J. DIFFERENTIAL GEOMETRY
17 (1982) 255–306

THREE-MANIFOLDS WITH POSITIVE RICCI CURVATURE

RICHARD S. HAMILTON

INDEX

1. Introduction

Our goal in this paper is to prove the following result.

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

All manifolds of constant curvature have been completely classified by Wolf [6]. For positive curvature in dimension three there is a pleasant variety of examples, of which the best known are the lens spaces $L_{p,q}$. Wolf gives five

Received December 21, 1981.

different types. By our theorem, these are the only compact three-manifolds which can carry metrics of strictly positive Ricci curvature. This answers affirmatively a conjecture in Bourguignon [1].

It is known by a theorem of Myers (see Cheeger and Ebin [2]) that a compact manifold of strictly positive Ricci curvature has finite fundamental group, so its universal cover is also compact and simply connected. The Poincaré conjecture would imply that the universal cover is the sphere. Then one version of the Smith conjecture would imply that the group of covering transformations is conjugate to a group of isometries in the standard metric, and the original space would admit a metric of constant positive curvature. Thus if both these famous conjectures were known to be true, our result would follow immediately. On the other hand if either of them fails, then there will be a compact three-manifold with finite fundamental group which does not admit a metric of strictly positive Ricci curvature.

The product manifold $S^2 \times S^1$ has a metric of nonnegative Ricci curvature, with two eigenvalues $+1$ and the third 0. It does not admit any metric of constant curvature, and hence represents an obstruction to improving the result.

Our method of proof is inspired by the ideas of Eells and Sampson [3]. We start with any metric g_{ij} of strictly positive Ricci curvature R_{ij} and try to improve it by means of a heat equation. It would be natural to try to minimize an energy functional. Unfortunately we cannot form any geometrically meaningful quadratic expression in the first derivatives of the g_{ij}, since they always vanish in normal coordinates. It has been suggested to use the integral $\int R\, d\mu$ of the scalar curvature as an energy. This leads to the evolution equation (with $n = \dim X$)

$$\frac{\partial}{\partial t} g_{ij} = \frac{2}{n} R g_{ij} - 2R_{ij},$$

which unfortunately will not have solutions even for a short time, since it implies a backward heat equation in R. To eliminate this problem, we solve instead the evolution equation

$$\frac{\partial}{\partial t} g_{ij} = \frac{2}{n} r g_{ij} - 2R_{ij},$$

where r is the average of the scalar curvature R,

$$r = \int R\, d\mu \Big/ \int d\mu.$$

This equation always has a solution at least for a short time on any compact manifold of any dimension for any initial value of the metric at $t = 0$. This involves some work, for the equation is not strictly parabolic, as its linearization involves some zero eigenvalues in the symbol. (But at least they are not negative, as is the case for the first equation.) We prove this result using the Nash-Moser inverse function theorem.

It is worth noting that the degeneracies are there because the equation is invariant under the full diffeomorphism group of X. This has the interesting consequence that any isometries which exist in the metric to begin with are preserved as the metric evolves. Hence if the initial metric is homogeneous or symmetric then it remains so. For such spaces the evolution may be described by the change in a finite number of parameters. For example, on the product space $S^2 \times S^1$ the factor S^2 shrinks and the factor S^1 expands. Our normalization r is chosen so that the volume is always preserved. We also note that if X has a fixed complex structure and if the initial metric is Kähler, then it will remain so.

The rest of our results are peculiar to three dimensions. The essential simplification here is that the full Riemannian curvature tensor R_{ijkl} can be recovered from the Ricci tensor R_{ij}, which is much smaller and easier to analyze. However, we have not used the Sobolev inequality in a delicate way, so there is hope that the method may also yield some results in higher dimensions.

For a compact three-manifold, we prove that if the initial metric has strictly positive Ricci curvature, then it continues so for all time, and converges as $t \to \infty$ to a metric of constant positive curvature. The proof of this result requires three a priori estimates peculiar to this problem. The first shows the Ricci curvature remains positive, the second, shows the eigenvalues of the Ricci tensor at each point approach each other, and the third shows the gradient of the scalar curvature R goes to zero, so that we can compare the curvature at distant points. All three of these estimates are consequences of the maximum principle for parabolic equations. Once these estimates are established, we can control all the higher derivatives by some straightforward interpolation inequalities.

We would like to express our gratitude to the Harwood Foundation for a generous grant for the research in this paper, and to Professor O'Donnell for many inspirational remarks.

2. Notations and conventions

We will use the old-fashioned index notation for tensors, since it is well-adapted to the intense computations we must perform. We denote vectors as v^i, covectors as v_j, and mixed tensors as T^{ij}_{klm} etc. The summation convention will always hold. The Riemannian metric is g_{ij}, its inverse is g^{ij}, and the induced measure is $d\mu = \mu(x)dx$ where $\mu(x) = \sqrt{\det g_{ij}}$. The Levi-Civita connection is given by the Christöffel symbols

$$\Gamma^h_{ij} = \frac{1}{2} g^{hk}\left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ik} - \frac{\partial}{\partial x^k} g_{ij} \right),$$

and the Riemannian curvature tensor is

$$R^h_{ijk} = \frac{\partial}{\partial x^i}\Gamma^h_{jk} - \frac{\partial}{\partial x^j}\Gamma^h_{ik} + \Gamma^h_{ip}\Gamma^p_{jk} - \Gamma^h_{jp}\Gamma^p_{ik}.$$

We lower the index to the middle position, so that

$$R_{ijkl} = g_{hk} R^h_{ijl}.$$

Then R_{ijkl} is anti-symmetric in the pairs i, j and k, l and symmetric in their interchange, and satisfies a Bianchi identity on the cyclic permutation of any three. For the sphere we have

$$R(u, v, u, v) = R_{ijkl} u^i v^j u^k v^l > 0,$$

which is the opposite of the usual convention, but more symmetric. The Ricci curvature is the contraction

$$R_{ik} = g^{jl} R_{ijkl},$$

and on the sphere we have

$$R(u, u) = R_{ij} u^i u^j > 0,$$

which agrees with the usual convention. The scalar curvature $R = g^{ij} R_{ij}$. We denote the covariant derivative of a vector v^j by

$$\partial_i v^j = \frac{\partial}{\partial x^i} v^j + \Gamma^j_{ik} v^k,$$

and this definition extends uniquely to tensors so as to preserve the product rule and contractions. For the interchange of two covariant derivatives we have

$$\partial_i \partial_j v^h - \partial_j \partial_i v^h = R^h_{ijk} v^k,$$
$$\partial_i \partial_j v_k - \partial_j \partial_i v_k = R_{ijkl} g^{lm} v_m,$$

and similar formulas for more complicated tensors. To see how to convert from the old coordinate notation to the new coordinate-free notation the reader should consider the formulas

for a vector v: $v = v^i \partial/\partial x^i$,
for a covector L: $L = L_i dx^i$,
for a pairing: $L(v) = L_i v^i$,
for a tensor: $T(\partial/\partial x^j, \partial/\partial x^k) = T^i_{jk} \partial/\partial x^i$,

for a covariant derivative:

$$\partial_v T\left(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right) = v^h \partial_h T^i_{jk} \frac{\partial}{\partial x^i}.$$

For any tensor T such as T^i_{jk} we define its length $| T^i_{jk} |$ by

$$| T^i_{jk} |^2 = g_{il} g^{jm} g^{kn} T^i_{jk} T^l_{mn},$$

and we define its Laplacian ΔT by

$$\Delta T^i_{jk} = g^{pq} \partial_p \partial_q T^i_{jk},$$

the trace of the second iterated covariant derivative. We hope these remarks will aid the reader in following the paper.

3. The evolution equation

We consider the evolution equation on X^n

$$(**) \qquad \frac{\partial}{\partial t} g_{ij} = \frac{2}{n} r g_{ij} - 2R_{ij},$$

where $r = \int R\, d\mu / \int d\mu$ is the average scalar curvature. The factor r serves to normalize the equation so that the volume is constant. To see this we observe that if $d\mu = \mu(x) dx$ is the measure then $\mu = \sqrt{\det g_{ij}}$ and

$$\frac{\partial}{\partial t} \log \mu = \frac{1}{2} g^{ij} \frac{\partial}{\partial t} g_{ij} = r - R,$$

$$\frac{\partial}{\partial t} \int d\mu = \int (r - R)\, d\mu = 0.$$

Now it is awkward to have the normalizing factor present until we really need it. Therefore we will deal first with the unnormalized evolution equation

$$(*) \qquad \frac{\partial}{\partial t} g_{ij} = -2R_{ij},$$

which is easier to handle. The two equations differ only by a change of scale in space and a change of parametrization in time. To see this we let t, g_{ij}, R_{ij}, R, r denote the variables for the unnormalized equation $(*)$ and $\tilde{t}$, $\tilde{g}_{ij}$, $\tilde{R}_{ij}$, $\tilde{R}$, $\tilde{r}$ the corresponding variables for the normalized equation. To make the conversion from $(*)$ to $(**)$, we first choose the normalization factor $\psi = \psi(t)$ so that if $\tilde{g}_{ij} = \psi g_{ij}$ then $\int d\tilde{\mu} = 1$, so that the new manifold has measure 1. Then we choose a new time scale $\tilde{t} = \int \psi(t)\, dt$. It is easy to see that

$$\tilde{R}_{ij} = R_{ij}, \quad \tilde{R} = \tfrac{1}{\psi} R, \quad \tilde{r} = \tfrac{1}{\psi} r,$$

and $\int d\tilde{\mu} = 1$ so $\int d\mu = \psi^{-n/2}$. Arguing as before we have $(\partial/\partial t)\log\mu = -R$ and so

$$\frac{d}{dt}\log\int d\mu = -r, \quad \frac{d}{dt}\log\psi = \frac{2}{n}r.$$

Then it follows that

$$\frac{\partial}{\partial\tilde{t}}\tilde{g}_{ij} = \frac{\partial}{\partial t}g_{ij} + \left(\frac{d}{dt}\log\psi\right)g_{ij} = \frac{2}{n}\tilde{r}\tilde{g}_{ij} - 2\tilde{R}_{ij}.$$

It is worth noting that for a sphere S^n the normalized equation is constant, while the unnormalized equation shrinks to a point in a finite time.

4. Solution for a short time

Consider the evolution equation $\partial g_{ij}/\partial t = E(g_{ij})$ where E is the second order nonlinear partial differential operator $E(g_{ij}) = -2R_{ij}$. The linearization of this equatioin is $\partial\tilde{g}_{ij}/\partial t = DE(g_{ij})\tilde{g}_{ij}$ where DE is the derivative of E and $\tilde{g}_{ij}$ is the variation in g_{ij}. We must compute DE, but all we need is its symbol. This is obtained by taking the highest order derivatives and replacing $\partial/\partial x^i$ by the Fourier transform variable ζ_i.

The variation $\tilde{g}_{ij}$ in the metric produces a variation $\tilde{\Gamma}^h_{jk}$ in the connection, and this produces a variation $\tilde{R}^h_{ijk}$ in the curvature. Working in normal coordinates where $\Gamma^h_{jk} = 0$ at a point and using the formulas in §2, we see that

$$\tilde{\Gamma}^h_{jk} = \tfrac{1}{2}g^{hl}(\partial_j\tilde{g}_{kl} + \partial_k\tilde{g}_{jl} - \partial_l\tilde{g}_{jk}),$$
$$\tilde{R}^h_{ijk} = \partial_i\tilde{\Gamma}^h_{jk} - \partial_j\tilde{\Gamma}^h_{ik}.$$

Now an interchange of two covariant derivatives produces a lower order term. Also the Ricci curvature is given by $R_{jk} = R^i_{ijk}$. Then it is easy to compute

$$DE(g_{jk})\tilde{g}_{jk} = -2\tilde{R}_{jk} = g^{hi}\left\{\frac{\partial^2\tilde{g}_{jk}}{\partial x^h\partial x^i} - \frac{\partial^2\tilde{g}_{hj}}{\partial x^i\partial x^k} - \frac{\partial^2\tilde{g}_{ik}}{\partial x^h\partial x^j} + \frac{\partial^2\tilde{g}_{hi}}{\partial x^j\partial x^k}\right\} + \cdots$$

where the dots denote lower order terms. The symbol of the linear differential operator $DE(g_{jk})$ in the direction ζ_i is

$$\sigma DE(g_{jk})(\zeta_i)\tilde{g}_{jk} = g^{hi}\{\zeta_h\zeta_i\tilde{g}_{jk} - \zeta_i\zeta_k\tilde{g}_{hj} - \zeta_h\zeta_j\tilde{g}_{ik} + \zeta_j\zeta_k\tilde{g}_{hi}\}.$$

To see what the symbol does, we can always choose coordinates at a point so that $g_{jk} = \delta_{jk} = (1 \text{ if } j = k,\ 0 \text{ otherwise})$, and without loss of generality since the function is homogeneous we may assume ζ_i has length 1, and rotate so

$\zeta_1 = 1$ and $\zeta_i = 0$ for $i \neq 1$. Then the effect of σDE on a tensor T_{jk} is

$$[\sigma DE(g)(\zeta)T]_{jk} = T_{jk} \quad \text{if } j \neq 1, k \neq 1,$$
$$[\sigma DE(g)(\zeta)T]_{1k} = 0 \quad \text{if } k \neq 1,$$
$$[\sigma DE(g)(\zeta)T]_{11} = T_{22} + T_{33} + \cdots + T_{nn}.$$

The presence of the zero eigenvalues shows that the equation is not strictly parabolic. There is actually a good reason for the presence of these zero eigenvalues. The first way to see it is to consider the steady state equation $R_{ij} = 0$. If the evolution equation were parabolic, the steady state equation would be elliptic, and its solution space would be finite dimensional. But the solutions of $R_{ij} = 0$ (when they exist) are invariant under the full diffeomorphism group, which is infinite dimensional.

The second way is to recall the second contracted Bianchi identity, which tells us

$$g^{ij}\partial_i R_{jk} = \tfrac{1}{2}\partial_k R.$$

For any tensor T_{jk} we define the linear operator $L(g_{hi})$, depending on the metric g_{hi} and its connection, by

$$L(g_{hi})T_{jk} = g^{ij}\left(\partial_i T_{jk} - \tfrac{1}{2}\partial_k T_{ij}\right).$$

Note that L has degree 1 in g_{hi} and degree 1 in T_{jk}. If $E(g_{jk}) = -2R_{jk}$ then

$$L(g_{jk})E(g_{jk}) = 0.$$

Taking first variations, we see that

$$L(g_{jk})DE(g_{jk})\tilde{g}_{jk} + DL(g_{jk})\{E(g_{jk}), \tilde{g}_{jk}\} = 0.$$

Now the operator in $\tilde{g}_{jk}$ given by DL is only of degree 1, so its symbol of degree 3 is zero, and 3 is the degree of the other term $L \circ DE$, because L has degree 1 and DE has degree 2. Therefore

$$\sigma L(g_{jk})(\zeta_i) \circ \sigma DE(g_{jk})(\zeta_i)\tilde{g}_{jk} = 0,$$

and the image of $\sigma DE(g_{jk})$ must lie in the null space of $\sigma L(g_{jk})$. This symbol is

$$\sigma L(g_{jk})(\zeta_i)T_{jk} = g^{ij}\left(\zeta_i T_{jk} - \tfrac{1}{2}\zeta_k T_{ij}\right).$$

Normalizing g_{jk} and ζ_i as before we have

$$[\sigma L(g)(\zeta)T]_k = T_{1k} \quad \text{if } k \neq 1,$$
$$[\sigma L(g)(\zeta)T]_1 = \tfrac{1}{2}(T_{11} - T_{22} - T_{33} - \cdots - T_{nn}).$$

The null space of $\sigma L(g)(\zeta)$ consists of all those symmetric tensors T_{jk} with $T_{11} = T_{22} + T_{33} + \cdots + T_{nn}$ and $T_{12} = T_{13} = \cdots = T_{1n} = 0$. It is clear that $\sigma DE(g)(\zeta)$ lies in this space. We can also see the following result.

4.1 Lemma. *The symbol $\sigma DE(g)(\zeta)$ acts as multiplication by $|\zeta|^2$ on the null space of the symbol $\sigma L(g)(\zeta)$.*

This shows that there are no degeneracies other than those implied by the second contracted Bianchi identity. The following theorem is then an immediate consequence of the general result in the next section.

4.2 Theorem. *The evolution equation $\partial g_{ij}/\partial t = -2 R_{ij}$ has a solution for a short time on any compact Riemannian manifold with any initial metric at $t = 0$.*

5. Evolution equations with an integrability condition

We shall consider evolution equations

$$\frac{\partial f}{\partial t} = E(f),$$

where $E(f)$ is a nonlinear differential operator of degree 2 in f. We suppose f belongs to an open set U in a vector bundle F over a compact manifold X, and $E(f)$ takes its values in F also. Then E is a smooth map

$$E: \mathcal{C}^\infty(X, U) \subseteq \mathcal{C}^\infty(X, F) \to \mathcal{C}^\infty(X, F)$$

of an open set in a Fréchet space to itself. In studying the evolution equation it is important to consider its linearization. Letting $\tilde{f}$ denote a variation in f, we get

$$\frac{\partial \tilde{f}}{\partial t} = DE(f)\tilde{f},$$

where the derivative $DE(f)\tilde{f}$ is a linear differential operator in $\tilde{f}$ of degree 2. We say E is parabolic if its linearization is parabolic around any f. This can be expressed in terms of the symbol $\sigma DE(f)(\xi)$, which is obtained by replacing each derivative $\partial/\partial x^j$ by ξ_j in the highest order terms. (For simplicity we omit the factor $i = \sqrt{-1}$.) If in local coordinates

$$\partial f^\alpha/\partial t = E^\alpha\left(x^i, f^\beta, f_i^\beta, f_{ij}^\beta\right),$$

then the symbol of $DE(f)$ is

$$\sigma DE(f)(\xi) = \frac{\partial E^\alpha}{\partial f_{ij}^\beta}\left(x^i, f^\beta, f_i^\beta, f_{ij}^\beta\right)\xi_i\xi_j.$$

The symbol is an automorphism of the vector bundle F to itself. Then $DE(f)$ is parabolic if all the eigenvalues of $\sigma DE(f)(\xi)$ have strictly positive real parts when $\xi \neq 0$. In this case it is well known that the evolution equation $\partial f/\partial t = E(f)$ has a unique smooth solution for the initial value problem $f = f_0$ at $t = 0$ for at least a short time interval $0 \leqslant t \leqslant \varepsilon$ (where ε may depend on f_0).

We shall consider problems where some of the eigenvalues of $\sigma DE(f)(\xi)$ are zero. This happens when $E(f)$ satisfies an integrability condition. Let $g = L(f)h$ be a differential operator of degree 1 on sections $f \in U \subseteq F$ and $h \in F$ with values g in another vector bundle G over X, such that the operator $Q(f) = L(f)E(f)$ only has degree at most 1 in f. We call $L(f)$ the integrability condition for $E(f)$. Taking a variation $\tilde{f}$ in f we see that

$$L(f)DE(f)\tilde{f} + DL(f)\{E(f), \tilde{f}\} = DQ(f)\tilde{f}.$$

Now the operators $DL(f)\{E(f), \tilde{f}\}$ and $DQ(f)\tilde{f}$ only have degree 1 in $\tilde{f}$, and hence the operator $L(f)DE(f)\tilde{f}$ also has degree 1 only. Therefore taking the symbols, $\sigma L(f)(\xi) \cdot \sigma DE(f)(\xi) = 0$. From this we see that

$$\operatorname{Im} \sigma DE(f)(\xi) \subseteq \operatorname{Null} \sigma L(f)(\xi).$$

If L is not trivial then $\sigma DE(f)(\xi)$ must have a null eigenspace. The most we can hope is that the restriction of $\sigma DE(f)(\xi)$ to Null $\sigma L(f)(\xi)$ is positive. We shall prove the following result.

5.1 Theorem. *Let $\partial f/\partial t = E(f)$ be an evolution equation with integrability condition $L(f)$. Suppose that*

(A) *$L(f)E(f) = Q(f)$ has degree* 1,

(B) *all the eigenvalues of the eigenspaces of $\sigma DE(f)(\xi)$ in* Null $\sigma L(f)(\xi)$ *have strictly positive real parts.*

Then the initial value problem $f = f_0$ at $t = 0$ has a unique smooth solution for a short time $0 \leqslant t \leqslant \varepsilon$ where ε may depend on f_0.

Proof. We shall use the Nash-Moser inverse function theorem (see [5] for a complete exposition by the author). We shall show that if $\partial \bar{f}/\partial t - E(\bar{f}) = \bar{h}$ is a solution of the evolution equation on $0 \leqslant t \leqslant 1$, with $\bar{f} = \bar{f}_0$ at $t = 0$, then for any f_0 near $\bar{f}_0$ and h near $\bar{h}$ there exists a unique solution of the equation $\partial f/\partial t - E(f) = h$ over the interval $0 \leqslant t \leqslant 1$ with $f = f_0$ at $t = 0$. To see that this implies the theorem, choose $\bar{f}$ to be any function whose formal Taylor series at $t = 0$ is what it must be to solve $\partial f/\partial t = E(f)$ with $f = f_0$ at $t = 0$, and let $\bar{h} = \partial \bar{f}/\partial t - E(\bar{f})$. Then the formal Taylor series of $\bar{h}$ at $t = 0$ is identically zero. By translating $\bar{h}$ a little, we can find h arbitrarily close to $\bar{h}$ and vanishing for a short time $0 \leqslant t \leqslant \varepsilon$. Then the solution of $\partial f/\partial t - E(f) = h$ with $f = f_0$ at $t = 0$ solves the equation up to time ε.

We can apply the Nash-Moser inverse function theorem to the operator

$$\mathcal{E}: \mathcal{C}^\infty(X \times [0,1], F) \to \mathcal{C}^\infty(X \times [0,1], F) \times \mathcal{C}^\infty(X, F),$$
$$\mathcal{E}(f) = (\partial f/\partial t - E(f), f \mid \{t = 0\}).$$

Its derivative is the operator

$$D\mathfrak{E}(f)\tilde{f} = \left(\partial\tilde{f}/\partial t - DE(f)\tilde{f},\ \tilde{f}\,|\,\{t=0\}\right).$$

We must show that the linearized equation $\partial\tilde{f}/\partial t - DE(f)\tilde{f} = \tilde{h}$ has a unique solution for the initial value problem $\tilde{f} = \tilde{f}_0$ at $t = 0$, and verify that the solution $\tilde{f}$ is a smooth tame function of $\tilde{h}$ and $\tilde{f}_0$.

We make the substitution $\tilde{g} = L(f)\tilde{f}$. Then $\tilde{g}$ will satisfy the evolution equation

$$\frac{\partial\tilde{g}}{\partial t} = L(f)\frac{\partial\tilde{f}}{\partial t} + DL(f)\left\{\tilde{f}, \frac{\partial f}{\partial t}\right\}.$$

However $\partial\tilde{f}/\partial t = DE(f)\tilde{f} + \tilde{h}$. Moreover differentiating the integrability condition $L(f)E(f) = Q(f)$ we get

$$L(f)DE(f)\tilde{f} + DL(f)\{E(f), \tilde{f}\} = DQ(f)\tilde{f}.$$

Then we get the equation

$$\frac{\partial\tilde{g}}{\partial t} - M(f)\tilde{f} = \tilde{k}$$

where $\tilde{k} = L(f)\tilde{h}$ and

$$M(f)\tilde{f} = DL(f)\left\{\tilde{f}, \frac{\partial f}{\partial t}\right\} - DL(f)\{E(f), \tilde{f}\} + DQ(f)\tilde{f}$$

is a linear differential operator in $\tilde{f}$ of degree 1 whose coefficients depend smoothly on f and its derivatives (possibly of degree 3 in space, or 1 in space and 1 in time).

If we choose a measure on X and inner products on the vector bundles F and G, we can form a differential operator $L^*(f)g = h$, of degree 1 in f and g, which is the adjoint of $L(f)$. Let us write

$$P(f) = DE(f) + L^*(f)L(f).$$

We claim that the equation $\partial\tilde{f}/\partial t = P(f)\tilde{f}$ is parabolic. To see this, we must examine the symbol

$$\sigma P(f)(\xi) = \sigma DE(f)(\xi) + \sigma L^*(f)(\xi)\cdot\sigma L(f)(\xi).$$

Suppose v is an eigenvector in F with eigenvalue λ. Then $\sigma P(f)(\xi)v = \lambda v$. But $\sigma L(f)(\xi)\cdot\sigma DE(f)(\xi) = 0$, so

$$\sigma L(f)(\xi)\cdot\sigma L^*(f)(\xi)\cdot\sigma L(f)(\xi)v = \lambda\sigma L(f)(\xi)v.$$

It follows that

$$|\sigma L^*(f)(\xi)\cdot\sigma L(f)(\xi)v|^2 = \lambda|\sigma L(f)(\xi)v|^2.$$

Now if $\sigma L^*(f)(\xi) \cdot \sigma L(f)(\xi)v = 0$ then $\sigma L(f)(\xi)v = 0$, and otherwise λ is real and strictly positive. When $\sigma L(f)(\xi)v = 0$ then $\sigma DE(f)(\xi)v = \lambda v$, and λ has strictly positive real part by our hypothesis (B). Thus $P(f)$ is parabolic.

We proceed to solve the system of equations

$$\frac{\partial \tilde{f}}{\partial t} - P(f)\tilde{f} + L^*(f)\tilde{g} = \tilde{h},$$
$$\frac{\partial \tilde{g}}{\partial t} - M(f)\tilde{f} = \tilde{k}$$

for the unknown functions $\tilde{f}$ and $\tilde{g}$ for given $\tilde{h}$ and $\tilde{k}$ and given f, with initial data $\tilde{f} = \tilde{f}_0$ and $\tilde{g} = \tilde{g}_0 = L(f_0)\tilde{f}_0$ at $t = 0$.

It follows from the theorem in the next section that the solution $(\tilde{f}, \tilde{g})$ exists and is unique, and is a smooth tame function of $(f, \tilde{h}, \tilde{k}, \tilde{f}_0, \tilde{g}_0)$. Then putting $\tilde{l} = \tilde{g} - L(f)\tilde{f}$ we see that $\tilde{l}$ satisfies the evolution equation

$$\frac{\partial \tilde{l}}{\partial t} = L(f)L^*(f)\tilde{l},$$

and $\tilde{l} = 0$ at $t = 0$. But then the obvious integral inequality

$$\frac{d}{dt}\int_X |\tilde{l}|^2\, d\mu + 2\int_X |L^*(f)\tilde{l}|^2\, d\mu = 0$$

proves that $\tilde{l} = 0$. Then it follows that $\partial\tilde{f}/\partial t - DE(f)\tilde{f} = \tilde{h}$. This completes the proof of the theorem, except for the result of the following section.

6. Weakly parabolic linear systems

Let X be a compact manifold and let F and G be vector bundles over X. We consider a system of linear evolution equations on $0 \leq t \leq T$ for sections f of F and g of G

$$\frac{\partial f}{\partial t} = Pf + Lg + h, \quad \frac{\partial g}{\partial t} = Mf + Ng + k,$$

where P, L, M and N are linear differential operators involving only space derivatives whose coefficients are smooth functions of both space and time. We assume P has degree 2, L and M have degree 1, and N has degree 0.

6. Theorem. *Suppose the equation $\partial f/\partial t = Pf$ is parabolic. Then for any given (f_0, g_0, h, k) there exists a unique smooth solution (f, g) of the system with $f = f_0$ and $g = g_0$ at $t = 0$.*

Proof. We can use the equation to solve formally for the Taylor series of f and g at $t = 0$. Choose functions $\bar{f}$ and $\bar{g}$ with the given Taylor series, and subtract them off from f and g. This reduces us to the case where f, g, h, k are all known to vanish for $t \leq 0$. We can then use the following regularization device.

We introduce a time lag $\delta > 0$ into the second equation, so that

$$\left(\frac{\partial g}{\partial t}\right)_{t+\delta} = (Mf + Ng + h)_t.$$

The resulting system clearly has a unique smooth solution on $0 \leq t \leq T$, for we can alternatively use the first and second equations separately to advance the solution on intervals of length $\delta > 0$. In the sequel we shall derive a priori estimates for the solutions f and g of the evolutionary system. These estimates also clearly hold for the delayed system and are independent of $\delta \to 0$. We leave the necessary modification to the reader. Then by passing to a convergent subsequence we get a solution for $\delta = 0$.

We turn to the a priori estimates. We introduce the following norms. For a section of F (or G) over X we let $|f|_n$ measure the L_2 norm of f and its derivatives up to degree n. For a time-dependent section f over $X \times [0, T]$ with $f = \{f_t: 0 \leq t \leq T\}$ we put

$$|f|_n^2 = \int_0^T |f_t|_n^2 \, dt,$$

so that $|f|_n$ measures space derivatives of degree $\leq n$ only. Then we put

$$\|f\|_n^2 = \sum_{2j \leq n} \|(\partial/\partial t)^j f\|_{n-2j},$$

which is a weighted norm counting one time derivative equal to two space derivatives. (We caution the reader that this weighted grading is not tamely equivalent to the usual one.) The differential operators P, L, M, N are all sections of some appropriate bundles over X, which could be interpreted in terms of jet bundles. We measure P, L, M, N in terms of norms $|[L]|_n$ where $[L]_n$ measures the supremum of L and its space derivatives up to degree n, and $|[\ ]|_n$ is the corresponding norm counting one time derivative equal to two space derivatives as before. (Note that the gradings $\|\ \|_n$ and $|[\ ]|_n$ are tamely equivalent. Also from a point of view of tamely equivalent gradings it does not really matter that for odd n our grading $\|\ \|_n$ has missed $\frac{1}{2}$ of a time derivative, compared to the usual one for parabolic equations. This allows us to avoid the nuisance of discussing fractional derivatives.)

6.2 Theorem. *Let the solution (f, g) of the system of evolution equations be written as a function*

$$(f, g) = S(P, L, M, N, h, k, f_0, g_0)$$

of the coefficients P, L, M, N, the data h, k and the initial values f_0, g_0. In the open set where P is parabolic the solution S is a smooth tame map in the gradings $\| \ \|_n$ on f, g, h, k and $| \ |_n$ on f_0 and g_0 and $|[\ ]|_n$ on P, L, M, N.

We shall prove these theorems by a sequence of lemmas.

6.3 Lemma. *If $\partial f/\partial t - Pf = h$ on $0 \leqslant t \leqslant T$ and $f = 0$ at $t = 0$ then we can find a constant C independent of θ such that for $0 \leqslant \theta \leqslant T$*

$$\int_0^\theta |f_t|_2^2\,dt \leqslant C\int_0^\theta |h_t|_0^2\,dt.$$

Proof. When $\theta = T$ this is a standard result for parabolic equations (see [4]). To see that C is independent of θ for $0 \leqslant \theta \leqslant T$ we use the following device. We extend P to be parabolic on the interval $-T \leqslant t \leqslant T$. Note that we may assume all the derivatives of f vanish at $t = 0$ also, for the set of such functions is dense in those with $f_0 = 0$ in the norm $\| \ \|_2$. Then we may extend f smoothly to be zero for $-T \leqslant t \leqslant 0$. Now we consider translations by $T - \theta$ of the original equation. Then P and f on $-T + \theta \leqslant t \leqslant \theta$ correspond to their translates on $0 \leqslant t \leqslant T$. Since the estimate above is coercive for P, it follows by the usual argument that the same constant C works for all operators in a neighborhood of P. Hence we can make one constant C work for any compact set of parabolic operators P. But the set of translates is compact, so the lemma follows.

6.4 Corollary. *If $\partial f/\partial t - Pf = h$ on $0 \leqslant t \leqslant T$ and $f = f_0$ at $t = 0$ then we can find a constant C independent of θ such that for $0 \leqslant \theta \leqslant T$*

$$\int_0^\theta |f_t|_2^2\,dt \leqslant C\int_0^\theta |h_t|_0^2\,dt + C|f_0|_1^2.$$

Proof. The norm $|f_0|_1$ is equivalent to the quotient norm $\inf\{\|f\|_2 \colon f = f_0$ at $t = 0\}$. It suffices to check this in local coordinates, where we can use the Fourier transform. Given $f_0(x)$ on $t = 0$, we define the extension $f(x, t)$ by letting

$$f(\xi, \tau) = \psi\left(\frac{\tau}{1 + |\xi|^2}\right) \cdot \frac{1}{1 + |\xi|^2}\hat{f}_0(\xi),$$

where $\psi(\tau)$ is a smooth function of compact support with $\int \psi(\tau)\,d\tau = 1$. Then f extends f_0 and if $\theta = \tau/(1 + |\xi|^2)$

$$\begin{aligned}\|f\|_2^2 &= \int\int\left(1 + |\xi|^2 + |\tau|\right)^2 |\hat{f}(\xi, \tau)|^2\,d\xi d\tau \\ &= \int|\psi(\theta)|^2(1 + |\theta|)^2\,d\theta \cdot \int\left(1 + |\xi|^2\right)|\hat{f}_0(\xi)|^2\,d\xi \\ &\leqslant C|f_0|_1^2.\end{aligned}$$

Conversely given on $t \geqslant 0$ with $f = f_0$ at $t = 0$, we can first extend f to $t \leqslant 0$ without increasing $\| f \|_2$ by more than a factor. Then using the Fourier transform, since

$$\hat{f}_0(\xi) = \int \hat{f}(\xi, \tau)\, d\tau$$

we have

$$\begin{aligned} |f_0|_1 &= \sup_{|g|_0 \leqslant 1} \int \left(1 + |\xi|^2\right)^{1/2} \hat{f}_0(\xi) \hat{g}(\xi)\, d\xi \\ &= \sup_{|g|_0 \leqslant 1} \int\int \left(1 + |\xi|^2\right)^{1/2} \hat{f}(\xi, \tau) \hat{g}(\xi)\, d\xi d\tau \\ &\leqslant \sup_{|g|_0 \leqslant 1} \left\{ \int\int \left(1 + |\xi|^2 + |\tau|\right)^2 |\hat{f}(\xi, \tau)|^2\, d\xi d\tau \right\}^{1/2} \\ &\quad \times \left\{ \int\int \frac{\left(1 + |\xi|^2\right) |\hat{g}(\xi)|^2}{\left(1 + |\xi|^2 + |\tau|\right)^2} d\xi d\tau \right\}^{1/2}. \end{aligned}$$

Now the first integral is bounded by $\| f \|_2^2$, and the second is bounded by a constant, since

$$\int \frac{1 + |\xi|^2}{\left(1 + |\xi|^2 + |\tau|\right)^2} d\tau = \int \frac{d\theta}{\left(1 + |\theta|\right)^2} = C < \infty.$$

Thus $|f_0|_1^2 \leqslant C \| f \|_2^2$, proving our assertion.

We can combine the extension operators in local coordinates to produce a linear extension operator $\mathcal{C}^\infty(X, F) \to \mathcal{C}^\infty(X \times [0, T], F)$ such that if f^* is the extension of f_0 then $f^* = f_0$ at $t = 0$ and $\| f^* \|_2 \leqslant C |f_0|_1$. Now given f satisfying $\partial f / \partial t - Pf = h$ and $f = f_0$ at $t = 0$, let f^* be the extension of f_0 constructed above and let $\partial f^* / \partial t - Pf^* = h^*$. If $f = f^* + \tilde{f}$ and $h = h^* + \tilde{h}$ then $\partial \tilde{f} / \partial t - P\tilde{f} = \tilde{h}$, and $\tilde{f} = 0$ at $t = 0$. We can then estimate $\tilde{f}$ using the previous lemma, so

$$\int_0^\theta |f_t|_0^2\, dt \leqslant \int_0^\theta |f_t^*|_0^2\, dt + \int_0^\theta |\tilde{f}_t|_0^2\, dt,$$

$$\int_0^\theta |f_t^*|_0^2\, dt \leqslant \int_0^T |f_t^*|_0^2\, dt \leqslant C \| f^* \|_2^2 \leqslant C |f_0|_1^2,$$

$$\int_0^\theta |\tilde{f}_t|_0^2\, dt \leqslant C \int_0^\theta |\tilde{h}_t|_0^2\, dt,$$

$$\int_0^\theta |\tilde{h}_t|_0^2\, dt \leqslant \int_0^\theta |h_t|_0^2\, dt + \int_0^\theta |h_t^*|_0^2\, dt,$$

$$\int_0^\theta |h_t^*|_0^2\,dt \leqslant \int_0^T |h_t^*|_0^2\,dt \leqslant C\|h^*\|_0^2 \leqslant C\|f^*\|_2^2 \leqslant C|f_0|_1^2.$$

Combining these estimates the result follows.

6.5 Lemma. *If $\partial g/\partial t = k$, then for $0 \leqslant \theta \leqslant T$*

$$|g_\theta|_1^2 \leqslant C\int_0^\theta |k_t|_1^2\,dt + C|g_0|_1^2.$$

Proof. Since

$$g_\theta = g_0 + \int_0^\theta k_t\,dt$$

and every norm is convex, we have

$$|g_\theta|_1 \leqslant |g_0|_1 + \int_0^\theta |k_t|_1\,dt.$$

But we also have

$$\left(\int_0^\theta |k_t|_1\,dt\right)^2 \leqslant \theta\int_0^\theta |k_t|_1^2\,dt.$$

Therefore the above estimate holds with a constant C independent of θ for $0 \leqslant \theta \leqslant T$.

Note that if there is a delay δ in time, so that $(\partial g/\partial t)_{t+\delta} = k_t$, and if g and k vanish for $t \leqslant 0$, then g also vanishes for $t \leqslant \delta$, and we have a better estimate

$$|g_{\theta+\delta}|_1^2 \leqslant C\int_{t=0}^\theta |k_t|_1^2\,dt,$$

for $0 \leqslant \theta \leqslant T - \delta$.

Now we assume f and g are solutions of the system of evolution equations $\partial f/\partial t = Pf + Lg + h$ and $\partial g/\partial t = Mf + Ng + k$ with $f = f_0$ and $g = g_0$ at $t = 0$. To simplify the following formulas we let

$$E = |h|_0 + |k|_1 + |f_0|_1 + |g_0|_1.$$

6.6 Lemma. *We have estimates for $0 \leqslant \theta \leqslant T$*

$$\int_0^\theta |f_t|_2^2\,dt \leqslant C\int_0^\theta |g_t|_1^2\,dt + CE^2,$$

$$|g_\theta|_1^2 \leqslant C\int_0^\theta \left(|f_t|_2^2 + |g_t|_1^2\right)dt + CE^2.$$

Proof. We apply our two previous estimates, replacing h by $Lg + h$ and k by $Mf + Ng + k$. Then

$$|Lg_t + h_t|_0 \leqslant C(|g_t|_1 + |h_t|_0),$$
$$|Mf_t + Ng_t + k_t|_1 \leqslant C(|f_t|_2 + |g_t|_1 + |k_t|_1)$$

and the result follows directly.

6.7 Corollary. *A solution of the system of evolution equations satisfies the a priori estimate*

$$|f|_2 + |g|_1 \leqslant C(|h|_0 + |k|_1 + |f_0|_1 + |g_0|_1).$$

Proof. By the above estimates we have

$$|g_\theta|_1^2 \leqslant C\int_0^\theta |g_t|_1^2\,dt + CE^2.$$

Then for any $\lambda > 0$ we have

$$\lambda\int_{\theta=0}^T e^{-\lambda\theta}|g_\theta|_1^2\,d\theta \leqslant C\int_{t=0}^T \left\{\int_{\theta=t}^T \lambda e^{\lambda(t-\theta)}\,d\theta\right\} e^{-\lambda t}|g_t|_1^2\,dt$$
$$+C\left\{\int_{\theta=0}^T \lambda e^{-\lambda\theta}\,d\theta\right\}E^2,$$

and since the bracketed integrals are $\leqslant 1$, we have

$$(\lambda - C)\int_{t=0}^T e^{-\lambda t}|g_t|_1^2\,dt \leqslant CE^2$$

with a constant C independent of λ. When $\lambda > C$ we get $|g|_1^2 \leqslant CE^2$. Then $|f|_2^2 \leqslant CE^2$ also.

Note that if there is a time delay δ, so that

$$(\partial g/\partial t)_{t+\delta} = (Mf + Ng + h)_t,$$

then we get a better estimate

$$|g_{\theta+\delta}|_1^2 \leqslant C\int_0^\theta |g_t|_1^2\,dt + CE^2,$$

and since

$$\int_{\theta=0}^{T-\delta} e^{-\lambda\theta}|g_{\theta+\delta}^2|_1\,d\theta = e^{\lambda\delta}\int_{\theta=0}^T e^{-\lambda\theta}|g_\theta|_1^2\,d\theta,$$

the same argument yields the same estimate with a constant independent of δ as $\delta \to 0$.

Next we show the same low-norm a priori estimate holds uniformly in a neighborhood of a given system. Fix operators $\overline{P}$, $\overline{L}$, $\overline{M}$, $\overline{N}$ and consider all operators P, L, M, N in a neighborhood

$$[P - \overline{P}]_0 + [L - \overline{L}]_0 + [M - \overline{M}]_1 + [N - \overline{N}]_1 \leqslant \delta.$$

If $\overline{P}$ is parabolic, and $\delta > 0$ is small enough, then so is P.

6.8 Lemma. *If $\delta > 0$ is small enough then for all systems P, L, M, N in the given neighborhood the a priori estimate*

$$|f|_2 + |g|_1 \leqslant C(|h|_0 + |k|_0 + |f_0|_1 + |g_0|_1)$$

holds with a fixed constant C.

Proof. If f and g solve the system of evolution equations for P, L, M, N, then

$$\frac{\partial f}{\partial t} = \bar{P}f + \bar{L}g + (P - \bar{P})f + (L - \bar{L})g + h,$$

$$\frac{\partial g}{\partial t} = \bar{M}f + \bar{N}g + (M - \bar{M})f + (N - \bar{N})g + k.$$

Applying the estimate for the fixed system $\bar{P}$, $\bar{L}$, $\bar{M}$, $\bar{N}$ we get

$$|f|_2 + |g|_1 \leqslant C(|(P - \bar{P})f + (L - \bar{L})g + h|_0 + |(M - \bar{M})f + (N - \bar{N})g + k|_1 + |f_0|_1 + |g_0|_1),$$

$$|f|_2 + |g|_1 \leqslant C\delta(|f|_2 + |g|_1) + C(|h|_0 + |k|_1 + |f_0|_1 + |g_0|_1).$$

When $\delta > 0$ is sufficiently small, the estimate follows.

We can now estimate higher space derivatives in the usual way by differentiating through the equation. Choose connections in the vector bundles F and G and let $\partial_v f$ denote the covariant derivative of the section f in the direction of a vector field v. There is then a natural way to define the covariant derivative of the linear differential operators so that (for example)

$$\partial_v(Lf) = L(\partial_v f) + (\partial_v L)f.$$

Note that $\partial_v L$ will be a differential operator of the same degree as L, formed by allowing the derivatives to fall on the coefficients. We will let Σ_v denote the sum over a finite number of vector fields which span the tangent space at each point of X.

6.9 Lemma. *For all solutions of all systems in the δ-neighborhood given before we have a priori estimates for all $n \geqslant 0$ of the form*

$$|f|_{n+2} + |g|_{n+1} \leqslant C(|h|_n + |k|_{n+1} + |f_0|_{n+1} + |g_0|_{n+1}) + C([P]_n + [L]_n + [M]_{n+1} + [N]_{n+1}) \times (|h|_0 + |k|_1 + |f_0|_1 + |g_0|_1).$$

Proof. This holds for $n = 0$. We proceed by induction. Suppose the estimate above holds up to some n. Differentiating through the equation, we have

$$\frac{\partial}{\partial t}\partial_v f = P\partial_v f + L\partial_v g + (\partial_v P)f + (\partial_v L)g + \partial_v h,$$

$$\frac{\partial}{\partial t}\partial_v g = M\partial_v f + N\partial_v g + (\partial_v M)f + (\partial_v N)g + \partial_v k.$$

For simplicity we write

$$A_n = |f|_{n+2} + |g|_{n+1},$$
$$B_n = [P]_n + [L]_n + [M]_{n+1} + [N]_{n+1},$$
$$E_n = |h|_n + |k|_{n+1} + |f_0|_{n+1} + |g_0|_{n+1},$$

in terms of which the induction hypothesis is $A_n \leqslant C(E_n + B_n E_0)$, and $B_0 \leqslant \delta \leqslant C$. Applying the induction hypothesis to the derived equation, we have

$$|\partial_v f|_{n+2} + |\partial_v g|_{n+1} \leqslant C(E_{n+1} + B_1 A_n + B_{n+1} A_0).$$

Now $A_0 \leqslant CE_0$ and $A_n \leqslant C(E_n + B_n E_0)$. Moreover by interpolation we have

$$B_1 B_n \leqslant CB_0 B_{n+1} \leqslant CB_{n+1},$$
$$B_1 E_n \leqslant C(B_0 E_{n+1} + B_{n+1} E_0),$$

and hence

$$A_{n+1} = |f|_{n+3} + |g|_{n+2} \leqslant \sum_v |\partial_v f|_{n+2} + |\partial_v g|_{n+1} + A_n$$
$$\leqslant C(E_{n+1} + B_{n+1} E_0),$$

which completes the induction.

Finally we can estimate time derivatives also simply by using the equations. We get the following result for the weighted gradings $\| \ \|_n$ and $|[\ ]|_n$ defined earlier, in which one time derivative counts for two space derivatives.

6.10 Lemma. *For all solutions of all systems in the δ-neighborhood given before we have a priori estimates for all $n \geqslant 0$ of the form*

$$\|f\|_{n+2} + \|g\|_{n+1} \leqslant C(\|h\|_n + \|k\|_{n+1} + |f_0|_{n+1} + |g_0|_{n+1})$$
$$+ C(|[P]|_n + |[L]|_n + |[M]|_{n+1} + |[N]|_{n+1})$$
$$\times (\|h\|_0 + \|k\|_1 + |f_0|_1 + |g_0|_1).$$

Proof. We must estimate the terms

$$|(\partial/\partial t)^j f|_{n-2j+2} + |(\partial/\partial t)^j g|_{n-2j+1}$$

for $2j \leqslant n$. We can do this for $j = 0$ as before. We proceed by induction on j. Suppose we have estimates up to some value of j. Then for the $j + 1$ terms we have

$$|(\partial/\partial t)^{j+1} f|_{n-2j} = |(\partial/\partial t)^j (Pf + Lg + h)|_{n-2j},$$
$$|(\partial/\partial t)^{j+1} g|_{n-2j-1} = |(\partial/\partial t)^j (Mf + Ng + k)|_{n-2j-1},$$

and by interpolation we need only consider the extreme cases where all the derivatives in both space and time fall entirely on P, L, M, N or entirely on f, g, h, k.

For the first terms we get

$$|(\partial/\partial t)^j f|_{n-2j+2} + |[P]|_n |f|_2 + |(\partial/\partial t)^j g|_{n-2j+1} + |[L]|_n |g|_1 + \|h\|_n$$

and for the second terms we get

$$\left|\left(\frac{\partial}{\partial t}\right)^j f\right|_{n-2j} + |[M]|_{n-1}|f|_1 + \left|\left(\frac{\partial}{\partial t}\right)^j g\right|_{n-2j-1} + |[N]|_{n-1}|g|_0 + \|k\|_{n-1},$$

which is even better than we need. The above can be bounded by terms

$$|(\partial/\partial t)^j f|_{n-2j+2} + |(\partial/\partial t)^j g|_{n-2j+1} + \|h\|_n + \|k\|_{n+1}$$
$$+ (|[P]|_n + |[L]|_n + |[M]|_{n+1} + |[N]|_{n+1})(|f|_2 + |g|_1).$$

We apply the induction hypothesis to the first part and our previous estimate to $|f|_2 + |g|_1$. This proves the lemma.

If the second equation contains a delay δ in time, we can still differentiate through the equation with respect to space or time, and the derived equation has the same form with the same delay. Hence the estimates in Lemmas 6.9 and 6.10 still hold with a constant C independent of δ as $\delta \to 0$. To prove existence for a single equation we do not have to keep track of how the constant depends on the coefficients P, L, M, N.

Now the last lemma clearly is a tame estimate on the solution map

$$(f, g) = \mathcal{S}(P, L, M, N, h, k, f_0, g_0)$$

in the weighted gradings. It follows that $\mathcal{S}$ is continuous, since the spaces $\mathcal{C}^\infty(X, F)$ and the others are all Montel spaces. Then it also follows that all the derivatives of $\mathcal{S}$ are tame also, by the formula for the derivative of an inverse.

7. Evolution of the curvature

The evolution equation $\partial g_{ij}/\partial t = -2R_{ij}$ for the metric implies a heat equation for the Riemannian curvature R_{ijkl} which we now derive. This equation will be the basis for all our a priori estimates on the evolution of the curvature. Recall we define

$$\Delta R_{ijkl} = g^{pq}\partial_p\partial_q R_{ijkl}.$$

Various second order derivatives of the curvature tensor are likely to differ by terms quadratic in the curvature tensor. To this end we introduce the tensors

$$B_{ijkl} = g^{pr}g^{qs}R_{piqj}R_{rksl}.$$

Note we have the obvious symmetries

$$B_{ijkl} = B_{jilk} = B_{klij}$$

but the other symmetries of the curvature tensor R_{ijkl} may fail to hold for B_{ijkl}.

7.1 Theorem. *The curvature tensor satisfies the evolution equation*

$$\frac{\partial}{\partial t}R_{ijkl} = \Delta R_{ijkl} + 2(B_{ijkl} - B_{ijlk} - B_{iljk} + B_{ikjl})$$
$$-g^{pq}(R_{pjkl}R_{qi} + R_{ipkl}R_{qj} + R_{ijpl}R_{qk} + R_{ijkp}R_{ql}).$$

Proof. Letting a prime denote differentiation with respect to time t, we see by considering the formulas for Γ_{jk}^i and R_{ijkl} in normal coordinates that for any evolution of a metric g_{ij} we have

$$\Gamma_{jl}'^h = \tfrac{1}{2}g^{hm}(\partial_j g'_{lm} + \partial_l g'_{jm} - \partial_m g'_{jl}),$$
$$R_{ijl}'^h = \partial_i \Gamma_{jl}'^h - \partial_j \Gamma_{il}'^h,$$
$$R'_{ijkl} = g_{hk}R_{ijl}'^h + g'_{hk}R_{ijl}^h.$$

Combining these results and the identity

$$\partial_i\partial_j g'_{kl} - \partial_j\partial_i g'_{kl} = g^{pq}(R_{ijkp}g'_{ql} + R_{ijlp}g'_{qk}),$$

we get the identity

$$R'_{ijkl} = -\tfrac{1}{2}(\partial_i\partial_k g'_{jl} - \partial_i\partial_l g'_{jk} - \partial_j\partial_k g'_{il} + \partial_j\partial_l g'_{ik})$$
$$+\tfrac{1}{2}g^{pq}(R_{ijkp}g'_{ql} + R_{ijpl}g'_{qk}),$$

which holds for any evolution of a metric. In our case $g'_{ij} = -2R_{ij}$, and substituting this gives

$$R'_{ijkl} = \partial_i\partial_k R_{jl} - \partial_i\partial_l R_{jk} - \partial_j\partial_k R_{il} + \partial_j\partial_l R_{ik}$$
$$-g^{pq}(R_{ijkp}R_{ql} + R_{ijpl}R_{qk}).$$

Then Theorem 7.1 is an immediate consequence of the following identity, which is independent of any evolution equation.

7.2 Lemma. *For any metric g_{ij} the curvature tensor R_{ijkl} satisfies the identity*

$$\Delta R_{ijkl} + 2(B_{ijkl} - B_{ijlk} - B_{iljk} + B_{ikjl})$$
$$= \partial_i\partial_k R_{jl} - \partial_i\partial_l R_{jk} - \partial_j\partial_k R_{il} + \partial_j\partial_l R_{ik}$$
$$+g^{pq}(R_{pjkl}R_{qi} + R_{ipkl}R_{qj}).$$

Proof. This formula is obtained from the second Bianchi identity

$$\partial_i R_{jklm} + \partial_j R_{kilm} + \partial_k R_{ijlm} = 0$$

by differentiating, exchanging derivatives, permuting indices and contracting. To begin we have

$$\Delta R_{ijkl} = g^{pq}\partial_p\partial_q R_{ijkl} = g^{pq}(\partial_p\partial_i R_{qjkl} - \partial_p\partial_j R_{qikl})$$

by differentiating the second Bianchi identity and contracting. We examine the first term, since the second is symmetric in i and j. Interchanging the order of derivatives we have

$$g^{pq}\partial_p\partial_i R_{qjkl} - g^{pq}\partial_i\partial_p R_{qjkl}$$
$$= g^{pq}g^{mn}\big(R_{piqm}R_{njkl} + R_{pijm}R_{qnkl} + R_{pikm}R_{qjnl} + R_{pilm}R_{qjkn}\big).$$

The first of these terms contracts to $g^{pq}R_{pjkl}R_{qi}$. On the second term we can use the first Bianchi identity to write it in terms of the tensor B_{ijkl}; thus

$$g^{pq}g^{mn}R_{pijm}R_{qnkl} = -B_{ijkl} + B_{ijlk},$$

and the last two terms are $-B_{ikjl} + B_{iljk}$. Moreover we have the contracted second Bianchi identity

$$g^{pq}\partial_p R_{qjkl} = \partial_k R_{jl} - \partial_l R_{jk},$$

to which we apply the derivative ∂_i. Then

$$\begin{aligned} g^{pq}\partial_p\partial_i R_{qjkl} = {}& \partial_i\partial_k R_{jl} - \partial_i\partial_l R_{jk} \\ & - \big(B_{ijkl} - B_{ijlk} - B_{iljk} + B_{ikjl}\big) \\ & + g^{pq}R_{pjkl}R_{qi}. \end{aligned}$$

Replacing this in our formula for ΔR_{ijkl} and doing the same for the term with i and j interchanged yields the formula in the lemma.

7.3 Corollary. *The Ricci curvature satisfies the evolution equation*

$$\frac{\partial}{\partial t}R_{ik} = \Delta R_{ik} + 2g^{pr}g^{qs}R_{piqk}R_{rs} - 2g^{pq}R_{pi}R_{qk}.$$

Proof. Recall $\Delta R_{ik} = g^{pq}\partial_p\partial_q R_{ik}$. We use the relation $R_{ik} = g^{jl}R_{ijkl}$ to contract the previous equation. Now

$$\big(g^{jl}\big)' = -g^{jp}g^{lq}g'_{pq}$$

by the usual formula for the derivative of the inverse of a matrix, and therefore

$$R'_{ik} = g^{jl}R'_{ijkl} + 2g^{jp}g^{lq}R_{ijkl}R_{pq}.$$

Substituting for R'_{ijkl} and making the obvious contractions yields

$$\begin{aligned} R'_{ik} = {}& \Delta R_{ik} + 2g^{jl}\big(B_{ijkl} - 2B_{ijlk}\big) \\ & + 2g^{pr}g^{qs}R_{piqk}R_{rs} - 2g^{pq}R_{pi}R_{qk}. \end{aligned}$$

Then the corollary follows from the following lemma.

7.4 Lemma. *For any metric g_{ij} the tensor B_{ijkl} satisfies the identity*

$$g^{jl}\big(B_{ijkl} - 2B_{ijlk}\big) = 0.$$

Proof. Using the Bianchi identity

$$\begin{aligned} g^{jl}B_{ijkl} &= g^{jl}g^{pr}g^{qs}R_{piqj}R_{rksl} \\ &= g^{jl}g^{pr}g^{qs}R_{pqij}R_{rskl} \\ &= g^{jl}g^{pr}g^{qs}(R_{piqj} - R_{pjqi})(R_{rksl} - R_{rlsk}) \\ &= 2g^{jl}(B_{ijkl} - B_{ijlk}) \end{aligned}$$

and the result follows.

7.5 Corollary. *The scalar curvature R satisfies the evolution equation*

$$\frac{\partial}{\partial t}R = \Delta R + 2g^{ij}g^{kl}R_{ik}R_{jl}.$$

Proof. Again we contract the previous equation. Since $R = g^{ik}R_{ik}$ we have

$$R' = g^{ik}R'_{ik} + 2g^{ij}g^{kl}R_{ik}R_{jl},$$

where the second term comes from $(g^{ik})'$. Then the equation for R'_{ik} immediately gives $g^{ik}R'_{ik} = \Delta R$.

7.6 Corollary. *If the scalar curvature $R > 0$ at $t = 0$, then it remains so.*

Proof. The term $g^{ij}g^{kl}R_{ik}R_{jl}$ is just the norm squared of the Ricci curvature, and hence is always positive. The result now follows from the maximum principle for the heat equation. This simple example is a model for our subsequent a priori estimates. It also shows why the evolution equation "prefers" positive curvature.

8. Curvature in dimension three

The Weyl conformal curvature tensor is defined as

$$\begin{aligned} W_{ijkl} = R_{ijkl} &- \frac{1}{n-2}(g_{ik}R_{jl} - g_{il}R_{jk} - g_{jk}R_{il} + g_{jl}R_{ik}) \\ &+ \frac{1}{(n-1)(n-2)}R(g_{ik}g_{jl} - g_{il}g_{jk}). \end{aligned}$$

This tensor is known to depend only on the conformal structure, so that if $\tilde{g}_{ij} = \psi g_{ij}$ then $\tilde{W}_{ijkl} = \psi W_{ijkl}$. In dimension $n \geqslant 4$ the conformal curvature tensor vanishes if and only if the metric g_{ij} is conformally flat. In dimension $n = 3$ this fails; instead there is a condition on the first derivative of the curvature, and the conformal curvature tensor always vanishes.

To see that $W_{ijkl} = 0$ in dimension three, observe first that it has all the symmetries of the Riemannian curvature tensor R_{ijkl}, so that

$$\begin{aligned} &W_{ijkl} = -W_{jikl} = -W_{ijlk} = W_{jilk} = W_{klij}, \\ &W_{ijkl} + W_{iklj} + W_{iljk} = 0, \end{aligned}$$

and in addition all its traces vanish, so

$$g^{ik}W_{ijkl} = 0.$$

Thus

$$W_{1111} + W_{1212} + W_{1313} = 0,$$

and so

$$W_{1212} = -W_{1313} = W_{2323} = -W_{2121} = -W_{1212},$$

which implies $W_{1212} = 0$. Moreover

$$W_{1213} + W_{2223} + W_{3233} = 0,$$

and so $W_{1213} = 0$ also. Hence in general any term $W_{ijkl} = 0$ unless i, j, k and l are all distinct. In dimension 3 there are only 3 possible choices for the indices, and the tensor must vanish identically.

This is just one special case of a general theory about tensors as representations of the orthogonal group $O(n)$. Any tensor decomposes as a sum of irreducible tensors, each of which is trace-free and has the maximum possible symmetry. Tensors with sufficiently exotic symmetries will always vanish in sufficiently low dimensions. In any case, we have the following result, which is well known.

8.1 Theorem. *In dimension three we have*

$$R_{ijkl} = g_{ik}R_{jl} - g_{il}R_{jk} - g_{jk}R_{il} + g_{jl}R_{ik} - \tfrac{1}{2}R(g_{ik}g_{jl} - g_{il}g_{jk}).$$

This result implies that we can recover the full Riemannian curvature tensor R_{ijkl} just from the Ricci curvature R_{ij}, which is much easier to handle. For example, we can always diagonalize R_{ij} at a point, so that

$$R_{ij} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{pmatrix},$$

where λ, μ, ν are the eigenvalues. Then the only nonzero components of R_{ijkl} are those of the form

$$R_{1212} = \tfrac{1}{2}(\lambda + \mu - \nu),$$

and those derived from it by permutation. Thus the condition for positive sectional curvature in three dimensions is that each eigenvalue of the Ricci tensor is smaller than the sum of the other two.

8.2 Corollary. *In dimension three a metric has positive sectional curvature if and only if* $R_{ij} < \frac{1}{2}Rg_{ij}$. *This shows that the condition of positive Ricci curvature is much weaker than that of positive sectional curvature.*

As a consequence of the formula for R_{ijkl} the evolution equation for the Ricci curvature R_{ij} takes a particularly simple form in dimension three. To simplify the formulas we introduce the following notation. We let

$$S_{il} = R^2_{il} = R_{ij}g^{jk}R_{kl},$$

$$T_{in} = R^3_{in} = R_{ij}g^{jk}R_{kl}g^{lm}R_{mn},$$

and we let S and T be the traces

$$S = g^{il}S_{il}, \quad T = g^{in}T_{in}.$$

Then in terms of the previous diagonalization

$$R_{ij} = \begin{pmatrix} \lambda & & \\ & \mu & \\ & & \nu \end{pmatrix}, \quad S_{ij} = \begin{pmatrix} \lambda^2 & & \\ & \mu^2 & \\ & & \nu^2 \end{pmatrix}, \quad T_{ij} = \begin{pmatrix} \lambda^3 & & \\ & \mu^3 & \\ & & \nu^3 \end{pmatrix},$$

$$R = \lambda + \mu + \nu, \quad S = \lambda^2 + \mu^2 + \nu^2, \quad T = \lambda^3 + \mu^3 + \nu^3.$$

We also introduce the tensor

$$Q_{ij} = 6S_{ij} - 3RR_{ij} + (R^2 - 2S)g_{ij},$$

whose entry in the top corner is

$$2\lambda^2 - \mu^2 - \nu^2 - \lambda\mu - \lambda\nu + 2\mu\nu,$$

and whose other entries may be obtained by permuting the eigenvalues. This tensor may seem somewhat bizarre, but is characterized by the following property.

8.3 Theorem. *The tensor Q_{ij} vanishes identically on any three dimensional symmetric Riemannian manifold. Any symmetric tensor T_{ij} which is quadratic in the Ricci curvature and has this property must be a scalar multiple of Q_{ij}.*

Proof. The Ricci curvature on a three dimensional symmetric space either (a) has all its eigenvalues equal, as for S^3, or else (b) has two equal eigenvalues and the third is zero, as for $S^2 \times S^1$. In either case it is easy to check that $Q_{ij} = 0$. Conversely any tensor of the given type which is quadratic in the Ricci tensor must be a linear combination of S_{ij}, Sg_{ij}, RR_{ij}, and R^2g_{ij}. Then considering the cases (a) and (b) gives enough conditions to show the tensor is a multiple of Q_{ij}.

8.4 Theorem. *In dimension three the Ricci tensor satisfies the evolution equation*

$$\frac{\partial}{\partial t}R_{ij} = \Delta R_{ij} - Q_{ij}.$$

Proof. This follows directly from substituting the formula in Theorem 8.1 for the Riemannian curvature into the formula in Corollary 7.3 for the evolution of the Ricci curvature. The reader can check for himself that

$$g^{pr}g^{qs}R_{piqk}R_{rs} = \tfrac{3}{2}RR_{ik} - 2S_{ik} + \left(S - \tfrac{1}{2}R^2\right)g_{ik},$$

and since $g^{pq}R_{pi}R_{qk} = S_{ik}$ the result holds.

9. Preserving positive Ricci curvature

We shall use the following result, which generalizes the maximum principle to tensors. We say that a symmetric tensor $M_{ij} \geqslant 0$ if $M_{ij}v^iv^j \geqslant 0$ for all vectors v^i. As usual, $\Delta M_{ij} = g^{pq}\partial_p\partial_q M_{ij}$. We let u^k be a vector field and we let g_{ij}, M_{ij} and N_{ij} be symmetric tensors on a compact manifold X which may all depend on time t.

We assume that $N_{ij} = p(M_{ij}, g_{ij})$ is a polynomial in M_{ij} formed by contracting products of M_{ij} with itself using the metric. We require that this polynomial satisfy the condition that whenever v^i is a null-eigenvector of M_{ij}, so that $M_{ij}v^i = 0$ for all j, then we have $N_{ij}v^iv^j \geqslant 0$. We prove the following result.

9.1 Theorem. *Suppose that on* $0 \leqslant t \leqslant T$

$$\frac{\partial}{\partial t}M_{ij} = \Delta M_{ij} + u^k\partial_k M_{ij} + N_{ij},$$

where $N_{ij} = p(M_{ij}, g_{ij})$ *satisfies the null-eigenvector condition above. If* $M_{ij} \geqslant 0$ *at* $t = 0$, *then it remains so on* $0 \leqslant t \leqslant T$.

Proof. We will show $M_{ij} \geqslant 0$ on $0 \leqslant t \leqslant \delta$ where $\delta > 0$ is small compared to a constant C depending only on $\max|M_{ij}|$. Then repeated application of this result will cover the entire interval in a finite number of steps. To this end, we let

$$\tilde{M}_{ij} = M_{ij} + \varepsilon(\delta + t)g_{ij},$$

and we claim $\tilde{M}_{ij} > 0$ on $0 \leqslant t \leqslant \delta$ for every $\varepsilon > 0$. Then letting $\varepsilon \to 0$ will finish the proof.

If not, there will be a first time θ with $0 < \theta \leqslant \delta$ where $\tilde{M}_{ij}$ acquires a null eigenvector v^i of unit length at some point $x \in X$. If $\tilde{N}_{ij} = p(\tilde{M}_{ij}, g_{ij})$, then by our null-eigenvector condition $\tilde{N}_{ij}v^iv^j \geqslant 0$ at (x, θ). Moreover

$$|\tilde{N}_{ij} - N_{ij}| \leqslant C|\tilde{M}_{ij} - M_{ij}|,$$

where the constant C depends only on $\max(|\tilde{M}_{ij}| + |M_{ij}|)$ since p is a polynomial. If we keep $\varepsilon, \delta \leqslant 1$, then $\max|\tilde{M}_{ij}|$ depends only on $\max|M_{ij}|$. Therefore

$$N_{ij}v^iv^j \geqslant -C\varepsilon\delta,$$

where C depends only on $\max|M_{ij}|$ and not on ε or δ.

We can extend v^i to a vector field in a neighborhood of x with $\partial_j v^i = 0$ at x, with v^i independent of t. Let

$$f = \tilde{M}_{ij} v^i v^j.$$

Then $f \geqslant 0$ on $0 \leqslant t \leqslant \theta$ and all of X, so $\partial f / \partial t < 0$ and $\partial_k f = 0$ and $\Delta f \geqslant 0$ at (x, θ), where $f = 0$. But

$$\frac{\partial f}{\partial t} = \left(\frac{\partial}{\partial t} M_{ij}\right) v^i v^j + \varepsilon,$$

and at (x, θ) where $\partial_j v^i = 0$ and $\tilde{M}_{ij} v^i = 0$,

$$\partial_k f = \partial_k M_{ij} v^i v^j, \quad \Delta f = \Delta M_{ij} v^i v^j.$$

From the evolution equation

$$\left(\frac{\partial}{\partial t} M_{ij}\right) v^i v^j = \Delta M_{ij} v^i v^j + u^k \partial_k M_{ij} v^i v^j + N_{ij} v^i v^j,$$

which shows that $N_{ij} v^i v^j \leqslant -\varepsilon$. Combining this with the previous estimate, we have $\varepsilon \leqslant C\delta\varepsilon$. This gives a contradiction when $C\delta < 1$.

We will assume now that the evolution equation has a solution on the interval $0 \leqslant t < T$.

9.2 Corollary. *If $R_{ij} \geqslant 0$ at $t = 0$ then $R_{ij} \geqslant 0$ on $0 \leqslant t < T$.*

Proof. We apply Theorem 9.1 with $u^k = 0$, $M_{ij} = R_{ij}$ and $N_{ij} = -Q_{ij}$. When M_{ij} has a null eigenvalue $\lambda = 0$ the corresponding eigenvalue of N_{ij} is $(\mu - \nu)^2 \geqslant 0$.

To get more precise control on R_{ij} we need the following computation.

9.3 Lemma. *If $R \neq 0$, then*

$$\frac{\partial}{\partial t}\left(\frac{R_{ij}}{R}\right) = \Delta\left(\frac{R_{ij}}{R}\right) + \frac{2}{R} g^{pq} \partial_p R \partial_q \left(\frac{R_{ij}}{R}\right) - \frac{R Q_{ij} + 2 S R_{ij}}{R^2}.$$

Proof. Since $\partial R_{ij} / \partial t = \Delta R_{ij} - Q_{ij}$ and $\partial R / \partial t = \Delta R + 2S$, we have

$$\frac{\partial}{\partial t}\left(\frac{R_{ij}}{R}\right) = \frac{R(\Delta R_{ij} - Q_{ij}) - R_{ij}(\Delta R + 2S)}{R^2}.$$

On the other hand

$$\Delta\left(\frac{R_{ij}}{R}\right) = \frac{R \Delta R_{ij} - R_{ij} \Delta R}{R^2} - \frac{2}{R} g^{pq} \partial_p R \partial_q \left(\frac{R_{ij}}{R}\right).$$

The lemma follows.

9.4 Theorem. *If $R \geqslant 0$ and $R_{ij} \geqslant \varepsilon R g_{ij}$ for some constant $\varepsilon > 0$ at $t = 0$, then both conditions continue to hold on $0 \leqslant t < T$.*

Proof. We saw that $R > 0$ continues to hold in Corollary 7.6. We apply Theorem 9.1 with

$$M_{ij} = \frac{R_{ij}}{R} - \varepsilon g_{ij}, \quad u^k = \frac{2}{R} g^{kl} \partial_l R,$$
$$N_{ij} = 2\varepsilon R_{ij} - \left(\frac{RQ_{ij} + 2SR_{ij}}{R^2} \right).$$

It is an immediate consequence of Lemma 9.3 that the equation in Theorem 9.1 is satisfied. Let us consider what happens to N_{ij} when M_{ij} acquires a null eigenvector. The analysis is easy since when R_{ij} is diagonal so are M_{ij} and N_{ij}. Suppose the null eigenvalue of M_{ij} occurs in the top position, corresponding to the eigenvalue λ of R_{ij}. Then $\lambda = \varepsilon(\lambda + \mu + \nu)$. The corresponding entry in $R^2 N_{ij}$ is

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

Using the previous identity to eliminate ε, and multiplying out and gathering terms, this entry becomes

$$(\lambda + \mu + \nu)\left[\lambda(\mu + \nu) + (\mu - \nu)^2\right] - 2\lambda(\lambda^2 + \mu^2 + \nu^2),$$

which further simplifies to

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

Now if $R_{ij} \geqslant \varepsilon R g_{ij}$ then $R \geqslant 3\varepsilon R$, and if $R > 0$ then $\varepsilon \leqslant \frac{1}{3}$. But then $\mu + \nu = (1/\varepsilon - 1)\lambda \geqslant 2\lambda$, so $\mu + \nu - 2\lambda \geqslant 0$. Therefore at any null eigenvector of M_{ij} the matrix N_{ij} is positive. The theorem follows.

It is easy to obtain a bound above on R_{ij}.

9.5 Lemma. *If $R_{ij} \geqslant 0$ then $R_{ij} \leqslant R g_{ij}$.*

Proof. Since $\lambda, \mu, \nu \geqslant 0$ we have $\lambda \leqslant \lambda + \mu + \nu$.

The consequence of these estimates is that when $R_{ij} > 0$ at $t = 0$ we have a uniform bound $\lambda/\mu \leqslant C$ on the ratio of any two eigenvalues of R_{ij} holding as long as the solution exists. This allows us to control all the curvature R_{ij} just in terms of the scalar curvature R. The following estimate is also interesting.

9.6 Theorem. *If $\varepsilon R g_{ij} \leqslant R_{ij} \leqslant \beta R g_{ij}$ for some constants ε and β with $0 < \varepsilon \leqslant \frac{1}{3} < \beta < 1$ at $t = 0$, then both conditions continue to hold on $0 \leqslant t < T$.*

Proof. Note that if $\varepsilon = \frac{1}{3}$ or $\beta = \frac{1}{3}$ then the manifold has constant curvature and the result is trivial, while $\beta = 1$ always holds. We apply Theorem 9.1 with

$$M_{ij} = \beta g_{ij} - \frac{R_{ij}}{R}, \quad u^k = \frac{2}{R} g^{kl} \partial_l R,$$

$$N_{ij} = \left(\frac{RQ_{ij} + 2SR_{ij}}{R^2}\right) - 2\beta R_{ij}.$$

It follows immediately from Lemma 9.3 that the equation in Theorem 9.1 is satisfied. Again we consider what happens to N_{ij} when M_{ij} acquires a null eigenvector in the top position where R_{ij} has eigenvalue λ. In this case $\lambda = \beta(\lambda + \mu + \nu)$. The top entry in $R^2 N_{ij}$ is

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

Eliminating β with the above identity and gathering terms, this reduces to

$$\lambda^2(2\lambda - \mu - \nu) - (\mu + \nu)(\mu - \nu)^2$$

which we can rearrange as

$$2\lambda^2(\lambda - \mu) + (\lambda^2 - \mu^2 + \nu^2)(\mu - \nu),$$

which is clearly positive if $\lambda \geqslant \mu \geqslant \nu \geqslant 0$. To handle the possibility that λ is not the largest eigenvalue we use a continuity argument. Let θ be the largest time on which $R_{ij} \leqslant (\beta + \delta)Rg_{ij}$, where δ will be chosen small compared to β and ε. If we can show $R_{ij} \leqslant \beta Rg_{ij}$ up to time θ, then we must have $\theta = T$. Now since $\lambda = \beta(\lambda + \mu + \nu)$ and $\beta \geqslant \frac{1}{3}$ we see λ cannot be the smallest eigenvalue. Assume $\mu \geqslant \lambda \geqslant \nu$. Up to time θ we have $\mu \leqslant (\beta + \delta)(\lambda + \mu + \nu)$, and by Theorem 9.4 we have $\nu \geqslant \varepsilon(\lambda + \mu + \nu)$. Since $\mu \geqslant \lambda = \beta(\lambda + \mu + \nu)$ and $\nu \leqslant \frac{1}{3}(\lambda + \mu + \nu)$ we have $\mu - \nu \geqslant (\beta - \frac{1}{3})(\lambda + \mu + \nu)$. If $\beta = \frac{1}{3}$ the manifold has constant curvature, and this case is easy to handle. Assume $\beta > \frac{1}{3}$. The entry of N_{ij} in question by algebraic rearrangement becomes

$$\nu^2(\mu - \nu) - (\mu - \lambda)\left[2\lambda^2 + (\lambda + \mu)(\mu - \nu)\right],$$

which is at least

$$\varepsilon^2(\beta - \tfrac{1}{3}) - \delta\left[2\beta^2 + (2\beta + \delta)(\beta + \delta)\right]$$

times $(\lambda + \mu + \nu)^2$. This expression will be positive if δ is small enough compared to β and ε. This completes the proof.

9.7 Corollary. *If the sectional curvature is positive at $t = 0$, then it remains so on $0 \leqslant t < T$.*

Proof. We say in Corollary 8.2 that the sectional curvature was positive if and only if $R_{ij} < \frac{1}{2}Rg_{ij}$. The same result holds for weakly positive sectional curvature, taking $\beta = \frac{1}{2}$.

10. Pinching the eigenvalues

The next estimate shows, after a fashion, that the eigenvalues of the Ricci tensor approach each other, at least at those points where the scalar curvature becomes large (for the unnormalized equation). Consider the expression $S - \frac{1}{3}R^2$ quadratic in the eigenvalues

$$S - \tfrac{1}{3}R^2 = \tfrac{1}{3}\left[(\lambda - \mu)^2 + (\lambda - \nu)^2 + (\mu - \nu)^2\right].$$

Clearly $S - \frac{1}{3}R^2 \geqslant 0$, and vanishes only when $\lambda = \mu = \nu$. Thus it measures how far the eigenvalues diverge from each other. If indeed the manifold is becoming spherical, we expect $S - \frac{1}{3}R^2$ to become small, at least compared to R^2 for the unnormalized equation. That is the content of our result. We assume as usual that our manifold is three dimensional, the initial metric has strictly positive Ricci curvature, and the (unnormalized) evolution equation has a solution on $0 \leqslant t < T$.

10.1 Theorem. *We can find a $\delta > 0$ and a constant $C < \infty$ depending only on the initial metric such that on $0 \leqslant t < T$ we have*

$$S - \tfrac{1}{3}R^2 \leqslant CR^{2-\delta}.$$

Proof. We take $\gamma = 2 - \delta$ with $1 < \gamma < 2$. The following equations follow from the equations for the evolution of the Ricci curvature and the scalar curvature.

$$\frac{\partial}{\partial t}R_{ij} = \Delta R_{ij} - Q_{ij}, \quad \frac{\partial}{\partial t}R = \Delta R + 2S.$$

Recall that

$$S = |R_{ij}|^2 = g^{ik}g^{jl}R_{ij}R_{kl} = \lambda^2 + \mu^2 + \nu^2,$$
$$T = g^{in}g^{jk}g^{lm}R_{ij}R_{kl}R_{mn} = \lambda^3 + \mu^3 + \nu^3,$$

and let

$$|\partial_i R_{jk}|^2 = g^{il}g^{jm}g^{kn}\partial_i R_{jk}\partial_l R_{mn},$$
$$C = \tfrac{1}{2}g^{ik}g^{jl}Q_{ij}R_{kl} = \tfrac{1}{2}(R^3 - 5RS + 6T)$$
$$= (\lambda^3 + \mu^3 + \nu^3) - (\lambda^2\mu + \lambda\mu^2 + \lambda^2\nu + \lambda\nu^2 + \mu^2\nu + \mu\nu^2) + 3\lambda\mu\nu$$

as the reader may compute. Note that C is a cubic expression in the eigenvalues which vanishes for any symmetric metric; one can show that this condition characterizes C up to a multiple.

10.2 Lemma. *The expression S satisfies the evolution equation*

$$\frac{\partial}{\partial t}S = \Delta S - 2|\partial_i R_{jk}|^2 + 4(T - C).$$

Proof. From the evolution equations for g_{ij} and R_{ij} we have

$$\frac{\partial}{\partial t}S = 2g^{ik}g^{jl}\Delta R_{ij}\cdot R_{kl} + 4(T - C),$$

while we also have

$$\Delta S = 2g^{ik}g^{jl}\Delta R_{ij}\cdot R_{kl} + 2\,|\,\partial_i R_{jk}\,|^2.$$

10.3 Lemma. *If $R > 0$ then for any γ*

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{S}{R^\gamma}\right) &= \Delta\left(\frac{S}{R^\gamma}\right) + \frac{2(\gamma-1)}{R}g^{pq}\partial_p R\cdot\partial_q\left(\frac{S}{R^\gamma}\right)\\ &\quad - \frac{2}{R^{\gamma+2}}\,|\,R\partial_i R_{jk} - \partial_i R\cdot R_{jk}\,|^2\\ &\quad - \frac{(2-\gamma)(\gamma-1)}{R^{\gamma+2}}S\,|\,\partial_i R\,|^2 + \frac{4R(T-C) - 2\gamma S^2}{R^{\gamma+1}}.\end{aligned}$$

Proof. We have

$$\frac{\partial}{\partial t}\left(\frac{S}{R^\gamma}\right) = \frac{R\frac{\partial S}{\partial t} - \gamma S\frac{\partial R}{\partial t}}{R^{\gamma+1}},\quad \partial_j\left(\frac{S}{R^\gamma}\right) = \frac{R\partial_j S - \gamma S\partial_j R}{R^{\gamma+1}},$$

$$\Delta\left(\frac{S}{R^\gamma}\right) = \frac{R\Delta S - \gamma S\Delta R}{R^{\gamma+1}} - \frac{2\gamma}{R^{\gamma+1}}g^{ij}\partial_i R\cdot\partial_j S + \frac{\gamma(\gamma+1)}{R^{\gamma+2}}S\,|\,\partial_i R\,|^2.$$

Introducing the obvious inner product of two tensors

$$\left\langle T_{ijk}, U_{ijk}\right\rangle = g^{il}g^{jm}g^{kn}T_{ijk}U_{lmn},$$

we have

$$\langle\partial_i R, \partial_i S\rangle = 2\langle\partial_i R_{jk}, \partial_i R\cdot R_{jk}\rangle,$$

$$\left\langle\partial_i R, \partial_i\left(\frac{S}{R^\gamma}\right)\right\rangle = \frac{1}{R^\gamma}\langle\partial_i R, \partial_i S\rangle - \frac{\gamma}{R^{\gamma+1}}S\,|\,\partial_i R\,|^2,$$

and we also have

$$S\,|\,\partial_i R\,|^2 = |\,\partial_i R\cdot R_{jk}\,|^2.$$

Thus the terms in the evolution equation for S/R^γ which are quadratic in the first derivatives of the curvature are equal to $1/R^{\gamma+2}$ times

$$\begin{aligned}&-2R^2\,|\,\partial_i R_{jk}\,|^2 + 2\gamma R\langle\partial_i R, \partial_i S\rangle - \gamma(\gamma+1)S\,|\,\partial_i R\,|^2\\ &\quad = -2\,|\,R\partial_i R_{jk} - \partial_i R\cdot R_{jk}\,|^2 + 2(\gamma-1)\big(R\langle\partial_i R, S\rangle - \gamma S\,|\,\partial_i R\,|^2\big)\\ &\qquad + (\gamma-2)(\gamma-1)S\,|\,\partial_i R\,|^2,\end{aligned}$$

and now the result follows directly.

10.4 Lemma. *If $R > 0$, then for any γ*

$$\frac{\partial}{\partial t} R^{2-\gamma} = \Delta R^{2-\gamma} + \frac{2(\gamma - 1)}{R} g^{pq} \partial_p R \partial_q (R^{2-\gamma})$$
$$- \frac{(2-\gamma)(\gamma-1)}{R^{\gamma+2}} R^2 |\partial_i R|^2 + 2(2-\gamma) R^{1-\gamma} S.$$

Proof. We have

$$\frac{\partial}{\partial t} R^{2-\gamma} = (2-\gamma) R^{1-\gamma}(\Delta R + 2S),$$
$$\Delta R^{2-\gamma} = (2-\gamma) R^{1-\gamma} \Delta R + (2-\gamma)(1-\gamma) R^{-\gamma} |\partial_i R|^2,$$
$$\frac{2(\gamma-1)}{R} g^{pq} \partial_p R \partial_q (R^{2-\gamma}) = 2(2-\gamma)(\gamma-1) R^{-\gamma} |\partial_i R|^2,$$

and the result follows.

10.5 Lemma. *If $f = S/R^{\gamma} - \frac{1}{3} R^{2-\gamma}$, then*

$$\frac{\partial f}{\partial t} = \Delta f + \frac{2(\gamma-1)}{R} g^{pq} \partial_p R \partial_q f - \frac{2}{R^{\gamma+2}} |R \partial_i R_{jk} - \partial_i R \cdot R_{jk}|^2$$
$$- \frac{(2-\gamma)(\gamma-1)}{R^{\gamma+2}} \left(S - \tfrac{1}{3} R^2\right) |\partial_i R|^2$$
$$+ \frac{2}{R^{\gamma+1}} \left[(2-\gamma) S \left(S - \tfrac{1}{3} R^2\right) - 2P\right],$$

where $P = S^2 + R(C - T)$.

Proof. This follows directly from Lemmas 10.3 and 10.4.

Now we must analyze the polynomial P, which is clearly a symmetric polynomial of degree 4 in λ, μ, ν.

10.6 Lemma.

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

Proof. Using our formulas for R, S, T, and C (given just before Lemma 10.2) we can multiply out to get

$$P = \left(\lambda^4 + \mu^4 + \nu^4\right) - \left(\lambda^3\mu + \lambda\mu^3 + \lambda^3\nu + \lambda\nu^3 + \mu^3\nu + \mu\nu^3\right)$$
$$+ \left(\lambda^2\mu\nu + \lambda\mu^2\nu + \lambda\mu\nu^2\right).$$

And if we multiply out the polynomial above we get the same thing. Note that the polynomial P vanishes for any symmetric metric, since it vanishes when $\lambda = \mu = \nu$ or when $\lambda = \mu$ and $\nu = 0$.

10.7 Lemma. *If $R > 0$ and $R_{ij} \geqslant \varepsilon R g_{ij}$ then $P \geqslant \varepsilon^2 S(S - \frac{1}{3} R^2)$.*

Proof. Since both sides are homogeneous of degree 4 in λ, μ, ν, it suffices to check the result on $S = \lambda^2 + \mu^2 + \nu^2 = 1$. Assume $\lambda \geqslant \mu \geqslant \nu > 0$. Since

$(\lambda+\mu+\nu)^2 \geqslant \lambda^2+\mu^2+\nu^2=1$, we have $\nu \geqslant \varepsilon(\lambda+\mu+\nu) \geqslant \varepsilon$ by the bound $R_{ij} \geqslant \varepsilon R g_{ij}$. Now we can rewrite P as

$$P=(\lambda-\mu)^2\left[\lambda^2+(\lambda+\mu)(\mu-\nu)\right]+\nu^2(\lambda-\nu)(\mu-\nu),$$

which makes it clear that

$$P \geqslant \lambda^2(\lambda-\mu)^2+\nu^2(\mu-\nu)^2,$$

and since $\lambda \geqslant \nu \geqslant \varepsilon$ we have

$$P \geqslant \varepsilon^2\left[(\lambda-\mu)^2+(\mu-\nu)^2\right].$$

On the other hand, since

$$(\lambda-\nu)^2=[(\lambda-\mu)+(\mu-\nu)]^2 \leqslant 2\left[(\lambda-\mu)^2+(\mu-\nu)^2\right],$$

we see that

$$S-\tfrac{1}{3}R^2=\tfrac{1}{3}\left[(\lambda-\mu)^2+(\lambda-\nu)^2+(\mu-\nu)^2\right] \leqslant (\lambda-\mu)^2+(\mu-\nu)^2,$$

and this proves the lemma.

10.8 Lemma. *If $\delta>0$ is chosen so small that $\delta \leqslant 2\varepsilon^2$, then with $\gamma=2-\delta$ and $f=S/R^\gamma-\frac{1}{3}R^{2-\gamma}$ we have*

$$\frac{\partial f}{\partial t} \leqslant \Delta f+u^k \partial_k f,$$

where $u^k=(2(\gamma-1)/R)g^{kl}\partial_l R$.

Proof. This follows from Lemma 10.5 and our estimate on P.

Now we can finish the proof of Theorem 10.1. Let $\delta>0$ be as above and choose C so that

$$\frac{S}{R^\gamma}-\tfrac{1}{3}R^{2-\gamma} \leqslant C$$

at $t=0$. Then $f \leqslant C$ at $t=0$, and by the maximum principle we have $f \leqslant C$ on $0 \leqslant t<T$ for the same C. Thus we have $S-\frac{1}{3}R^2 \leqslant CR^{2-\delta}$ as desired.

11. The gradient of the scalar curvature

Again we assume our manifold is compact and three dimensional, the initial metric has strictly positive Ricci curvature, and the unnormalized evolution equation has a solution on $0 \leqslant t<T$.

11.1 Theorem. *For every $\eta > 0$ we can find a constant $C(\eta)$ depending only on η and the initial value of the metric, such that on $0 \leqslant t < T$ we have*

$$|\partial_i R|^2 \leqslant \eta R^3 + C(\eta).$$

Proof. We start with the evolution equation for $|\partial_i R|^2 = g^{ij}\partial_i R \partial_j R$.

11.2 Lemma. *The gradient squared of the scalar curvature satisfies the evolution equation*

$$\frac{\partial}{\partial t}|\partial_i R|^2 = \Delta|\partial_i R|^2 - 2|\partial_i\partial_j R|^2 + 4g^{ij}\partial_i R\partial_j S.$$

Proof. We compute

$$\begin{aligned}\frac{\partial}{\partial t}|\partial_i R|^2 &= 2g^{ij}\partial_i \Delta R \cdot \partial_j R + 4g^{ij}\partial_i R\partial_j S\\ &\quad + 2g^{ik}g^{jl}R_{kl}\partial_i R\partial_j R,\\ \Delta|\partial_i R|^2 &= 2g^{ij}\Delta\partial_i R \cdot \partial_j R + 2|\partial_i\partial_j R|^2,\\ \Delta\partial_i R &= \partial_i \Delta R + g^{jk}R_{ij}\partial_k R,\end{aligned}$$

and the result follows easily by combining terms.

11.3 Lemma. *We have the evolution equation*

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{|\partial_i R|^2}{R}\right) &= \Delta\left(\frac{|\partial_i R|^2}{R}\right) - \frac{2}{R^3}|R\partial_i\partial_j R - \partial_i R\partial_j R|^2\\ &\quad + \frac{4}{R}g^{ij}\partial_i R\partial_j S - \frac{2S}{R^2}|\partial_i R|^2.\end{aligned}$$

Proof. We compute

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{|\partial_i R|^2}{R}\right) &= \frac{R\Delta|\partial_i R|^2 - |\partial_i R|^2\Delta R}{R^2} - \frac{2}{R^3}|R\partial_i\partial_j R|^2\\ &\quad + \frac{4}{R}g^{ij}\partial_i R\partial_j S - \frac{2S}{R^2}|\partial_i R|^2,\end{aligned}$$

while

$$\begin{aligned}\Delta\left(\frac{|\partial_i R|^2}{R}\right) &= \frac{R\Delta|\partial_i R|^2 - |\partial_i R|^2\Delta R}{R^2}\\ &\quad - \frac{4}{R^3}\langle R\partial_i\partial_j R, \partial_i R\partial_j R\rangle + \frac{2}{R^3}|\partial_i R\,\partial_j R|^2,\end{aligned}$$

and the result follows.

11.4 Lemma. *We have evolution equations*

$$\frac{\partial}{\partial t}R^2 = \Delta R^2 - 2|\partial_i R|^2 + 4RS,$$

$$\frac{\partial}{\partial t}S = \Delta S - 2|\partial_i R_{jk}|^2 + 4(T - C),$$

$$\frac{\partial}{\partial t}\left(S - \tfrac{1}{3}R^2\right) = \Delta\left(S - \tfrac{1}{3}R^2\right) - 2\left(|\partial_i R_{jk}|^2 - \tfrac{1}{3}|\partial_i R|^2\right) + 4Q,$$

where $Q = T - \frac{1}{3}RS - C$.

Proof. The first follows from $\partial R/\partial t = \Delta R + 2S$, the second is Lemma 10.2, and the third follows by subtraction.

11.5 Lemma. $Q \leqslant R(S - \frac{1}{3}R^2)$.

Proof. Recall that the polynomial $P = S^2 + R(C - T) \geqslant 0$ from Lemma 10.5. Then since $S \leqslant R^2$ we have

$$QR \leqslant P + QR = S\left(S - \tfrac{1}{3}R^2\right) \leqslant R^2\left(S - \tfrac{1}{3}R^2\right),$$

and the result follows by dividing by R.

Now since $\partial_i R = g^{jk}\partial_i R_{jk}$, it is trivial to see that

$$|\partial_i R|^2 \leqslant 3|\partial_i R_{jk}|^2,$$

since $(a + b + c)^2 \leqslant 3(a^2 + b^2 + c^2)$. It is a little surprising that a slightly better estimate holds.

11.6 Lemma. $|\partial_i R|^2 \leqslant \frac{20}{7}|\partial_i R_{jk}|^2$.

Proof. This is a consequence of the contracted second Bianchi identity, which says,

$$g^{ij}\partial_i R_{jk} = \tfrac{1}{2}\partial_k R.$$

It is always a good idea to try writing a tensor as a sum of irreducible components. Write

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

where

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

Then it is easy to compute

$$|E_{ijk}|^2 = \tfrac{7}{20}|\partial_i R|^2.$$

On the other hand, we can check that the tensor $F_{ijk} = \partial_i R_{jk} - E_{ijk}$ is trace-free, so that

$$g^{ij}F_{ijk} = 0, \quad g^{ik}F_{ijk} = 0, \quad g^{jk}F_{ijk} = 0.$$

Indeed that was how we figured out what E_{ijk} should be. It follows that $\langle E_{ijk}, F_{ijk}\rangle = 0$ and the tensors E_{ijk} and F_{ijk} are perpendicular. Then

$$|\partial_i R_{jk}|^2 = |E_{ijk}|^2 + |F_{ijk}|^2 \geqslant \tfrac{7}{20} |\partial_i R|^2,$$

and this proves the lemma.

11.7 Lemma. *We have the estimate*

$$\frac{\partial}{\partial t}\left(S - \tfrac{1}{3}R^2\right) \leqslant \Delta\left(S - \tfrac{1}{3}R^2\right) - \tfrac{2}{21} |\partial_i R_{jk}|^2 + 4R\left(S - \tfrac{1}{3}R^2\right).$$

Proof. This follows from Lemmas 11.4, 11.5, and 11.6.

Now we return to the equation for $|\partial_i R|^2/R$ in Lemma 11.3. The problem with this equation is the term $g^{ij}\partial_i R \partial_j S$, which we estimate as follows.

11.8 Lemma. *We have*

$$g^{ij}\partial_i R \partial_j S \leqslant 4R |\partial_i R_{jk}|^2.$$

Proof. We use the Cauchy-Schwartz inequality

$$g^{ij}\partial_i R \partial_j S = 2\langle \partial_i R \cdot R_{jk}, \partial_i R_{jk}\rangle \leqslant 2 |\partial_i R| |R_{jk}| |\partial_i R_{jk}| ,$$

and $|R_{jk}|^2 = S \leqslant R^2$ and $|\partial_i R|^2 \leqslant 3 |\partial_i R_{jk}|^2$. We take $\sqrt{3} \leqslant 2$ to avoid a square root.

11.9 Lemma. *We have the estimate for* $\eta \leqslant \frac{1}{3}$

$$\frac{\partial}{\partial t}\left(\frac{|\partial_i R|^2}{R} - \eta R^2\right) \leqslant \Delta\left(\frac{|\partial_i R|^2}{R} - \eta R^2\right) + 16 |\partial_i R_{jk}|^2 - \tfrac{4}{3}\eta R^3.$$

Proof. We use the equation in Lemma 11.3 and the first equation in Lemma 11.4, multiplied by η. Since $S \geqslant \frac{1}{3}R^2$ the term $2(S/R^2)|\partial_i R|^2$ dominates $2\eta |\partial_i R|^2$ for $\eta \leqslant \frac{1}{3}$. We bound the term $g^{ij}\partial_i R \partial_j S$ by Lemma 11.8.

Now we want to combine Lemmas 11.7 and 11.9. The idea is to add enough of $S - \frac{1}{3}R^2$ to $|\partial_i R|^2/R$ to cancel off the term $|\partial_i R_{jk}|^2$, and then use Theorem 10.1 to make $R(S - \frac{1}{3}R^2)$ small compared to R^3. Note $168 \cdot \frac{2}{21} = 16$.

11.10 Lemma. *Let* $F = |\partial_i R|^2/R - \eta R^2 + 168(S - \frac{1}{3}R^2)$. *Then for any* η *with* $0 < \eta \leqslant \frac{1}{3}$ *we can find a constant* $C(\eta)$ *depending only on* η *and the initial value of the metric at* $t = 0$ *such that*

$$\frac{\partial F}{\partial t} \leqslant \Delta F + C(\eta).$$

Proof. Using Lemmas 11.7 and 11.9, the terms which are left are

$$672R\left(S - \tfrac{1}{3}R^2\right) - \tfrac{4}{3}\eta R^3.$$

Now by Theorem 10.1 we can find a constant C and some $\delta > 0$ depending only on the initial value of the metric, such that $S - \frac{1}{3}R^2 \leqslant CR^{2-\delta}$. Then with a constant $C(\eta)$ depending only on C, δ, and η we have

$$CR^{3-\delta} - \tfrac{4}{3}\eta R^3 \leqslant C(\eta)$$

and the result follows.

It remains only to find a bound on T, the time for which the solution exists. Since $S \geqslant \frac{1}{3}R^2$ we have

$$\frac{\partial R}{\partial t} \geqslant \Delta R + \frac{2}{3}R^2$$

which forces the minimum value of the scalar curvature R to go to infinity in a finite time.

11.11 Lemma. *If $R \geqslant \rho$ at $t = 0$ for some constant $\rho > 0$ then $T \leqslant 3/(2\rho)$.*

Proof. The solution of the ordinary differential equation

$$\frac{df}{dt} = \frac{2}{3}f^2 \quad \text{with } f = \rho \text{ at } t = 0$$

is given by

$$f = \frac{3\rho}{3 - 2\rho t}.$$

Taking f as a function on $X \times t$ constant in X we have

$$\frac{\partial}{\partial t}(R - f) \geqslant \Delta(R - f) + \frac{2}{3}(R + f)(R - f),$$

and the maximum principle implies $R - f \geqslant 0$ on $0 \leqslant t < T$. Since $f \to \infty$ as $t \to 3/\rho$, we must have $T \leqslant 3/\rho$.

Now the equation $\partial F/\partial t \leqslant \Delta F + C(\eta)$ implies $\max F_t \leqslant \max F_0 + C(\eta)t$. Then our bound on T shows that $F \leqslant C(\eta)$ for some (possibly larger) constant $C(\eta)$ depending only on $\eta > 0$ and the initial value of the metric (which determines F_0). This gives

$$F = |\partial_i R|^2/R - \eta R^2 + 168\left(S - \tfrac{1}{3}R^2\right) \leqslant C(\eta),$$
$$|\partial_i R|^2 \leqslant \eta R^3 + C(\eta)R,$$

and of course

$$\eta R^3 + C(\eta)R \leqslant 2\eta R^3 + \tilde{C}(\eta)$$

for some constant $\tilde{C}(\eta)$. Since $\eta > 0$ is arbitrary, this proves the result.

12. Interpolation inequalities for tensors

Let $T = \{T_{j\cdots k}\}$ denote a tensor covariant in any number of indices. We adopt the extended summation convention that if a pair of indices is repeated on the bottom, we should sum over that index in an orthonormal basis with respect to the metric g_{ij}. We let $\partial T = \{\partial_i T_{j\cdots k}\}$ be the covariant derivative with respect to the Levi-Civita connection Γ_{ij}^p associated to g_{ij}, and we let $\partial^2 T = \partial_h \partial_i T_{j\cdots k}$ be the second (iterated) covariant derivative. We also let $d\mu = \mu(x)dx$ be the volume form associated to the metric. The tensor $T = T_{j\cdots k}$ has length $|T|$ given by

$$|T|^2 = T_{j\cdots k} T_{j\cdots k},$$

and $|\partial T|$ and $|\partial^2 T|$ are defined analogously. We prove the following interpolation inequality by integration by parts.

12.1 Theorem. *Let X be a compact Riemannian manifold of dimension m and let $T = T_{j\cdots k}$ be any tensor on X. Suppose*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} \quad \text{with } r \geqslant 1.$$

Then

$$\left\{\int |\partial T|^{2r} d\mu\right\}^{1/r} \leqslant (2r-2+m)\left\{\int |\partial^2 T|^p d\mu\right\}^{1/p}\left\{\int |T|^q d\mu\right\}^{1/q}.$$

Proof. For simplicity we take $T = \{T_j\}$, since the more general case involves nothing extra but is more cumbersome to write. Integrating by parts

$$\begin{aligned}\int |\partial T|^{2r} d\mu &= \int \partial_i T_j \cdot \partial_i T_j \cdot (\partial_k T_l \cdot \partial_k T_l)^{r-1} d\mu \\ &= -\int T_j \cdot \partial_i \partial_i T_j \cdot |\partial T|^{2r-2} d\mu \\ &\quad -2(r-1)\int \langle T_j \partial_i \partial_k T_l, \partial_i T_j \cdot \partial_k T_l\rangle |\partial T|^{2r-4} d\mu.\end{aligned}$$

Now

$$|T_j \cdot \partial_i \partial_i T_j| \leqslant m|T||\partial^2 T|,$$

$$\langle T_j \partial_i \partial_k T_l, \partial_i T_j \cdot \partial_k T_l\rangle \leqslant |T||\partial^2 T||\partial T|^2,$$

and therefore

$$\int |\partial T|^{2r} d\mu \leqslant (2r-2+m)\int |T||\partial^2 T||\partial T|^{2r-2} d\mu.$$

We can estimate the last integral using Hölder's inequality with

$$\frac{1}{p}+\frac{1}{q}+\frac{r-1}{r}=1,$$

and we get

$$\int |\partial T|^{2r}\,d\mu$$
$$\leqslant (2r-2+m)\left\{\int |\partial^2 T|^p\,d\mu\right\}^{1/p}\left\{\int |T|^q\,d\mu\right\}^{1/q}\left\{\int |\partial T|^{2r}\,d\mu\right\}^{1-1/r},$$

and hence

$$\left\{\int |\partial T|^{2r}\,d\mu\right\}^{1/r}\leqslant (2r-2+m)\left\{\int |\partial^2 T|^p\,d\mu\right\}^{1/p}\left\{\int |T|^q\,d\mu\right\}^{1/q}.$$

12.2 Corollary. *If $p \geqslant 1$ we have*

$$\left\{\int |\partial T|^{2p}\,d\mu\right\}^{1/p}\leqslant (2p-2+m)\max_X |T|\cdot\left\{\int |\partial^2 T|^p\,d\mu\right\}^{1/p}.$$

Proof. Take $q=\infty$ in the previous argument.

Next we need a result on convexity, which is geometrically obvious.

12.3 Lemma. *Let $f(k)$ be a real valued function of the integer k for $0 \leqslant k \leqslant n$. If*

$$f(k) \leqslant \tfrac{1}{2}[f(k-1)+f(k+1)],$$

then

$$f(k) \leqslant (1-\tfrac{k}{n})f(0)+\tfrac{k}{n}f(n).$$

Proof. If we replace $f(k)$ by

$$\tilde{f}(k)=f(k)-(1-\tfrac{k}{n})f(0)-\tfrac{k}{n}f(n),$$

the same hypothesis holds. Thus we may assume $f(0)=0$ and $f(n)=0$. Let $g(k)=f(k)-f(k-1)$ for $1 \leqslant k \leqslant n$. Then our hypothesis states that $g(k) \leqslant g(k+1)$. Choose the integer m so that

$$g(1) \leqslant \cdots \leqslant g(m) \leqslant 0 \leqslant g(m+1) \leqslant \cdots g(n).$$

For any k

$$f(k)=\sum_{i=1}^{k} g(i) = -\sum_{i=k+1}^{n} g(i).$$

When $k \leqslant m$ the first representation is negative, and when $k \geqslant m$ the second is. This proves $f(k) \leqslant 0$ for $0 \leqslant k \leqslant n$, which is the desired conclusion for $f(0) = 0$ and $f(n) = 0$.

12.4 Corollary. *If $f(k)$ satisfies*

$$f(k) \leqslant \tfrac{1}{2}[f(k-1) + f(k+1)] + C,$$

then

$$f(k) \leqslant (1 - \tfrac{k}{n})f(0) + \tfrac{k}{n}f(n) + Ck(n-k).$$

Proof. Apply the previous result to $g(k) = f(k) + Ck^2$.

12.5 Corollary. *If $f(k)$ satisfies*

$$f(k) \leqslant Cf(k-1)^{1/2}f(k+1)^{1/2},$$

then

$$f(k) \leqslant C^{k(n-k)}f(0)^{1-k/n}f(n)^{k/n}.$$

Proof. Apply the previous result to $g(k) = \log f(k)$.

We let $\partial^n T = \{\partial_{i_1} \cdots \partial_{i_n} T_{j \cdots k}\}$ be the nth iterated covariant derivative of the tensor T.

12.6 Corollary. *If T is any tensor and if $1 \leqslant i \leqslant n-1$ then with a constant $C = C(n, m)$ depending only on n and $m = \dim X$ and independent of the metric g_{ij} or the connection Γ_{ij}^k we have the estimate*

$$\int |\partial^i T|^{2n/i}\, d\mu \leqslant C \max_X |T|^{2(n/i-1)} \int |\partial^n T|^2\, d\mu.$$

Proof. Applying the previous estimate to the tensor $\partial^{i-1}T$ when $2 \leqslant i \leqslant n-1$ with

$$p = \frac{2n}{i+1}, \quad q = \frac{2n}{i-1}, \quad r = \frac{n}{i} > 1,$$

we get

$$\left\{\int |\partial^i T|^{2n/i}\, d\mu\right\}^{i/n} \leqslant C\left\{\int |\partial^{i+1}T|^{2n/(i+1)}\, d\mu\right\}^{(i+1)/2n}\left\{\int |\partial^{i-1}T|^{2n/(i-1)}\, d\mu\right\}^{(i-1)/2n},$$

where $C = 2n/i - 2 + m$ depends only on m and n. Or when $i = 1$ we have

$$\left\{\int |\partial T|^{2n}\, d\mu\right\}^{1/n} \leqslant C\left\{\int |\partial^2 T|^n\right\}^{1/n} \cdot \max_X |T|$$

with $C = 2n - 2 + m$. Let

$$f(0) = \max_X |T|, \quad f(i) = \left\{\int |\partial^i T|^{2n/i}\right\}^{i/2n}, \qquad 1 \leq i \leq n.$$

Then we have

$$f(i) \leq Cf(i+1)^{1/2} f(i-1)^{1/2}$$

from the previous estimates. Therefore

$$f(i) \leq Cf(0)^{1-i/n} f(n)^{i/n}$$

with a constant C depending only on m. This proves the theorem, since

$$\left\{\int |\partial^i T|^{2n/i}\, d\mu\right\}^{i/2n} \leq C \max_X |T|^{1-i/n} \left\{\int |\partial^n T|^2\, d\mu\right\}^{i/2n}.$$

12.7 Corollary. *If T is any tensor then with a constant $C = C(n, m)$ depending only on n and $m = \dim X$ and independent of the metric g_{ij} and the connection Γ_{ij}^k we have the estimate for $0 \leq i \leq n$*

$$\int |\partial^i T|^2\, d\mu \leq C\left\{\int |\partial^n T|^2\, d\mu\right\}^{i/n} \left\{\int |T|^2\, d\mu\right\}^{1-i/n}.$$

Proof. If we apply Theorem 12.1 to the tensor $\partial^{i-1}T$ with $p = q = 2$ and $r = 1$ we get

$$\int |\partial^i T|^2\, d\mu \leq C\left\{\int |\partial^{i+1} T|^2\, d\mu\right\}^{1/2} \left\{\int |\partial^{i-1} T|^2\, d\mu\right\}^{1/2},$$

and the result now follows from Corollary 12.5.

13. Higher derivatives of the curvature

If A and B are two tensors we write $A * B$ for any linear combination of tensors formed by contraction on $A_{i\ldots j} B_{k\ldots l}$ using the g^{ik}. To avoid confusion between the Riemannian, Ricci and scalar curvatures we let

$$Rm = \{R_{ijkl}\} \quad \text{and} Rc = \{R_{ij}\}.$$

As before $\partial^n T$ is the nth iterated covariant derivative of a tensor T.

We want to derive the evolution equation for the nth covariant derivative $\partial^n Rm$ of the Riemannian curvature. To that end the following lemma is useful.

13.1 Lemma. *If A and B are tensors satisfying the evolution equation*

$$\frac{\partial}{\partial t}A = \Delta A + B,$$

then the covariant derivative ∂A satisfies an equation of the form

$$\frac{\partial}{\partial t}\partial A = \Delta\partial A + Rm * \partial A + A * \partial Rm + \partial B.$$

(In dimension 3 we may substitute Rc.)

Proof. The covariant derivative ∂ involves the Christöffel symbols Γ^i_{jk}, and their time derivative is

$$\Gamma'^i_{jk} = \tfrac{1}{2}g^{il}\{\partial_j g'_{kl} + \partial_k g'_{jl} - \partial_l g'_{jk}\},$$

which may be expressed in terms of ∂Rc since $g'_{ij} = -2R_{ij}$. Then

$$\frac{\partial}{\partial t}\partial A = \partial\frac{\partial A}{\partial t} + \partial Rm * A.$$

Now by interchanging derivatives

$$\partial\Delta A = \Delta\partial A + \partial Rm * A + Rm * \partial A,$$

and this completes the proof.

13.2 Theorem. *The nth covariant derivative $\partial^n Rm$ of the Riemannian curvature satisfies an evolution equation of the form*

$$\frac{\partial}{\partial t}\partial^n Rm = \Delta\partial^n Rm + \sum_{i+j=n}\partial^i Rm * \partial^j Rm.$$

Proof. If $n = 0$ we know this is true by Theorem 7.1, which gives the explicit form of the quadratic term. We proceed by induction on n, using the previous lemma. This gives

$$\frac{\partial}{\partial t}\partial^{n+1}Rm = \Delta\partial^{n+1}Rm + Rm * \partial^{n+1}Rm + \partial^n Rm * \partial^1 Rm$$
$$+\partial\left(\sum_{i+j=n}\partial^i Rm * \partial^j Rm\right),$$

and the result follows by the distributive rule for ∂.

13.3 Corollary. *For any n we have an evolution equation*

$$\frac{\partial}{\partial t}|\partial^n Rm|^2 = \Delta|\partial^n Rm|^2 - 2|\partial^{n+1}Rm|^2 + \sum_{i+j=n}\partial^i Rm * \partial^j Rm * \partial^n Rm.$$

Proof. This follows from the previous theorem. We have

$$\frac{\partial}{\partial t}|\partial^n Rm|^2 = 2\left\langle \partial^n Rm, \frac{\partial}{\partial t}\partial^n Rm \right\rangle + Rm * \partial^n Rm * \partial^n Rm,$$

where the extra terms come from the variation of the g^{ij} defining the norm $|\ |^2$. The usual computation gives

$$\Delta|\partial^n Rm|^2 = 2\langle \partial^n Rm, \Delta\partial^n Rm\rangle + 2|\partial^{n+1} Rm|^2,$$

and the result follows.

13.4 Theorem. *We have the estimate*

$$\frac{d}{dt}\int_X |\partial^n Rm|^2\, d\mu + 2\int_X |\partial^{n+1} Rm|^2\, d\mu \leqslant C \max_X |Rm| \int_X |\partial^n Rm|^2\, d\mu$$

with a constant C independent of the metric, depending only on the number n of derivatives and the dimension m of X.

Proof. Since $\int \Delta f\, d\mu = 0$ for any function f, if we integrate the equation in the previous corollary over X we only need to estimate the terms

$$\int_X |\partial^i Rm||\partial^j Rm||\partial^n Rm|\, d\mu$$

$$\leqslant \left\{\int_X |\partial^i Rm|^{2n/i}\, d\mu\right\}^{i/2n} \left\{\int_X |\partial^j Rm|^{2n/j}\, d\mu\right\}^{j/2n} \left\{\int_X |\partial^n Rm|^2\, d\mu\right\}^{1/2}$$

with $i + j = n$. By our interpolation result of Corollary 12.6 we have

$$\left\{\int_X |\partial^i Rm|^{2n/i}\, d\mu\right\}^{i/2n} \leqslant C \max_X |Rm|^{1-i/n} \left\{\int_X |\partial^n Rm|^2\, d\mu\right\}^{i/2n},$$

and doing the same for j the theorem follows. Recall the constant in Corollary 12.6 depends only on n and m.

14. Long time existence

Let X be a compact manifold of any dimension and let us be given any initial metric at $t = 0$.

14.1 Theorem. *The evolution equation*

$$\frac{\partial}{\partial t} g_{ij} = -2R_{ij}$$

has a unique solution on a maximal time interval $0 \leqslant t < T \leqslant \infty$. *If* $T < \infty$ *then* $\max_X |R_{ijkl}| \to \infty$ *as* $t \to T$.

Proof. Since we already know short time existence and uniqueness by the Nash-Moser inverse function theorem, we can take the maximum time interval

$0 \leqslant t < T$ on which the solution exists. We will show that if $T < \infty$ and $|R_{ijkl}| \leqslant C$ as $t \to T$, then the metric g_{ij} converges as $t \to T$ to a limit metric (which is strictly positive-definite), and all the derivatives converge also, showing the limit metric is smooth. We could then use the short time existence result to continue the solution past T, showing T is not maximal.

14.2 Lemma. *Let g_{ij} be a time-dependent metric on X for $0 \leqslant t < T \leqslant \infty$. Suppose*

$$\int_{t=0}^{T} \max_{X} |g'_{ij}|\, dt \leqslant C < \infty.$$

Then the metrics $g_{ij}(t)$ for all different times are equivalent, and they converge as $t \to T$ uniformly to a positive-definite metric tensor $g_{ij}(T)$ which is continuous and also equivalent.

Proof. Notice the argument is slightly subtle, since we measure the size of g'_{ij} with respect to g_{ij} which is changing;

$$|g'_{ij}|^2 = g^{ik} g^{jl} g'_{ij} g'_{kl}.$$

Fix a tangent vector $v \in TX$ at a point $x \in X$ and let

$$|v|_t^2 = g_{ij}(x, t) v^i v^j.$$

Then we take

$$\frac{d}{dt} |v|_t^2 = g'_{ij} v^i v^j,$$

and it follows by Cauchy-Schwartz that

$$\left| \frac{d}{dt} \log |v|_t^2 \right| \leqslant |g'_{ij}|.$$

Then for $0 \leqslant \tau \leqslant \theta < T$ we have

$$|\log |v|_\theta^2 / |v|_\tau^2| \leqslant \int_\tau^\theta |g'_{ij}|^2\, dt.$$

If the improper integral is finite, we see that all the metrics are equivalent. Moreover $|v_t|^2$ converges uniformly to a continuous function $|v|_T^2$ as $t \to T$ and $|v|_T \neq 0$ if $v \neq 0$. Since the parallelogram law

$$|v + w|^2 + |v - w|^2 = 2\big(|v|^2 + |w|^2\big)$$

continues to hold in the limit, the limiting norm comes from an inner product $g_{ij}(T)$, using the rule

$$g(v, w) = \tfrac{1}{4}\big(|v + w|^2 - |v - w|^2\big).$$

This completes the proof.

14.3 Lemma. *If $|Rm| \leqslant C$ on $0 \leqslant t < T$ and $T < \infty$, then for any n we can find a constant C_n with*

$$\int_X |\partial^n Rm|^2 \, d\mu \leqslant C_n.$$

Proof. This follows directly from Theorem 13.4 and the observation that if $df/dt \leqslant Cf$ on a finite time interval then we can bound f in terms of its initial data. (Hint: let $\tilde{f} = e^{-Ct} f$.)

We wish next to derive supremum norm estimates on $\partial^n Rm$. Since they are tensors, we use the following trick. First note that from the interpolation inequality in Corollary 12.6 we immediately get estimates

$$\int_X |\partial^n Rm|^p \, d\mu \leqslant C_{n,p},$$

for all n and $p < \infty$. Now let $E_n = |\partial^n Rm|^2$. Then for all $p < \infty$ we have estimates

$$\int_X (|E_n|^p + |\partial E_n|^p) \, d\mu \leqslant \tilde{C}_{n,p}$$

since E_n and ∂E_n can be expressed in terms of Rm and its covariant derivatives. But E_n is just a function, and by Sobolev's inequality if $p > n$

$$\max |f|^p \leqslant C_t \int_X (|f|^p + |\partial f|^p) \, d\mu.$$

Of course the constant C_t depends on the metric $g_{ij}(t)$ and hence on time t. But it does not depend on the derivatives of the g_{ij}, since it enters the expression on the right only through $|\partial f|^2 = g^{ij} \partial_i f \partial_j f$ and the measure $d\mu = \mu(x) dx$ with $\mu(x) = \sqrt{\det g_{ij}}$. The derivative $\partial_i f = \partial f / \partial x^i$ is independent of the connection Γ_{ij}^k. Thus for functions the constant C_t is uniformly bounded as $t \to T$, since the metrics are all equivalent by Lemma 14.2. Applying this estimate to E_n we get the following result.

14.4 Lemma. *If $|Rm| \leqslant C_0$ on $0 \leqslant t < T$ and $T < \infty$ then $|\partial^n Rm| \leqslant C_n$ for all n. The constant C_n depends only on the initial value of the metric and the constant C_0.*

Of course the estimates on $Rm = \{R_{ijkl}\}$ imply ones on $Rc = \{R_{ij}\}$. Since $\partial g_{ij}/\partial t = -2R_{ij}$, it is easy to see that the $g_{ij}(t)$ have all their derivatives bounded, and converge to the limit metric $g_{ij}(T)$ in the C^∞ topology as $t \to T$. This completes the proof of Theorem 14.1.

15. Controlling $R_{\max}/R_{\min}$

We return to the case where X is a compact three-manifold and the initial metric has strictly positive Ricci curvature. The unnormalized evolution equation $\partial g_{ij}/\partial t = -2R_{ij}$ will have a solution on a maximal time interval $0 \leqslant t < T$. We know from Lemma 11.11 that $T < \infty$, and then from Theorem 14.1 that $\max|R_{ij}| \to \infty$ as $t \to T$. Since $|R_{ij}|^2 = S \leqslant R^2$, we have $R_{\max} \to \infty$ as $t \to T$, where $R_{\max}$ is the maximum value of R and $R_{\min}$ will be its minimum value. We want to estimate $R_{\max}/R_{\min}$.

15.1 Theorem. *We have* $R_{\max}/R_{\min} \to 1$ *as* $t \to T$.

Proof. By Theorem 11.1 we know that for every $\eta > 0$ we can find a constant $C(\eta)$ with

$$|\partial_i R| \leqslant \tfrac{1}{2}\eta^2 R^{3/2} + C(\eta)$$

on $0 \leqslant t < T$. Since $R_{\max} \to \infty$ as $t \to T$, we can find θ with $C(\eta) \leqslant \frac{1}{2}\eta^2 R_{\max}^{3/2}$ for $\theta \leqslant t < T$. Then $|\partial_i R| \leqslant \eta^2 R_{\max}^{3/2}$ for $t \geqslant \theta$.

Fix a point $x \in X$ where R assumes its maximum. Then on any geodesic out of x of length at most $s = 1/\eta R_{\max}^{1/2}$ we have $R \geqslant (1-\eta)R_{\max}$. We claim that when $\eta > 0$ is small enough then this includes all of X. For $R_{ij} \geqslant \varepsilon R g_{ij}$ for some $\varepsilon > 0$. It follows that every geodesic from x of length s has a conjugate point when η is small by the following well-known theorem of Myers, which the reader will find in Cheeger and Ebin [2, Theorem 1.26(1)].

15.2 Theorem (*Myers*). *If* $R_{ij} \geqslant (m-1)Hg_{ij}$ *along a geodesic of length at least* $\pi H^{-1/2}$ *on a manifold of dimension* m *then the geodesic has conjugate points.*

Thus we can reach every point of X by a geodesic of length at most s, and hence $R_{\min} \geqslant (1-\eta)R_{\max}$. It follows that $R_{\max}/R_{\min} \to 1$ as $t \to T$.

15.3 Theorem. *We have*

$$\int_0^T R_{\max}\, dt = \infty.$$

Proof. Choose a function $f(t)$ equal to $R_{\max}$ at $t = 0$ and solving the ordinary differential equation

$$\frac{df}{dt} = 2R_{\max} f,$$

which is possible since $R_{\max}$ is a continuous function of t. Since $S \leqslant R^2$, we have

$$\frac{\partial}{\partial t}(R-f) \leqslant \Delta(R-f) + 2R_{\max}(R-f),$$

and hence $R \leqslant f$ on $0 \leqslant t < T$ by the maximum principle. Since $R_{\max} \to \infty$ as $t \to T$, we have $f \to \infty$ also. But

$$\log f(t)/f(0) = 2\int_0^t R_{\max}(\theta)\, d\theta,$$

and hence the integral diverges as $t \to T$.

15.4 Corollary. *If r is the average scalar curvature, then*

$$\int_0^T r\, dt = \infty.$$

Proof. We have $R_{\min} \leqslant r \leqslant R_{\max}$ and $R_{\max}/R_{\min} \to 1$ as $t \to T$.

15.5 Theorem. $S/R^2 - \frac{1}{3} \to 0$ *as* $t \to T$.

Proof. By Theorem 10.1 we have

$$S/R^2 - \tfrac{1}{3} \leqslant CR^{-\delta},$$

and $R_{\min} \to \infty$ (since $R_{\max} \to \infty$ and $R_{\max}/R_{\min} \to 1$).

16. Estimating the normalized equation

Next we consider how to convert our estimates for the unnormalized equation

$$(*) \qquad \frac{\partial}{\partial t} g_{ij} = -2R_{ij}$$

into estimates on the normalized equation

$$(**) \qquad \frac{\partial}{\partial t} \tilde{g}_{ij} = \tfrac{2}{3}\tilde{r}\tilde{g}_{ij} - 2\tilde{R}_{ij}.$$

Let $(*)$ have a solution on a maximal interval $0 \leqslant t < T$ and let $(**)$ have a corresponding solution on $0 \leqslant \tilde{t} < \tilde{T}$ related by the transformation equations given in §3.

16.1 Lemma. $\tilde{R}_{\max}/\tilde{R}_{\min} \to 1$ *as* $\tilde{t} \to \tilde{T}$.

Proof. Since we are dilating by a constant, the ratio is unchanged.

16.2 Lemma. $\tilde{R}_{ij} \geqslant \varepsilon \tilde{R}\tilde{g}_{ij}$ *for some* $\varepsilon > 0$.

Proof. Again both sides stretch equally under dilations.

16.3 Lemma. $\tilde{R}_{\max} \leqslant C < \infty$ *on* $0 \leqslant \tilde{t} < \tilde{T}$.

Proof. Let the metric $\tilde{g}_{ij}$ have volume $\tilde{V}$ and diameter $\tilde{d}$. Then $\tilde{V} \leqslant C\tilde{d}^3$, and since $\tilde{R}_{ij} \geqslant \varepsilon \tilde{R}\tilde{g}_{ij}$ we have $\tilde{d} \leqslant C\tilde{R}_{\min}^{-1/2}$ by Myer's Theorem 15.2. Thus $\tilde{V}\tilde{R}_{\min}^{3/2} \leqslant C$. But for the normalized equation the volume $\tilde{V} = 1$. Thus $\tilde{R}_{\min} \leqslant C$. Then $\tilde{R}_{\max} \leqslant C$ also from Lemma 16.1.

16.4 Lemma. $\tilde{T} = \infty$.

Proof. Since $d\tilde{t}/dt = \psi$ and $\psi\tilde{r} = r$ we have

$$\int_0^{\tilde{T}} \tilde{r}\, d\tilde{t} = \int_0^{T} r\, dt = \infty$$

by Corollary 15.4. But $\tilde{r} \leqslant \tilde{R}_{\max} \leqslant C$, so we must have $\tilde{T} = \infty$.

16.5 Lemma. *$\tilde{S}/\tilde{R}^2 - \frac{1}{3} \to 0$ as $\tilde{t} \to \infty$.*

Proof. Again this follows from Theorem 15.5 since the expression is invariant under dilation.

Since we have the relation

$$\tilde{S} - \tfrac{1}{3}\tilde{R}^2 = \tfrac{1}{3}\left[(\tilde{\lambda} - \tilde{\mu})^2 + (\tilde{\lambda} - \tilde{\nu})^2 + (\tilde{\mu} - \tilde{\nu})^2\right],$$

it follows that the ratio $\tilde{\lambda}/\tilde{\mu}$ of any two eigenvalues of $\tilde{R}_{ij}$ converges to 1 as $\tilde{t} \to \infty$. Since $\tilde{R}_{\max}/\tilde{R}_{\min} \to 1$ as $t \to \infty$ also, it must eventually happen that the sectional curvature is $\frac{1}{4}$ pinched, or indeed as pinched as we like. At this point it follows from the Sphere Theorem (see Cheeger and Ebin [2, Theorem 6.1]) that the universal cover of X is a sphere. However, we shall only borrow a lemma.

16.6 Lemma. (*Klingenberg*). *Let X be a simply connected manifold of dimension* 3 *or more whose sectional curvature is pinched between K and $\frac{1}{4}K$. Then the injectivity radius of X is at least $\pi/\sqrt{K}$.*

Proof. See Cheeger and Ebin [2, Theorem 5.10].

We apply this result to the universal cover Y of X. The constant K will be proportional to $R_{\min}$. The volume is at least some multiple of the injectivity radius. Thus we get an estimate $\tilde{R}_{\min}^{-3/2} \leqslant C \cdot \widetilde{\mathrm{Vol}}(Y)$. But X has volume one for the normalized equation, and then the volume of its universal cover Y is just the number of elements in the fundamental group of X (which is finite by Myer's theorem). This gives a proof of the following.

16.7 Lemma. *We can find $\varepsilon > 0$ such that $\tilde{R}_{\min} \geqslant \varepsilon$ on $0 \leqslant \tilde{t} < \infty$.*

17. Exponential convergence

We start with a principle for converting from the unnormalized to the normalized evolution equation. Let P and Q be two expressions formed from the metric and curvature tensors, and let $\tilde{P}$ and $\tilde{Q}$ be the corresponding expressions for the normalized equation. Since they differ by dilations, they differ by a power of ψ. We say P has degree n if $\tilde{P} = \psi^n P$. Thus g_{ij} has degree 1, R_{ij} has degree 0, R has degree -1, and S has degree -2.

17.1 Lemma. *Suppose P satisfies*

$$\frac{\partial P}{\partial t} = \Delta P + Q$$

for the unnormalized equation, and P has degree n. Then Q has degree $n - 1$, *and for the normalized equation*

$$\frac{\partial \tilde{P}}{\partial \tilde{t}} = \tilde{\Delta}\tilde{P} + \tilde{Q} + \frac{2}{3}n\tilde{r}\tilde{P}.$$

Proof. We see Q has degree $n - 1$ since $\partial\tilde{t}/\partial t = \psi$ and $\Delta = \psi\tilde{\Delta}$. Then

$$\psi\frac{\partial}{\partial\tilde{t}}(\psi^{-n}\tilde{P}) = \psi\tilde{\Delta}(\psi^{-n}\tilde{P}) + \psi^{-n+1}\tilde{Q},$$

$$\frac{\partial}{\partial\tilde{t}}\tilde{P} = \tilde{\Delta}\tilde{P} + \tilde{Q} + \frac{n}{\psi}\frac{\partial\psi}{\partial\tilde{t}}\tilde{P}.$$

But from §3 we know $d\log\psi/dt = \frac{2}{3}r$, so $d\log\psi/d\tilde{t} = \frac{2}{3}\tilde{r}$. This prove the lemma.

Now from §16 we know that the normalized equation (**) has a solution on $0 \leqslant t < \infty$ with

$$0 < \varepsilon \leqslant \tilde{R}_{\min} \leqslant \tilde{R}_{\max} \leqslant C,$$

$$\varepsilon\tilde{R}\tilde{g}_{ij} \leqslant \tilde{R}_{ij} \leqslant \tilde{R}\tilde{g}_{ij},$$

$$\tilde{R}_{\max}/\tilde{R}_{\min} \to 1 \quad \text{and} \quad \tilde{S}/\tilde{R}^2 - \tfrac{1}{3} \to 0 \qquad \text{as } t \to \infty.$$

We want to show the convergence is exponential.

17.2 Lemma. *We can find constants* $C < \infty$ *and* $\delta > 0$ *such that*

$$\tilde{S} - \tfrac{1}{3}\tilde{R}^2 \leqslant Ce^{-\delta\tilde{t}}.$$

Proof. We let $\tilde{f} = \tilde{S}/\tilde{R}^2 - \frac{1}{3}$. Note $\tilde{f}$ has degree 0. Then by Lemma 10.5 with $\gamma = 2$ we have

$$\frac{\partial\tilde{f}}{\partial\tilde{t}} \leqslant \tilde{\Delta}\tilde{f} + \frac{2}{\tilde{R}}\tilde{g}^{pq}\partial_p\tilde{R}\partial_q\tilde{f} - 4\tilde{P}/\tilde{R}^3,$$

and by Lemma 10.7 we have

$$\tilde{P} \geqslant \varepsilon^2\tilde{S}(\tilde{S} - \tfrac{1}{3}\tilde{R}^2).$$

This makes

$$\frac{\partial\tilde{f}}{\partial\tilde{t}} \leqslant \tilde{\Delta}\tilde{f} + \tilde{u}^k\partial_k\tilde{f} - \delta\tilde{f}$$

with $\tilde{u}^q = 2\tilde{g}^{pq}\partial_p\tilde{R}/\tilde{R}$ and $\delta = 4\varepsilon^2/3C$ since

$$4\tilde{P}/\tilde{R}^3 \geqslant 4\varepsilon^2\tilde{S}\tilde{f}/\tilde{R}^3 \geqslant 4\varepsilon^2\tilde{f}/3\tilde{R} \geqslant \delta\tilde{f}$$

with $\tilde{S} \geqslant \frac{1}{3}\tilde{R}^2$ and $\tilde{R} \leqslant C$. But then

$$\frac{\partial}{\partial\tilde{t}}\left(e^{\delta\tilde{t}}\tilde{f}\right) \leqslant \tilde{\Delta}\left(e^{\delta\tilde{t}}\tilde{f}\right) + \tilde{u}^k\partial_k\left(e^{\delta\tilde{t}}\tilde{f}\right),$$

and by the maximum principle $e^{\delta\tilde{t}}\tilde{f} \leqslant C$. Thus $\tilde{f} \leqslant Ce^{-\delta\tilde{t}}$. Since $\tilde{R}$ is bounded above and below, this is equivalent to the theorem.

17.3 Corollary. $|\tilde{R}_{ij} - \frac{1}{3}\tilde{R}\tilde{g}_{ij}| \leqslant Ce^{-\delta\tilde{t}}$.

Proof. The eigenvalues of the matrix are of the form

$$\lambda - \tfrac{1}{3}(\lambda + \mu + \nu) = \tfrac{1}{3}[(\lambda - \mu) + (\lambda - \nu)],$$

while

$$\tilde{S} - \tfrac{1}{3}\tilde{R}^2 = \tfrac{1}{3}\left[(\lambda - \mu)^2 + (\lambda - \nu)^2 + (\mu - \nu)^2\right].$$

The estimate follows.

17.4 Lemma. *We can find constants* $C < \infty$ *and* $\delta > 0$ *such that*

$$\tilde{R}_{\max} - \tilde{R}_{\min} \leqslant Ce^{-\delta\tilde{t}}.$$

Proof. This time we let

$$\tilde{F} = |\partial_i\tilde{R}|^2/\tilde{R} + 168\left(\tilde{S} - \tfrac{1}{3}\tilde{R}^2\right).$$

Then F has degree -2, and from Lemmas 11.7 and 11.9 (with $\eta = 0$) we get

$$\frac{\partial}{\partial\tilde{t}}\tilde{F} \leqslant \Delta\tilde{F} + 672\tilde{R}\left(\tilde{S} - \tfrac{1}{3}\tilde{R}^2\right) - \frac{4}{3}\tilde{r}\tilde{F},$$

since Lemma 17.1 also works for inequalities. Using our estimate from Lemma 17.2

$$\frac{\partial}{\partial\tilde{t}}\tilde{F} \leqslant \Delta\tilde{F} + Ce^{-\delta\tilde{t}} - \delta\tilde{F}$$

for some $C < \infty$, $\delta > 0$ and $\varepsilon > 0$, since $\tilde{R} \leqslant C$ and $\tilde{r} \geqslant \tilde{R}_{\min} \geqslant \varepsilon > 0$. But this makes

$$\frac{\partial}{\partial\tilde{t}}\left(e^{\delta\tilde{t}}\tilde{F} - C\tilde{t}\right) \leqslant \tilde{\Delta}\left(e^{\delta\tilde{t}}\tilde{F} - C\tilde{t}\right),$$

and by the maximum principle we have $e^{\delta\tilde{t}}\tilde{F} - C\tilde{t} \leqslant C$. Then $\tilde{F} \leqslant C(1 + \tilde{t})e^{-\delta\tilde{t}}$, and since $\delta > 0$ is arbitrary this proves the theorem (by taking a slightly smaller δ).

17.5 Corollary. $|\tilde{R}_{ij} - \frac{1}{3}\tilde{r}\tilde{g}_{ij}| \leq Ce^{-\delta\tilde{t}}$.

Proof. This follows from the last two lemmas since

$$|\tilde{R}_{ij} - \tfrac{1}{3}\tilde{r}\tilde{g}_{ij}| \leq |\tilde{R}_{ij} - \tfrac{1}{3}\tilde{R}\tilde{g}_{ij}| + |\tilde{R} - \tilde{r}|/\sqrt{3}.$$

Using Lemma 14.2, we get the following result.

17.6 Theorem. *The metrics* $\tilde{g}_{ij}(\tilde{t})$ *are all equivalent, and converge as* $\tilde{t} \to \infty$ *uniformly to a continuous positive-definite metric* $\tilde{g}_{ij}(\infty)$.

To estimate higher derivatives we return to Theorem 13.4. Notice that all three terms have the same degree of homogeneity, and hence the same result holds for the normalized evolution equation. Since in three dimensions the Riemannian curvature Rm is entirely determined by the Ricci curvature Rc, we have the estimate

$$\frac{d}{d\tilde{t}}\int_X |\partial^n\tilde{R}c|^2\,d\tilde{\mu} + 2\int_X |\partial^{n+1}\tilde{R}c|^2\,d\tilde{\mu} \leq C\max_X |\tilde{R}c| \int_X |\partial^n\tilde{R}c|^2\,d\tilde{\mu}$$

and $\max_X |\tilde{R}c| \leq C$. We introduce the tensor $\tilde{E} = \{\tilde{E}_{ij}\}$ defined by

$$\tilde{E}_{ij} = \tilde{R}_{ij} - \tfrac{1}{3}\tilde{r}\tilde{g}_{ij},$$

and observe that $\partial^n\tilde{R}c = \partial^n\tilde{E}$ for $n > 0$, since $\tilde{r}$ is constant. Then interpolating by Corollary 12.7

$$\int_X |\partial^n\tilde{R}c|^2\,d\tilde{\mu} \leq C\left\{\int_X |\partial^{n+1}\tilde{R}c|^2\,d\tilde{\mu}\right\}^{n/(n+1)}\left\{\int_X |\tilde{E}|^2\,d\tilde{\mu}\right\}^{1/(n+1)}.$$

Now for any $\varepsilon > 0$ and all $x, y > 0$ we have

$$x^n y \leq C\varepsilon x^{n+1} + C\varepsilon^{-n}y^{n+1},$$

and applying this above gives

$$\int_X |\partial^n\tilde{R}c|^2\,d\tilde{\mu} \leq C\varepsilon\int_X |\partial^{n+1}\tilde{R}c|^2\,d\tilde{\mu} + C\varepsilon^{-n}\int_X |\tilde{E}|^2\,d\tilde{\mu}.$$

Then we get the following result.

17.7 Lemma. *For every n we have*

$$\int_X |\partial^n\tilde{R}c|^2\,d\tilde{\mu} \leq C$$

with a constant depending on n.

Proof. If we choose $\varepsilon > 0$ so small that $C\varepsilon \leq 2$ we can substitute this in the previous equation and get

$$\frac{d}{dt}\int_X |\partial^n\tilde{R}c|^2\,d\tilde{\mu} \leq C\int_X |\tilde{E}|^2\,d\tilde{\mu}.$$

But we know $|\tilde{E}| \leqslant Ce^{-\delta\tilde{t}}$ for some $\delta > 0$ by Corollary 17.5, and the lemma follows.

Next we use the interpolation estimate of Corollary 12.6, which immediately gives the following result.

17.8 Lemma. *For every $n > 0$ and every $p < \infty$ we have*

$$\int_X |\partial^n \tilde{R}c|^p \, d\tilde{\mu} \leqslant Ce^{-\delta\tilde{t}}$$

for some constants $C < \infty$ and $\delta > 0$ depending on n and p.

Proof. This follows immediately from Corollary 12.6 since for $1 \leqslant i \leqslant n-1$

$$\int_X |\partial^i \tilde{R}c|^{2n/i} \, d\tilde{\mu} \leqslant C \max_X |\tilde{E}|^{2(n/i-1)} \int_X |\partial^n \tilde{R}c|^2 \, d\tilde{\mu},$$

and the maximum norm of $\tilde{E}$ decreases exponentially while the L_2 norm of $\partial^n \tilde{R}c$ is bounded.

17.9 Theorem. *For every $n > 0$ we have*

$$\max_X |\partial^n \tilde{R}c| \leqslant Ce^{-\delta\tilde{t}}$$

for some constants $C < \infty$ and $\delta > 0$ depending on n.

Proof. We repeat the argument of Lemma 14.4. The function $\tilde{E}_n = |\partial^n \tilde{R}c|^2$ is exponentially decreasing in L_p norm for all $p < \infty$ as are its first derivatives. Since the metrics $\tilde{g}_{ij}(\tilde{t})$ are all equivalent as $\tilde{t} \to \infty$, we can apply the Sobolev estimate with a uniform constant to show the supremum norm of E_n is also exponentially decreasing.

17.10 Corollary. *As $\tilde{t} \to \infty$ the metrics $\tilde{g}_{ij}(\tilde{t})$ converge to the limit metric $\tilde{g}_{ij}(\infty)$ in the C^∞ topology. Hence $\tilde{g}_{ij}(\infty)$ is smooth, and the curvatures $\tilde{R}_{ij}(\tilde{t})$ converge to the curvature $\tilde{R}_{ij}(\infty)$.*

Proof. This follows directly from the previous result since

$$\frac{\partial}{\partial \tilde{t}} \tilde{g}_{ij} = \frac{2}{3} \tilde{r} \tilde{g}_{ij} - 2\tilde{R}_{ij}.$$

17.11 Corollary. *The limit metric $\tilde{g}_{ij}(\infty)$ has constant positive curvature.*

Proof. By Corollary 17.5 the tensor $\tilde{R}_{ij} - \frac{1}{3}\tilde{r}\tilde{g}_{ij}$ converges uniformly to zero. This proves Main Theorem 1.1.

References

[1] J. P. Bourguignon, *Ricci curvature and Einstein metrics*, Lecture Notes in Math., Vol. 838, Springer, Berlin, p. 298.

[2] J. Cheeger & D. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland, Amsterdam, 1975, p. 174.

[3] J. Eells & J. H. Sampson, *Harmonic mappings of Riemannian manifolds*, Amer. J. Math. **86** (1964) 109–160.

[4] R. S. Hamilton, *Harmonic maps of manifolds with boundary*, Lecture Notes in Math., Vol. 471, Springer, Berlin, 1975, p. 168.

[5] ______, *The inverse function theorem of Nash and Moser* (new version), preprint, Cornell University, p. 294.

[6] J. A. Wolf, *Spaces of constant curvature*, McGraw-Hill, New York, 1967, p. 408.CORNELL

CORNELL UNIVERSITY