rc}}. \tag{6.34}$$

5. Local pinching estimates

Now we are ready to state the pinching results which are true for 3-manifolds with positive Ricci curvature. The first estimate proves that curvature pinching is preserved, whereas the second shows that it improves, hence that a solution to the Ricci flow on a 3-manifold with positive Ricci curvature is nearly Einstein at any point where its scalar curvature is large. These are pointwise estimates; below, we shall develop techniques to compare curvatures at different points of a solution.

As in Section 4, we let $\lambda(t)\geq\mu(t)\geq\nu(t)$ denote the eigenvalues of the curvature operator Rm of a solution $(\mathcal{M}^3,g(t))$ to the Ricci flow on a closed 3-manifold, recalling that these are twice the sectional curvatures.

LEMMA 6.28 (Ricci pinching is preserved.). *Let $(\mathcal{M}^3,g(t))$ be a solution of the Ricci flow on a closed 3-manifold such that the initial metric g_0 has strictly positive Ricci curvature. If there exists a constant $C<\infty$ such that*

$$\lambda\leq C(\nu+\mu) \tag{6.35}$$

at $t=0$, then this inequality persists as long as the solution exists.

PROOF. Recall that $\lambda\geq\frac{1}{2}(\nu+\mu)>0$, so that we can take logarithms. Compute

$$\begin{aligned}\frac{d}{dt}\log\left(\frac{\lambda}{\nu+\mu}\right)&=\frac{1}{\lambda(\nu+\mu)}\left((\nu+\mu)\frac{d\lambda}{dt}-\lambda\frac{d}{dt}(\nu+\mu)\right)\\&=\frac{1}{\lambda(\nu+\mu)}\left((\nu+\mu)\left(\lambda^2+\mu\nu\right)-\lambda\left(\nu^2+\lambda\mu+\mu^2+\lambda\nu\right)\right)\\&=\frac{\mu^2(\nu-\lambda)+\nu^2(\mu-\lambda)}{\lambda(\nu+\mu)}\leq 0.\end{aligned} \tag{6.36}$$

Let $\lambda(\mathbb{P}) \geq \mu(\mathbb{P}) \geq \nu(\mathbb{P})$ denote the eigenvalues of

$$\mathbb{P} \in \left(\wedge^2 T^* \mathcal{M}^3 \otimes_S \wedge^2 T^* \mathcal{M}^3\right)_x .$$

Define $\mathcal{K} \subset \left(\wedge^2 T^* \mathcal{M}^3 \otimes_S \wedge^2 T^* \mathcal{M}^3\right)_x$ by

$$\mathcal{K} \doteqdot \{\mathbb{P} : \lambda(\mathbb{P}) - C(\nu(\mathbb{P}) + \mu(\mathbb{P})) \leq 0\}.$$

It is easy to see that $\mathcal{K}$ is invariant under parallel translation. $\mathcal{K}$ is convex in each fiber, because the function

$$\lambda(\mathbb{P}) - C(\nu(\mathbb{P}) + \mu(\mathbb{P})) = \max_{|U|=1} \mathbb{P}(U,U) + C \max_{\substack{|V|=|W|=1 \\ \langle V,W\rangle = 0}} [-\mathbb{P}(V,V) - \mathbb{P}(W,W)]$$

is convex. Let $\mathbb{M}(t) \in \left(\wedge^2 T^* \mathcal{M}^3 \otimes_S \wedge^2 T^* \mathcal{M}^3\right)_x$ be the quadratic form corresponding to $\operatorname{Rm}[g(t)]$, and choose $C < \infty$ so large that $\mathbb{M}(0) \in \mathcal{K}$ for all $x \in \mathcal{M}^3$. By (6.36), $\mathbb{M}$ remains in $\mathcal{K}$. □

This estimate implies a global bound which will be useful later.

COROLLARY 6.29. *Let $(\mathcal{M}^3, g(t))$ be a solution of the Ricci flow on a closed 3-manifold of initially positive Ricci curvature, and define $R_{\min}(t) \doteqdot \inf_{x \in \mathcal{M}^3} R(x,t)$. Then there exists $\beta > 0$ depending only on g_0 such that at all points of $\mathcal{M}^3$,*

$$\operatorname{Rc} \geq 2\beta^2 R g \geq 2\beta^2 R_{\min} g.$$

PROOF. By the lemma and formula (6.33), there exists $C > 0$ depending only on g_0 such that the inequality

$$\operatorname{Rc} \geq \frac{\mu+\nu}{2} g \geq \frac{\lambda}{2C} g \geq \frac{\lambda+\mu+\nu}{6C} g = \frac{R}{6C} g \geq \frac{R_{\min}}{6C} g$$

holds everywhere on $\mathcal{M}^3$. □

Lemma 6.28 also helps us prove the following result, which shows that the metric is nearly Einstein at points where the scalar curvature is large.

THEOREM 6.30 (Ricci pinching is improved.). *Let $(\mathcal{M}^3, g(t))$ be a solution of the Ricci flow on a closed 3-manifold such that g_0 has strictly positive Ricci curvature. Then there exist positive constants $\delta < 1$ and C depending only on g_0 such that*

$$\frac{\lambda - \nu}{\nu + \mu} \leq \frac{C}{(\nu+\mu)^\delta}.$$

In particular, the left-hand side is invariant under homotheties of the metric, while the right-hand side tends to 0 as $\nu + \mu \to \infty$.

REMARK 6.31. Theorem 6.30 is equivalent to the following statement, which was proved in Hamilton's original paper [**58**]: *There exist constants $\bar{\delta} > 0$ and $C < \infty$ depending only on g_0 such that*

$$\frac{|\overset{\circ}{\operatorname{Rm}}|^2}{R^2} = 4\frac{|\overset{\circ}{\operatorname{Rc}}|^2}{R^2} \leq C R^{-\bar{\delta}}, \tag{6.37}$$

where as in (6.34), $\overset{\circ}{\mathrm{Rm}}$ *and* $\overset{\circ}{\mathrm{Rc}}$ *are the trace-free parts of the Riemann curvature operator and Ricci tensor, respectively.* Once again, the left-hand side is scale invariant, while the right-hand side tends to 0 uniformly if $\inf_{x \in \mathcal{M}^3} R(x,t) \to \infty$.

We shall give two proofs of the theorem: the first uses the maximum principle for systems and is in the spirit of [**59**], whereas the second is closer in spirit to the original argument of [**58**].

First proof of Theorem 6.30. Because

$$\frac{d}{dt}(\lambda - \nu) = (\lambda - \nu)(\lambda - \mu + \nu),$$

we may assume that $\lambda(\mathbb{M}) > \nu(\mathbb{M})$ for $\mathbb{M}(t) \in \left(\wedge^2 T^* \mathcal{M}^3 \otimes_S \wedge^2 T^* \mathcal{M}^3\right)_x$. We calculate

$$\frac{d}{dt}\log(\lambda - \nu) = \lambda - \mu + \nu$$

and

$$\frac{d}{dt}\log(\nu + \mu) = \lambda + \frac{\nu^2 + \mu^2}{\nu + \mu}.$$

Hence

$$\begin{aligned}
\frac{d}{dt}\log\left[\frac{\lambda - \nu}{(\nu + \mu)^{1-\delta}}\right] &= (\lambda - \mu + \nu) - (1 - \delta)\left(\lambda + \frac{\nu^2 + \mu^2}{\nu + \mu}\right) \\
&= \delta\lambda + (\nu - \mu) - (1 - \delta)\frac{\nu^2 + \mu^2}{\nu + \mu} \\
&\leq \delta\lambda - \frac{1}{2}(1 - \delta)(\nu + \mu),
\end{aligned}$$

where we used $\nu \leq \mu$ and $\nu^2 + \mu^2 \geq \frac{1}{2}(\nu + \mu)^2$ to obtain the last inequality. By Lemma 6.28, the right-hand side is nonpositive for δ chosen small enough. Now define the convex set

$$\mathcal{K} = \left\{\mathbb{P} : [\lambda(\mathbb{P}) - \nu(\mathbb{P})] - C[\nu(\mathbb{P}) + \mu(\mathbb{P})]^{1-\delta} \leq 0\right\}$$

and continue as before. □

Second proof of Theorem 6.30. By Lemma 6.34 (below), the function

$$f \doteqdot \frac{|\overset{\circ}{\mathrm{Rm}}|^2}{R^{2-\varepsilon}}$$

satisfies the evolution equation

$$\frac{\partial f}{\partial t} \leq \Delta f + 2(1 - \varepsilon)\langle \nabla f, \nabla(\log R)\rangle + 2Q. \tag{6.38}$$

Here

$$Q \doteqdot \frac{1}{R^{3-\varepsilon}}\left(\varepsilon|\mathrm{Rm}|^2 |\overset{\circ}{\mathrm{Rm}}|^2 - P\right),$$

where P may be written in terms of the eigenvalues λ, μ, ν of Rm as

$$P \doteqdot \lambda^2 (\mu - \nu)^2 + \mu^2 (\lambda - \nu)^2 + \nu^2 (\lambda - \mu)^2 \geq 0. \tag{6.39}$$

By the maximum principle, we only need to show that $Q \leq 0$ to establish the lemma. Notice that the good term is $-P$, while the bad term $|\mathrm{Rm}|^2 |\overset{\circ}{\mathrm{Rm}}|^2$ is scaled by ε. So to prove $Q \leq 0$, one first observes that $P \geq \delta^2 |\mathrm{Rm}|^2 |\overset{\circ}{\mathrm{Rm}}|^2$ if $\mathrm{Rc} \geq \delta R \cdot g$ for some constant $\delta > 0$. Then one uses Corollary 6.28 to conclude that the inequality $\mathrm{Rc} \geq \delta R \cdot g$ is preserved if it holds at $t = 0$. Thus we can take $\varepsilon = \delta^2$ for a sufficiently small positive constant $\delta = \delta(g_0)$. □

REMARK 6.32. The hypothesis of strictly positive Ricci curvature is necessary for either proof to work, since $Q > 0$ if $\lambda > 0$ and $\mu = \nu = 0$ at a point. This failure is related to the fact that $\mathcal{S}^1 \times \mathcal{S}^2$ accepts no metric of positive Ricci curvature. (Indeed, by Myers' Theorem, a complete manifold $\mathcal{M}^n$ admits such a metric only if its universal cover is compact.)

Equation (6.38) is derived in the following two lemmas:

LEMMA 6.33. *If φ is a nonnegative function and ψ is a positive function on space and time, then*

$$\begin{aligned}\left(\frac{\partial}{\partial t} - \Delta\right)\left(\frac{\varphi^\alpha}{\psi^\beta}\right) &= \alpha \frac{\varphi^{\alpha-1}}{\psi^\beta}\left(\frac{\partial}{\partial t} - \Delta\right)\varphi - \beta \frac{\varphi^\alpha}{\psi^{\beta+1}}\left(\frac{\partial}{\partial t} - \Delta\right)\psi \\ &\quad - \alpha(\alpha-1)\frac{\varphi^{\alpha-2}}{\psi^\beta}|\nabla\varphi|^2 - \beta(\beta+1)\frac{\varphi^\alpha}{\psi^{\beta+2}}|\nabla\psi|^2 \\ &\quad + 2\alpha\beta\frac{\varphi^{\alpha-1}}{\psi^{\beta+1}}\langle\nabla\varphi, \nabla\psi\rangle.\end{aligned}$$

PROOF. Straightforward calculation. □

LEMMA 6.34. *The function $f \doteqdot R^{\varepsilon-2}|\overset{\circ}{\mathrm{Rm}}|^2$ satisfies the differential inequality*

$$\frac{\partial f}{\partial t} \leq \Delta f + \frac{2(1-\varepsilon)}{R}\langle\nabla f, \nabla R\rangle + 2Q.$$

PROOF. Recall that in the time-dependent orthonormal moving frame $\{e_i\}$, we have

$$\frac{\partial}{\partial t}\mathrm{Rm} = \Delta\,\mathrm{Rm} + \mathrm{Rm}^2 + \mathrm{Rm}^\#.$$

Hence

$$\begin{aligned}\frac{\partial}{\partial t}|\mathrm{Rm}|^2 &= 2\left\langle\Delta\,\mathrm{Rm} + \mathrm{Rm}^2 + \mathrm{Rm}^\#, \mathrm{Rm}\right\rangle \\ &= \Delta|\mathrm{Rm}|^2 - 2|\nabla\,\mathrm{Rm}|^2 + 2\left\langle\mathrm{Rm}^2 + \mathrm{Rm}^\#, \mathrm{Rm}\right\rangle \\ &= \Delta|\mathrm{Rm}|^2 - 2|\nabla\,\mathrm{Rm}|^2 + 2\left(\lambda^3 + \mu^3 + \nu^3 + 3\lambda\mu\nu\right).\end{aligned}$$

Since the scalar curvature is $R = \operatorname{tr}(\mathrm{Rm}) = \operatorname{tr}(\mathrm{Rc}) = \lambda + \mu + \nu$, it follows from (6.33) that

$$\frac{\partial}{\partial t} R = \Delta R + 2\,|\mathrm{Rc}|^2 = \Delta R + \lambda^2 + \mu^2 + \nu^2 + \lambda\mu + \lambda\nu + \mu\nu, \tag{6.40}$$

hence that

$$\begin{aligned}\frac{\partial}{\partial t}\left(R^2\right) &= \Delta\left(R^2\right) - 2\,|\nabla R|^2 \\ &\quad + 2\,(\lambda+\mu+\nu)\left(\lambda^2+\mu^2+\nu^2+\lambda\mu+\lambda\nu+\mu\nu\right).\end{aligned}$$

Now since $|\mathring{\mathrm{Rm}}|^2 = |\mathrm{Rm}|^2 - R^2/3$, we compute that

$$\begin{aligned}\frac{\partial}{\partial t}|\mathring{\mathrm{Rm}}|^2 &= \Delta\,|\mathrm{Rm}|^2 - \frac{1}{3}\Delta\left(R^2\right) - 2\,|\nabla \mathrm{Rm}|^2 + \frac{2}{3}\,|\nabla R|^2 \\ &\quad + 2\left(\lambda^3+\mu^3+\nu^3+3\lambda\mu\nu\right) \\ &\quad - \frac{2}{3}\,(\lambda+\mu+\nu)\left(\lambda^2+\mu^2+\nu^2+\lambda\mu+\lambda\nu+\mu\nu\right) \\ &= \Delta|\mathring{\mathrm{Rm}}|^2 - 2|\nabla\mathring{\mathrm{Rm}}|^2 \\ &\quad + \frac{4}{3}\left(\lambda^3+\mu^3+\nu^3+3\lambda\mu\nu\right) \\ &\quad - \frac{4}{3}\left(\lambda^2\mu+\lambda^2\nu+\mu^2\lambda+\mu^2\nu+\nu^2\lambda+\nu^2\mu\right).\end{aligned}$$

Taking $\varphi = |\mathring{\mathrm{Rm}}|^2$, $\psi = R$, $\alpha = 1$, and $\beta = 2-\varepsilon$ in Lemma 6.33, we get

$$\begin{aligned}\frac{\partial}{\partial t} f &= \Delta f - \frac{2}{R^{2-\varepsilon}}|\nabla\mathring{\mathrm{Rm}}|^2 \\ &\quad + \frac{4}{3R^{2-\varepsilon}}\left\{\begin{array}{c}\lambda^3+\mu^3+\nu^3+3\lambda\mu\nu \\ -\left(\lambda^2\mu+\lambda^2\nu+\mu^2\lambda+\mu^2\nu+\nu^2\lambda+\nu^2\mu\right)\end{array}\right\} \\ &\quad - (2-\varepsilon)\,\frac{|\mathring{\mathrm{Rm}}|^2}{R^{3-\varepsilon}}\left(\lambda^2+\mu^2+\nu^2+\lambda\mu+\lambda\nu+\mu\nu\right) \\ &\quad - (2-\varepsilon)\,(3-\varepsilon)\,\frac{|\mathring{\mathrm{Rm}}|^2\,|\nabla R|^2}{R^{4-\varepsilon}} + \frac{2\,(2-\varepsilon)}{R^{3-\varepsilon}}\left\langle \nabla|\mathring{\mathrm{Rm}}|^2, \nabla R\right\rangle.\end{aligned}$$

Simplifying and combining terms, we obtain

$$\begin{aligned}\frac{\partial}{\partial t} f &= \Delta f + \frac{2\,(1-\varepsilon)}{R}\left\langle \nabla\left(\frac{|\mathring{\mathrm{Rm}}|^2}{R^{2-\varepsilon}}\right), \nabla R\right\rangle + 2\varepsilon\frac{|\mathring{\mathrm{Rm}}|^2\,|\mathrm{Rc}|^2}{R^{3-\varepsilon}} \\ &\quad - \frac{2}{R^{4-\varepsilon}}\left|R\left(\nabla\mathring{\mathrm{Rm}}\right) - \nabla R\otimes\mathring{\mathrm{Rm}}\right|^2 - \varepsilon\,(1-\varepsilon)\,\frac{|\mathring{\mathrm{Rm}}|^2\,|\nabla R|^2}{R^{4-\varepsilon}} \\ &\quad - \frac{4}{R^{3-\varepsilon}}\left[\lambda^2\mu^2+\lambda^2\nu^2+\mu^2\nu^2-\left(\lambda^2\mu\nu+\mu^2\lambda\nu+\nu^2\lambda\mu\right)\right].\end{aligned}$$

The result follows when we note that $|\mathrm{Rc}|^2 \leq |\mathrm{Rm}|^2$, discard the negative terms on the second line, and compute

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

□

6. The gradient estimate for the scalar curvature

In this section we obtain a gradient estimate for the scalar curvature. This estimate is important because it enables us to compare curvatures at different points, whereas the pinching estimate of Section 5 is a pointwise estimate which compares curvatures at the same point.

As motivation, let us recall **Schur's Lemma**. If g is an Einstein metric on a manifold $\mathcal{M}^n$, then $\mathrm{Rc} = f \cdot g$ for some function f on $\mathcal{M}^n$ and one has

$$\nabla_k R = \nabla_k \left(g^{ij} R_{ij} \right) = n \nabla_k f.$$

On the other hand, the contracted second Bianchi identity implies that

$$\nabla_k R = 2 \nabla^j R_{jk} = 2 \nabla^j \left(f g_{jk} \right) = 2 \nabla_k f.$$

So we have $(n-2) \nabla f \equiv 0$, which shows that $f = R/n$ is constant in dimensions $n > 2$.

Now suppose we have a solution $\left(\mathcal{M}^3, g(t) \right)$ of the Ricci flow on a closed 3-manifold. The pinching estimate (6.37) proved in Section 5 can be written in the form

$$\frac{\left| \mathrm{Rc} - \frac{1}{3} R g \right|^2}{R^2} \leq C R^{-\bar{\delta}}, \tag{6.41}$$

where the left-hand side is a scale-invariant quantity which measures how far the metric is from being Einstein, and the right-hand side is small when the scalar curvature is large. So if the scalar curvature becomes uniformly large, it is natural to expect its gradient to approach zero uniformly. It is also natural to expect that the contracted second Bianchi identity will be key to proving this result. Both expectations are in fact correct, and we shall now establish the following result.

THEOREM 6.35. *Let $\left(\mathcal{M}^3, g(t) \right)$ be a solution of the Ricci flow on a closed 3-manifold with $g(0) = g_0$. If $\mathrm{Rc}(g_0) > 0$, then there exist $\bar{\beta}, \bar{\delta} > 0$ depending only on g_0 such that for any $\beta \in \left[0, \bar{\beta} \right]$, there exists C depending only on β and g_0 such that*

$$\frac{|\nabla R|^2}{R^3} \leq \beta R^{-\bar{\delta}/2} + C R^{-3}.$$

Here, the left-hand side is a scale invariant quantity, while the right-hand side is small when the scalar curvature is large.

To prove this estimate, we need to compute several evolution equations, the first of which is for the square of the norm of the gradient of the scalar curvature.

LEMMA 6.36. *If $(\mathcal{M}^n, g(t))$ is a solution of the Ricci flow, then*

$$\frac{\partial}{\partial t} |\nabla R|^2 = \Delta |\nabla R|^2 - 2|\nabla\nabla R|^2 + 4\left\langle \nabla R, \nabla |\mathrm{Rc}|^2 \right\rangle. \tag{6.42}$$

PROOF. By (6.6), one has

$$\begin{aligned}\frac{\partial}{\partial t} |\nabla R|^2 &= \frac{\partial}{\partial t}\left(g^{ij}\nabla_i R \nabla_j R\right) \\ &= 2\,\mathrm{Rc}(\nabla R, \nabla R) + 2\left\langle \nabla R, \nabla\left(\Delta R + 2|\mathrm{Rc}|^2\right)\right\rangle,\end{aligned}$$

whence the lemma follows from the Bochner–Weitzenböck formula

$$\Delta |\nabla R|^2 = 2|\nabla\nabla R|^2 + 2\langle \nabla R, \Delta\nabla R\rangle + 2\,\mathrm{Rc}(\nabla R, \nabla R).$$

□

Next we calculate the evolution of $|\nabla R|^2$ divided by R. This choice of power is not strongly intuitive, but yields a useful equation.

LEMMA 6.37. *Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow such that $R > 0$ initially. Then for as long as the solution exists,*

$$\frac{\partial}{\partial t}\left(\frac{|\nabla R|^2}{R}\right) = \Delta\left(\frac{|\nabla R|^2}{R}\right) - 2R\left|\nabla\left(\frac{\nabla R}{R}\right)\right|^2 \tag{6.43a}$$

$$- 2\frac{|\nabla R|^2}{R^2}|\mathrm{Rc}|^2 + \frac{4}{R}\left\langle \nabla R, \nabla |\mathrm{Rc}|^2\right\rangle. \tag{6.43b}$$

PROOF. By Corollary 6.8, the inequality $R > 0$ is preserved. Using (6.6) and (6.42), one computes that

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{|\nabla R|^2}{R}\right) &= \frac{1}{R}\cdot\frac{\partial}{\partial t}|\nabla R|^2 - \frac{|\nabla R|^2}{R^2}\cdot\frac{\partial}{\partial t}R \\ &= \frac{1}{R}\left(\Delta |\nabla R|^2 - 2|\nabla\nabla R|^2 + 4\left\langle \nabla R, \nabla |\mathrm{Rc}|^2\right\rangle\right) \\ &\quad - \frac{|\nabla R|^2}{R^2}\left(\Delta R + 2|\mathrm{Rc}|^2\right).\end{aligned}$$

Since for any smooth functions u and v,

$$\Delta\left(\frac{u}{v}\right) = \frac{\Delta u}{v} - \frac{u\Delta v}{v^2} - \frac{2}{v^2}\langle \nabla u, \nabla v\rangle + 2\frac{u}{v^3}|\nabla v|^2,$$

we have

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{|\nabla R|^2}{R}\right) &= \Delta\left(\frac{|\nabla R|^2}{R}\right) - 2\left(\frac{|\nabla R|^4}{R^3} - \frac{\left\langle \nabla |\nabla R|^2, \nabla R\right\rangle}{R^2} + \frac{|\nabla\nabla R|^2}{R}\right) \\ &\quad - \frac{2}{R^2}|\nabla R|^2|\mathrm{Rc}|^2 + \frac{4}{R}\left\langle \nabla R, \nabla |\mathrm{Rc}|^2\right\rangle.\end{aligned}$$

The lemma follows by completing the square. □

The evolution equation for $\frac{|\nabla R|^2}{R}$ has only one bad (potentially positive) term on the right-hand side, namely $\frac{4}{R}\left\langle \nabla R, \nabla |\mathrm{Rc}|^2 \right\rangle$. In dimension $n = 3$, we can remedy this by adding the quantity $|\mathrm{Rc}|^2 - \frac{1}{3}R^2$ to $\frac{|\nabla R|^2}{R}$, thereby introducing a good (negative) term that cancels out the bad term. Before proving this, we compute two more evolution equations valid for a general n-dimensional solution.

LEMMA 6.38. *If $(\mathcal{M}^n, g(t))$ is a solution of the Ricci flow, then*

$$\frac{\partial}{\partial t} R^2 = \Delta \left(R^2\right) - 2|\nabla R|^2 + 4R|\mathrm{Rc}|^2 \tag{6.44}$$

and

$$\frac{\partial}{\partial t}|\mathrm{Rc}|^2 = \Delta |\mathrm{Rc}|^2 - 2|\nabla \mathrm{Rc}|^2 + 4R_{ijk\ell}R^{i\ell}R^{jk}. \tag{6.45}$$

PROOF. It follows from (6.6) that the square of the scalar curvature evolves by

$$\begin{aligned} \frac{\partial}{\partial t} R^2 &= 2R\left(\Delta R + 2|\mathrm{Rc}|^2\right) \\ &= \Delta\left(R^2\right) - 2|\nabla R|^2 + 4R|\mathrm{Rc}|^2. \end{aligned}$$

On the other hand, (6.7) implies that the square of the norm of the Ricci tensor evolves by

$$\begin{aligned} \frac{\partial}{\partial t}|\mathrm{Rc}|^2 &= \frac{\partial}{\partial t}\left(g^{ij}g^{k\ell}R_{ik}R_{j\ell}\right) \\ &= 4\operatorname{tr}_g\left(\mathrm{Rc}^3\right) + 2\left\langle \mathrm{Rc}, \Delta_L \mathrm{Rc}\right\rangle \\ &= \Delta|\mathrm{Rc}|^2 - 2|\nabla \mathrm{Rc}|^2 + 4R_{ijk\ell}R^{i\ell}R^{jk}. \end{aligned}$$

□

COROLLARY 6.39. *If $\left(\mathcal{M}^3, g(t)\right)$ is a solution of the Ricci flow on a 3-manifold, then*

$$\begin{aligned} \frac{\partial}{\partial t}\left(|\mathrm{Rc}|^2 - \frac{1}{3}R^2\right) &= \Delta\left(|\mathrm{Rc}|^2 - \frac{1}{3}R^2\right) - 2\left(|\nabla \mathrm{Rc}|^2 - \frac{1}{3}|\nabla R|^2\right) \\ &\quad - 8\operatorname{tr}_g(\mathrm{Rc}^3) + \frac{26}{3}R\,|\mathrm{Rc}|^2 - 2R^3. \end{aligned}$$

PROOF. Formula (6.9) implies that (6.45) can be written in dimension $n = 3$ as

$$\frac{\partial}{\partial t}|\mathrm{Rc}|^2 = \Delta|\mathrm{Rc}|^2 - 2|\nabla \mathrm{Rc}|^2 - 2R^3 - 8\operatorname{tr}_g\left(\mathrm{Rc}^3\right) + 10R|\mathrm{Rc}|^2.$$

Alternatively, one may obtain this formula by using (6.10). □

The useful term in the evolution equation for $|\mathrm{Rc}|^2 - \frac{1}{3}R^2$ is the quantity $-2(|\nabla \mathrm{Rc}|^2 - \frac{1}{3}|\nabla R|^2)$. Our next task is to show that this term dominates the bad term $\frac{4}{R}\langle \nabla R, \nabla |Rc|^2\rangle$ in the evolution equation (6.43) for $\frac{|\nabla R|^2}{R}$. It is easy to see that $|\nabla \mathrm{Rc}|^2 - \frac{1}{3}|\nabla R|^2 \geq 0$, but we need a slightly better estimate.

LEMMA 6.40. *In dimension* $n = 3$,

$$|\nabla \mathrm{Rc}|^2 - \frac{1}{3}|\nabla R|^2 \geq \frac{1}{37}|\nabla \mathrm{Rc}|^2 .$$

PROOF. Define a $(3,0)$-tensor X by $X_{ijk} \doteqdot \nabla_i R_{jk} - \frac{1}{3}\nabla_i R\, g_{jk}$ and a $(1,0)$-tensor Y by $Y_k \doteqdot g^{ij}X_{ijk}$. The contracted second Bianchi identity shows that

$$|Y|^2 = \left(\nabla^i R_{ik} - \frac{1}{3}\nabla_k R\right)\left(\nabla^j R_j^k - \frac{1}{3}\nabla^k R\right) = \frac{1}{36}|\nabla R|^2 .$$

But the standard estimate

$$|Z|^2 \geq \frac{1}{n}\left(\operatorname{tr}_g Z\right)^2$$

for any $(2,0)$-tensor Z (not necessarily symmetric) implies in dimension $n = 3$ that

$$\frac{1}{3}|Y|^2 \leq |X|^2 = |\nabla \mathrm{Rc}|^2 - \frac{1}{3}|\nabla R|^2 .$$

Hence we get

$$|\nabla \mathrm{Rc}|^2 \geq \frac{1}{3}\left(1 + \frac{1}{36}\right)|\nabla R|^2 = \frac{37}{108}|\nabla R|^2, \tag{6.46}$$

from which the result follows easily. □

REMARK 6.41. By decomposing the $(3,0)$-tensor $\nabla \mathrm{Rc}$ into irreducible components, Hamilton proved a stronger inequality

$$|\nabla \mathrm{Rc}|^2 - \frac{1}{3}|\nabla R|^2 \geq \frac{1}{21}|\nabla \mathrm{Rc}|^2$$

in Lemma 11.6 of [**58**]; but the estimate above suffices for the proof of Theorem 6.35. Hamilton observes that $\nabla \mathrm{Rc}$ may be written as the sum of two irreducible components

$$\nabla_i R_{jk} = A_{ijk} + B_{ijk},$$

where

$$A_{ijk} = \frac{1}{20}\left[(\nabla_k R)\, g_{ij} + (\nabla_j R)\, g_{ik}\right] + \frac{3}{10}(\nabla_i R)\, g_{jk}.$$

The constants $1/20$ and $3/10$ are determined by the condition that all traces of B_{ijk} are zero. The contracted second Bianchi identity implies that $g^{jk}A_{ijk} = \nabla_i R$ and $g^{ij}A_{ijk} = g^{ij}A_{ikj} = \frac{1}{2}\nabla_k R$. It follows that $\langle A, B\rangle = 0$. Calculating that

$$|A|^2 = \left(\frac{8}{400} + \frac{27}{100} + \frac{12}{200}\right)|\nabla R|^2 = \frac{7}{20}|\nabla R|^2,$$

one concludes that

$$|\nabla \operatorname{Rc}|^2 = |A|^2 + |B|^2 \geq \frac{7}{20} |\nabla R|^2 ,$$

whence Hamilton's estimate follows.

Combining Corollary 6.39 with Lemma 6.40 gives the following differential inequality.

COROLLARY 6.42. *If* $(\mathcal{M}^3, g(t))$ *is a solution of the Ricci flow on a* 3*-manifold, then*

$$\frac{\partial}{\partial t}\left(|\operatorname{Rc}|^2 - \frac{1}{3}R^2\right) \leq \Delta\left(|\operatorname{Rc}|^2 - \frac{1}{3}R^2\right) - \frac{2}{37}|\nabla \operatorname{Rc}|^2$$
$$- 8 \operatorname{tr}_g\left(\operatorname{Rc}^3\right) + \frac{26}{3}R\,|\operatorname{Rc}|^2 - 2R^3.$$

Now we return to equation (6.43) for the evolution of $|\nabla R|^2/R$. On any manifold of positive Ricci curvature, one can estimate $|\operatorname{Rc}| \leq R$. Then estimating

$$\left|\nabla |\operatorname{Rc}|^2\right| \leq 2\,|\nabla \operatorname{Rc}| \cdot |\operatorname{Rc}|$$

and recalling (6.46), we apply the Cauchy–Schwarz inequality to the bad term $\frac{4}{R}\left\langle \nabla R, \nabla |\operatorname{Rc}|^2 \right\rangle$ on the right-hand side of (6.43), obtaining

$$\begin{aligned}\frac{4}{R}\left\langle \nabla R, \nabla |\operatorname{Rc}|^2 \right\rangle &\leq \frac{4}{R}|\nabla R| \cdot \left|\nabla |\operatorname{Rc}|^2\right| \\ &\leq 8\,|\nabla R| \cdot |\nabla \operatorname{Rc}| \frac{|\operatorname{Rc}|}{R} \\ &\leq 8\sqrt{3}\,|\nabla \operatorname{Rc}|^2 . \qquad (6.47)\end{aligned}$$

This motivates us to consider the quantity

$$V \doteqdot \frac{|\nabla R|^2}{R} + \frac{37}{2}\left(8\sqrt{3}+1\right)\left(|\operatorname{Rc}|^2 - \frac{1}{3}R^2\right) \qquad (6.48)$$

which gives an upper bound for $|\nabla R|^2/R$. Combining Lemma 6.42 and Corollary 6.42 with estimate (6.47) shows that V satisfies the following differential inequality.

LEMMA 6.43. *If* $(\mathcal{M}^3, g(t))$ *is a solution of the Ricci flow on a* 3*-manifold whose Ricci curvature is positive initially, then*

$$\frac{\partial}{\partial t}V \leq \Delta V - 2R\left|\nabla\left(\frac{\nabla R}{R}\right)\right|^2 - 2\frac{|\operatorname{Rc}|^2}{R^2}|\nabla R|^2 - |\nabla \operatorname{Rc}|^2$$
$$+ \frac{37}{2}\left(8\sqrt{3}+1\right)\left(\frac{26}{3}R\,|\operatorname{Rc}|^2 - 8\operatorname{tr}_g\left(\operatorname{Rc}^3\right) - 2R^3\right).$$

The only potentially positive term in the evolution equation for V is

$$W \doteqdot \frac{26}{3} R \left|\mathrm{Rc}\right|^2 - 8 \operatorname{tr}_g \left(\mathrm{Rc}^3\right) - 2R^3.$$

One expects this quantity to be small when the metric is close to Einstein, because the formula derived in Corollary 6.39 shows that W vanishes identically on an Einstein 3-manifold. We now make this expectation precise.

LEMMA 6.44. *On a 3-manifold of positive Ricci curvature, one has*

$$W \leq \frac{50}{3} R \left(\left|\mathrm{Rc}\right|^2 - \frac{1}{3} R^2 \right).$$

PROOF. If we define

$$X \doteqdot -8 \left\langle \mathrm{Rc} - \frac{1}{3} Rg, \mathrm{Rc}^2 \right\rangle,$$

then $X = \frac{8}{3} R \left|\mathrm{Rc}\right|^2 - 8 \operatorname{tr}_g \left(\mathrm{Rc}^3\right)$, and we may write W as

$$W = X + 6R \left(\left|\mathrm{Rc}\right|^2 - \frac{1}{3} R^2 \right).$$

Define a $(2,0)$-tensor Y by $Y \doteqdot \mathrm{Rc}^2 - \frac{1}{9} R^2 g$, observing that

$$Y_{ij} \doteqdot R_i^k R_{kj} - \frac{1}{9} R^2 g_{ij} = g^{k\ell} \left(R_{ik} - \frac{1}{3} R g_{ik} \right) \left(R_{j\ell} + \frac{1}{3} R g_{j\ell} \right).$$

Then using Cauchy–Schwarz, we can estimate X in dimension $n = 3$ by

$$\begin{aligned} X &= -8 \left\langle \mathrm{Rc} - \frac{1}{3} Rg, Y \right\rangle \\ &\leq 8 \left| \mathrm{Rc} + \frac{1}{3} Rg \right| \cdot \left| \mathrm{Rc} - \frac{1}{3} Rg \right|^2 \\ &\leq \frac{32}{3} R \cdot \left| \mathrm{Rc} - \frac{1}{3} Rg \right|^2 = \frac{32}{3} R \left(\left|\mathrm{Rc}\right|^2 - \frac{1}{3} R^2 \right). \end{aligned}$$

We used the positivity of Rc to get the last inequality. The lemma follows. □

We are ready to prove the main result of this section.

PROOF OF THEOREM 6.35. If $\left(\mathcal{M}^3, g(t)\right)$ is a solution of the Ricci flow on a closed 3-manifold whose Ricci curvature is positive initially, we may combine the results of Lemmas 6.43 and 6.44 to obtain

$$\frac{\partial}{\partial t} V \leq \Delta V - \left|\nabla \mathrm{Rc}\right|^2 + \frac{7400\sqrt{3} + 925}{3} R \left(\left|\mathrm{Rc}\right|^2 - \frac{1}{3} R^2 \right).$$

Applying the result of Theorem 6.30 in the form (6.41) lets us estimate this further as

$$\frac{\partial}{\partial t} V \leq \Delta V - \left|\nabla \mathrm{Rc}\right|^2 + C R^{3-2\gamma},$$

where C and $\gamma \equiv \bar{\delta}/2$ depend only on g_0. On the other hand, it follows from formula (6.6) that

$$\frac{\partial}{\partial t} R^{2-\gamma} = \Delta \left(R^{2-\gamma} \right) - (2-\gamma)(1-\gamma) R^{-\gamma} |\nabla R|^2 + 2(2-\gamma) R^{1-\gamma} |\mathrm{Rc}|^2 .$$

Choose $\bar{\beta}$ depending only on g_0 such that

$$0 < \bar{\beta} \leq \frac{(R_{\min}(0))^\gamma}{3(2-\gamma)(1-\gamma)},$$

and recall that

$$|\nabla R|^2 \leq 3 |\nabla \mathrm{Rc}|^2 .$$

(This is a consequence of (6.46), but is easy to prove directly.) Then for any $\beta \in [0, \bar{\beta}]$, one has an estimate

$$\begin{aligned} \frac{\partial}{\partial t} \left(V - \beta R^{2-\gamma} \right) &\leq \Delta \left(V - \beta R^{2-\gamma} \right) \\ &\quad + \left[\beta (2-\gamma)(1-\gamma) R^{-\gamma} |\nabla R|^2 - |\nabla \mathrm{Rc}|^2 \right] \\ &\quad + C R^{3-2\gamma} - 2\beta (2-\gamma) R^{1-\gamma} |\mathrm{Rc}|^2 \\ &\leq \Delta \left(V - \beta R^{2-\gamma} \right) + C_1, \end{aligned}$$

where C_1 depends only on β and g_0. By the maximum principle, we conclude that

$$V - \beta R^{2-\gamma} \leq C_1 t.$$

But formula (6.6) implies that

$$\frac{\partial}{\partial t} R \geq \Delta R + \frac{2}{3} R^2;$$

as we shall see in Lemma 6.53, this forces the solution to become singular within a finite time T depending only on g_0. Hence we conclude that

$$\frac{|\nabla R|^2}{R} \leq V \leq \beta R^{2-\gamma} + C_1 T.$$

□

7. Higher derivative estimates and long-time existence

If g_0 is a smooth metric on a compact manifold $\mathcal{M}^n$, Theorem 3.13 implies that a unique solution $g(t)$ of the Ricci flow satisfying $g(0) = g_0$ exists for a short time. It follows that there is a maximal time interval $0 \leq t < T \leq \infty$ on which the solution exists. If $T < \infty$, it will be important for us to understand what goes wrong. The goal of this section is to prove the following result.

THEOREM 6.45. *If g_0 is a smooth metric on a compact manifold $\mathcal{M}^n$, the unnormalized Ricci flow with $g(0) = g_0$ has a unique solution $g(t)$ on a maximal time interval $0 \leq t < T \leq \infty$. Moreover, if $T < \infty$, then*

$$\lim_{t \nearrow T} \left(\sup_{x \in \mathcal{M}^n} |\mathrm{Rm}(x,t)| \right) = \infty.$$

We shall in fact prove the contrapositive of the theorem. Namely, we will show that if the maximum curvature were to remain bounded along a sequence of times approaching T, the solution could be extended past T. In order to make this argument rigorous, we must first discuss *a priori* estimates for any solution whose curvatures remain suitably bounded.

7.1. Higher derivative estimates. In Section 5 of Chapter 3, we saw how the Ricci flow may be regarded heuristically as a nonlinear heat equation for a Riemannian metric. This viewpoint admits a rigorous interpretation in the sense that the curvatures of an evolving metric all satisfy parabolic equations. In particular, bounds on the curvature of an initial metric g_0 automatically induce *a priori* bounds on all derivatives of the curvature for a short time. We call these **Bernstein–Bando–Shi (BBS) derivative estimates**, because they follow the strategy introduced by Bernstein [**16, 17, 18**] for proving gradient bounds via the maximum principle, and were derived for the Ricci flow in [**9**] and [**117, 118**]. (See also [**11**] and Section 7 of [**63**].)

Because the BBS estimates for the Ricci flow are not surprising in that they follow the natural parabolic scaling in which time scales like distance squared, we shall postpone our technical discussion of them until Chapter 7. (Certain local analogues of the BBS estimates will also be discussed in the planned successor to this volume.) It will be enough for our present needs to assume the following result from Chapter 7.

THEOREM 6.46 (Theorem 7.1). *Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow for which the maximum principle holds. (This is true for instance if $\mathcal{M}^n$ is compact.) Then for each $\alpha > 0$ and every $m \in \mathbb{N}$, there exists a constant C_m depending only on m, and n, and* $\max\{\alpha, 1\}$*such that if*

$$|\mathrm{Rm}(x,t)|_g \leq K \quad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in [0, \frac{\alpha}{K}],$$

then

$$|\nabla^m \mathrm{Rm}(x,t)|_g \leq \frac{C_m K}{t^{m/2}} \quad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in (0, \frac{\alpha}{K}].$$

As stated above, the BBS derivative estimates deteriorate as $t \searrow 0$. This is due to the fact that bounds on the Riemann tensor of an initial metric imply nothing about its derivatives. Furthermore, the BBS estimates above allow the factors C_j to grow as the time interval over which we estimate becomes larger. Because our ultimate goal is to establish long-time existence of the flow, we want a practical criterion that will let us know when a solution which exists up to an arbitrarily large time T can be extended past T. We can accomplish this by the following strategy: when we have an upper bound on $|\mathrm{Rm}|$ which is uniform in time, we can apply Theorem 6.46 at large times by taking a fixed small step backwards in time. This technique lets us prove the following consequence of the theorem.

COROLLARY 6.47. *Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow for which the weak maximum principle holds. If there are $\beta > 0$ and $K > 0$ such that*

$$|\mathrm{Rm}(x,t)|_g \leq K \quad \textit{for all } x \in \mathcal{M} \textit{ and } t \in [0,T],$$

where $T > \beta/K$, then there exists for each $m \in \mathbb{N}$ a constant C_m depending only on m, n, and $\min\{\beta, 1\}$ such that

$$|\nabla^m \mathrm{Rm}(x,t)|_g \leq C_m K^{1+\frac{m}{2}} \quad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in \left[\frac{\min\{\beta,1\}}{K}, T\right].$$

PROOF. Let $\beta_1 \doteqdot \min\{\beta, 1\}$. Fix $t_0 \in [\beta_1/K, T]$ and set $T_0 \doteqdot t_0 - \beta_1/K$. Let $\bar{t} \doteqdot t - T_0$, and let $\bar{g}(\bar{t})$ solve the Cauchy problem

$$\frac{\partial}{\partial \bar{t}}\bar{g} = -2\overline{\mathrm{Rc}}$$
$$\bar{g}(0) = g(T_0).$$

Then by uniqueness of solutions to the Ricci flow, $\bar{g}(\bar{t}) = g(\bar{t} + T_0) = g(t)$ for $\bar{t} \in [0, \beta_1/K]$. Thus by hypothesis on the solution $g(t)$, we have

$$\left|\overline{\mathrm{Rm}}(x,\bar{t})\right|_{\bar{g}} \leq K$$

for all $x \in \mathcal{M}^n$ and $\bar{t} \in [0, \beta_1/K]$. Applying Theorem 6.46 with $\alpha = \beta_1$, we get a family of constants $\bar{C}_m$ depending only on m and n such that

$$\left|\bar{\nabla}^m \overline{\mathrm{Rm}}(x,\bar{t})\right|_{\bar{g}} \leq \frac{\bar{C}_m K}{\bar{t}^{m/2}}$$

for all $x \in \mathcal{M}^n$ and $\bar{t} \in (0, \beta_1/K]$. Now when $\bar{t} \in \left[\frac{\beta_1}{2K}, \frac{\beta_1}{K}\right]$, we have

$$\bar{t}^{m/2} \geq \beta_1^{m/2} 2^{-m/2} K^{-m/2}.$$

Taking $\bar{t} = \beta_1/K$, we find in particular that

$$|\nabla^m \mathrm{Rm}(x,t_0)|_g \leq \left(\frac{2^{m/2}\bar{C}_m}{\beta_1^{m/2}}\right) K^{1+m/2} \qquad \text{for all } x \in \mathcal{M}^n.$$

Since $t_0 \in [\beta_1/K, T]$ was arbitrary, the result follows. □

The main objective of this subsection is to obtain the following gradient estimates. They are stated with respect to a half-open time interval $[0,T)$ because one of their applications (moreover, the one of current interest) is to help us study obstacles to long-time existence of the flow.

PROPOSITION 6.48. *Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow on a compact manifold with a fixed background metric $\bar{g}$ and connection $\bar{\nabla}$. If there exists $K > 0$ such that*

$$|\mathrm{Rm}(x,t)|_g \leq K \qquad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in [0,T),$$

then there exists for every $m \in \mathbb{N}$ a constant C_m depending on m, n, K, T, g_0, and the pair $(\bar{g}, \bar{\nabla})$ such that

$$\left|\bar{\nabla}^m g(x,t)\right|_{\bar{g}} \leq C_m \qquad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in [0,T).$$

The first step in proving the proposition (and ultimately in establishing long-time existence of the normalized flow) is to obtain a sufficient condition for the metrics composing a smooth one-parameter family to be uniformly equivalent. Recall that one writes $A \leq B$ for symmetric 2-tensors A and B if $B - A$ is a nonnegative definite quadratic form, that is if $(B-A)(V,V) \geq 0$ for all vectors V.

LEMMA 6.49. *Let $\mathcal{M}^n$ be a closed manifold. For $0 \leq t < T \leq \infty$, let $g(t)$ be a one-parameter family of metrics on $\mathcal{M}^n$ depending smoothly on both space and time. If there exists a constant $C < \infty$ such that*

$$\int_0^T \left|\frac{\partial}{\partial t} g(x,t)\right|_{g(t)} dt \leq C$$

for all $x \in \mathcal{M}^n$, then

$$e^{-C} g(x,0) \leq g(x,t) \leq e^C g(x,0)$$

for all $x \in \mathcal{M}^n$ and $t \in [0,T)$. Furthermore, as $t \nearrow T$, the metrics $g(t)$ converge uniformly to a continuous metric $g(T)$ such that for all $x \in \mathcal{M}^n$,

$$e^{-C} g(x,0) \leq g(x,T) \leq e^C g(x,0).$$

PROOF. Let $x \in \mathcal{M}^n$, $t_0 \in [0,T)$, and $V \in T_x\mathcal{M}^n$ be arbitrary. Then using the fact that $|A(U,U)| \leq |A|_g$ for any 2-tensor A and unit vector U, we obtain

$$\begin{aligned}
\left|\log\left(\frac{g_{(x,t_0)}(V,V)}{g_{(x,0)}(V,V)}\right)\right| &= \left|\int_0^{t_0} \frac{\partial}{\partial t}\left[\log g_{(x,t)}(V,V)\right] dt\right| \\
&= \left|\int_0^{t_0} \frac{\frac{\partial}{\partial t} g_{(x,t)}(V,V)}{g_{(x,t)}(V,V)} dt\right| \\
&\leq \int_0^{t_0} \left|\frac{\partial}{\partial t} g_{(x,t)}\left(\frac{V}{|V|}, \frac{V}{|V|}\right)\right| dt \\
&\leq \int_0^{t_0} \left|\frac{\partial}{\partial t} g_{(x,t)}\right|_{g(t)} dt \leq C.
\end{aligned}$$

The uniform bounds on $g(t)$ follow from exponentiation.

Since $\mathcal{M}^n$ is compact, $\int_t^T \left| \frac{\partial g_{ij}}{\partial t}(x,\tau) \right|_{g(t)} d\tau \to 0$ as $t \nearrow T$ uniformly in $x \in \mathcal{M}^n$. Hence the function $f : T\mathcal{M}^n \to \mathbb{R}$ defined by

$$f(x,V) \doteqdot \lim_{t\to T} g_{(x,t)}(V,V)$$

exists and is continuous in $x \in \mathcal{M}^n$ and $V \in T_x\mathcal{M}^n$. By polarization, we can define a $(2,0)$-tensor $g_{(x,T)}$ on $\mathcal{M}^n$ by

$$g_{(x,T)}(V,W) = \frac{1}{4}\left[f(x,V+W) - f(x,V-W)\right],$$

noting that $g_{(x,T)}(V,W) = \lim_{t\to T} g_{(x,t)}(V,W)$. The uniform bounds on $g(t)$ ensure that

$$e^{-C} g_{(x,0)}(V,V) \le f(x,V) \le e^{C} g_{(x,0)}(V,V),$$

hence that $g_{(x,T)}$ is a Riemannian metric. □

In particular, the lemma implies that if $(\mathcal{M}^n, g(t))$ is a solution of the Ricci flow satisfying a uniform curvature bound on a finite time interval $[0,T)$, then all metrics in the family $\{g(t) : 0 \le t < T\}$ are uniformly equivalent.

COROLLARY 6.50. *Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow. If there exists a constant K such that $|\mathrm{Rc}| \le K$ on the time interval $[0,T]$, then*

$$e^{-2KT} g(x,0) \le g(x,t) \le e^{2KT} g(x,0)$$

for all $x \in \mathcal{M}^n$ and $t \in [0,T]$.

We are now ready to prove the proposition.

PROOF OF PROPOSITION 6.48. Since $\mathcal{M}^n$ is compact, it is covered by a finite atlas in which we have uniform estimates on the derivatives of the local charts. Henceforth, we fix such a chart $\varphi : \mathcal{U} \to \mathbb{R}^n$. Since $\bar{g}$ and the connection $\bar{\nabla}$ are fixed, it will suffice to prove that for each $m \in \mathbb{N}$, the ordinary derivatives of g of order m satisfy an estimate

$$|\partial^m g(x,t)| \le C_m \qquad \text{for all } x \in \mathcal{U} \text{ and } t \in t \in [0,T),$$

where the norm $|\cdot| \equiv |\cdot|_\delta$ is taken with respect to the Euclidean metric δ in $\mathcal{U}$, and C_m depends on only on m, n, K, T, and g_0. Adopting this point of view, we shall in particular regard Γ as a tensor in $\mathcal{U}$, namely as the difference of the Levi-Civita connection of g and the background (flat) connection in $\mathcal{U}$. The proof will be by complete induction on $m \in \mathbb{N}$.

In the estimates which follow, C will denote a generic constant that may change from one inequality to the next, but which depends only on m, n, K, T, and g_0. Choose $\beta = \beta(K,T)$ so that $0 < \beta < \min\{KT, 1\}$. By Corollary 6.50, we have uniform pointwise bounds from above and below for the metrics $g(t)$ on the time interval $[0,T)$. To estimate the first derivatives of the metric, we begin by writing

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{\partial}{\partial x^i}g_{jk}\right) &= \frac{\partial}{\partial x^i}\left(\frac{\partial}{\partial t}g_{jk}\right) = -2\frac{\partial}{\partial x^i}R_{jk} \\ &= -2\left(\nabla_i R_{jk} + \Gamma_{ij}^{\ell}R_{\ell k} + \Gamma_{ik}^{\ell}R_{j\ell}\right).\end{aligned}$$

Because $|\mathrm{Rm}(x,t)|_g \leq K$ by hypothesis, this implies that

$$\left|\frac{\partial}{\partial t}\partial g\right| = 2\left|\partial\,\mathrm{Rc}\right| \leq 2\left|\nabla\,\mathrm{Rc}\right| + CK\left|\Gamma\right|. \tag{6.49}$$

By equation (6.1), the tensor $\frac{\partial}{\partial t}\Gamma$ is given by

$$\frac{\partial}{\partial t}\Gamma_{ij}^{k} = -g^{k\ell}\left(\nabla_i R_{j\ell} + \nabla_j R_{i\ell} - \nabla_\ell R_{ij}\right),$$

whence we get the estimate

$$\left|\frac{\partial}{\partial t}\Gamma\right| \leq C\left|\nabla\,\mathrm{Rc}\right|.$$

Now by Corollary 6.47, there exists $B = B(m,n,K,\beta)$ such that the bound $|\nabla\,\mathrm{Rc}| \leq B$ holds on the time interval $(\beta/K, T)$. Since $|\Gamma|$ is bounded on $[0,\beta/K]$ by some $A = A(K,\beta,g_0)$, we see that

$$|\Gamma(x,t)| \leq A + BC(T-\beta/K) \leq C \tag{6.50}$$

for all $x \in \mathcal{M}^n$ and $t \in [0,T)$. Since $|\nabla\,\mathrm{Rc}|$ is bounded on $[0,\beta/K]$ by some $D = D(K,\beta,g_0)$, we conclude by (6.49) that

$$|\partial g| \leq |\partial g_0| + CD\frac{\beta}{K} + (2B+C)(T-\beta/K) \leq C.$$

To do the general (inductive) step, let $\alpha = (a_1,\ldots,a_r)$ be any multi-index with $|\alpha| = m$. Then since

$$\frac{\partial}{\partial t}\left(\frac{\partial^{|\alpha|}}{\partial x^\alpha}g_{ij}\right) = -2\left(\frac{\partial^{|\alpha|}}{\partial x^\alpha}R_{ij}\right),$$

it will suffice to bound $\left|\partial^{|\alpha|}\,\mathrm{Rc}\right|$. A moment's reflection reveals that we can estimate

$$\left|\partial^m\,\mathrm{Rc}\right| \leq \sum_{i=0}^{m} c_i\left|\Gamma^i\right|\left|\nabla^{m-i}\,\mathrm{Rc}\right| + \sum_{i=1}^{m-1} c_i'\left|\partial^i\Gamma\right|\left|\partial^{m-1-i}\,\mathrm{Rc}\right|, \tag{6.51}$$

where the constants c_i, c_i' depend only on m and n. For instance, the case $m = 2$ corresponds to the explicit formula

$$\begin{aligned}\partial_i\partial_j R_{k\ell} &= \nabla_i\nabla_j R_{k\ell} \\ &\quad + \left(\Gamma_{ij}^p \nabla_p R_{kl} + \Gamma_{ik}^p \nabla_j R_{p\ell} + \Gamma_{i\ell}^p \nabla_j R_{kp} + \Gamma_{jk}^p \nabla_i R_{p\ell} + \Gamma_{j\ell}^p \nabla_i R_{kp}\right) \\ &\quad + \left(\Gamma_{ip}^q \Gamma_{jl}^p R_{kq} + \Gamma_{ip}^q \Gamma_{jk}^p R_{q\ell} + \Gamma_{i\ell}^p \Gamma_{jk}^q R_{qp} + \Gamma_{ik}^p \Gamma_{j\ell}^q R_{pq}\right) \\ &\quad + \left(\partial_i \Gamma_{jk}^p R_{p\ell} + \partial_i \Gamma_{j\ell}^p R_{kp}\right).\end{aligned}$$

Applying Corollary 6.47 and estimate (6.50), we see that

$$\sum_{i=0}^{m} \left|\Gamma^i\right| \left|\nabla^{m-i}\operatorname{Rc}\right| \leq C \sup_{0\leq t\leq \beta/K}\left[\sum_{i=0}^{m}\left|\nabla^{m-i}\operatorname{Rc}\right|\right] + C\sum_{i=0}^{m} C_{m-i}K^{1+\frac{m-i}{2}} \leq C.$$

By the inductive hypothesis, we may assume that $|\partial_t \partial^p g|$, hence $|\partial^p \operatorname{Rc}|$, has been estimated for all $0 \leq p < m-1$. By (6.51), this implies in particular that $\left|\partial^i\Gamma\right|$ has been bounded for all $1 \leq i \leq m-2$. So to finish the proof, it will suffice to estimate $\left|\partial^{m-1}\Gamma\right|$. We shall accomplish this by taking a time derivative and integrating. By equation (6.1), we have

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\frac{\partial^{m-1}}{\partial x^{p_1}\cdots\partial x^{p_{m-1}}}\Gamma_{ij}^k\right) \\ &= \frac{\partial^{m-1}}{\partial x^{p_1}\cdots\partial x^{p_{m-1}}}\left(\frac{\partial}{\partial t}\Gamma_{ij}^k\right) \\ &= \frac{\partial^{m-1}}{\partial x^{p_1}\cdots\partial x^{p_{m-1}}}\left[g^{k\ell}\left(-\nabla_i R_{j\ell} - \nabla_j R_{i\ell} + \nabla_\ell R_{ij}\right)\right].\end{aligned}$$

So by the inductive hypothesis and Corollary 6.50,

$$\begin{aligned}\left|\frac{\partial}{\partial t}\partial^{m-1}\Gamma\right| &\leq C\sum_{i=0}^{m-1}\left|\partial^{m-1-i}\left(g^{-1}\right)\right|\left|\partial^i\nabla\operatorname{Rc}\right| \\ &\leq C\sum_{i=0}^{m-1}\left|\partial^{m-1-i}g\right|\left|\partial^i\nabla\operatorname{Rc}\right| \\ &\leq C\sum_{i=0}^{m-1}\left|\partial^i\nabla\operatorname{Rc}\right|.\end{aligned} \tag{6.52}$$

Here we used the fact that bounds on g and its derivatives induce bounds on the derivatives of g^{-1} via the identity

$$\frac{\partial}{\partial x^i}g^{jk} = -g^{jp}g^{kq}\frac{\partial}{\partial x^i}g_{pq}.$$

Reasoning as in the derivation of (6.51) with the tensor Rc replaced by $\nabla \operatorname{Rc}$, we observe that

$$\left|\partial^i \nabla \operatorname{Rc}\right| \leq \sum_{j=0}^{i} \bar{c}_j \left|\Gamma^j\right| \left|\nabla^{i+1-j} \operatorname{Rc}\right| + \sum_{j=1}^{i-1} \bar{c}'_j \left|\partial^j \Gamma\right| \left|\partial^{i-j} \operatorname{Rc}\right|,$$

where the constants $\bar{c}_j, \bar{c}''_j$ depend only on p and n. Applying this estimate to (6.52) then shows that

$$\left|\frac{\partial}{\partial t} \partial^{m-1} \Gamma\right| \leq C \sum_{i=0}^{m-1} \left(\sum_{j=0}^{i} \left|\Gamma^j\right| \left|\nabla^{i+1-j} \operatorname{Rc}\right| + \sum_{j=1}^{i-1} \left|\partial^j \Gamma\right| \left|\partial^{i-j} \operatorname{Rc}\right| \right),$$

where all of the terms on the right-hand side have already been bounded. Thus we have $\left|\frac{\partial}{\partial t} \partial^{m-1} \Gamma\right| \leq C$ and hence $\left|\partial^{m-1} \Gamma\right| \leq C + CT$. This completes the proof. □

An examination of the proof makes it evident that we have also established the following result.

COROLLARY 6.51. *Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow on a compact manifold with a fixed background metric $\bar{g}$ and connection $\bar{\nabla}$. If there exists $K > 0$ such that*

$$|\operatorname{Rm}(x,t)|_g \leq K \quad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in [0,T),$$

then there exists for every $m \in \mathbb{N}$ a constant C'_m depending on m, n, K, T, g_0, and the pair $(\bar{g}, \bar{\nabla})$ such that

$$\left|\bar{\nabla}^m \operatorname{Rc}(x,t)\right|_{\bar{g}} \leq C'_m \quad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in [0,T).$$

REMARK 6.52. As an alternative to the argument used to prove Proposition 6.48, one may begin by writing the identity

$$0 = \nabla_i g_{jk} = \frac{\partial}{\partial x^i} g_{jk} - \Gamma_{ij}^{\ell} g_{\ell k} - \Gamma_{ik}^{\ell} g_{j\ell}$$

in the form

$$\frac{\partial}{\partial x^i} g_{jk} = \Gamma_{ij}^{\ell} g_{\ell k} + \Gamma_{ik}^{\ell} g_{j\ell}.$$

Then using the uniform bounds on $g(t)$ given by Corollary 6.50, one gets the estimate

$$\left|\frac{\partial}{\partial x^i} g_{jk}\right| \leq C |\Gamma|$$

and proceeds as before.

7.2. Long-time existence. Now we are ready to show that the curvature becoming unbounded is the only obstacle to long-time existence of the flow.

PROOF OF THEOREM 6.45. Define $M(t) \doteqdot \sup_{x \in \mathcal{M}^n} |\mathrm{Rm}(x,t)|$. We shall first prove the claim that

$$\limsup_{t \nearrow T} (M(t)) = \infty.$$

By Theorem 3.13, a unique solution $g(t)$ of the Ricci flow satisfying $g(0) = g_0$ exists for a short time. Suppose that the solution exists on a maximal finite time interval $[0,T)$. To obtain a contradiction, suppose further that there is a positive constant K such that

$$\sup_{0 \le t < T} M(t) \le K.$$

Fix a local coordinate patch $\mathcal{U}$ around an arbitrary point $x \in \mathcal{M}^n$, and let $\tau \in (0,T)$ be arbitrary as well. Then by Lemma 6.49, a continuous limit metric $g(T)$ exists and is given in component form by the integral formula

$$g_{ij}(x,T) = g_{ij}(x,\tau) - 2\int_\tau^T R_{ij}(x,t)\, dt.$$

Let $\alpha = (a_1, \dots, a_r)$ be any multi-index with $|\alpha| = m \in \mathbb{N}$. By Proposition 6.48 and Corollary 6.51, both $\frac{\partial^m}{\partial x^\alpha} g_{ij}$ and $\frac{\partial^m}{\partial x^\alpha} R_{ij}$ are uniformly bounded on $\mathcal{U} \times [0,T)$. Thus we can write

$$\left(\frac{\partial^m}{\partial x^\alpha} g_{ij}\right)(x,T) = \left(\frac{\partial^m}{\partial x^\alpha} g_{ij}\right)(x,\tau) - 2\int_\tau^T \left(\frac{\partial^m}{\partial x^\alpha} R_{ij}\right)(x,t)\, dt,$$

which shows that $|\partial^\alpha g| \le C$ for some positive constant C, hence that $g(T)$ is a C^∞ metric, and also

$$\left|\left(\frac{\partial^m}{\partial x^\alpha} g_{ij}\right)(x,T) - \left(\frac{\partial^m}{\partial x^\alpha} g_{ij}\right)(x,\tau)\right| \le C(T-\tau),$$

which shows that $g(\tau) \to g(T)$ uniformly in any C^m norm as $t \nearrow T$.

Now since $g(T)$ is smooth, Theorem 3.13 implies there is a solution of the Ricci flow $\bar{g}(t)$ that satisfies $\bar{g}(0) = g(T)$ and exists for a short time $0 \le t < \varepsilon$. Since $g(\tau) \to g(T)$ smoothly, if follows that

$$\tilde{g}(t) \doteqdot \begin{cases} g(t) & 0 \le t < T \\ \bar{g}(t-T) & T \le t < T+\varepsilon \end{cases}$$

is a solution of the Ricci flow that is smooth in space and time (in particular near T) and satisfies $g(0) = g_0$. This contradicts the assumption that T is maximal, and proves the claim.

Finally, we prove the theorem by replacing the lim sup of the claim with a *bona fide* limit. Suppose the theorem is false. Then there exists $K_0 < \infty$

and a sequence of times $t_i \nearrow T$ such that $M(t_i) \leq K_0$. A consequence of the differential inequality

$$\frac{\partial}{\partial t} |\mathrm{Rm}|^2 \leq \Delta |\mathrm{Rm}|^2 - 2 |\nabla \mathrm{Rm}|^2 + C_n |\mathrm{Rm}|^3$$

we shall derive in Lemma 7.4 is the doubling-time estimate of Corollary 7.5, which implies that there exists $c > 0$ depending only on the dimension n such that

$$M(t) \leq 2M(t_i) \leq 2K_0$$

for all times t satisfying

$$t_i \leq t < \min\left\{T, t_i + \frac{c}{K_0}\right\}.$$

Since $t_i \nearrow T$ as $i \to \infty$, there exists i_0 large enough that $t_{i_0} + c/K_0 \geq T$. But then this implies that

$$\sup_{t_{i_0} \leq t < T} M(t) \leq 2K_0,$$

which contradicts the claim established above and hence proves the theorem. □

8. Finite-time blowup

In Section 9 below, we shall shift our attention to the normalized Ricci flow in order to complete our proof of Theorem 6.3. Along the way, we will need to prove that the unique solution of the normalized Ricci flow starting on a closed 3-manifold $(\mathcal{M}^3, g_0)$ of positive Ricci curvature exists for all positive time. Paradoxically, the key to proving this fact is obtaining a better understanding of exactly how the corresponding unnormalized solution becomes singular. We begin with the observation that a finite-time singularity is inevitable.

LEMMA 6.53. *Let $(\mathcal{M}^n, g(t) : 0 \leq t < T)$ be a solution of the unnormalized Ricci flow for which the weak maximum principle holds. If there are $t_0 \geq 0$ and $\rho > 0$ such that*

$$\inf_{x \in \mathcal{M}^n} R(x, t_0) = \rho,$$

then $g(t)$ becomes singular in finite time.

PROOF. By equation (6.6), we have

$$\frac{\partial}{\partial t} R = \Delta R + 2 |\mathrm{Rc}|^2 \geq \Delta R + \frac{2}{n} R^2.$$

Consider the solution

$$r(t) = \frac{1}{\frac{1}{\rho} - \frac{2}{n}(t - t_0)}$$

to the initial value problem

$$\begin{cases} dr/dt &= 2r^2/n \\ r(t_0) &= \rho \end{cases}.$$

By the maximum principle, $R_{\inf}(t) \doteqdot \inf_{x\in\mathcal{M}^n} R(x,t)$ satisfies the inequality $R_{\inf}(t) \geq r(t)$ for as long as both solutions exist. □

Once we know a solution becomes singular in finite time, Theorem 6.45 tells us that its Riemannian curvatures must blow up as we approach the singularity time.

COROLLARY 6.54. *Any solution $(\mathcal{M}^3, g(t))$ of the unnormalized Ricci flow on a compact manifold whose Ricci curvature is positive initially exists on a maximal time interval $0 \leq t < T < \infty$ and has the property that*

$$\lim_{t\nearrow T}\left(\sup_{x\in\mathcal{M}^3}|\mathrm{Rm}(x,t)|\right) = \infty.$$

By combining this fact with results of Sections 5 and 6, we can obtain global pinching estimates for the curvature.

LEMMA 6.55. *Let $(\mathcal{M}^3, g(t))$ be a solution of the unnormalized Ricci flow on a compact manifold whose Ricci curvature is positive initially. Then the solution becomes singular at some time $T < \infty$. Moreover, it obeys the following* a priori *estimates.*

(1) *There exist positive constants C and γ depending only on the initial data such that*

$$\frac{R_{\min}}{R_{\max}} \geq 1 - CR_{\max}^{-\gamma}$$

for all times $0 \leq t < T$. In particular, $R_{\min}/R_{\max} \to 1$ as $t \nearrow T$.

(2) *For $x \in \mathcal{M}^3$ and $t \in [0,T)$, let $\lambda(x,t) \geq \mu(x,t) \geq \nu(x,t)$ denote the eigenvalues of the curvature operator at (x,t). Then for any $\varepsilon \in (0,1)$, there exists $T_\varepsilon \in [0,T)$ such that for all times $T_\varepsilon \leq t < T$, one has*

$$\min_{x\in\mathcal{M}^3}\nu(x,t) \geq (1-\varepsilon)\max_{y\in\mathcal{M}^3}\lambda(y,t) > 0.$$

In particular, the solution eventually attains positive sectional curvature everywhere.

PROOF. By Lemma 6.53, we know the solution becomes singular at some time $T < \infty$. Since $c|\mathrm{Rc}| \leq |\mathrm{Rm}| \leq C|\mathrm{Rc}|$ in dimension $n = 3$, it follows from Theorem 6.45 that

$$\lim_{t\nearrow T}\left(\sup_{x\in\mathcal{M}^3}|\mathrm{Rc}(x,t)|\right) = \infty. \tag{6.53}$$

To prove claim (1), define

$$R_{\min}(t) \doteqdot \inf_{x\in\mathcal{M}^3} R(x,t) \qquad \text{and} \qquad R_{\max}(t) \doteqdot \sup_{x\in\mathcal{M}^3} R(x,t).$$

By Theorem 6.35, there are positive constants A, B, and α such that

$$|\nabla R|^2 \leq \frac{1}{2}A^2 R_{\max}^{3-2\alpha} + B^2.$$

Since $|\mathrm{Rc}|^2 \leq R^2$ and R is positive, it follows from (6.53) that there exists $\tau \in [0, T)$ such that

$$|\nabla R| \leq A R_{\max}^{3/2-\alpha}$$

at all times $t \in (\tau, T)$. Consider any time $t \in (\tau, T)$. Since $\mathcal{M}^3$ is compact, there exist $\bar{x}(t) \in \mathcal{M}^3$ such that $R_{\max}(t) = R(\bar{x}, t)$. Given $\varepsilon > 0$, consider the geodesic ball $B(\bar{x}, L)$, where

$$L(t) \doteqdot \frac{1}{\varepsilon \sqrt{R_{\max}(t)}}.$$

If γ is any minimizing geodesic from $\bar{x}$ to $\underline{x} \in B(\bar{x}, L)$, we can estimate that

$$R_{\max} - R(\underline{x}) \leq \int_\gamma |\nabla R| \, ds \leq ALR_{\max}^{3/2-\alpha} \leq \frac{A}{\varepsilon} R_{\max}^{1-\alpha}.$$

Hence on $B(\bar{x}, L)$, we have the lower bound

$$R \geq R_{\max}\left(1 - \frac{A}{\varepsilon} R_{\max}^{-\alpha}\right). \tag{6.54}$$

It follows that there exists $\bar{t} \in (\tau, T)$ depending only on A, α, and ε such that

$$R \geq (1 - \varepsilon) R_{\max} \tag{6.55}$$

on $B(\bar{x}, L)$ for all $t \in [\bar{t}, T)$. Now by Corollary 6.29, there exists $\beta > 0$ depending only on g_0 such that the inequality

$$\mathrm{Rc} \geq 2\beta^2 R g$$

holds at all points of $\mathcal{M}^3$. By Myers' Theorem, this implies that the minimizing geodesic γ emanating from $\bar{x}$ must encounter a conjugate point within the distance

$$\frac{\pi}{\beta\sqrt{\inf_{B(\bar{x},L)} R}}.$$

Now when $\varepsilon > 0$ is sufficiently small, estimate (6.55) implies that

$$\frac{\pi}{\beta\sqrt{\inf_{B(\bar{x},L)} R}} \leq \frac{\pi}{\beta\sqrt{(1-\varepsilon) R_{\max}}} \leq \frac{1}{\varepsilon\sqrt{R_{\max}}} = L,$$

hence that $B(\bar{x}, L)$ is all of $\mathcal{M}^3$.

To prove claim (2), recall that Theorem 6.30 implies there exist positive constants C and $\delta < 1$ depending only on g_0 such that

$$\nu \geq \lambda - C(\mu + \nu)^{1-\delta} \geq \lambda - C(\lambda + \mu + \nu)^{1-\delta}$$

at all points on the manifold. Thus we have the pointwise inequality

$$\frac{\nu}{\lambda} \geq 1 - 3CR^{-\delta} \geq 1 - 3CR_{\min}^{-\delta}. \tag{6.56}$$

Let $x, y \in \mathcal{M}^3$ and $\eta > 0$ be given. Then by (6.53), (6.55), and (6.56), there exists $T_\eta \in [0, T)$ such that for all times $T_\eta \leq t < T$ one has

$$\begin{aligned}
\nu(x,t) &\geq (1-\eta)\,\lambda(x,t) \\
&\geq \frac{1-\eta}{3} R(x,t) \\
&\geq \frac{(1-\eta)^2}{3} R(y,t) \\
&\geq \frac{(1-\eta)^2}{3} \left[\lambda(y,t) + 2(1-\eta)\,\lambda(y,t)\right] \\
&\geq (1-\eta)^3\,\lambda(y,t).
\end{aligned}$$

The claim follows by taking the infimum over all $x \in \mathcal{M}^3$ and the supremum over all $y \in \mathcal{M}^3$. □

This result lets us conclude in particular that $g(t)$ is approaching an Einstein metric uniformly as $t \nearrow T$.

Corollary 6.56. *If $\left(\mathcal{M}^3, g(t)\right)$ is a solution of the unnormalized Ricci flow on a compact manifold whose Ricci curvature is positive initially, then*

$$\lim_{t \nearrow T} \left(\sup_{x \in \mathcal{M}^3} \frac{|\overset{\circ}{\mathrm{Rc}}|^2}{R^2} \right) = 0.$$

Proof. Applying the local pinching estimate of Theorem 6.30 in the form (6.37) shows that there are positive constants C and $\bar{\delta}$ such that

$$\frac{|\overset{\circ}{\mathrm{Rc}}|^2}{R^2} \leq C R^{-\bar{\delta}} \leq R_{\min}^{-\bar{\delta}}.$$

Since $\lim_{t \nearrow T} R_{\max}(t)$ by Theorem 6.54 and $\lim_{t \nearrow T} (R_{\min}/R_{\max}) = 1$ by Lemma 6.55, the result follows. □

9. Properties of the normalized Ricci flow

We now have at our disposal all the facts about the unnormalized Ricci flow that we will need to prove Theorem 6.3. But since the theorem is stated for the normalized flow, we must convert to that flow in order to complete its proof. In this section, we shall study how this is done and will then show that the normalized flow exists for all time and asymptotically approaches an Einstein metric. In the final part of this chapter (Section 10, below) we will prove that the convergence is in fact exponential.

9.1. Equivalence of the flows under rescaling. We begin by showing that the unnormalized and normalized Ricci flows differ only by a rescaling of space and time. Specifically, suppose that $(\mathcal{M}^n, g(t))$ is a solution of the unnormalized Ricci flow

$$\frac{\partial}{\partial t} g = -2\,\mathrm{Rc}$$

on a manifold of finite volume. To convert to the normalized flow, define dilating factors $\psi(t) > 0$ so that the metrics $\bar{g}(t) = \psi(t) \cdot g(t)$ have constant volume

$$\int_{\mathcal{M}^n} d\bar{\mu} \equiv 1,$$

and put $\bar{t} = \int_0^t \psi(\tau)\, d\tau$. Then $d\bar{t}/dt = \psi(t)$, while the geometries of g and $\bar{g}$ are related by the following lemma.

LEMMA 6.57. *Let $(\mathcal{M}^n, g)$ be a Riemannian manifold. If $\bar{g} = \psi g$ for some $\psi > 0$, then the following relations result.*

1. *The Levi-Civita connections of $\bar{g}$ and g are related by $\bar{\Gamma}_{ij}^k = \Gamma_{ij}^k$.*
2. *The $(3,1)$-Riemann curvature tensors of $\bar{g}$ and g are related by $\bar{R}_{ijk}^{\ell} = R_{ijk}^{\ell}$.*
3. *The $(4,0)$-Riemann curvature tensors of $\bar{g}$ and g are related by $\bar{R}_{ijk\ell} = \psi R_{ijk\ell}$.*
4. *The Ricci curvature tensors of $\bar{g}$ and g are related by $\bar{R}_{ij} = R_{ij}$.*
5. *The scalar curvatures of $\bar{g}$ and g are related by $\bar{R} = \psi^{-1} R$.*
6. *The volume elements of $\bar{g}$ and g are related by $d\bar{\mu} = \psi^{n/2}\, d\mu$.*

Now writing equation 6.5 in the form

$$\frac{\partial}{\partial t} \log \det g = g^{ij} \frac{\partial}{\partial t}(g_{ij}) = -2R,$$

we see that

$$\frac{d}{dt} \int_{\mathcal{M}^n} d\mu = -\int_{\mathcal{M}^n} R\, d\mu,$$

hence that ψ is a smooth function of time. Thus we get the evolution equation

$$\frac{\partial}{\partial \bar{t}} \bar{g} = \frac{dt}{d\bar{t}} \frac{\partial}{\partial t} (\psi g) = -2\overline{\mathrm{Rc}} + \left(\frac{1}{\psi^2} \frac{d\psi}{dt} \right) \bar{g}. \tag{6.57}$$

Denote the average scalar curvature of $\bar{g}(t)$ by

$$\bar{\rho}(t) \doteqdot \frac{\int_{\mathcal{M}^n} \bar{R}\, d\bar{\mu}}{\int_{\mathcal{M}^n} d\bar{\mu}} = \int_{\mathcal{M}^n} \bar{R}\, d\bar{\mu}.$$

Then observing as above that

$$\frac{\partial}{\partial t} \log \det (\psi g) = \frac{1}{\psi} g^{ij} \frac{\partial}{\partial t} (\psi g_{ij}) = -2R + \frac{n}{\psi} \frac{d\psi}{dt},$$

we obtain

$$\begin{aligned} 0 = \frac{d}{dt} \int d\bar{\mu} &= \int \frac{\partial}{\partial t} \log \sqrt{\det (\psi g)}\, d\bar{\mu} \\ &= \int \left(-\psi \bar{R} + \frac{n}{2\psi} \frac{\partial \psi}{\partial t} \right) d\bar{\mu} \\ &= -\psi \bar{\rho} + \frac{n}{2\psi} \frac{d\psi}{dt}. \end{aligned}$$

This shows that $\frac{2\bar{\rho}}{n} = \frac{1}{\psi^2}\frac{d\psi}{dt}$, hence by (6.57) that $\bar{g}$ evolves by the normalized Ricci flow

$$\frac{\partial}{\partial \bar{t}}\bar{g} = -2\overline{\mathrm{Rc}} + \frac{2\bar{\rho}}{n}\bar{g}.$$

9.2. Long-time existence of the normalized flow. Let $(\mathcal{M}^3, g(t))$ be the unique solution of the unnormalized Ricci flow starting on a 3-manifold of positive Ricci curvature. By Theorem 6.54, $g(t)$ becomes singular in the sense that it exists on a maximal time interval $0 \leq t < T < \infty$ and satisfies $\lim_{t \nearrow T} R_{\max}(t) = \infty$, where $R_{\max}(t) \doteqdot \sup_{x \in \mathcal{M}^3} R(x,t)$. Let $(\mathcal{M}^3, \bar{g}(\bar{t}))$ be the corresponding normalized solution constructed in Subsection 9.1. Then $\bar{g}(\bar{t})$ exists on a maximal time interval $0 \leq \bar{t} < \bar{T} \leq \infty$. We shall now prove that $\bar{g}$ enjoys long-time existence, namely that $\bar{T} = \infty$.

THEOREM 6.58. *If $(\mathcal{M}^3, g(t))$ is a solution of the unnormalized Ricci flow starting on a compact manifold of positive Ricci curvature, then the corresponding normalized solution $(\mathcal{M}^3, \bar{g}(\bar{t}))$ exists for all positive time.*

We begin with some preliminary results. Define

$$\bar{R}_{\min}(\bar{t}) \doteqdot \inf_{x \in \mathcal{M}^3} \bar{R}(x, \bar{t})$$

and

$$\bar{R}_{\max}(\bar{t}) \doteqdot \sup_{x \in \mathcal{M}^3} \bar{R}(x, \bar{t}).$$

LEMMA 6.59. *If $(\mathcal{M}^3, g(t))$ is a solution of the Ricci flow with initially positive Ricci curvature, then*

$$\int_0^T R_{\max}(t)\, dt = \infty.$$

PROOF. Since $R_{\max}$ is a continuous function of $t \in [0, T)$, there is a unique solution ρ to the initial value problem

$$\begin{cases} d\rho/dt &= 2R_{\max} \cdot \rho \\ \rho(0) &= R_{\max}(0) \end{cases}.$$

Since $g(t)$ has positive Ricci curvature for as long as it exists, we have $|\mathrm{Rc}|^2 \leq R^2 \leq R \cdot R_{\max}$, which implies that

$$\begin{aligned}\frac{\partial}{\partial t}(R - \rho) &= \Delta (R - \rho) + 2\left(|\mathrm{Rc}|^2 - R_{\max} \cdot \rho\right) \\ &\leq \Delta (R - \rho) + 2R_{\max}(R - \rho).\end{aligned}$$

So by the maximum principle, we have $R \leq \rho$ for all points $x \in \mathcal{M}^3$ and times $0 \leq t < T$. This implies in particular that $\rho \geq R_{\max}$, hence by Theorem 6.54 that $\lim_{\tau \nearrow T} \rho(\tau) = \infty$. Since for $0 \leq \tau < T$, we have

$$\log \frac{\rho(\tau)}{\rho(0)} = \int_0^\tau \frac{d}{dt} \log \rho \, dt = 2\int_0^\tau R_{\max}(t)\, dt,$$

the integral on the right-hand side must diverge. $\square$

LEMMA 6.60. *There exists a positive constant $\bar{C}$ such that*

$$\bar{R}_{\max} \le \bar{C}.$$

PROOF. Let $\bar{L}(\bar{t})$ and $\bar{V}(\bar{t})$ denote the diameter and volume of $\bar{g}(\bar{t})$, respectively. Since $\overline{\mathrm{Rc}} = \mathrm{Rc} > 0$, the Bishop–Guenther volume comparison theorem implies that

$$1 \equiv \bar{V} \le \frac{4}{3}\pi \bar{L}^3. \tag{6.58}$$

On the other hand, Corollary 6.29 shows there is a positive constant β depending only on g_0 such that

$$\overline{\mathrm{Rc}} = \mathrm{Rc} \ge 2\beta^2 R_{\min} g = 2\beta^2 \bar{R}_{\min} \bar{g}.$$

So by Myers' Theorem,

$$\bar{L} \le \frac{\pi}{\beta\sqrt{\bar{R}_{\min}}}. \tag{6.59}$$

Since

$$\lim_{\bar{t} \nearrow \bar{T}} \frac{\bar{R}_{\min}(\bar{t})}{\bar{R}_{\max}(\bar{t})} = \lim_{t \nearrow T} \frac{R_{\min}(t)}{R_{\max}(t)} = 1$$

by Lemma 6.55, there exists a positive constant C such that

$$\frac{\bar{R}_{\min}}{\bar{R}_{\max}} \ge \frac{1}{C}. \tag{6.60}$$

Combining estimates (6.58), (6.59), and (6.60), we get the desired result:

$$\bar{R}_{\max} \le C\bar{R}_{\min} \le C\left(\frac{\pi}{\beta\bar{L}}\right)^2 \le C\left(\frac{\pi}{\beta}\right)^2\left(\frac{4\pi}{3}\right)^{2/3}.$$

□

We are now ready to prove the main result of this section.

PROOF OF THEOREM 6.58. Let

$$\rho(t) \doteqdot \frac{\int_{\mathcal{M}^3} R\, d\mu}{\int_{\mathcal{M}^3} d\mu}$$

and

$$\bar{\rho}(\bar{t}) \doteqdot \frac{\int_{\mathcal{M}^3} \bar{R}\, d\bar{\mu}}{\int_{\mathcal{M}^3} d\bar{\mu}} = \int_{\mathcal{M}^3} \bar{R}\, d\bar{\mu}$$

denote the average scalar curvatures of the unnormalized and normalized solutions, respectively. Recall that $d\bar{t} = \psi dt$. By Lemma 6.57, we have

$\bar{R} = R/\psi$, and $d\bar{\mu} = \psi^{n/2} d\mu = \psi^{3/2} d\mu$. Thus for any $\tau \in [0, T)$, we have

$$\begin{aligned}\int_0^\tau \rho(t)\, dt &= \int_0^\tau \frac{\int_{\mathcal{M}^3} R\, d\mu}{\int_{\mathcal{M}^3} d\mu}\, dt \\ &= \int_0^{\bar{\tau}} \frac{\int_{\mathcal{M}^3} \psi \bar{R} \psi^{-3/2}\, d\bar{\mu}}{\int_{\mathcal{M}^3} \psi^{-3/2}\, d\bar{\mu}} \psi^{-1}\, d\bar{t} \\ &= \int_0^{\bar{\tau}} \bar{\rho}(\bar{t})\, d\bar{t},\end{aligned}$$

where $\bar{\tau} = \int_0^\tau \psi(t)\, dt$. Now since

$$R_{\min}(t) \leq \rho(t) \leq R_{\max}(t),$$

we apply Lemmas 6.55 and 6.59 to conclude that the left-hand side diverges as $\tau \nearrow T$. Hence the right-hand side must also diverge as $\bar{\tau} \nearrow \bar{T}$. But by Lemma 6.60, there exists $\bar{C} > 0$ such that

$$\int_0^{\bar{\tau}} \bar{\rho}(\bar{t})\, d\bar{t} \leq \int_0^{\bar{\tau}} \bar{R}_{\max}(\bar{t})\, d\bar{t} \leq \bar{C}\bar{\tau}.$$

Hence $\bar{T} = \infty$. □

Noting that the expression in Corollary 6.56 is invariant under dilation of the metric, we obtain the following convergence result for free.

COROLLARY 6.61. *If $(\mathcal{M}^3, g(t))$ is a solution of the unnormalized Ricci flow on a compact manifold whose Ricci curvature is positive initially, then the corresponding normalized solution $\bar{g}(\bar{t})$ exists for all positive time and asymptotically approaches an Einstein metric uniformly in the sense that*

$$\lim_{\bar{t}\to\infty} \left(\sup_{x \in \mathcal{M}^3} \frac{|\overset{\circ}{\overline{\mathrm{Rc}}}|^2}{\bar{R}^2} \right) = 0.$$

In Section 10, we shall show that the convergence is exponential.

9.3. Evolution of curvature under the normalized flow. To augment the scaling arguments employed above, it will be useful in the sequel to observe directly how various geometric quantities evolve under the normalized Ricci flow. To wit, suppose that we are given a solution of the normalized Ricci flow

$$\frac{\partial}{\partial t} g = -2\,\mathrm{Rc} + rg$$

on a manifold $\mathcal{M}^n$, where r is the function of time alone defined by

$$r(t) \doteqdot \frac{2}{n}\rho(t) \equiv \frac{2}{n} \frac{\int_{\mathcal{M}^n} R\, d\mu}{\int_{\mathcal{M}^n} d\mu}. \tag{6.61}$$

(Note that we are now dropping the $\bar{g}(\bar{t})$ notation used above.) The easiest way to obtain the evolution equations for quantities depending on g is by first making a more general observation.

LEMMA 6.62. *Suppose that $g(t)$ is a smooth one-parameter family of metrics on a manifold $\mathcal{M}^n$ such that*

$$\frac{\partial}{\partial t} g = \varphi g,$$

where φ is a function of time alone.

(1) *The Levi-Civita connection Γ of g is independent of time.*
(2) *The $(3,1)$-Riemann curvature tensor of g is independent of time.*
(3) *The $(4,0)$-Riemann curvature tensor* Rm *of g evolves by*

$$\frac{\partial}{\partial t} R_{ijk\ell} = \varphi R_{ijk\ell}.$$

(4) *The Ricci tensor* Rc *of g is independent of time.*
(5) *The scalar curvature R of g evolves by*

$$\frac{\partial}{\partial t} R = -\varphi R.$$

PROOF. All the formulas are immediate consequences of Lemma 6.5 except for (3), which follows from (2) when we write

$$\frac{\partial}{\partial t}\left(R_{ijk\ell}\right) = \frac{\partial}{\partial t}\left(g_{\ell m} R_{ijk}^{m}\right) = \varphi R_{ijk\ell}.$$

□

REMARK 6.63. The formulas above may also be deduced from Lemma 6.57 by observing how the various geometric quantities under consideration respond to scaling the metric. (Compare with Lemma 17.1 of [**58**].)

COROLLARY 6.64. *Let $(\mathcal{M}^n, g(t))$ be a solution of the normalized Ricci flow, and let $r(t)$ be defined by (6.61).*

(1) *The Levi-Civita connection Γ of g evolves by*

$$\frac{\partial}{\partial t}\Gamma_{ij}^{k} = -\nabla_i R_j^k - \nabla_j R_i^k + \nabla^k R_{ij}.$$

(2) *The $(3,1)$-Riemann curvature tensor evolves by*

$$\begin{aligned}\frac{\partial}{\partial t} R_{ijk}^{\ell} = \Delta R_{ijk}^{\ell} &+ g^{pq}\left(R_{ijp}^{r} R_{rqk}^{\ell} - 2R_{pik}^{r} R_{jqr}^{\ell} + 2R_{pir}^{\ell} R_{jqk}^{r}\right) \\ &- R_i^p R_{pjk}^{\ell} - R_j^p R_{ipk}^{\ell} - R_k^p R_{ijp}^{\ell} + R_p^{\ell} R_{ijk}^{p}.\end{aligned}$$

(3) *The $(4,0)$-Riemann curvature tensor* Rm *of g evolves by*

$$\begin{aligned}\frac{\partial}{\partial t} R_{ijk\ell} = \Delta R_{ijk\ell} &+ 2\left(B_{ijk\ell} - B_{ij\ell k} - B_{i\ell jk} + B_{ikj\ell}\right) \\ &+ R_{k\ell j}^{m} R_{mi} - R_{k\ell i}^{m} R_{mj} + R_{ij\ell}^{m} R_{mk} - R_{ijk}^{m} R_{m\ell} + rR_{ijk\ell},\end{aligned}$$

where the tensor B is defined in (6.15).

(4) *The Ricci tensor of g evolves by*

$$\frac{\partial}{\partial t} R_{jk} = \Delta_L R_{jk} = \Delta R_{jk} + 2 R_{pjkq} R^{pq} - 2 R_j^\ell R_{\ell k}.$$

(5) *The scalar curvature R of g evolves by*

$$\frac{\partial}{\partial t} R = \Delta R + 2 |\mathrm{Rc}|^2 - rR.$$

10. Exponential convergence

We may now assume we are given a solution $(\mathcal{M}^3, g(t))$ of the normalized Ricci flow

$$\frac{\partial}{\partial t} g = -2\,\mathrm{Rc} + rg$$

starting on a closed 3-manifold of positive Ricci curvature, where $r(t)$ is given by formula (6.61). By Theorem 6.58, we know that $g(t)$ exists for all positive time. By Lemma 6.8, Corollary 6.11, and Lemma 6.57, we know that the Ricci and scalar curvatures of $g(t)$ remain strictly positive. By Lemma 6.60, we know that there exists a positive constant A such that

$$0 < R \leq A, \tag{6.62}$$

hence such that

$$0 < r \leq AV, \tag{6.63}$$

where V is the constant volume of $(\mathcal{M}^3, g(t))$. And by Corollary 6.58, we know $g(t)$ becomes asymptotically Einstein in the sense that

$$\lim_{t \to \infty} \left(\sup_{x \in \mathcal{M}^3} \frac{|\overset{\circ}{\mathrm{Rc}}|^2}{R^2} \right) = 0.$$

In order to show that this convergence is exponential, we shall follow the method of Section 4 (which uses techniques from the later paper [**59**] instead of the method originally employed in [**58**]). As in Section 2, let V be a vector bundle over $\mathcal{M}^n$ isomorphic to $T\mathcal{M}^n$. Corresponding to the evolution of $g(t)$ by the normalized Ricci flow, we construct a one-parameter family of bundle isometries $\iota(t) : V \to T\mathcal{M}^n$ evolving by

$$\frac{\partial}{\partial t} \iota = \mathrm{Rc} \circ \iota - \frac{r}{2} \iota.$$

Then recalling Corollary 6.64, we compute that the analogue of formula (6.21) for the evolution of the components (R_{abcd}) of $\iota^* \mathrm{Rm}$ is

$$\frac{\partial}{\partial t} R_{abcd} = \Delta_D R_{abcd} + 2 (B_{abcd} - B_{abdc} + B_{acbd} - B_{adbc}) - r R_{abcd}.$$

It follows that the quadratic form $\mathbb{M}$ defined by the curvature operator on each fiber of the bundle $\wedge^2 T\mathcal{M}^3$ evolves by the ODE

$$\frac{d}{dt} \mathbb{M} = \mathbb{M}^2 + \mathbb{M}^\# - rM.$$

In particular, the eigenvalues $\lambda \geq \mu \geq \nu$ of $\mathbb{M}$ (which are twice the sectional curvatures) evolve by

$$\frac{d}{dt}\lambda = \lambda^2 + \mu\nu - r\lambda$$
$$\frac{d}{dt}\mu = \mu^2 + \lambda\nu - r\mu$$
$$\frac{d}{dt}\nu = \nu^2 + \lambda\mu - r\nu.$$

It is easy to check that

$$\begin{aligned} [\lambda(\mu+\nu)] \frac{d}{dt} \log \frac{\lambda}{\mu+\nu} &= (\mu+\nu)\left(\lambda^2 + \mu\nu\right) - r\lambda(\mu+\nu) \\ &\quad - \lambda\left[\mu^2 + \nu^2 + \lambda(\mu+\nu)\right] + r\lambda(\mu+\nu), \end{aligned}$$

hence that the proof of Lemma 6.28 goes through for the normalized flow exactly as written. Thus we can immediately state the following result.

LEMMA 6.65. *Let $(\mathcal{M}^3, g(t))$ be a solution of the normalized Ricci flow on a closed 3-manifold of initially positive Ricci curvature. Then there exists a constant B such that*

$$\lambda \leq B(\mu+\nu)$$

for all positive time.

REMARK 6.66. One can also conclude directly that the estimate of Lemma 6.28 applies to the normalized Ricci flow by writing it in the scale-invariant form

$$\frac{\lambda}{\mu+\nu} \leq B.$$

We shall need one more simple observation.

LEMMA 6.67. *Let $(\mathcal{M}^3, g(t))$ be a solution of the normalized Ricci flow on a closed 3-manifold of initially positive Ricci curvature. Then there exists $\varepsilon > 0$ depending only on g_0 such that $R \geq \varepsilon$ for all positive time.*

PROOF. By (6.6), R satisfies the differential inequality $\partial R/\partial t \geq \Delta R$. Hence $R_{\min}(t) \geq R_{\min}(0) > 0$. □

Now we are ready for the key estimate of this section.

PROPOSITION 6.68. *Let $(\mathcal{M}^3, g(t))$ be a solution of the normalized Ricci flow on a closed 3-manifold of initially positive Ricci curvature. Then there exist constants $\alpha \in (0,1)$, β, and C depending only on g_0 such that*

$$\lambda - \nu \leq C(\mu+\nu)^{1-\alpha} e^{-\beta t}$$

for all positive time.

PROOF. Following the proof of Theorem 6.30, we compute that

$$\frac{d}{dt}\log(\lambda-\nu)=\lambda-\mu+\nu-r$$

and

$$\frac{d}{dt}\log(\mu+\nu)=\lambda+\frac{\mu^2+\nu^2}{\mu+\nu}-r.$$

The calculation

$$\frac{d}{dt}(\mu-\nu)=(\mu+\nu-\lambda-r)(\mu-\nu)$$

shows that $\mu-\nu$ stays positive. Thus we can compute that

$$\frac{d}{dt}\log\left[e^{\beta t}\frac{\lambda-\nu}{(\mu+\nu)^{1-\alpha}}\right]=\alpha(\lambda-r)+\beta-(\mu-\nu)-(1-\alpha)\frac{\mu^2+\nu^2}{\mu+\nu}.$$

Recalling estimates (6.62) and (6.63), and applying Lemma 6.65, we estimate this as

$$\frac{d}{dt}\log\left[e^{\beta t}\frac{\lambda-\nu}{(\mu+\nu)^{1-\alpha}}\right]\le\alpha B(\mu+\nu)+\beta-\frac{1-\alpha}{2}(\mu+\nu).$$

If $\alpha>0$ is chosen smaller than $1/[2(1+2B)]$, then

$$\frac{d}{dt}\log\left[e^{\beta t}\frac{\lambda-\nu}{(\mu+\nu)^{1-\alpha}}\right]\le\beta-\frac{1-\alpha(1+2B)}{2}(\mu+\nu)\le\beta-\frac{\mu+\nu}{4}.$$

By Lemmas 6.65 and (6.67), there exists $\varepsilon>0$ such that

$$4\varepsilon\le\lambda+\mu+\nu\le(1+B)(\mu+\nu).$$

Hence

$$\frac{d}{dt}\log\left[e^{\beta t}\frac{\lambda-\nu}{(\mu+\nu)^{1-\alpha}}\right]\le\beta-\frac{\varepsilon}{1+B}.$$

For sufficiently small $\beta>0$ depending only on B and ε, the right-hand side is negative. So there exists a constant $C=C(g_0)$ large enough that $\mathbb{M}(t)$ starts inside and hence remains in the convex set

$$\mathcal{K}=\left\{\mathbb{P}:e^{\beta t}[\lambda(\mathbb{P})-\nu(\mathbb{P})]-C[\mu(\mathbb{P})+\nu(\mathbb{P})]^{1-\alpha}\le 0\right\}.$$

□

By Lemma 6.65 and estimate (6.62), the proposition implies that

$$\lambda-\nu\le CR^{1-\alpha}e^{-\beta t}\le ACe^{-\beta t},$$

whence formula (6.34) and Lemma 6.67 imply that $g(t)$ is exponentially becoming a metric of constant positive sectional curvature.

COROLLARY 6.69. *If $(\mathcal{M}^3,g(t))$ is a solution of the normalized Ricci flow on a closed 3-manifold of initially positive Ricci curvature, then there exist positive constants β and B such that*

$$\left|\mathrm{Rc}-\frac{1}{3}R\cdot g\right|=|\overset{\circ}{\mathrm{Rc}}|\le Be^{-\beta t}$$

for all positive time.

From here it is not hard to show that all derivatives of the curvature decay exponentially. This implies that $g(t)$ converges exponentially fast in every C^m norm to a smooth Einstein metric g_∞; for details, the reader is referred to Section 17 of [**58**]. (Also see Section 5 of Chapter 5.)

Notes and commentary

The pinching estimate in Lemma 6.28 has an analogue for the mean curvature flow [**74**]. This is also true for Theorem 6.30; however, the proof of the analogous estimate for the mean curvature flow requires integral estimates. In dimension $n = 4$, pinching estimates for the Ricci flow we obtained by Hamilton under the assumption of positive curvature operator [**59**]. Higher dimensional estimates under stronger pinching assumptions were obtained independently by Huisken [**75**], Margerin [**95, 96**], and Nishikawa [**102, 103**].

CHAPTER 7

Derivative estimates

We saw in Chapter 3 that the Ricci flow is equivalent via DeTurck's trick to a quasilinear parabolic PDE, and may in fact be regarded heuristically as a nonlinear heat equation for a Riemannian metric. Furthermore, we saw in Chapter 6 that the intrinsically defined curvatures of a Riemannian metric evolving by the Ricci flow all obey parabolic equations with quadratic nonlinearities. Knowing this, anyone familiar with the smoothing properties of parabolic equations would expect that appropriate bounds on the geometry of a given Riemannian manifold $(\mathcal{M}^n, g_0)$ would induce *a priori* bounds on the geometry of the unique solution $g(t)$ of the Ricci flow such that $g(0) = g_0$. Moreover, one would even expect the geometry to improve, at least for a short time.

In this chapter, we verify those expectations by proving global short-time derivative estimates for all derivatives of the curvature of a solution to the Ricci flow. As we remarked in Chapter 6, we call these **Bernstein–Bando–Shi estimates** (briefly, BBS estimates) because they follow Bernstein's technique of using the maximum principle to establish gradient estimates, and were applied to the Ricci flow in papers authored independently by Bando and Shi.

1. Global estimates and their consequences

Local derivative estimates are important for performing dimension reduction and in other cases where one wants to take a 'semi-global limit' of a sequence of geodesic balls whose radii go to infinity, or a 'local limit' of a sequence of geodesic balls of uniform size. Local derivative estimates will be discussed in the successor to this volume. In the present chapter, we will prove those global derivative estimates that are important for establishing long-time existence of the flow, as was discussed in Section 7 of Chapter 6. Our main goal for this chapter is to prove the following result.

THEOREM 7.1. *Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow for which the maximum principle holds. (This is true in particular if $\mathcal{M}^n$ is compact.) Then for each $\alpha > 0$ and every $m \in \mathbb{N}$, there exists a constant C_m depending only on m, and n, and* $\max\{\alpha, 1\}$ *such that if*

$$|\mathrm{Rm}(x,t)|_{g(x,t)} \leq K \quad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in [0, \frac{\alpha}{K}],$$

then

$$|\nabla^m \operatorname{Rm}(x,t)|_{g(x,t)} \leq \frac{C_m K}{t^{m/2}} \quad \textit{for all } x \in \mathcal{M}^n \textit{ and } t \in (0, \frac{\alpha}{K}].$$

Note that the estimates in Theorem 7.1 follow the natural parabolic scaling in which time behaves like distance squared. Note too that we have stated the estimates in a form that deteriorates as $t \searrow 0$. This is the best one can do without making further assumptions on the initial metric. Indeed, it is easy to construct examples of rotationally symmetric metrics on $\mathcal{S}^n$ with $|R| \leq 1$ and $|\nabla R|$ arbitrarily large by writing the metric as a warped product. (Recall that we used warped product metrics on spheres in Section 5 of Chapter 2.)

Theorem 7.1 has several important consequences, of which the following two are particularly useful.

COROLLARY 7.2 (Long-time existence). *If g_0 is a smooth metric on a compact manifold $\mathcal{M}^n$, the unique solution $g(t)$ of the Ricci flow such that $g(0) = g_0$ exists on a maximal time interval $0 \leq t < T \leq \infty$. Moreover, $T < \infty$ only if*

$$\lim_{t \nearrow T} \left(\sup_{x \in \mathcal{M}^n} |\operatorname{Rm}(x,t)| \right) = \infty.$$

This result equivalent to Theorem 6.45.

COROLLARY 7.3 (Uniform bounds for sequences). *Let $\{(\mathcal{M}_i^n, g_i^0) : i \in \mathbb{N}\}$ be a sequence of compact Riemannian manifolds. If their curvatures are uniformly bounded in the sense that*

$$\left|\operatorname{Rm}\left[g_i^0\right]\right|_{g_i^0} \leq K$$

for some constant $K < \infty$ independent of $i \in \mathbb{N}$, then there exists a constant $c > 0$ depending only on n such that for each $i \in \mathbb{N}$, the unique solution $g_i(t)$ to the Ricci flow such that $g_i(0) = g_i^0$ exists on $\mathcal{M}_i^n$ for the uniform time interval $t \in [0, c/K]$. Moreover, there is for each $m \in \mathbb{N}$ a constant $C_m < \infty$ depending only on m and n (that is, independent of $i \in \mathbb{N}$) such that

$$\sup_{x \in \mathcal{M}_i^n} |\nabla^m \operatorname{Rm}[g_i(t)]|_{g_i(t)} \leq \frac{C_m K}{\left(\sqrt{t}\right)^m}.$$

In particular, for any time $t_0 \in (0, c/K)$, there is for each $m \in \mathbb{N}$ some $C_m' = C_m'(m, n, K, t_0)$ such that

$$\sup_{\mathcal{M}^n \times [t_0, c/K]} |\nabla^m \operatorname{Rm}[g_i]|_{g_i} \leq C_m'.$$

This fact is an immediate consequence of Theorem 7.1 and the minimal existence time result (Corollary 7.7) we shall prove below. An important application of Corollary 7.3 is to obtain convergence in each C^m norm when one proves the compactness theorem for the Ricci flow, which we state in Section 3 of this chapter.

Before proving Theorem 7.1, we will calculate the evolution of the square of the norm of the curvature tensor. Besides being a prototype for the sort of equations we will encounter in the proof of the theorem, this result has useful applications of its own.

LEMMA 7.4. *If $(\mathcal{M}^n, g(t))$ is a solution of the Ricci flow, then the square of the norm of its curvature tensor evolves by*

$$\frac{\partial}{\partial t}|\mathrm{Rm}|^2 = \Delta|\mathrm{Rm}|^2 - 2|\nabla \mathrm{Rm}|^2 + 4g^{ri}g^{sj}g^{pk}g^{q\ell}R_{rspq}\left(B_{ijk\ell} - B_{ij\ell k} + B_{ikj\ell} - B_{i\ell jk}\right),$$

where

$$B_{ijk\ell} = -R_{pij}^{q}R_{q\ell k}^{p}.$$

In particular, one has

$$\frac{\partial}{\partial t}|\mathrm{Rm}|^2 \leq \Delta|\mathrm{Rm}|^2 - 2|\nabla \mathrm{Rm}|^2 + C|\mathrm{Rm}|^3,$$

where C is a constant depending only on the dimension n.

PROOF. By formula (6.17), the $(4,0)$-Riemann curvature tensor evolves by

$$\frac{\partial}{\partial t}R_{ijk\ell} = \Delta R_{ijk\ell} + 2\left(B_{ijk\ell} - B_{ij\ell k} + B_{ikj\ell} - B_{i\ell jk}\right) \tag{7.1a}$$

$$- \left(R_i^pR_{pjk\ell} + R_j^pR_{ipk\ell} + R_k^pR_{ijp\ell} + R_\ell^pR_{ijkp}\right). \tag{7.1b}$$

It is easy to check that

$$\begin{aligned}\frac{\partial}{\partial t}|\mathrm{Rm}|^2 &= \frac{\partial}{\partial t}\left(g^{ri}g^{sj}g^{pk}g^{q\ell}R_{rspq}R_{ijk\ell}\right)\\ &= 2g^{ri}g^{sj}g^{pk}g^{q\ell}R_{rspq}\left[\Delta R_{ijk\ell} + 2\left(B_{ijk\ell} - B_{ij\ell k} + B_{ikj\ell} - B_{i\ell jk}\right)\right];\end{aligned}$$

indeed, the terms that arise when we differentiate g^{-1} exactly cancel the terms that appear in line (7.1b). Since

$$\Delta|\mathrm{Rm}|^2 = 2g^{ri}g^{sj}g^{pk}g^{q\ell}R_{rspq}\Delta R_{ijk\ell} + 2|\nabla \mathrm{Rm}|^2,$$

the result follows. □

The following important consequence of Lemma 7.4 explains why the assumption in Theorem 7.1 that $|\mathrm{Rm}|$ is bounded for a short time is a reasonable one.

COROLLARY 7.5 (Doubling-time estimate). *There exists $c > 0$ depending only on the dimension n such that if $(\mathcal{M}^n, g(t) : 0 \leq t \leq \tau)$ is a solution of the Ricci flow on a compact manifold and*

$$M(t) \doteqdot \sup_{x\in\mathcal{M}^n}|\mathrm{Rm}(x,t)|_{g(x,t)},$$

then

$$M(t) \leq 2M(0) \quad \textit{for all times} \quad 0 \leq t \leq \min\left\{\tau, \frac{\mathrm{c}}{M(0)}\right\}.$$

PROOF. By Lemma 7.4, $M(t)$ is a Lipschitz function of time which satisfies

$$\frac{dM}{dt} \leq \frac{CM^3}{2M} = \frac{C}{2}M^2$$

in the sense of the lim sup of forward difference quotients, where C depends only on n. This implies that

$$M(t) \leq \frac{1}{\frac{1}{M(0)} - \frac{C}{2}t}$$

as long as $t \in [0, \tau]$ satisfies $t < 2/(CM(0))$. Let $\mathrm{c} = 1/C$. Then one has $M(t) \leq 2M(0)$ for all times t satisfying $0 \leq t \leq \min\{\tau, c/M(0)\}$. □

REMARK 7.6. We encountered a special case of this result in Lemma 5.45.

Combining Corollaries 7.2 and 7.5 yields the following useful observation.

COROLLARY 7.7 (Minimal existence-time). *If $(\mathcal{M}^n, g_0)$ is a Riemannian manifold such that $|\mathrm{Rm}[g_0]|_{g_0} \leq K$, then the unique solution $g(t)$ of the Ricci flow with $g(0) = g_0$ exists at least for $t \in [0, \mathrm{c}/K]$, where $\mathrm{c} > 0$ is a constant depending only on n.*

2. Proving the global estimates

In proving Theorem 7.1, we shall have to estimate the time derivative of the quantity $\left|\nabla^k \mathrm{Rm}\right|^2$ on a solution $(\mathcal{M}^n, g(t))$ of the Ricci flow. This is the generalization to n dimensions of the computations in Chapter 5 where we considered the evolution of $\left|\nabla^k R\right|^2$ on a surface. To prepare for the calculations at hand, let us consider a simpler but representative problem. If $Q(t)$ is a 1-parameter family of $(1,0)$-tensor fields on a solution $(\mathcal{M}^n, g(t))$ of the Ricci flow, then formula (6.1) implies that

$$\begin{aligned}\frac{\partial}{\partial t}\nabla_i Q_j &= \frac{\partial}{\partial t}\left(\frac{\partial}{\partial x^i}Q_j - \Gamma_{ij}^k Q_k\right)\\ &= \nabla_i\left(\frac{\partial}{\partial t}Q\right)_j + \left(\nabla_i R_j^k + \nabla_j R_i^k - \nabla^k R_{ij}\right)Q_k.\end{aligned}$$

Hence by Lemma 3.1, we have

$$\begin{aligned}\frac{\partial}{\partial t}|\nabla Q|^2 &= \frac{\partial}{\partial t}\left(g^{ik}g^{j\ell}\nabla_i Q_j \nabla_k Q_\ell\right)\\ &= 2\nabla^i Q^j \nabla_i\left(\frac{\partial}{\partial t}Q\right)_j + 2\nabla^i Q^j\left(\nabla_i R_j^k + \nabla_j R_i^k - \nabla^k R_{ij}\right)Q_k\\ &\quad + 2R^{ik}\nabla_i Q^j \nabla_k Q_j + 2R^{j\ell}\nabla^i Q_j \nabla_i Q_\ell.\end{aligned}$$

Simply put, in taking the time derivative of a quantity such as $|\nabla Q|^2$, one must take into account the evolution of the metric and its Levi-Civita connection as well as the evolution of the tensor itself.

To avoid a notational quagmire, we adopt the following convention in the proof. If A and B are two tensors on a Riemannian manifold, we denote by $A * B$ any quantity obtained from $A \otimes B$ by one or more of these operations: (1) summation over pairs of matching upper and lower indices, (2) contraction on upper indices with respect to the metric, (3) contraction on lower indices with respect to the metric inverse, and (4) multiplication by constants depending only on n and the ranks of A and B. We also denote by A^{*k} any k-fold product $A * \cdots * A$. For example, this convention lets us write the conclusion of Lemma 7.4 in the form

$$\frac{\partial}{\partial t}|\mathrm{Rm}|^2 = \Delta|\mathrm{Rm}|^2 - 2|\nabla \mathrm{Rm}|^2 + (\mathrm{Rm})^{*3}.$$

Proof of Theorem 7.1. The proof is by complete induction on m. We first consider the case $m = 1$. The evolution equation for $|\nabla \mathrm{Rm}|^2$ is of the form

$$\frac{\partial}{\partial t}|\nabla \mathrm{Rm}|^2 = 2\left\langle \nabla\left(\frac{\partial}{\partial t}\mathrm{Rm}\right), \nabla \mathrm{Rm}\right\rangle + \nabla \mathrm{Rc} * \mathrm{Rm} * \nabla \mathrm{Rm} + \mathrm{Rc} * (\nabla \mathrm{Rm})^{*2}.$$

The standard technique of commuting derivatives and using the second Bianchi identity shows that for any tensor A, the commutator $[\nabla, \Delta] A$ is given by

$$[\nabla, \Delta] A = \nabla\Delta A - \Delta\nabla A = \mathrm{Rm} * \nabla A + \nabla \mathrm{Rc} * A.$$

By applying formula (6.17) and replacing instances of Rc with Rm, we get

$$\nabla\left(\frac{\partial}{\partial t}\mathrm{Rm}\right) = \nabla\left(\Delta \mathrm{Rm} + (\mathrm{Rm})^{*2}\right) = \Delta\nabla \mathrm{Rm} + \mathrm{Rm} * \nabla \mathrm{Rm}.$$

Because

$$\Delta|A|^2 = \Delta\langle A, A\rangle = 2\langle \Delta A, A\rangle + 2\langle \nabla A, \nabla A\rangle$$

for any tensor A, we conclude that

$$\frac{\partial}{\partial t}|\nabla \mathrm{Rm}|^2 = \Delta|\nabla \mathrm{Rm}|^2 - 2\left|\nabla^2 \mathrm{Rm}\right|^2 + \mathrm{Rm} * (\nabla \mathrm{Rm})^{*2}. \tag{7.2}$$

There are two obstacles to our deriving a satisfactory estimate for $|\nabla \mathrm{Rm}|^2$ directly from equation (7.2). The first difficulty is the potentially bad term $\mathrm{Rm} * (\nabla \mathrm{Rm})^{*2}$ on the right-hand side. The second problem is that our assumptions give us no control on $|\nabla \mathrm{Rm}|^2$ at $t = 0$. To circumvent these obstacles, we define

$$F \doteqdot t|\nabla \mathrm{Rm}|^2 + \beta|\mathrm{Rm}|^2,$$

where β is a constant to be chosen below. The strategy for this choice is that when t is small, it will allow us to control the bad term we get when we differentiate $|\nabla \operatorname{Rm}|^2$ by the good term $-2\beta |\nabla \operatorname{Rm}|^2$ we get when we differentiate $|\operatorname{Rm}|^2$. Moreover, we get an upper bound $F \leq \beta K^2$ at $t = 0$ by hypothesis on $|\operatorname{Rm}|$, without needing further assumptions on the initial metric. Combining equation (7.2) with the result of Lemma 7.4 and discarding the term $-2t \left|\nabla^2 \operatorname{Rm}\right|^2$, we observe that F satisfies the differential inequality

$$\frac{\partial}{\partial t} F \leq \Delta F + (1 + c_1 t |\operatorname{Rm}| - 2\beta) |\nabla \operatorname{Rm}|^2 + c_2 \beta |\operatorname{Rm}|^3,$$

where the constants c_1, c_2 depend only on n. By hypothesis, $|\operatorname{Rm}| \leq K$ for all $t \in [0, \alpha/K]$. So for such times,

$$\frac{\partial}{\partial t} F \leq \Delta F + (1 + c_1 \alpha - 2\beta) |\nabla \operatorname{Rm}|^2 + c_2 \beta K^3.$$

Choose $\beta \geq (1 + c_1 \alpha)/2$, noting that β depends only on n and $\max\{\alpha, 1\}$. Then for $t \in [0, \alpha/K]$,

$$\frac{\partial}{\partial t} F \leq \Delta F + c_2 \beta K^3.$$

Thus by the parabolic maximum principle, we have

$$\sup_{x \in \mathcal{M}^n} F(x,t) \leq \beta K^2 + c_2 \beta K^3 t \leq (1 + c_2 \alpha) \beta K^2 \leq C_1^2 K^2$$

for $0 \leq t \leq \alpha/K$, where C_1 is a constant depending only on n and $\max\{\alpha, 1\}$. Hence

$$|\nabla \operatorname{Rm}| \leq \sqrt{\frac{F}{t}} \leq \frac{C_1 K}{t^{1/2}} \qquad \text{for} \quad 0 < t \leq \frac{\alpha}{K}.$$

This completes the proof of the case $m = 1$.

Now by induction, we may assume that we have estimated $\left|\nabla^j \operatorname{Rm}\right|$ for all $1 \leq j < m$. Let $1 \leq k \leq m$. We begin by computing

$$\frac{\partial}{\partial t} \left|\nabla^k \operatorname{Rm}\right|^2 = 2 \left\langle \frac{\partial}{\partial t} \left(\nabla^k \operatorname{Rm}\right), \nabla^k R \right\rangle + \operatorname{Rc} * \left(\nabla^k \operatorname{Rm}\right)^{*2}. \tag{7.3}$$

To evaluate the time derivative on the right-hand side, we again recall formulas (6.1) and (6.17) and write

$$\begin{aligned}
\frac{\partial}{\partial t} \left(\nabla^k \operatorname{Rm}\right) &= \nabla^k \left(\frac{\partial}{\partial t} \operatorname{Rm}\right) + \sum_{j=0}^{k-1} \nabla^j \left(\nabla \operatorname{Rc} * \nabla^{k-1-j} \operatorname{Rm}\right) \\
&= \nabla^k \left(\Delta \operatorname{Rm} + (\operatorname{Rm})^{*2}\right) + \sum_{j=1}^{k} \nabla^j \operatorname{Rm} * \nabla^{k-j} \operatorname{Rm} \\
&= \nabla^k \Delta \operatorname{Rm} + \sum_{j=0}^{k} \nabla^j \operatorname{Rm} * \nabla^{k-j} \operatorname{Rm}.
\end{aligned}$$

In order to put (7.3) into the form of a heat equation, we need $\Delta\nabla^k \mathrm{Rm}$. Since for any tensor A, the commutator $\left[\nabla^k, \Delta\right] A$ is given by

$$\left[\nabla^k, \Delta\right] A = \nabla^k \Delta A - \Delta \nabla^k A = \sum_{j=0}^{k} \nabla^j \mathrm{Rm} * \nabla^{k-j} A,$$

we find that

$$\frac{\partial}{\partial t} \nabla^k \mathrm{Rm} = \Delta \nabla^k \mathrm{Rm} + \sum_{j=0}^{k} \nabla^j \mathrm{Rm} * \nabla^{k-j} \mathrm{Rm}.$$

Substituting this formula into (7.3) and recalling that

$$2 \langle \Delta A, A \rangle = \Delta |A|^2 - 2 |\nabla A|^2,$$

we obtain

$$\frac{\partial}{\partial t} \left|\nabla^k \mathrm{Rm}\right|^2 = \Delta \left|\nabla^k \mathrm{Rm}\right|^2 - 2 \left|\nabla^{k+1} \mathrm{Rm}\right|^2 \tag{7.4a}$$

$$+ \sum_{j=0}^{k} \nabla^j \mathrm{Rm} * \nabla^{k-j} \mathrm{Rm} * \nabla^k \mathrm{Rm}. \tag{7.4b}$$

Now applying identity (7.4) in the case $k = m$, we get a differential inequality

$$\frac{\partial}{\partial t} |\nabla^m \mathrm{Rm}|^2 \leq \Delta |\nabla^m \mathrm{Rm}|^2 + \sum_{j=0}^{m} c_{mj} \left|\nabla^j \mathrm{Rm}\right| \cdot \left|\nabla^{m-j} \mathrm{Rm}\right| \cdot |\nabla^m \mathrm{Rm}|,$$

where the constants c_{mj} depend only on j, m, and n. The inductive hypothesis then gives an estimate

$$\begin{aligned} \frac{\partial}{\partial t} |\nabla^m \mathrm{Rm}|^2 &\leq \Delta |\nabla^m \mathrm{Rm}|^2 + (c_{m0} + c_{mm}) K |\nabla^m \mathrm{Rm}|^2 \\ &\quad + \left(\sum_{j=1}^{m-1} c_{mj} \frac{C_j}{t^{j/2}} \frac{C_{m-j}}{t^{(m-j)/2}} \right) K^2 |\nabla^m \mathrm{Rm}| \\ &\leq \Delta |\nabla^m \mathrm{Rm}|^2 + K \left(C'_m |\nabla^m \mathrm{Rm}|^2 + \frac{C''_m}{t^{m/2}} K |\nabla^m \mathrm{Rm}| \right) \end{aligned}$$

on the time interval $0 < t \leq \alpha/K$, where the constants C'_m and C''_m depend only on m and n. Completing the square on the right-hand side and using the fact that $(a+b)^2 \leq 2\left(a^2 + b^2\right)$, we obtain $\tilde{C}_m$ depending only on m and n such that

$$\frac{\partial}{\partial t} |\nabla^m \mathrm{Rm}|^2 \leq \Delta |\nabla^m \mathrm{Rm}|^2 + \tilde{C}_m K \left(|\nabla^m \mathrm{Rm}|^2 + \frac{K^2}{t^m} \right). \tag{7.5}$$

As in the case $m = 1$, we shall not try to control $|\nabla^m \mathrm{Rm}|^2$ directly from this equation; instead, we define

$$G = t^m |\nabla^m \mathrm{Rm}|^2 + \beta_m \sum_{k=1}^{m} \frac{(m-1)!}{(m-k)!} t^{m-k} \left|\nabla^{m-k} \mathrm{Rm}\right|^2,$$

where β_m is a constant to be chosen momentarily. We will explain the full strategy behind this choice after we estimate the evolution equation satisfied by G. For now, we simply observe that our assumption on $|\mathrm{Rm}|$ implies that $G \leq \beta_m (m-1)!\, K^2$ at $t = 0$. By (7.4) and the inductive hypothesis, there are constants $\bar{C}_k$ depending only on k and n such that for all $1 \leq k < m$ we have

$$\frac{\partial}{\partial t}\left|\nabla^k \mathrm{Rm}\right|^2 \leq \Delta\left|\nabla^k \mathrm{Rm}\right|^2 - 2\left|\nabla^{k+1} \mathrm{Rm}\right|^2 + \frac{\bar{C}_k K^3}{t^k} \tag{7.6}$$

on the time interval $0 < t \leq \alpha/K$. It will be significant that we retain the good term $-2\left|\nabla^{k+1} \mathrm{Rm}\right|^2$ that appears in the inequality (7.6) for $k < m$, although we discarded it in estimate (7.5). Indeed, using formulas (7.5) and (7.6), we compute that G satisfies the differential inequality

$$\begin{aligned} \frac{\partial}{\partial t} G \leq{} & \Delta G + \bar{C}_m K t^m \left|\nabla^m \mathrm{Rm}\right|^2 + \bar{C}_m K^3 + m t^{m-1}\left|\nabla^m \mathrm{Rm}\right|^2 \\ & + \beta_m \sum_{k=1}^{m} \frac{(m-1)!}{(m-k)!} \left\{ \begin{array}{c} -2t^{m-k}\left|\nabla^{m-k+1} \mathrm{Rm}\right|^2 + \bar{C}_{m-k} K^3 \\ + (m-k)\, t^{m-k-1}\left|\nabla^{m-k} \mathrm{Rm}\right|^2 \end{array} \right\}. \end{aligned}$$

This reveals why we defined G as we did: the good terms

$$-2\frac{(m-1)!}{(m-k)!} t^{m-k}\left|\nabla^{m-k+1} \mathrm{Rm}\right|^2$$

that we get when we differentiate $\left|\nabla^{m-k} \mathrm{Rm}\right|^2$ compensate for the bad terms

$$\frac{(m-1)!}{(m-k+1)!}(m-k+1)\, t^{m-k}\left|\nabla^{m-k+1} \mathrm{Rm}\right|^2$$

we got when we differentiated $t^{m-(k-1)}$. In this way, we get the estimate

$$\frac{\partial}{\partial t} G \leq \Delta G + \left(\bar{C}_m K t + m - 2\beta_m\right) t^{m-1}\left|\nabla^m \mathrm{Rm}\right|^2 + \left(\bar{C}_m + \beta \bar{C}'_m\right) K^3,$$

where

$$\bar{C}'_m \doteqdot \sum_{k=1}^{m} \frac{(m-1)!}{(m-k)!} \bar{C}_{m-k}.$$

Choose $\beta_m \geq \left(\bar{C}_m \alpha + m\right)/2$, noting that β_m only depends on m, n, and $\max\{\alpha, 1\}$. Then for $t \in [0, \alpha/K]$, we have

$$\frac{\partial}{\partial t} G \leq \Delta G + \left(\bar{C}_m + \beta \bar{C}'_m\right) K^3.$$

Since $G \leq \beta_m (m-1)!\, K^2$ at $t = 0$, the maximum principle implies that

$$\sup_{x \in \mathcal{M}^n} G(x,t) \leq \beta_m (m-1)!\, K^2 + \left(\bar{C}_m + \beta \bar{C}'_m\right) K^3 t \leq C_m^2 K^2$$

for $0 \leq t \leq \alpha/K$, where $C_m \doteqdot \sqrt{\beta_m (m-1)! + \alpha \left(\bar{C}_m + \beta \bar{C}'_m\right)}$ depends only on m, n, and $\max\{\alpha, 1\}$. Thus we have

$$|\nabla^m \mathrm{Rm}| \leq \sqrt{\frac{G}{t^m}} \leq \frac{C_m K}{t^{m/2}} \qquad \text{for} \quad 0 < t \leq \frac{\alpha}{K}.$$

This completes the inductive step, hence our proof of Theorem 7.1. □

3. The Compactness Theorem

The **Compactness Theorem** for the Ricci flow [**64**] tells us that any sequence of complete solutions to the Ricci flow having curvatures uniformly bounded from above and below and injectivity radii uniformly bounded from below contains a convergent subsequence. This result has its roots in the convergence theory developed by Cheeger and Gromov. In many contexts where this theory is applied, regularity is a crucial issue. By contrast, the proof of the Compactness Theorem for the Ricci flow is greatly aided by the fact that sequences of solutions to the Ricci flow enjoy excellent regularity properties. (See Corollary 7.3 above.) Indeed, it is precisely because bounds on the curvature of a solution to the Ricci flow imply bounds on all derivatives of the curvature that the compactness theorem produces C^∞ convergence on compact sets. For this reason, we state it in this chapter. We plan to present a detailed proof in a successor to this volume.

The most important application of the compactness theorem is the formation of singularity models (final time limit flows). These are complete nonflat solutions of the Ricci flow that occur as limits of sequences of dilations about a singularity. Indeed, we already used it for this purpose in Section 15 of Chapter 5. We shall discuss singularity models in greater generality in Chapter 8.

THEOREM 7.8 (Compactness Theorem). *Let*

$$\{\mathcal{M}_i^n, g_i(t), O_i, F_i : i \in \mathbb{N}\}$$

be a sequence of complete solutions to the Ricci flow existing for $t \in (\alpha, \omega)$, where $-\infty \leq \alpha < 0 < \omega \leq \infty$. Each solution is marked by an origin $O_i \in \mathcal{M}_i^n$ and a frame $F_i = \left\{e_1^i, \ldots, e_n^i\right\}$ at O_i which is orthonormal with respect to $g_i(0)$. Suppose that there exists $K < \infty$ such that the sectional curvatures of the sequence are uniformly bounded by K in the sense that

$$\sup_{\mathcal{M}_i^n \times (\alpha, \omega)} |\mathrm{sect}[g_i]| \leq K$$

for all $i \in \mathbb{N}$. Suppose further that there exists $\delta > 0$ such that the injectivity radii of the sequence are bounded at $O_i \in \mathcal{M}_i^n$ and $t = 0$ in the sense that

$$\mathrm{inj}_{g_i(0)}(O_i) \geq \delta$$

for all $i \in \mathbb{N}$. Then there exists a subsequence which converges in the pointed category to a complete solution

$$\{\mathcal{M}_\infty^n, g_\infty(t), O_\infty, F_\infty\}$$

of the Ricci flow existing for $t \in (\alpha, \omega)$ with the properties that

$$\sup_{\mathcal{M}^n_\infty \times (\alpha,\omega)} |\mathrm{sect}\,[g_\infty]| \leq K$$

and

$$\mathrm{inj}_{\,g_\infty(0)}\,(O_\infty) \geq \delta.$$

The convergence is in the following specific sense: there exist a sequence of open sets $\mathcal{U}_i \subseteq \mathcal{M}^n_\infty$ such that $O_\infty \in \mathcal{U}_i$ for all $i \in \mathbb{N}$ and such that any compact subset of $\mathcal{M}^n_\infty$ is contained in $\mathcal{U}_i$ for all i sufficiently large, a sequence of open sets $\mathcal{V}_i \in \mathcal{M}^n_i$ such that $O_i \in \mathcal{V}_i$ for all $i \in \mathbb{N}$, and a sequence of diffeomorphisms $\varphi_i : \mathcal{U}_i \to \mathcal{V}_i$ such that $\varphi_i(O_\infty) = O_i$ and $(\varphi_i)_*(F_\infty) = F_i$ for all $i \in \mathbb{N}$; these diffeomorphisms have the property that the pullbacks $\varphi_i(g_i)$ converge uniformly to g_∞ in every C^m norm on any compact subset of $\mathcal{M}^n_\infty \times (\alpha, \omega)$.

Notes and commentary

Two good references for derivative estimates that apply to general parabolic equations are [**94**] and [**89**]. The reader may also be interesting in consulting some of Bernstein's original papers, notably [**16, 17, 18**]. For the Ricci flow in particular, the early Bemelmans–Min-Oo–Ruh paper [**11**], Bando's paper [**9**], and Shi's papers [**117, 118**] are all useful references. Our approach is a modification of the method outlined in Section 7 of [**63**].

The convergence theory of Cheeger and Gromov and in particular the Gromov–Hausdorff convergence of Riemannian manifolds is currently an active and rich area of research. An interested reader is urged to consult the influential book [**53**] and the fine survey [**109**]. Good examples of convergence results in the literature can be found in [**2**], [**26**], [**45**], [**50, 51**], [**108**], and [**130, 131**].

CHAPTER 8

Singularities and the limits of their dilations

In this chapter, we shall survey the standard classification of maximal solutions to the Ricci flow. From the perspective of this classification, one regards a solution $(\mathcal{M}^n, g(t))$ that exists up to a maximal time $T \leq \infty$ as becoming singular at T, because it cannot be extended further in time. So each maximal solution may in this way be thought of as a singular solution. We shall then study how singularities may be removed by dilations. In many cases, one can pass to a limit flow, called a *singularity model,* whose properties can yield information about the geometry of the original manifold near the singularity just prior to its formation. We discussed some examples of singularity models in Chapter 2. Singularity models always exist on infinite time intervals. In dimension $n = 3$, singularity models have other special properties (in particular nonnegative sectional curvature) that imply that they are simple topologically, and thus make it reasonable to expect that one could obtain a classification of such limit solutions that is adequate to derive geometric and topological conclusions about the underlying manifold $\mathcal{M}^3$ of the original singular solution $g(t)$. In Chapter 9, we shall see some of the reasons why singularity models in dimension three are so nice.

1. Classifying maximal solutions

Consider a solution $(\mathcal{M}^n, g(t))$ of the Ricci flow which exists on a maximal time interval $[0, T)$, where $T \in (0, \infty]$. We call T the **singularity time.** This divides solutions into two categories: those where $T < \infty$ and those where $T = \infty$. In both cases, there are further subdivisions into two types of singularities.

To motivate these subdivisions, we look to the simplest solutions, those which are the fixed points of the normalized Ricci flow. So let us consider the evolution $(\mathcal{M}^n, g(t))$ of an Einstein metric by the Ricci flow. There are three possibilities: the scalar curvature R may be positive, zero or negative. If $R(x, 0) > 0$, it follows from the evolution equation

$$\frac{\partial R}{\partial t} = \Delta R + 2\,|\mathrm{Rc}|^2 = \Delta R + \frac{2}{n} R^2$$

and the fact that R is constant in space that $T < \infty$; in fact,

$$R(x, t) \equiv \frac{n}{2(T - t)}$$

for all $x \in \mathcal{M}^n$. Then since $|\mathrm{Rm}(t)| = CR$, where $C > 0$ is a constant depending only on $g(0)$, we have

$$|\mathrm{Rm}(t)|(T-t) \equiv C$$

for another positive constant C.

Analogously, if $R(x,0) < 0$, then

$$R(x,t) \equiv \frac{n}{nR(0)^{-1} - 2t},$$

so $T = \infty$ and

$$|\mathrm{Rm}(t)|(C_1 + t) \equiv C_2$$

for some $C_1, C_2 > 0$. Finally if $R(x,0) \equiv 0$, then $R \equiv 0$, $T = \infty$ and

$$|\mathrm{Rm}(t)| \equiv C \geq 0.$$

These considerations lead us to subdivide solutions of the Ricci flow into two types: those whose curvatures are bounded above by a constant times the curvature of an Einstein solution with $R > 0$ or $R < 0$ (Type I or III) and those whose curvatures are not (Type IIa or IIb). More precisely, the classification of maximal solutions $(\mathcal{M}^n, g(t))$ by their **singularity type** is as follows.

If the singularity time T is finite, Theorem 6.45 tells us that the curvature becomes unbounded,

$$\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}| = \infty,$$

and classify the solution by its **curvature blow-up rate** — the rate at which the curvature approaches infinity as time approaches T. (Curvature blow-up rates were discussed briefly in Subsection 6.2 of Chapter 2.)

DEFINITION 8.1. Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow that exists up to a maximal time $T \leq \infty$.

- One says $(\mathcal{M}^n, g(t))$ encounters a **Type I Singularity** if $T < \infty$ and
$$\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}(\cdot,t)|(T-t) < \infty.$$
- One says $(\mathcal{M}^n, g(t))$ encounters a **Type IIa Singularity** if $T < \infty$ and
$$\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}(\cdot,t)|(T-t) = \infty.$$
- One says $(\mathcal{M}^n, g(t))$ encounters a **Type IIb Singularity** if $T = \infty$ and
$$\sup_{\mathcal{M}^n \times [0,\infty)} |\mathrm{Rm}(\cdot,t)|\, t = \infty.$$
- One says $(\mathcal{M}^n, g(t))$ encounters a **Type III Singularity** if $T = \infty$ and
$$\sup_{\mathcal{M}^n \times [0,\infty)} |\mathrm{Rm}(\cdot,t)|\, t < \infty.$$

An Einstein solution with $R > 0$ is Type I; an Einstein solution with $R < 0$ is Type III; and a nonflat Einstein solution with $R = 0$ is Type IIb. Although Type IIa singularities are strongly conjectured to exist, it is believed (but not proved) that they will occur only for very special initial data. The degenerate neckpinch discussed heuristically in Section 6 of Chapter 2 is simplest example in which a Type IIa singularity is expected to form.

REMARK 8.2. Perelman's recent work [**105, 106, 107**] does not explicitly distinguish between Type I and Type IIa singularities. Nonetheless, the concepts of rapidly-forming and slowly-forming singularities are important in so many aspects of the analysis of geometric evolution equations that we will endeavor to help the reader be aware of the distinctions between these singularity types for the Ricci flow.

2. Singularity models

Singularity models are complete nonflat solutions which occur as limits of dilations about singularities. Any such solution exists for an infinite time interval. Each type of singular solution (I, IIa, IIb, and III) has its own singularity model.

For insight, we turn again to the Einstein metrics. Notice that an Einstein solution with $R < 0$ is an **immortal solution**: a solution that exists on the future time interval $[0, \infty)$. Indeed, we may extend such a solution to its maximal time interval of existence via the relation

$$g(t) = \left(1 - \frac{2R(0)}{n} t\right) g(0), \qquad t \in \left(\frac{n}{2R(0)}, \infty\right).$$

An Einstein solution with $R > 0$ on a time interval $[0, T)$ with $T < \infty$ may be uniquely extended backward in time to an **ancient solution**: a solution that exists on the past time interval $(-\infty, T)$. Here

$$g(t) = \left(1 - \frac{t}{T}\right) g(0), \qquad t \in (-\infty, T).$$

Finally, an Einstein solution with $R = 0$ (namely, a Ricci-flat solution) may be uniquely extended to an **eternal solution**: a solution that exists for all time $(-\infty, \infty)$. Clearly, a Ricci-flat metric is static:

$$g(t) = g(0) \qquad t \in (-\infty, +\infty).$$

Singularity models will turn out to be solutions of one of these three types — ancient, eternal, or immortal. (Recall that we saw other examples of ancient, eternal, and immortal solutions in Chapter 2.)

DEFINITION 8.3. Let $(\mathcal{M}_\infty^n, g_\infty(t))$ be a limit solution of the Ricci flow.

- One says $(\mathcal{M}_\infty^n, g_\infty(t))$ is an **ancient Type I singularity model** if it exists on a time interval $(-\infty, \omega)$ containing $t = 0$ and satisfies the curvature bound

$$\sup_{\mathcal{M}^n_\infty \times (-\infty,0]} |\mathrm{Rm}_\infty(\cdot,t)| \cdot |t| < \infty.$$

- One says $(\mathcal{M}^n_\infty, g_\infty(t))$ is an **ancient Type II singularity model** if it exists on a time interval $(-\infty,\omega)$ containing $t=0$ and satisfies the curvature bound

$$\sup_{\mathcal{M}^n_\infty \times (-\infty,0]} |\mathrm{Rm}_\infty(\cdot,t)| \cdot |t| = \infty.$$

- One says $(\mathcal{M}^n_\infty, g_\infty(t))$ is an **eternal Type II singularity model** if it exists for all $t \in (-\infty,\infty)$ and satisfies the curvature bound

$$\sup_{\mathcal{M}^n_\infty \times (-\infty,\infty)} |\mathrm{Rm}_\infty(\cdot,t)| < \infty.$$

- One says $(\mathcal{M}^n_\infty, g_\infty(t))$ is an **immortal Type III singularity model** if it exists on a time interval $(-\alpha,\infty)$ containing $t=0$ and satisfies the curvature bound

$$\sup_{\mathcal{M}^n_\infty \times [0,\infty)} |\mathrm{Rm}_\infty(\cdot,t)| \cdot t < \infty.$$

REMARK 8.4. As we shall see when we study dilation below, it is always possible to apply a more careful 'point picking argument' and thereby to choose the sequence of points and times about which one dilates so that the singularity model satisfies

$$\sup_{x\in\mathcal{M}^n} |\mathrm{Rm}_\infty(x,0)| = |\mathrm{Rm}_\infty(y,0)| \tag{8.1}$$

for some $y \in \mathcal{M}^n$. If the curvature operator of the solution $(\mathcal{M}^n_\infty, g_\infty(t))$ is nonnegative, one can replace this with

$$\sup_{x\in\mathcal{M}^n} R_\infty(x,0) = R_\infty(y,0). \tag{8.2}$$

The latter condition allows application of the strong maximum principle and a differential Harnack inequality. (We introduced Harnack estimates in Section 10 of Chapter 5 and will see them again in Section 6 of Chapter 9. We will make a thorough study of differential Harnack estimates in the successor to this volume.) Condition (8.2) is useful primarily for Type II singularities. (For example, see Section 6.)

REMARK 8.5. By Theorem 9.4, any limit $\left(\mathcal{M}^3_\infty, g_\infty(t)\right)$ of a finite-time singularity in dimension $n=3$ has nonnegative curvature operator. So in dimension 3, one can always satisfy (8.2) by making appropriate choices of points and times about which to dilate.

REMARK 8.6. The distinction between ancient and eternal Type II models should be carefully noted. It arises because any limit of a Type II singularity is noncompact. In order to study its geometry at spatial infinity, one may take a second limit by a procedure called *dimension reduction.* (Dimension reduction will be discussed briefly in Section 4 and thoroughly in a planned successor to this volume.) This technique produces ancient solutions which are either Type I or Type II.

EXAMPLE 8.7. The Rosenau solution introduced in Section 3.3 is an example of a Type II ancient solution that is not eternal but has a 'backward limit' (as described in Section 6 below) that is an eternal solution. In fact, this limit is the cigar soliton studied in Section 2.1.

REMARK 8.8. It is finite-time singularities (Type I and Type IIa) that are most important in applications of the Ricci flow program towards proving the Geometrization Conjecture for closed 3-manifolds.

EXAMPLE 8.9. In dimension $n = 3$, the neckpinch (Section 5 of Chapter 2) and the conjectured degenerate neckpinch (Section 6 of Chapter 2) are the canonical examples of Type I and Type IIa singularities, respectively.

3. Parabolic dilations

The basic idea of dilating about a finite-time singularity is to choose a sequence of points and times where the norm of the curvature tends to infinity and is comparable to its maximum in sufficiently large spatial and temporal neighborhoods of the chosen points and times. Heuristically, one trains a microscope at those neighborhoods and magnifies so that the norm of the curvature is uniformly bounded from above and is equal to 1 at the chosen points. Because Perelman's *No Local Collapsing* theorem [**105**] provides an appropriate injectivity estimate in those neighborhoods, one can then obtain a nontrivial limit solution of the Ricci flow. Complete solutions which arise as such limits have special properties which suggest that they can be classified in a way that will yield information about the singularity. We shall make these ideas precise in what follows. The case of an infinite time singularity is analogous, and will be considered in Section 5.

3.1. How to choose a sequence of points and times. If $(\mathcal{M}^n, g(t))$ is a solution of the Ricci flow on a maximal time interval $0 \leq t < T \leq \infty$, we may study its properties by dilating the solution about a sequence of points and times (x_i, t_i), where $x_i \in \mathcal{M}^n$ and $t_i \nearrow T$. Of course, there are many possible sequences of points and times around which to dilate, and the limits we get may depend on our choices. However, each sequence should at the very least satisfy certain conditions which we outline below. The guiding principle is that if we want to get a complete smooth limit flow, then we want the norm of the curvature of each dilated solution to be suitably bounded in an appropriate parabolic neighborhood of each (x_i, t_i).

In order to develop intuition for the dilation criteria introduced below, recall that if the singularity time T is finite, then Theorem 6.54 tells us that

$$\lim_{t_i \nearrow T} \sup_{\mathcal{M}^n} |\mathrm{Rm}(\cdot, t_i)| = \infty \tag{8.3}$$

regardless of whether the singularity is forming quickly (Type I) or slowly (Type IIa). In light of this fact, one may choose the sequence of times t_i so that the maximum of the curvature at time t_i is comparable to its maximum over sufficiently large previous time intervals. Similarly, one may choose the sequence of points x_i so that the curvature at (x_i, t_i) is comparable to its global maximum at time t_i. Occasionally (for example, in dimension reduction) one chooses points x_i so that the curvature at (x_i, t_i) is comparable to its local maximum over a sufficiently large ball. (In some cases, it may be possible to get a complete limit for *any* sequence (x_i, t_i) such that $|\mathrm{Rm}(x_i, t_i)| \to \infty$ if one has good enough bounds on $|\nabla \mathrm{Rm}|$; we shall not discuss such limits in this volume, however.)

Motivated by these considerations, we adopt the following definitions for any solution $(\mathcal{M}^n, g(t))$ of the Ricci flow on a maximal time interval $0 \leq t < T \leq \infty$. Notice that they reflect the way parabolic PDE naturally equate time with distance squared.

DEFINITION 8.10. We say a sequence (x_i, t_i) is **globally curvature essential** if $t_i \nearrow T \leq \infty$ and there exists a constant $C \geq 1$ and a sequence of radii $r_i \in \left(0, \sqrt{t_i}\right)$ such that

$$\sup \left\{ |\mathrm{Rm}(x,t)| : x \in \bar{B}_{g(t)}(x_i, r_i),\ t \in \left[t_i - r_i^2, t_i\right] \right\} \leq C\, |\mathrm{Rm}(x_i, t_i)| \tag{8.4}$$

for all $i \in \mathbb{N}$, where

$$\lim_{i \to \infty} r_i^2 \, |\mathrm{Rm}(x_i, t_i)| = \infty. \tag{8.5}$$

DEFINITION 8.11. We say a sequence (x_i, t_i) is **locally curvature essential** if $t_i \nearrow T \leq \infty$ and there exists a constant $C < \infty$ and a sequence of radii $r_i \in \left(0, \sqrt{t_i}\right)$ such that for all $i \in \mathbb{N}$, one has

$$\sup \left\{ |\mathrm{Rm}(x,t)| : x \in \bar{B}_{g(t)}(x_i, r_i),\ t \in \left[t_i - r_i^2, t_i\right] \right\} \leq C r_i^{-2}. \tag{8.6}$$

In this volume, we shall deal mainly with globally essential dilation sequences. In a later volume, we hope to discuss the relevance of locally essential sequences to the surgery arguments in [**106**].

REMARK 8.12. In some cases, one works with a special case of condition (8.4) in which

$$\sup \left\{ |\mathrm{Rm}(x,t)| : x \in \mathcal{M}^n,\ t \in [0, t_i] \right\} \leq C\, |\mathrm{Rm}(x_i, t_i)|. \tag{8.7}$$

REMARK 8.13. By (8.4), any smooth limit of a globally curvature essential sequence formed by the parabolic dilations (8.8) has its curvature bounded by some function $K(t)$. Moreover, property (8.5) ensures that any such limit is complete. A smooth limit of a locally curvature essential sequence will have bounded curvature on a metric ball of radius 1 for times $-1 \leq t \leq 0$.

Once we have chosen a sequence (x_i, t_i) of points and times such that $t_i \nearrow T \in (0, \infty]$, we consider a sequence of **parabolic dilations**: the solutions $(\mathcal{M}^n, g_i(t))$ defined by

$$g_i(t) \doteqdot |\mathrm{Rm}\,(x_i, t_i)| \cdot g\left(t_i + \frac{t}{|\mathrm{Rm}\,(x_i, t_i)|}\right) \tag{8.8}$$

that exist for

$$-t_i\,|\mathrm{Rm}\,(x_i, t_i)| \leq t < (T - t_i)\,|\mathrm{Rm}\,(x_i, t_i)|\,.$$

(The right endpoint is ∞ if $T = \infty$.) The translation in time ensures that $g_i(0)$ is a homothetic multiple of $g(t_i)$, and the dilation in time ensures that $g_i(t)$ is still a solution of the Ricci flow. The spatial dilation is chosen so that the curvature $\mathrm{Rm}\,[g_i]$ of the new metric g_i has norm 1 at the new 'origin' x_i and the new time 0, namely so that

$$|\mathrm{Rm}\,[g_i]\,(x_i, 0)|_{g_i} = 1. \tag{8.9}$$

This guarantees that the limit of the pointed solutions $(\mathcal{M}^n, g_i(t), x_i)$, if it exists, will not be flat.

REMARK 8.14. In dimension 3, one can replace $|\mathrm{Rm}|$ by R in the discussion above. Indeed, the ODE estimate for the curvature obtained in Lemma 9.10 implies that there exist constants $C, C' > 0$ such that

$$R \geq C\,|\mathrm{Rm}| - C'$$

in dimension $n = 3$. But in any dimension n, there exists C_n depending on n such that

$$R \leq C_n\,|\mathrm{Rm}|\,.$$

Hence the scalar curvature R is equivalent to $|\mathrm{Rm}|$ whenever the latter is large enough.

If a sequence (x_i, t_i) is globally curvature essential, then estimate (8.4) implies that $|\mathrm{Rm}\,[g_i]\,(x, t)|$ will be bounded on arbitrarily large balls as $i \to \infty$. So in order for the limit solution to exist, it suffices to have

$$\mathrm{inj}\,[g_i]\,(x_i, 0) \geq \mathrm{c} > 0$$

for some constant $\mathrm{c} > 0$ independent of i. A sufficient condition is given in the following.

DEFINITION 8.15. A solution $(\mathcal{M}^n, g(t))$ of the Ricci flow on a time interval $[0, T)$ is said to satisfy a **global injectivity radius estimate on the scale of its maximum curvature** if there exists a constant $\mathrm{c} > 0$ such that

$$\mathrm{inj}\,(x, t)^2 \geq \frac{c}{\sup_{\mathcal{M}^n} |\mathrm{Rm}(\cdot, t)|}$$

for all $(x, t) \in \mathcal{M}^n \times [0, T)$.

In more advanced applications to be considered in a later volume, one may replace this by a local condition.

DEFINITION 8.16. A solution $(\mathcal{M}^n, g(t))$ of the Ricci flow on a time interval $[0,T)$ is **κ-noncollapsed on the scale ρ** if for every metric ball $B_{g(t)}(x_0, r)$ of radius $r < \rho$ in which the curvature is bounded from above in the sense that

$$\sup\left\{|\mathrm{Rm}(x,t)| : x \in B_{g(t)}(x_0, r),\ 0 \leq t < T\right\} \leq r^{-2},$$

one has a lower bound on volume of the type

$$\mathrm{Vol}\left[B_{g(t)}(x_0, r)\right] \geq \kappa r^n.$$

If $(\mathcal{M}^n, g(t))$ obeys a suitable injectivity radius estimate, the Compactness Theorem from Section 3 of Chapter 7 shows that a subsequence of the pointed sequence $(\mathcal{M}^n, g_i(t), x_i)$ converges to a complete pointed solution $(\mathcal{M}^n_\infty, g_\infty(t), x_\infty)$ of the Ricci flow on an ancient time interval $-\infty < t < \omega \leq \infty$. This singularity model satisfies

$$|\mathrm{Rm}[g_\infty](x,t)|_{g_\infty} \leq C$$

for all $(x,t) \in \mathcal{M}^n_\infty \times (-\infty, 0]$.

If $(\mathcal{M}^n, g(t))$ becomes singular in finite time — in other words, whenever $(\mathcal{M}^n, g(t))$ encounters a Type I or Type IIa singularity — Perelman's *No Local Collapsing* Theorem [**105**] provides such an estimate. In particular, Perelman's result implies that there exists $\kappa > 0$ such that the singularity model $(\mathcal{M}^n_\infty, g_\infty(t), x_\infty)$ is κ-noncollapsed at all scales.

THEOREM 8.17 (Perelman). *Let $(\mathcal{M}^n, g(t))$ be a solution of the Ricci flow that encounters a singularity at some time $T < \infty$. Then there exists a constant $c > 0$ independent of t and a subsequence (x_i, t_i) such that*

$$\mathrm{inj}(x_i, t_i) \geq \frac{c}{\sqrt{\max_{\mathcal{M}^n} |\mathrm{Rm}(\cdot, t_i)|}}.$$

In forthcoming work, we plan to provide a detailed explanation of the proof.

REMARK 8.18. A global injectivity radius estimate on the scale of the maximum curvature was originally established for Type I singularities by Hamilton's isoperimetric estimate. (See §23 of [**63**].)

4. Dilations of finite-time singularities

There is a lower bound for the curvature blow-up rate which applies both to Type I and Type IIa singularities.

LEMMA 8.19. *Let $\mathcal{M}^n$ be a compact manifold. If $0 \leq t < T < \infty$ is the maximal interval of existence of $(\mathcal{M}^n, g(t))$, there exists a constant $c_0 > 0$ depending only on n such that*

$$\sup_{x \in \mathcal{M}^n} |\mathrm{Rm}(x,t)| \geq \frac{c_0}{T-t}. \tag{8.10}$$

PROOF. Recall from Lemma 7.4 that

$$\frac{\partial}{\partial t} |\mathrm{Rm}|^2 \leq \Delta |\mathrm{Rm}|^2 + C |\mathrm{Rm}|^3$$

for some C depending only on n. Define

$$K(t) \doteqdot \sup_{x \in \mathcal{M}^n} |\mathrm{Rm}(x,t)|^2 .$$

By the maximum principle, we have $dK/dt \leq CK^{3/2}$, which implies that

$$\frac{d}{dt} K^{-1/2} \geq -\frac{C}{2}.$$

Integrating this inequality from t to $\tau \in (t,T)$ and using the fact that

$$\liminf_{\tau \to T} K(\tau)^{-1/2} = 0,$$

we obtain

$$K^{-1/2}(t) \leq \frac{C}{2}(T-t).$$

Hence $\sup_{x \in \mathcal{M}^n} |\mathrm{Rm}(x,t)| \geq 2/[C(T-t)]$. □

4.1. Type I limits of Type I singularities. Given a Type I singular solution $(\mathcal{M}^n, g(t))$ on $[0,T)$, define S by

$$\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}(x,t)| (T-t) \doteqdot S < \infty. \tag{8.11}$$

Let (x_i, t_i) be any sequence that is globally curvature essential. Then $t_i \nearrow T$, and the norm of the curvature at (x_i, t_i) is comparable to its maximum at time t in the sense of (8.4). Consider the dilated solutions $(\mathcal{M}^n, g_i(t))$ of the Ricci flow defined by (8.8), and recall that $\mathrm{Rm}_i \doteqdot \mathrm{Rm}[g_i]$ satisfies $|\mathrm{Rm}_i(x_i, 0)| = 1$ for all i.

For simplicity, we will assume the global bound (8.7); the general case is essentially the same. By using definitions (8.8) and (8.11), estimate (8.4), and Lemma 8.19, we find that the solutions $(\mathcal{M}^n, g_i(t))$ obey the uniform bound

$$\begin{aligned}
|\mathrm{Rm}_i(x,t)| &= \frac{1}{|\mathrm{Rm}(x_i,t_i)|} \left| \mathrm{Rm}\left(t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right) \right| \\
&\leq C(T-t_i) \frac{S}{T - \left(t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)} \\
&= CS \left(1 - \frac{t}{(T-t_i)|\mathrm{Rm}(x_i,t_i)|}\right)^{-1} \\
&\leq CS \left(1 + \frac{|t|}{S}\right)^{-1}
\end{aligned}$$

for all $x \in \mathcal{M}^n$ and $t \in [-\alpha_i, 0]$, where $\alpha_i \doteqdot t_i |\mathrm{Rm}(x_i,t_i)|$. This implies the estimate

$$\sup_{\mathcal{M}^n \times [-\alpha_i, 0]} |t| \cdot |\mathrm{Rm}_i(x,t)| \leq CS^2 < \infty,$$

which is the curvature bound we need for a subsequence to converge to a Type I singularity model.

By classifying the limits of dilations of all sequences (x_i, t_i) chosen in this manner, one can understand a Type I singularity at all fixed relative scales in the interval $(0, 1]$, where the **relative curvature scale** of a point $x \in \mathcal{M}^n$ at time $t \in [0, T)$ is defined to be the ratio

$$\frac{|\mathrm{Rm}\,(x,t)|}{\sup_{\mathcal{M}^n} |\mathrm{Rm}\,(\cdot,t)|}.$$

There are certain properties common to any such limit.

PROPOSITION 8.20. *Let $(\mathcal{M}^n, g(t))$ exhibit a Type I singularity at $T \in (0,\infty)$, and let (x_i, t_i) be any globally curvature essential sequence. Then any limit $(\mathcal{M}^n_\infty, g_\infty(t))$ is an ancient Type I singularity model that exhibits a Type I singularity at some time $\omega < \infty$.*

PROOF. By the Type I singularity condition and Lemma 8.19, there exist positive constants $c_0 < C$ such that

$$\frac{c_0}{T-t} \leq \sup_{\mathcal{M}^n} |\mathrm{Rm}\,(\cdot,t)| \leq \frac{C}{T-t}.$$

Then by our choice of (x_i, t_i), there is $c > 0$ such that

$$\frac{c}{T-t_i} \leq |\mathrm{Rm}\,(x_i,t_i)| \leq \frac{C}{T-t_i}.$$

The solutions $(\mathcal{M}^n, g_i(t))$ defined by (8.8) exist for $-\alpha_i \leq t < \omega_i$, where

$$\alpha_i = t_i\,|\mathrm{Rm}\,(x_i,t_i)| > 0,$$
$$\omega_i = (T - t_i)\,|\mathrm{Rm}\,(x_i,t_i)| > 0.$$

For any $t \in (\alpha_i, \omega_i)$, the curvature Rm_i of g_i obeys the estimate

$$\frac{c}{C}\frac{\omega_i}{\omega_i - t} \leq |\mathrm{Rm}_i\,(x,t)| \leq \frac{C}{c}\frac{\omega_i}{\omega_i - t}. \tag{8.12}$$

Indeed, one has

$$\begin{aligned} |\mathrm{Rm}_i\,(x,t)| &= \frac{1}{|\mathrm{Rm}\,(x_i,t_i)|}\left|\mathrm{Rm}\left(t_i + \frac{t}{|\mathrm{Rm}\,(x_i,t_i)|}\right)\right| \\ &\leq \frac{T-t_i}{c}\frac{C}{T-t_i - t\,|\mathrm{Rm}\,(x_i,t_i)|^{-1}} \\ &= \frac{C}{c}\frac{\omega_i}{\omega_i - t}, \end{aligned}$$

with the other inequality being obtained similarly.

Now since $c \leq \omega_i \leq C$ for all i, we can choose a subsequence (x_i, t_i) such that $\omega \doteqdot \lim_{i\to\infty} \omega_i$ exists. Since $\alpha_i \geq T\,|\mathrm{Rm}\,(x_i,t_i)| - C$, we have $\alpha_i \to \infty$ by (8.3). So any limit $(\mathcal{M}^n_\infty, g_\infty(t))$ is defined on $(-\infty, \omega)$. For each time $t \in (-\infty, \omega)$, there is I_t so large that $t \in (-\alpha_i, \omega_i)$ for all $i \geq I_t$, whence it follows from (8.12) that $(\mathcal{M}_\infty^n, g_\infty(t))$ exhibits a Type I singularity at $\omega \in [c, C]$. $\square$

REMARK 8.21. A consequence of the results in Chapter 6 is that any solution starting at a compact 3-manifold of positive Ricci curvature encounters a Type I singularity. Furthermore, any Type I limit formed at that singularity will be a shrinking spherical space form.

We can satisfy the stronger condition (8.1) by being more careful in picking the sequence of points and times. In particular, a necessary but not sufficient condition is that the relative scales of the points (x_i, t_i) tend to 1 in the sense that

$$\lim_{i\to\infty} \frac{|\mathrm{Rm}(x_i,t_i)|}{\sup_{\mathcal{M}^n} |\mathrm{Rm}(t_i)|} = 1. \tag{8.13}$$

Note that when $\mathcal{M}^n$ is compact, we may choose x_i so that

$$|\mathrm{Rm}(x_i,t_i)| = \max_{\mathcal{M}^n} |\mathrm{Rm}(\cdot,t_i)|,$$

but this is not necessary (and not always possible when $\mathcal{M}^n$ is noncompact). Define

$$\omega \doteqdot \limsup_{t\nearrow T} |\mathrm{Rm}(\cdot,t)|(T-t). \tag{8.14}$$

By Lemma 8.19 and definition (8.11), we have $\omega \in (0,\infty)$. Take any sequence of points and times (x_i,t_i) with $t_i \nearrow T$ such that

$$|\mathrm{Rm}(x_i,t_i)|(T-t_i) \doteqdot \omega_i \to \omega. \tag{8.15}$$

Assuming (8.13) holds, this condition is equivalent to the requirement that the times t_i satisfy

$$\sup_{\mathcal{M}^n} |\mathrm{Rm}(\cdot,t_i)|(T-t_i) \to \omega.$$

In other words, we want the Type I rescaling factors to tend to their lim sup.

Consider the dilated solutions $g_i(t)$ of the Ricci flow defined by (8.8). By definition of ω, there is for every $\varepsilon > 0$ a time $t_\varepsilon \in [0,T)$ such that

$$|\mathrm{Rm}(x,t)|(T-t) \le \omega + \varepsilon$$

for all $x \in \mathcal{M}^n$ and $t \in [t_\varepsilon, T)$. The curvature norm of g_i then satisfies

$$\begin{aligned} |\mathrm{Rm}_i(x,t)| &= \frac{1}{|\mathrm{Rm}(x_i,t_i)|}\left|\mathrm{Rm}\left(x, t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\right| \\ &= \frac{\left|\mathrm{Rm}\left(x, t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\right|\left(T - t_i - \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)}{|\mathrm{Rm}(x_i,t_i)|(T-t_i) - t} \\ &\le \frac{\omega+\varepsilon}{\omega_i - t}, \end{aligned} \tag{8.16}$$

if

$$t_\varepsilon \le t_i + t\,|\mathrm{Rm}(x_i,t_i)|^{-1} < T,$$

hence if

$$t \in [-|\mathrm{Rm}(x_i, t_i)|(t_i - t_\varepsilon), \omega_i).$$

Note that $\omega_i \to \omega$ and that

$$\lim_{i \to \infty} |\mathrm{Rm}(x_i, t_i)|(t_i - t_\varepsilon) = \infty$$

for any $\varepsilon > 0$. So for a pointed limit solution $(\mathcal{M}_\infty^n, g_\infty(t), x_\infty)$ of a subsequence of $(\mathcal{M}^n, g_i(t), x_i)$, we can let $i \to \infty$ and then $\varepsilon \searrow 0$ in estimate (8.16) to conclude that

$$|\mathrm{Rm}_\infty(x, t)| \leq \frac{\omega}{\omega - t}$$

for all $x \in \mathcal{M}_\infty^n$ and $t \in (-\infty, \omega)$. Because $|\mathrm{Rm}_i(x_i, 0)| = 1$ for all i, the limit satisfies $|\mathrm{Rm}_\infty(x_\infty, 0)| = 1$.

4.2. Type II limits of Type IIa singularities. An important source of intuition into the expected behavior of Type IIa singularities is given by the conjectured degenerate neckpinch described in Section 6 of Chapter 2. Recall that the Type IIa condition says that

$$\sup_{\mathcal{M}^n \times [0,T)} |\mathrm{Rm}(\cdot, t)|(T - t) = \infty.$$

This will allow us to select a sequence of points and times so that the dilated solutions have uniformly bounded curvatures on larger and larger time intervals both forwards and backwards in time.

First choose any sequence of times $T_i \nearrow T$. Given any $c \in (0, 1)$, let (x_i, t_i) be a sequence of points and times such that $t_i \to T$ and

$$|\mathrm{Rm}(x_i, t_i)|(T_i - t_i) \geq c \sup_{\mathcal{M}^n \times [0,T_i]} |\mathrm{Rm}(x, t)|(T_i - t). \tag{8.17}$$

In other words, if we were doing Type I rescalings relative to the intervals $[0, T_i]$, we would choose the curvatures at (x_i, t_i) to be comparable to their maxima on $\mathcal{M}^n \times [0, T_i]$. Since $T_i < T$, such points and times always exist. When $\mathcal{M}^n$ is compact, they also exist for $c = 1$.

Consider the dilated solutions given by (8.8). Condition (8.17) implies that for

$$t \in [-t_i |\mathrm{Rm}(x_i, t_i)|, (T_i - t_i)|\mathrm{Rm}(x_i, t_i)|)$$

we have

$$\begin{aligned} &|\mathrm{Rm}_i(x, t)| \left(T_i - t_i - \frac{t}{|\mathrm{Rm}(x_i, t_i)|} \right) \\ &= \frac{\left| \mathrm{Rm}\left(x, t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right) \right| \left(T_i - t_i - \frac{t}{|\mathrm{Rm}(x_i,t_i)|} \right)}{|\mathrm{Rm}(x_i, t_i)|(T_i - t_i)} (T_i - t_i) \\ &\leq \frac{1}{c}(T_i - t_i), \end{aligned}$$

or equivalently,

$$\left|\mathrm{Rm}_i(x,t)\right| \leq \frac{1}{c}\frac{(T_i - t_i)\left|\mathrm{Rm}(x_i,t_i)\right|}{(T_i - t_i)\left|\mathrm{Rm}(x_i,t_i)\right| - t}.$$

Since $T_i \to T$, the Type IIa condition implies that

$$\lim_{i\to\infty}\left[\sup_{\mathcal{M}^n\times[0,T_i]}\left|\mathrm{Rm}(x,t)\right|(T_i - t)\right] = \infty.$$

Thus, by our choices of x_i and t_i we have $\left|\mathrm{Rm}(x_i,t_i)\right|(T_i - t_i) \to \infty$.

Applying Theorem 8.17 and the Compactness Theorem, one obtains a limit solution $(\mathcal{M}^n_\infty, g_\infty(t))$ that exists for all $t \in (-\infty,\infty)$ with uniformly bounded curvature

$$\sup_{\mathcal{M}^n_\infty\times(-\infty,\infty)}\left|\mathrm{Rm}_\infty\right| \leq \frac{1}{c}. \tag{8.18}$$

As in the Type I case, a more carefully chosen sequence (x_i, t_i) will guarantee that the pointed limit solution $(\mathcal{M}^n_\infty, g_\infty(t), x_\infty)$ attains its maximum curvature at $(x_\infty, 0)$. Again choose any sequence of times $T_i \nearrow T$. The basic idea is to choose points and times (x_i, t_i) such that

$$\lim_{i\to\infty}\frac{\left|\mathrm{Rm}(x_i,t_i)\right|(T_i - t_i)}{\sup_{\mathcal{M}^n\times[0,T_i]}\left|\mathrm{Rm}(x,t)\right|(T_i - t)} = 1. \tag{8.19}$$

We shall give the proof for the case that $\mathcal{M}^n$ is compact, leaving the general case to the reader. Then since $T_i < T$, there exist points and times (x_i, t_i) such that equality holds in (8.19). By repeating the argument of Section 4.2 with $c = 1$, we obtain the estimate

$$\left|\mathrm{Rm}_i(x,t)\right| \leq \frac{(T_i - t_i)\left|\mathrm{Rm}(x_i,t_i)\right|}{(T_i - t_i)\left|\mathrm{Rm}(x_i,t_i)\right| - t}$$

for all $x \in \mathcal{M}^n$ and

$$t \in \left[-t_i\left|\mathrm{Rm}(x_i,t_i)\right|, (T_i - t_i)\left|\mathrm{Rm}(x_i,t_i)\right|\right),$$

where we have

$$(T_i - t_i)\left|\mathrm{Rm}(x_i,t_i)\right| \to \infty.$$

Hence the pointed limit $(\mathcal{M}^n_\infty, g_\infty(t), x_\infty)$ is defined for all $t \in (-\infty,\infty)$ and satisfies the uniform curvature bound

$$\sup_{\mathcal{M}^n_\infty\times(-\infty,\infty)}\left|\mathrm{Rm}_\infty\right|_{g_\infty} \leq 1. \tag{8.20}$$

Moreover, equality holds at $(x_\infty, 0)$, because $\left|\mathrm{Rm}_i(x_i,0)\right| = 1$.

REMARK 8.22. When $n = 3$, we have the estimate

$$C\left|\mathrm{Rm}\right| - C' \leq R \leq C_3\left|\mathrm{Rm}\right|$$

of Lemma 9.10, which implies that

$$\sup_{\mathcal{M}^3\times[0,T)} R(x,t)(T - t) = \infty.$$

Thus we may replace $|\mathrm{Rm}|$ by R in the discussion above, whence it follows that

$$R_\infty(x,t) \le 1 = R_\infty(x_\infty, 0)$$

for a pointed Type II limit $(\mathcal{M}^3_\infty, g_\infty(t), x_\infty)$. The advantage of the scalar curvature R_∞ attaining its maximum in space and time is as follows. A stronger consequence of Theorem 9.4 is Corollary 9.7, which says that any type of limit of dilations about a finite-time singularity has a non-negative curvature operator. When R_∞ attains its maximum, this enables the application of a differential Harnack estimate together with the strong maximum principle.

EXAMPLE 8.23. An example of a rotationally symmetric Type II limit is the Bryant soliton [**21**] on $\mathbb{R}^3$.

5. Dilations of infinite-time singularities

5.1. Type II limits of Type IIb singularities. This case is dual to the Type IIa case. Recall that the Type IIb condition is

$$\sup_{\mathcal{M}^n \times [0,\infty)} t\,|\mathrm{Rm}(x,t)| = \infty. \tag{8.21}$$

Analogous to Type IIa, choose any sequence $T_j \nearrow \infty$. If we follow what we did before, it is natural to choose a sequence (x_i, t_i) such that

$$t_i\,|\mathrm{Rm}(x_i,t_i)| \doteqdot \alpha_i \to \infty$$

and

$$\frac{t_i\,|\mathrm{Rm}(x_i,t_i)|}{\sup_{\mathcal{M}^n\times[0,T_i]} t\,|\mathrm{Rm}(x,t)|} \ge c > 0.$$

This is always possible but may not be a good choice, because it does not guarantee that the curvatures of the dilated solutions can be uniformly bounded on finite time intervals. To see this, notice that the dilated solution $g_i(t)$ is defined for $t \in [-\alpha_i, \infty)$ and satisfies

$$\begin{aligned}
|\mathrm{Rm}_i(x,t)| &= \frac{1}{|\mathrm{Rm}(x_i,t_i)|}\left|\mathrm{Rm}\left(x, t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\right| \\
&= \frac{1}{t_i\,|\mathrm{Rm}(x_i,t_i)|}\left|\mathrm{Rm}\left(x, t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\right| \\
&\qquad \times \left(t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\frac{t_i\,|\mathrm{Rm}(x_i,t_i)|}{t_i\,|\mathrm{Rm}(x_i,t_i)| + t} \\
&\le \frac{1}{c}\frac{\alpha_i}{\alpha_i + t} \qquad (8.22)
\end{aligned}$$

for all $x \in \mathcal{M}^n$ and t such that

$$0 \le t_i + t\,|\mathrm{Rm}(x_i,t_i)|^{-1} \le T_i,$$

hence for all $t \in [-\alpha_i, \omega_i]$, where

$$\omega_i \doteqdot (T_i - t_i)\,|\mathrm{Rm}(x_i,t_i)|.$$

Now $\alpha_i \to \infty$, but we do not know that ω_i also tends to infinity. This motivates finding a better sequence of points and times (x_i, t_i).

Refining our choice by the method of Section 4.2, we pick points (x_i, t_i) so that

$$\frac{t_i (T_i - t_i) |\mathrm{Rm}(x_i, t_i)|}{\sup_{\mathcal{M}^n \times [0,T_i]} [t (T_i - t) |\mathrm{Rm}(x,t)|]} \doteqdot 1 - \delta_i \to 1.$$

This guarantees that $\alpha_i \to \infty$ and $\omega_i \to \infty$. Indeed, it follows from (8.21) that

$$\begin{aligned}
\frac{1}{\alpha_i^{-1} + \omega_i^{-1}} = \frac{\alpha_i \omega_i}{\alpha_i + \omega_i} &= \frac{t_i |\mathrm{Rm}(x_i, t_i)| (T_i - t_i)}{T_i} \\
&= \frac{1}{T_i} \sup_{\mathcal{M}^n \times [0,T_i]} t |\mathrm{Rm}(x,t)| (T_i - t) \\
&\geq \frac{1}{2} \sup_{\mathcal{M}^n \times [0,T_i/2]} t |\mathrm{Rm}(x,t)| \to \infty,
\end{aligned}$$

hence that $\alpha_i^{-1} \to 0$ and $\omega_i^{-1} \to 0$. Then as in (8.22), we get

$$\begin{aligned}
&|\mathrm{Rm}_i(x,t)| \\
&= \frac{\left(t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\left(T_i - t_i - \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\left|\mathrm{Rm}\left(x, t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\right|}{t_i (T_i - t_i) |\mathrm{Rm}(x_i, t_i)|} \\
&\qquad \times \frac{t_i |\mathrm{Rm}(x_i, t_i)| (T_i - t_i)}{(t_i |\mathrm{Rm}(x_i, t_i)| + t)\left(T_i - t_i - \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)} \\
&\leq \frac{1}{1 - \delta_i} \times \frac{\alpha_i}{\alpha_i + t} \frac{\omega_i}{\omega_i - t}
\end{aligned}$$

for all $x \in \mathcal{M}^n$ and $t \in [-\alpha_i, \omega_i)$. Since $\delta_i \to 0$, $\alpha_i \to \infty$, and $\omega_i \to \infty$, we conclude that the pointed limit solution $(\mathcal{M}^n_\infty, g_\infty(t), x_\infty)$, if it exists, is defined for all $t \in (-\infty, \infty)$ and satisfies the curvature bound

$$\sup_{\mathcal{M}^n_\infty \times (-\infty,\infty)} |\mathrm{Rm}_\infty| \leq 1 = |\mathrm{Rm}_\infty(x_\infty, 0)|.$$

5.2. Type III limits of Type III singularities. Recall that a solution $(\mathcal{M}^n, g(t))$ of the Ricci flow forms a Type III singularity at $T = \infty$ if the solution exists for $t \in [0, \infty)$ and satisfies

$$\sup_{\mathcal{M}^n \times [0,\infty)} t |\mathrm{Rm}(\cdot, t)| < \infty.$$

Define

$$\alpha \doteqdot \limsup_{t \to \infty} \left(t \sup_{\mathcal{M}^n} |\mathrm{Rm}(\cdot, t)| \right) \in [0, \infty). \tag{8.23}$$

The nonnegative number α is analogous to ω defined in equation (8.14). We shall now demonstrate that one may assume α is strictly positive. We begin by noting a consequence of the formula

$$\frac{d}{dt}L_t(\gamma_t) = \frac{1}{2}\int_{\gamma_t} \frac{\partial g}{\partial t}(S,S)\, ds - \int_{\gamma_t} \langle \nabla_S S, V\rangle\, ds \tag{8.24}$$

derived in Lemma 3.11.

COROLLARY 8.24. *Let γ_t be a time-dependent family of curves with fixed endpoints in a solution $(\mathcal{M}^n, g(t))$ of the Ricci flow. If γ_t is a constant path or if each γ_t is geodesic with respect to $g(t)$, then*

$$\frac{d}{dt}L_t(\gamma_t) = -\int_{\gamma_t} \operatorname{Rc}(S,S)\, ds.$$

PROOF. The last term in formula (8.24) vanishes if the paths γ_t are independent of time or variations through geodesics. □

PROPOSITION 8.25. *Let $(\mathcal{M}^n, g(t))$ be any solution of the Ricci flow on $[0,\infty)$. There exists a constant $\varepsilon > 0$ depending only on n such that $\alpha \geq \varepsilon$ if $\mathcal{M}^n$ is not a nilmanifold.*

PROOF. We claim there is $\varepsilon = \varepsilon(n) > 0$ small enough so that if

$$\alpha = \limsup_{t\to\infty}\left(t \sup_{\mathcal{M}^n} |\operatorname{Rm}(\cdot,t)|\right) \leq \varepsilon, \tag{8.25}$$

then there exist $C < \infty$, $\delta > 0$, and $T_\varepsilon < \infty$ depending only on n such that

$$\operatorname{diam}(\mathcal{M}^n, g(t)) \leq C t^{1/2-\delta} \tag{8.26}$$

for $t \geq T_\varepsilon$. Setting

$$S \doteqdot \sup_{\mathcal{M}^n\times[0,\infty)}\left(t \sup_{\mathcal{M}^n} |\operatorname{Rm}(\cdot,t)|\right) < \infty,$$

the claim implies that for all $t \in (0,\infty)$,

$$\sup_{\mathcal{M}^n} |\operatorname{Rm}(\cdot,t)| \left[\operatorname{diam}(\mathcal{M}^n, g(t))\right]^2 \leq SC^2 t^{-2\delta},$$

hence that $\mathcal{M}^n$ is a nilmanifold, because $SC^2t^{-2\delta} \to 0$ as $t \to \infty$.

To prove the claim, let $\gamma : [a,b] \to \mathcal{M}^n$ be a fixed path and put

$$L(t) \doteqdot \underset{g(t)}{\operatorname{length}}\, \gamma.$$

Then

$$\frac{dL}{dt} = -\int_\gamma \operatorname{Rc}(T,T)\, ds$$

by Corollary 8.24, and so

$$\left|\frac{dL}{dt}\right| \leq \int_\gamma |\operatorname{Rc}|_{g(t)}\, ds.$$

If (8.25) holds for some $\varepsilon > 0$, there exists a time $T_\varepsilon < \infty$ such that for all $t \geq T_\varepsilon$ we have $t \sup_{\mathcal{M}^n} |\mathrm{Rm}(\cdot, t)| \leq 2\varepsilon$. In particular,

$$\sup_{\mathcal{M}^n} |\mathrm{Rc}(\cdot, t)| \leq \frac{C\varepsilon}{t},$$

for all $t \geq T_\varepsilon$, where C depends only on n. Hence

$$\frac{dL}{dt}(\tau) \leq \left|\frac{dL}{dt}(\tau)\right| \leq \frac{C\varepsilon}{\tau} L(\tau)$$

for $\tau \geq T_\varepsilon$. Integrating this inequality from time T_ε to time $t > T_\varepsilon$ implies

$$L(t) \leq L(T_\varepsilon)\left(\frac{t}{T_\varepsilon}\right)^{C\varepsilon} = L(T_\varepsilon) T_\varepsilon^{-C\varepsilon} \cdot t^{C\varepsilon}.$$

In particular, if we choose $0 < \varepsilon < 1/(2C)$, then any two points in $(\mathcal{M}^n, g(t))$ can be joined by a path of length

$$L(t) \leq [\mathrm{diam}(\mathcal{M}^n, g(T_\varepsilon))] T_\varepsilon^{-C\varepsilon} \cdot t^{1/2-\delta}$$

where $\delta \doteqdot 1/2 - C\varepsilon > 0$. This implies (8.26) and proves the claim. □

Because 3-dimensional nilmanifolds are known to be geometrizable, we now describe how to obtain Type III limits under the assumption $\alpha > 0$. By definition of α, there exist sequences (x_i, t_i) with $t_i \nearrow \infty$ such that

$$t_i |\mathrm{Rm}(x_i, t_i)| \doteqdot \alpha_i \to \alpha.$$

Choose any such sequence. Also by definition of α, there is for any $\varepsilon > 0$ a time $T_\varepsilon \in [0, \infty)$ such that

$$t |\mathrm{Rm}(x, t)| \leq \alpha + \varepsilon$$

for all $x \in \mathcal{M}^n$ and $t \in [T_\varepsilon, \infty)$. The dilated solutions $g_i(t)$ exist on the time intervals $[-\alpha_i, \infty)$ and satisfy

$$\begin{aligned}
|\mathrm{Rm}_i(x, t)| &= \frac{1}{|\mathrm{Rm}(x_i, t_i)|}\left|\mathrm{Rm}\left(x, t_i + \frac{t}{|\mathrm{Rm}(x_i, t_i)|}\right)\right| \\
&= \frac{\left|\mathrm{Rm}\left(x, t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)\right| \left(t_i + \frac{t}{|\mathrm{Rm}(x_i,t_i)|}\right)}{|\mathrm{Rm}(x_i, t_i)| t_i + t} \\
&\leq \frac{\alpha + \varepsilon}{\alpha_i + t},
\end{aligned}$$

if

$$t_i + t |\mathrm{Rm}(x_i, t_i)|^{-1} \geq T_\varepsilon,$$

that is if

$$t \geq |\mathrm{Rm}(x_i, t_i)| (T_\varepsilon - t_i).$$

Note that for any fixed $\varepsilon > 0$, we have

$$|\mathrm{Rm}(x_i, t_i)| (T_\varepsilon - t_i) \to -\alpha$$

and

$$\frac{\alpha + \varepsilon}{\alpha_i + t} \to \frac{\alpha + \varepsilon}{\alpha + t}$$

for $t > -\alpha$. Hence the pointed limit solution $(\mathcal{M}_\infty^n, g_\infty(t), x_\infty)$, if it exists, is defined for $t \in (-\alpha, \infty)$ and satisfies

$$\sup_{\mathcal{M}_\infty^n \times (-\alpha,\infty)} |\mathrm{Rm}_\infty| \leq \frac{\alpha}{\alpha + t} = |\mathrm{Rm}_\infty(x_\infty, 0)|.$$

6. Taking limits backwards in time

Assume that we have taken the limit of a sequence of dilations and obtained a Type II ancient solution $(\mathcal{M}_\infty^n, g_\infty(t))$ for $t \in (-\infty, \omega)$, where

$$\sup_{\mathcal{M}_\infty^n \times (-\infty,0]} |t| \cdot |\mathrm{Rm}_\infty(x,t)| = \infty. \tag{8.27}$$

Type II ancient limit solutions typically arise from taking a second limit, as in dimension reduction. However, such limit solutions might not attain their maximum curvatures, hence might not satisfy condition (8.1). When $n = 3$, these limit solutions have nonnegative sectional curvature, which allows for the application of the Harnack estimate. But to combine the Harnack estimate with the strong maximum principle, it is necessary that the limit solution attain its maximum scalar curvature. This motivates us to take a another (third) limit to obtain a solution where this maximum is attained. We can do this by translating and dilating about a suitable sequence of times $t_i \searrow -\infty$ and points $x_i \in \mathcal{M}_\infty^n$ where the curvatures almost attain their maximums.

So let $(\mathcal{M}_\infty^n, g_\infty(t))$ be a complete Type II ancient solution defined for $t \in (-\infty, \omega)$. Let $\varepsilon_i < 1$ be any sequence of constants such that $\varepsilon_i \searrow 0$, and choose any sequence of times $T_i \to -\infty$ and corresponding points and times $(x_i, t_i) \in \mathcal{M}_\infty^n \times (T_i, 0)$ such that

$$|t_i|(t_i - T_i)|\mathrm{Rm}_\infty(x_i,t_i)| \geq (1-\varepsilon_i) \sup_{\mathcal{M}_\infty^n \times [T_i,0]} |t|(t - T_i)|\mathrm{Rm}_\infty(x,t)|. \tag{8.28}$$

This is very similar to how we chose the points and times in the case of a Type IIb singularity. Set

$$\begin{aligned} \alpha_i &\doteqdot (t_i - T_i)|\mathrm{Rm}_\infty(x_i,t_i)| \geq 0, \\ \omega_i &\doteqdot -t_i|\mathrm{Rm}_\infty(x_i,t_i)| \geq 0, \end{aligned}$$

and observe that $\lim_{i\to\infty} \alpha_i = \lim_{i\to\infty} \omega_i = \infty$. Indeed, by assumption (8.27) and estimate (8.28), we have

$$\begin{aligned} \frac{1}{1/\alpha_i + 1/\omega_i} = \frac{\alpha_i \omega_i}{\alpha_i + \omega_i} &= \frac{|t_i|(t_i - T_i)|\mathrm{Rm}_\infty(x_i,t_i)|}{|T_i|} \\ &\geq (1-\varepsilon_i) \sup_{\mathcal{M}_\infty^n \times [T_i,0]} \frac{|t|(t - T_i)|\mathrm{Rm}_\infty(x,t)|}{|T_i|} \to \infty \end{aligned}$$

as $i \to \infty$, because $T_i \to -\infty$.

Now consider the dilated solutions

$$(g_\infty)_i(t) = |\mathrm{Rm}_\infty(x_i,t_i)| \cdot g_\infty\left(t_i + \frac{t}{|\mathrm{Rm}_\infty(x_i,t_i)|}\right)$$

defined as before. Each solution $(g_\infty)_i$ exists on the dilated time interval

$$(-\infty, \ (\omega - t_i) \left|\mathrm{Rm}_\infty (x_i, t_i)\right|)$$

which contains the subinterval $[-\alpha_i, \omega_i]$ on which we have a good estimate of the curvature. Indeed, (8.28) implies that for all $x \in \mathcal{M}^n_\infty$ and $t \in [-\alpha_i, \omega_i]$, one has

$$\begin{aligned}
\left|(\mathrm{Rm}_\infty)_i (x,t)\right| &= \frac{1}{\left|\mathrm{Rm}_\infty (x_i, t_i)\right|} \left|\mathrm{Rm}_\infty \left(x, t_i + \frac{t}{\left|\mathrm{Rm}_\infty (x_i,t_i)\right|}\right)\right| \\
&\leq \frac{(t_i - T_i)\left|t_i\right|}{(1-\varepsilon_i)\left(t_i + \frac{t}{\left|\mathrm{Rm}_\infty (x_i,t_i)\right|} - T_i\right)\left|t_i + \frac{t}{\left|\mathrm{Rm}_\infty (x_i,t_i)\right|}\right|} \\
&= \frac{\alpha_i \omega_i}{(1-\varepsilon_i)(\alpha_i + t)(\omega_i - t)}.
\end{aligned}$$

Because $\alpha_i \to \infty$, $\omega_i \to \infty$, and $\varepsilon_i \searrow 0$, we conclude that the pointed limit solution $(\mathcal{M}^n_{2\infty}, g_{2\infty}(t), x_{2\infty})$ is an eternal solution that satisfies

$$\sup_{\mathcal{M}^n_{2\infty} \times (-\infty,\infty)} \left|\mathrm{Rm}_{2\infty}\right| \leq 1 = \left|\mathrm{Rm}_{2\infty}(x_{2\infty}, 0)\right|.$$

Notes and commentary

The main reference for this chapter is Section 16 of [**63**]. Limit solutions are obtained by the application of the Compactness Theorem we reviewed in Section 3 of Chapter 7. Limits formed from dilations about sequences of points and times that are not curvature essential will be considered in the next volume.

CHAPTER 9

Type I singularities

Suppose that $(\mathcal{M}^3, g(t))$ is a solution of the Ricci flow which becomes singular at some time $T < \infty$. In order to extract geometric and topological information about $\mathcal{M}^3$ from the solution $g(t)$, one wants to perform geometric-topological surgeries on $\mathcal{M}^3$ as the singularity develops in such a way that the maximum curvature of the solution is reduced sufficiently. In order to perform such surgeries, one needs a good understanding of the possible singularities that can arise and of the possible limits of dilations about them. Ultimately, one then wants to argue that only geometrically recognizable pieces will remain after finitely many surgeries.

This is the program for proving the Geometrization Conjecture that was designed by Hamilton and advanced by Perelman [**105, 106, 107**]. We will further discuss the details of this program in a successor to this volume. In this chapter, we will begin to explore some of the reasons why singularities are expected to be topologically and geometrically tractable in dimension 3, and why near a 'typical' singularity, one expects to see a neck: a piece of the manifold which is geometrically close to a quotient of the shrinking round product cylinder. This heuristic notion will be made precise in Section 4 below. (Recall that we made a rigorous study of neckpinch singularities under certain symmetry assumptions in Section 5 of Chapter 2.)

1. Intuition

Singularities of the Ricci flow $(\mathcal{M}^n, g(t))$ in high dimensions are expected to be very complex. In dimension $n = 3$ however, there are three observations that lead one to expect singularities to be relatively tractable.

The first observation is the pinching estimate of Theorem 9.4, below. This estimate says that at any point and time where a sectional curvature is negative and large in absolute value, one finds a much larger positive sectional curvature. It implies in particular that any singularity model must have nonnegative sectional curvature at $t = 0$.

The second observation which restricts the possible singularity models one may see in dimension $n = 3$ is the fact that $\mathfrak{so}(2)$ is the only proper nontrivial Lie subalgebra of $\mathfrak{so}(3)$. This fact says that at the origin of any singularity model, the eigenvalues of the rescaled curvature operator must conform to either the signature $(+,+,+)$ or else $(0,0,+)$. The signature $(0,+,+)$ is ruled out by the Lie algebra structure of $\mathfrak{so}(3)$. The other possible pattern, $(0, 0, 0,)$, may be ruled out by choosing a dilation sequence so that (8.9) holds, thereby ensuring that the model one obtains is not flat.

The third observation is the fact that a strong maximum principle holds for tensors. (See Chapter 4.) Because the curvature operator of any singularity model $\left(\mathcal{M}_{\infty}^{3}, g_{\infty}(t), x_{\infty}\right)$ is nonnegative at $t = 0$ and has either the signature $(+, +, +)$ or else $(0, 0, +)$ at the single point $(x_{\infty}, 0)$, the strong maximum principle says that the curvature operator must possess either the signature $(+, +, +)$ or else $(0, 0, +)$ respectively at *all* points $x \in \mathcal{M}_{\infty}^{3}$ and times $t > 0$ such that the limit solution $g_{\infty}(t)$ exists.

The standard example of a singularity of signature $(+, +, +)$ in dimension 3 is the self-similar solution formed by the shrinking round 3-sphere

$$4(T - t)\, g_{\text{can}}, \tag{9.1}$$

which we encountered back in Section 5. Singularities of signature $(0, 0, +)$ are expected to resemble the self-similar 'cylinder solution'

$$ds^2 + 2(T - t)\, g_{\text{can}} \tag{9.2}$$

on $\mathbb{R} \times \mathcal{S}^2$. We saw rigorous examples of such singularities and their asymptotic behaviors in Section 5.

When viewed at an appropriate length scale, the geometry of a solution to the Ricci flow $\left(\mathcal{M}^3, g(t)\right)$ that becomes singular at time $T < \infty$ closely resembles the singularity model $\left(\mathcal{M}_{\infty}^{3}, g_{\infty}(t), x_{\infty}\right)$ one obtains by blowing-up the singularity, at least for points near the singularity and times just before T. The intuitions outlined above, therefore, support the following heuristic picture. This picture is intended as a guide to our intuition in the remainder of this chapter.

(1) If a solution $\left(\mathcal{M}^3, g(t)\right)$ of the unnormalized Ricci flow encounters a Type I singularity, then after performing dilations correctly and obtaining an injectivity radius estimate, one expects to see a limit which is a quotient of either
   (a) a compact shrinking round sphere $\left(\mathcal{S}^3, g(t)\right)$, where $g(t)$ is given by (9.1), or
   (b) a noncompact shrinking cylinder $\left(\mathbb{R} \times \mathcal{S}^2, g(t)\right)$, where $g(t)$ is given by (9.2).
(2) If a solution $\left(\mathcal{M}^3, g(t)\right)$ of the unnormalized Ricci flow encounters a Type IIa singularity, then after performing dilations correctly and obtaining an injectivity radius estimate, one expects to see a quotient of one of the following noncompact limits:
   (a) a steady self-similar solution $\left(\mathbb{R}^3, g(t)\right)$ of positive sectional curvature,
   (b) a shrinking cylinder $\left(\mathbb{R} \times \mathcal{S}^2, g(t)\right)$ as in Case 1b above, or

(c) a cigar product $(\mathbb{R}^3, g(t))$, where $g(t)$ is the self-similar solution corresponding to the soliton metric introduced in Subsection 2.1.

For any limit in Case 2a, one could perform a technique called *dimension reduction* in the hope of obtaining an ancient solution. (Dimension reduction is introduced in Section 4 below and will be discussed further in a planned successor to this volume. It involves taking a limit around a suitable sequence of points tending to spatial infinity.) If the ancient solution one obtains is not in fact an eternal solution which attains its maximum curvature, one can then take a third limit about a suitable sequence of points and times tending to $-\infty$ where the curvature is sufficiently near its maximum. Having done so, one expects to see either Case 2b or 2c above.

If one were in Case 1a, then one could conclude from compactness of the limit that the underlying manifold of the original solution must have been $\mathcal{S}^3$ or one of its quotients. In the other cases, the singularity model should give local information about the original solution near the singularity just prior to its formation.

REMARK 9.1. The recent work [**105**] of Perelman rules out Case 2c for singularities that form in finite time.

2. Positive curvature is preserved

In this chapter, we shall make heavy use of the maximum principles for systems discussed in Section 3 of Chapter 4. In so doing, we will follow the methods introduced in Section 4 of Chapter 6. In order to review these techniques, we will first apply them to derive two results which were obtained earlier by classical methods in Sections 1 and 3 of Chapter 6.

LEMMA 9.2 (Positive sectional curvature is preserved.). *If $(\mathcal{M}^3, g(t))$ is a solution of the Ricci flow such that the initial metric $g(0)$ has positive (nonnegative) sectional curvature, then the metrics $g(t)$ have positive (nonnegative) sectional curvature for all $t > 0$ that the solution exists.*

PROOF. We show that positive sectional curvature is preserved at each $x \in \mathcal{M}^3$; the proof that nonnegative sectional curvature is preserved is entirely analogous. Let $\lambda(\mathbb{P}) \geq \mu(\mathbb{P}) \geq \nu(\mathbb{P})$ denote the eigenvalues of any

$$\mathbb{P} \in \left(\wedge^2 T^*\mathcal{M}^3 \otimes_S \wedge^2 T^*\mathcal{M}^3\right)_x.$$

Let $\mathbb{M}(t) \in \left(\wedge^2 T^*\mathcal{M}^3 \otimes_S \wedge^2 T^*\mathcal{M}^3\right)_x$ be the quadratic form corresponding to $\mathrm{Rm}[g]$. Define the subset $\mathcal{K} \subset \left(\wedge^2 T^*\mathcal{M}^3 \otimes_S \wedge^2 T^*\mathcal{M}^3\right)_x$ by

$$\mathcal{K} \doteqdot \{\mathbb{P} : \nu(\mathbb{P}) > 0\},$$

and notice that $\mathbb{M}(t) \in \mathcal{K}$ if and only if all sectional curvatures are positive at (x,t). It is easy to see that $\mathcal{K}$ is invariant under parallel translation. $\mathcal{K}$ is convex in each fiber because if

$$\rho : \left(\wedge^2 T^*\mathcal{M}^3 \otimes_S \wedge^2 T^*\mathcal{M}^3\right)_x \to \mathbb{R}$$

by

$$\rho : \mathbb{P} \mapsto \min_{|V|=1} \mathbb{P}(V,V),$$

then $\nu(\mathbb{P}) = \rho(\mathbb{P})$ and

$$\rho(s\mathbb{P} + (1-s)\mathbb{Q}) \geq s\rho(\mathbb{P}) + (1-s)\rho(\mathbb{Q}).$$

The fact that $\mathbb{M}$ remains in $\mathcal{K}$ under the ODE (6.31) follows from (6.32), because

$$\frac{d}{dt}\nu = \nu^2 + \lambda\mu > 0$$

whenever $\nu > 0$. The lemma now follows from the maximum principle for systems, which says that if $\mathrm{Rm}(x,0) \geq 0$, then $\mathrm{Rm}(x,t) \geq 0$ for all $t \geq 0$ such that the solution exists. □

The fact that positive sectional curvature is preserved in dimension 3 is a special case of the more general result that positivity of the curvature operator is preserved in all dimensions. A result which is special to 3 dimensions is the following:

LEMMA 9.3 (Positive Ricci curvature is preserved.). *If $(\mathcal{M}^3, g(t))$ is a solution of the Ricci flow such that the initial metric $g(0)$ has positive (nonnegative) Ricci curvature, then the metrics $g(t)$ have positive (nonnegative) Ricci curvature for all $t > 0$ that the solution exists.*

PROOF. The proof is analogous to the previous lemma. Here we define $\mathcal{K}$ by

$$\mathcal{K} = \{\mathbb{P} : \nu(\mathbb{P}) + \mu(\mathbb{P}) > 0\},$$

so that the quadratic form $\mathbb{M}$ corresponding to $\mathrm{Rm}[g]$ lies in $\mathcal{K}$ if and only if $\mathrm{Rc}[g] > 0$. $\mathcal{K}$ is fiberwise convex, because the function

$$\nu(\mathbb{P}) + \mu(\mathbb{P}) = \min_{\substack{|V|=|W|=1 \\ \langle V,W\rangle=0}} [\mathbb{P}(V,V) + \mathbb{P}(W,W)]$$

is concave in each fiber. If $\nu + \mu > 0$, we have $\lambda \geq \mu \geq (\nu+\mu)/2 > 0$ and hence

$$\frac{d}{dt}(\nu+\mu) = \nu^2 + \lambda\mu + \mu^2 + \lambda\nu > \lambda(\nu+\mu) > 0.$$

The lemma follows again from the maximum principle for systems. □

3. Positive sectional curvature dominates

The Ricci flow prefers positive curvature, in the sense that it preserves positive curvature operator in any dimension. Moreover, we saw in Section 2 that both positive sectional curvature and positive Ricci curvature are preserved in dimension 3. Thus it is natural to investigate the extent to which the sectional curvatures of an arbitrary initial metric tend to become positive under the flow. By the maximum principle, it suffices to study the corresponding system (6.32) of ODE. We are interested in the case that the smallest eigenvalue ν is negative. To simplify and motivate the qualitative study of (6.32), we first consider the subcase that $\mu = \nu < 0$. We then have the following system of two ODE:

$$\frac{d\lambda}{dt} = \lambda^2 + \nu^2$$
$$\frac{d\nu}{dt} = \nu^2 + \lambda\nu.$$

For convenience, we set $\pi = -\nu > 0$ and consider the system

$$\frac{d\lambda}{dt} = \lambda^2 + \pi^2$$
$$\frac{d\pi}{dt} = \lambda\pi - \pi^2$$

under the assumption that $\lambda > \nu = -\pi$ initially. Since

$$\frac{d}{dt}(\lambda + \pi) = \lambda(\lambda + \pi),$$

this condition is preserved. Similarly, $\pi > 0$ as long as the solution exists.

If $\lambda \leq 0$, then $d\pi/dt \leq -\pi^2 < 0$. If $\pi \geq \lambda$, then $d\pi/dt = \pi(\lambda - \pi) \leq 0$ but $d\lambda/dt > 0$. Hence π is non-increasing except possibly when $\pi < \lambda$, and it follows that $\pi \leq \max\{\lambda, C\}$. To improve on this estimate, note that $d\lambda/dt > 0$, so that we may write

$$\frac{d\pi}{d\lambda} = \frac{\lambda\pi - \pi^2}{\lambda^2 + \pi^2}.$$

This is a homogeneous equation. The standard substitution $\pi = \lambda\zeta$ gives

$$\zeta + \lambda\frac{d\zeta}{d\lambda} = \frac{\zeta - \zeta^2}{1 + \zeta^2}.$$

Using partial fractions and integrating, we obtain

$$\int \frac{d\lambda}{\lambda} + \int \left(\frac{1}{\zeta^2} - \frac{1}{\zeta} + \frac{2}{1+\zeta}\right) d\zeta = 0,$$

which yields

$$\log|\lambda| - \frac{1}{\zeta} - \log|\zeta| + 2\log|1+\zeta| = C.$$

Since $\zeta = \pi/\lambda$ and $\lambda + \pi > 0$, we may write this in terms of λ and π as

$$\log\pi = \frac{\lambda}{\pi} + 2\log\left(\frac{\pi}{\pi + \lambda}\right) + C.$$

If $\pi(t)$ is sufficiently large and positive, this leads to a contradiction unless

$$\lambda(t) > \pi(t)\log\pi(t).$$

In particular, λ is positive and much larger than π. On the other hand, if π is sufficiently small and positive, we get a contradiction unless $-\pi < \lambda < 0$.

This discussion motivates the following theorem, which reveals the precise sense in which all sectional curvatures of a complete 3-manifold evolving by the Ricci flow are dominated by the positive sectional curvatures.

Namely, large negative sectional curvatures can occur only in the presence of much larger positive sectional curvatures:

THEOREM 9.4. *Let $(\mathcal{M}^3, g(t))$ be any solution of the Ricci flow on a closed 3-manifold for $0 \leq t < T$. Let $\nu(x,t)$ denote the smallest eigenvalue of the curvature operator. If $\inf_{x \in \mathcal{M}^3} \nu(x,0) \geq -1$, then at any point $(x,t) \in \mathcal{M}^3 \times [0,T)$ where $\nu(x,t) < 0$, the scalar curvature is estimated by*

$$R \geq |\nu| \left(\log |\nu| + \log(1+t) - 3\right).$$

Note that one can always achieve $\inf_{x \in \mathcal{M}^3} \nu(x,0) \geq -1$ simply by scaling $g(0)$ by a sufficiently large constant. Note also that if $\nu \leq -e^6$, we have

$$R \geq \frac{1}{2} |\nu| \log |\nu|.$$

In particular, $R \gg |\nu|$ when $\nu \ll -1$. Thus whenever we encounter a large negative sectional curvature at some point and time (x,t), we find a much larger positive sectional curvature at the same point and time. As we shall see, this estimate implies that limits (if they exist) of sequences of dilations about a finite-time singularity have nonnegative sectional curvature.

Before proving the theorem, we establish two technical lemmas.

LEMMA 9.5. *At each fixed time t, the subsets*

$$\mathcal{K} \subset \left(\wedge^2 T^* \mathcal{M}^3 \otimes_S \wedge^2 T^* \mathcal{M}^3\right)_x$$

defined by

$$\mathcal{K} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{c} \operatorname{tr} \mathbb{P} \geq -3(1+t), \\ \\ \text{and if } \nu(\mathbb{P}) \leq -1/(1+t), \text{ then} \\ \\ \operatorname{tr} \mathbb{P} \geq |\nu(\mathbb{P})| \left(\log |\nu(\mathbb{P})| + \log(1+t) - 3\right) \end{array} \right\}$$

are invariant under parallel translation and are convex in each fiber.

PROOF. Invariance under parallel translation is clear. To prove convexity at each fixed $x \in \mathcal{M}^3$ and $t \in [0,T)$, first consider the region $K \subset \mathbb{R}^2$ given by

$$K = \left\{ (u,v) \;\middle|\; \begin{array}{c} v \geq -3(1+t), \\ \\ v \geq -3u, \\ \\ \text{and if } u \geq 1/(1+t), \text{ then} \\ \\ v \geq u\left(\log u + \log(1+t) - 3\right) \end{array} \right\}.$$

See figure (1) and notice that all three curves shown intersect when $u = 1/(1+t)$. The region K is bounded on the left by the line $v = -3u$, below by the line $v = -3/(1+t)$, and on the right by the curve $v = u(\log u + \log(1+t) - 3)$. It is clear that K is a convex subset of $\mathbb{R}^2$.

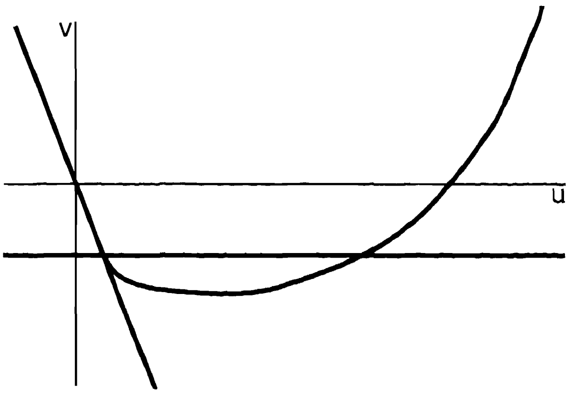

FIGURE 1. Boundary of the region K in the (u,v)-plane.

Now consider the map $\Phi : \left(\wedge^2 T^*\mathcal{M}^3 \otimes_S \wedge^2 T^*\mathcal{M}^3\right) \to \mathbb{R}^2$ defined by

$$\Phi : \mathbb{P} \mapsto (u(\mathbb{P}), v(\mathbb{P})) \doteqdot (|\nu|, \operatorname{tr}\mathbb{P})$$

and observe that $\mathbb{P} \in \mathcal{K}$ if and only if $\Phi(\mathbb{P}) \in K$. (The condition $v \geq -3u$ is satisfied automatically, because $\operatorname{tr}\mathbb{P} \geq 3\nu(\mathbb{P})$.) Thus the proof will be complete if we show that $\Phi(s\mathbb{P} + (1-s)\mathbb{Q}) \in K$ for all $\mathbb{P}, \mathbb{Q} \in \mathcal{K}$ and $s \in [0,1]$. Since trace is a linear functional,

$$v(s\mathbb{P} + (1-s)\mathbb{Q}) = sv(\mathbb{P}) + (1-s)v(\mathbb{Q}).$$

Because $v \geq -3u$ and the function $-u(\mathbb{P}) = \min_{|W|=1} \mathbb{P}(W,W)$ is concave,

$$-\frac{v(s\mathbb{P} + (1-s)\mathbb{Q})}{3} \leq u(s\mathbb{P} + (1-s)\mathbb{Q}) \leq su(\mathbb{P}) + (1-s)u(\mathbb{Q}). \tag{9.3}$$

Together, these relations show that $\Phi(s\mathbb{P} + (1-s)\mathbb{Q}) \in H$, where H is the rhombus with vertices $(u(\mathbb{P}), v(\mathbb{P}))$, $(u(\mathbb{Q}), v(\mathbb{Q}))$, $\left(-\frac{1}{3}v(\mathbb{P}), v(\mathbb{P})\right)$, and $\left(-\frac{1}{3}v(\mathbb{Q}), v(\mathbb{Q})\right)$. Because each vertex of H lies in the convex set K, we have $H \subset K$. □

LEMMA 9.6. *Given any $x \in \mathcal{M}^3$ and quadratic form*

$$\mathbb{P} \in \left(\wedge^2 T^*\mathcal{M}^3 \otimes_S \wedge^2 T^*\mathcal{M}^3\right)_x$$

with eigenvalues $\lambda(\mathbb{P}) \geq \mu(\mathbb{P}) \geq \nu(\mathbb{P})$, where $\nu(\mathbb{P}) < 0$, define

$$\Omega(\mathbb{P}) \doteqdot \frac{\operatorname{tr}\mathbb{P}}{-\nu(\mathbb{P})} - \log(-\nu(\mathbb{P})).$$

Let $\mathbb{M}(t)$ correspond to $\operatorname{Rm}[g]$. *Then wherever $\Omega(\mathbb{M})$ is defined,*

$$\frac{d}{dt}\Omega(\mathbb{M}) \geq -\nu(\mathbb{M}).$$

PROOF. By (6.32), we calculate

$$\begin{aligned}\nu^2\frac{d}{dt}\Omega(\mathbb{M}) &= -\nu\frac{d}{dt}(\operatorname{tr}\mathbb{M}) + (\operatorname{tr}\mathbb{M})\frac{d}{dt}\nu - \nu\frac{d\nu}{dt}\\ &= -\nu\frac{d}{dt}(\lambda+\mu+\nu) + (\lambda+\mu)\frac{d}{dt}\nu\\ &= -\nu^3 - \nu\left(\lambda^2+\mu^2+\lambda\mu\right) + (\lambda+\mu)\lambda\mu.\end{aligned}$$

Thus it suffices to prove that

$$\pi \doteqdot -\nu\left(\lambda^2+\mu^2+\lambda\mu\right) + (\lambda+\mu)\lambda\mu \geq 0$$

whenever $\nu < 0$. There are two cases: if $\mu < 0$, then we write

$$\pi = (\mu-\nu)\left(\lambda^2+\mu^2+\lambda\mu\right) - \mu^3 \geq -\mu^3 > 0,$$

while if $\mu \geq 0$, we have

$$\pi = \lambda^2(\mu-\nu) - \nu\mu^2 + \lambda\mu(\mu-\nu) \geq 0.$$

□

PROOF OF THEOREM 9.4. Fix any $x \in \mathcal{M}^3$ with $\nu(x,0) < 0$. Let $\mathbb{M}$ be the quadratic form corresponding to $\operatorname{Rm}[g]$, and let

$$\mathcal{K} \subset \left(\wedge^2 T^*\mathcal{M}^3 \otimes_S \wedge^2 T^*\mathcal{M}^3\right)_x$$

be the time-dependent set defined in Lemma 9.5. Then since

$$\mathcal{K}(0) = \left\{\mathbb{P} \;\middle|\; \begin{array}{ll}\operatorname{tr}\mathbb{P} \geq -3 & \text{and}\\ \operatorname{tr}\mathbb{P} \geq |\nu(\mathbb{P})|\left(\log|\nu(\mathbb{P})| - 3\right) & \text{if } \nu(\mathbb{P}) \leq -1\end{array}\right\},$$

we have $\mathbb{M}(0) \in \mathcal{K}$ by the hypothesis $\inf \nu(\cdot,0) \geq -1$. We claim $\mathbb{M}(t)$ remains in $\mathcal{K}$ as long as $\nu < 0$. The theorem follows directly from the claim, which implies that

$$\frac{R}{-\nu} - \log(-\nu) \geq \frac{\frac{-3}{1+t}}{\frac{1}{1+t}} - \log\left(\frac{1}{1+t}\right) = \log(1+t) - 3$$

when $0 < -\nu < 1/(1+t)$, and

$$\frac{R}{-\nu} \geq \log(-\nu) + \log(1+t) - 3$$

when $-\nu \geq 1/(1+t)$.

The proof that $\mathbb{M}(t) \in \mathcal{K}$ uses the maximum principle for systems. The inequality $\operatorname{tr}\mathbb{M} \geq -3/(1+t)$ is preserved, because

$$\frac{d}{dt}(\operatorname{tr}\mathbb{M}) - \frac{1}{3}(\operatorname{tr}\mathbb{M})^2 \geq \frac{1}{3}\left(\lambda^2+\mu^2+\nu^2\right) \geq 0.$$

And if $|\nu| \geq 1/(1+t)$, then Lemma 9.6 implies that $\frac{d}{dt}\Omega(\mathbb{M}) \geq 1/(1+t)$, which is equivalent to the inequality

$$\frac{d}{dt}\left[\frac{\operatorname{tr}\mathbb{M}}{|\nu|} - \log|\nu| - \log(1+t)\right] \geq 0.$$

□

COROLLARY 9.7. *In dimension $n = 3$, every Type I limit of a Type I singularity or Type II limit of a Type IIa singularity has nonnegative sectional curvature for as long as it exists.*

PROOF. Let $(\mathcal{M}^3, g(t))$ be a singular solution with blow-up time $T < \infty$. As in Section 3.1 of Chapter 8, let (x_i, t_i) be a sequence of points and times chosen such that $t_i \nearrow T$ and

$$|\mathrm{Rm}(x_i, t_i)| \geq c_2 \sup_{x \in \mathcal{M}^3} |\mathrm{Rm}(x, t_i)| \geq c_2 c_1 \sup_{(x,t) \in \mathcal{M}^3 \times [s_i, t_i]} |\mathrm{Rm}(x, t)|,$$

where $(t_i - s_i) \sup_{x \in \mathcal{M}^3} |\mathrm{Rm}(x, t_i)| \to \infty$. Suppose that a subsequence of the pointed solutions $(\mathcal{M}^3, g_i(t), x_i)$ defined by

$$g_i(t) \doteqdot |\mathrm{Rm}(x_i, t_i)| \cdot g\left(t_i + \frac{t}{|\mathrm{Rm}(x_i, t_i)|}\right)$$

converges to a complete pointed solution $(\mathcal{M}^3_\infty, g_\infty(t), x_\infty)$ of the Ricci flow. Then by Proposition 8.20 and (8.18), $g_\infty(t)$ is an ancient solution satisfying

$$\sup_{x \in \mathcal{M}^3_\infty} \left| \underset{\infty}{\mathrm{Rm}}(x, t) \right| < \infty$$

for all $t \leq 0$. By Corollary 9.8, the sectional curvatures of g_∞ are nonnegative. $\square$

Another consequence of the pinching estimate in Theorem 9.4 is that ancient 3-dimensional solutions of the Ricci flow have nonnegative sectional curvature.

COROLLARY 9.8. *Let $(\mathcal{M}^3, g(t))$ be a complete ancient solution of the Ricci flow. Assume that there exists a continuous function $K(t)$ such that $|\mathrm{sect}[g(t)]| \leq K(t)$. Then $g(t)$ has nonnegative sectional curvature for as long as it exists.*

PROOF. If $(\mathcal{M}^3, g(t))$ is a solution of the Ricci flow that exists at least for $0 \leq t < T$ with

$$0 > \nu_0 \doteqdot \inf_{x \in \mathcal{M}^3} \nu(x, 0),$$

then the family of metrics $\tilde{g}(t)$ defined on $\mathcal{M}^3$ by

$$\tilde{g}(t) \doteqdot |\nu_0| \cdot g\left(\frac{t}{|\nu_0|}\right)$$

is a solution of the Ricci flow with $\tilde{\nu} \geq -1$ at $t = 0$. So at any $(x, |\nu_0|\, t) \in \mathcal{M}^3 \times [0, T)$, Theorem 9.4 gives the estimate

$$\tilde{R}(x, |\nu_0|\, t) \geq |\tilde{\nu}(x, |\nu_0|\, t)| \cdot [\log |\tilde{\nu}(x, |\nu_0|\, t)| + \log(1 + |\nu_0|\, t) - 3]$$

wherever $\tilde{\nu}(x, |\nu_0|\, t) < 0$. (Maximum principles work on noncompact as well as on compact manifolds. We will provide a complete proof of this fact in the sequel to this volume.) In terms of the solution $g(t)$, this shows that

$$R(x,t) \geq |\nu(x,t)| \cdot \left[\log\left(\frac{|\nu(x,t)|}{|\nu_0|}\right) + \log(1+|\nu_0|\, t) - 3\right]$$
$$= |\nu(x,t)| \cdot \left[\log|\nu(x,t)| + \log\left(|\nu_0|^{-1} + t\right) - 3\right]$$

wherever $\nu(x,t) < 0$. In particular, if $\nu(x,t) < 0$ at some point $x \in \mathcal{M}^3$ and time $t > 0$, then

$$R(x,t) > |\nu(x,t)| (\log|\nu(x,t)| + \log t - 3).$$

Now if $g(t)$ exists for $-\alpha \leq t < \omega$, we can translate time to conclude that

$$R(x,t) > |\nu(x,t)| (\log|\nu(x,t)| + \log(t+\alpha) - 3)$$

wherever $\nu(x,t) < 0$. But if $g(t)$ is ancient, this leads to a contradiction, since $\lim_{\alpha\to\infty} \log(t+\alpha) = \infty$. Hence $\nu(x,t) \geq 0$ for all x and t. □

4. Necklike points in finite-time singularities

Suppose $(\mathcal{M}^3, g)$ is a closed solution of the Ricci flow which forms a singularity at $T < \infty$, in other words, a singularity of Type I or Type IIa. The main result of this section asserts that either $\mathcal{M}^3$ is a spherical space form or else $\mathcal{M}^3$ admits a 'dimension reduction' somewhere near the singularity, in the sense that at any time sufficiently close to T, we can find a point in $\mathcal{M}^3$ near which the geometry is arbitrarily close to the product of a surface and a line. When the singularity is Type I, the surface arising from a limit of dilations about a subsequence of such points and times approaching T is a constant-curvature 2-sphere. In this case, the limit is a quotient of $\mathcal{S}^2 \times \mathbb{R}$, and we say that $\mathcal{M}^3$ develops a 'neck.' (See Section 5 Section of Chapter 2.)

To make the notion of **dimension reduction** precise, we shall say that (x,t) is a **Type I** c**-essential point** if

$$|\mathrm{Rm}(x,t)| \geq \frac{c}{T-t} > 0.$$

We say that (x,t) is a δ**-necklike point** if there exists a unit 2-form θ at (x,t) such that

$$|\mathrm{Rm} - R(\theta \otimes \theta)| \leq \delta |\mathrm{Rm}|.$$

THEOREM 9.9. *Let $(\mathcal{M}^3, g(t))$ be a closed solution of the unnormalized Ricci flow on a maximal time interval $0 \leq t < T < \infty$. If the normalized flow does not converge to a metric of constant positive sectional curvature, then there exists a constant $c > 0$ such that for all $\tau \in [0,T)$ and $\delta > 0$, there are $x \in \mathcal{M}^3$ and $t \in [\tau, T)$ such that (x,t) is a Type I c-essential point and a δ-necklike point.*

Observe that the theorem is equivalent to the following statement, which we will prove:

If for every $c > 0$*, there exist some* $\tau < T$ *and* $\delta > 0$ *such that there are no Type I* c*-essential* δ*-necklike points after time* τ*, then the normalized solution* $\tilde{g}(t)$ *converges to a spherical space form.*

Our strategy will be to modify the second proof of Theorem 6.30 in order to show that the pinching of the curvature operator improves sufficiently under the flow when this hypothesis holds. Recall that we considered the function

$$f \doteqdot \frac{|\overset{\circ}{\mathrm{Rm}}|^2}{R^{2-\varepsilon}}, \tag{9.4}$$

and saw that f stays bounded for ε sufficiently small when $(\mathcal{M}^3, g)$ has strictly positive Ricci curvature. The first difficulty we face is that we no longer know $R > 0$. But if we choose constants $\rho \geq 0$, $c > 0$ so that $R + \rho \geq c > 0$ at $t = 0$, we have $R + \rho \geq c$ for $0 \leq t < T$ by the maximum principle, because

$$\frac{\partial}{\partial t}(R + \rho) = \Delta(R + \rho) + 2|\mathrm{Rc}|^2 .$$

Thus we may replace R by $R+\rho$ in the denominator of (9.4). This introduces an additional bad term in the evolution equation for (9.4). To remedy this, we introduce the time factor $(T-t)^\varepsilon$ and define

$$F \doteqdot (T-t)^\varepsilon \frac{|\overset{\circ}{\mathrm{Rm}}|^2}{(R+\rho)^{2-\varepsilon}} = [(T-t)(R+\rho)]^\varepsilon \frac{|\overset{\circ}{\mathrm{Rm}}|^2}{(R+\rho)^2}. \tag{9.5}$$

Since we must allow for arbitrary initial metrics, we shall have to work harder to show that F stays bounded. However, the multiplicative factor $(T-t)^\varepsilon$ helps, because its time derivative is negative. In fact, a key step in proving the theorem will be to show that $F \to 0$ as $t \to T$ if $(\mathcal{M}^3, g)$ has no essential necklike points. Notice that $(R+\rho)^{-2} |\overset{\circ}{\mathrm{Rm}}|^2$ is scale invariant if we ignore the constant ρ, and that $(T-t)(R+\rho)$ is invariant under Type I rescaling.

Our first step is to derive an analog of Lemma 6.34.

LEMMA 9.10. *There exists a constant* $C = C(g(0))$ *such that*

$$|\mathrm{Rm}| \leq C (R + \rho) .$$

PROOF. We may assume without loss of generality that $\nu \leq \mu \leq \lambda$; and since $|\mathrm{Rm}|$ scales like R, we may assume that $\min_{x \in \mathcal{M}^3} \nu(x, 0) \geq -1$.

Set $A = e^6$ and $B = A/c$, so that $B(R+\rho) \geq A$. Let N denote the number of negative sectional curvatures at a point (x,t). There are two cases: If $-A \leq \nu$, it is easy to see that

$$|\mathrm{Rm}| \leq |\nu| + |\mu| + |\lambda| \leq \nu + \mu + \lambda + 2NA \leq (1 + 2NB)(R + \rho) .$$

If $\nu < -A$, then $|\nu| \leq R/3$ by Theorem 9.4; so $1 \leq N \leq 2$, and we have

$$|\mathrm{Rm}| \leq |\nu| + |\mu| + |\lambda| \leq \nu + \mu + \lambda + 2N|\nu| \leq \left(1 + \frac{2}{3}N\right) R.$$

□

LEMMA 9.11. *The function F satisfies a differential inequality of the form*

$$\frac{\partial}{\partial t} F \leq \Delta F + \frac{2(1-\varepsilon)}{R+\rho} \langle \nabla F, \nabla R \rangle + H,$$

where

$$H \doteqdot \frac{(T-t)^{\varepsilon}}{(R+\rho)^{3-\varepsilon}} \left[(B_1 + B_2 - G_1) |\mathrm{Rm}| |\overset{\circ}{\mathrm{Rm}}|^2 - G_2 \right].$$

Here, $B_1 \doteqdot C_1 \rho$ and $B_2 \doteqdot C_2 \varepsilon |\mathrm{Rm}|$ are bad terms, while $G_1 \doteqdot \frac{\varepsilon}{T-t}$ and $G_2 \doteqdot 2P$ are good terms. The function P is defined in (6.39).

PROOF. Recall that

$$|\overset{\circ}{\mathrm{Rm}}|^2 = \frac{1}{3} \left[(\lambda - \mu)^2 + (\lambda - \nu)^2 + (\mu - \nu)^2 \right]$$

and set

$$\begin{aligned} \Sigma &\doteqdot \lambda^3 + \mu^3 + \nu^3 + 3\lambda\mu\nu - \left(\lambda^2 \mu + \lambda^2 \nu + \mu^2 \lambda + \mu^2 \nu + \nu^2 \lambda + \nu^2 \mu \right) \\ &= \frac{3}{2} R |\overset{\circ}{\mathrm{Rm}}|^2 - \lambda (\mu - \nu)^2 - \mu (\lambda - \nu)^2 - \nu (\lambda - \mu)^2 . \end{aligned}$$

Taking $\varphi = (T-t)^{\varepsilon} |\overset{\circ}{\mathrm{Rm}}|^2$, $\psi = R + \rho$, $\alpha = 1$, and $\beta = 2 - \varepsilon$ in Lemma 6.33, we get

$$\begin{aligned} \frac{\partial}{\partial t} F = \Delta F &+ \frac{(T-t)^{\varepsilon}}{(R+\rho)^{2-\varepsilon}} \left(\frac{4}{3} \Sigma - 2 |\nabla \overset{\circ}{\mathrm{Rm}}|^2 - \frac{\varepsilon |\overset{\circ}{\mathrm{Rm}}|^2}{T-t} \right) \\ &- (2-\varepsilon) \frac{(T-t)^{\varepsilon} |\overset{\circ}{\mathrm{Rm}}|^2}{(R+\rho)^{3-\varepsilon}} \left(\lambda^2 + \mu^2 + \nu^2 + \lambda\mu + \lambda\nu + \mu\nu \right) \\ &- (2-\varepsilon)(3-\varepsilon) \frac{(T-t)^{\varepsilon} |\overset{\circ}{\mathrm{Rm}}|^2 |\nabla R|^2}{(R+\rho)^{4-\varepsilon}} \\ &+ 2(2-\varepsilon) \frac{(T-t)^{\varepsilon}}{(R+\rho)^{3-\varepsilon}} \left\langle \nabla |\overset{\circ}{\mathrm{Rm}}|^2, \nabla R \right\rangle . \end{aligned}$$

Collecting terms as in Lemma 6.34, we obtain

$$\begin{aligned}
\frac{\partial}{\partial t}F &= \Delta F + \frac{2(1-\varepsilon)}{R+\rho}\langle \nabla F, \nabla R\rangle \\
&\quad - \frac{(T-t)^{\varepsilon}}{(R+\rho)^{4-\varepsilon}}\left\{\begin{array}{c} 2\left|(R+\rho)(\nabla \overset{\circ}{\mathrm{Rm}}) - \nabla R \otimes \overset{\circ}{\mathrm{Rm}}\right|^2 \\ +\varepsilon(1-\varepsilon)|\overset{\circ}{\mathrm{Rm}}|^2|\nabla R|^2 \end{array}\right\} \\
&\quad + 2(T-t)^{\varepsilon}\frac{\varepsilon|\overset{\circ}{\mathrm{Rm}}|^2|\mathrm{Rc}|^2 - P}{(R+\rho)^{3-\varepsilon}} - \frac{\varepsilon|\overset{\circ}{\mathrm{Rm}}|^2}{(R+\rho)^{2-\varepsilon}(T-t)^{1-\varepsilon}} + \frac{4}{3}\frac{(T-t)^{\varepsilon}\rho\Sigma}{(R+\rho)^{3-\varepsilon}} \\
&\leq \Delta F + \frac{2(1-\varepsilon)}{R+\rho}\langle \nabla F, \nabla R\rangle \\
&\quad + \frac{(T-t)^{\varepsilon}}{(R+\rho)^{3-\varepsilon}}\left[\left(2\varepsilon|\mathrm{Rc}|^2 - \frac{\varepsilon(R+\rho)}{T-t}\right)|\overset{\circ}{\mathrm{Rm}}|^2 + \frac{4}{3}\rho\Sigma - 2P\right].
\end{aligned}$$

The result follows when we estimate $|\mathrm{Rc}|^2 \leq |\mathrm{Rm}|^2 \leq C(R+\rho)^2$ and

$$\Sigma \leq \frac{9}{2}(|\lambda|+|\mu|+|\nu|)|\overset{\circ}{\mathrm{Rm}}|^2 \leq \frac{27}{2}|\mathrm{Rm}|\,|\overset{\circ}{\mathrm{Rm}}|^2.$$

□

Notice that the bad terms $B_1 + B_2 = C_1\rho + C_2\varepsilon|\mathrm{Rm}|$ come from the $1/(R+\rho)^{2-\varepsilon}$ factor in F, whereas the good term $-G_1 = -\varepsilon/(T-t)$ comes from the $(T-t)^{\varepsilon}$ factor. The good term $-G_2 = -2P$ would be the only term present if $\rho = \varepsilon = 0$.

The next two lemmas constitute the second step in the proof of Theorem 9.9 and show that the pinching of the curvature operator improves if there are no essential necklike points.

LEMMA 9.12. *If* $|\mathrm{Rm} - R(\theta \otimes \theta)|^2 \geq \delta|\mathrm{Rm}|^2$ *for some* $\delta \in (0,1)$, *then*

$$P \geq \frac{\delta}{96(3-\delta)}|\mathrm{Rm}|^2|\overset{\circ}{\mathrm{Rm}}|^2.$$

PROOF. We may assume without loss of generality that $|\lambda| \geq |\mu| \geq |\nu|$. The hypothesis implies that

$$\mu^2 + \nu^2 + \mu\nu \geq \frac{\delta}{2}\left(\lambda^2 + \mu^2 + \nu^2\right),$$

and hence that $\mu^2 + \nu^2 \geq \frac{\delta}{3-\delta}\lambda^2$. Since $|\mu| \geq |\nu|$ by assumption, we have

$$\begin{aligned}
P &= \lambda^2(\mu-\nu)^2 + \mu^2(\lambda-\nu)^2 + \nu^2(\lambda-\mu)^2 \\
&\geq \lambda^2(\mu-\nu)^2 + \frac{\delta}{2(3-\delta)}\lambda^2(\lambda-\nu)^2 + \nu^2(\lambda-\mu)^2.
\end{aligned}$$

Now notice that

$$|\mathrm{Rm}|^2 = \lambda^2 + \mu^2 + \nu^2 \leq 3\lambda^2$$

and

$$|\overset{\circ}{\mathrm{Rm}}|^2 = \frac{1}{3}\left[(\lambda-\mu)^2 + (\lambda-\nu)^2 + (\mu-\nu)^2\right] \le \frac{4}{3}\left(\lambda^2+\mu^2+\nu^2\right) \le 4\lambda^2.$$

So if $|\nu| \le |\lambda|/2$, we have

$$P \ge \frac{\delta}{2(3-\delta)}\lambda^2(\lambda-\nu)^2 \ge \frac{\delta}{8(3-\delta)}\lambda^4 \ge \frac{\delta}{96(3-\delta)}|\mathrm{Rm}|^2|\overset{\circ}{\mathrm{Rm}}|^2,$$

while if $|\nu| > |\lambda|/2$, we get

$$\begin{aligned} P &> \lambda^2(\mu-\nu)^2 + \frac{\delta}{2(3-\delta)}\lambda^2(\lambda-\nu)^2 + \frac{1}{4}\lambda^2(\lambda-\mu)^2 \\ &\ge \frac{\delta}{2(3-\delta)}\lambda^2\left[(\mu-\nu)^2 + (\lambda-\nu)^2 + (\lambda-\mu)^2\right] \\ &\ge \frac{\delta}{2(3-\delta)}|\mathrm{Rm}|^2|\overset{\circ}{\mathrm{Rm}}|^2, \end{aligned}$$

because $\frac{1}{4} > \frac{\delta}{2(3-\delta)}$. □

LEMMA 9.13. *Assume that for every $c > 0$, there exist $\tau \in [0,T)$ and $\delta > 0$ such that for every $x \in \mathcal{M}^3$ and $t \in [\tau, T)$, either*

$$|\mathrm{Rm}(x,t)| < \frac{c}{T-t} \tag{9.6}$$

or

$$|\mathrm{Rm} - R(\theta\otimes\theta)| > \delta|\mathrm{Rm}| \tag{9.7}$$

for every unit 2-form θ at (x,t). Then $F \to 0$ uniformly as $t \to T$.

PROOF. By Lemma 9.11, F satisfies the differential inequality

$$\frac{\partial}{\partial t}F \le \Delta F + \frac{2(1-\varepsilon)}{R+\rho}\langle\nabla F, \nabla R\rangle + H, \tag{9.8}$$

where

$$H \doteqdot \frac{(T-t)^\varepsilon}{(R+\rho)^{3-\varepsilon}}\left[(B_1+B_2-G_1)|\mathrm{Rm}||\overset{\circ}{\mathrm{Rm}}|^2 - G_2\right].$$

The proof is in three steps.

We first show that $H \le 0$; by the maximum principle, this will prove that F is nonincreasing. Observe that for every $\varepsilon > 0$ and $\rho < \infty$, there exists $\tau \in [0,T)$ such that

$$B_1 = C_1\rho \le \frac{1}{3}\frac{\varepsilon}{T-t} = \frac{1}{3}G_1$$

for all $t \in [\tau, T)$. Now if condition (9.6) holds at (x,t) with $c = 1/(3C_2)$, then

$$B_2 = C_2\varepsilon|\mathrm{Rm}(x,t)| \le \frac{1}{3}\frac{\varepsilon}{T-t} = \frac{1}{3}G_1.$$

On the other hand, if condition (9.7) holds at (x,t), Lemma 9.12 implies there is $\eta = \eta(\delta)$ such that

$$P \geq \eta \left|\mathrm{Rm}\right|^2 |\overset{\circ}{\mathrm{Rm}}|^2,$$

whence it follows that choosing $\varepsilon \leq \frac{\eta}{2C_2}$ yields

$$B_2\left(\left|\mathrm{Rm}\right| |\overset{\circ}{\mathrm{Rm}}|^2\right) = C_2\varepsilon \left|\mathrm{Rm}\right|^2 |\overset{\circ}{\mathrm{Rm}}|^2 \leq P = \frac{1}{2}G_2.$$

In either case, we have used at most 2/3 of G_1 and at most 1/2 of G_2. Therefore

$$H \leq -\frac{(T-t)^{\varepsilon}}{(R+\rho)^{3-\varepsilon}}\left[\frac{1}{3}\frac{\varepsilon}{T-t}\left|\mathrm{Rm}\right| |\overset{\circ}{\mathrm{Rm}}|^2 + P\right] \leq 0. \tag{9.9}$$

We next improve estimate (9.8) Recall that either (9.6) or (9.7) holds at (x,t). In the first case, the inequality $R \leq C_3 \left|\mathrm{Rm}\right|$ implies there is $\tau < T$ such that $(R(x,t)+\rho)(T-t) \leq 2cC_3$ for $t \in [\tau, T)$, whence we get

$$H \leq -\frac{F}{(T-t)(R+\rho)}\left[\frac{\varepsilon}{3}\left|\mathrm{Rm}\right|\right] \leq -\frac{\varepsilon}{6cC_3}\left|\mathrm{Rm}\right| F$$

from (9.9). If (9.6) fails, then (9.7) holds and we have both conditions $\left|\mathrm{Rm}(x,t)\right| > c/(T-t)$ and $P \geq \eta\left|\mathrm{Rm}\right|^2 |\overset{\circ}{\mathrm{Rm}}|^2$. Thus by our choice of τ, we obtain

$$\begin{aligned} H &\leq -\frac{(T-t)^{\varepsilon}}{(R+\rho)^{3-\varepsilon}}\left[\eta\left|\mathrm{Rm}\right|^2 |\overset{\circ}{\mathrm{Rm}}|^2\right] \\ &= -\eta\frac{\left|\mathrm{Rm}\right|^2}{(R+\rho)}F \leq -\eta\frac{\left|\mathrm{Rm}\right| c}{(R+\rho)(T-t)}F \leq -\frac{\eta}{2C_3}\left|\mathrm{Rm}\right| F. \end{aligned}$$

Hence in either case we have

$$H \leq -C(\delta, c, C_2, C_3)\left|\mathrm{Rm}\right| F,$$

and thus

$$\frac{\partial}{\partial t}F \leq \Delta F + \frac{2(1-\varepsilon)}{R+\rho}\langle \nabla F, \nabla R\rangle - C\left|\mathrm{Rm}\right| F.$$

Now we can show that F actually tends to zero uniformly. Since F is bounded above, we may define $F_{\max}(t) \doteqdot \sup_{\mathcal{M}^3} F(\cdot, t)$. Recall that in dimension 3, there is a constant $A < \infty$ such that $|\overset{\circ}{\mathrm{Rm}}|^2 \leq \left|\mathrm{Rm}\right|^2 < A(R+\rho)^2$. Let $\alpha > 0$ be given, and consider any (x,t) with $\tau \leq t < T$. Observe that $(T-t)\cdot(R+\rho)(x,t) \leq \alpha$ only if

$$F(x,t) = \left[(T-t)(R+\rho)\right]^{\varepsilon}\frac{|\overset{\circ}{\mathrm{Rm}}|^2}{(R+\rho)^2} < A\alpha^{\varepsilon}.$$

Thus if $F(x,t) \geq A\alpha^{\varepsilon}$, we must have

$$\left|\mathrm{Rm}\right| \geq c(R+\rho) > \frac{c\alpha}{T-t}$$

and hence

$$\frac{\partial F}{\partial t} \leq \Delta F + \frac{2(1-\varepsilon)}{R+\rho}\langle \nabla F, \nabla R\rangle - \frac{Cc\alpha}{T-t}F.$$

So the maximum principle implies that whenever $F_{\max}(t) \geq A\alpha^{\varepsilon}$ at some $t \in [\tau, T)$, we have

$$\frac{d}{dt}\left[(T-t)^{-Cc\alpha} F_{\max}(t)\right] \leq 0$$

in the sense of lim sup of forward difference quotients. In particular, we get

$$\frac{d}{dt}F_{\max} \leq -\frac{CcA\alpha^{1+\varepsilon}}{T-t} \doteqdot -\frac{\beta}{T-t}$$

if $F_{\max}(t) \geq A\alpha^{\varepsilon}$. Since F is nonincreasing, this implies that for all $t \in [\tau, T)$, we have

$$F_{\max}(t) \leq \max\left\{A\alpha^{\varepsilon},\ F_{\max}(\tau) + \beta \log \frac{T-t}{T-\tau}\right\}.$$

It follows that $F_{\max}(t) \leq A\alpha^{\varepsilon}$ for t sufficiently close to T. □

To finish the proof of Theorem 9.9, we need to verify that the solution satisfies an injectivity radius estimate on the scale of its maximum curvature. This follows from Perelman's *No Local Collapsing* result (Theorem 8.17). Armed with this fact, we now proceed to the final step of the proof.

Proof of Theorem 9.9. We are finally ready to take a limit of dilations about the singularity at time $T < \infty$, even though we do not know whether the singularity is of Type I or Type IIa. In either case, choose a sequence (x_i, t_i) as in Section 3.1 of Chapter 8, taking care that the curvature at each (x_i, t_i) is uniformly comparable to its maximum at time t_i. For $c_0 \in (0,1)$, set

$$\mathcal{M}_i^{c_0} \doteqdot \left\{x \in \mathcal{M}^3 : |\mathrm{Rm}(x,t_i)| \geq c_0 \max_{\mathcal{M}^3} |\mathrm{Rm}(\cdot, t_i)|\right\}.$$

Then by Lemma 8.19, there is $c_1 > 0$ such that for all $x \in \mathcal{M}_i^{c_0}$,

$$|\mathrm{Rm}(x,t_i)| \geq \frac{c_0 c_1}{T-t_i}.$$

That is, every point in $\mathcal{M}_i^{c_0}$ is Type I c_0c_1-essential. For such points, we have $(T-t_i)(R(x,t_i)+\rho) \geq c > 0$ by Lemma 9.10.

By the compactness theorem stated in Section 3 of Chapter 7 and Corollary 9.7, the pointed sequence $\left(\mathcal{M}^3, g_i(t), x_i\right)$ limits to a complete ancient solution $\left(\mathcal{M}^3_\infty, g_\infty(t), x_\infty\right)$ of the Ricci flow having bounded nonnegative sectional curvature. Since we assumed there are eventually no Type I c-essential δ-necklike points, the limit actually has positive sectional curvature. By construction of that limit, there is for every point $y \in \mathcal{M}^3_\infty$ a subsequence of points $\{y_i\}$ in $\mathcal{M}^3$, orthonormal frames $\{\mathcal{F}_i\}$ at (y_i, t_i), and an orthonormal frame $\mathcal{F}_\infty$ at $(y, 0)$ such that

$$\lim_{i\to\infty} \frac{\mathrm{Rm}\,(y_i, t_i)}{|\mathrm{Rm}\,(x_i, t_i)|} = \mathrm{Rm}_{\,\infty}\,(y, 0)\,,$$

where $\mathrm{Rm}_\infty \doteqdot \mathrm{Rm}\,[g_\infty]$ and the curvature operators are represented in the respective orthonormal frames. Because $\mathrm{Rm}_\infty\,(y, 0) > 0$, we have $y_i \in \mathcal{M}_i^{c_0}$ for some $c_0 > 0$ independent of i. Because $t_i \to T$, Lemma 9.13 implies that

$$\lim_{i\to\infty} \left[(T - t_i)\,(R\,(y_i, t_i) + \rho)\right]^\varepsilon \frac{|\overset{\circ}{\mathrm{Rm}}\,(y_i, t_i)\,|^2}{(R\,(y_i, t_i) + \rho)^2} = 0.$$

Then since $(T - t_i)\,(R\,(y_i, t_i) + \rho) \geq c > 0$ and $R\,(y_i, t_i) \to \infty$, we conclude that

$$0 = \lim_{i\to\infty} \frac{|\overset{\circ}{\mathrm{Rm}}\,(y_i, t_i)\,|^2}{R\,(y_i, t_i)^2} = \frac{|\overset{\circ}{\mathrm{Rm}}_\infty\,(y, 0)\,|^2}{(R_\infty\,(y, 0))^2}.$$

Since $y \in \mathcal{M}^3$ was arbitrary and $R_\infty > 0$, this implies that $\overset{\circ}{\mathrm{Rm}}_\infty\,(\cdot, 0) \equiv 0$. Then Schur's lemma implies that $g_\infty\,(0)$ is a metric of constant positive sectional curvature on $\mathcal{M}^3_\infty$. By Myers' Theorem, $\mathcal{M}^3_\infty$ is compact, hence diffeomorphic to $\mathcal{M}^3$. It follows that $g\,(t)$ has uniformly positive sectional curvature for t close enough to T, hence that $\tilde{g}\,(t)$ limits to a metric of constant positive sectional curvature. □

REMARK 9.14. Mao-Pei Tsui has suggested an alternate proof of the fact that F remains bounded if for every $c > 0$, there are $\tau < T$ and $\delta > 0$ such that no c-essential δ-necklike points exist after time τ. The argument uses the enhanced maximum principle for systems (Theorem 4.10) and resembles our first proof of Theorem 6.30. It is simpler than but basically equivalent to the argument given above in Lemma 9.13.

ALTERNATE PROOF THAT F IS BOUNDED. Recall that $\lambda \geq \mu \geq \nu$ and that $\lambda + \mu + \nu + \rho = R + \rho \geq c > 0$. Since

$$\frac{d}{dt}\,(\lambda + \mu + \nu + \rho) = \lambda^2 + \mu^2 + \nu^2 + \lambda\mu + \lambda\nu + \mu\nu,$$

we have

$$\frac{d}{dt} \log\,(\lambda + \mu + \nu + \rho) = \lambda + \frac{\mu^2 + \nu^2 + \mu\nu - \rho\lambda}{\lambda + \mu + \nu + \rho}.$$

Fix a point where $\lambda > \nu$ and define

$$\Phi \doteqdot (T - t)^\varepsilon \frac{\lambda - \nu}{(\lambda + \mu + \nu + \rho)^{1-\varepsilon}}.$$

Then since $\frac{d}{dt} \log\,(\lambda - \nu) = \lambda + \nu - \mu$, we compute

$$\frac{d}{dt} \log \Phi = \varepsilon\lambda - \frac{\varepsilon}{T - t} - (\mu - \nu) + (1 - \varepsilon)\,\frac{\rho\lambda - \left(\mu^2 + \nu^2 + \mu\nu\right)}{\lambda + \mu + \nu + \rho}.$$

We claim there is a constant $\Gamma < \infty$ depending only on $g(0)$ such that

$$\frac{\lambda}{\lambda + \mu + \nu + \rho} \leq \Gamma.$$

Indeed, since the left-hand side is scale invariant, we may assume without loss of generality that

$$\min_{x \in \mathcal{M}^3} \nu(x, 0) \geq -1.$$

Let $A = e^6$. If $-3A \leq \lambda \leq 3A$, then

$$\frac{\lambda}{\lambda + \mu + \nu + \rho} \leq \frac{3A}{c}.$$

If not, then by Theorem 9.4 we have either $\nu \geq -A$ and $\lambda > 3A$ or else $\nu < -A$ and $3|\nu| \leq R < \lambda$; in either case,

$$\frac{\lambda}{\lambda + \mu + \nu + \rho} < \frac{\lambda}{\lambda/3 + \rho} \leq 3.$$

Now for $\varepsilon > 0$ to be chosen later, we may by the claim above choose τ_ε large enough so that when $\tau_\varepsilon \leq t < T$, we have

$$(1-\varepsilon)\rho\frac{\lambda}{\lambda + \mu + \nu + \rho} \leq (1-\varepsilon)\rho\Gamma \leq \frac{\varepsilon}{2(T-t)}$$

and hence

$$\frac{d}{dt}\log\Phi \leq \varepsilon\lambda - \frac{\varepsilon}{2(T-t)} - \frac{1-\varepsilon}{2}\frac{\mu^2+\nu^2}{\lambda+\mu+\nu+\rho}.$$

By hypothesis, for every $c > 0$ there are $\delta(c) > 0$ and $\tau_c < T$ such that $\mathbb{M}(t) \notin \mathcal{X}(t)$ for $\tau_c \leq t < T$, where

$$\mathcal{X}(t) \doteqdot \left\{\mathbb{P} : (T-t)\lambda(\mathbb{P}) \geq c > 0 \quad \text{and} \quad \mu(\mathbb{P})^2 + \nu(\mathbb{P})^2 \leq \delta \cdot \lambda(\mathbb{P})^2\right\}$$

is the *avoidance set* discussed in Section 3 of Chapter 4. Choose $C(\varepsilon, \mathrm{c}) < \infty$ large enough so that $\mathbb{M}(\tau) \in \mathcal{K}(\tau)$, where $\tau = \max\{\tau_\varepsilon, \tau_c\}$ and

$$\mathcal{K}(t) \doteqdot \left\{\mathbb{P} : \frac{(T-t)^\varepsilon(\lambda(\mathbb{P}) - \nu(\mathbb{P}))}{(\lambda(\mathbb{P}) + \mu(\mathbb{P}) + \nu(\mathbb{P}) + \rho)^{1-\varepsilon}} \leq C\right\}.$$

Notice that $\mathcal{K}(t)$ is convex. We next claim that when $c = 1/2$, there is $\varepsilon > 0$ small enough that $\frac{d}{dt}\log\Phi \leq 0$ whenever $\tau \leq t < T$ and $\mathbb{M}(t) \in \mathcal{K}(t) \setminus \mathcal{X}(t)$. This claim implies that $\mathbb{M}(t)$ remains in $\mathcal{K}(t) \setminus \mathcal{X}(t)$, hence that $\Phi \leq C$. By Theorem 4.10, this implies that F is bounded above, because

$$|\overset{\circ}{\mathrm{Rm}}|^2 = \frac{1}{3}\left[(\lambda-\mu)^2 + (\lambda-\nu)^2 + (\mu-\nu)^2\right] \leq (\lambda-\nu)^2.$$

To prove the claim, notice that $\mathbb{M}(t) \notin \mathcal{X}(t)$ only if $\lambda(T-t) < \mathrm{c}$ or $\mu^2 + \nu^2 > \delta\lambda^2$. In the first case, we have

$$\frac{d}{dt}\log\Phi \leq \varepsilon\lambda - \frac{\varepsilon}{2(T-t)} - \frac{1-\varepsilon}{2}\frac{\mu^2+\nu^2}{\lambda+\mu+\nu+\rho} \leq \varepsilon\left(\lambda - \frac{1}{2(T-t)}\right) < 0.$$

On the other hand, if $\lambda(T-t) \geq c$, then $\mu^2+\nu^2 > \delta\lambda^2$ and we have

$$\frac{d}{dt}\log\Phi < \varepsilon\lambda - \frac{\varepsilon}{2(T-t)} - \frac{1-\varepsilon}{2}\frac{\delta\lambda^2}{\lambda+\mu+\nu+\rho} \leq 0$$

for $\varepsilon > 0$ small enough. □

5. Necklike points in ancient solutions

The objective of this section is to show that an ancient solution of positive sectional curvature is either isometric to a spherical space form or else contains points at arbitrarily ancient times where its geometry is arbitrarily close to the product of a surface and a line. We begin with some general observations about ancient solutions — including the fact that every ancient solution (of any dimension) has nonnegative scalar curvature.

LEMMA 9.15. *Let $(\mathcal{M}^n, g(t))$ be a complete ancient solution of the Ricci flow. Assume that $R_{\min}(t) \doteqdot \inf_{x\in\mathcal{M}^n} R(x,t)$ is finite for all $t \leq 0$ and that there is a continuous function $K(t)$ such that $|\mathrm{sect}[g(t)]| \leq K(t)$. Then $g(t)$ has nonnegative scalar curvature for as long as it exists.*

PROOF. Recall that the parabolic maximum principle works on noncompact as well as on compact manifolds. Recall also that

$$\frac{\partial R}{\partial t} = \Delta R + 2|\mathrm{Rc}|^2 \geq \Delta R + \frac{2}{n}R^2.$$

So if R ever becomes nonnegative, it remains so for as long as the solution exists. On the other hand, if $R_{\min}(t_1) < 0$, then $R_{\min}(t_0) < 0$ for all $t_0 \leq t_1$, and one has

$$R_{\min}(t) \geq \frac{n}{n(R_{\min}(t_0))^{-1} - 2(t-t_0)} \geq -\frac{n}{2(t-t_0)}$$

for all $t > t_0$. Since the solution is ancient, we may let $t_0 \searrow -\infty$, obtaining

$$R_{\min}(t) \geq -\lim_{t_0\to-\infty}\frac{n}{2(t-t_0)} = 0.$$

□

REMARK 9.16. In dimension $n=3$, we can strengthen this result, as we proved in Corollary 9.8.

In Lemma 8.19, we derived a lower bound for the blow-up rate of solutions that develop singularities in finite time. The analogous result for ancient solutions is the following:

LEMMA 9.17. *Let $(\mathcal{M}^n, g(t))$ be an ancient solution with nonnegative Ricci curvature. Then there exists a constant $c_0 > 0$ depending only on n such that*

$$\liminf_{t\to-\infty}|t|\sup_{x\in\mathcal{M}^n}|\mathrm{Rm}(x,t)| \geq c_0.$$

PROOF. We may assume that $\sup_{x\in\mathcal{M}^n} |\mathrm{Rm}(x,t)|$ is finite for $t \leq 0$. Because $\mathrm{Rc} \geq 0$, we can estimate the growth of the scalar curvature R by

$$\frac{\partial}{\partial t}R = \Delta R + 2|\mathrm{Rc}|^2 \leq \Delta R + 2R^2.$$

Hence $R_{\max}(t) \doteqdot \sup_{x\in\mathcal{M}^n} R(x,t) \leq C(n) \sup_{x\in\mathcal{M}^n} |\mathrm{Rm}(x,t)|$ satisfies the differential inequality

$$\frac{d}{dt}(R_{\max})^{-1} \geq -2.$$

So for all $t < 0$, we have

$$R_{\max}(t) \geq \frac{1}{(R_{\max}(0))^{-1} - 2t}.$$

Letting $t \to -\infty$, we obtain

$$\liminf_{t\to-\infty} (|t| R_{\max}(t)) \geq \frac{1}{2},$$

which implies the result. □

REMARK 9.18. Because nonnegative Ricci curvature is preserved only in dimension $n = 3$, one usually must assume that the curvature operator is nonnegative in order to apply the lemma. (See Section 2.)

We are now ready to state and prove our main assertion. We shall say that (x,t) is an **ancient Type I c-essential point** if

$$|\mathrm{Rm}(x,t)| \cdot |t| \geq c > 0.$$

By modifying the proof of Theorem 9.9, we obtain the following:

THEOREM 9.19. *Let $(\mathcal{M}^3, g(t))$ be a complete ancient solution of the Ricci flow with positive sectional curvature. Suppose that*

$$\sup_{\mathcal{M}^3\times(-\infty,0]} |t|^\gamma R(x,t) < \infty$$

for some $\gamma > 0$. Then either $(\mathcal{M}^3, g(t))$ is isometric to a spherical space form, or else there exists a constant $c > 0$ such that for all $\tau \in (-\infty, 0]$ and $\delta > 0$, there are $x \in \mathcal{M}^3$ and $t \in (-\infty, \tau)$ such that (x,t) is an ancient Type I c-essential point and a δ-necklike point.

PROOF. We will show that if for every $c > 0$ there are $\tau \in (-\infty, 0]$ and $\delta > 0$ such that there are no ancient Type I c-essential δ-necklike points before time τ, then $(\mathcal{M}^3, g(t))$ is a shrinking spherical space form.

By hypothesis on the ancient solution, there is $\gamma > 0$ small enough that

$$K \doteqdot \sup_{\mathcal{M}^3\times(-\infty,0]} |t|^\gamma R(x,t) < \infty.$$

(When $\gamma = 1$, K corresponds to the bound in the definition of ancient Type I singularity models.) Since the scalar curvature of $(\mathcal{M}^3, g(t))$ is positive, the function

$$G \doteqdot |t|^{\gamma\varepsilon/2} \frac{|\overset{\circ}{\mathrm{Rm}}|^2}{R^{2-\varepsilon}}$$

is well-defined. Since the sectional curvatures are positive, we have the estimates

$$G \le CR^{\varepsilon} |t|^{\gamma\varepsilon/2} \le CK^{\varepsilon} |t|^{-\gamma\varepsilon/2},$$

which show that G is bounded for all times $-\infty < t \le 0$ and satisfies

$$\lim_{t\to-\infty} \max_{x\in\mathcal{M}^3} G(x,t) = 0. \tag{9.10}$$

Note that our choice of G is similar to taking $T = 0$ and $\rho = 0$ in the function F studied in Theorem 9.9, except that we now consider the possibly smaller power $(T-t)^{\gamma} = |t|^{\gamma}$. This modification does not significantly harm the evolution of G: we will once again be able to show that if there are no essential necklike points, then there exists $\varepsilon > 0$ small enough that G is a subsolution of the heat equation, hence that its maximum cannot increase. The advantage of the modification is that we can then use (9.10) to conclude that $G \equiv 0$, hence that $(\mathcal{M}^3, g(t))$ is complete and locally isometric to a round $\mathcal{S}^3$, hence that it is compact and globally isometric to a spherical space form $\mathcal{S}^3/\Gamma$.

If there are no ancient c-essential δ-necklike points on the time interval $(-\infty, \tau]$, then for every $x \in \mathcal{M}^3$ and $t \in (-\infty, \tau)$ either

$$|\mathrm{Rm}(x,t)| \cdot |t| < c \tag{9.11}$$

or we have

$$|\mathrm{Rm} - R(\theta \otimes \theta)| > \delta |\mathrm{Rm}| \tag{9.12}$$

for every unit 2-form θ at (x,t). Taking $\varphi = (-t)^{\gamma\varepsilon/2} |\overset{\circ}{\mathrm{Rm}}|^2$, $\psi = R$, $\alpha = 1$, and $\beta = 2 - \varepsilon$ in Lemma 6.33 and estimating as in Lemma 9.11, we get the differential inequality

$$\frac{\partial}{\partial t} G \le \Delta G + \frac{2(1-\varepsilon)}{R} \langle \nabla G, \nabla R \rangle + 2J,$$

where

$$J \doteqdot \frac{|t|^{\gamma\varepsilon/2}}{R^{3-\varepsilon}} \left[\varepsilon |\overset{\circ}{\mathrm{Rm}}|^2 \left(|\mathrm{Rm}|^2 - \frac{\gamma R}{4|t|} \right) - P \right]$$

and $P \ge 0$ is defined in (6.39). Fix any (x,t) with $t < \tau \le 0$. If (9.11) holds there with $c \le \gamma/8$, then we have

$$|\mathrm{Rm}|^2 - \frac{\gamma R}{4|t|} \le R\left(|\mathrm{Rm}| - \frac{\gamma}{4|t|} \right) < R\left(\frac{c}{|t|} - \frac{\gamma}{4|t|} \right) \le -\frac{\gamma R}{8|t|}$$

and hence

$$J < -\frac{\gamma\varepsilon}{8|t|} G.$$

If (9.12) holds, then by Lemma 9.12 there is $\eta(\delta)$ such that

$$P \geq \eta |\mathrm{Rm}|^2 |\overset{\circ}{\mathrm{Rm}}|^2,$$

whence it follows that taking $\varepsilon \leq \eta$ gives

$$J \leq -\frac{\gamma\varepsilon}{4|t|} G.$$

Thus in either case, G is a subsolution of the heat equation for all times $-\infty < t < \tau$, because

$$\frac{\partial}{\partial t} G \leq \Delta G + \frac{2(1-\varepsilon)}{R} \langle \nabla G, \nabla R \rangle - \frac{\gamma\varepsilon}{2|t|} G.$$

□

6. Type I ancient solutions on surfaces

The objective of this section is to prove that the shrinking round 2-sphere is the only nonflat Type I ancient solution of the Ricci flow on a surface. Because of dimension reduction, this fact plays an important role in the classification of 3-dimensional singularities.

Before we classify Type I ancient solutions on surfaces, we shall establish a result which is of considerable independent interest. It follows from a generalization of the LYH differential Harnack estimates introduced in Section 10 of Chapter 5. We will discuss the full family of LYH differential Harnack estimates for the Ricci flow in a chapter of the successor to this volume. For now, we simply state the following result from [**61**].

PROPOSITION 9.20 (LYH estimate, trace version). *If $(\mathcal{M}^n, g(t))$ is a solution of the unnormalized Ricci flow on a compact manifold with initially positive curvature operator, then for any vector field X on $\mathcal{M}^n$ and all times $t > 0$ such that the solution exists, one has*

$$0 \leq \frac{\partial}{\partial t} R + \frac{R}{t} + 2\langle \nabla R, X \rangle + 2\,\mathrm{Rc}(X, X).$$

COROLLARY 9.21. *If $(\mathcal{M}^n, g(t))$ is a solution of the unnormalized Ricci flow on a compact manifold with initially positive curvature operator, then the function tR is pointwise nondecreasing for all $t \geq 0$ that the solution exists. If $(\mathcal{M}^n, g(t))$ is also ancient, then R itself is pointwise nondecreasing.*

PROOF. Taking $X = 0$, we have

$$\frac{\partial}{\partial t}(tR) = t\left(\frac{\partial}{\partial t} R + \frac{R}{t}\right) \geq 0$$

for all $t \geq 0$ such that the solution exists. This proves the first assertion. To prove the second assertion, notice that one has

$$\frac{\partial}{\partial t} R + \frac{R}{t+\alpha} \geq 0$$

whenever the solution exists for $t \in [-\alpha, \omega)$. When the solution is ancient, letting $\alpha \to \infty$ gives the result. $\square$

In the successor to this volume, we will also prove the following statement.

LEMMA 9.22. *If $(\mathcal{N}^2, h(t))$ is a complete ancient Type I solution of the Ricci flow with strictly positive scalar curvature, then $\mathcal{N}^2$ is compact.*

Assuming this for now, we will conclude this volume by characterizing all complete ancient solutions of the Ricci flow on surfaces. We begin with a classification of all complete ancient Type I 2-dimensional solutions.

PROPOSITION 9.23. *A complete ancient Type I solution $(\mathcal{N}^2, h(t))$ of the Ricci flow on a surface is a quotient of either a shrinking round $\mathcal{S}^2$ or else a flat $\mathbb{R}^2$.*

PROOF. Since $(\mathcal{N}^2, h(t))$ is an ancient Type I solution, it exists on a maximal time interval $(-\infty, \omega)$ containing $t = 0$ and satisfies a curvature bound of the form

$$|R(x,t)| \leq C_n |\mathrm{Rm}(x,t)| \leq \frac{C}{|t|}$$

for all $x \in \mathcal{M}^3$ and $t \in (-\infty, 0)$. By Lemma 9.15, $R \geq 0$ on $\mathcal{N}^2 \times (-\infty, \omega)$. By the strong maximum principle, $(\mathcal{N}^2, h(t))$ is either flat and hence a quotient of $\mathbb{R}^2$ or else has strictly positive scalar curvature. From now on, we assume the latter.

By Lemma 9.22 (below) $\mathcal{N}^2$ is compact, hence diffeomorphic to either $\mathcal{S}^2$ or $\mathbb{RP}^2$. By passing to the twofold cover if necessary, we may assume $\mathcal{N}^2 \approx \mathcal{S}^2$. The area A of $\mathcal{N}^2$ evolves by

$$\frac{d}{dt}A = -\int_{\mathcal{N}^2} R\,dA = -4\pi\chi(\mathcal{N}^2),$$

where $\chi(\mathcal{N}^2)$ is the Euler characteristic. Since $A(t) \to 0$ as $t \to \omega$ by [60], we have

$$A(t) = 4\pi\chi(\mathcal{N}^2)\cdot(\omega - t) = 8\pi(\omega - t).$$

Recall (Section 8 of Chapter 5) that the entropy

$$\begin{aligned} E(h(t)) &\doteqdot \int_{\mathcal{N}^2} R\log(RA)\,dA \\ &= \int_{\mathcal{N}^2} R\log[R(\omega - t)]\,dA + 4\pi\chi(\mathcal{N}^2)\log\left[4\pi\chi(\mathcal{N}^2)\right] \end{aligned}$$

is a scale-invariant and nonincreasing function of time. Because $(\mathcal{N}^2, h(t))$ is an ancient Type I solution, we have

$$\sup_{\mathcal{N}^2\times(-\infty,\omega)} R(x,t)\cdot(\omega - t) \leq C < \infty,$$

whence it follows that the limit

$$E_{-\infty} \doteqdot \lim_{t\to-\infty} E(h(t)) \tag{9.13}$$

exists and is finite.

We want to use the compactness theorem to take a limit backwards in time and get a new compact ancient solution of constant entropy. To do this, we need a diameter bound. Since $(\mathcal{N}^2, h(t))$ is an orientable positively-curved surface such that

$$R_{\max}(t) \doteqdot \max_{\mathcal{N}^2} R(\cdot, t) \leq C \cdot |t|^{-1}$$

for $t \in (-\infty, 0)$, Klingenberg's Theorem (Theorem 5.9 of [**27**]) implies there exists $c_1 > 0$ independent of time such that

$$\operatorname{inj}\left(\mathcal{N}^2, h(t)\right) \geq \frac{\pi}{\sqrt{R_{\max}(t)/2}} \geq c_1\sqrt{|t|}.$$

Then the area comparison theorem (§3.4 of [**25**]) implies there is $c_2 > 0$ independent of time such that

$$\begin{aligned} 4\pi\chi\left(\mathcal{N}^2\right)\cdot(\omega - t) = A(t) &\geq c_2 \cdot \operatorname{diam}\left(\mathcal{N}^2, h(t)\right)\cdot \operatorname{inj}\left(\mathcal{N}^2, h(t)\right) \\ &\geq c_1 c_2\sqrt{|t|}\cdot \operatorname{diam}\left(\mathcal{N}^2, h(t)\right), \end{aligned}$$

hence $C_3 < \infty$ such that

$$\operatorname{diam}\left(\mathcal{N}^2, h(t)\right) \leq C_3 \cdot \sqrt{|t|} \tag{9.14}$$

for all $t \in (-\infty, -1]$.

Now take any sequence of points and times (x_i, t_i) with

$$R(x_i, t_i) = R_{\max}(t_i)$$

and $t_i \searrow -\infty$. Consider the dilated solutions

$$h_i(t) \doteqdot R(x_i, t_i)\cdot h\left(t_i + \frac{t}{R(x_i, t_i)}\right).$$

At each fixed $x \in \mathcal{N}^2$, Corollary 9.21 implies that the function $t \mapsto R(x,t)$ is nondecreasing in time. Thus the scalar curvature R of h satisfies

$$\max_{\mathcal{N}^2\times(-\infty, t_i]} R(x,t) = R(x_i, t_i)$$

and the scalar curvature R_i of h_i satisfies

$$\max_{\mathcal{N}^2\times(-\infty, 0]} R_i(x,t) = 1 = R_i(x_i, 0).$$

Then applying Klingenberg's Theorem again, we obtain the injectivity radius estimate

$$\operatorname{inj}\left(\mathcal{N}^2, h_i(0)\right) \geq \frac{\pi}{\sqrt{1/2}} = \sqrt{2}\pi.$$

By (9.14) and the Type I hypothesis, there is $C_4 < \infty$ giving the diameter bound

$$\operatorname{diam}\left(\mathcal{N}^2, h_i(0)\right) = \sqrt{R(x_i, t_i)}\operatorname{diam}\left(\mathcal{N}^2, h(t_i)\right) \leq C_3\sqrt{R(x_i, t_i)\,|t_i|} \leq C_4.$$

Hence by the Compactness Theorem (Section 3 of Chapter 7), there is a subsequence $\left(\mathcal{N}^2, h_i(t), x_i\right)$ that limits to an ancient solution

$$\left(\mathcal{N}^2, h_{-\infty}(0), x_{-\infty}\right)$$

of the Ricci flow on a compact surface of positive curvature. By (9.13), the scale-invariant entropy of the limit satisfies

$$E\left(h_{-\infty}(t)\right) = \lim_{i\to\infty} E\left(h_i(t)\right) = \lim_{i\to\infty} E\left(h\left(t_i + \frac{t}{R(x_i, t_i)}\right)\right) \equiv E_{-\infty},$$

because $\lim_{i\to\infty}\left(t_i + t/R(x_i, t_i)\right) = -\infty$ for all $t \in (-\infty, 0]$. In particular, the entropy of the limit is independent of time. Thus by Proposition 5.40, $h_{-\infty}(t)$ is a shrinking round 2-sphere. Since constant-curvature metrics minimize entropy among all metrics on $\mathcal{S}^2$, we have $E_{-\infty} \leq E(h(t))$. But since $E(h(t))$ is nonincreasing, $E_{-\infty} \geq E(h(t))$. It follows that $E(h(t)) \equiv E_{-\infty}$, hence that $h(t)$ is a shrinking 2-sphere of constant curvature. □

We now consider the case that the solution is Type II. Recall that we discussed backwards limits in Section 6 of Chapter 8. The self-similar solution corresponding to the cigar soliton was introduced in Section 2 of Chapter 2.

PROPOSITION 9.24. *Let $\left(\mathcal{M}^2, g(t)\right)$ be a complete Type II ancient solution of the Ricci flow defined on an interval $(-\infty, \omega)$, where $\omega > 0$. Assume there exists a function $K(t)$ such that $|R| \leq K(t)$. Then either $g(t)$ is flat or else there exists a backwards limit that is the self-similar cigar solution.*

PROOF. By Lemma 9.15, we have $R \geq 0$. The strong maximum principle implies that either $R \equiv 0$ or else $R > 0$. From now on, we assume the latter. By Corollary 9.21, the scalar curvature is pointwise nondecreasing. Hence we have the uniform bound $R \leq K(0)$ on $(-\infty, 0]$. Now if $\mathcal{M}^2$ is compact, then Klingenberg's Theorem implies that its injectivity radius is bounded from below by $\pi/\sqrt{K(0)}$. If $\mathcal{M}^2$ is noncompact, we get the same bound from Theorem B.65. So in either case, we can take a limit backwards in time as in Section 6 of Chapter 8. Since the solution is Type II, we obtain an eternal solution $\left(\mathcal{M}^2_\infty, g_\infty(t)\right)$ satisfying $0 < R_\infty \leq 1$ and $R(x_\infty, 0) = 1$. By Lemma 5.96, $\left(\mathcal{M}^2_\infty, g_\infty(t)\right)$ is isometric to the cigar solution $\left(\mathbb{R}^2, g_\Sigma(t)\right)$. □

COROLLARY 9.25. *Let $\left(\mathcal{M}^2, g(t)\right)$ be a complete ancient solution defined on $(-\infty, \omega)$, where $\omega > 0$. Assume that its curvature is bounded by some function of time alone. Then either the solution is flat, or it is a round shrinking sphere, or there exists a backwards limit that is the cigar.*

Notes and commentary

The main references for this chapter are [**58**], [**59**], and [**63**]. The estimate that the curvatures tend to positive in Section 3 was proved in Theorem 4.1 of [**66**]. Similar estimates, without the time term, were earlier proved independently by Ivey [**77**] and Hamilton (Theorem 24.4 of [**63**]). Theorem

9.9 is essentially Theorem 24.7 of [**63**]; the minor extension is that one assumes that the point is Type I c-essential rather than that the singularity is Type I and the point is c-essential. The alternate proof we gave of the boundedness of F was suggested to us by Mao-Pei Tsui. The nonnegativity of the sectional curvatures of ancient solutions was observed by Hamilton. Theorem 9.19, which does not seem to appear explicitly in print, is analogous to Corollary B3.5 in [**65**]. Proposition 9.23 is Theorem 26.1 of [**63**].

APPENDIX A

The Ricci calculus

From the viewpoint of Riemannian geometry, the Ricci flow is a very natural evolution equation, because it can be formulated entirely in terms of intrinsically-defined geometric quantities and thus enjoys full diffeomorphism invariance. Nevertheless, when studying the properties of the Ricci flow as a PDE and in particular computing the evolution equations satisfied by these intrinsic geometric quantities, it is often more convenient to use (classical) local coordinates rather than (presumably more modern) invariant notation. Consequently, we provide here a brief review of some aspects of the Ricci calculus, concentrating on its application to covariant differentiation in local coordinates.

1. Component representations of tensor fields

If $\mathcal{M}^n$ is a smooth manifold and $\pi : \mathcal{E} \to \mathcal{M}^n$ is any differentiable vector bundle, standard constructions in geometry produce its dual bundle $\mathcal{E}^*$, various tensor products like $\mathcal{E} \otimes \mathcal{E}^* \otimes \mathcal{E}^*$, symmetric products like $\mathcal{E} \otimes_S \mathcal{E}$, and exterior products like $\wedge^p(\mathcal{E})$. Everything we shall say below extends readily, *mutatis mutandis,* to this general situation. However, to keep the exposition concise and concrete, we will specialize to the case that $\mathcal{E} = T\mathcal{M}^n$ is the tangent bundle of $\mathcal{M}^n$.

Let $(\mathcal{M}^n, g)$ be a (smooth) Riemannian manifold. Then a (p, q)-tensor A is a smooth section of the vector bundle

$$T_p^q \mathcal{M}^n \doteqdot (\otimes^p T^* \mathcal{M}^n) \bigotimes (\otimes^q T\mathcal{M}^n)$$

over $\mathcal{M}^n$. We denote this by writing $A \in C^\infty(T_p^q \mathcal{M}^n)$. For example, a vector field is a smooth section of $T_0^1 \mathcal{M}^n = T\mathcal{M}^n$, while a covector field is a smooth section of $T_1^0 \mathcal{M}^n = T^*\mathcal{M}^n$. If $\mathcal{U} \subseteq \mathcal{M}^n$ is an open set and $\varphi : \mathcal{U} \to \mathbb{R}^n$ is a chart inducing a system of local coordinates $(x^1, \ldots, x^n)$ on $\mathcal{U}$, we define the **component representation** $\left(A_{j_1 \cdots j_p}^{k_1 \cdots k_q}\right)$ of A with respect to the chart $\varphi : \mathcal{U} \to \mathbb{R}^n$ by

$$A = A_{j_1 \cdots j_p}^{k_1 \cdots k_q} \, dx^{j_1} \otimes \cdots \otimes dx^{j_p} \otimes \frac{\partial}{\partial x^{k_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{k_q}}.$$

Equivalently,

$$A_{j_1 \cdots j_p}^{k_1 \cdots k_q} = A\left(\frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_p}}; \; dx^{k_1}, \ldots, dx^{k_q}\right).$$

Although there is no natural isomorphism (in the functorial sense) between the bundles $T\mathcal{M}^n$ and $T^*\mathcal{M}^n$, the metric

$$g = g_{ij}\, dx^i \otimes dx^j$$

and its inverse

$$g^{-1} = g^{ij}\, \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}$$

induce canonical isomorphisms $T_p^q \mathcal{M}^n \to T_{p'}^{q'} \mathcal{M}^n$ whenever $p+q = p'+q'$. In particular, the metric dual of a vector field $X \in C^\infty(T\mathcal{M}^n)$ is the covector field $X_\flat \in C^\infty(T^*\mathcal{M}^n)$ defined by

$$X_\flat = \left(X^i g_{ij}\right)\, dx^j;$$

and the metric dual of a covector field $\theta \in C^\infty(T^*\mathcal{M}^n)$ is the vector field $\theta^\sharp \in C^\infty(T\mathcal{M}^n)$ defined by

$$\theta^\sharp = \left(\theta_i g^{ij}\right)\, \frac{\partial}{\partial x^j}.$$

In formulas involving the components of familiar tensors in local coordinates, we will often use these isomorphisms without explicit mention, for instance writing R_i^k for $g^{jk} R_{ij}$ or $R^{k\ell}$ for $g^{ik} g^{j\ell} R_{ij}$.

2. First-order differential operators on tensors

Differentiation of functions on a smooth manifold is easy to define. If $X \in C^\infty(T\mathcal{M}^n)$ is a vector field and $f : \mathcal{M}^n \to \mathbb{R}$ is a differentiable function, then X acts on f to produce the function $X(f)$; in local coordinates

$$X(f) = X^i \frac{\partial f}{\partial x^i}.$$

(Here and everywhere in this book, the Einstein summation convention is in effect.)

To differentiate tensors requires a **connection**. It is a standard fact in geometry that a connection in a vector bundle is equivalent to a definition of parallel transport, which is in turn equivalent to a definition of covariant differentiation. Thus in the present context, we may in particular identify the Levi-Civita connection Γ of g with the **covariant derivative**, which is the unique operator

$$\nabla : C^\infty(T\mathcal{M}^n) \to C^\infty(T\mathcal{M}^n \otimes T^*\mathcal{M}^n)$$

with the property that

$$X(g(Y,Z)) \equiv X(\langle Y, Z\rangle) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

for all vector fields $X, Y, Z \in C^\infty(T\mathcal{M}^n)$, where

$$\nabla_Y X = (\nabla X)(Y) \in C^\infty(T\mathcal{M}^n).$$

In terms of the Christoffel symbols $\left(\Gamma_{ij}^k\right)$ induced by the Levi-Civita connection in the chart $\varphi : \mathcal{U} \to \mathbb{R}^n$, the covariant derivative of a vector field X is the $(1,1)$-tensor with components

$$\nabla_i X^k = (\nabla X)\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial}{\partial x^i} X^k + \Gamma_{ij}^k X^j.$$

In particular, it is customary in local coordinates to denote the operator $\nabla_{\frac{\partial}{\partial x^i}}$ simply by ∇_i.

The Levi-Civita connection also defines a covariant derivative ∇^* for covectors. This is the unique operator

$$\nabla^* : C^\infty\left(T^*\mathcal{M}^n\right) \to C^\infty\left(T^*\mathcal{M}^n \otimes T^*\mathcal{M}^n\right)$$

with the property that

$$X\left(\theta\left(Y\right)\right) = \left(\nabla_X^* \theta\right)\left(Y\right) + \theta\left(\nabla_X Y\right)$$

for all $X, Y \in C^\infty\left(T\mathcal{M}^n\right)$ and $\theta \in C^\infty\left(T^*\mathcal{M}^n\right)$, where

$$\nabla_X^* \theta = \left(\nabla^* \theta\right)\left(X\right) \in C^\infty\left(T^*\mathcal{M}^n\right).$$

More generally, the Levi-Civita connection defines a covariant derivative $\nabla^{(p,q)}$ on each tensor bundle $T_p^q\mathcal{M}^n$. This is the first-order differential operator

$$\nabla^{(p,q)} : C^\infty\left(T_p^q\mathcal{M}^n\right) \to C^\infty\left(T_{p+1}^q\mathcal{M}^n\right)$$

defined by the requirement that

$$\begin{aligned} X\left(A\left(Y_1, \ldots, Y_p;\ \theta_1, \ldots, \theta_q\right)\right) = {} & \left(\nabla_X^{(p,q)} A\right)\left(Y_1, \ldots, Y_p;\ \theta_1, \ldots, \theta_q\right) \\ & + \sum_{i=1}^p A\left(Y_1, \ldots, \nabla_X Y_i, \ldots, Y_p;\ \theta_1, \ldots, \theta_q\right) \\ & + \sum_{j=1}^q A\left(Y_1, \ldots, Y_p;\ \theta_1, \ldots, \nabla_X^* \theta_j, \ldots, \theta_q\right) \end{aligned}$$

for all (p,q)-tensors A, all vector fields $Y_1, \ldots, Y_p$, and all covector fields $\theta_1, \ldots, \theta_q$, where

$$\left(\nabla_X^{(p,q)} A\right)\left(Y_1, \ldots, Y_p;\ \theta_1, \ldots, \theta_q\right) = \left(\nabla^{(p,q)} A\right)\left(X, Y_1, \ldots, Y_p;\ \theta_1, \ldots, \theta_q\right) \in \mathbb{R}.$$

Note that $\nabla^{(0,1)} = \nabla$ and $\nabla^{(1,0)} = \nabla^*$. The component representation of the covariant derivative is defined by

$$\nabla_i^{(p,q)} A_{j_1 \cdots j_p}^{k_1 \cdots k_q} \doteqdot (\nabla A)\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_p}};\ dx^{k_1}, \ldots, dx^{k_q}\right).$$

REMARK A.1. It is customary and harmless to abuse notation and denote all $\nabla^{(p,q)}$ simply by ∇. We adopt this convention throughout this volume.

The **Lie derivative** is defined on vector fields so that

$$\mathcal{L}_X Y = [X, Y]$$

for all $X, Y \in C^\infty(T\mathcal{M}^n)$, where $[X, Y]$ is the unique vector field such that

$$[X, Y](f) = X(Y(f)) - Y(X(f))$$

for all differentiable functions $f : \mathcal{M}^n \to \mathbb{R}$. If θ is a covector field and X, Y are vector fields, the Lie derivative of θ is given by

$$(\mathcal{L}_X \theta)(Y) = X(\theta(Y)) - \theta([X, Y]).$$

Note that

$$(\mathcal{L}_X \theta)(Y) = d\theta(X, Y) + Y(\theta(X)),$$

where d is the exterior derivative defined in Section 3 below. Note too that even though the Lie derivative is independent of any Riemannian metric g on $\mathcal{M}^n$, the Lie derivative of a vector field can be calculated with respect to the covariant derivative ∇ induced by g by the formula

$$\mathcal{L}_X Y = [X, Y] = \nabla_X Y - \nabla_Y X. \tag{A.1}$$

Similarly, the Lie derivative of a covector field may by calculated by

$$(\mathcal{L}_X \theta)(Y) = (\nabla_X \theta)(Y) + \theta(\nabla_Y X).$$

More generally, the Lie derivative of a (p, q)-tensor field A satisfies

$$(\mathcal{L}_X A)(Y_1, \ldots, Y_p;\ \theta_1, \ldots, \theta_q) \tag{A.2a}$$
$$= X(A(Y_1, \ldots, Y_p;\ \theta_1, \ldots, \theta_q)) \tag{A.2b}$$
$$- \textstyle\sum_{1 \le i \le p} A(Y_1, \ldots, [X, Y_i], \ldots, Y_p;\ \theta_1, \ldots, \theta_q) \tag{A.2c}$$
$$- \textstyle\sum_{1 \le j \le q} A(Y_1, \ldots, Y_p;\ \theta_1, \ldots, \mathcal{L}_X \theta_j, \ldots, \theta_q) \tag{A.2d}$$

for all vector fields X and $Y_1, \ldots, Y_p$, and all covector fields $\theta_1, \ldots, \theta_q$. A specific case of some interest for us is the observation that the Lie derivative of the metric itself satisfies

$$(\mathcal{L}_X g)(Y, Z) = g(\nabla_Y X, Z) + g(Y, \nabla_Z X)$$

for all vector fields X, Y, Z. Hence in local coordinates, $\mathcal{L}_X g$ is just the symmetrized first covariant derivative of X, namely

$$(\mathcal{L}_X g)_{ij} = (\mathcal{L}_X g)\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \nabla_i X_j + \nabla_j X_i.$$

A first-order differential operator of particular interest in Chapter 3 is the **divergence**

$$\delta : C^\infty(T^*\mathcal{M}^n \otimes_S T^*\mathcal{M}^n) \to C^\infty(T^*\mathcal{M}^n)$$

defined for all symmetric $(2, 0)$-tensors A and all vector fields X by

$$(\delta A)(X) = -(\operatorname{div} A)(X) = \sum_{i=1}^{n} (\nabla A)(e_i, e_i, X),$$

where $\{e_i\}_{i=1}^n$ is a (local) orthonormal frame field. If $\mathcal{M}^n$ is compact, the formal adjoint of δ with respect to the L^2 inner product

$$(\cdot,\cdot) \doteqdot \int_{\mathcal{M}^n} \langle \cdot,\cdot \rangle \, d\mu$$

induced by the metric g is denoted by

$$\delta^* : C^\infty(T^*\mathcal{M}^n) \to C^\infty(T^*\mathcal{M}^n \otimes_S T^*\mathcal{M}^n).$$

By integrating by parts, it is easy to verify that

$$(\delta^*\theta) = \frac{1}{2}\mathcal{L}_{\theta^\sharp} g$$

for all covector fields θ, where the Lie derivative $\mathcal{L}$ and metric dual $\theta^\sharp$ are defined above. Notice that this formula can be used to define δ^* even when $\mathcal{M}^n$ is not compact.

3. First-order differential operators on forms

A p-form is a smooth section of the bundle $\wedge^p(T^*\mathcal{M}^n)$, namely an element of

$$\Omega^p(\mathcal{M}^n) \doteqdot C^\infty(\wedge^p(T^*\mathcal{M}^n)).$$

The **exterior derivative** is the family of operators

$$d \equiv d_p : \Omega^p(\mathcal{M}^n) \to \Omega^{p+1}(\mathcal{M}^n)$$

defined for all p-forms θ and vector fields $Y_1, \ldots, Y_p$ by

$$\begin{aligned} &d\theta(Y_0, Y_1, \ldots, Y_p) \\ &\doteqdot \textstyle\sum_{0\le i\le p} (-1)^i \theta\left(Y_0, Y_1, \ldots, \hat{Y}_i, \ldots, Y_p\right) \\ &+ \textstyle\sum_{0\le i<j\le p} (-1)^{i+j} \theta\left([Y_i, Y_j], Y_0, Y_1, \ldots, \hat{Y}_i, \ldots, \hat{Y}_j, \ldots, Y_p\right), \end{aligned}$$

where the symbol $\hat{Y}_i$ means that term is omitted. Although independent of any Riemannian metric g on $\mathcal{M}^n$, the exterior derivative may be expressed with respect to the covariant derivative ∇ defined by the Levi-Civita connection of that metric by

$$d\theta(Y_0, Y_1, \ldots, Y_p) = \textstyle\sum_{i=0}^p (-1)^i (\nabla_{Y_i}\theta)\left(Y_0, Y_1, \ldots, \hat{Y}_i, \ldots, Y_p\right).$$

If $\mathcal{M}^n$ is compact, the formal adjoint of d with respect to the L^2 inner product $(\cdot,\cdot)$ induced by the metric g is the family of operators

$$\delta \equiv \delta_{p+1} : \Omega^{p+1}(\mathcal{M}^n) \to \Omega^p(\mathcal{M}^n)$$

defined by the requirement that

$$(d\theta, \eta) = (\theta, \delta\eta)$$

for all $\theta \in \Omega^p(\mathcal{M}^n)$ and $\eta \in \Omega^{p+1}$. By integrating by parts, it is easy to verify that

$$(\delta\eta)(X_1, \ldots, X_p) = -\sum_{i=1}^{n} (\nabla\eta)(e_i, e_i, X_1, \ldots, X_p)$$

for all vector fields $X_1, \ldots, X_p$, where $\{e_i\}_{i=1}^n$ is a (local) orthonormal frame field. Note that this formula can be used to define δ even when $\mathcal{M}^n$ is not compact.

4. Second-order differential operators

A number of second-order differential operators are of considerable importance to our study of the Ricci flow.

The **rough Laplacian** denotes the family of operators

$$\Delta : C^\infty(T_p^q\mathcal{M}^n) \to C^\infty(T_p^q\mathcal{M}^n)$$

defined by

$$(\Delta A)(Y_1, \ldots, Y_p;\ \theta_1, \ldots, \theta_q) = \sum_{i=1}^{n} (\nabla^2 A)(e_i, e_i, Y_1, \ldots, Y_p;\ \theta_1, \ldots, \theta_q)$$

for all (p,q)-tensors A, all vector fields $Y_1, \ldots, Y_p$, and all covector fields $\theta_1, \ldots, \theta_q$, where $\{e_i\}_{i=1}^n$ is a (local) orthonormal frame field. In local coordinates, we write

$$\Delta A_{j_1 \cdots j_p}^{k_1 \cdots k_q} \doteqdot (\Delta A)\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_p}};\ dx^{k_1}, \ldots, dx^{k_q}\right).$$

The **Hodge–de Rham Laplacian** (frequently called the **Laplace–Beltrami operator**) denotes the family of maps

$$-\Delta_d : \Omega^p(\mathcal{M}^n) \to \Omega^p(\mathcal{M}^n)$$

defined by

$$-\Delta_d = d\delta + \delta\delta = d_{p-1}\delta_p + \delta_{p+1}d_p,$$

where d and δ are given in Section 3 above. It differs from the rough Laplacian by curvature terms which depend on the index p of the space $\Omega^p(\mathcal{M}^n)$. For instance, if θ is a 1-form (equivalently, a covector field), then the components of $\Delta_d\theta$ in local coordinates are given by the formula

$$\Delta_d\theta = \left(\Delta\theta_i - R_i^j\theta_j\right)\, dx^i.$$

If α is a 2-form, then

$$\text{(A.3)}\quad \Delta_d\alpha = \left(\Delta\alpha_{ij} + g^{kp}g^{\ell q}R_{ijk\ell}\alpha_{qp} - g^{k\ell}R_{ik}\alpha_{\ell j} - g^{k\ell}R_{jk}\alpha_{i\ell}\right)\, dx^i \wedge dx^j.$$

Finally, the **Lichnerowicz Laplacian**, introduced in [**93**], is an operator

$$\Delta_L : C^\infty(T^*\mathcal{M}^n \otimes_S T^*\mathcal{M}^n) \to C^\infty(T^*\mathcal{M}^n \otimes_S T^*\mathcal{M}^n)$$

defined so that whenever $(\mathcal{M}^n, g)$ is Ricci-parallel — that is, whenever the $(3,0)$-tensor $\nabla \operatorname{Rc}$ vanishes identically, as happens in particular if g is Einstein — the following diagram commutes:

$$\begin{array}{ccc} C^\infty\left(T^*\mathcal{M}^n \otimes_S T^*\mathcal{M}^n\right) & \xrightarrow{\Delta_L} & C^\infty\left(T^*\mathcal{M}^n \otimes_S T^*\mathcal{M}^n\right) \\ \Big\downarrow \delta & & \Big\uparrow \delta^* \\ C^\infty\left(T^*\mathcal{M}^n\right) \cong \Omega^1\left(\mathcal{M}^n\right) & \xrightarrow{\Delta_d} & \Omega^1\left(\mathcal{M}^n\right) \cong C^\infty\left(T^*\mathcal{M}^n\right). \end{array}$$

If A is a symmetric $(2,0)$-tensor, then the components of the Lichnerowicz Laplacian $\Delta_L A$ in local coordinates are given by the formula

(A.4)
$$\Delta_L A = \left(\Delta A_{ij} + 2g^{kp}g^{\ell q}R_{ik\ell j}A_{pq} - g^{k\ell}R_{ik}\,A_{\ell j} - g^{k\ell}R_{jk}\,A_{i\ell}\right)\,dx^i \otimes dx^j.$$

REMARK A.2. As we saw in Chapter 3, the Lichnerowicz Laplacian is essentially the linearization of the Ricci flow operator.

REMARK A.3. The Lichnerowicz Laplacian on symmetric 2-tensors is formally the same as the Hodge–de Rham Laplacian on 2-forms. Indeed, an easy application of the first Bianchi identity shows that equation (A.3) is equivalent to

$$\Delta_d \alpha = \left(\Delta \alpha_{ij} + 2g^{kp}g^{\ell q}R_{ik\ell j}\alpha_{pq} - g^{k\ell}R_{ik}\alpha_{\ell j} - g^{k\ell}R_{jk}\alpha_{i\ell}\right)\;dx^i \wedge dx^j,$$

which is formally the same as equation (A.4).

5. Notation for higher derivatives

A delicate notational issue arises when considering the operators

$$\nabla^k : C^\infty\left(T_p^q\mathcal{M}^n\right) \to C^\infty\left(T_{p+k}^q\mathcal{M}^n\right)$$

defined for $k > 1$. If $A \in C^\infty\left(T_p^q\mathcal{M}^n\right)$ and $X_1, \ldots, X_k$ are vector fields, we adopt the convention (which is not universally followed!) that

$$\nabla_{X_1}\nabla_{X_2}\cdots\nabla_{X_k}A$$

is the unique (p,q)-tensor field such that

$$\begin{aligned} &\left(\nabla_{X_1}\nabla_{X_2}\cdots\nabla_{X_k}A\right)\left(Y_1,\ldots,Y_p;\ \theta_1,\ldots,\theta_q\right) \\ &= \left(\nabla^k A\right)\left(X_1, X_2, \ldots, X_k,\ Y_1,\ldots,Y_p;\ \theta_1,\ldots,\theta_q\right) \end{aligned}$$

for all vector fields $Y_1, \ldots, Y_p$ and all covector fields $\theta_1, \ldots, \theta_q$. Note that

$$\left(\nabla^k A\right)\left(X_1,\ldots,X_k\right) \doteqdot \nabla_{X_1}\nabla_{X_2}\cdots\nabla_{X_k}A \neq \nabla_{X_1}\left(\nabla_{X_2}\left(\cdots\left(\nabla_{X_k}A\right)\right)\right)$$

in general. For example, in our convention,

(A.5) $$\nabla_X\left(\nabla_Y A\right) = \nabla_X\nabla_Y A + \nabla_{\nabla_X Y}A = \left(\nabla^2 A\right)(X,Y) + (\nabla A)\left(\nabla_X Y\right).$$

6. Commuting covariant derivatives

Recall that the $(3,1)$-Riemann curvature tensor is defined so that

$$R(X,Y)Z = \left(\nabla^2 Z\right)(X,Y) - \left(\nabla^2 Z\right)(Y,X) \in C^\infty(T\mathcal{M}^n) \tag{A.6}$$

for all vector fields X, Y, Z. Following the the notational convention of Section 5 above, we write this as

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z. \tag{A.7}$$

One advantage of this formula (hence of the convention adopted in Section 5) is that it makes the geometric significance of the Riemannian curvature clear by demonstrating its fundamental role in commuting covariant derivatives. By (A.1) and (A.5), one may also write equation (A.6) in the more familiar form

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

The reader is warned that this form is used *without parentheses* by those authors who do not follow the convention we adopted in Section 5. In local coordinates, all brackets $\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right]$ vanish identically, so that regardless of which convention one adopts, one can write unambiguously

$$R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} = R_{ijk}^{\ell}\frac{\partial}{\partial x^\ell} = (\nabla_i\nabla_j - \nabla_j\nabla_i)\left(\frac{\partial}{\partial x^k}\right).$$

The commutator formulas for covariant derivatives implied by equation (A.6) are collectively known as the **Ricci identities**. For example, if X is a vector field, then in local coordinates,

$$[\nabla_i, \nabla_j]X^\ell \equiv \nabla_i\nabla_j X^\ell - \nabla_j\nabla_i X^\ell = R_{ijk}^{\ell}X^k.$$

The commutator for a covector field θ (equivalently, a 1-form) may be obtained from this by using the metric dual $\theta^\sharp$, yielding

$$\begin{aligned}[\nabla_i, \nabla_j]\theta_k &\equiv \nabla_i\nabla_j\theta_k - \nabla_j\nabla_i\theta_k \\ &= g_{k\ell}[\nabla_i, \nabla_j]\left(\theta^\sharp\right)^\ell = g_{k\ell}R_{ijm}^{\ell}\left(g^{rm}\theta_r\right) = -R_{ijk}^{\ell}\theta_\ell.\end{aligned}$$

More generally, if A is any (p,q)-tensor field, one has the commutator

$$\begin{aligned}[\nabla_i, \nabla_j]A_{k_1\cdots k_p}^{\ell_1\cdots\ell_q} &\equiv \nabla_i\nabla_j A_{k_1\cdots k_p}^{\ell_1\cdots\ell_q} - \nabla_j\nabla_i A_{k_1\cdots k_p}^{\ell_1\cdots\ell_q} \\ &= \sum_{r=1}^{q} R_{ijm}^{\ell_r} A_{k_1\cdots k_p}^{\ell_1\cdots\ell_{r-1}\,m\,\ell_{r+1}\cdots\ell_q} - \sum_{s=1}^{p} R_{ijk_s}^{m} A_{k_1\cdots k_{s-1}\,m\,k_{s+1}\cdots k_p}^{\ell_1\cdots\ell_q}.\end{aligned}$$

Notes and commentary

The contents of this appendix are standard parts of classical Riemannian geometry. For the convenience of the reader, we have reviewed them here in order to establish an unambiguous notation.

APPENDIX B

Some results in comparison geometry

We begin this appendix by recalling basic results in comparison geometry that are useful in the study of limits of dilations about a singularity of the Ricci flow. Because these are standard results, we shall in some cases omit or merely sketch the proofs. We assume throughout that $(\mathcal{M}^n, g)$ is a complete Riemannian manifold.

1. Some results in local geometry

We begin with a collection of results that hold at small length scales, where a complete Riemannian manifold $(\mathcal{M}^n, g)$ is 'almost Euclidean'.

1.1. The Gauss Lemma and radial vector fields. Recall that for each point $p \in \mathcal{M}^n$ and tangent vector $V \in T_p\mathcal{M}^n$, there is a unique constant-speed geodesic

$$\gamma_V : [0, \infty) \to \mathcal{M}^n$$

emanating from p such that $\gamma_V(0) = p$ and $\dot{\gamma}_V(0) = V$. Given $p \in \mathcal{M}^n$, let $\exp_p : T_p\mathcal{M}^n \to \mathcal{M}^n$ denote the exponential map defined by

$$\exp_p(V) \doteqdot \gamma_V(1).$$

By the uniqueness of geodesics with given initial data, one has

$$\gamma_V(t) = \gamma_{tV}(1),$$

so that

$$\gamma_V(t) = \exp_p(tV).$$

Observe that $\left(\exp_p\right)_* : T_{\vec{0}}(T_p\mathcal{M}^n) \to T_p\mathcal{M}^n$ is the natural isomorphism between the two vector spaces. (We may in fact regard it as the identity map.) So by the inverse function theorem, there exists $\varepsilon > 0$ such that

$$\exp_p\big|_{B(\vec{0},\varepsilon)} : B(\vec{0}, \varepsilon) \to \exp_p\left(B(\vec{0}, \varepsilon)\right) \subset \mathcal{M}^n$$

is a diffeomorphism, where $B(\vec{V}, r)$ denotes the open ball with center $\vec{V} \in T_p\mathcal{M}^n$ and radius r. In fact, even more is true; $\exp_p$ preserves orthogonality to radial directions, and hence the length of radial vectors.

LEMMA B.1 (Gauss). *If $p \in \mathcal{M}^n$ and $V, W \in T_p\mathcal{M}^n$ are such that $\langle V, W\rangle = 0$, then*

$$\left\langle \left(\exp_p\right)_*(V_{tV}),\ \left(\exp_p\right)_*(W_{tV}) \right\rangle = 0.$$

Here $V_{tV}, W_{tV} \in T_{tV}(T_p\mathcal{M}^n)$ are the vectors identified with $V, W \in T_p\mathcal{M}^n$ respectively, and

$$(\exp_p)_* : T_{tV}(T_p\mathcal{M}^n) \to T_{\exp_p(tV)}\mathcal{M}^n.$$

Let ε be as above. The radial function

$$r : B(\vec{0}, \varepsilon) \to [0, \varepsilon)$$

given by

$$r(V) \doteqdot |V|$$

defines a function

$$f : \exp_p\left(B(\vec{0}, \varepsilon)\right) \to [0, \varepsilon)$$

where

$$f(x) \doteqdot r\left(\left(\exp_p\big|_{B(\vec{0},\varepsilon)}\right)^{-1}(x)\right).$$

The observation $f\left(\exp_p V\right) = |V|$ shows that f is just the distance function induced by g on $\exp_p\left(B(\vec{0}, \varepsilon)\right)$. The function f is smooth everywhere except at p, so that ∇f exists on $\exp_p\left(B(\vec{0}, \varepsilon)\right) \setminus \{p\}$.

On the other hand, there is a unit radial vector field $\vec{R}$ defined on $T_p\mathcal{M}^n \setminus \{\vec{0}\}$ by

$$\vec{R}_V \doteqdot \frac{V}{|V|}.$$

Define a vector field $\partial/\partial r$ on $\exp_p\left(B(\vec{0}, \varepsilon)\right) \setminus \{p\}$ by

$$\left(\frac{\partial}{\partial r}\right)_{\exp_p(V)} \doteqdot (\exp_p)_*\left(\vec{R}_V\right).$$

Since

$$(\exp_p)_*\left(\vec{R}_V\right) = \frac{1}{|V|}\frac{d}{dt}(\gamma_V(t))\bigg|_{t=1},$$

we have

$$\left|\frac{\partial}{\partial r}\right| \equiv 1.$$

The Gauss Lemma implies the following result.

COROLLARY B.2. *On $\exp_p\left(B(\vec{0}, \varepsilon)\right) \setminus \{p\}$, the vector fields ∇f and $\partial/\partial r$ are identical:*

$$\nabla f = \frac{\partial}{\partial r}.$$

PROOF. For every $V \in B(\vec{0}, \varepsilon) \setminus \{\vec{0}\}$ and every $Y \in T_{\exp_p(V)}\mathcal{M}^n$, there is a unique orthogonal decomposition

$$Y = a\frac{\partial}{\partial r} + Z$$

obtained by taking $a \doteqdot \langle Y, \partial/\partial r\rangle$ and $Z \doteqdot Y - a\,(\partial/\partial r)$. Then $\langle Z, \partial/\partial r\rangle = 0$ and

$$\langle Y, \nabla f\rangle = Y\,(f) = a\frac{\partial}{\partial r}\,(f) + Z\,(f)\,.$$

We claim that

$$\frac{\partial}{\partial r}\,(f) = 1 \qquad \text{and} \qquad Z\,(f) = 0.$$

The corollary will follow from these formulas, because they imply that

$$\langle Y, \nabla f\rangle = a \doteqdot \left\langle Y, \frac{\partial}{\partial r}\right\rangle$$

for all $Y \in T_{\exp_p(V)}\mathcal{M}^n$ and all $V \in B(\vec{0}, \varepsilon)\backslash\{\vec{0}\}$.

To first formula of the claim is proved by the observation

$$\frac{\partial}{\partial r}\,(f) = \left(\left(\exp_p\right)_*\left(\vec{R}\right)\right)\left(r \circ \left(\exp_p|_{B(\vec{0},\varepsilon)}\right)^{-1}\right) = \vec{R}\,(r) = 1.$$

To get the second formula, we note that

$$\begin{aligned} Z\,(f) &= Z\left(r \circ \left(\exp_p|_{B(\vec{0},\varepsilon)}\right)^{-1}\right) \\ &= \left[\left(\left(\exp_p|_{B(\vec{0},\varepsilon)}\right)^{-1}\right)_*(Z)\right](r) \\ &= \left\langle\left(\left(\exp_p|_{B(\vec{0},\varepsilon)}\right)^{-1}\right)_*(Z)\,,\ \vec{R}\right\rangle. \end{aligned}$$

But

$$\left\langle Z, \left(\exp_p\right)_*\left(\vec{R}\right)\right\rangle = \left\langle Z, \frac{\partial}{\partial r}\right\rangle = 0.$$

By the Gauss lemma, $\left(\exp_p\right)_*$ maps $\vec{R}^\perp$ to $(\partial/\partial r)^\perp$; since $(\exp_p|_{B(\vec{0},\varepsilon)})_*$ is an isomorphism, it follows that $((\exp_p|_{B(\vec{0},\varepsilon)})^{-1})_*$ maps $(\partial/\partial r)^\perp$ to $\vec{R}^\perp$, hence that

$$Z\,(f) = \left\langle\left(\left(\exp_p|_{B(\vec{0},\varepsilon)}\right)^{-1}\right)_*(Z)\,,\ \vec{R}\right\rangle = 0.$$

□

1.2. Conjugate points. If $\varphi : \mathcal{N}^n \to \mathcal{P}^n$ is a smooth map between differentiable manifolds of the same dimension, we say $\varphi_* : T_x\mathcal{N}^n \to T_{\varphi(x)}\mathcal{P}^n$ is **singular** if it is not an isomorphism. In this case we say that $x \in \mathcal{N}^n$ is a **critical point** of φ and that $\varphi\,(x)$ is a **critical value** of φ.

DEFINITION B.3. A point $q \in \mathcal{M}^n$ is a **conjugate point** of $p \in \mathcal{M}^n$ if q is a critical value of

$$\exp_p : T_p\mathcal{M}^n \to \mathcal{M}^n,$$

namely, if $q = \exp_p(V)$ for some $V \in T_p\mathcal{M}^n$ such that

$$\left(\exp_p\right)_* : T_V\left(T_p\mathcal{M}^n\right) \to T_{\exp_p(V)}\mathcal{M}^n$$

is singular. In this case, we say q is conjugate to p along the geodesic γ_V defined by $\gamma_V : t \longmapsto \exp_p(tV)$.

DEFINITION B.4. The **conjugate radius** $r_c \in (0,\infty]$ of a point $p \in \mathcal{M}^n$ is

$$r_c(p) \doteqdot \sup\left\{r : \left(\exp_p\right)_* \text{ is nonsingular in } B(\vec{0},r)\right\}.$$

REMARK B.5. $\exp_p|_{B(\vec{0},r_c)}$ is an immersion of the open ball $B(\vec{0},r_c)$.

Intuitively, a conjugate point is one where distinct geodesics come together. The following result (which uses terminology from Definition B.66) makes this notion precise.

LEMMA B.6. *Let $(\mathcal{M}^n, g)$ be a complete Riemannian manifold, and let $\{\Gamma_i\}_{i\in\mathbb{N}}$ be a sequence of nondegenerate proper geodesic 2-gons,*

$$\Gamma_i = \left\{\gamma_1^i : \left[0,\ell_1^i\right] \to \mathcal{M}^n,\ \gamma_2^i : \left[0,\ell_2^i\right] \to \mathcal{M}^n\right\},$$

such that $\gamma_1^i(0) = \gamma_2^i\left(\ell_2^i\right) \doteqdot p_i$ and $\gamma_2^i(0) = \gamma_1^i\left(\ell_1^i\right) \doteqdot q_i$. Suppose that the following limits all exist:

$$\begin{aligned}
p_\infty &\doteqdot \lim_{i\to\infty} p_i \in \mathcal{M}^n,\\
q_\infty &\doteqdot \lim_{i\to\infty} q_i \in \mathcal{M}^n,\\
V_\infty &\doteqdot \lim_{i\to\infty} \dot\gamma_1^i(0) = -\lim_{i\to\infty} \dot\gamma_2^i\left(\ell_2^i\right) \in T_{p_\infty}\mathcal{M}^n\\
\ell_\infty &\doteqdot \lim_{i\to\infty} \ell_1^i = \lim_{i\to\infty} \ell_2^i.
\end{aligned}$$

Then $\ell_\infty > 0$, and q_∞ is a conjugate point to p_∞ along the limit geodesic $\gamma_\infty : [0,\ell_\infty] \to \mathcal{M}^n$ determined by $\dot\gamma_\infty(0) = V_\infty$.

PROOF. That $\ell_\infty > 0$ follows from the lower bound for the injectivity radius of $(\mathcal{M}^n, g)$ in a compact set. Now suppose that $q_\infty = \exp_{p_\infty}(\ell_\infty V_\infty)$ is not a conjugate point of p_∞ along the geodesic $\gamma_\infty : [0,\ell_\infty] \to \mathcal{M}^n$. Then

$$\left(\exp_{p_\infty}\right)_* : T_{\ell_\infty V_\infty}\left(T_{p_\infty}\mathcal{M}^n\right) \to T_{q_\infty}\mathcal{M}^n$$

is nonsingular. Since $\exp : T\mathcal{M}^n \to \mathcal{M}^n$ is a smooth map, there exists a neighborhood $\mathcal{U}$ of p_∞ in $\mathcal{M}^n$ and a neighborhood $\mathcal{V}$ of $\ell_\infty V_\infty$ in $T_{p_\infty}\mathcal{M}^n$ such that

$$\left(\exp_p\right)_* : T_W\left(T_p\mathcal{M}^n\right) \to T_{\exp_p(W)}\mathcal{M}^n$$

is nonsingular for all $p \in \mathcal{U}$ and $W \in \mathcal{V} \subset T_p\mathcal{M}^n$. Here we used parallel translation along geodesics emanating from p_∞ to identify $T_{p_\infty}\mathcal{M}^n$ with $T_p\mathcal{M}^n$ for all $p \in \mathcal{U}$, which allows us to regard $\mathcal{V}$ as a subset of $T_p\mathcal{M}^n$. Thus it follows from the inverse function theorem that there exists a neighborhood $\mathcal{V}' \subseteq \mathcal{V}$ and some $\varepsilon > 0$ such that

$$\exp_p|_{B(\ell_\infty V_\infty,\varepsilon)} : B\left(\ell_\infty V_\infty, \varepsilon\right) \to \mathcal{M}^n$$

is injective for all $p \in \mathcal{V}'$. Now observe that for all sufficiently large i, we have

$$p_i \in \mathcal{V}',$$
$$\ell_1^i \dot{\gamma}_1^i (0) \in B (\ell_\infty V_\infty, \varepsilon) \subset T_{p_i} \mathcal{M}^n,$$
$$-\ell_2^i \dot{\gamma}_2^i (\ell_2^i) \in B (\ell_\infty V_\infty, \varepsilon) \subset T_{p_i} \mathcal{M}^n.$$

Since $\ell_1^i \dot{\gamma}_1^i (0) \neq -\ell_2^i \dot{\gamma}_2^i (\ell_2^i)$ but $\exp_{p_i} \left[\ell_1^i \dot{\gamma}_1^i (0)\right] = \exp_{p_i} \left[-\ell_2^i \dot{\gamma}_2^i (\ell_2^i)\right] = q_i$, this contradicts the injectivity of $\exp_p \big|_{B(\ell_\infty V_\infty, \varepsilon)}$. □

REMARK B.7. It is important to note that this result can fail in the context of sequences $(\mathcal{M}_i^n, g_i)$ of Riemannian manifolds — in particular, when each Γ_i is a nondegenerate proper geodesic 2-gon in $(\mathcal{M}_i^n, g_i)$. The failure can occur even if the sequence $(\mathcal{M}_i^n, g_i)$ has uniformly bounded geometry. This fact becomes highly relevant when one wants to develop injectivity radius estimates for such sequences in order to pass to a limit $(\mathcal{M}_\infty^n, g_\infty)$.

EXAMPLE B.8. Consider a sequence $\{\mathcal{T}_i^2\}_{i=\mathbb{Z}_+}$ of collapsing flat tori with fundamental domains

$$[-i, i] \times [-1/i, 1/i] \subset \mathbb{R}^2.$$

Take $O_i = (0,0)$, and define constant-speed geodesics

$$\alpha_i, \beta_i : \left[0, \sqrt{i^2 - 1}/i\right] \to \mathcal{T}_i^2$$

by

$$\alpha_i (s) = \left(s, \csc^{-1} (i) \cdot s\right) \qquad \text{and} \qquad \beta_i (s) = \left(s, -\csc^{-1} (i) \cdot s\right).$$

Then $\operatorname{length} \alpha_i = \operatorname{length} \beta_i = 1$ for all i, but their limit in the universal cover $\mathbb{R}^2$ is just the segment $s \mapsto (s, 0)$ for $0 \leq s \leq 1$, which has no conjugate points.

EXAMPLE B.9. Consider $\mathcal{S}_1^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$. For each $i \in \mathbb{N}$, let $(\mathcal{M}_i^2, g_i)$ denote $\mathcal{S}_1^2 / \mathbb{Z}_i$, where $\mathbb{Z}_i$ acts on $\mathcal{S}_1^2$ by rotation by angle $2\pi/i$ around a fixed axis, say the x-axis. The surface $\mathcal{M}_i^2$ is an orbifold with two cone points of order i. Simply for the sake of visualization, isometrically embed $\mathcal{M}_i^2$ into $\mathbb{R}^3$ as a surface of revolution about the x-axis centered at the origin. This embedding will be smooth except at the two cone points which are at opposite ends of the surface and lie on the x-axis. Since $\mathcal{M}_i^2$ is collapsing, its length along the x-axis tends to π. For i large enough, consider two points in $\mathcal{M}_i^2$ defined as follows. Let $p_i^1 = (x_i^1, y_i^1, z_i^1)$ be the unique point on $\mathcal{M}_i^2$ with $x_i^1 = -1$, $y_i^1 = 0$, and $z_i^1 < 0$. Let $p_i^2 = (x_i^2, y_i^2, z_i^2)$ be the unique point on $\mathcal{M}_i^2$ with $x_i^2 = +1$, $y_i^2 = 0$, and $z_i^1 > 0$. There are exactly two minimal geodesics α_i and β_i joining p_i^1 and p_i^2 in $\mathcal{M}_i^2$. These geodesics have the same length, and the angle between them at p_i^1 (and p_i^2) tends to 0 as $i \to \infty$. However, the limit

$$\gamma_\infty = \lim_{i \to \infty} \alpha_i = \lim_{i \to \infty} \beta_i$$

in the limit geodesic tube, which has constant curvature 1, is a geodesic of length 2. In particular, γ_∞ does not have any conjugate points along it.

1.3. Jacobi fields. We have already seen how distinct geodesics can come together at a conjugate point. The infinitesimal version of this idea is captured in the concept of a Jacobi field.

DEFINITION B.10. A **Jacobi field** is a variation vector field of a one-parameter family of geodesics.

LEMMA B.11. *$\exp_p(V)$ is conjugate to p if and only if there exists $W \in T_p\mathcal{M}^n\setminus\{0\}$ such that $\langle W, V\rangle = 0$ and $(\exp_p)_*(W_V) = \vec{0}$, where $W_V \in T_V(T_p\mathcal{M}^n)$.*

PROOF. Sufficiency is clear. To prove necessity, it is enough to observe that $\dot\gamma_V(1) \in T_{\exp_p(V)}\mathcal{M}^n\setminus\{\vec{0}\}$. □

PROPOSITION B.12. *Let $(\mathcal{M}^n, g)$ be complete and connected. Then q is conjugate to p along a geodesic γ if and only if there exists a nonzero Jacobi field J along γ with*

$$J(0) = J(1) = 0.$$

PROOF. By Hopf–Rinow, $q = \exp_p(V)$ for some $V \in T_p\mathcal{M}^n\setminus\{0\}$. Take any $W \in T_p\mathcal{M}^n\setminus\{0\}$ such that $\langle W, V\rangle = 0$. For s small, consider the one-parameter family of geodesics

$$\gamma_s : t \mapsto \exp_p[t(V + sW)].$$

Note that $\gamma \equiv \gamma_0$ is a geodesic such that $\gamma(0) = p$ and $\gamma(1) = q$. The corresponding Jacobi field is given by

$$\begin{aligned} J_W(t) &\doteqdot \left.\frac{\partial}{\partial s}\gamma_s(t)\right|_{s=0} \\ &= \left.\frac{\partial}{\partial s}\exp_p(tV + stW)\right|_{s=0} = (\exp_p)_*(tW_{tV}) \in T_{\gamma(t)}\mathcal{M}^n. \end{aligned}$$

Since $J_W(0) = 0$ and

$$J_W(1) = (\exp_p)_*(W_V),$$

the result follows from the lemma. □

By using the calculus of variations, one obtains the **Jacobi equation**.

LEMMA B.13. *J is a Jacobi field along a geodesic γ with unit tangent vector T if and only if*

$$\nabla_T\nabla_T J = R(T, J)T.$$

COROLLARY B.14. *Any Jacobi field J along a geodesic γ with unit tangent vector T admits the unique orthogonal decomposition*

$$J = J_0 + (at + b)T,$$

where $a, b \in \mathbb{R}$, and J_0 is a Jacobi field such that

$$\langle J_0, T\rangle \equiv 0.$$

PROOF. Since γ is a geodesic, we have $\nabla_T T = 0$ and hence

$$\left.\frac{d^2}{dt^2}\langle J, T\rangle\right|_{t=0} = \langle \nabla_T \nabla_T J, T\rangle = \langle R(T, J)\, T, T\rangle = 0.$$

□

1.4. Focal points and the normal bundle. Focal points are a natural generalization of the notion of conjugate points.

DEFINITION B.15. If $\iota : \Sigma^k \looparrowright \mathcal{M}^n$ is an immersion of a smooth manifold Σ^k, the **normal bundle** of Σ^k is the subbundle $N\Sigma^k$ of the pullback bundle $\iota^* T\mathcal{M}^n$ whose fiber over each $p \in \Sigma^k$ is

$$N_p\Sigma^k = \left\{V \in T_p\mathcal{M}^n : \langle V, W\rangle = 0 \text{ for all } W \in \iota_*\left(T_{\iota^{-1}(p)}\Sigma^k\right)\right\}.$$

Consider

$$\exp_p\big|_{N_p\Sigma^K} : N_p\Sigma^K \to \mathcal{M}^n.$$

Taking the union over all $p \in \Sigma^K$, one obtains a map of the total space

$$\exp|_{N\Sigma} : N\Sigma^k \to \mathcal{M}^n$$

defined for all pairs (p, V) with $p \in \Sigma^k$ and $V \in N_p\Sigma^k$ by

$$\exp|_{N\Sigma^k}(p, V) \doteqdot \exp_p\big|_{N_p\Sigma^k}(V).$$

Notice that $\exp|_{N\Sigma^k} : N\Sigma^k \to \mathcal{M}^n$ is a smooth map between manifolds of the same dimension.

EXAMPLE B.16. If $\Sigma^n = \mathcal{M}^n$, then $N\mathcal{M}^n = \mathcal{M}^n \times \{\vec{0}\}$ is naturally identified with $\mathcal{M}^n$, so that we may regard $\exp|_{N\mathcal{M}^n} : N\mathcal{M}^n \to \mathcal{M}^n$ as the identity map.

EXAMPLE B.17. If $\Sigma^0 = \{p\}$ is a point, then $T\{p\} = \{\vec{0}\}$, so that $N\{p\} = T_p\mathcal{M}^n$ and $\exp|_{N\{p\}} = \exp_p$.

DEFINITION B.18. Let $\iota : \Sigma^k \looparrowright \mathcal{M}^n$ be an immersion. One says $q \in \mathcal{M}^n$ is a **focal point** of $\iota\left(\Sigma^k\right)$ if q is a critical value of $\exp|_{N\Sigma^k}$. In particular, one says q is a focal point of Σ^k at p if it is the image of some critical point in $N_p\Sigma^k$.

Notice that $q \in \mathcal{M}^n$ is a focal point of $\Sigma^k \subseteq \mathcal{M}^n$ only if $q \notin \Sigma^k$.

EXAMPLE B.19. Let $\mathcal{M}^n = \mathbb{R}^n$ and let $\Sigma^{n-1} = S_r^{n-1}(p)$ be the $(n-1)$-sphere of radius $r > 0$ centered at $p \in \mathcal{M}^n$. Then the unique focal point of Σ^{n-1} is p.

1.5. The Rauch comparison theorem. If one has suitable bounds on the sectional curvatures of a Riemannian manifold $(\mathcal{M}^n, g)$, one can compare distances in $\mathcal{M}^n$ with a those in a convenient model space.

THEOREM B.20 (Rauch). *Let $(\mathcal{M}^n, g)$ and $\left(\overline{\mathcal{M}}^n, \bar{g}\right)$ be complete Riemannian manifolds. Let*

$$\gamma : [0, \ell] \to \mathcal{M}^n$$
$$\bar{\gamma} : [0, \ell] \to \overline{\mathcal{M}}^n$$

be unit-speed geodesics of the same length such that

$$\operatorname{sect}(\Pi) \leq \overline{\operatorname{sect}}\left(\bar{\Pi}\right)$$

for any 2-plane Π containing $T \doteqdot \dot{\gamma}$ and 2-plane $\bar{\Pi}$ containing $\bar{T} \doteqdot \dot{\bar{\gamma}}$. Let J and $\bar{J}$ be Jacobi fields along γ and $\bar{\gamma}$ respectively, such that $J(0)$ and $\bar{J}(0)$ are tangent to γ and $\bar{\gamma}$ respectively, and

$$|J(0)|_g = \left|\bar{J}(0)\right|_{\bar{g}}.$$

Suppose that

$$|\nabla_T J(0)|_g = \left|\bar{\nabla}_{\bar{T}} \bar{J}(0)\right|_{\bar{g}}$$

and

$$\langle \nabla_T J(0), T \rangle_g = \left\langle \bar{\nabla}_{\bar{T}} \bar{J}(0), \bar{T} \right\rangle_{\bar{g}}.$$

If $\bar{\gamma}(t)$ is not conjugate to $\bar{\gamma}(0)$ for any $t \in [0, \ell]$, then for all $t \in [0, \ell]$,

$$|J(t)|_g \geq \left|\bar{J}(t)\right|_{\bar{g}}.$$

Note that it is often convenient to apply the theorem when one knows that $\sup(\operatorname{sect}(\mathcal{M}^n, g)) \leq \inf\left(\operatorname{sect}\left(\overline{\mathcal{M}}^n, \bar{g}\right)\right)$ and takes Jacobi fields J, $\bar{J}$ such that $J(0) = \bar{J}(0) = 0$, $|\nabla_T J(0)|_g = \left|\bar{\nabla}_{\bar{T}} \bar{J}(0)\right|_{\bar{g}}$, and $\langle J, T \rangle_g \equiv \left\langle \bar{J}, \bar{T} \right\rangle_{\bar{g}} \equiv 0$. Note too that the condition $\langle J, T \rangle_g \equiv 0$ implies in particular that

$$0 \equiv \frac{d}{dt} \langle J, T \rangle_g = \langle \nabla_T J, T \rangle_g,$$

because γ is a geodesic.

By taking $\left(\overline{\mathcal{M}}^n, \bar{g}\right)$ to be the Euclidean sphere $\left(\mathcal{S}^n_{1/\sqrt{K}}, g_{\text{can}}\right)$ of radius $1/\sqrt{K}$, one obtains the following result.

COROLLARY B.21. *Let $K > 0$, and suppose $(\mathcal{M}^n, g)$ is a complete Riemannian manifold with $\operatorname{sect}(g) \leq K$. If J is a Jacobi field along a geodesic γ in $\mathcal{M}^n$ with unit tangent T such that*

$$J(0) = 0 \qquad \text{and} \qquad \langle \nabla_T J(0), T \rangle = 0,$$

then

$$|J(t)| \geq \frac{|\nabla_T J(0)|}{\sqrt{K}} \sin\left(\sqrt{K} t\right) > 0$$

for all $t \in \left[0, \pi/\sqrt{K}\right)$. In particular, there does not exist a conjugate point in $B\left(p, \pi/\sqrt{K}\right) = \left\{q \in \mathcal{M}^n : d(p, q) < \pi/\sqrt{K}\right\}$.

COROLLARY B.22. *Let $(\mathcal{M}^n, g)$ be a complete Riemannian manifold with sectional curvatures bounded above by $K > 0$. Let β be a geodesic path from p to q in $\mathcal{M}^n$ with $L(\beta) < \pi/\sqrt{K}$. Then for any points p' and q' sufficiently near p and q respectively, there is a unique geodesic path β' from p' to q' which is close to β.*

PROOF. By the previous corollary, the map

$$\exp_p\big|_{B(\vec{0},\pi/\sqrt{K})} : B(\vec{0},\pi/\sqrt{K}) \subset T_p\mathcal{M}^n \to \mathcal{M}^n$$

is an immersion. We may assume that the geodesic $\beta : [0,1] \to \mathcal{M}^n$ is parameterized so that $\beta(s) = \exp_p(sV)$ for some $V \in B(\vec{0},\pi/\sqrt{K})$ with $\exp_p(V) = q$. Since $\exp_p\big|_{B(\vec{0},\pi/\sqrt{K})}$ is a local diffeomorphism, there is for all q' sufficiently close to q a unique $V' \in B(\vec{0},\pi/\sqrt{K})$ such that $s \mapsto \exp_p(sV')$ is the unique geodesic path from p to q' lying close to β. Since for all $p' = \exp_p W$ sufficiently close to p, the map

$$\left(\exp_p\right)_* : T_p\mathcal{M}^n \cong T_W\left(T_p\mathcal{M}^n\right) \to T_{p'}\mathcal{M}^n$$

is an isomorphism, there exists a unique vector $V'' \in T_{p'}\mathcal{M}^n$ such that

$$\exp_{p'}\left(V''\right) = \exp_p\left(V'\right) = q'.$$

The path $\beta' : [0,1] \to \mathcal{M}^n$ defined by $\beta'(s) \doteqdot \exp_{p'}(sV'')$ is the unique geodesic near β joining p' to q'. $\square$

2. Distinguishing between local geometry and global geometry

In Section 1, we observed that the exponential map is always an embedding of a sufficiently small ball. We now study larger length scales, in order to investigate some of the ways that the topology of a complete Riemannian manifold $(\mathcal{M}^n, g)$ affects its geometry. In this section, interpret *geodesic* to mean *unit-speed geodesic.*

2.1. Cut points and the injectivity radius. Given $p \in \mathcal{M}^n$ and a geodesic $\gamma : [0,\infty) \to \mathcal{M}^n$ with $\gamma(0) = p$, define

$$S_\gamma \doteqdot \{t \in [0,\infty) : d(\gamma(0), \gamma(t)) = t\}.$$

A sufficiently short geodesic always minimizes distance. Its behavior at larger length scales is described by the following elementary observation.

LEMMA B.23. *Either $d(\gamma(0), \gamma(t)) = t$ for all $t \in [0,\infty)$, or else there exists a unique t_γ such that $d(\gamma(0), \gamma(t)) = t$ for all $t \in [0, t_\gamma]$ and $d(\gamma(0), \gamma(t)) < t$ for all $t > t_\gamma$. In short,*

$$\textit{either} \quad S_\gamma = [0,\infty) \quad \textit{or else} \quad S_\gamma = [0, t_\gamma].$$

PROOF. If $d(\gamma(0), \gamma(\bar{t})) < \bar{t}$ for some $\bar{t} > 0$, then for all $t \geq \bar{t}$,

$$d(\gamma(0), \gamma(t)) \leq d(\gamma(0), \gamma(\bar{t})) + d(\gamma(\bar{t}), \gamma(t)) < \bar{t} + (t - \bar{t}) = t.$$

Take $t_\gamma \doteqdot \inf\{t \geq 0 : d(\gamma(0), \gamma(t)) < t\}$. $\square$

DEFINITION B.24. If $\gamma : [0,\infty) \to \mathcal{M}^n$ is a geodesic with $\gamma(0) = p$, we call $\gamma(t_\gamma)$ the **cut point of p along γ.**

Notice that if $\mathcal{M}^n$ is compact, there is for each $p \in \mathcal{M}^n$ a unique cut point of p along every geodesic γ emanating from p.

DEFINITION B.25. The union

$$\mathrm{Cut}(p) \doteqdot \{\gamma(t_\gamma) : \gamma : [0,\infty) \to \mathcal{M}^n \text{ is a geodesic with } \gamma(0) = p\}$$

of all cut points along all geodesics emanating from $p \in \mathcal{M}^n$ is called the **cut locus** of p.

The existence of a conjugate point is sufficient (but not necessary) for a geodesic to fail to minimize distance.

LEMMA B.26. *Let $\gamma : [0,\ell] \to \mathcal{M}^n$ be a geodesic. If there is $\tau \in (0,\ell)$ such that $\gamma(\tau)$ is conjugate to $\gamma(0)$ along γ, then there exists a proper variation*

$$\{\Gamma_s(t) : -\varepsilon < s < \varepsilon,\ 0 \le t \le \ell\}$$

with $\Gamma_s(0) \equiv \gamma(0)$, $\Gamma_s(\ell) \equiv \gamma(\ell)$, and $\Gamma_0 = \gamma$ such that for some $s \neq 0$,

$$L(\Gamma_s) < L(\gamma).$$

Briefly: a geodesic fails to minimize distance past its first conjugate point.

COROLLARY B.27. *If $\gamma(t)$ is conjugate to $\gamma(0)$ along γ, then $t \ge t_\gamma$.*

If $\exp_p$ does not fail to be an immersion at the cut locus, it must at least fail to be an embedding there. This statement can be made more precise.

LEMMA B.28. *$\gamma(t)$ is a cut point of $\gamma(0)$ along the unit-speed geodesic γ if and only if $t > 0$ is the smallest value such that either of the following (non-exclusive) conditions hold:*

(1) *The point $\gamma(t)$ is conjugate to $\gamma(0)$ along γ.*
(2) *There exists a unit-speed geodesic $\beta : [0,t] \to \mathcal{M}^n$ such that $\beta(0) = \gamma(0)$, $\beta(t) = \gamma(t)$, and $L(\beta) = L\left(\gamma|_{[0,t]}\right)$, but $\dot\beta(0) \neq \dot\gamma(0)$. That is, β and $\gamma|_{[0,t]}$ are distinct geodesic paths of equal length from their common starting point $\gamma(0)$ to their common end point $\gamma(t)$.*

Given $p \in \mathcal{M}^n$, denote the unit sphere in $T_p\mathcal{M}^n$ by

$$\mathcal{S}_p^{n-1}\mathcal{M}^n \doteqdot \{V \in T_p\mathcal{M}^n : |V| = 1\},$$

and define $\rho : \mathcal{S}_p^{n-1}\mathcal{M}^n \to (0,\infty]$ by

$$\rho(V) \doteqdot \begin{cases} d(p,\gamma_V(t_{\gamma_V})) = t_{\gamma_V} & \text{if } t_{\gamma_V} \text{ exists} \\ \infty & \text{otherwise.} \end{cases}$$

That is, $\rho(V)$ is the distance from p to the cut point of p along γ_V.

LEMMA B.29. *For each $p \in \mathcal{M}^n$,*

$$\rho : \mathcal{S}_p^{n-1}\mathcal{M}^n \to (0,\infty]$$

is a continuous function. That is to say:

(1) $\mathcal{O} \doteqdot \{V \in \mathcal{S}_p^{n-1}\mathcal{M}^n : \rho(V) < \infty\}$ *is an open subset of* $\mathcal{S}_p^{n-1}\mathcal{M}^n$;
(2) $\rho|_{\mathcal{O}} : \mathcal{O} \to (0, \infty)$ *is continuous; and*
(3) *for any* $V \in \mathcal{S}_p^{n-1}\mathcal{M}_p^n \backslash \mathcal{O}$ *and every* $L \in (0, \infty)$, *there exists a neighborhood* $\mathcal{U}$ *of* V *in* $\mathcal{S}_p^{n-1}\mathcal{M}^n$ *such that* $\rho(U) > L$ *for all* $U \in \mathcal{U}$.

PROOF. The proof is in two steps.

We shall first show that if a geodesic emanating from $p \in \mathcal{M}^n$ stops minimizing at some point, then nearby geodesics emanating from p also stop minimizing near the same point. If $V \in \mathcal{O}$ and $\varepsilon > 0$, the definition of $\rho(V)$ implies that

$$d\left(p, \exp_p\left[(\rho(V) + \varepsilon) V\right]\right) < \rho(V) + \varepsilon.$$

Define $\delta \in (0, \infty)$ by

$$\delta \doteqdot \rho(V) + \varepsilon - d\left(p, \exp_p\left[(\rho(V) + \varepsilon) V\right]\right).$$

By continuity of the exponential map, there exists a neighborhood $\mathcal{W}_V$ of V in $\mathcal{S}_p^{n-1}\mathcal{M}^n$ such that for every $W \in \mathcal{W}_V$, one has

$$d\left(\exp_p\left[(\rho(V) + \varepsilon) W\right], \exp_p\left[(\rho(V) + \varepsilon) V\right]\right) < \frac{\delta}{2}.$$

By the triangle inequality, we obtain

$$\left|d\left(p, \exp_p\left[(\rho(V) + \varepsilon) W\right]\right) - d\left(p, \exp_p\left[(\rho(V) + \varepsilon) V\right]\right)\right| < \frac{\delta}{2}$$

and thus

$$\begin{aligned} d\left(p, \exp_p\left[(\rho(V) + \varepsilon) W\right]\right) &< d\left(p, \exp_p\left[(\rho(V) + \varepsilon) V\right]\right) + \frac{\delta}{2} \\ &= \rho(V) + \varepsilon - \frac{\delta}{2} < \rho(V) + \varepsilon. \end{aligned}$$

We conclude that $\rho(W) < \rho(V) + \varepsilon$ for all $W \in \mathcal{W}_V$. This proves both that ρ is upper semicontinuous at V and that $\mathcal{W}_V \subseteq \mathcal{O}$, hence that $\mathcal{O}$ is open.

Now suppose that $\{V_i\}$ is a sequence from $\mathcal{S}^{n-1}\mathcal{M}_p^n$ such that $V_i \to V$ but $\rho(V_i) \nrightarrow \rho(V)$. We shall derive a contradiction and thus complete the proof of the lemma. By passing to a subsequence, we may assume there is ρ_∞ in the compact set $[0, \infty]$ such that $\rho(V_i) \to \rho_\infty \neq \rho(V)$. Since ρ is upper semicontinuous at V, we may in fact assume $\rho_\infty < \rho(V)$, which implies in particular that $\exp_p(\rho_\infty V)$ is not a cut point of p along the geodesic $\exp_p(tV)$. By passing to a further subsequence, we may by Lemma B.28 assume either that each $\exp_p(\rho(V_i) V_i)$ is a singular point of $(\exp_p)_*$ or else that there are $V_i' \in \mathcal{S}_p^{n-1}\mathcal{M}^n$ such that $V_i' \neq V_i$ for any i, but $\exp_p(\rho(V_i) V_i') = \exp_p(\rho(V_i) V_i)$. The first case is impossible, since it implies that $(\exp_p)_*$ is singular at $\rho_\infty V$. In the second case, observe that $\exp_p$ is an embedding of a sufficiently small neighborhood $\mathcal{V}$ of V in $T_p\mathcal{M}^n$. So for all large enough i, each V_i' lies outside $\mathcal{V} \cap \mathcal{S}_p^{n-1}\mathcal{M}^n$. By passing to a further subsequence, we may thus assume $V_i' \to V_\infty' \neq V$. By continuity of the exponential map, this implies that $\exp_p(\rho_\infty V_\infty') = \exp_p(\rho_\infty V)$ and

$$\begin{aligned} d\left(p, \exp_p\left(\rho_\infty V_\infty'\right)\right) &= \lim_{i\to\infty} d\left(p, \exp_p\left(\rho\left(V_i\right) V_i'\right)\right) \\ &= \lim_{i\to\infty} d\left(p, \exp_p\left(\rho\left(V_i\right) V_i\right)\right) = d\left(p, \exp_p\left(\rho_\infty V\right)\right), \end{aligned}$$

which contradicts the fact that $\exp_p(\rho_\infty V)$ is not a cut point of p along the geodesic $\exp_p(tV)$. □

COROLLARY B.30. $\mathrm{Cut}(p)$ *is a closed set for each* $p \in \mathcal{M}^n$.

Given $p \in \mathcal{M}^n$, define

$$C_p \doteqdot \left\{V \in T_p\mathcal{M}^n : d\left(p, \exp_p(V)\right) = |V|\right\}.$$

Recalling that $\gamma_V(t) = \exp_p(tV)$, we observe that

$$\begin{aligned} C_p &= \left\{tV : V \in \mathcal{S}_p^{n-1}\mathcal{M}^n \text{ and } d\left(p, \gamma_V(t)\right) = t\right\} \\ &= \left\{tV : V \in \mathcal{S}_p^{n-1}\mathcal{M}^n \text{ and } t \leq d\left(p, \gamma_V\left(t_{\gamma_V}\right)\right)\right\}, \end{aligned}$$

hence conclude that C_p is closed.

DEFINITION B.31. If $p \in \mathcal{M}^n$, we call $\partial C_p \subset T_p\mathcal{M}^n$ the **cut locus in the tangent space** $T_p\mathcal{M}^n$.

Note that ∂C_p may be the empty set $\emptyset$. But in any case,

$$\mathrm{Cut}(p) = \exp_p(\partial C_p).$$

Moreover, $\mathcal{M}^n \backslash \mathrm{Cut}(p)$ is homeomorphic to $C_p \backslash \partial C_p$.

LEMMA B.32. *For each* $p \in \mathcal{M}^n$, *the map*

$$\exp_p\big|_{\mathrm{int}\, C_p} : C_p \backslash \partial C_p \to \mathcal{M}^n \backslash \mathrm{Cut}(p)$$

is an embedding.

Note that if $\mathcal{M}^n$ is compact, then $\mathrm{int}\, C_p$, hence $\mathcal{M}^n \backslash \mathrm{Cut}(p)$, is homeomorphic to an open n-ball.

DEFINITION B.33. The **injectivity radius** $\mathrm{inj}(p)$ at a point $p \in \mathcal{M}^n$ is

$$\mathrm{inj}(p) \doteqdot \sup\left\{r > 0 : \exp_p\big|_{B(\vec{0},r)} : B\left(\vec{0}, r\right) \to \mathcal{M}^n \text{ is an embedding}\right\}.$$

The **injectivity radius** $\mathrm{inj}(\mathcal{M}^n, g)$ is

$$\mathrm{inj}(\mathcal{M}^n, g) \doteqdot \inf\left\{\mathrm{inj}(p) : p \in \mathcal{M}^n\right\}.$$

When we want to take limits of a sequence of Riemannian manifolds, it is of fundamental importance to be able to estimate the injectivity radius from below. A basic estimate is the following.

LEMMA B.34 (Klingenberg). *If* $\mathcal{M}^n$ *is compact and* $\mathrm{sect}(g) \leq K$ *for some constant* $K > 0$, *then*

$$\mathrm{inj}(\mathcal{M}^n, g) \geq \min\left\{\frac{\pi}{\sqrt{K}}, \frac{1}{2}\left(\begin{array}{c}\textit{the length of the shortest} \\ \textit{smooth closed geodesic in } \mathcal{M}^n\end{array}\right)\right\}.$$

2.2. Lifting the metric by the exponential map. Let $r_c = r_c(p)$ be the conjugate radius of $p \in \mathcal{M}^n$. Then

$$\exp_p\big|_{B(\vec{0},r_c)} : B(\vec{0}, r_c) \to \mathcal{M}^n$$

is an immersion. Hence

$$\tilde{g} \doteqdot \left(\exp_p\big|_{B(\vec{0},r_c)}\right)^* g$$

is a smooth metric on $B(\vec{0}, r_c) \subset T_p\mathcal{M}^n$.

LEMMA B.35. *There are no cut points of $\vec{0}$ in the open Riemannian manifold*

$$\left(B(\vec{0}, r_c), \tilde{g}\right).$$

PROOF. Let $\tilde{d}$ denote distance in $B\left(\vec{0}, r_c\right)$ measured with respect to the metric $\tilde{g}$. It will suffice to show that $\tilde{d}\left(\vec{0}, V\right) = |V|$ for all $V \in B(\vec{0}, r_c)$. The unit speed geodesics of $\left(B(\vec{0}, r_c), \tilde{g}\right)$ emanating from $\vec{0}$ are the lines defined for $V \in S_p^{n-1}\mathcal{M}^n \subset T_p\mathcal{M}^n$ and $t \in [0, r_c)$ by $\tilde{\gamma}_V : t \longmapsto tV$. We have

$$L_{\tilde{g}}\left(\tilde{\gamma}_V\big|_{[0,1]}\right) = |V|,$$

and hence

$$\tilde{d}\left(\vec{0}, V\right) \leq |V|.$$

To see that equality holds, let $\omega : [0, a] \to B(\vec{0}, r_c)$ be any path with $\omega(0) = \vec{0}$ and $\omega(a) = V$. By Corollary B.2 of the Gauss lemma, we have

$$\nabla_{\tilde{g}} r = \left(\exp_p^{-1}\right)_* (\nabla_g f) = \left(\exp_p^{-1}\right)_* \left(\frac{\partial}{\partial r}\right) = \vec{R}.$$

Since $\left|\vec{R}\right|_{\tilde{g}} = 1$, it follows that

$$\begin{aligned}
L_{\tilde{g}}(\omega) &\doteqdot \int_0^a |\dot{\omega}(t)|_{\tilde{g}}\, dt \\
&\geq \int_0^a \left\langle \dot{\omega}(t), \vec{R} \right\rangle_{\tilde{g}} dt = \int_0^a \langle \dot{\omega}(t), \nabla_{\tilde{g}} r \rangle_{\tilde{g}}\, dt \\
&= \int_0^a \frac{d}{dt}\left[r(\omega(t))\right] dt = r(\omega(a)) - r(\omega(0)) = r(V) = |V|.
\end{aligned}$$

Hence

$$\tilde{d}\left(V, \vec{0}\right) \geq |V|.$$

□

Notice that the lemma implies

$$\operatorname{inj}_{\tilde{g}}\left(\vec{0}\right) = r_c.$$

Applying Corollary B.21, we obtain the following conclusion.

COROLLARY B.36. *If $(\mathcal{M}^n, g)$ is a complete Riemannian manifold with sectional curvatures bounded above by $K > 0$, then for any $p \in \mathcal{M}^n$,*

$$\operatorname{inj}_{\tilde{g}}(\vec{0}) \geq \frac{\pi}{\sqrt{K}}.$$

2.3. A Laplacian comparison theorem for distance functions. Given any smooth function f on $(\mathcal{M}^n, g)$, the **Bochner–Weitzenböck formula** says

$$\frac{1}{2}\Delta |\nabla f|^2 = |\nabla\nabla f|^2 + \langle \nabla f, \nabla (\Delta f)\rangle + \operatorname{Rc}(\nabla f, \nabla f).$$

This formula is of fundamental importance and is used in the proof of the Li–Yau gradient estimate [**91**] for the first eigenfunction and the Li–Yau differential Harnack inequality [**92**] for positive solutions of the heat equation, among many other results. It is easily proved using Ricci calculus:

$$\begin{aligned}\frac{1}{2}\Delta |\nabla f|^2 &= \frac{1}{2}\nabla_i\nabla^i\left(\nabla_j f \nabla^j f\right) = \nabla_i\left(\nabla_j\nabla^i f \nabla^j f\right)\\ &= \nabla_i\nabla_j\nabla^i f\nabla^j f + \nabla^i\nabla_j f\nabla_i\nabla^j f\\ &= \nabla_j\nabla_i\nabla^i f\nabla^j f + R_{jk}\nabla^j f\nabla^k f + \nabla_i\nabla_j f\nabla^i\nabla^j f\\ &= \langle\nabla\Delta f, \nabla f\rangle + \operatorname{Rc}(\nabla f, \nabla f) + |\nabla\nabla f|^2.\end{aligned}$$

We say $r : \mathcal{M}^n \to [0, \infty)$ is a **generalized distance function** if

$$|\nabla r|^2 = 1 \tag{B.1}$$

at all points where r is smooth. Generalized distance functions have the property that the integral curves of ∇r are geodesics.

LEMMA B.37. *If r is a generalized distance function, then wherever r is smooth,*

$$\nabla_{\nabla r}(\nabla r) \equiv 0.$$

PROOF. Differentiating (B.1) shows that for all $i = 1, \ldots, n$,

$$0 = \nabla_i |\nabla r|^2 = \nabla_i\left(\nabla_j r\nabla^j r\right) = 2\nabla^j r\nabla_j\nabla_i r = 2\left(\nabla_{\nabla r}\nabla r\right)_i.$$

□

COROLLARY B.38. *Wherever r is smooth,*

$$|\nabla\nabla r|^2 \geq \frac{1}{n-1}(\Delta r)^2.$$

PROOF. $\nabla\nabla r$ has a zero eigenvalue, because

$$(\nabla\nabla r)(\nabla r, \nabla r) = \nabla^i r\nabla^j r\nabla_i\nabla_j r = \langle\nabla r, \nabla_{\nabla r}\nabla r\rangle = 0.$$

Hence the standard estimate becomes

$$(\Delta r)^2 = (\operatorname{tr}_g\nabla\nabla r)^2 \leq (n-1)|\nabla\nabla r|^2.$$

□

Assume there is $B \in \mathbb{R}$ such that $\mathrm{Rc} \geq (n-1)\, Bg$. Wherever r is a smooth generalized distance function, taking $f = r$ in the Bochner–Weitzenböck formula gives the estimate

$$\begin{aligned} 0 &= |\nabla\nabla r|^2 + \langle \nabla r, \nabla(\Delta r)\rangle + \mathrm{Rc}(\nabla r, \nabla r) \\ &\geq \frac{1}{n-1}(\Delta r)^2 + (\nabla r)(\Delta r) + (n-1)\, B, \end{aligned} \tag{B.2}$$

since $\langle \nabla r, \nabla(\Delta r)\rangle = d(\Delta r)(\nabla r) = (\nabla r)(\Delta r)$. If we define

$$v \doteqdot \frac{1}{n-1}\Delta r,$$

then (B.2) is equivalent to the inequality

$$dv(r) = (\nabla r)(v) \leq -v^2 - B. \tag{B.3}$$

This suggests we compare $v(r)$ with the solutions of the Riccati equation

$$\frac{dw}{dr} = -w^2 - B, \qquad \lim_{r\to 0+} w = \infty,$$

namely

$$(B>0) \qquad w(r) = \sqrt{B}\cot\left(\sqrt{B}r\right)$$

$$(B=0) \qquad w(r) = \frac{1}{r}$$

$$(B<0) \qquad w(r) = \sqrt{-B}\coth\left(\sqrt{-B}r\right).$$

We can indeed make this heuristic rigorous in the following case.

PROPOSITION B.39. *Let $(\mathcal{M}^n, g)$ be a complete Riemannian manifold such that $\mathrm{Rc} \geq (n-1)\, Bg$ for some $B \in \mathbb{R}$. If $r(x) \doteqdot d(p,x)$ is the metric distance from some fixed $p \in \mathcal{M}^n$, then on a neighborhood of p contained in$\mathcal{M}^n \backslash (\{p\} \cup \mathrm{Cut}(p))$, we have the estimate*

$$\frac{\Delta r}{n-1} \leq \begin{cases} \sqrt{B}\cot\left(\sqrt{B}r\right) & \text{if } B>0 \\ 1/r & \text{if } B=0 \\ \sqrt{-B}\coth\left(\sqrt{-B}r\right) & \text{if } B<0. \end{cases}$$

PROOF. Note that r is smooth on $\mathcal{M}^n \backslash (\{p\} \cup \mathrm{Cut}(p))$, and Δr satisfies

$$\lim_{r\to 0+} r\Delta r = n-1.$$

Thus the function defined by

$$u(x) \doteqdot \begin{cases} \frac{n-1}{\Delta r} & \text{if } x \neq p \\ 0 & \text{if } x = p \end{cases}$$

is continuous in a neighborhood $\mathcal{U} \subseteq \mathcal{M}^n \setminus \mathrm{Cut}(p)$ of p and is smooth on $\mathcal{U} \setminus \{p\}$. Let

$$\gamma : [0, \mathrm{inj}(p)) \to \mathcal{M}^n$$

be a unit-speed geodesic such that $\gamma(0) = p$ and $\dot\gamma = \nabla r$, and fix any $r \in (0, \mathrm{inj}(p))$. By (B.3), we obtain the differential inequality

$$1 \leq \frac{du(r)}{1 + Bu^2} = \frac{(\nabla r)(u)}{1 + Bu^2},$$

which implies

$$r \leq \int_0^r \frac{\dot\gamma(t)(u)}{1 + B \cdot u(\gamma(t))^2} \, dt$$

So if $B > 0$, we get

$$r \leq \int_0^r \frac{d}{dt}\left(\frac{1}{\sqrt{B}} \tan^{-1}\left(\sqrt{B} u(\gamma(t))\right)\right) dt = \frac{\tan^{-1}\left(\sqrt{B} u(\gamma(r))\right)}{\sqrt{B}};$$

while if $B = 0$, we have

$$r \leq \int_0^r \frac{du}{dt} \, dt = u(\gamma(r));$$

and finally if $B < 0$, we obtain

$$r \leq \int_0^r \frac{d}{dt}\left(\frac{1}{\sqrt{-B}} \tanh^{-1}\left(\sqrt{-B} u(\gamma(t))\right)\right) dt = \frac{\tanh^{-1}\left(\sqrt{-B} u(\gamma(r))\right)}{\sqrt{-B}}.$$

The result now follows easily from the monotonicity of $\tan^{-1}$ and $\tanh^{-1}$. $\square$

2.4. The Toponogov comparison theorem. The Rauch comparison theorem works at infinitesimal length scales to compare the geometry of a Riemannian manifold $(\mathcal{M}^n, g)$ with model geometries of constant curvature. It has a powerful analog at global length scales: the Toponogov comparison theorem.

THEOREM B.40 (Toponogov Comparison Theorem). *Let $(\mathcal{M}^n, g)$ be a complete Riemannian manifold with sectional curvatures bounded below by $H \in \mathbb{R}$.*

Triangle version (SSS): *Let Δ be a geodesic triangle with vertices (p, q, r), sides $\overline{qr}$, $\overline{rp}$, $\overline{pq}$ of lengths*

$$a = \mathrm{length}(\overline{qr}), \; b = \mathrm{length}(\overline{rp}), \; c = \mathrm{length}(\overline{pq})$$

satisfying $a \leq b + c$, $b \leq a + c$, $c \leq a + b$ (for example, when all of the geodesic sides are minimal), and interior angles $\alpha = \measuredangle rpq$, $\beta = \measuredangle pqr$, $\gamma = \measuredangle qrp$, where $\alpha, \beta, \gamma \in [0, \pi]$. Assume that $c \leq \pi/\sqrt{H}$ if $H > 0$. (No assumption on c is needed if $H \leq 0$.) If the geodesics $\overline{qr}$ and $\overline{rp}$ are minimal, then there exists a geodesic triangle $\bar\Delta = (\bar p, \bar q, \bar r)$ in the complete simply-connected space *of constant sectional curvature H with the same side lengths (a,b,c) such that*

$$\alpha \geq \bar{\alpha} \doteqdot \measuredangle \bar{r}\bar{p}\bar{q}$$
$$\beta \geq \bar{\beta} \doteqdot \measuredangle \bar{p}\bar{q}\bar{r}.$$

Hinge version (SAS): *Let $\angle$ be a geodesic hinge with vertices (p,q,r), sides $\overline{qr}$ and $\overline{rp}$, and interior angle $\measuredangle qrp \in [0,\pi]$ in $\mathcal{M}^n$. Suppose that $\overline{qr}$ is minimal and that* $\operatorname{length}(\overline{rp}) \leq \pi/\sqrt{H}$ *if $H > 0$. Let $\angle'$ be a geodesic hinge with vertices (p',q',r') in the complete simply-connected space of constant sectional curvature H with the same side lengths and same angle. Then one may compare the distances between the endpoints of the hinges as follows:*

$$\operatorname{dist}(p,q) \leq \operatorname{dist}(p',q').$$

3. Busemann functions

In this section, we develop some tools to study a complete noncompact Riemannian manifold $(\mathcal{M}^n, g)$ at very large length scales, in order to understand its geometry 'at infinity'.

3.1. Definition and basic properties.

DEFINITION B.41. A unit speed geodesic $\gamma : [0,\infty) \longrightarrow \mathcal{M}^n$ is a **ray** if each segment is minimal, namely, if $\gamma|_{[a,b]}$ is minimal for all $0 \leq a < b < \infty$. We say γ is a **ray emanating from** $O \in \mathcal{M}$ if γ is a ray with $\gamma(0) = O$.

LEMMA B.42. *For any point $O \in \mathcal{M}^n$, there exists a ray γ emanating from O.*

PROOF. Choose any sequence of points $p_i \in \mathcal{M}^n$ such that $d(O,p_i) \nearrow \infty$. For each i, choose a minimal geodesic segment γ_i joining O and p_i. Let

$$V_i = d\gamma_i/dt\,(0) \in T_O\mathcal{M}^n.$$

Since the unit sphere in $T_O\mathcal{M}^n$ is compact and $|V_i| \equiv 1$, there exists a subsequence such that the limit

$$V \doteqdot \lim_{i\to\infty} V_i \tag{B.4}$$

exists. Let $\gamma : [0,\infty) \to \mathcal{M}^n$ be the unique unit-speed geodesic with $\gamma(0) = O$ and $d\gamma/dt\,(0) = V$. Recalling that the solution of an ODE is a continuous function of its initial data, we note that (B.4) implies the images $\gamma_i \to \gamma$ uniformly on compact subsets of $\mathcal{M}^n$. Because each segment of every γ_i is minimal, it follows that each segment of γ is minimal. □

DEFINITION B.43. If γ is a ray emanating from $O \in \mathcal{M}^n$, the **pre-Busemann function** $b_{\gamma,s} : \mathcal{M}^n \to \mathbb{R}$ associated to the ray γ and $s \in [0,\infty)$ is defined by

$$b_{\gamma,s}(x) \doteqdot d(O,\gamma(s)) - d(\gamma(s),x) = s - d(\gamma(s),x).$$

LEMMA B.44. *The functions $b_{\gamma,s}$ are uniformly bounded: for all $x \in \mathcal{M}^n$ and $s \geq 0$,*

$$|b_{\gamma,s}(x)| \leq d(x, O).$$

The functions $b_{\gamma,s}$ are uniformly Lipschitz with Lipschitz constant 1: for all $x, y \in \mathcal{M}^n$ and $s \geq 0$,

$$|b_{\gamma,s}(x) - b_{\gamma,s}(y)| \leq d(x, y).$$

And the functions $b_{\gamma,s}(x)$ are monotone increasing in s for each $x \in \mathcal{M}^n$: if $s < t$, then

$$b_{\gamma,s}(x) \leq b_{\gamma,t}(x).$$

PROOF. All three statements are consequences of the triangle inequality. The first is immediate, and the second follows from observation

$$|b_{\gamma,s}(x) - b_{\gamma,s}(y)| = |d(y, \gamma(s)) - d(x, \gamma(s))| \leq d(x, y).$$

To prove the third, we note that $d(\gamma(s), \gamma(t)) = t - s$ and observe that

$$\begin{aligned} b_{\gamma,t}(x) &\doteqdot t - d(x, \gamma(t)) \\ &\geq s + (t - s) - d(x, \gamma(s)) - d(\gamma(s), \gamma(t)) \\ &= s - d(x, \gamma(s)) \doteqdot b_{\gamma,s}(x). \end{aligned}$$

□

The monotonicity of the pre-Busemann functions in the parameter s enables us to make the following

DEFINITION B.45. The **Busemann function** $b_\gamma : \mathcal{M}^n \to \mathbb{R}$ **associated to the ray** γ is

$$b_\gamma(x) \doteqdot \lim_{s \to \infty} b_{\gamma,s}(x).$$

Since the family $\{b_{\gamma,s}\}$ is uniformly Lipschitz and uniformly bounded above, we can immediately make the following observations.

LEMMA B.46. *The Busemann function b_γ associated to a ray γ emanating from $O \in \mathcal{M}^n$ is bounded above: for all $x \in \mathcal{M}^n$,*

$$|b_\gamma(x)| \leq d(x, O). \tag{B.5}$$

And b_γ is uniformly Lipschitz with Lipschitz constant 1: for all $x, y \in \mathcal{M}^n$,

$$|b_\gamma(x) - b_\gamma(y)| \leq d(x, y). \tag{B.6}$$

Intuitively, $b_\gamma(x)$ measures how far out toward infinity x is in the direction of γ. One could also regard b_γ as a renormalized distance function from what one might think of as the 'point' $\gamma(\infty)$. For example, in Euclidean space, the Busemann functions are the affine projections. In particular, if

$\gamma : [0,\infty) \to \mathbb{R}^n$ is a ray with $\gamma(0) = O$ and $d\gamma/dt(0) = V$, then the associated Busemann function is easily computed using the law of cosines:

$$\begin{aligned} b_\gamma(x) &= \lim_{s\to\infty} \left(d(O, O+sV) - d(O+sV, x) \right) \\ &= \lim_{s\to\infty} \left(s - \sqrt{s^2 + |x-O|^2 - 2s\langle x - O, V\rangle} \right) \\ &= \langle x - O, V\rangle . \end{aligned}$$

DEFINITION B.47. The **Busemann function $\bar{b} : \mathcal{M}^n \to \mathbb{R}$ associated to the point $O \in \mathcal{M}^n$** is

$$\bar{b} \doteqdot \sup_\gamma b_\gamma,$$

where the supremum is taken over all rays γ emanating from O.

This definition enables us to associated a Busemann function to a point. In Euclidean space, the Busemann function associated to a point $O \in \mathbb{R}^n$ is just the distance function, because for all $x \in \mathbb{R}^n$,

$$\bar{b}(x) = \left\langle x - O, \frac{x-O}{|x-O|} \right\rangle = |x - O| = d(x, O).$$

LEMMA B.48. *The Busemann function $\bar{b}$ associated to $O \in \mathcal{M}^n$ is bounded above: for all $x \in \mathcal{M}^n$,*

$$\left|\bar{b}(x)\right| \le d(x, O).$$

And $\bar{b}$ is uniformly Lipschitz with Lipschitz constant 1: for all $x, y \in \mathcal{M}^n$,

$$\left|\bar{b}(x) - \bar{b}(y)\right| \le d(x, y).$$

PROOF. The first statement follows from (B.5). To prove the second, note that for any $x, y \in \mathcal{M}^n$ and $\varepsilon > 0$, there exists a ray γ emanating from O such that

$$\bar{b}(x) - \bar{b}(y) \le b_\gamma(x) + \varepsilon - \bar{b}(y) \le b_\gamma(x) + \varepsilon - b_\gamma(y).$$

Hence by (B.6),

$$\bar{b}(x) - \bar{b}(y) \le d(x, y) + \varepsilon.$$

□

The following result says that any sufficiently long minimal geodesic segment emanating from a point O can be well approximated by a ray emanating from O.

LEMMA B.49. *Given $O \in \mathcal{M}^n$, define $\theta : [0,\infty) \to [0,\pi]$ by*

$$\theta(r) = \sup_{\sigma\in\mathcal{S}(r)} \inf_{\rho\in\mathcal{R}} \measuredangle_O (\dot\sigma(0), \dot\rho(0)),$$

where $\mathcal{S}(r)$ is the set of all minimal geodesic segments σ of length $L(\sigma) \ge r$ emanating from O, and $\mathcal{R}$ is the set of rays emanating from O. Then

$$\lim_{r\to\infty} \theta(r) = 0.$$

PROOF. If not, there exists $\varepsilon > 0$, a sequence of points $p_i \in \mathcal{M}^n$ with $d(p_i, O) \nearrow \infty$, and minimal geodesic segments σ_i joining O and p_i such that

$$\measuredangle_O(\dot{\sigma}_i(0), \dot{\rho}(0)) \geq \varepsilon$$

for each i and all rays ρ emanating from O. By compactness of the unit sphere in $T_O\mathcal{M}^n$, there exists a subsequence such that $\lim_{i\to\infty} \dot{\sigma}_i(0) \doteqdot V$ exists. Let $\sigma_\infty : [0, \infty) \to \mathcal{M}^n$ be the unique geodesic with $\sigma_\infty(0) = O$ and $\dot{\sigma}_\infty(0) = V$. Arguing as in Lemma B.42, we see that σ_∞ is a ray. In particular, the condition

$$\measuredangle_O(\dot{\sigma}_i(0), \dot{\sigma}(0)) \geq \varepsilon$$

is impossible. □

Recall that we already have a good upper bound for the Busemann function associated to a point:

$$\bar{b}(x) \leq d(x, O).$$

The previous lemma lets us construct a lower bound for $\bar{b}$ in the event that $\mathcal{M}^n$ has nonnegative sectional curvature.

COROLLARY B.50. *If $(\mathcal{M}^n, g)$ is a complete noncompact Riemannian manifold of nonnegative sectional curvature, then*

$$\bar{b}(x) \geq d(x, O)(1 - \theta(d(x, O))).$$

PROOF. Given any point $x \in \mathcal{M}^n$, let α be a minimal geodesic segment joining O and x. By the lemma, there exists a ray γ emanating from O such that

$$\measuredangle(\dot{\gamma}(0), \dot{\alpha}(0)) \leq \theta(d(x, O)).$$

Set $y \doteqdot \gamma(d(x, O))$. By the Toponogov comparison theorem applied to the hinge with vertex O and sides α joining O to x and $\gamma|_{[0,d(x,O)]}$ joining O to y, we have

$$d(x, y) \leq \theta(d(x, O)) \cdot d(x, O),$$

where the right-hand side is the length of an arc of angle $\theta(d(x, O))$ in a circle in the Euclidean plane of radius $d(x, O)$. Thus since $b_\gamma(y) = d(x, O)$, we get

$$\bar{b}(x) \geq b_\gamma(x) \geq b_\gamma(y) - d(x, y) \geq d(x, O)[1 - \theta(d(x, O))],$$

where we used the fact that b_γ has Lipschitz constant 1 to obtain the second inequality. □

The gist of the this is that the Busemann function $\bar{b}$ associated to a point O is similar to the distance function to O. However, from some points of view, $\bar{b}$ has better properties. The property we shall be most interested in is convexity.

3.2. Constructing totally convex half spaces. An important use of Busemann functions is to define arbitrarily large 'totally convex subsets' in a complete noncompact manifold of nonnegative sectional curvature. As a first step in their construction, we make the following definitions.

DEFINITION B.51. Given a ray $\gamma : [0, \infty) \to \mathcal{M}^n$ emanating from a point $O \in \mathcal{M}^n$, the open **right half space** is the union

$$\mathbb{B}_\gamma \doteqdot \bigcup_{s \in (0,\infty)} B(\gamma(s), s),$$

where $B(\gamma(s), s) \doteqdot \{x \in \mathcal{M}^n : d(x, \gamma(s)) < s\}$. The closed **left half space** $\mathbb{H}_\gamma$ is its complement:

$$\mathbb{H}_\gamma \doteqdot \mathcal{M}^n \backslash \mathbb{B}_\gamma = \mathcal{M}^n \Big\backslash \bigcup_{s \in (0,\infty)} B(\gamma(s), s).$$

The motivation for the term 'left half space' is the following result.

LEMMA B.52. *$b_\gamma(x) \leq 0$ if and only if $x \in \mathbb{H}_\gamma$.*

PROOF. By Lemma B.44, $b_{\gamma,s}$ is monotone increasing in s. So $b_\gamma(x) \leq 0$ if and only if $b_{\gamma,s}(x) \leq 0$ for all $s \in (0, \infty)$. But by definition

$$b_{\gamma,s}(x) \leq 0 \quad \text{if and only if} \quad x \notin B(\gamma(s), s).$$

Hence $b_\gamma(x) \leq 0$ if and only if $x \notin \bigcup_{s \in (0,\infty)} B(\gamma(s), s)$. $\square$

DEFINITION B.53. A set $X \subseteq \mathcal{M}^n$ is **totally convex** if for every $x, y \in X$ and every minimizing geodesic α joining x and y, we have $\alpha \subseteq X$.

Our interest in half spaces is explained by the following result.

PROPOSITION B.54. *If γ is any ray in a complete noncompact Riemannian manifold $(\mathcal{M}^n, g)$ of nonnegative sectional curvature, then the closed left half-space $\mathbb{H}_\gamma$ is totally convex.*

We shall give two proofs of the proposition, which are quite different in character. The first uses the second variation formula, while the second uses the Toponogov comparison theorem.

PROOF OF PROPOSITION B.54 IN THE CASE $\operatorname{sect}(g) > 0$. If $\mathbb{H}_\gamma$ is not totally convex, there are $x, y \in \mathbb{H}_\gamma$ and a unit-speed geodesic $\alpha : [0, \ell] \to \mathcal{M}^n$ such that $\alpha(0) = x$ and $\alpha(\ell) = y$, but $\alpha \not\subseteq \mathbb{H}_\gamma$. (We abuse notation by writing α to denote its image $\alpha([0, \ell])$.) First observe that $B(\gamma(t), t) \supseteq B(\gamma(s), s)$ for all $t \geq s > 0$, since by Lemma B.44,

$$p \in B(\gamma(s), s) \quad \Rightarrow \quad 0 < b_{\gamma,s}(p) \leq b_{\gamma,t}(p) \quad \Rightarrow \quad p \in B(\gamma(t), t).$$

Then note that by definition of $\mathbb{H}_\gamma$, we have $\alpha \cap B(\gamma(s_0), s_0) \neq \emptyset$ for some $s_0 > 0$, hence for all $s \geq s_0$. Fix any $s \geq \max\{s_0, 1\}$. Since the image of α is compact, there is $\tau_s \in [0, \ell]$ and a minimal geodesic $\beta_s : [0, \ell_s] \to \mathcal{M}^n$ such that $\beta_s(0) = \alpha(\tau_s)$ and $\beta_s(\ell_s) = \gamma(s)$, where $\ell_s \doteqdot L(\beta_s) = d(\alpha, \gamma(s)) < s$.

In fact, $\tau_s \in (0, \ell)$, because $\alpha(0), \alpha(\ell) \in \mathbb{H}_\gamma$. Thus we can apply the first variation formula to conclude that $\dot{\alpha}(\tau_s) \perp \dot{\beta}_s(0)$ at $\alpha(\tau_s) = \beta_s(0)$. Now parallel translate the unit vector $\dot{\alpha}(\tau_s)$ to get a unit vector field U along β_s, and define the variation vector field

$$V(\beta_s(\sigma)) \doteqdot \frac{L(\beta_s) - \sigma}{L(\beta_s)} U(\beta_s(\sigma)), \qquad 0 \le \sigma \le L(\beta_s).$$

Note that V is the Jacobi field of a family of geodesics joining α to $\gamma(s)$; in particular, $V(\beta_s(0)) = \dot{\alpha}(\tau_s)$ and $V(\beta_s(\ell_s)) = 0$. Because β_s is minimal among all geodesics joining α to $\gamma(s)$, applying the second variation formula with $S \doteqdot (\beta_s)_* (d/d\sigma)$ yields

$$0 \le \int_0^{L(\beta_s)} \left(|\nabla_S V|^2 - \langle R(S,V)V, S\rangle \right) d\sigma.$$

Since $\nabla_S V = -\frac{1}{L(\beta_s)} U$, we have

$$\int_0^{L(\beta_s)} |\nabla_S V|^2 \, d\sigma = \frac{1}{L(\beta_s)^2} \int_0^{L(\beta_s)} d\sigma = \frac{1}{L(\beta_s)}.$$

But since $\operatorname{sect}(g) > 0$, there is $\varepsilon > 0$ depending only on $\alpha \subset \mathcal{M}^n$ such that

$$\begin{aligned}
\int_0^{L(\beta_s)} \langle R(S,V)V, S\rangle \, d\sigma &\ge \int_0^1 \langle R(S,V)V, S\rangle \, d\sigma \\
&\ge \varepsilon \int_0^1 \left(\frac{L(\beta_s) - \sigma}{L(\beta_s)} \right)^2 d\sigma \\
&\ge \varepsilon \frac{L(\beta_s) - 1}{L(\beta_s)}.
\end{aligned}$$

Hence $\mathbb{H}_\gamma$ can fail to be totally convex only if

$$0 \le \frac{1}{L(\beta_s)} - \varepsilon \frac{L(\beta_s) - 1}{L(\beta_s)},$$

which is equivalent to the condition

$$\varepsilon (L(\beta_s) - 1) \le 1$$

holding for all $s \ge s_0 > 0$. Since $L(\beta_s) \to \infty$ as $s \to \infty$, this is impossible. □

Proof of Proposition B.54 in the case $\operatorname{sect}(g) \ge 0$. If $\mathbb{H}_\gamma$ is not totally convex, there are $x, y \in \mathbb{H}_\gamma$ and a unit-speed geodesic $\alpha : [0, \ell] \to \mathcal{M}^n$ such that $\alpha(0) = x$ and $\alpha(\ell) = y$, but $\alpha \not\subseteq \mathbb{H}_\gamma$. As in the proof above, there is $s_0 > 0$ such that $\alpha \cap B(\gamma(s), s) \neq \emptyset$ for all $s \ge s_0$. Since $\alpha([0, \ell])$ is compact, there are for any $s \ge s_0$ some $\tau_s \in [0, \ell]$ and a minimal geodesic $\beta_{s,\tau_s} : [0, \ell_{s,\tau_s}] \to \mathcal{M}^n$ such that $\beta_{s,\tau_s}(0) = \alpha(\tau_s)$ and $\beta_{s,\tau_s}(\ell_{s,\tau_s}) = \gamma(s)$, where $\ell_{s,\tau_s} \doteqdot L(\beta_{s,\tau_s}) = d(\alpha, \gamma(s)) < s$. Define

$$\varepsilon \doteqdot s_0 - d(\alpha(\tau_{s_0}), \gamma(s_0)) = s_0 - d(\alpha, \gamma(s_0)) > 0$$

and notice that the triangle inequality proves for all $s \geq s_0$ that

$$\begin{aligned} \ell_{s,\tau_s} &= d(\alpha, \gamma(s)) \leq d(\alpha(\tau_{s_0}), \gamma(s)) \\ &\leq d(\alpha(\tau_{s_0}), \gamma(s_0)) + d(\gamma(s_0), \gamma(s)) = s - \varepsilon. \end{aligned} \tag{B.7}$$

Let $\beta_{s,0} : [0, \ell_{s,0}] \to \mathcal{M}^n$ be a minimal geodesic such that $\beta_{s,0} = \alpha(0) = x$ and $\beta_{s,0}(\ell_{s,0}) = \gamma(s)$, where $\ell_{s,0} \doteqdot L(\beta_{s,0}) = d(x, \gamma(s))$. Because $\alpha(0), \alpha(\ell) \in \mathbb{H}_\gamma = \mathcal{M}^n \setminus \cup_{s \in (0,\infty)} B(\gamma(s), s)$, we have $\tau_s \in (0, \ell)$ and

$$\ell_{s,\tau_s} < s \leq \ell_{s,0}. \tag{B.8}$$

In particular, we can apply the first variation formula at τ_s to conclude that $\dot{\alpha}(\tau_s) \perp \dot{\beta}_{s,\tau_s}(0)$. Thus we have constructed a right triangle with vertices $\gamma(s), \alpha(0), \alpha(\tau_s)$ and sides $\beta_{s,0}, \alpha|_{[0,\tau_s]}, \beta_{s,\tau_s}$. Since $\operatorname{sect}(g) \geq 0$, the Toponogov comparison theorem implies that the hypotenuse length satisfies the inequality

$$\ell_{s,0} = L(\beta_{s,0}) \leq \sqrt{L\left(\alpha|_{[0,\tau_s]}\right)^2 + L(\beta_{s,\tau_s})^2} = \sqrt{\tau_s^2 + \ell_{s,\tau_s}^2}. \tag{B.9}$$

Recalling that $\tau_s \in [0, \ell]$ and combining inequalities (B.7), (B.8), and (B.9), we see that $\mathbb{H}_\gamma$ can fail to be totally convex only if

$$s^2 \leq \tau_s^2 + \ell_{s,\tau_s}^2 \leq \ell^2 + \ell_{s,\tau_s}^2 \leq \ell^2 + (s - \varepsilon)^2$$

for all $s \geq s_0$. Since ℓ and ε are independent of s, this is impossible. □

3.3. Constructing totally convex sublevel sets. We are now ready to construct the sublevel sets of the Busemann function associated to a point $O \in \mathcal{M}^n$.

DEFINITION B.55. If $\gamma : [0, \infty) \to \mathcal{M}^n$ is a ray and $s \in [0, \infty)$, the **shifted ray** $\gamma_s : [0, \infty) \to \mathcal{M}^n$ is defined for all $s' \in [0, \infty)$ by

$$\gamma_s(s') \doteqdot \gamma(s' + s).$$

Note that $\gamma_0 = \gamma$. It is worth noting some elementary properties of the half spaces associated to a shifted ray.

LEMMA B.56. *If $s \leq t$, then $\mathbb{B}_{\gamma_t} \subseteq \mathbb{B}_{\gamma_s}$.*

PROOF. For all $r \in [0, \infty)$, we have the inclusion

$$B(\gamma(t + r), r) \subseteq B(\gamma(s + (t - s + r)), t - s + r) \subseteq \mathbb{B}_{\gamma_s}.$$

Hence

$$\mathbb{B}_{\gamma_t} \doteqdot \bigcup_{r \in (0,\infty)} B(\gamma(t + r), r) \subseteq \mathbb{B}_{\gamma_s}.$$

□

LEMMA B.57. *Whenever $0 < s < t$,*

$$\mathbb{B}_{\gamma_s} = \{x \in \mathcal{M}^n : d(x, \mathbb{B}_{\gamma_t}) < t - s\}.$$

PROOF. We prove both inclusions $\subseteq$ and $\supseteq$ to obtain equality.

($\subseteq$) If $x \in \mathbb{B}_{\gamma_s}$, there exists $r > 0$ such that $d(x, \gamma(r+s)) < r$. We may assume $r > t - s$, because $d(x, \gamma(q+s)) < q$ for all $q \geq r$. Then we have

$$d(x, B(\gamma(s+r), r-(t-s))) < r - (r-(t-s)) = t - s.$$

This proves that $d(x, \mathbb{B}_{\gamma_t}) < t - s$, because

$$B(\gamma(s+r), r-(t-s)) = B(t + (s+r-t), s+r-t) \subseteq \mathbb{B}_{\gamma_t}.$$

($\supseteq$) If $d(x, \mathbb{B}_{\gamma_t}) < t - s$ for some $x \in \mathcal{M}^n$, then there exists $y \in \mathbb{B}_{\gamma_t}$ such that $d(x,y) < t - s$. By definition of $\mathbb{B}_{\gamma_t}$, this implies there is $r \in (0, \infty)$ such that $y \in B(\gamma_t(r), r)$. Thus we have

$$\begin{aligned} d(x, \gamma(s + (t-s+r))) &= d(x, \gamma(t+r)) \\ &\leq d(x,y) + d(y, \gamma(t+r)) < t - s + r, \end{aligned}$$

and hence $x \in B(\gamma(s + (t-s+r)), t-s+r) \subseteq \mathbb{B}_{\gamma_s}$. $\square$

DEFINITION B.58. Given a point $O \in \mathcal{M}^n$ and $s \in [0, \infty)$, the **sublevel set of the Busemann function associated to** O is

$$C_s \doteqdot \bigcap_{\gamma \in \mathcal{R}(O)} \mathbb{H}_{\gamma_s},$$

where $\mathcal{R}(O)$ is the set of all rays emanating from O.

Note that

$$C_s = \mathcal{M}^n \setminus \bigcup_{\gamma \in \mathcal{R}(O)} \mathbb{B}_{\gamma_s}.$$

The next several results explain both the name and the usefulness of C_s.

LEMMA B.59. *For every choice of origin $O \in \mathcal{M}^n$ and $s \in [0, \infty)$, the sublevel set of the Busemann function associated to O is given by*

$$C_s = \left\{ x \in \mathcal{M}^n : \bar{b}(x) \leq s \right\},$$

where $\bar{b}$ is the Busemann function associated to O. In particular, we have

$$C_s \subseteq C_t$$

whenever $0 \leq s \leq t$, and

$$\partial C_s = \left\{ x \in \mathcal{M}^n : \bar{b}(x) = s \right\}$$

because $\bar{b}$ is continuous.

PROOF. Observe that $b_{\gamma_s} = b_\gamma - s$, since for all $x \in \mathcal{M}^n$,

$$\begin{aligned} b_{\gamma_s}(x) &\doteqdot \lim_{t \to \infty} b_{\gamma_s, t}(x) = \lim_{t \to \infty} (t - d(\gamma_s(t), x)) \\ &= -s + \lim_{t \to \infty} (s + t - d(\gamma(s+t), x)) = -s + b_\gamma(x). \end{aligned}$$

Hence by Lemma B.52,

$$\begin{aligned} C_s &\doteqdot \bigcap_{\gamma \in \mathcal{R}(O)} \mathbb{H}_{\gamma_s} = \bigcap_{\gamma \in \mathcal{R}(O)} \{x \in \mathcal{M}^n : b_{\gamma_s}(x) \leq 0\} \\ &= \{x \in \mathcal{M}^n : b_\gamma(x) \leq s \text{ for all } \gamma \in \mathcal{R}(O)\} \\ &= \left\{x \in \mathcal{M}^n : \bar{b}(x) \leq s\right\}. \end{aligned}$$

□

COROLLARY B.60. *If $s > 0$, then C_s contains the closed ball of radius s at O.*

PROOF. By Lemma B.48, we have $\bar{b}(x) \leq d(O, x)$, and hence

$$C_s = \left\{x \in \mathcal{M}^n : \bar{b}(x) \leq s\right\} \supseteq \{x \in \mathcal{M}^n : d(O, x) \leq s\} \doteqdot \bar{B}(O, s).$$

□

COROLLARY B.61. *For any choice of origin $O \in \mathcal{M}^n$, one has $\cup_{s \geq 0} C_s = \mathcal{M}^n$.*

PROOF. $\mathcal{M}^n = \cup_{s \geq 0} B(O, s) \subseteq \cup_{s \geq 0} C_s$. □

PROPOSITION B.62. *For every choice of origin $O \in \mathcal{M}^n$ and $s \in [0, \infty)$, the set C_s is compact and totally convex.*

PROOF. By Proposition B.54, C_s is the intersection of closed and totally convex sets, hence is itself closed and totally convex.

Suppose C_s is not compact. Then there exists a sequence of points $p_i \in C_s$ with $d(O, p_i) \to \infty$ as $i \to \infty$. For each i, let $\beta_i : [0, d(O, p_i)] \to \mathcal{M}^n$ be a unit-speed minimal geodesic from O to p_i. Since C_s is totally convex, each $\beta_i \subset C_s$. After passing to a subsequence, we may assume the unit tangent vectors $\dot{\beta}_i(0)$ converge to a unit vector $V_\infty \in T_O\mathcal{M}^n$. Let $\beta_\infty : [0, \infty) \to \mathcal{M}^n$ denote the geodesic with $\dot{\beta}_\infty(0) = V_\infty$. As in Lemma B.42, we observe that β_∞ is a ray such that the images $\beta_i \to \beta_\infty$ uniformly on compact sets. Now for any $t \in (0, \infty)$, consider the point $\beta_\infty(s + t)$. Then $\beta_i(s + t) \in C_s$ is defined for all i large enough, and

$$\lim_{i \to \infty} d(\beta_i(s + t), \beta_\infty(s + t)) = 0.$$

Since C_s is closed, this implies $\beta_\infty(s + t) \in C_s$, which contradicts the fact that $\beta_\infty(s + t) \in \mathbb{B}_{(\beta_\infty)_s}$. □

COROLLARY B.63. *The Busemann function $\bar{b}$ associated to a point $O \in \mathcal{M}^n$ is bounded below.*

PROOF. $\bar{b}(x)$ is continuous and $C_s = \left\{x \in \mathcal{M}^n : \bar{b}(x) \leq s\right\}$ is compact.

□

We have seen that the totally convex sets C_s exhaust $\mathcal{M}^n$. The next result refines Lemma B.59 and gives the sense in which the level sets of $\bar{b}$ are parallel.

PROPOSITION B.64. *For any choice of origin $O \in \mathcal{M}^n$ and all $s < t$,*

$$C_s = \{x \in C_t : d(x, \partial C_t) \geq t - s\}. \tag{B.10}$$

In particular, $O \in \partial C_0$ and

$$\partial C_s = \{x \in C_t : d(x, \partial C_t) = t - s\}.$$

PROOF. By Lemma B.59, we have $O \in \partial C_0 = \{x \in \mathcal{M}^n : \bar{b}(x) = 0\}$, because $\bar{b}(O) = 0$. And since the distance function is continuous, the characterization of ∂C_s given here follows from (B.10). So it will suffice to prove both inclusions $\subseteq$ and $\supseteq$ in (B.10).

($\subseteq$) Suppose $x \in C_t$ and $d(x, \partial C_t) < t - s$. We want to show that $x \notin C_s$. Because

$$d\left(x, \cup_{\gamma \in \mathcal{R}(O)} \mathbb{B}_{\gamma_t}\right) = d(x, \mathcal{M}^n \backslash C_t) = d(x, \partial C_t) < t - s,$$

there exists a ray β emanating from O such that $d(x, \mathbb{B}_{\beta_t}) < t - s$. By Lemma B.57, this proves that $x \in \mathbb{B}_{\beta_s}$, hence that $x \notin C_s$.

($\supseteq$) Suppose $x \in C_t \backslash C_s$. We want to show that $d(x, \partial C_t) < t - s$. Since $x \notin C_s = \cap_{\gamma \in \mathcal{R}(O)} \mathbb{H}_{\gamma_s}$, there is a ray β emanating from O such that $x \in \mathbb{B}_{\beta_s}$. By Lemma B.57, this proves that $d(x, \mathbb{B}_{\beta_t}) < t - s$. Hence

$$d(x, \partial C_t) = d(x, \mathcal{M}^n \backslash C_t) = d\left(x, \cup_{\gamma \in \mathcal{R}(O)} \mathbb{B}_{\gamma_t}\right) < t - s.$$

□

4. Estimating injectivity radius in positive curvature

The objective of this section is to prove the following result.

THEOREM B.65. *Let $(\mathcal{M}^n, g)$ be a complete noncompact Riemannian manifold of positive sectional curvatures bounded above by some $K \in (0, \infty)$. Then its injectivity radius satisfies*

$$\operatorname{inj}(\mathcal{M}^n, g) \geq \frac{\pi}{\sqrt{K}}.$$

Before proving the theorem, we will establish some preliminary results. Assume for now only that the sectional curvatures of $(\mathcal{M}^n, g)$ are bounded above by $K > 0$.

DEFINITION B.66. If $k \in \mathbb{N}^*$, a **proper geodesic k-gon** is a collection

$$\Gamma = \{\gamma_i : [0, \ell_i] \to \mathcal{M}^n : i = 1, \ldots, k\}$$

of unit-speed geodesic paths between k pairwise-distinct vertices $p_i \in \mathcal{M}^n$ such that $p_i = \gamma_i(0) = \gamma_{i-1}(\ell_{i-1})$ for each i, where all indices are interpreted modulo k. The **total length** of a proper geodesic k-gon is

$$L(\Gamma) \doteqdot \sum_{i=1}^{k} L(\gamma_i).$$

Γ is a **nondegenerate proper geodesic k-gon** if $\measuredangle_{p_i}(-\dot{\gamma}_{i-1}, \dot{\gamma}_i) \neq 0$ for each $i = 1, \ldots, k$; if $k = 1$, we interpret this to mean $L(\Gamma) > 0$.

Note that a proper j-gon may be regarded as a proper k-gon for any $k > j$ merely by choosing extra vertices. In fact, it will be easier to deal with limits if we work with a more general collection of objects.

DEFINITION B.67. A **(nondegenerate) geodesic k-gon** is a (nondegenerate) proper geodesic j-gon for some $j = 1, \ldots, k$.

If the theorem is false, there is a nondegenerate geodesic 2-gon Γ in $\mathcal{M}^n$ of total length $2\lambda < \pi/\sqrt{K}$. Let $O \in \mathcal{M}^n$ be any choice of origin, and let C_r denote the sublevel sets of the Busemann function based at O. By Corollary B.61, the collection $\{C_r : 0 < r < \infty\}$ exhausts $\mathcal{M}^n$. Thus there is $s \in (0, \infty)$ such that $\Gamma \subset C_s$. By Proposition B.62, C_s is compact. So define

$$\Lambda \doteqdot \{\text{nondegenerate geodesic 2-gons in } C_s \text{ of total length } \leq 2\lambda\}$$

and

$$\Xi \doteqdot \{\text{geodesic 2-gons in } C_s \text{ of total length } \leq 2\lambda\}.$$

Let $\mathcal{S}^{n-1}C_s$ denote the unit sphere bundle of $\mathcal{M}^n$ restricted to C_s. Observe that

$$\Lambda \subseteq \Xi \subseteq \left(\mathcal{S}^{n-1}C_s \times [0, 2\lambda]\right) \times \left(\mathcal{S}^{n-1}C_s \times [0, 2\lambda]\right),$$

because any $\Gamma \in \Xi$ is described uniquely by data $(p, V, \ell, p', V', \ell')$, where γ_1 is the geodesic path determined by $\gamma_1(0) = p$, $\dot{\gamma}_1(0) = V \in T_p\mathcal{M}^n$, and $L(\gamma_1) = \ell \in [0, 2\lambda]$, while γ_2 is the geodesic path determined by $\gamma_2(0) = p'$, $\dot{\gamma}_2(0) = V' \in T_{p'}\mathcal{M}^n$, and $L(\gamma_2) = \ell' \in [0, 2\lambda]$. ($V$ may be chosen arbitrarily if $\ell = 0$, likewise for V' and ℓ'.) It is easy to see that Ξ is closed in the induced topology, hence is compact.

LEMMA B.68. *Let $(\mathcal{M}^n, g)$ be a complete noncompact manifold with sectional curvatures bounded above by K. Let $\Lambda \subseteq \Xi$ be defined as above. Then Λ is closed, hence is compact.*

PROOF. Suppose $\{\alpha_i\}$ is a sequence from Λ such that $\alpha_i \to \alpha_\infty \in \Xi$. Equivalently,

$$(p_i, V_i, \ell_i, p'_i, V'_i, \ell'_i) \to (p_\infty, V_\infty, \ell_\infty, p'_\infty, V'_\infty, \ell'_\infty).$$

We first claim α_∞ is not a degenerate proper 1-gon. This can happen only if $\ell_\infty = \ell'_\infty = 0$, hence only if $p_\infty = p'_\infty$. But since C_s is compact, we see that this is impossible, because

$$0 < \operatorname{inj}(C_s) \leq \operatorname{inj}(p_i) = d(p_i, \operatorname{cut}(p_i)) \leq \max\{\ell_i, \ell'_i\} \searrow 0$$

as $i \to \infty$.

We next claim that α_∞ is not a degenerate proper 2-gon. This can happen only if the path β from p_∞ to $p'_\infty \neq p_\infty$ is the path β' from p'_∞ to p_∞, traced in the opposite direction. Then for i sufficiently large, there are distinct but nearby geodesics β_i from p_i to p'_i and β'_i from p'_i to p_i, such that $\max\{L(\beta_i), L(\beta'_i)\} \leq \frac{2}{3}\pi/\sqrt{K}$. But this is impossible, because the sectional curvature hypothesis implies $\exp_{p_i} : B(\vec{0}, \pi/\sqrt{K}) \to \mathcal{M}^n$ is an immersion.

Now by the claims above, we know α_∞ is a nondegenerate proper 1-gon or 2-gon, hence belongs to Λ. It follows that Λ is a closed subset of the compact set Ξ. $\square$

This compactness property implies there is $\beta \in \Lambda$ whose total length realizes $L(\beta) = \inf_{\Gamma \in \Lambda} L(\Gamma)$. The extra hypothesis of positive sectional curvature guarantees that β is smooth:

LEMMA B.69. *Let $(\mathcal{M}^n, g)$ be a complete noncompact Riemannian manifold of positive sectional curvatures bounded above by $K > 0$. Then any $\beta \in \Lambda$ such that $L(\beta) = \inf_{\Gamma \in \Lambda} L(\Gamma)$ is a smooth geodesic loop.*

PROOF. By choosing a vertex at $L(\beta)/2$ if necessary, we may suppose without loss of generality that β is a nondegenerate proper 2-gon corresponding to the data $(p, V, \ell, p', V', \ell')$. Let γ denote the path from p to p', and let γ' denote the path from p' to p. We will show that β is smooth at p', hence at p by relabeling.

Define a 1-parameter family

$$\{\beta_t : 0 \leq t < \ell_0 \doteqdot \min\{\ell, \ell'\}\}$$

of nondegenerate proper 2-gons β_t by taking $p'_t \doteqdot \gamma(\ell - t)$ and letting γ_t denote the truncated path $\gamma|_{[0,\ell-t]}$. Then by Corollary B.22, there is a unique geodesic γ'_t near γ' joining p'_t to p. By Proposition B.62, C_s is totally convex; since $p, p'_t \in C_s$, it follows that γ'_t lies in C_s, hence that $\beta_t \in \Lambda$ for all $t \in [0, \ell_0)$. Notice that β can fail to be smooth at p' only if $\dot\gamma(\ell) \neq \dot\gamma'(\ell')$. But the variation vector field of β_t is a Jacobi field V along γ' with $V(0) = -\dot\gamma(\ell)$ and $V(\ell') = 0$. Thus by the first variation formula, we take a one-sided derivative to obtain

$$\begin{aligned}\left.\frac{d}{dt} L(\beta_t)\right|_{t=0} &= \langle V(0), \dot\gamma'(0) + \dot\gamma(\ell)\rangle = -\langle \dot\gamma(\ell), \dot\gamma'(0)\rangle - \langle \dot\gamma(\ell), \dot\gamma(\ell)\rangle \\ &= -\langle \dot\gamma(\ell), \dot\gamma'(0)\rangle - 1 < 0,\end{aligned}$$

because $|\dot\gamma| = |\dot\gamma'| = 1$. This contradicts the minimality of $L(\beta)$ in Λ unless β is smooth at p'. $\square$

PROOF OF THEOREM B.65. We have already shown that the result can fail only if there is a smooth geodesic loop β of length less than $\pi/\sqrt{K}$ contained in a compact totally geodesic set C_s based at $O \in \mathcal{M}^n$. Let α be any ray emanating from O. For $t \in (1, \infty)$ to be chosen later, join the loop β to the point $\alpha(t)$ by a minimal geodesic $\gamma_t : [0, \ell_t] \to \mathcal{M}^n$, where $\gamma_t(0) \in \beta$, $\gamma_t(\ell_t) = \alpha(t)$, and $\ell_t \doteqdot d(\alpha(t), \beta)$. Since β is smooth, we can apply the first variation formula to conclude that $\dot\gamma_t \perp \dot\beta$ at $\gamma_t(0) \in \beta$. As in the first proof of Proposition B.54, we shall obtain a contradiction by applying a second variation argument to γ_t among curves joining β to $\alpha(t)$.

Parallel translate $\dot{\beta}(\gamma_t(0))$ to get a unit vector field U along γ_t, and define

$$V(\gamma_t(\tau)) \doteqdot \frac{\ell_t - \tau}{\ell_t} U(\gamma_t(\tau)), \qquad 0 \leq \tau \leq \ell_t.$$

Then V is a Jacobi field with $V(\gamma_t(0)) = \dot{\beta}(\gamma_t(0))$ and $V(\gamma_t(\ell_t)) = 0$. Because γ_t is minimal among geodesics from β to $\alpha(t)$, the second variation formula applied with $T \doteqdot \dot{\gamma}_t$ yields

$$0 \leq \int_0^{\ell_t} \left(|\nabla_T V|^2 - \langle R(T,V)V,T\rangle \right) d\tau.$$

Noting that $\nabla_T V = -\frac{1}{\ell_t} U$, we calculate that the first term is

$$\int_0^{\ell_t} |\nabla_T V|^2 \, d\tau = \frac{1}{\ell_t^2} \int_0^{\ell_t} d\tau = \frac{1}{\ell_t}.$$

Since $\sec(g) > 0$, there is $\varepsilon > 0$ depending only on $\beta \subset C_s \subset \mathcal{M}^n$ such that

$$\begin{aligned} \int_0^{\ell_t} \langle R(T,V)V,T\rangle \, d\tau &\geq \int_0^1 \langle R(T,V)V,T\rangle \, d\tau \\ &\geq \varepsilon \int_0^1 \left(\frac{\ell_t - \tau}{\ell_t} \right)^2 d\tau \geq \varepsilon \frac{\ell_t - 1}{\ell_t}. \end{aligned}$$

Hence if there is a smooth geodesic loop β of length less than $\pi/\sqrt{K}$ contained in C_s, we must have

$$0 \leq \frac{1 + \varepsilon - \varepsilon \ell_t}{\ell_t}$$

holding for all choices of $t \in (1, \infty)$. Since $\ell_t \to \infty$ as $t \to \infty$, this is impossible. □

Notes and commentary

The main reference for comparison geometry is Cheeger and Ebin's book [**27**]. A more recent survey is the book [**54**] edited by Grove and Petersen.

Bibliography

[1] Altschuler, Steven; Angenent, Sigurd B.; Giga, Yoshikazu. *Mean curvature flow through singularities for surfaces of rotation.* J. Geom. Anal. **5** (1995), no. 3, 293-358.

[2] Anderson, Michael T. *Convergence and rigidity of manifolds under Ricci curvature bounds.* Invent. Math. **102** (1990), no. 2, 429-445.

[3] Andrews, Ben. *Entropy estimates for evolving hypersurfaces.* Comm. Anal. Geom. **2** (1994), no. 1, 53-64.

[4] Andrews, Ben. Personal communication.

[5] Angenent, Sigurd B. *The zero set of a solution of a parabolic equation.* J. Reine Angew. Math. **390** (1988), 79-96.

[6] Angenent, Sigurd B.; Knopf, Dan. *An example of neckpinching for Ricci flow on* S^{n+1}. Preprint.

[7] Angenent, Sigurd B.; Velázquez, J. J. L. *Degenerate neckpinches in mean curvature flow.* J. Reine Angew. Math. **482** (1997), 15-66.

[8] Aronson, D. G. *The porous medium equation.* Nonlinear diffusion problems (Montecatini Terme, 1985), 1-46, Lecture Notes in Math., 1224, Springer, Berlin, 1986.

[9] Bando, Shigetoshi. *Real analyticity of solutions of Hamilton's equation.* Math. Z. **195** (1987), no. 1, 93-97.

[10] Bartz, J.; Struwe, Michael.; Ye, Rugang. *A new approach to the Ricci flow on* S^2. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **21** (1994), no. 3, 475-482.

[11] Bemelmans, Josef; Min-Oo; Ruh, Ernst A. *Smoothing Riemannian metrics.* Math. Z. **188** (1984), no. 1, 69-74.

[12] Berger, Marcel. *Les variétés Riemanniennes* 1/4*-pincées.* (French) Ann Scuola Norm. Sup. Pisa (3) **14** 1960 161-170.

[13] Berger, Marcel. *Sur quelques variétés riemanniennes suffisamment pincées.* (French) Bull. Soc. Math. France **88** 1960 57-71.

[14] Berger, Marcel. *On the diameter of some Riemannian manifolds.* Technical report, University of California Berkeley (1962).

[15] Berger, M.; Ebin, D. *Some decompositions of the space of symmetric tensors on a Riemannian manifold.* J. Differential Geometry **3** (1969) 379-392.

[16] Bernstein, Sergi N. *Sur la généralisation du problème de Dirichlet II.* Math. Ann. **69** (1910) 82-136.

[17] Bernstein, Sergi N. *Sur les équations du calcul des variations.* Ann. Sci. École Norm. Sup. **29** (1912) 431-485.

[18] Bernstein, Sergi N. *A limitation on the moduli of a sequence of derivatives of solutions of equations of parabolic type.* (Russian) Dokl. Akad. Nauk SSSR **18** (1938) 385-388.

[19] Bertozzi, Andrea L. *The mathematics of moving contact lines in thin liquid films.* Notices Amer. Math. Soc. **45** (1998), no. 6, 689-697.

[20] Besse, Arthur L. *Einstein manifolds.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], **10.** Springer-Verlag, Berlin, 1987.

[21] Bryant, Robert. *Local existence of gradient Ricci solitons.* Preprint.

[22] Casson, Andrew; Jungreis, Douglas. *Convergence groups and Seifert fibered 3-manifolds.* Invent. Math. **118** (1994), no. 3, 441–456.

[23] Cao, Huai-Dong. *Existence of gradient Kähler-Ricci solitons.* Elliptic and parabolic methods in geometry (Minneapolis, MN, 1994), 1–16, A K Peters, Wellesley, MA, 1996.

[24] Cao, Huai-Dong. *Limits of solutions to the Kähler-Ricci flow.* J. Differential Geom. **45** (1997), no. 2, 257–272.

[25] Chavel, Isaac. *Riemannian geometry — a modern introduction.* Cambridge Tracts in Mathematics, **108**. Cambridge University Press, Cambridge, 1993.

[26] Cheeger, Jeff. *Finiteness theorems for Riemannian manifolds.* Amer. J. Math. **92** (1970) 61–74.

[27] Cheeger, Jeff; Ebin, David G. *Comparison theorems in Riemannian geometry.* North-Holland Mathematical Library, Vol. 9. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1975.

[28] Chow, Bennett. *The Ricci flow on the 2-sphere.* J. Differential Geom. **33** (1991), no. 2, 325–334.

[29] Chow, Bennett. *On the entropy estimate for the Ricci flow on compact 2-orbifolds.* J. Differential Geom. **33** (1991), no. 2, 597–600.

[30] Chow, Bennett. *A gradient estimate for the Ricci-Kähler flow.* Ann. Global Anal. Geom. **19** (2001), no. 4, 321–325.

[31] Chow, Bennett. *On Harnack's inequality and entropy for the Gaussian curvature flow.* Comm. Pure Appl. Math. **44** (1991), no. 4, 469–483.

[32] Chow, Bennett; Hamitlon, Richard S. *The Cross Curvature Flow of 3-manifolds with Negative Sectional Curvature.* `arXiv:math.DG/0309008`.

[33] Chow, Bennett; Lu, Peng. *The time-dependent maximum principle for systems of parabolic equations subject to an avoidance set.* Pacific J. Math.. To appear. Also see: `arXiv:math.DG/0211209`.

[34] Chow, Bennett; Wu, Lang-Fang. *The Ricci flow on compact 2-orbifolds with curvature negative somewhere.* Comm. Pure Appl. Math. **44** (1991), no. 3, 275–286.

[35] Daskalopoulos, Panagiota; Hamilton, Richard S. *Geometric estimates for the logarithmic fast diffusion equation.* Preprint.

[36] DeTurck, Dennis M. *Deforming metrics in the direction of their Ricci tensors.* J. Differential Geom. **18** (1983), no. 1, 157–162.

[37] DeTurck, Dennis M. *Deforming metrics in the direction of their Ricci tensors, improved version,* Collected Papers on Ricci Flow, ed. H.-D. Cao, B. Chow, S.-C. Chu, and S.-T. Yau. Internat. Press, Somerville, MA, 2003.

[38] Ebin, D. *The manifold of Riemannian metrics.* 1970 Global Analysis (Proc. Sympos. Pure Math., Vol. XV, Berkeley CA, 1968) 11–40, Amer. Math. Soc., Providence, RI.

[39] Epstein, D. B. A. *Periodic flows on three-manifolds.* Ann. of Math. (2) **95** 1972 66–82.

[40] Feldman, Mikhail; Ilmanen, Tom; Knopf, Dan. *Rotationally symmetric shrinking and expanding gradient Kähler-Ricci solitons.* J. Differential Geom. To appear.

[41] Filippas, Stathis; Kohn, Robert V. *Refined asymptotics for the blowup of $u_t - \Delta u = u^p$.* Comm. Pure Appl. Math. **45** (1992), no. 7, 821–869.

[42] Gabai, David. *Convergence groups are Fuchsian groups.* Ann. of Math. (2) **136** (1992), no. 3, 447–510.

[43] Gabai, David. *Homotopy hyperbolic 3-manifolds are virtually hyperbolic.* J. Amer. Math. Soc. **7** (1994), no. 1, 193–198.

[44] Gage, M.; Hamilton, Richard S. *The heat equation shrinking convex plane curves.* J. Differential Geom. **23** (1986), no. 1, 69–96.

[45] Gao, L. Zhiyong. *Convergence of Riemannian manifolds; Ricci and $L^{n/2}$-curvature pinching.* J. Differential Geom. **32** (1990), no. 2, 349–381.

[46] Giga, Yoshikazu; Kohn, Robert V. *Asymptotically self-similar blow-up of semilinear heat equations.* Comm. Pure Appl. Math. **38** (1985), no. 3, 297–319.

[47] Giga, Yoshikazu; Kohn, Robert V. *Characterizing blowup using similarity variables.* Indiana Univ. Math. J. **36** (1987), no. 1, 1–40.

[48] Giga, Yoshikazu; Kohn, Robert V. *Nondegeneracy of blowup for semilinear heat equations.* Comm. Pure Appl. Math. **42** (1989), no. 6, 845–884.

[49] Grayson, Matthew A. *Shortening embedded curves.* Ann. of Math. (2) **129** (1989), no. 1, 71–111.

[50] Greene, R. E.; Wu, H. *Lipschitz convergence of Riemannian manifolds.* Pacific J. Math. **131** (1988), no. 1, 119–141.

[51] Greene, R. E.; Wu, H. Addendum to: "Lipschitz convergence of Riemannian manifolds" Pacific J. Math. **140** (1989), no. 2, 398.

[52] Gromoll, Detlef; Meyer, Wolfgang. *On complete open manifolds of positive curvature.* Ann. of Math. (2) **90** 1969 75-90.

[53] Gromov, Misha. *Metric structures for Riemannian and non-Riemannian spaces.* Based on the 1981 French original. With appendices by M. Katz, P. Pansu and S. Semmes. Translated from the French by Sean Michael Bates. Progress in Mathematics, **152**. Birkhäuser Boston, Inc., Boston, MA, 1999.

[54] Grove, Karsten; Petersen, Peter. (editors) *Comparison geometry.* Papers from the Special Year in Differential Geometry held in Berkeley, CA, 1993-94. Mathematical Sciences Research Institute Publications, **30**. Cambridge University Press, Cambridge, 1997.

[55] Gutperle, Michael; Headrick, Matthew; Minwalla, Shiraz; Schomerus, Volker. *Space-time Energy Decreases under World-sheet RG Flow.* `arXiv:hep-th/0211063`.

[56] Hamilton, Richard S. *Harmonic maps of manifolds with boundary.* Lecture Notes in Mathematics, Vol. 471. Springer-Verlag, Berlin-New York, 1975.

[57] Hamilton, Richard S. *The inverse function theorem of Nash and Moser.* Bull. Amer. Math. Soc. (N.S.) **7** (1982), no. 1, 65–222.

[58] Hamilton, Richard S. *Three-manifolds with positive Ricci curvature.* J. Differential Geom. **17** (1982), no. 2, 255–306.

[59] Hamilton, Richard S. *Four-manifolds with positive curvature operator.* J. Differential Geom. **24** (1986), no. 2, 153–179.

[60] Hamilton, Richard S. *The Ricci flow on surfaces.* Mathematics and general relativity (Santa Cruz, CA, 1986), 237–262, Contemp. Math., **71**, Amer. Math. Soc., Providence, RI, 1988.

[61] Hamilton, Richard S. *The Harnack estimate for the Ricci flow.* J. Differential Geom. **37** (1993), no. 1, 225–243.

[62] Hamilton, Richard S. *An isoperimetric estimate for the Ricci flow on the two-sphere.* Modern methods in complex analysis (Princeton, NJ, 1992), 191–200, Ann. of Math. Stud., **137**, Princeton Univ. Press, Princeton, NJ, 1995.

[63] Hamilton, Richard S. *The formation of singularities in the Ricci flow.* Surveys in differential geometry, Vol. II (Cambridge, MA, 1993), 7–136, Internat. Press, Cambridge, MA, 1995.

[64] Hamilton, Richard S. *A compactness property for solutions of the Ricci flow.* Amer. J. Math. **117** (1995), no. 3, 545–572.

[65] Hamilton, Richard S. *Four-manifolds with positive isotropic curvature.* Comm. Anal. Geom. **5** (1997), no. 1, 1–92.

[66] Hamilton, Richard S. *Non-singular solutions of the Ricci flow on three-manifolds.* Comm. Anal. Geom. **7** (1999), no. 4, 695–729.

[67] Hamilton, Richard; Isenberg, James. *Quasi-convergence of Ricci flow for a class of metrics.* Comm. Anal. Geom. **1** (1993), no. 3-4, 543–559.

[68] Hamilton, Richard S.; Yau, Shing-Tung. *The Harnack estimate for the Ricci flow on a surface—revisited.* Asian J. Math. **1** (1997), no. 3, 418–421.

[69] Hass, Joel; Morgan, Frank. *Geodesics and soap bubbles in surfaces.* Math. Z. **223** (1996), no. 2, 185–196.

[70] Hatcher, Allen. *Basic topology of 3-manifolds.* `http://www.math.cornell.edu/~hatcher/#3M`.

[71] Hempel, John *3-Manifolds.* Ann. of Math. Studies, No. 86. Princeton University Press, Princeton, N. J.; University of Tokyo Press, Tokyo, 1976.

[72] Hsu, Shu-Yu. *Large time behaviour of solutions of the Ricci flow equation on* $\mathbb{R}^2$. Pacific J. Math. **197** (2001), no. 1, 25–41.

[73] Hsu, Shu-Yu. *Global existence and uniqueness of solutions of the Ricci flow equation.* Differential Integral Equations **14** (2001), no. 3, 305–320.

[74] Huisken, Gerhard. *Flow by mean curvature of convex surfaces into spheres.* J. Differential Geom. **20** (1984), no. 1, 237–266.

[75] Huisken, Gerhard. *Ricci deformation of the metric on a Riemannian manifold.* J. Differential Geom. **21** (1985), no. 1, 47–62.

[76] Isenberg, James; Jackson, Martin. *Ricci flow of locally homogeneous geometries on closed manifolds.* J. Differential Geom. **35** (1992), no. 3, 723–741.

[77] Ivey, Thomas. *Ricci solitons on compact Kähler surfaces.* Proc. Amer. Math. Soc. **125** (1997), no. 4, 1203–1208.

[78] Jaco, William. *Lectures on three-manifold topology.* CBMS Regional Conference Series in Mathematics, 43. American Mathematical Society, Providence, R.I., 1980.

[79] Jaco, William; Shalen, Peter B. *Seifert fibered spaces in 3-manifolds.* Geometric topology (Proc. Georgia Topology Conf., Athens, Ga., 1977), pp. 91–99, Academic Press, New York-London, 1979.

[80] Johannson, Klaus. *Homotopy equivalences of 3-manifolds with boundaries.* Lecture Notes in Mathematics, **761** Springer, Berlin, 1979.

[81] Kazdan, Jerry L. *Another proof of Bianchi's identity in Riemannian geometry.* Proc. Amer. Math. Soc. **81** (1981), no. 2, 341–342

[82] Kazdan, Jerry L.; Warner, F. W. *Curvature functions for compact 2-manifolds.* Ann. of Math. (2) **99** (1974), 14–47.

[83] Klingenberg, Wilhelm. *Contributions to Riemannian geometry in the large.* Ann. of Math. (2) **69** 1959 654–666.

[84] Klingenberg, Wilhelm. *Über Riemannsche Mannigfaltigkeiten mit positiver Krümmung.* (German) Comment. Math. Helv. **35** 1961 47–54.

[85] Kneser, H. *Geschloßene Flächen in dreidimensionalen Mannigfaltikeiten,* Jahresbericht der Deut. Math. Verein. **38** (1929) 248–260.

[86] Knopf, Dan; McLeod, Kevin. *Quasi-convergence of model geometries under the Ricci flow.* Comm. Anal. Geom. **9** (2001), no. 4, 879–919.

[87] Kobayashi, Shoshichi. *Transformation groups in differential geometry.* Reprint of the 1972 edition. Classics in Mathematics. Springer-Verlag, Berlin, 1995.

[88] Koiso, Norihito. *On rotationally symmetric Hamilton's equation for Kähler-Einstein metrics.* Recent topics in differential and analytic geometry, 327–337, Adv. Stud. Pure Math., 18-I, Academic Press, Boston, MA, 1990.

[89] Ladyženskaja, O. A.; Solonnikov, V. A.; Uralćeva, N. N. *Linear and quasilinear equations of parabolic type.* (Russian) Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23 American Mathematical Society, Providence, R.I. 1967.

[90] Leviton, P. R. A.; Rubinstein, J. H. *Deforming Riemannian metrics on the 2-sphere.* Miniconference on geometry and partial differential equations (Canberra, 1985), 123–127, Proc. Centre Math. Anal. Austral. Nat. Univ., **10**, Austral. Nat. Univ., Canberra, 1986.

[91] Li, Peter; Yau, Shing Tung. *Estimates of eigenvalues of a compact Riemannian manifold.* Geometry of the Laplace operator (Proc. Sympos. Pure Math., Univ. Hawaii, Honolulu, Hawaii, 1979), 205–239, Proc. Sympos. Pure Math., XXXVI, Amer. Math. Soc., Providence, R.I., 1980.

[92] Li, Peter; Yau, Shing-Tung. *On the parabolic kernel of the Schrödinger operator.* Acta Math. **156** (1986), no. 3-4, 153–201.

[93] Lichnerowicz, André. *Propagateurs et commutateurs en relativité générale.* (French) Inst. Hautes Études Sci. Publ. Math. No. 10, 1961.

[94] Lieberman, Gary M. *Second order parabolic differential equations.* World Scientifici Publishing Co., River Edge, NJ, 1996.

[95] Margerin, Christophe. *Un théorème optimal pour le pincement faible en dimension* 4. (French) [An optimal theorem for a weakly pinched 4-manifold] C. R. Acad. Sci. Paris Sér. I Math. **303** (1986), no. 17, 877–880.

[96] Margerin, Christophe. *Pointwise pinched manifolds are space forms.* Geometric measure theory and the calculus of variations (Arcata, Calif., 1984), 307–328, Proc. Sympos. Pure Math., **44**, Amer. Math. Soc., Providence, RI, 1986.

[97] Milnor, John. *A unique factorization theorem for 3-manifolds,* Amer. J. Math. **84** (1962) 1–7.

[98] Milnor, John. *Curvatures of left invariant metrics on Lie groups.* Advances in Math. **21** (1976), no. 3, 293–329.

[99] Morgan, Frank. *Geometric measure theory. A beginner's guide.* Third edition. Academic Press, Inc., San Diego, CA, 2000.

[100] Myers, Sumner; Steenrod, Norman. *The group of isometries of a Riemannian manifold.* Annals Math. **40** (1939) 400–416.

[101] Ni, Lei. *Ricci flow and nonnegativity of curvature.* `arXiv:math.DG/0305246`.

[102] Nishikawa, Seiki. *Deformation of Riemannian metrics and manifolds with bounded curvature ratios.* Geometric measure theory and the calculus of variations (Arcata, Calif., 1984), 343–352, Proc. Sympos. Pure Math., **44**, Amer. Math. Soc., Providence, RI, 1986.

[103] Nishikawa, Seiki. *On deformation of Riemannian metrics and manifolds with positive curvature operator.* Curvature and topology of Riemannian manifolds (Katata, 1985), 202–211, Lecture Notes in Math., **1201**, Springer, Berlin, 1986.

[104] Palais, Richard S. *Foundations of global non-linear analysis.* W. A. Benjamin, Inc., New York-Amsterdam 1968.

[105] Perelman, Grisha. *The entropy formula for the Ricci flow and its geometric applications.* `arXiv:math.DG/0211159`.

[106] Perelman, Grisha. *Ricci flow with surgery on three-manifolds.* `arXiv:math.DG/0303109`.

[107] Perelman, Grisha. *Finite extinction time for the solutions to the Ricci flow on certain three-manifolds.* `arXiv:math.DG/0307245`.

[108] Peters, Stefan. *Convergence of Riemannian manifolds.* Compositio Math. **62** (1987), no. 1, 3–16.

[109] Petersen, Peter. *Convergence theorems in Riemannian geometry.* Comparison geometry (Berkeley, CA, 1993–94), 167–202, Math. Sci. Res. Inst. Publ., **30**, Cambridge Univ. Press, Cambridge, 1997.

[110] Protter, Murray H.; Weinberger, Hans F. *Maximum principles in differential equations.* Corrected reprint of the 1967 original. Springer-Verlag, New York, 1984.

[111] Rosenau, Philip. *On fast and super-fast diffusion.* Phys. Rev. Lett. **74** (1995) 1056–1059.

[112] Schoen, Richard M. *A report on some recent progress on nonlinear problems in geometry,* Surveys in differential geometry, Vol. I (Cambridge, MA, 1990), 201–241, Lehigh Univ., Bethlehem, PA, 1991.

[113] Schoen, R.; Yau, S.-T. *Lectures on differential geometry.* Lecture notes prepared by Wei Yue Ding, Kung Ching Chang [Gong Qing Zhang], Jia Qing Zhong and Yi Chao Xu. Translated from the Chinese by Ding and S. Y. Cheng. Preface translated from the Chinese by Kaising Tso. Conference Proceedings and Lecture Notes in Geometry and Topology, I. International Press, Cambridge, MA, 1994.

[114] Scott, Peter. *A new proof of the annulus and torus theorems.* Amer. J. Math. **102** (1980), no. 2, 241–277.

[115] Scott, Peter. *The geometries of 3-manifolds.* Bull. London Math. Soc. **15** (1983), no. 5, 401–487.

[116] Seifert, Herbert. *Topologie dreidimensionaler gefaserter Räume.* Acta. Math. **60** (1933) 147–238.

[117] Shi, Wan-Xiong. *Deforming the metric on complete Riemannian manifolds.* J. Differential Geom. **30** (1989), no. 1, 223–301.

[118] Shi, Wan-Xiong. *Ricci deformation of the metric on complete noncompact Riemannian manifolds.* J. Differential Geom. **30** (1989), no. 2, 303–394.

[119] Simon, Miles. *A class of Riemannian manifolds that pinch when evolved by Ricci flow.* Manuscripta Math. **101** (2000), no. 1, 89–114.

[120] Singer, I. M. *Infinitesimally homogeneous spaces.* Comm. Pure Appl. Math. **13** 1960 685–697.

[121] Smoller, Joel. *Shock waves and reaction-diffusion equations.* Second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **258**. Springer-Verlag, New York, 1994.

[122] Thurston, William P. *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry.* Bull. Amer. Math. Soc. (N.S.) **6** (1982), no. 3, 357–381.

[123] Thurston, William P. *Three-dimensional geometry and topology.* Vol. 1. Edited by Silvio Levy. Princeton Mathematical Series, **35**. Princeton University Press, Princeton, NJ, 1997.

[124] Waldhausen, Friedholm. *Gruppen mit Zentrum und 3-dimensionale Mannigfaltigkeiten.* Topology **6** 1967 505–517.

[125] Witten, Edward. *String theory and black holes.* Phys. Rev. D (3) **44** (1991), no. 2, 314–324.

[126] Wolf, Joseph A. *Spaces of constant curvature.* Fifth edition. Publish or Perish, Inc., Houston, TX, 1984.

[127] Wu, Lang-Fang. *The Ricci flow on 2-orbifolds with positive curvature.* J. Differential Geom. **33** (1991), no. 2, 575–596.

[128] Wu, Lang-Fang. *The Ricci flow on complete* $\mathbb{R}^2$. Comm. Anal. Geom. **1** (1993), no. 3-4, 439–472.

[129] Wu, Lang-Fang. *A new result for the porous medium equation derived from the Ricci flow.* Bull. Amer. Math. Soc. (N.S.) **28** (1993), no. 1, 90–94.

[130] Yang, Deane. *Convergence of Riemannian manifolds with integral bounds on curvature. I.* Ann. Sci. École Norm. Sup. (4) **25** (1992), no. 1, 77–105.

[131] Yang, Deane. *Convergence of Riemannian manifolds with integral bounds on curvature. II.* Ann. Sci. École Norm. Sup. (4) **25** (1992), no. 2, 179–199.

Index