

Ricci Flow: The Foundations via Optimal Transportation

Peter Topping

**** DRAFT VERSION ****

Errors please to p.m.topping@warwick.ac.uk

June 26, 2009 (@4mpm)

<http://www.warwick.ac.uk/~maseq>

Contents

Preface	2
Prerequisites	3
1 Introduction to Ricci flow and diffusions	4
1.1 The equation	4
1.2 Rapid overview of Ricci flow theory, and motivations	5
1.3 Ricci solitons	7
1.4 Diffusion and the heat equation on manifolds and flows	8
2 Optimal Transportation	10
2.1 Wasserstein distance	10
2.2 Contractivity for diffusions on evolving manifolds	10
3 The Canonical Solitons	13
3.1 Definition of the Canonical Shrinking Soliton	13
3.2 Perelman's \mathcal{L} -length	15
3.3 \mathcal{L} -Optimal Transportation	16
3.4 Contractivity of \mathcal{L} -Wasserstein distance	17
3.5 Applications of \mathcal{L} -Wasserstein contractivity	19
3.5.1 Recovering Perelman's \mathcal{W} -entropy	19
3.5.2 No local collapsing	21
4 Harnack inequalities	23
4.1 The Canonical Expanding Soliton	23
4.2 Harnack estimates generated by the Canonical Expanding Soliton	25
4.3 Applications to understanding blown-up Ricci flows	27
Bibliography	29

Preface

Since the creation of Ricci flow by Hamilton in 1982, a rich theory has been developed in order to understand the behaviour of the flow, and to analyse the singularities which may occur, and these developments have had profound applications, most famously to the Poincaré conjecture. At the heart of the theory lie a large number of *a priori* estimates and geometric constructions which include most notably the Harnack estimates of Hamilton, the \mathcal{L} -length of Perelman (in the spirit of Li-Yau), the logarithmic Sobolev inequality arising from Perelman's \mathcal{W} -entropy, and the reduced volume of Perelman, amongst others.

The objective of these lectures is to explain this theory from the point of view of Optimal Transportation. As I explain in Chapter 2, Ricci flow and Optimal Transportation combine rather well, and we will see fundamental but elementary aspects of this when we see in Theorem 2.2.1 how diffusions contract under reverse-time Ricci flow. However, the key to the whole theory is to realise to which object one should apply this result: not the original Ricci flow, but a new Ricci flow derived from the original one, on a base manifold of one higher dimension, that we call the Canonical Soliton. In this way essentially the entire foundational theory of Ricci flow mentioned above drops out naturally.

Throughout the lectures I emphasise the intuition; the objective is to demonstrate how one can discover the theory rather than treat it as a black box which just happens to work.

These lectures were given at the summer school on 'Optimal Transportation: Theory and Applications' at Grenoble in June 2009, and I thank Yann Ollivier, Herve Pajot and Cedric Villani for their invitation. I would also like to thank Esther Cabezas-Rivas and Dan Jane for many useful comments.

Peter Topping, Warwick, June 2009.

Prerequisites

It would be helpful, but not essential, to have some basic idea of what Ricci flow is before the talks. The relevant reference is [22], and I will refer back to that monograph during these lectures.

I will assume the basic notions of manifolds and Riemannian geometry, including Riemannian metrics, Riemannian distance, geodesics and geodesic balls. To understand the details of the lectures, one would need also to understand a little about curvature, including the curvature tensor Rm , Ricci tensor Ric , scalar curvature R , sectional curvature and hopefully also the curvature operator \mathcal{R} . Occasionally we will pull back tensors by diffeomorphisms and consider Lie derivatives.

We will need the most basic Bochner formula (which follows from a short computation in Riemannian geometry).

I will assume you know how to integrate vector fields on manifolds (i.e. follow flow-lines) and understand the corresponding notion of ‘complete vector field’.

From PDE theory we will assume very little, except basic facts and intuition concerning the heat equation. In particular, we need the parabolic maximum principle.

I will try to recap the theory of Optimal Transportation that we require such as the definition of Wasserstein distance, or develop it during the lectures.

Chapter 1

Introduction to Ricci flow and diffusions

1.1 The equation

The Ricci flow is a way of deforming a Riemannian manifold under a nonlinear evolution equation in order to process or improve it. To be a Ricci flow, we ask a smooth one-parameter family $g(t)$ of Riemannian metrics on a manifold \mathcal{M} to satisfy the nonlinear partial differential equation

$$\frac{\partial g}{\partial t} = -2 \operatorname{Ric}(g) \tag{1.1.1}$$

where $\operatorname{Ric}(g)$ is the Ricci curvature of the metric $g(t)$, a section of $\operatorname{Sym}^2 T^* \mathcal{M}$.

There are a number of points to recall about Ricci curvature which may help here and later in these notes. First, if \mathcal{M} were two-dimensional, then $\operatorname{Ric}(g)$ is just the Gaussian curvature times the metric g . More generally, one should view $-\operatorname{Ric}(g)$ as some sort of Laplacian of g , as discussed in [22, §1.1], so the Ricci flow is like a heat equation for g (albeit nonlinear, and liable to develop singularities). Generally in the subject of these notes, it is useful to view $\operatorname{Ric}(g)$ as a quantity which controls *volume growth*. In particular, it governs the change in a volume element along a geodesic.

The interpretation of Ricci flow as some sort of diffusion equation for a Riemannian metric is brought out when one computes the evolution equation governing the curvatures of a Ricci flow. For example, the scalar curvature R (which is the trace of the Ricci tensor) satisfies

$$\frac{\partial R}{\partial t} = \Delta_{g(t)} R + 2|\operatorname{Ric}|^2, \tag{1.1.2}$$

while the full curvature tensor Rm and the Ricci tensor Ric satisfy related equations [22, §2.5].

Note that $|\text{Ric}|$ is the norm of Ric using the norm induced by g (i.e. $|\text{Ric}|^2 = R_{ij}R^{ij}$). For further explanation see [22, Chapter 2].

One useful consequence of (1.1.2), thanks to the maximum principle (whenever it applies) is that positive (or weakly positive) scalar curvature is preserved. That is, if $R(g(0)) \geq 0$ then $R(g(t)) \geq 0$ for $t > 0$. This is just one instance of a recurring theme; many positive curvature conditions are preserved. For example positive curvature operator $\mathcal{R}(g) \geq 0$. (See [22] for the definition of curvature operator.) Other recent breakthroughs [3], [15] have involved showing that the notion of *positive isotropic curvature* is preserved, and we will briefly return to this notion in Section 4.2.

Before proceeding, we note some simple examples of Ricci flows relevant to the next section. First, if we deform a round sphere under the Ricci flow then it will shrink (owing to its positive curvature) and disappear to nothing within a finite time. (See Section 1.3 for more details.) Also, the Cartesian product of two Ricci flows is another Ricci flow, and in particular, a cylinder $S^n \times \mathbb{R}$ can be evolved under Ricci flow simply by shrinking the S^n component.

1.2 Rapid overview of Ricci flow theory, and motivations

Ricci flow has become a powerful tool for understanding the topology of manifolds, particularly 3-dimensional manifolds because it is in that dimension that the singularities of Ricci flow can be analysed to a sufficiently fine degree. A typical singularity occurs because a part of the manifold ‘pinches a neck’.¹

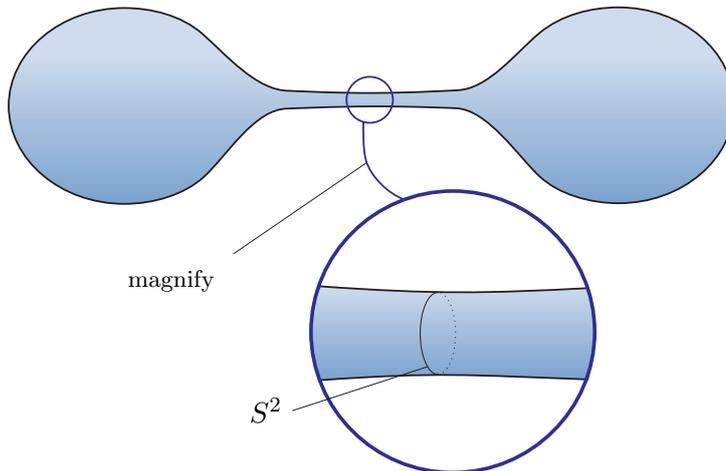


Figure 1.1: Blowing up.

At a singularity, the curvature blows up. More precisely, the maximum sectional curvature of $g(t)$ becomes unbounded as t increases to some singular time T . One would like to look at this flow through a microscope (i.e. rescale parabolically, thus reducing the curvature) and try to

¹Thanks to Neil Course for drawing Figure 1.1.

extract a limiting flow. For example in the case of a neck pinch, the limit flow should be a cylinder $S^2 \times \mathbb{R}$ with the S^2 part shrinking as time increases (see Figure 1.1).

The technical details of this process need not concern us here, except that it turns out to be possible to find this limit only if the flow does not degenerate in a particular way. One needs to rule out ‘local collapsing’ where balls of a given radius (and controlled curvature) have much smaller volume than they would have in Euclidean space, or more precisely, show that:

There exists $\kappa > 0$ such that for any $t < T$ and $r \in (0, 1]$, if the sectional curvatures of $g(t)$ are bounded in norm by r^{-2} in some geodesic ball $B(x, r)$ then the volume of $B(x, r)$ is bounded below by κr^n .

In these lectures we will develop the elegant theory which guarantees this (and many other things besides) due to Perelman [17], via the theory of optimal transportation. This essentially requires us to find a certain log-Sobolev inequality on our Ricci flow, and plug in a carefully chosen function.

Now suppose we have extracted a limit Ricci flow (which we will also call $g(t)$), for example the shrinking cylinder mentioned above. This turns out to be defined for all $t < 0$, and when \mathcal{M} is 3-dimensional, a result of Hamilton and Ivey [11] tells us that it has positive curvature. It is a crucial step to analyse the geometry of limits such as $g(t)$, and in these lectures we will see a number of important techniques and tools required in this analysis springing out of the optimal transportation viewpoint, including Perelman’s \mathcal{L} -length and the theory thereof. A further big step is to be able to relate the (scalar) curvatures at different points in space and time on such a flow and in this direction we will be led very naturally to the following *Harnack* inequality which first arises in the work of Hamilton [10]:

For $x_1, x_2 \in \mathcal{M}$ and times $t_1 < t_2 < 0$ on a 3-dimensional limit Ricci flow, we have

$$\frac{R(x_2, t_2)}{R(x_1, t_1)} \geq \exp\left(-\frac{d_{g(t_1)}^2(x_1, x_2)}{2(t_2 - t_1)}\right).$$

One further issue in the analysis of these limit flows is to understand what they look like at very large scales, very far back in time. The theory developed in these lectures can also be used to show that they look like solitons. (The original reference for this is [17, §11.2].)

1.3 Ricci solitons

The most elementary examples of Ricci flows are induced by Einstein metrics – i.e. metrics g_0 on \mathcal{M} which satisfy $\text{Ric}(g_0) = \lambda g_0$, for $\lambda \in \mathbb{R}$. An obvious Ricci flow starting at g_0 is given by²

$$g(t) = (1 - 2\lambda t)g_0. \quad (1.3.1)$$

For example, a round sphere would shrink homothetically to nothing in a finite time, and a hyperbolic manifold would expand homothetically indefinitely.

A more general notion of self-similar solution is that of a *Ricci soliton flow*, for which the manifold $(\mathcal{M}, g(t))$ is again isometric to a homothetic scaling of $(\mathcal{M}, g(0))$, although this time not necessarily via the identity map.

Definition 1.3.1. We call the combination (\mathcal{M}, g, f) of a Riemannian manifold (\mathcal{M}, g) together with a smooth function $f : \mathcal{M} \rightarrow \mathbb{R}$ a (*gradient*) *Ricci soliton* if it satisfies the equation

$$\text{Ric}(g) + \text{Hess}_g(f) = \lambda g. \quad (1.3.2)$$

Moreover, we call (\mathcal{M}, g, f) either *steady*, *expanding* or *shrinking* depending on whether $\lambda = 0$, $\lambda < 0$ or $\lambda > 0$ respectively.

The significance of such a combination of manifold and function is that it induces a Ricci flow generalising the Einstein case. It turns out (clarified only recently [25]) that if the manifold is complete, then simply by virtue of f satisfying the equation (1.3.2), the vector field ∇f is *complete*. (That is, one can follow the flow-lines of this vector field without falling off the manifold.) In particular, the vector field generates a family of diffeomorphisms $\psi_t : \mathcal{M} \rightarrow \mathcal{M}$ with ψ_0 the identity, according to

$$\frac{\partial}{\partial t}(\psi_t(y)) = \frac{\nabla f(\psi_t(y))}{1 - 2\lambda t} \quad (1.3.3)$$

for $t < \frac{1}{2\lambda}$ if $\lambda > 0$ and $t > \frac{1}{2\lambda}$ if $\lambda < 0$. We may then define

$$G(t) := (1 - 2\lambda t)\psi_t^*(g), \quad (1.3.4)$$

and compute (see also [22, §1.2])

$$\begin{aligned} \frac{\partial G(t)}{\partial t} &= -2\lambda\psi_t^*(g) + (1 - 2\lambda t)\psi_t^*\left(\mathcal{L}_{\frac{\nabla f(\psi_t(y))}{1-2\lambda t}}g\right) \\ &= \psi_t^*(-2\lambda g + \mathcal{L}_{\nabla f(\psi_t(y))}g) \\ &= \psi_t^*(-2\lambda g + 2\text{Hess}_g(f)) \\ &= \psi_t^*(-2\text{Ric}(g)) \\ &= -2\text{Ric}(\psi_t^*g) \\ &= -2\text{Ric}(G(t)) \end{aligned}$$

to see that $G(t)$ is in fact a Ricci flow generalising (1.3.1); this motivates:

²Recall that scaling a metric does not change the Ricci tensor: $\text{Ric}(\lambda g) = \text{Ric}(g)$ for any $\lambda > 0$.

Definition 1.3.2. We call a flow $G(t)$ constructed in this way a *Ricci soliton flow*.

Ricci solitons occur throughout the theory of Ricci flow. Our Canonical Solitons alluded to earlier will be exotic examples.

1.4 Diffusion and the heat equation on manifolds and flows

Consider a solution $u : \mathcal{M} \times [0, T] \rightarrow [0, \infty)$ to the heat equation

$$\frac{\partial u}{\partial t} = \Delta u,$$

on a closed Riemannian manifold (\mathcal{M}, g) , with $\int_{\mathcal{M}} u(\cdot, 0) d\mu_g = 1$, where μ_g is the Riemannian volume measure. Since

$$\frac{d}{dt} \int_{\mathcal{M}} u(\cdot, t) d\mu_g = \int_{\mathcal{M}} \Delta u d\mu_g = 0,$$

the integral of u over \mathcal{M} remains one for each $t \in [0, T]$, and in particular, u can be viewed as a probability density of a measure $\nu(t)$ defined by $d\nu(t) := u(\cdot, t) d\mu_g$. This one-parameter family of measures $\nu(t)$ represents the probability distribution of a particle moving on the manifold under Brownian motion, with initial probability distribution $\nu(0)$, and will be an example of what we will call a ‘diffusion’ in these lectures.

In practice, we need to generalise these familiar notions to the case of diffusion on an *evolving* manifold.

Definition 1.4.1. Suppose $g(\tau)$ is a family of Riemannian metrics on \mathcal{M} , for $\tau \in [0, T]$. We call a flow of measures $\nu(\tau)$ a **diffusion** (representing the probability distribution of a Brownian particle as above) if $d\nu(\tau) = u(\cdot, \tau) d\mu_{g(\tau)}$ with u satisfying the *conjugate* heat equation

$$\frac{\partial u}{\partial \tau} = \Delta_{g(\tau)} u - \left(\frac{1}{2} \text{tr} \frac{\partial g}{\partial \tau} \right) u. \quad (1.4.1)$$

The single most important situation is that where $g(\tau)$ solves the equation $\frac{\partial g}{\partial \tau} = 2\text{Ric}(g(\tau))$, which makes $g(\cdot)$ a Ricci flow with respect to a ‘reverse’ time coordinate τ (that is, $\tau = C - t$ for some constant C). In that case, u would satisfy

$$\frac{\partial u}{\partial \tau} = \Delta_{g(\tau)} u - Ru, \quad (1.4.2)$$

where $R = \text{tr Ric}$ is the scalar curvature as before.

Note that the extra term in (1.4.1) compensates for the fact that the Riemannian volume measure $\mu_{g(\tau)}$ is no longer constant in time – in fact $\frac{\partial}{\partial \tau} d\mu_{g(\tau)} = \frac{1}{2} \text{tr} \frac{\partial g}{\partial \tau} d\mu_{g(\tau)}$ – to keep $\nu(\tau)$ a

probability measure:

$$\frac{d}{d\tau} \int_{\mathcal{M}} \nu(\tau) = \frac{d}{d\tau} \int_{\mathcal{M}} u(\cdot, \tau) d\mu_{g(\tau)} = \int_{\mathcal{M}} \left(\frac{\partial u}{\partial \tau} d\mu_{g(\tau)} + u \frac{\partial}{\partial \tau} d\mu_{g(\tau)} \right) = \int_{\mathcal{M}} \Delta_{g(\tau)} u d\mu_{g(\tau)} = 0.$$

We will also need the following more general statement which justifies the terminology *conjugate* heat equation above

Proposition 1.4.2. *If $g(\tau)$ is a flow of Riemannian metrics for $\tau \in [a, b]$, and $\nu(\tau)$ is a diffusion (as defined above) then given any $f : \mathcal{M} \times [a, b] \rightarrow \mathbb{R}$ solving $-\frac{\partial f}{\partial \tau} = \Delta_{g(\tau)} f$, we have*

$$\frac{d}{d\tau} \int_{\mathcal{M}} f(\cdot, \tau) d\nu(\tau) = 0 \tag{1.4.3}$$

In the sequel, we will see that the notion of diffusion discussed in this section interacts extremely well both with Ricci flow and with Optimal Transportation.

Chapter 2

Optimal Transportation

2.1 Wasserstein distance

Suppose that (\mathcal{M}, g) is a closed (compact, no boundary) Riemannian manifold, and ν_1 and ν_2 are two Borel probability measures on \mathcal{M} .¹ For $p \in [1, \infty)$, the p -Wasserstein distance W_p between ν_1 and ν_2 is defined to be

$$W_p^g(\nu_1, \nu_2) := \left[\inf_{\pi \in \Gamma(\nu_1, \nu_2)} \int_{\mathcal{M} \times \mathcal{M}} d^p(x, y) d\pi(x, y) \right]^{\frac{1}{p}}, \quad (2.1.1)$$

where $d(\cdot, \cdot)$ is the Riemannian distance function induced by g and $\Gamma(\nu_1, \nu_2)$ is the space of probability measures on $\mathcal{M} \times \mathcal{M}$ with marginals ν_1 and ν_2 . (In other words, $\pi \in \Gamma(\nu_1, \nu_2)$ if the push forward² of π under the projections $\mathcal{M} \times \mathcal{M} \mapsto \mathcal{M}$ onto the first and second components are ν_1 and ν_2 respectively. Or put another way, if $A \subset \mathcal{M}$ then $\pi(A \times \mathcal{M}) = \nu_1(A)$ and $\pi(\mathcal{M} \times A) = \nu_2(A)$.)

2.2 Contractivity for diffusions on evolving manifolds

A basic principle in the subject is that two probability measures evolving under an appropriate diffusion equation should get closer in the Wasserstein sense provided the manifold satisfies some curvature condition, most famously positive Ricci curvature. This principle has many incarnations in the literature; see Sturm and von Renesse [20] and the references therein. The geometric side of the subject can be followed back to the work of Otto and Villani [16].

¹From now on, all maps, measures and sets considered will be assumed to be Borel.

² Given a map $F : M \rightarrow \hat{M}$ between manifolds and a probability measure ν on M , the push-forward $F_{\#}\nu$ of ν under F is the probability measure on \hat{M} , defined by $(F_{\#}\nu)[V] = \nu[F^{-1}(V)]$ for all $V \subset \hat{M}$.

In [14], McCann and the author showed that this type of contractivity on an *evolving* manifold $(\mathcal{M}, g(\tau))$ characterises super-solutions of the Ricci flow (parametrised backwards in time) by which we mean solutions to

$$\frac{\partial g}{\partial \tau} \leq 2 \operatorname{Ric}(g(\tau)). \quad (2.2.1)$$

Here the relevant notion of diffusion on an evolving manifold $(\mathcal{M}, g(\tau))$ was given in Definition 1.4.1.

Theorem 2.2.1. (cf. [14, Theorem 2].) *Suppose that \mathcal{M} is a closed manifold equipped with a smooth family of metrics $g(\tau)$ for $\tau \in [\tau_1, \tau_2] \subset \mathbb{R}$. Then the following are equivalent:*

- (A) $g(\tau)$ is a super Ricci flow (i.e. satisfies (2.2.1));
- (B₁) whenever $\tau_1 < a < b < \tau_2$ and $\nu_1(\tau), \nu_2(\tau)$ are diffusions (Definition 1.4.1) for $\tau \in (a, b)$, the function $\tau \mapsto W_1^{g(\tau)}(\nu_1(\tau), \nu_2(\tau))$ is weakly decreasing in $\tau \in (a, b)$;
- (B₂) whenever $\tau_1 < a < b < \tau_2$ and $\nu_1(\tau), \nu_2(\tau)$ are diffusions (Definition 1.4.1) for $\tau \in (a, b)$, the function $\tau \mapsto W_2^{g(\tau)}(\nu_1(\tau), \nu_2(\tau))$ is weakly decreasing in $\tau \in (a, b)$;
- (C) whenever $\tau_1 < a < b < \tau_2$ and $f : \mathcal{M} \times (a, b) \rightarrow \mathbb{R}$ is a solution to $-\frac{\partial f}{\partial \tau} = \Delta_{g(\tau)} f$, the Lipschitz constant of $f(\cdot, \tau)$ with respect to $g(\tau)$ (i.e. the function $\tau \mapsto \sup_M |\nabla f(\cdot, \tau)|$) is weakly increasing in τ .

The characterisation $(A) \iff (B_2) \iff (C)$ was proved in [14]. It was since extended to handle the W_p distance for arbitrary p [1]. The only other case relevant to us is the case $p = 1$ (i.e. (B₁)); the extension to $p = 1$ was first observed by Tom Ilmanen.

By taking flows $g(\tau)$ which are independent of τ in the theorem above, one recovers an existing theory of contractivity on manifolds of weakly positive Ricci curvature (see [20] and the references therein).

In some sense, the most important proof is the implication $(A) \implies (B_2)$ [14] because the techniques therein extend to give the rigorous proof of the key result, Theorem 3.4.1, which we describe in the next chapter. However, here will describe the more elementary (but elegant) implications $(A) \implies (C)$ and $(A) \implies (B_1)$ and use them in the next chapter to give a *formal* derivation of Theorem 3.4.1.

Proof. $(A) \implies (C)$. Define $t = b - \tau$ so that f satisfies the standard heat equation $\frac{\partial f}{\partial t} = \Delta f$ on the super Ricci flow $g(\cdot)$. We remark that it is important that the heat equation ‘flows’ at the same rate as the Ricci flow, by which we mean that if f satisfied $\frac{\partial f}{\partial t} = c \Delta f$ for $c \neq 1$, then the implication would not work.

We compute (taking care with signs which can be confusing because g is the notation used

both for the metric on $T\mathcal{M}$ and $T^*\mathcal{M}$)

$$\begin{aligned}
\frac{\partial|\nabla f|^2}{\partial t} &= \frac{\partial|df|^2}{\partial t} \\
&= -\frac{\partial g}{\partial t}(df, df) + 2\langle d\frac{\partial f}{\partial t}, df \rangle \\
&= \frac{\partial g}{\partial \tau}(\nabla f, \nabla f) + 2\langle d\Delta f, df \rangle \\
&= \frac{\partial g}{\partial \tau}(\nabla f, \nabla f) - 2\text{Ric}(\nabla f, \nabla f) + \Delta|\nabla f|^2 - 2|\text{Hess}(f)|^2,
\end{aligned} \tag{2.2.2}$$

by the Bochner formula. By the definition (2.2.1) of super Ricci flow, we see that

$$\frac{\partial|\nabla f|^2}{\partial t} \leq \Delta|\nabla f|^2,$$

and therefore by the maximum principle (see for example [22, §3.1]) we conclude that $\sup_{\mathcal{M}} |\nabla f|^2$ is a decreasing function of t . \square

Proof. (A) \implies (B₁). Suppose that $g(\tau)$, $\nu_1(\tau)$ and $\nu_2(\tau)$ are defined for τ in a neighbourhood of $\tau_0 \in (a, b)$. By Kantorovich-Rubinstein duality (see e.g. [24, §1.2.1]) we have

$$\begin{aligned}
W_1^{g(\tau)}(\nu_1(\tau), \nu_2(\tau)) &:= \max \left\{ \int_{\mathcal{M}} \varphi d\nu_1(\tau) - \int_{\mathcal{M}} \varphi d\nu_2(\tau) \mid \varphi : \mathcal{M} \rightarrow \mathbb{R} \text{ is Lipschitz} \right. \\
&\quad \left. \text{and } \|\varphi\|_{Lip} \leq 1 \text{ with respect to } g(\tau) \right\}.
\end{aligned} \tag{2.2.3}$$

Let $\varphi_0 : \mathcal{M} \rightarrow \mathbb{R}$ be a function which achieves the maximum in this variational problem at time τ_0 , and extend φ_0 to a function $\varphi : \mathcal{M} \times [\tau_0 - \varepsilon, \tau_0] \rightarrow \mathbb{R}$ for some $\varepsilon > 0$ by solving the equation

$$-\frac{\partial \varphi}{\partial \tau} = \Delta_{g(\tau)} \varphi.$$

By the implication (A) \implies (C) of the theorem (proved above) for all $\tau \in [\tau_0 - \varepsilon, \tau_0]$ we have $\|\varphi(\cdot, \tau)\|_{Lip} \leq 1$ and therefore $\varphi(\cdot, \tau)$ can be used as a competitor in the variational problem (2.2.3) to see that

$$W_1^{g(\tau)}(\nu_1(\tau), \nu_2(\tau)) \geq \int_{\mathcal{M}} \varphi(\cdot, \tau) d\nu_1(\tau) - \int_{\mathcal{M}} \varphi(\cdot, \tau) d\nu_2(\tau).$$

But by Proposition 1.4.2 we know that the functions

$$\tau \mapsto \int_{\mathcal{M}} \varphi(\cdot, \tau) d\nu_1(\tau) \quad \text{and} \quad \tau \mapsto \int_{\mathcal{M}} \varphi(\cdot, \tau) d\nu_2(\tau)$$

are each independent of τ , so we deduce that

$$W_1^{g(\tau)}(\nu_1(\tau), \nu_2(\tau)) \geq W_1^{g(\tau_0)}(\nu_1(\tau_0), \nu_2(\tau_0)).$$

\square

As for the implication (A) \implies (B₂), one considers Wasserstein geodesics between $\nu_1(\tau)$ and $\nu_2(\tau)$ and analyses the extent to which the Boltzmann entropy is convex along these geodesics. See [14] for the full story.

Chapter 3

The Canonical Solitons

3.1 Definition of the Canonical Shrinking Soliton

In the last chapter, we saw an elementary but fundamental result combining Ricci flow and optimal transportation. It turns out that there is a surprising way of applying it to obtain more exotic and powerful corollaries. The trick, when trying to understand a Ricci flow $(\mathcal{M}^n, g(t))$, $t \in [0, T]$, is to not apply the result directly to that Ricci flow, but to allow that Ricci flow to induce a more elaborate Ricci flow on the $(n + 1)$ -dimensional manifold $\mathcal{M} \times [0, T]$ which we engineer to be a soliton flow, and to apply the results of the previous section (and other fundamental results) to that instead.

We call these new $(n + 1)$ -dimensional Ricci flows *Canonical Soliton Flows* and by considering them from different viewpoints, we can construct the main foundations of Ricci flow in a natural way. We get steady, shrinking and expanding Canonical Solitons for each Ricci flow, each of which have different applications. To be precise, we get approximate solitons which can be viewed as exact in the limit of some parameter as we now clarify in the shrinking case.

Theorem 3.1.1. *(Cabezas-Rivas and Topping, [4]) Suppose $g(\tau)$ is a (reverse) Ricci flow – i.e. a solution of $\frac{\partial g}{\partial \tau} = 2 \text{Ric}(g(\tau))$ – defined for τ within a time interval $(a, b) \subset (0, \infty)$, on a manifold \mathcal{M} of dimension $n \in \mathbb{N}$. Suppose $\Omega \subset\subset \mathcal{M}$, $I \subset\subset (a, b)$, and $N \in \mathbb{N}$ is sufficiently large to give a positive definite metric \hat{g} on $\Omega \times I \subset \hat{\mathcal{M}} := \mathcal{M} \times (a, b)$ defined by*

$$\hat{g}_{ij} = \frac{g_{ij}}{\tau}; \quad \hat{g}_{00} = \frac{N}{2\tau^3} + \frac{R}{\tau} - \frac{n}{2\tau^2}; \quad \hat{g}_{0i} = 0,$$

where i, j are coordinate indices on the \mathcal{M} factor, 0 represents the index of the time coordinate $\tau \in (a, b)$, and the scalar curvature of g is written as R .

Then up to errors of order $\frac{1}{N}$, the metric \hat{g} is a gradient shrinking Ricci soliton on the higher

dimensional space $\hat{\mathcal{M}}$:

$$\text{Ric}(\hat{g}) + \text{Hess}_{\hat{g}}\left(\frac{N}{2\tau}\right) \simeq \frac{1}{2}\hat{g}, \quad (3.1.1)$$

by which we mean that on $\Omega \times I \subset \hat{\mathcal{M}}$, the quantity

$$N \left[\text{Ric}(\hat{g}) + \text{Hess}_{\hat{g}}\left(\frac{N}{2\tau}\right) - \frac{1}{2}\hat{g} \right]$$

is bounded independently of N , with respect to any fixed metric on $\hat{\mathcal{M}}$.

We call \hat{g} the Canonical Shrinking Soliton associated with $g(t)$.

Notice that given a Ricci flow over the time interval (a, b) , we are turning the old time parameter τ into an additional space parameter; it then makes sense to introduce a new time parameter (called s in what follows) and consider Ricci flow using $\mathcal{M} \times (a, b)$ as the underlying manifold by taking the Ricci soliton flow $G(s)$ as described in Section 1.3; this is what we call the Canonical Shrinking Soliton Flow.

The essential point of this theorem is the form of the metric \hat{g} . The verification of the result is a calculation. Some parts of that calculation are given in the proposition below; more involved calculations will be required in Chapter 4 for the Canonical *Expanding* Soliton, and can be found in [5].

Proposition 3.1.2. (See [4].) *Fixing $\tau > 0$, a time at which the Ricci flow exists, and fixing local coordinates $\{x^i\}$ in a neighbourhood U of some $p \in \mathcal{M}$, then in any neighbourhood $V \subset\subset U \times (a, b)$ of (p, τ) , we have*

$$\hat{R}_{ij} \simeq R_{ij}; \quad \hat{R}_{i0} \simeq -\frac{1}{2}\nabla_i R; \quad \hat{R}_{00} \simeq -\frac{R_\tau}{2} - \frac{R}{2\tau},$$

where \simeq denotes equality of the coefficients up to an error bounded in magnitude by $\frac{C}{N}$, with $C > 0$ a constant independent of N (but depending on V and the choice of coordinates). Moreover, we have

$$\hat{\nabla}_{ij}^2\left(\frac{N}{2\tau}\right) \simeq \frac{g_{ij}}{2\tau} - R_{ij}; \quad \hat{\nabla}_{i0}^2\left(\frac{N}{2\tau}\right) \simeq \frac{\nabla_i R}{2}; \quad \hat{\nabla}_{00}^2\left(\frac{N}{2\tau}\right) \simeq \frac{N}{4\tau^3} + \frac{R}{\tau} - \frac{n}{4\tau^2} + \frac{R_\tau}{2}.$$

By combining the formulae of Proposition 3.1.2 and the definition of \hat{g} , we deduce Theorem 3.1.1.

The creation of the Canonical Solitons follows a chain of previous work [6, 7, 17] which we discuss further in Chapter 4.

3.2 Perelman's \mathcal{L} -length

It is a general principle that standard concepts and theorems in Riemannian geometry applied to a Canonical Soliton should yield something interesting about the original Ricci flow. The first such instance of this is that considering Riemannian length/distance on $(\hat{\mathcal{M}}, \hat{g})$ induces the fundamental concept of \mathcal{L} -length due to Perelman [17] based on the ideas of Li-Yau [12].

Suppose that $x, y \in \mathcal{M}$ and $[\tau_1, \tau_2] \subset (0, \infty)$ lies within the time domain on which the (reverse) Ricci flow is defined. Consider paths $\Gamma : [\tau_1, \tau_2] \rightarrow \hat{\mathcal{M}}$ connecting (x, τ_1) and (y, τ_2) in $\hat{\mathcal{M}}$ of the form $\Gamma(\tau) = (\gamma(\tau), \tau)$, where $\gamma : [\tau_1, \tau_2] \rightarrow \mathcal{M}$ satisfies $\gamma(\tau_1) = x$ and $\gamma(\tau_2) = y$. Then

$$\begin{aligned}
\text{Length}(\Gamma) &= \int_{\tau_1}^{\tau_2} \left| \gamma'(\tau) + \frac{\partial}{\partial \tau} \right|_{\hat{g}} d\tau \\
&= \int_{\tau_1}^{\tau_2} \left[\frac{|\gamma'|_{g(\tau)}^2}{\tau} + \frac{N}{2\tau^3} + \frac{R}{\tau} - \frac{n}{2\tau^2} \right]^{\frac{1}{2}} d\tau \\
&= \int_{\tau_1}^{\tau_2} \left[\frac{N}{2\tau^3} \right]^{\frac{1}{2}} \left(1 + \frac{2\tau^2 |\gamma'|_{g(\tau)}^2}{N} + \frac{2\tau^2 R}{N} - \frac{\tau n}{N} \right)^{\frac{1}{2}} d\tau \tag{3.2.1} \\
&= \int_{\tau_1}^{\tau_2} \left[\frac{N}{2\tau^3} \right]^{\frac{1}{2}} \left(1 + \frac{\tau^2 |\gamma'|^2}{N} + \frac{\tau^2 R}{N} - \frac{\tau n}{2N} + O\left(\frac{1}{N^2}\right) \right) d\tau \\
&= \sqrt{2N} (\tau_1^{-\frac{1}{2}} - \tau_2^{-\frac{1}{2}}) + \frac{1}{\sqrt{2N}} \hat{\mathcal{L}}(\gamma) + O\left(\frac{1}{N^{3/2}}\right),
\end{aligned}$$

where

$$\hat{\mathcal{L}}(\gamma) := \int_{\tau_1}^{\tau_2} \sqrt{\tau} \left(R(\gamma(\tau), \tau) + |\gamma'(\tau)|^2 - \frac{n}{2\tau} \right) d\tau. \tag{3.2.2}$$

Thus we get an expansion for the length of Γ in terms of N . Unsurprisingly, given the extreme stretching of the $\frac{\partial}{\partial \tau}$ direction in the Canonical Soliton, the leading order term only depends on the times τ_1 and τ_2 , not on the points x and y . However, the quantity $\hat{\mathcal{L}}$ thrown up as the next order term turns out to be one of the most important concepts in the theory of Ricci flow. Traditionally, one would write

$$\hat{\mathcal{L}}(\gamma) = \mathcal{L}(\gamma) - n(\sqrt{\tau_2} - \sqrt{\tau_1}),$$

where

$$\mathcal{L}(\gamma) := \int_{\tau_1}^{\tau_2} \sqrt{\tau} (R(\gamma(\tau), \tau) + |\gamma'(\tau)|^2) d\tau$$

is Perelman's \mathcal{L} -length [17, §7].

Just as in Riemannian geometry, one can use such a notion of length to generate a notion of distance:

$$\hat{Q}(x, \tau_1; y, \tau_2) = \inf_{\gamma} \hat{\mathcal{L}}(\gamma),$$

where the infimum is again over curves γ as above, and for consistency with prior notation, we write $Q(x, \tau_1; y, \tau_2) = \inf_{\gamma} \mathcal{L}(\gamma)$. Note, however, that these lengths and distances are liable to

be negative if the Ricci flow has any negative scalar curvature. For τ_2 only marginally larger than τ_1 , the scalar curvature part of the definition of $\hat{\mathcal{L}}$ becomes irrelevant and one recovers Riemannian distance in that

$$\lim_{\tau_2 \downarrow \tau_1} 2(\sqrt{\tau_2} - \sqrt{\tau_1})\hat{Q}(x, \tau_1; y, \tau_2) = d_{g(\tau_1)}^2(x, y). \quad (3.2.3)$$

The considerations above suggest that minimising paths Γ should essentially arise from \mathcal{L} -geodesics γ for large N , and that

$$d_{\hat{g}}((x, \tau_1), (y, \tau_2)) = \sqrt{2N}(\tau_1^{-\frac{1}{2}} - \tau_2^{-\frac{1}{2}}) + \frac{1}{\sqrt{2N}}\hat{Q}(x, \tau_1; y, \tau_2) + O(\frac{1}{N^{3/2}}). \quad (3.2.4)$$

3.3 \mathcal{L} -Optimal Transportation

The Canonical Shrinking Soliton also leads naturally to the notion of \mathcal{L} -optimal transportation and the theory concerning it (both of which first arose in [23], before the notion of Canonical Soliton had been discovered). Roughly speaking, \mathcal{L} -optimal transportation is optimal transportation in which the cost function arises from Perelman's \mathcal{L} -length.

The key point is that we will try to learn something significant about a Ricci flow by exploiting Theorem 2.2.1, but instead of applying it to the Ricci flow directly, we will apply it to the Ricci flow on space-time induced by its Canonical Soliton.

Consider two probability measures ν_1 and ν_2 on the underlying manifold \mathcal{M} of a (reverse) Ricci flow $g(\tau)$ defined for $\tau \in I \subset (0, \infty)$. We can view these as measures $\hat{\nu}_1$ and $\hat{\nu}_2$ on space-time $\mathcal{M} \times I$ supported on time slices τ_1 and τ_2 by defining $\hat{\nu}_i$ to be the push-forward of ν_i under F_{τ_i} where for $\tau \in I$, the map $F_\tau : \mathcal{M} \rightarrow \mathcal{M} \times I$ is defined by $x \mapsto (x, \tau)$.

If we now equip space-time $\mathcal{M} \times I$ with the Canonical Shrinking Soliton Riemannian metric (for some large N) then we can consider the W_1 Wasserstein distance between $\hat{\nu}_1$ and $\hat{\nu}_2$. By definition (2.1.1) of W_1 , the expansion (3.2.4) suggests that

$$W_1^{\hat{g}}(\hat{\nu}_1, \hat{\nu}_2) = \sqrt{2N}(\tau_1^{-\frac{1}{2}} - \tau_2^{-\frac{1}{2}}) + \frac{1}{\sqrt{2N}} \inf_{\pi \in \Gamma(\nu_1, \nu_2)} \int_{\mathcal{M} \times \mathcal{M}} \hat{Q}(x, \tau_1; y, \tau_2) d\pi(x, y) + O(\frac{1}{N^{3/2}}), \quad (3.3.1)$$

motivating the following:

Definition 3.3.1. (From [23].) The \mathcal{L} -Wasserstein distance between ν_1 at time τ_1 and ν_2 at time τ_2 is defined to be

$$\mathcal{D}(\nu_1, \tau_1; \nu_2, \tau_2) := \inf_{\pi \in \Gamma(\nu_1, \nu_2)} \int_{\mathcal{M} \times \mathcal{M}} \hat{Q}(x, \tau_1; y, \tau_2) d\pi(x, y). \quad (3.3.2)$$

In the next section we will see that this distance has useful contractivity properties when applied to two diffusions at appropriate times.

3.4 Contractivity of \mathcal{L} -Wasserstein distance

In this section we want to apply Theorem 2.2.1 to the Canonical Soliton Flow in a heuristic fashion, in order to arrive at a \mathcal{L} -Wasserstein contractivity result which one can then go back and prove rigorously. We will be very cavalier with the distinction between ‘approximate’ and ‘exact,’ and will be very loose in making precise the time intervals being considered. (You might like to consider initially only Ricci flows defined on the largest possible time interval $(0, \infty)$.) With hindsight we will see that this is not too important.

By the theory of Ricci solitons described in Section 1.3, we can introduce a new time variable s (since τ is now a spatial coordinate on $\hat{\mathcal{M}}$) and consider the (approximate, reverse) Canonical Soliton Ricci flow $G(s)$ on $\hat{\mathcal{M}} := \mathcal{M} \times (a, b)$ starting at time $s = 1$ with the metric $G(1) = \hat{g}$, given by

$$G(s) := s \psi_s^*(\hat{g}) \tag{3.4.1}$$

where $\psi_s : \hat{\mathcal{M}} \rightarrow \hat{\mathcal{M}}$ is the family of diffeomorphisms with ψ_1 the identity, obtained by integrating the vector field $X := -\frac{1}{s} \hat{\nabla}(\frac{N}{2\tau})$ with $\hat{\nabla}$ representing the gradient with respect to \hat{g} . (Note that for convenience we have shifted time compared with the exposition in Section 1.3.)

We may compute

$$X = \frac{N}{2s\tau^2} \hat{g}_{00}^{-1} \frac{\partial}{\partial \tau} = \frac{\tau}{s} \frac{\partial}{\partial \tau} + O(\frac{1}{N}), \tag{3.4.2}$$

and by neglecting the error of order $\frac{1}{N}$ this integrates to

$$\psi_s(x, \tau) \simeq (x, s\tau). \tag{3.4.3}$$

Therefore the flow $G(s)$ arises by pulling back \hat{g} by a dilation of time and scaling appropriately.

If we then view $G(s)$ as a flow which essentially satisfies part (A) of Theorem 2.2.1, then we can hope to exploit the equivalent part (B_1) of the same theorem. The question is then which diffusions we should consider in (B_1) .

The answer is to take diffusions starting (at $s = 1$) at singular measures supported on time-slices. Suppose $\bar{\tau}_1, \bar{\tau}_2 \in (a, b)$, with $\bar{\tau}_1 < \bar{\tau}_2$ and $\tilde{\nu}_1, \tilde{\nu}_2$ are two measures on \mathcal{M} . Then the initial measures for the diffusions will be $(F_{\bar{\tau}_k})_{\#}(\tilde{\nu}_k)$ for $k = 1, 2$, which are supported on the slices $\mathcal{M} \times \{\bar{\tau}_k\}$.

Because of the extreme stretching of the τ direction in the metric \hat{g} (recall that \hat{g}_{00} is of order N) there is essentially no diffusion of the measures in the τ direction, and we view them at later times $s > 1$ as remaining supported in the time slices $\mathcal{M} \times \{\bar{\tau}_k\}$ and write them $(F_{\bar{\tau}_k})_{\#}(\nu_k(s\bar{\tau}_k))$ for some flows of measures $\nu_k(\cdot)$ with $\nu_k(\bar{\tau}_k) = \tilde{\nu}_k$. These singular measures then evolve mainly under diffusion in the \mathcal{M} factor, under the Laplacian induced by $G_{ij}(s)$. By (3.4.1), (3.4.3) and the definition of \hat{g} from Theorem 3.1.1, on the time-slice $\mathcal{M} \times \{\bar{\tau}_k\}$ we have

$$G_{ij}(s) = s[\psi_s^*(\hat{g})]_{ij} = s[\hat{g}|_{\mathcal{M} \times \{s\bar{\tau}_k\}}]_{ij} = \frac{g_{ij}(s\bar{\tau}_k)}{\bar{\tau}_k}.$$

Therefore, the measures evolve by diffusion in the \mathcal{M} factor under the operator

$$\Delta_{\frac{g(s\bar{\tau}_k)}{\bar{\tau}_k}} = \bar{\tau}_k \Delta_{g(s\bar{\tau}_k)},$$

which implies that the flows of measures $\nu_k(\cdot)$ mentioned above must themselves be diffusions (but on the original Ricci flow $g(\tau)$) with $\nu_k(\bar{\tau}_k) = \tilde{\nu}_k$.

We are now in a position to apply part (B_1) of Theorem 2.2.1 to deduce that our two singular diffusions on the evolving manifold $(\hat{\mathcal{M}}, G(s))$ should get closer in the W_1 sense as s increases. In other words, we have (formally) deduced that

$$s \mapsto W_1^{G(s)}((F_{\bar{\tau}_1})_{\#}\nu_1(s\bar{\tau}_1), (F_{\bar{\tau}_2})_{\#}\nu_2(s\bar{\tau}_2))$$

is weakly decreasing.

If we now push forward this whole construction under the diffeomorphisms ψ_s (recalling the abbreviation $\hat{\nu}_k(\tau) := (F_\tau)_{\#}\nu_k(\tau)$) then we find that

$$s \mapsto W_1^{s\hat{g}}(\hat{\nu}_1(s\bar{\tau}_1), \hat{\nu}_2(s\bar{\tau}_2)) = s^{\frac{1}{2}} W_1^{\hat{g}}(\hat{\nu}_1(s\bar{\tau}_1), \hat{\nu}_2(s\bar{\tau}_2)) \quad (3.4.4)$$

is weakly decreasing, where $\hat{\nu}_i(\tau) := (F_\tau)_{\#}\nu_i(\tau)$.

By the expansion (3.3.1) and the definition of (3.3.2), this implies that

$$s \mapsto \sqrt{2N}(\bar{\tau}_1^{-\frac{1}{2}} - \bar{\tau}_2^{-\frac{1}{2}}) + \frac{s^{\frac{1}{2}}}{\sqrt{2N}} \mathcal{D}(\nu_1(s\bar{\tau}_1), s\bar{\tau}_1; \nu_2(s\bar{\tau}_2), s\bar{\tau}_2)$$

is monotonically decreasing, and hence so is

$$s \mapsto s^{\frac{1}{2}} \mathcal{D}(\nu_1(s\bar{\tau}_1), s\bar{\tau}_1; \nu_2(s\bar{\tau}_2), s\bar{\tau}_2).$$

These heuristic arguments suggest the following theorem which was rigorously proved in [23]:

Theorem 3.4.1. *(Equivalent to [23, Theorem 1.1].) Suppose that $g(\tau)$ is a Ricci flow on a closed manifold \mathcal{M} over an open time interval containing $[\bar{\tau}_1, \bar{\tau}_2]$, and suppose that $\nu_1(\tau)$ and $\nu_2(\tau)$ are two diffusions (in the same sense as in Theorem 2.2.1) defined for τ in neighbourhoods of $\bar{\tau}_1$ and $\bar{\tau}_2$ respectively. Then the distance between the diffusions decays in the sense that for $s \geq 1$ sufficiently close to 1,*

$$\mathcal{D}(\nu_1(s\bar{\tau}_1), s\bar{\tau}_1; \nu_2(s\bar{\tau}_2), s\bar{\tau}_2) \leq s^{-\frac{1}{2}} \mathcal{D}(\nu_1(\bar{\tau}_1), \bar{\tau}_1; \nu_2(\bar{\tau}_2), \bar{\tau}_2).$$

The rigorous proof of this result involves introducing a notion of \mathcal{L} -Wasserstein geodesic through a Ricci flow, and considering the convexity of the Boltzmann entropy along this geodesic (see [23]) in a manner related to the arguments in [14].

3.5 Applications of \mathcal{L} -Wasserstein contractivity

From the \mathcal{L} -optimal transportation construction in Theorem 3.4.1 one can [23] recover Perelman's monotonic quantities (both involving entropies and \mathcal{L} -length) which are central in his work on Ricci flow [17, 18, 19], with the case of Perelman's reduced volume appearing in [13]. We will highlight the case of Perelman's \mathcal{W} -entropy since it illustrates how entropies and \mathcal{L} -length interact and will solve the degeneration issue discussed in Section 1.2. But first we indicate how Theorem 3.4.1 generalises the main implication $(A) \implies (B_2)$ in Theorem 2.2.1.

As mentioned in (3.2.3), by inspection of the definition of $\hat{\mathcal{L}}$ in (3.2.2) we see that when τ_2 is only marginally larger than τ_1 , the $\sqrt{\tau}|\gamma'(\tau)|^2$ part of the integrand dominates and

$$\lim_{\tau_2 \downarrow \tau_1} 2(\sqrt{\tau_2} - \sqrt{\tau_1})\hat{Q}(x, \tau_1; y, \tau_2) = d_{g(\tau_1)}^2(x, y).$$

This suggests that for τ_2 only marginally larger than τ_1 ,

$$2(\sqrt{\tau_2} - \sqrt{\tau_1})\mathcal{D}(\nu_1, \tau_1; \nu_2, \tau_2) \simeq \left[W_2^{g(\tau_1)}(\nu_1, \nu_2) \right]^2, \quad (3.5.1)$$

and since by Theorem 3.4.1 the function

$$s \mapsto 2(\sqrt{s\bar{\tau}_2} - \sqrt{s\bar{\tau}_1})\mathcal{D}(\nu_1(s\bar{\tau}_1), s\bar{\tau}_1; \nu_2(s\bar{\tau}_2), s\bar{\tau}_2) \quad (3.5.2)$$

is decreasing, we are led to property (B_2) :

Corollary 3.5.1. (*McCann and Topping [14]*) *Given two diffusions $\nu_1(\tau)$ and $\nu_2(\tau)$ (Definition 1.4.1) on a reverse Ricci flow $g(\tau)$, the function*

$$\tau \rightarrow W_2^{g(\tau)}(\nu_1(\tau), \nu_2(\tau))$$

is (weakly) decreasing in τ .

Unlike some of the heuristic arguments in these lectures, the one above is relatively easy to make rigorous [23].

3.5.1 Recovering Perelman's \mathcal{W} -entropy

We now want to take the limit $\bar{\tau}_2 \downarrow \bar{\tau}_1$ again in Theorem 3.4.1, but this time with $\nu_1(\tau) = \nu_2(\tau) =: \nu(\tau)$ the same diffusion. In this case, by (3.5.1), the monotonic quantity of (3.5.2) is zero to leading order in $\bar{\tau}_2 - \bar{\tau}_1$. The trick is to look at the next order term in the expansion, which will also be monotonic.

We will need to consider the infinitesimal version of the \mathcal{L} -Wasserstein distance \mathcal{D} implied by the following lemma, which is a variation on a well-known infinitesimal description of the

usual W_2 distance as an H^{-1} norm (see [24]). Given a smooth family of positive probability measures $\nu(\tau)$ on a closed manifold \mathcal{M} , for τ in some neighbourhood of τ_1 , we call a vector field $X \in \Gamma(T\mathcal{M})$ an *advection field* for $\nu(\tau)$ at $\tau = \tau_1$ if there exists a smooth family of diffeomorphisms $\psi_\tau : \mathcal{M} \rightarrow \mathcal{M}$, for τ in a neighbourhood of τ_1 , with ψ_{τ_1} the identity, and such that $(\psi_\tau)_\# \nu(\tau_1) = \nu(\tau)$ and $X = \frac{\partial \psi}{\partial \tau} |_{\tau=\tau_1}$.

Lemma 3.5.2. (Lemma 1.3 from [23].) *Suppose $g(\tau)$ is a (reverse) Ricci flow, and $\nu(\tau)$ is a smooth family of positive probability measures on a closed manifold \mathcal{M} , for τ in some neighbourhood of $\tau_1 \in \mathbb{R}$. Then as $\tau_2 \downarrow \tau_1$,*

$$\mathcal{D}(\nu(\tau_1), \tau_1; \nu(\tau_2), \tau_2) = (\tau_2 - \tau_1) \left[\inf_X \sqrt{\tau_1} \int_{\mathcal{M}} \left(R(\cdot, \tau_1) + |X|_{g(\tau_1)}^2 - \frac{n}{2\tau_1} \right) d\nu(\tau_1) \right] + o(\tau_2 - \tau_1), \quad (3.5.3)$$

where the infimum is taken over all advection fields X for $\nu(\tau)$ at $\tau = \tau_1$.

We call the coefficient of $(\tau_2 - \tau_1)$ in (3.5.3) (the part within square brackets) the infinitesimal \mathcal{L} -Wasserstein distance, or \mathcal{L} -Wasserstein *speed* of $\nu(\tau)$, with respect to $g(\tau)$, at $\tau = \tau_1$.

One can give an expression for X as the gradient of the solution to a certain elliptic equation (see [23]) but here it suffices to note that by considering Hodge decompositions, we find that there can be only one advection field which is the gradient of a function, and that must be optimal. But if $\nu(\tau)$ is a diffusion, then Fourier's law tells us that $\nabla \ln u$ is an advection field, so it must be optimal.

We now combine these considerations with Theorem 3.4.1 specialised to the situation that $\nu_1(\tau) = \nu_2(\tau) =: \nu(\tau)$. Writing $\bar{\tau}_2 = (1 + \eta)\bar{\tau}_1$ (for small η) we can express the monotonic quantity from Theorem 3.4.1 as

$$\begin{aligned} s^{\frac{1}{2}} \mathcal{D}(\nu_1(s\bar{\tau}_1), s\bar{\tau}_1; \nu_2(s\bar{\tau}_2), s\bar{\tau}_2) &= \eta \left(\tau_1^2 \int_{\mathcal{M}} \left(R + |\nabla \ln u|^2 - \frac{n}{2\tau_1} \right) d\nu(\tau_1) \right) \bar{\tau}_1^{-\frac{1}{2}} + o(\eta) \\ &= \eta \left(\tau_1^2 \mathcal{F}(\tau_1) - \frac{n\tau_1}{2} \right) \bar{\tau}_1^{-\frac{1}{2}} + o(\eta), \end{aligned} \quad (3.5.4)$$

where $\tau_1 := s\bar{\tau}_1$ and where for each τ ,

$$\mathcal{F} = \int_{\mathcal{M}} (R + |\nabla \ln u|^2) u d\mu$$

is Perelman's \mathcal{F} -information functional [17], [22, §6.2]. Theorem 3.4.1 then tells us that

$$\left(\tau^2 \mathcal{F}(\tau) - \frac{n\tau}{2} \right) \text{ is weakly decreasing in } \tau. \quad (3.5.5)$$

In terms of the normalised Boltzmann entropy

$$E(\tau) := - \int_{\mathcal{M}} u \ln u d\mu - \frac{n}{2}(1 + \ln(4\pi\tau)),$$

(normalised so that the heat kernel in Euclidean space would have zero entropy for all time) the decreasing quantity in (3.5.5) can be written $\tau^2 \frac{dE}{d\tau}$. Therefore

$$0 \geq \frac{d}{d\tau}(\tau^2 \frac{dE}{d\tau}) = \tau \frac{d^2}{d\tau^2}(\tau E),$$

and we see that (τE) is convex in τ and in particular, the quantity

$$W(\tau) := \frac{d}{d\tau}(\tau E),$$

is weakly decreasing in τ . A short calculation gives the explicit formula

$$W(\tau) = \int_{\mathcal{M}} [\tau(|\nabla \ln u|^2 + R) - \ln u] u d\mu - n - \frac{n}{2} \ln(4\pi\tau),$$

which is the entropy quantity discovered by Perelman [17].

3.5.2 No local collapsing

In this section we sketch how the monotonicity of the previous section rules out the local collapsing that was discussed in Section 1.2. Recall, we have a Ricci flow defined on some time interval $[0, T)$, and we have to show that there exists $\kappa > 0$ such that for any $\bar{t} \in [0, T)$ and $r \in (0, 1]$, if the sectional curvatures of $g(\bar{t})$ are bounded in norm by r^{-2} in some geodesic ball $B(x, r)$ then the volume of $B(x, r)$ is bounded below by κr^n .

The main step is to prove the following log-Sobolev inequality:

Proposition 3.5.3. *For all $\bar{t} \in [0, T)$, $r \in (0, 1]$ and $\varphi \in C^{0,1}(\mathcal{M})$, there holds on $(\mathcal{M}, g(\bar{t}))$*

$$\int_{\mathcal{M}} \varphi^2 \ln \varphi^2 - \left(\int_{\mathcal{M}} \varphi^2 \right) \ln \left(\int_{\mathcal{M}} \varphi^2 \right) \leq r^2 \int_{\mathcal{M}} (4|\nabla \varphi|^2 + R\varphi^2) + (C - \ln(r^n)) \left(\int_{\mathcal{M}} \varphi^2 \right), \quad (3.5.6)$$

where C depends only on n, T and $(\mathcal{M}, g(0))$.

Proof. (Sketch. For more details, see [22, Chapter 8].) If we view Perelman's entropy from the previous section as a functional¹

$$\mathcal{W}(g, u, \tau) := \int_{\mathcal{M}} [\tau(|\nabla \ln u|^2 + R) - \ln u] u d\mu - n - \frac{n}{2} \ln(4\pi\tau),$$

on probability densities u and numbers $\tau > 0$, for a given metric g , then as we shall see, an inequality

$$\mathcal{W}(g, \cdot, \tau) \geq -C$$

can be viewed as a log-Sobolev inequality on (\mathcal{M}, g) at 'scale' τ .

¹Note the distinction between W and \mathcal{W} !

All closed manifolds admit some log-Sobolev inequality, but if the manifold evolves, then that log-Sobolev inequality could degenerate, i.e. the constants in the inequality could just get worse and worse as we approach a singular time, and the inequality would become unusable. The monotonicity from the previous section will rule out that possibility.

To analyse the flow at our time $\bar{t} \in [0, T)$, we define a reverse time parameter $\tau = \bar{t} + r^2 - t$ and look at the flow over the range $\tau \in [r^2, \bar{t} + r^2]$.

Let us accept the log-Sobolev inequality at $t = 0$

$$\mathcal{W}(g(\tau = \bar{t} + r^2), \cdot, \bar{t} + r^2) \geq -C_0,$$

(noting that $\tau = \bar{t} + r^2$ corresponds to $t = 0$) where C_0 depends on T (i.e. an upper bound for \bar{t}) and $(\mathcal{M}, g(t = 0))$. In particular, if we let $u : \mathcal{M} \times [r^2, \bar{t} + r^2] \rightarrow [0, \infty)$ be the solution of the conjugate heat equation (1.4.2) with $u(\cdot, r^2) = (\int \varphi^2)^{-1} \varphi^2$, then we have $W(\bar{t} + r^2) \geq -C_0$. Appealing to the monotonicity of W , we then also have $W(r^2) \geq -C_0$. Unravelling the definition of $W(r^2)$ using the fact that we know $u(\cdot, r^2)$ gives the proposition. \square

To rule out local collapsing using (3.5.6) is essentially a question of finding the right function φ to plug in to the inequality. Given a ball $B(x, r)$ at time \bar{t} , with volume V , satisfying the assumption that all sectional curvatures are less than r^{-2} , we should take φ to be the function identically equal to 1 on $B(x, r/2)$, equal to zero outside $B(x, r)$, and equal to $2 - 2d(\cdot, x)/r$ on the remaining ‘annulus.’

If we check the order of magnitude of each term in (3.5.6) (we need the fact that $B(x, r/2)$ will also have volume of order V which follows from a theorem of Bishop and Gromov from comparison geometry [8] which we do not discuss here) we find that

$$\int_{\mathcal{M}} \varphi^2 \sim V; \quad - \int_{\mathcal{M}} \varphi^2 \ln \varphi^2 \lesssim V; \quad r^2 \int_{\mathcal{M}} (4|\nabla \varphi|^2 + R\varphi^2) \lesssim V$$

and thus (3.5.6) gives a positive lower bound for $\frac{V}{r^n}$ as desired.

More refined and complete arguments along the lines of this section can be found in [22] and [21].

Bibliography

- [1] M. Arnaudon, A. Coulibaly and A. Thalmaier, *Horizontal diffusion in C^1 path space*. Preprint.
- [2] S. Brendle, *A generalization of Hamilton's differential Harnack inequality for the Ricci flow*. J. Differential Geom. **82** (2009), 207–227.
- [3] S. Brendle and R. Schoen, *Manifolds with 1/4-pinched curvature are space forms*. J. Amer. Math. Soc. **22** (2009) 287–307.
- [4] E. Cabezas-Rivas and P.M. Topping, *The Canonical Shrinking Soliton associated to a Ricci flow*. Preprint (2008) <http://www.warwick.ac.uk/~maseq>
- [5] E. Cabezas-Rivas and P.M. Topping, *The Canonical Expanding Soliton and Harnack inequalities for Ricci flow*. In preparation.
- [6] B. Chow and S.-C. Chu, *A geometric interpretation of Hamilton's Harnack inequality for the Ricci flow*. Math. Res. Lett. **2** (1995) 701–718.
- [7] B. Chow and D. Knopf, *New Li-Yau-Hamilton inequalities for the Ricci Flow via the space-time approach*. J. Differential Geom. **60** (2002), 1–54.
- [8] S. GALLOT, D. HULIN and J. LAFONTAINE, *Riemannian geometry*. (Second edition) Springer-Verlag, 1993.
- [9] R.S. Hamilton, *Three-manifolds with positive Ricci curvature*. J. Differential Geometry **17** (1982) 255–306.
- [10] R.S. Hamilton, *The Harnack estimate for the Ricci flow*. J. Differential Geom. **37** (1993) 225–243.
- [11] R.S. Hamilton, *The formation of singularities in the Ricci flow*. Surveys in differential geometry, Vol. II (Cambridge, MA, 1993) 7–136, Internat. Press, Cambridge, MA, 1995.
- [12] P. Li and S.-T. Yau, *On the parabolic kernel of the Schrödinger operator*. Acta Math. **156** (1986) 153–201.
- [13] J. Lott, *Optimal transport and Perelman's reduced volume*. Preprint (2008). <http://arXiv.org/abs/0804.0343v1>

- [14] R.J. McCann and P.M. Topping, *Ricci flow, entropy and optimal transportation*. To appear, Amer. J. Math. <http://www.warwick.ac.uk/~maseq>
- [15] H. Nguyen, *Invariant curvature cones and the Ricci SSow*. PhD thesis, Australian National University (2007).
- [16] F. Otto and C. Villani. *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*. J. Funct. Anal. **173** (2000) 361–400.
- [17] G. Perelman *The entropy formula for the Ricci flow and its geometric applications*. <http://arXiv.org/abs/math/0211159v1> (2002).
- [18] G. Perelman *Ricci flow with surgery on three-manifolds*. <http://arxiv.org/abs/math/0303109v1> (2003).
- [19] G. Perelman *Finite extinction time for the solutions to the Ricci flow on certain three-manifolds*. <http://arXiv.org/abs/math/0307245v1> (2003).
- [20] K.-T. Sturm and M.-K. von Renesse. *Transport inequalities, gradient estimates, entropy and Ricci curvature*. Comm. Pure Appl. Math. **58** (2005) 923–940.
- [21] P. M. Topping, *Diameter control under Ricci flow*. Comm. Anal. Geom. **13** (2005) 1039–1055
- [22] P.M. Topping, ‘Lectures on the Ricci flow.’ L.M.S. Lecture note series **325** C.U.P. (2006) <http://www.warwick.ac.uk/~maseq/RFnotes.html>
- [23] P.M. Topping, *\mathcal{L} -optimal transportation for Ricci flow*. Preprint (2007) to appear in Journal für die reine und angewandte Mathematik (Crelle’s journal). <http://www.warwick.ac.uk/~maseq>
- [24] C. Villani, Topics in optimal transportation. ‘Grad. Stud. Math.’, **58** A.M.S. 2003.
- [25] Z.-H. Zhang, *On the completeness of gradient Ricci solitons*. Preprint (2008).