J. DIFFERENTIAL GEOMETRY
37 (1993) 225-243

THE HARNACK ESTIMATE FOR THE RICCI FLOW

RICHARD S. HAMILTON

1. The result

1.1. Main Theorem. *Let g_{ij} be a complete solution with bounded curvature to the Ricci flow*

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

on a manifold M for t in some time interval $0 < t < T$ and suppose g_{ij} has a weakly positive curvature operator, so that

$$R_{ijkl}U_{ij}U_{kl} \geq 0$$

for all two-forms U_{ij}. Let

$$P_{ijk} = D_iR_{jk} - D_jR_{ik}$$

and let

$$M_{ij} = \Delta R_{ij} - \frac{1}{2}D_iD_jR + 2R_{ikjl}R_{kl} - R_{ik}R_{jk} + \frac{1}{2t}R_{ij}.$$

Then for any one-form W_i and any two-form U_{ij} we have

$$M_{ij}W_iW_j + 2P_{ijk}U_{ij}W_k + R_{ijkl}U_{ij}U_{kl} \geq 0.$$

1.2. Corollary. *For any one-form V_i we have*

$$\frac{\partial R}{\partial t} + \frac{R}{t} + 2D_iR \cdot V_i + 2R_{ij}V_iV_j \geq 0.$$

The corollary follows immediately by taking

$$U_{ij} = \tfrac{1}{2}(V_iW_j - V_jW_i)$$

and tracing over W_i.

The existence of inequalities on the second derivatives of solutions of parabolic equations was first noted by Peter Li and S.-T. Yau [12] for the scalar heat flow on a Riemannian manifold. The author has observed a similar phenomenon for the matrix of second derivatives in the scalar heat

Received December 4, 1991. The author's research was supported by NSF grant #DMS 90-03333.

flow [9] and for the Ricci flow on a surface [6] and for the mean curvature flow [10]. Ben Chow has derived similar results for the Gauss curvature flow [3] and the Yamabe flow [4]. These are called Harnack estimates because they can be integrated along paths in space-time to produce comparisons of the solution between different points at earlier and later times, as in the classical Harnack estimates.

1.3. Corollary. *Let X_1 and X_2 be two points in M and let t_1 and t_2 be two different times with $0 < t_1 < t_2$. Let $c(X_1, X_2, t_1)$ be the distance between X_1 and X_2 at time t_1. Then*

$$R(X_2, t_2) \geq \frac{t_1}{t_2} e^{-d(X_1, X_2, t_1)^2/2(t_2-t_1)} R(X_1, t_1).$$

Proof. Take the geodesic path $X(t)$ from X_1 to X_2 at time t_1 parametrized proportional to arc length with parameter t starting at X_1 at time t_1 and ending at X_2 at time t_2. At time t_1 the constant velocity is

$$d(X_1, X_2, t_1)/(t_2 - t_1).$$

Now consider the path $(X(t), t)$ in space-time. Since the curvature is weakly positive the metric g_{ij} will be weakly shrinking, so if we take the velocity vector $V^i = \frac{d}{dt} X^i$ its length at time $t \geq t_1$ will be no more than its length was at time t_1. Thus

$$g_{ij}(X(t), t) V^i V^j \leq d(X_1, X_2, t_1)^2/(t_2 - t_1)^2.$$

Now from Corollary 1.2 (and replacing V by $\frac{1}{2}V$) we have the estimate

$$\frac{dR}{dt} = \frac{\partial R}{\partial t} + D_i R \cdot V^i \geq -\frac{R}{t} - \frac{1}{2} R_{ij} V^i V^j,$$

where the total derivative dR/dt is the rate of change of R along the path in space-time. Now $R_{ij} \leq R g_{ij}$ for weakly positive Ricci curvature, so

$$\frac{d}{dt} \log R \geq -\frac{1}{t} - \frac{1}{2} g_{ij} V^i V^j.$$

We use the estimate on the length of V given above and integrate over time to get

$$\log \frac{R(X_2, t_2)}{R(X_1, t_1)} \geq \log \frac{t_1}{t_2} - \frac{d(X_1, X_2, t_1)^2}{2(t_2 - t_1)}.$$

Now exponentiating and rearranging gives the desired result.

2. The notation

We interpret covariant differentiation in terms of vector fields on the frame bundle. Let M be the manifold, X a point in M, and $Y =$

$(Y_1, Y_2, \cdots, Y_n)$ a frame at X consisting of a basis for the space TM_x of tangent vectors at X. A symmetric connection Γ on the tangent bundle determines a choice of horizontal tangent vectors on the bundle $F(M)$ of all frames. We let D_a be the vector field on the frame bundle which is horizontal and projects onto the vector Y_a in M when we are at the point Y in $F(M)$. If f is a function on the frame bundle, we denote by $D_a f$ the derivative of f in the direction of the vector field D_a. Any tensor gives rise to a system of functions on the frame bundle. For example if $V: TM \to R$ is a covector then

$$V = \{V_a\} \quad \text{where } V(Y_a) = V_a,$$

or if $V: TM \times TM \to TM$ is a tensor then

$$V = \{V_{ab}^c\} \quad \text{where } V(Y_a, Y_b) = V_{ab}^c Y_c.$$

We can thus interpret the covariant derivative as applying the vector field D_a to the component functions of the tensor V. In the first case $V = \{V_a\}$ we have

$$DV = \{D_a V_b\} \quad \text{where } DV(Y_a)(Y_b) = D_a V_b,$$

and in the second case

$$DV = \{D_a V_{bc}^d\} \quad \text{where } DV(Y_a)(Y_b, Y_c) = D_a V_{bc}^d Y_d.$$

The same applies to any tensor.

In local coordinates $X = \{x^i\}$ and $Y = \{y_a^i\}$ where $Y_a = y_a^i \partial/\partial x^i$. The vector fields are given locally by

$$D_a = y_a^i \left[\frac{\partial}{\partial x^i} - \Gamma_{ij}^k(x) y_b^i \frac{\partial}{\partial y_b^k} \right],$$

where $\Gamma_{ij}^k(x)$ are the Christoffel symbols of the connection. We then have

$$D_a V_b = y_a^i y_b^j D_i V_j,$$

where

$$D_i V_j = \frac{\partial V_j}{\partial x^i} - \Gamma_{ij}^k(x) V^k$$

is the usual local formula for covariant derivative.

We also have vector fields tangent to the fibres of the frame bundle representing the action of the change of frame group $Gl(n)$. These are the vector fields

$$\nabla_b^a = y_b^i \frac{\partial}{\partial y_a^i}$$

representing the skewing of the a-axis into the b-axis for $a \neq b$, or the stretching of the axis when $a = b$. The action of ∇_b^a on tensors is easy to describe. For example, if $V_c = y_c^j V_j(x)$ comes from a covector then $\nabla_b^a V_c = I_c^a V_b$, while if $V_{ab} = y_a^i y_b^j V_{ij}(x)$ then $\nabla_b^a V_{cd} = I_c^a V_{bd} + I_d^a V_{cb}$ and so on.

It is important to compute the commutators of these vector fields. The D_a and ∇_a^b form a basis for the tangent vectors on the frame bundle $F(M)$. The curvature tensor is given by R_{abd}^c where

$$R_{abd}^c y_c^k = R_{ijl}^k y_a^i y_b^j y_d^l$$

and as usual

$$R_{ijl}^k = \frac{\partial}{\partial x^i}\Gamma_{jl}^k - \frac{\partial}{\partial x^j}\Gamma_{il}^k + \Gamma_{im}^k \Gamma_{jl}^m - \Gamma_{jm}^k \Gamma_{il}^m .$$

Then we can easily compute the commutator relation

$$D_a D_b - D_b D_a = R_{abd}^c \nabla_c^d .$$

We also have

$$\nabla_b^a D_c - D_c \nabla_b^a = I_c^a D_b$$

and

$$\nabla_b^a \nabla_d^c - \nabla_d^c \nabla_b^a = I_d^a \nabla_b^c - I_b^c \nabla_d^a .$$

Suppose now that we have a Riemannian metric g with Levi-Civita connection Γ. The metric defines the system of functions $g_{ab} = g(Y_a, Y_b)$ on the frame bundle $F(M)$. The orthonormal frame bundle is the subbundle $OF(M)$ where $g_{ab} = I_{ab}$. The vector fields ∇_{ab} are not tangent. For this purpose we introduce the vector fields

$$\delta_{ab} = g_{ac}\nabla_b^c - g_{bc}\nabla_a^c ,$$

which represent infinitesimal notations of the ab-planes. Then D_a and δ_{bc} form a basis for the vector fields on $OF(M)$. For covectors we have

$$\delta_{ab} V_c = g_{ac} V_b - g_{bc} V_a ,$$

and similar formulas hold for other tensors. Thus

$$\delta_{ab} V_{cd} = g_{ac} V_{bd} + g_{ad} V_{cb} - g_{bc} V_{ad} - g_{bd} V_{ca}$$

for a 2-tensor, and in particular $\delta_{ab} g_{cd} = 0$ which shows that δ_{ab} is indeed tangent to the subbundle of orthonormal frames where $g_{ab} = I_{ab}$.

The commutator of D_a and D_b is given by

$$D_a D_b - D_b D_a = \tfrac{1}{2} R_{abcd} \delta_{cd} ,$$

where $R_{abcd} = g_{ce} R^e_{abd}$. Thus for covectors we have

$$D_a D_b V_c - D_b D_a V_c = R_{abcd} V_d ,$$

and similar formulas hold for other tensors. [Note we can sum over lowered indices since $g_{ab} = I_{ab}$.] The other commutators are

$$\delta_{ab} D_c - D_c \delta_{ab} = g_{ac} D_b - g_{bc} D_a$$

and

$$\delta_{ab}\delta_{cd} - \delta_{cd}\delta_{ab} = g_{ac}\delta_{bd} + g_{bd}\delta_{ac} - g_{ad}\delta_{bc} - g_{bc}\delta_{ad} .$$

For example in three dimensions we have the usual formulas

$$[\delta_{12}, \delta_{13}] = \delta_{23}, \quad [\delta_{13}, \delta_{23}] = \delta_{12}, \quad [\delta_{23}, \delta_{12}] = \delta_{13},$$

which generalize as above.

When we come to look at the Ricci flow, we must cross the manifold with the time axis t. We can then just cross the whole frame bundle with t also. There is thus a single new vector field $\frac{\partial}{\partial t}$ on the frame bundle cross time. If the metric were constant, the orthonormal frame bundle would also be constant. But since the metric varies according to the formula

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

we see that the orthonormal frame bundle where $g_{ab} = I_{ab}$ will now vary with time. Therefore we modify the timelike vector field to make it tangent to the orthonormal frame bundle. We let

$$D_t = \frac{\partial}{\partial t} + R_{ab} g^{bc} \nabla^a_c ,$$

so that D_t is the unique vector field which is tangent to the orthonormal frame bundle and has the property that $D_t - \frac{\partial}{\partial t}$ is a space-like vector orthogonal to the orthonormal frame bundle [in the metric on $F(M)$ where D_a and ∇^b_c are an orthonormal basis]. We can then define the time covariant derivative of a tensor by differentiating its components in the direction D_t. Thus for example for a covector V_a we have

$$D_t V_a = \frac{\partial}{\partial t} V_a + R_{ab} g^{bc} V_c$$

and similar formulas for other tensors. Indeed

$$D_t g_{ab} = \frac{\partial}{\partial t} g_{ab} + R_{ac} g^{cd} g_{bd} + R_{bc} g^{cd} g_{ad} = 0,$$

which shows D_t is tangent to the orthonormal frame bundle. (We obtain the same formulas as from the procedure in [8] where instead we vary the orthonormal frame.) We can compute the time derivative of the Riemannian curvature tensor as

$$D_t R_{abcd} = \Delta R_{abcd} + 2(B_{abcd} - B_{abdc} + B_{acbd} - B_{adbc}),$$

where $\Delta = D_e D_e$ and $B_{abcd} = R_{aebf} R_{cedf}$ as in [8].

For our computations we will need the commutators of D_t and D_a and δ_{ab}. Recall that

$$D_a = y_a^i \left\{ \frac{\partial}{\partial x^i} - \Gamma_{ij}^k(x) y_b^i \frac{\partial}{\partial y_b^k} \right\}.$$

For the Ricci flow we have

$$\frac{\partial}{\partial t} \Gamma_{ij}^k = -g^{kl}(D_i R_{jl} + D_j R_{il} - D_l R_{ij}),$$

so on the frame bundle we compute

$$\frac{\partial}{\partial t} D_a - D_a \frac{\partial}{\partial t} = g^{cd}(D_a R_{bc} + D_b R_{ac} - D_c R_{ab}) \nabla_d^b.$$

Since

$$D_t = \frac{\partial}{\partial t} + R_{bc} g^{cd} \nabla_d^b$$

and we know the commutator of D_a with ∇_d^b, we can compute

$$D_t D_a - D_a D_t = R_{ab} D_b + D_b R_{ac} \cdot \delta_{bc}.$$

We can also compute the commutator

$$\Delta D_a - D_a \Delta = R_{ab} D_b + D_b R_{ac} \delta_{bc} - R_{abcd} D_b \delta_{cd},$$

and get the important formula

$$(D_t - \Delta) D_a - D_a (D_t - \Delta) = R_{abcd} D_b \delta_{cd}$$

for commuting the evolution operator $D_t - \Delta$ with the covariant derivative D_a. For example,

$$(D_t - \Delta) D_a f - D_a (D_t - \Delta) f = 0$$

for functions f on M, and

$$(D_t - \Delta) D_a V_b - D_a (D_t - \Delta) V_b = 2 R_{acbd} D_c V_d$$

for covectors V_b, while

$$(D_t - \Delta) D_a V_{bc} - D_a (D_t - \Delta) V_{bc} = 2 R_{adbe} D_d V_{ec} + 2 R_{adce} D_d V_{be}$$

for two-tensors V_{bc}, and similar formulas hold for higher tensors (as we must expect from the product rule).

Finally, since $\nabla^a_b = y^i_b \partial/\partial y^i_a$ we see that

$$\frac{\partial}{\partial t}\nabla^a_b - \nabla^a_b \frac{\partial}{\partial t} = 0.$$

Then using

$$D_t = \frac{\partial}{\partial t} + R_{cd} g^{da} \nabla^c_a \quad \text{and} \quad \delta_{ab} = g_{af}\nabla^f_b - g_{bf}\nabla^f_a$$

we also compute

$$D_t\delta_{ab} - \delta_{ab}D_t = 0.$$

We summarize these results.

2.1. Theorem. *On the orthonormal frame bundle the metric is given by* $g_{ab} = I_{ab}$. *A basis for the tangent vectors to the orthonormal frame bundle is given by the horizontal spacelike tangent vectors* D_a, *the vertical spacelike rotation tangent vectors* δ_{bc}, *and the timelike vector* D_t. *The commutators are given by*

$$D_aD_b - D_bD_a = \tfrac{1}{2}R_{abcd}\delta_{cd},$$
$$\delta_{ab}D_c - D_c\delta_{ab} = g_{ac}D_b - g_{bc}D_a,$$
$$\delta_{ab}\delta_{cd} - \delta_{cd}\delta_{ab} = g_{ac}\delta_{bd} + g_{bd}\delta_{ac} - g_{ad}\delta_{bc} - g_{bc}\delta_{ad},$$
$$D_tD_a - D_aD_t = R_{ab}D_b + D_bR_{ac}\delta_{bc},$$
$$D_t\delta_{ab} - \delta_{ab}D_t = 0.$$

We also have

$$(D_t - \Delta)D_a - D_a(D_t - \Delta) = R_{abcd}D_b\delta_{cd}.$$

The action of δ_{ab} *on a covector* V_c *is given by*

$$\delta_{ab}V_c = g_{ac}V_b - g_{bc}V_a$$

and extends to other tensors by the product rule. Thus

$$D_aD_bV_c - D_bD_aV_c = R_{abcd}V_d$$

and

$$(D_t - \Delta)D_aV_b - D_a(D_t - \Delta)V_b = 2R_{acbd}D_cV_d,$$

and these formulas also extend to other tensors by the product rule. The curvature tensor itself evolves by

$$(D_t - \Delta)R_{abcd} = 2(B_{abcd} - B_{abdc} + B_{acbd} - B_{adbc}),$$

where $B_{abcd} = R_{aebf}R_{cedf}$.

3. The solitons

We call a solution to an evolution equation a soliton if it moves under a one-parameter subgroup of the invariance group of the equation. For example, the invariance group of the Ricci flow contains the full group of diffeomorphisms. A Ricci soliton is a solution to the Ricci flow which changes only by a diffeomorphism. Thus there will be a one-parameter family of diffeomorphisms such that the metric $g_{ab}(X, t)$ at times t is obtained from the metric $g_{ab}(X, 0)$ at time 0 by the diffeomorphism $X \to \varphi(X, t)$. If the subgroup is obtained by exponentiating the vector field V_a then we will have

$$D_a V_b + D_b V_a = 2R_{ab}$$

since the motion of g is by the Lie derivative of the vector field to be a soliton, and by the Ricci tensor to solve the Ricci flow. Conversely if we start with a metric which satisfies this equation, it must evolve under the Ricci flow by composing with the one-parameter family of diffeomorphisms obtained by exponentiating V_a. If we have a soliton where $V_a = D_a f$ is the gradient of a function f, we say it is a gradient soliton. In this case

$$D_a D_b f = R_{ab}$$

so we have a gradient soliton precisely when the Ricci tensor is the Hessian of a function. We are indebted to Robert Bryant and Thomas Ivey for an extensive study of the Ricci soliton and gradient soliton equations in terms of the calculus of exterior differential systems [1], [11].

These are steady solitons, which exist for $-\infty < t < \infty$. There are also homothetically expanding solitons for $0 < t < \infty$, and shrinking solitons for $-\infty < t < 0$, corresponding to the fact that the Ricci flow equation is also invariant under a one-parameter group of homotheties, where time dilates like space squared. For a homothetically expanding soliton we have

$$D_a V_b + D_b V_a = 2R_{ab} + \frac{1}{t} g_{ab}$$

and the opposite sign for a shrinking one. The expanding gradient solitons are closely related to the Harnack inequality, because it becomes an equality in this case. Indeed this is how we found the correct Harnack expression. Moreover it is a great aid in doing the calculations to check at each step that we get equality on the expanding gradient solitons.

Here is how we derive the Harnack expression Z. On an expanding gradient soliton

$$D_a V_b = R_{ab} + \frac{1}{2t} g_{ab}$$

since $V_a = D_a f$ implies $D_a V_b = D_b V_a$. Differentiating and commuting give the first order relations

$$D_a R_{bc} - D_b R_{ac} = R_{abcd} V_d,$$

and differentiating again gives

$$D_aD_bR_{cd} - D_aD_cR_{bd} = D_aR_{bcde}V_e + R_{ae}R_{bcde} + \frac{1}{2t}R_{bcda}.$$

We take the trace of this on a and b to conclude that

$$M_{ab} + P_{cab}V_c = 0,$$

where

$$M_{ab} = \Delta R_{ab} - \frac{1}{2}D_aD_bR + 2R_{acbd}R_{cd} - R_{ac}R_{bc} + \frac{1}{2t}R_{ab}$$

and

$$P_{abc} = D_aR_{bc} - D_bR_{ac}$$

as before. The first relation was then

$$P_{cba}V_c + R_{acbd}V_cV_d = 0,$$

and in order to get a good expression we add the two equations to make

$$M_{ab} + (P_{cab} + P_{cba})V_c + R_{acbd}V_cV_d = 0.$$

We apply this to an arbitrary vector W_a and get

$$M_{ab}W_aW_b + (P_{cab} + P_{cba})W_aW_bV_c + R_{acbd}W_aV_cW_bV_d = 0.$$

If we write

$$U_{ab} = \tfrac{1}{2}(V_aW_b - V_bW_a)$$

for the wedge product of V and W, the above can be rearranged as

$$Z = M_{ab}W_aW_b + 2P_{abc}U_{ab}W_c + R_{abcd}U_{ab}U_{cd} = 0,$$

which shows that the Harnack inequality becomes an equality on an expanding gradient soliton.

Since there are other expressions which vanish, one may ask how we come to select this one. One important criterion is that if $Z \geq 0$ for all choices of W and U then when $Z = 0$ on the soliton we must also have $\partial Z/\partial W = 0$ and $\partial Z/\partial U = 0$. This dictates that we need to take the trace of the second derivative expression, since otherwise we cannot mix it with the first derivative expression, and it also shows we must take an equal amount of each.

The author has written down a steady gradient soliton in dimension two (see [6]) given by

$$ds^2 = \frac{dx^2 + dy^2}{1 + x^2 + y^2},$$

which has positive curvature and one end of finite circumference like a cylinder. In dimension three Robert Bryant [2] has proved the existence of a rotationally symmetric complete steady gradient soliton by proving existence of a solution to an ODE which is hard to solve in closed form. It has positive curvature and opens more like a paraboloid. There seems little doubt that the same techniques will prove existence of both steady and expanding rotationally symmetric complete solitons of positive curvature operator in all dimensions. The proof of the Harnack inequality by no means depends on such an existence result. We use it only as inspiration.

As an example of this, on the soliton we chose

$$U_{ab} = \frac{1}{2}(V_a W_b - V_b W_a)$$

and we have

$$D_a V_b = R_{ab} + \frac{1}{2t} g_{ab}.$$

If we take the arbitrary section W_a such that $D_a W_b = 0$ at a point, then we get

$$D_a U_{bc} = \frac{1}{2}(R_{ab} W_c - R_{ac} W_b) + \frac{1}{4t}(g_{ab} W_c - g_{ac} W_b),$$

which is a choice we will make in the proof of the Harnack inequality. Now you will know where it comes from.

4. The computation

We assume we have a solution to the Ricci flow and let

$$P_{abc} = D_a R_{bc} - D_b R_{ac}$$

and

$$M_{ab} = \Delta R_{ab} - \frac{1}{2} D_a D_b R + 2R_{acbd} R_{cd} - R_{ac} R_{bc} + \frac{1}{2t} R_{ab},$$

and form the quadratic

$$Z = M_{ab} W_a W_b + 2P_{abc} U_{ab} W_c + R_{abcd} U_{ab} U_{cd}$$

where W_a is a one-form and U_{ab} is a two-form, depending on space and time.

4.1. Theorem. *At a point where*

$$(D_t - \Delta) W_a = \frac{1}{t} W_a, \; (D_t - \Delta) U_{ab} = 0,$$

$$D_a W_b = 0 \quad \textit{and} \quad D_a U_{bc} = \frac{1}{2}(R_{ab} W_c - R_{ac} W_b) + \frac{1}{4t}(g_{ab} W_c - g_{ac} W_b),$$

we have

$$\begin{aligned}(D_t - \Delta)Z &= 2R_{acbd}M_{cd}W_aW_b - 2P_{acd}P_{bdc}W_aW_b \\ &\quad + 8R_{adce}P_{dbe}U_{ab}W_c + 4R_{aecf}R_{bedf}U_{ab}U_{cd} \\ &\quad + [P_{abc}W_c + R_{abcd}U_{cd}][P_{abe}W_e + R_{abef}U_{ef}].\end{aligned}$$

Proof. First we must compute the evolution of the coefficients M_{ab}, P_{abc}, and R_{abcd}.

4.2. Lemma.

$$(D_t - \Delta)R_{abcd} = 2(B_{abcd} - B_{abdc} + B_{acbd} - B_{adbc}),$$

where B_{abcd} *is the quadratic*

$$B_{abcd} = R_{aebf}R_{cedf}.$$

Proof. This is the standard formula.

4.3. Lemma.

$$(D_t - \Delta)P_{abc} = 2R_{adbe}P_{dec} + 2R_{adce}P_{dbe} + 2R_{bdce}P_{ade} - 2R_{de}D_dR_{abce}.$$

Proof. The evolution of the Ricci tensor is given by

$$(D_t - \Delta)R_{bc} = 2R_{bdce}R_{de}$$

and from the commutator formula

$$(D_t - \Delta)D_aR_{bc} = D_a(D_t - \Delta)R_{bc} + 2R_{adbe}D_dR_{ec} + 2R_{adce}D_dR_{be}.$$

Evaluating the first term on the right gives

$$\begin{aligned}(D_t - \Delta)D_aR_{bc} &= 2R_{adbe}D_dR_{ec} + 2R_{adce}D_dR_{eb} \\ &\quad + 2R_{bdce}D_aR_{de} + 2R_{de}D_aR_{bdce}.\end{aligned}$$

Now we use

$$P_{abc} = D_aR_{bc} - D_bR_{ac}$$

and the second Bianchi identity

$$D_aR_{bdce} + D_bR_{dace} + D_dR_{abce} = 0$$

to complete the lemma.

4.4. Lemma.

$$\begin{aligned}(D_t - \Delta)M_{ab} &= 2R_{acbd}M_{cd} + 2R_{cd}[D_cP_{dab} + D_cP_{dba} \\ &\quad + 2P_{acd}P_{bcd} - 4P_{acd}P_{bdc} + 2R_{cd}R_{ce}R_{adbe} - \frac{1}{2t^2}R_{ab}.\end{aligned}$$

Proof. It is easiest to start from

$$M_{ab} = D_cP_{cab} + R_{acbd}R_{cd} + \frac{1}{2t}R_{ab},$$

which can be easily checked. The commutation formula gives

$$(D_t - \Delta)D_c P_{cab} = D_c(D_t - \Delta)P_{cab} + 2R_{de}D_d P_{eab} + 2R_{cdae}D_d P_{ceb} + 2R_{cdbe}D_d P_{cae}.$$

We now use the identities

$$P_{abc} + P_{bca} + P_{cab} = 0, \quad D_c R_{cdab} = P_{abd},$$
$$D_c D_d R_{cabe} = D_d P_{bea} + R_{df} R_{fabe} + B_{daeb} + B_{deab} - B_{dabe} - B_{dbae},$$

and

$$R_{cdae}[D_c R_{deb} + D_d R_{ceb}] = 0$$

which comes from the symmetry and antisymmetry in c and d to conclude that

$$(D_t - \Delta)D_c P_{cab} = 2R_{acbd}D_e P_{ecd} + 2R_{cd}(D_c P_{dab} + D_c P_{dba}) + 2P_{acd}P_{bcd} - 4P_{acd}P_{bdc} + 2R_{cd}R_{ce}R_{adbe} + 2D_e R_{acbd}D_e R_{cd} - 2R_{cd}(B_{abcd} + B_{acbd} - B_{acdb} - B_{adcb})$$

using the fact that the tensor $B_{abcd} = R_{aebf}R_{cedf}$ has the symmetries

$$B_{abcd} = B_{cdab} = B_{badc} = B_{dcba}.$$

Also

$$(D_t - \Delta)[R_{acbd}R_{cd}] = 2R_{acbd}[R_{cedf}R_{ef}] - 2D_e R_{acbd}D_e R_{cd} + 2R_{cd}(B_{abcd} + B_{acbd} - B_{acdb} - B_{adcb})$$

and

$$(D_t - \Delta)\left[\frac{1}{2t}R_{ab}\right] = 2R_{acbd}\left[\frac{1}{2t}R_{cd}\right] - \frac{1}{2t^2}R_{ab}.$$

Adding these reuslts completes the lemma, once we use

$$M_{cd} = D_e P_{ecd} + R_{cedf}R_{ef} + \frac{1}{2t}R_{cd}$$

and observe the obvious cancellations.

Now when

$$Z = M_{ab}W_a W_b + 2P_{abc}U_{ab}W_c + R_{abcd}U_{ab}U_{cd}$$

and we are at a point where

$$D_a W_b = 0 \quad \text{and} \quad (D_t - \Delta)U_{ab} = 0,$$

we compute

$$\begin{aligned}(D_t - \Delta)Z = {} & [(D_t - \Delta)M_{ab}]W_a W_b \\ & + 2[(D_t - \Delta)P_{abc}]U_{ab}W_c + [(D_t - \Delta)R_{abcd}]U_{ab}U_{cd} \\ & + 2[M_{ac}W_a + P_{abc}U_{ab}](D_t - \Delta)W_c \\ & - 4D_e P_{abc} D_e U_{ab} W_c - 4D_e R_{abcd} D_e U_{ab} U_{cd} \\ & - 2R_{abcd} D_e U_{ab} D_e U_{cd}.\end{aligned}$$

If we substitute the computed values for $(D_t - \Delta)M_{ab}$, $(D_t - \Delta]P_{abc}$, and $(D_t - \Delta)R_{abcd}$ from the lemmas, and the chosen values for $(D_t - \Delta)W_a$ and $D_a U_{bc}$, the result in the theorem follows from the obvious cancellations.

We now give the idea of the proof, in a form which is not quite rigorous, but shows why it works. We shall have to mess the formula up a bit to sneak an $\varepsilon > 0$ in, as is usual in making maximum principle arguments work, so it is best to see what is really happening before it gets too messy.

If the manifold is compact and the curvature operator is strictly positive, then the quadratic form Z will be strictly positive for small time. If it ever becomes negative, there will be a first time t° when it is zero, and this will happen at some point X° and in the direction of some eigenvector W_a° and U_{ab}°. We can extend these in space-time to sections W_a and U_{ab} with $W_a = W_a^\circ$ and $U_{ab} = U_{ab}^\circ$, and we can do so however we please and still have $Z \geq 0$ up to time t°. In particular we can make the first derivatives in space and time anything we like, so we can extend first in space to make

$$D_a W_b = 0$$

and

$$D_a U_{bc} = \frac{1}{2}(R_{ab}W_c - R_{ac}W_b) + \frac{1}{4t}(g_{ab}W_c - g_{ac}W_b),$$

and then, knowing ΔW_a and ΔU_{ab}, we can extend in time to make

$$(D_t - \Delta)W_a = \frac{1}{t}W_a \quad \text{and} \quad (D_t - \Delta)U_{ab} = 0.$$

Actually, if we compute carefully, it turns out that at a null eigenvector of $\widehat{Z}$ it does not matter what $(D_t - \Delta)W_a$ or $(D_t - \Delta)U_{ab}$ is anyway, but it is easier to prescribe than compute. In the evolution of Z the quadratic term

$$[P_{abc}W_c + R_{abcd}U_{cd}][P_{abe}W_e + R_{abef}U_{ef}]$$

is clearly nonnegative. Actually at a null eigenvector,

$$P_{abc}W_c + R_{abcd}U_{cd} = 0$$

anyway. The fact that $(D_t - \Delta)Z \geq 0$ at (X°, t°) then follows from the lemma below. Since $Z \geq 0$ everywhere at time t° we get $\Delta Z \geq 0$, and so $D_t Z \geq 0$. Thus wherever Z first becomes zero it is nonincreasing. This sort of shows Z stays ≥ 0, but we need to sneak an $\varepsilon > 0$ in somewhere, as we mentioned before. The crucial step is the following.

4.5. Lemma. *If the quadratic form in W and U*

$$Z = M_{ab}W_aW_b + 2P_{abc}U_{ab}W_c + R_{abcd}U_{ab}U_{cd}$$

is weakly positive, so is the quadratic form

$$\begin{aligned} Q = {} & 2R_{acbd}M_{cd}W_aW_b - 2P_{acd}P_{bdc}W_aW_b \\ & + 8R_{adce}P_{dbe}U_{ab}W_c + 4R_{aecf}R_{bedf}U_{ab}U_{cd}. \end{aligned}$$

Proof. A weakly positive quadratic form can always be written as a sum of squares of linear forms. This is equivalent to diagonalizing a symmetric matrix and writing each nonnegative eigenvalue as a square. Write

$$Z = \sum_N (X_a^N W_a + Y_{ab}^N U_{ab})^2.$$

This makes

$$M_{ac} = \sum_N X_a^N X_c^N, \qquad P_{abc} = \sum_N Y_{ab}^N X_c^N,$$
$$R_{abcd} = \sum_N Y_{ab}^N Y_{cd}^N,$$

where X_a^N is a one-form, and Y_{ab}^N is a two-form for each N. It is then easy to compute

$$Q = \sum_{M,N} (Y_{ac}^M X_c^N W_a - Y_{ac}^N X_c^M W_a - 2Y_{ac}^M Y_{bc}^N U_{ab})^2,$$

which shows Q is also a sum of squares of linear forms and hence is a weakly positive quadratic form.

5. The argument

We assume that we have a complete solution to the Ricci flow with bounded curvature and nonnegative curvature operator. We may easily assume we are working on a closed time interval $0 \leq t \leq T$, for if we only start with a solution for $0 < t < T$ we can pass to the interval $\varepsilon \leq t \leq T_\varepsilon$ and let $\varepsilon \to 0$. By the work of W. X. Shi [13] we can then also assume bounds on the covariant derivatives of the curvature. Note that the final conclusion is independent of these bounds. In what follows we will let C denote various constants which depend only on the dimension, the time interval T, and bounds on the curvature $|Rm|$ and its derivatives $|DR_m|$ and $|D^2 Rm|$. The constants will vary from line to line, and to be precise could be indexed by the order of occurrence.

The idea of the proof is to perturb the expression Z slightly to $\widehat{Z}$ so as to make $\widehat{Z}$ very positive if $t \to 0$, or if the point $X \to \infty$ in case the manifold is not compact. Also our perturbation must add a little positive push, so that wherever $\widehat{Z}$ first acquires a zero it is strictly increasing. It then follows that $\widehat{Z}$ never could make it to zero after all. Since we can take $\widehat{Z}$ as close to Z as we like on compact sets in space-time avoiding $t = 0$, we get $Z \geq 0$ as desired.

We take $\widehat{Z}$ in the form

$$\widehat{Z} = \widehat{M}_{ab} W_a W_b + 2P_{abc} U_{ab} W_c + \widehat{R}_{abcd} U_{ab} U_{cd},$$

where we take

$$\widehat{M}_{ab} = M_{ab} + \frac{1}{t}\varphi g_{ab},$$

$$\widehat{R}_{abcd} = R_{abcd} + \frac{1}{2}\psi(g_{ac}g_{bd} - g_{ad}g_{bc}),$$

for suitably chosen functions φ and ψ. In fact we will later choose

$$\varphi = \frac{\varepsilon}{\sqrt{t}} e^{At} f(X) \quad \text{and} \quad \psi = \delta e^{Bt}$$

with A and B large and ε and δ small, and where $f(X)$ is a function of position only such that $f(X) \to \infty$ as $X \to \infty$ but the derivatives of f are bounded. In case the manifold is compact we just take $f = 1$.

First we review the construction of f, as in Greene and Wu [5] and W. X. Shi [13].

5.1. Lemma. *There exists a smooth function f such that $f \geq 1$ everywhere and $f(X) \to \infty$ as $X \to \infty$ but all the covariant derivatives are bounded, so that $|Df| \leq C$ and $|D^2 f| \leq C$ for a constant C.*

Proof. Let $d(X)$ be the distance from some fixed point at time zero, let $p(V)$ be a smooth function on Euclidean space which is rotationally symmetric with support in a small ball, and let

$$f(X) = \int_{T_X} p(V) d(\exp_X V)\, dV$$

be the integral over the tangent space T_X at X. If the size of the support of $p(V)$ is small compared to the maximum curvature, it is well known that this smoothing gives a function with $f(X) \to \infty$ as $X \to \infty$, while its derivatives are bounded (See Greene and Wu [5]). We can also easily bound f from below, and by dilating $p(V)$ by a constant we take $f \geq 1$. Finally (as W. X. Shi does in [13]) we can bound covariant derivatives of f at times $t > 0$ using the standard estimate on the change in the metric and the change in the connection, all of which is easily controlled.

5.2. Lemma. *Given any constant C, any $\eta > 0$, and any compact set K in space-time we can find functions $\psi = \psi(t)$ depending on time alone and $\varphi = \varphi(X, t)$ depending on both space and time such that*

(1) $\psi \leq \eta$ everywhere, and $\psi \geq \delta$ for some $\delta > 0$,

(2) $\varphi \leq \eta$ on the set K, and $\varphi \geq \varepsilon$ for some $\varepsilon > 0$, while $\varphi(X, t) \to \infty$ if $X \to \infty$ in the sense that the sets $\varphi \leq M$ are all compact in space-time for $0 \leq t \leq T$ (if the manifold is compact this condition is vacuous),

(3) $(D_t - \Delta)\varphi > C\varphi$,

(4) $D_t\psi > C\psi$,

(5) $\varphi \geq C\psi$.

Proof. We look first for φ in the form

$$\varphi(X, t) = \varepsilon e^{At} f(X)$$

with f as before. Since $\Delta f \leq C$ and $f \geq 1$ we get $\Delta\varphi \leq C\varphi$, and so to make (3) work we only need $D_t\varphi > C\varphi$ with a different C. But this works if we pick $A > C$. To make (2) work we need

$$\varepsilon \leq \eta e^{-AT} \max_K f(X),$$

which we can do.

Then we look for ψ in the form $\psi(t) = \delta e^{Bt}$ and find that (4) works when $B > C$. To make (1) work we take $\delta < \eta e^{-BT}$ and to make (5) work we take $\delta < \varepsilon e^{-BT}/C$ and use $e^{At} \geq 1$ and $f \geq 1$. This proves the lemma.

Now we study the evolution of $\widehat{Z}$. We observe the extra terms which are added, which are few because $D_a W_b = 0$ and $(D_t - \Delta)U_{ab} = 0$ at our point as before, and ψ depends only on t so $D_a\psi = 0$ and $\Delta\psi = 0$. The extra terms come (1) when $(D_t - \Delta)$ falls on φ/t, (2) when $(D_t - \Delta)$ falls on ψ, (3) when $(D_t - \Delta)$ falls on W_a and there is an extra φ/t in $\widehat{M}_{ab}$, and (4) when the Δ distributes as one derivative on each of U_{ab} and U_{cd} and there is an extra ψ in $\widehat{R}_{abcd}$. This gives us

$$\begin{aligned}(D_t - \Delta)\widehat{Z} = (D_t - \Delta)Z &+ \frac{1}{t}\left[(D_t - \Delta)\varphi - \frac{1}{t}\varphi\right] |W|^2 \\ &+ \frac{2}{t}\varphi W_a (D_t - \Delta)W_a + (D_t\psi)|U|^2 + \psi|D_a U_{bc}|^2 .\end{aligned}$$

If again we take

$$[D_t - \Delta]W_a = \frac{1}{t}W_a,$$
$$D_a U_{bc} = \frac{1}{2}(R_{ab}W_c - R_{ac}W_b) + \frac{1}{4t}(g_{ab}W_c - g_{ac}W_b)$$

at our point, substitute above and use $|Rm| \le C$ and $t \le C$, we get

$$(D_t - \Delta)\widehat{Z} \ge (D_t - \Delta)Z + \frac{1}{t}\left[(D_t - \Delta)\varphi + \frac{1}{t}\varphi - \frac{C}{t}\psi\right]|W|^2 + (D_t\psi)|U|^2.$$

Now starting from our previous computation of $(D_t - \Delta)Z$ we replace M_{ab} by $\widehat{M}_{ab}$ and R_{abcd} by $\widehat{R}_{abcd}$ and bound the resulting errors. This gives

$$\begin{aligned}(D_t - \Delta)Z \ge\ & 2\widehat{R}_{acbd}\widehat{M}_{cd}W_aW_b - 2P_{acd}P_{bdc}W_aW_b \\ & + 8\widehat{R}_{adce}P_{dbe}U_{ab}W_c + 4\widehat{R}_{aecf}\widehat{R}_{bedf}U_{ab}U_{cd} \\ & + [P_{abc}W_c + \widehat{R}_{abcd}U_{cd}][P_{abe}W_e + \widehat{R}_{abef}U_{ef}] \\ & - \frac{C}{t}(\varphi + \psi + \varphi\psi)|W|^2 - C\psi|U|\,|W| - C(\psi^2 + \psi)|U|^2.\end{aligned}$$

We can simplify some of these errors. First

$$\psi|U|\,|W| \le \psi|W|^2 + \psi|U|^2$$

gets rid of the cross-term. Then using $\psi \le \varphi$ (we even have $C\psi \le \varphi$) and $\psi \le 1$ (we even have $\psi \le \eta$ and since we want η small we can make sure $\eta \le 1$) the errors reduce to

$$\frac{C}{t}\varphi|W|^2 + C\psi|U|^2.$$

Combining this with the preceding calculation we get the following result.

Theorem. *Let φ and ψ be as in Lemma 2, and let W_a be a one-form and U_{ab} a two-form which at a given point satisfy*

$$(D_t - \Delta)W_a = \frac{1}{t}W_a, \quad (D_t - \Delta)U_{ab} = 0, \quad D_aW_b = 0$$

and

$$D_aU_{bc} = \frac{1}{2}(R_{ab}W_c - R_{ac}W_b) + \frac{1}{4t}(g_{ab}W_c - g_{ac}W_b).$$

Let

$$\widehat{M}_{ab} = M_{ab} + \frac{1}{t}\varphi g_{ab},$$
$$\widehat{R}_{abcd} = R_{abcd} + \frac{1}{2}\psi(g_{ac}g_{bd} - g_{ad}g_{bc})$$

and form the quadratic

$$\widehat{Z} = \widehat{M}_{ab} W_a W_b + 2P_{abc} U_{ab} W_c + \widehat{R}_{abcd} U_{ab} U_{cd}.$$

Then

$$\begin{aligned}(D_t - \Delta)\widehat{Z} \geq\ & 2\widehat{R}_{acbd} \widehat{M}_{cd} W_a W_b - 2P_{acd} P_{bdc} W_a W_b \\ & + 8\widehat{R}_{adce} P_{dbe} U_{ab} W_c + 4\widehat{R}_{aecf} \widehat{R}_{bedf} U_{ab} U_{ad} \\ & + [P_{abc} W_c + \widehat{R}_{abcd} U_{cd}][P_{abe} W_e + \widehat{R}_{abef} U_{ef}] \\ & + \frac{1}{t}\left[(D_t - \Delta)\varphi + \frac{1}{t}\varphi - \frac{C}{t}\psi - C\varphi\right] |W|^2 \\ & + [D_t \psi - C\psi] |U|^2.\end{aligned}$$

Now we finish the argument rigorously along the previous lines. Note that R_{abcd} is nonnegative, P_{abc} is bounded and M_{ab} is the sum of a bounded term plus $1/t$ times a nonnegative one. It follows that

$$Z \geq -C|W|^2 - C|W|\,|U|$$

and hence

$$\widehat{Z} \geq \left(\frac{1}{t}\varphi - C\right) |W|^2 - C|W|\,|U| + \psi |U|^2.$$

Now $\psi \geq \delta$ while φ/t is big when either t is small or the point X is outside a compact set. So we see that the quadratic form $\widehat{Z}$ is strictly positive outside of a compact set in space-time that avoids $t = 0$. We claim of course that $\widehat{Z}$ is in fact always strictly positive, or we would get a contradiction.

For suppose we look at the first time $t^\circ > 0$ where $\widehat{Z}$ has a zero eigenvector, occurring at some point X° in the direction of the one-form W_a° and the two-form U_{ab}°. Extend then to sections W_a and U_{ab} with $W_a = W_a^\circ$ and $U_{ab} = U_{ab}^\circ$ at X° in such a way that we have

$$(D_t - \Delta) W_a = \frac{1}{t} W_a, \quad (D_t - \Delta) U_{ab} = 0, \quad D_a W_b = 0$$

and

$$D_a U_{bc} = \frac{1}{2}(R_{ab} W_a - R_{ac} W_b) + \frac{1}{4t}(g_{ab} W_c - g_{ac} W_b)$$

at the point X°. Then arguing on $(D_t - \Delta)\widehat{Z}$ we see as before that the quadratic terms are nonnegative at a zero eigenvector. But now the estimates of Lemma 2 give $(D_t - \Delta)\widehat{Z} > 0$ (since either $W \neq 0$ or $U \neq 0$ for an eigenvector). But $\widehat{Z} \geq 0$ everywhere at the time t°, so $\Delta\widehat{Z} \geq 0$. This makes $D_t \widehat{Z} > 0$ at X° at the time t° when $\widehat{Z} = 0$. But then a short time before $\widehat{Z}$ must have been negative at the point X°, for some choice of W_a and U_{ab} coming from our extension. Since this is a contradiction, it follows that $\widehat{Z} > 0$ always and everywhere.

Now it remains to let $\eta \to 0$ in Lemma 2, and we get $Z \geq 0$ in the limit. This finishes the rigorous proof of the Harnack inequality.

References

[1] R. Bryant, *Local solvability of the gradient Ricci flow equation*, preprint.
[2] ——, *Existence of a gradient Ricci soliton in dimension three*, preprint.
[3] B. Chow, *The Harnack estimate for the Gauss curvature flow*, preprint.
[4] ——, *The Harnack estimate for the Yamabe flow*, preprint.
[5] R. Greene & H. Wu, C^∞ *convex functions and manifolds of positive curvature*, Acta Math. **13** (1976) 209–245.
[6] R. Hamilton, *The Ricci flow on surfaces.*
[7] ——, *Three-manifolds with positive Ricci curvature*, J. Differential Geometry **17** (1982) 255–306.
[8] ——, *Four-manifolds with positive curvature operator*, J. Differential Geometry **24** (1986) 153–179.
[9] ——, *A matrix Harnack estimate for the heat equation*, in preparation.
[10] ——, *The Harnack estimate for the mean curvature flow*, in preparation.
[11] T. Ivey, *Local existence of Ricci solitons in dimension three*, preprint.
[12] P. Li & S.-T. Yau, *Inequalities for the parabolic Schrodinger operator*, Acta Math.
[13] W.-X. Shi, *Ricci deformation of the metric on complete noncompact Riemannian manifolds*, J. Differential Geometry **30** (1989) 303–394.

UNIVERSITY OF CALIFORNIA, LA JOLLA