Uniformisation as a Bridge Between Ricci Flow and General Relativity in Two Spatial Dimensions

Christo Morison

Supervisors: Prof. N.P. Landsman and Prof. H.B. Posthuma
Second Reader: Dr. R.R.J. Bocklandt

Korteweg-de Vries Institute voor Wiskunde
Universiteit van Amsterdam
Nederland

August 21, 2020

Abstract

The normalised version of Ricci flow is surveyed—with proofs of short- and long-time existence and convergence—then used to prove the Uniformisation Theorem, which states that any closed and connected Riemannian 2-manifold is conformal to a manifold of constant curvature (this proof was first completed by Chen, Lu and Tian after extensive work by Hamilton and Chow). This result is used to inspect the Einstein equations of General Relativity in $(2+1)$ dimensions, where the spacetime is split using the ADM formalism and the spacelike Cauchy hypersurface is classified by genus using the Uniformisation Theorem. Following the work of Moncrief, the Einstein equations are reduced to dynamics on the (finite-dimensional) cotangent bundle of the Teichmüller space of the Cauchy hypersurface. Various comparisons are drawn between Ricci flow and General Relativity.

To solve a clearly posed challenging problem, even one without any apparent contact with physical reality, is always a pleasure.

YVONNE CHOQUET-BRUHAT, 1923-

You know, people think mathematics is complicated. Mathematics is the simple bit. It's the stuff we can understand. It's cats that are complicated. I mean, what is it in those little molecules and stuff that make one cat behave differently than another, or that make a cat? And how do you define a cat? I have no idea.

JOHN CONWAY, 1937-2020

When I was a boy in England long ago, people who travelled on trains with dogs had to pay for a dog ticket. The question arose whether I needed to buy a dog ticket when I was travelling with a tortoise. The conductor on the train gave me the answer: "Cats is dogs and rabbits is dogs but tortoises is insects and travel free according."

FREEMAN DYSON, 1923-2020

*   *   *

Acknowledgements

The majority of this thesis was written during a global pandemic. My heartfelt thanks go out to all healthcare workers and other essential service providers who have kept the world afloat. I would also like to thank my supervisors, Klaas and Hessel, for their guidance and advice, despite our meetings taking place virtually. Most importantly, I am grateful to my parents for their endless support and to E.L. and T.S. for being my constant companions during this peculiar time.

This document was typeset in LaTeX in the Pazo Math font.

Contents

Chapter 1

Introduction

At the beginning of the twentieth century, Poincaré mused (in [Poi04]) whether every closed, simply-connected 3-manifold is homeomorphic to the 3-sphere.[1] Nearly a century of work by various mathematicians resulted in many advances in the direction of resolving the conjecture, including the invention of Ricci flow by Hamilton in 1982 (see [Ham82]). Though intended as as a tool to resolve the Geometrisation of 3-manifolds (as named by Thurston in 1982; see [And04] for a historical survey), Ricci flow was also applicable to the Uniformisation Theorem (first conjectured by Poincaré and Klein in 1882 [Poi82] and 1883 [Kle83], respectively), which can be loosely thought of as a lower-dimensional version of Geometrisation. It can be stated in the context of Riemannian 2-manifolds as follows.

Theorem (Uniformisation)**.** *Every complete Riemannian metric on a closed* 2*-manifold is conformal to a complete metric with constant curvature.*

Though work by Hamilton in [Ham88], Chow in [Cho91] and others furthered the understanding of Ricci flow in 2 dimensions, not until Chen, Lu and Tian's brief paper [CLT06] appeared in 2006 was the Uniformisation Theorem completely proved by Ricci flow methods. The proof leans heavily on the existence, uniqueness and convergence of the normalised Ricci flow, which takes the following form:

$$\partial_t g(t) = (r - R(t))g(t) \quad \text{with} \quad g(0) = g_0,$$

where $g(t)$ is a one-parameter family of Riemannian metrics on a given Riemannian 2-manifold (Σ, g_0) with associated Ricci scalars $R(t)$ and average Ricci scalar r (which is independent of t as a consequence of the Gauss-Bonnet Theorem). The main idea is to prove the following result, which will imply the Uniformisation Theorem.

Theorem. *On a closed and connected* 2*-dimensional Riemannian manifold* (Σ, g_0)*, there exists a unique solution (which exists for all time) of Riemannian metrics* $g(t)$ *to the normalised Ricci flow with initial metric* $g(0) = g_0$ *such that* $g(t)$ *converges uniformly as* $t \to \infty$ *(in any* C^k *norm) to a metric* g_∞ *of constant curvature.*

This result will take a considerable portion of the text to prove. Short-time existence (and uniqueness) is established by comparing the normalised Ricci flow to a modified

[1] The original question, translated from the French by Gray in [Gra13], is as follows:

> One question remains to be dealt with: Is it possible for the fundamental group of V to reduce to the identity without V being simply-connected?

formulation called the Ricci-DeTurck flow which is (strongly) parabolic and thus whose solutions exist and are unique by standard parabolic theory. Long-time existence is argued by a standard *a priori* argument, which relies on bounds that are established later on. The proof of convergence is then split into various cases, depending on the sign of the average Ricci scalar r. The cases $r \leq 0$ are comparatively straightforward, using a reaction-diffusion maximum principle applied to equations dictating the evolution of various quantities to show that $R(t) \to r$ as $t \to \infty$ (in which case the limiting metric will have constant curvature, as desired). When $r > 0$, the flow will be shown to tend metrics to gradient Ricci solitons, which are metrics that satisfy the following:

$$(R(t) - r)g(t) = 2\nabla\nabla f(t),$$

where $f(t)$ is called the (scalar) potential function of the curvature. The key result from Chen, Lu and Tian was proving that gradient Ricci solitons have constant curvature.[2] Other tools will be necessary, including more uses of the reaction-diffusion maximum principle, as well as so-called Harnack inequalities.

Another system of partial differential equations whose solution is a semi-Riemannian metric also received great attention in the twentieth century: the Einstein equations of General Relativity, first published by their namesake in [Ein15]. Though these equations—as shown by Choquet-Bruhat (see [CB52] and subsequently [CBR83])—are (quasi-linear) hyperbolic, rather than (weakly) parabolic like Ricci flow, the initial goal for this project was to connect the two subjects, either by finding intersections or by using tools and perspectives invented for one to help better understand the other.

Some researchers have published papers that link the two subjects: Fischer, for example, in his work with Moncrief on reducing the Einstein equations in $(3+1)$ dimensions to dynamics on a cotangent bundle of Teichmüller space (see [FM96a], [FM96b] and [FM96c]) as well as his exploration in [Fis04] of what he calls conformal Ricci flow. More recently, Kröncke has studied the stability of Einstein manifolds and asymptotically locally Euclidean metrics under Ricci flow (see [Krö13], [Krö15] and [DK20], the last of which is co-authored by Deruelle) as well as the constant mean curvature Einstein flow in [FK15] and [FK18], both of which are with Fajman.[3] From a less mathematical perspective, Woolgar studied applications of Ricci flow in physics in [Woo08]. Many of the fruits borne from these studies were not quite accessible (yet), so a focus for this project was made on the simplest case: 2 spatial dimensions. Upon exploration of $(2+1)$-dimensional General Relativity—a higher dimension since in relativity time is intrinsic to the manifold rather than a flow parameter—the Uniformisation Theorem was found to play a significant role in understanding the different cases of solutions to the Einstein equations and was thus chosen to be the linking subject of this thesis.

The ADM (Hamiltonian) formalism introduced in 1959 by Arnowitt, Deser and Misner (see [ADM59]) allows spacetime to be split into 'space' and 'time.' With the Uniformisation Theorem in hand, the $(2+1)$-dimensional context can be classified by conformal class into three cases: positive, zero or negative curvature. As shown in the late 1980s by Moncrief in [Mon86] and [Mon89] (with later developments added in [Mon07]), the $(2+1)$-dimensional Einstein equations can be seen as dynamics on the cotangent bundle of the Teichmüller space[4] $\mathcal{T}^*\mathscr{T}$ of the 2-dimensional spatial portion of spacetime, called

[2] Hamilton and Chow had relied upon the Uniformisation Theorem for their proofs of this fact.

[3] Kröncke also proved a generalisation of our 2-dimensional Ricci flow convergence result with Branding for a general connection (that is, which may have non-vanishing torsion) in [BK17].

[4] The Teichmüller space $\mathscr{T}$ of a Riemann surface has existed in complex geometry since the first half of the twentieth century.

the Cauchy hypersurface.[5] This built on the usage of Teichmüller theory in Riemannian geometry surveyed by Fischer and Tromba in [FT84], wherein the space of Riemannian metrics $\mathscr{M}$ for a given manifold Σ is seen as an infinite-dimensional (Hilbert) manifold, and connected to the space of complex structures $\mathscr{C}$ (on Σ, seen as a Riemann surface) via the space of almost complex structures $\mathscr{A}$. Moncrief proved that the Hamiltonian in $(2+1)$-dimensional General Relativity is the area functional of the Cauchy hypersurface, finding its explicit form in the case of a toroidal (zero curvature) spacetime.

The ADM formalism and Moncrief's reduction to dynamics on $\mathcal{T}^*\mathscr{T}$ is presented in this text, with details that were omitted from the original papers presented in a more comprehensive manner. The definition of the ADM Hamiltonian in terms of the so-called Teichmüller parameters (which parametrise the Teichmüller space) and the eventual definition of Einstein flow are the endpoints of this study. The text will reinforce the perspective of the space of metrics $\mathscr{M}$ as being interesting: in Ricci flow, a solution $g(t)$ can be thought of as a curve $g : [0,\infty) \to \mathscr{M}$ from g_0 to g_∞; in General Relativity, the conformal solution space modulo diffeomorphisms is exactly the cotangent bundle of the Teichmüller space of the Cauchy hypersurface, which is of finite (known) dimension. Though the leading goal of informing General Relativity with knowledge from Ricci flow (or vice versa) was not directly accomplished, surveys of the two subjects in 2 spatial dimensions established firm understandings of both their current pools of research and their future prospects for symbiosis.

* * *

Conventions. (Note: an index of pertinent symbols can be found at the end of the text.)

For the most part, we have attempted to follow the conventions and notation of the literature, while attempting to standardise our fonts. For example, a script font is always for a space of functions—though the converse is not always true: for example, we write $\mathcal{C}^\infty$ (not $\mathscr{C}^\infty$) for smooth functions.

Importantly, we use the term 'positive' to refer to a quantity that cannot be zero (sometimes this is called 'strictly positive,' though we will not use this term here), with 'non-negative' referring to greater-than-or-equal-to zero (and vice versa for negative and non-positive).

We will employ the **Einstein summation convention**: *repeated indices are summed over* throughout the text. In Chapter 4 we will also use **prescript notation**: *the pre-superscript denotes the dimensionality of an object*.

Finally, crucial statements in the main text will be italicised for emphasis, and significant terms will be emboldened when first introduced. Footnotes will frequently contain further (albeit non-crucial) information or commentary, and may be less formal than the main text.

Overview. (Note: each chapter will include a detailed overview at its onset.)

Chapter 2 will provide necessary background and review material for the chapters to follow, including an introduction to semi-Riemannian geometry and the Uniformisation

[5] Moncrief's work on the same subject continued with Fischer into $(3+1)$ dimensions in [FM96a], as mentioned above.

Theorem. Chapter 3 will introduce Ricci flow, and upon focusing on the 2-dimensional case, the Uniformisation Theorem will be proved using the normalised Ricci flow. Chapter 4 explores General Relativity via the Einstein equations, using the ADM (Hamiltonian) formalism to split spacetime into 'space' and 'time' before proving that in $(2+1)$ dimensions the solution space of the Einstein equations is finite-dimensional. Chapter 5 briefly concludes and gives three possible future directions to explore in comparing Ricci flow and General Relativity. Appendix A provides a supplement to some constructions in semi-Riemannian geometry introduced in Section 2.1 and encountered throughout the text. Appendix B contains an introduction to the content written for someone with minimal mathematical background.[6] Bibliography and Indices (of Terms and of Symbols) follow.

Reference Guide. (Note: each chapter will include a summary of pertinent references.)

Much of this text draws from the following resources: differential geometry textbooks by Lee ([Lee06] and [Lee13]) as well as Kobayashi and Nomizu ([KN63] and [KN69]); introductory Ricci flow textbooks by Chow and Knopf [CK04] and Topping [Top06]; the work by Hamilton on 2-dimensional Ricci flow [Ham88], with the final step of the proof of the Uniformisation Theorem in the paper by Chen, Lu and Tian [CLT06]; a standard text on General Relativity by Wald [Wal10], as well as one specific to $(2+1)$ dimensions by Carlip [Car03] and one focusing on splitting spacetime by Gourgoulhon [Gou12]; the original paper on Hamiltonian formalism by Arnowitt, Deser and Misner [ADM59]; the paper by Fischer and Tromba [FT84] on Teichmüller theorem in Riemannian geometry, complemented by textbook on the subject by Tromba [Tro92]; and the work on reduction to Teichmüller space by Moncrief in the papers [Mon86], [Mon89] and [Mon07].

[6] The original contents of this thesis were completed on July 29, 2020; the addition of Appendix B and the subsequent updating of the date on the title page were the only changes made on August 21, 2020.

Chapter 2

Background

Before embarking upon our journey through Ricci flow and General Relativity, we need to be familiar with their common language: semi-Riemannian geometry. The five following sections deal with introductory definitions and examples that will be of utmost importance in the chapters to come. Note that while the bulk of this text treats 2 and 3 dimensions, we will begin in a more general setting, though with particular emphasis on our dimensions of interest.

Overview. Section 2.1 will define tensors, metrics, geodesics, the Levi-Civita connection and conformal Killing fields.[1] Section 2.2 discusses curvature via the Riemann tensor, the Ricci tensor and scalar, the sectional curvature, and how all of they are related. Section 2.3 surveys partial differential equations, the parabolic (reaction-diffusion) maximum principle and Sobolev spaces. Section 2.4 introduces Hilbert manifolds and the space of Riemannian metrics. Section 2.5 states the Uniformisation Theorem and its translation to Riemannian geometry.

Reference Guide. Basic introductions to differential and semi-Riemannian geometry can be found in [Lee13] and [Lee06], respectively. More advanced differential geometry is found in the classic texts [KN63] and [KN69]. An advanced text on Riemannian geometry is [Cha06]. A standard text on partial differential equations is [Eva10]. Hilbert manifolds, including the space of metrics, the space of almost complex structures and other infinite-dimensional spaces of pertinence in Riemannian geometry are discussed in [Tro92].

2.1 Introductory Semi-Riemannian Geometry

We will write $\mathcal{M}$ for a smooth, orientable, connected and closed n-dimensional manifold. In the case where $n = 2$, we will frequently write Σ, and sometimes call it a surface. The **space of smooth functions** on $\mathcal{M}$ is written $\mathcal{C}^\infty(\mathcal{M})$ and a generic element of it is usually called f.[2] The **diffeomorphism group** of $\mathcal{M}$ is written $\mathscr{D}(\mathcal{M})$, often shortened to $\mathscr{D}$ when no confusion may arise.

A generic point is usually written $p \in \mathcal{M}$, and **coordinates** are written $(x^i)_{n=1}^n$. Partial derivatives $(\partial_i)_{i=1}^n$ along these directions and their duals $(\mathrm{d}x^i)_{i=1}^n$ form bases of the **tangent** and **cotangent spaces** $\mathcal{T}_p\mathcal{M}$ and $\mathcal{T}_p^*\mathcal{M}$, respectively.[3] The unions of these spaces

[1] Appendix A contains a more thorough introduction to the topics found in Section 2.1.

[2] Unless otherwise noted, we assume that every function (no matter on what space) is smooth.

[3] We use the standard notation $\partial_i := \partial/\partial x^i$ and $\mathrm{d}x^j(\partial_i) = \delta_i^j$.

form the **tangent** and **cotangent bundles** ($\mathcal{TM}$ and $\mathcal{T}^*\mathcal{M}$), whose sections[4] are called vector and covector fields, though this is usually shortened to **vectors** and **covectors**.[5] Similarly, we will often drop the point where the vector is defined: V will be written for $V_p \in \mathcal{T}_p\mathcal{M}$. Two useful operations are the pull-back and the push-forward, defined as follows.

Definition 2.1.1 (Pull-back and push-forward)**.** Let $\mathcal{M}$ and $\mathcal{N}$ be manifolds. Then,

1. For $\psi : \mathcal{M} \to \mathcal{N}$ a smooth map and $f \in \mathcal{C}^\infty(\mathcal{N})$, the **pull-back** of f by ψ, written $\psi^* f$, is given by
$$\psi^* f := \psi \circ f$$
and is in $\mathcal{C}^\infty(\mathcal{M})$;

2. For $\psi : \mathcal{M} \to \mathcal{N}$ a smooth map between manifolds, the **push-forward** (or differential) of ψ at a point $p \in \mathcal{M}$ is a linear map $(\psi_*)_p : \mathcal{T}_p\mathcal{M} \to \mathcal{T}_{\psi(p)}\mathcal{N}$ given by
$$(\psi_*)_p(V_p)(f) := V_p(f \circ \psi),$$
for $V_p \in \mathcal{T}_p\mathcal{M}$ and $f \in \mathcal{C}^\infty(\mathcal{M})$.

We now define the fundamental objects of semi-Riemannian geometry: tensors.

Definition 2.1.2 (Tensor)**.** The (k,l)**-tensor bundle** $\mathcal{T}^{(k,l)}\mathcal{M}$ over $\mathcal{M}$ is the tensor product bundle given as

$$\mathcal{T}^{(k,l)}\mathcal{M} := \underbrace{\mathcal{T}^*\mathcal{M} \otimes \cdots \otimes \mathcal{T}^*\mathcal{M}}_{k \text{ times}} \otimes \underbrace{\mathcal{TM} \otimes \cdots \otimes \mathcal{TM}}_{l \text{ times}}.$$

Sections of this bundle are known as (k,l)**-tensors**, or (k,l)-tensor fields, and are said to be of **rank-**$(k+l)$. They can be written in local coordinates around $p \in \mathcal{M}$ as

$$\tau_p = \tau^{b_1\cdots b_l}_{a_1\cdots a_k}\Big|_p \partial_{b_1} \otimes \cdots \otimes \partial_{b_l} \otimes \mathrm{d}x^{a_1} \otimes \cdots \otimes \mathrm{d}x^{a_k},$$

for $\tau^{b_1\cdots b_l}_{a_1\cdots a_k}$ the function such that at $p \in \mathcal{M}$ we have

$$\tau^{b_1\cdots b_l}_{a_1\cdots a_k}\Big|_p = \tau_p(\mathrm{d}x^{a_1}, \ldots, \mathrm{d}x^{a_k}, \partial_{b_1}, \ldots, \partial_{b_l}),$$

where we use the fact that elements of the form $\partial_{b_1} \otimes \cdots \otimes \partial_{b_l} \otimes \mathrm{d}x^{a_1} \otimes \cdots \otimes \mathrm{d}x^{a_k}$ form a basis, where all indices run from 1 to n.[6] Note that these constructions do not rely on the basis chosen, which is a fact about tensors that we will rely upon heavily. We will drop the tensor product from our notation, calling instead things of the form ζ_j vectors, ω^i covectors, and τ^{mn}_{ijk} (or things of the sort) tensors. Note that the above discussion describes a correspondence between maps that are $(k+l)$-linear over $\mathcal{C}^\infty(\mathcal{M})$ and (k,l)-tensors, so we can equivalently think of them as maps

$$\tau : \underbrace{\mathcal{TM} \times \cdots \times \mathcal{TM}}_{k \text{ times}} \times \underbrace{\mathcal{T}^*\mathcal{M} \times \cdots \times \mathcal{T}^*\mathcal{M}}_{l \text{ times}} \to \mathbb{R}. \tag{2.1}$$

[4]Recall that a **section** of a vector bundle $\pi : \mathcal{E} \to \mathcal{M}$ is a right-inverse of π, generically written s, living in the space $\Gamma(\mathcal{E})$. See Section A.2 in the appendix for details.

[5]Sections of the bundle $\wedge^k\mathcal{T}^*\mathcal{M}$ are called k-forms, whose space is $\Omega^k(\mathcal{M})$.

[6]Note that by linear algebraic properties of the wedge product, k**-forms** are $(k,0)$-tensors that are anti-symmetric under the exchange of any two indices.

We can now define three operations on tensors: tensoring, tracing and Lie derivation.

Definition 2.1.3 (Operations on tensors)**.** Two tensors τ and ϑ can be **tensored** by the pointwise multiplication of the maps in (2.1): $(\tau\vartheta)^{dj}_{abci} := \tau^{d}_{abc}\vartheta^{j}_{i}$.

Tensors can also be **traced**: $\tau^{d}_{adc} = \vartheta_{ac}$, where we have employed the **Einstein summation convention**: *repeated indices are summed over*. Note that we will use the same symbol for the traced tensor (as with the Ricci and Riemann tensors, soon to be defined) since there is no ambiguity on which tensor bundles they live because their number of indices is different.

The final operation on tensors that we will use is the usual **Lie derivative** $\mathcal{L}_V$ (along some vector field V), which satisfies the following properties:

1. For a vector field V on a manifold $\mathcal{M}$ and $f \in \mathcal{C}^\infty(\mathcal{M})$, we have $\mathcal{L}_V f = V(f)$;
2. For vector fields V and W, we have $\mathcal{L}_V W = [V, W]$, the commutator of vector fields;
3. The Lie derivative obeys a Leibniz-type rule with respect to tensoring:

   $$\mathcal{L}_V(\tau \otimes \vartheta) = (\mathcal{L}_V\tau) \otimes \vartheta + \tau \otimes (\mathcal{L}_V\vartheta)\,,$$

   for tensors τ and ϑ and a vector field V;
4. The Lie derivative commutes with all tracing operations.

We now define the metric, the most crucial tensor in our repertoire.

Definition 2.1.4 (Metric)**.** A **metric** on a tangent bundle is a $(2,0)$-tensor field g that is symmetric and non-degenerate. That is, for every point $p \in \mathcal{M}$ we have a bilinear map $g_p : \mathcal{T}_p\mathcal{M} \times \mathcal{T}_p\mathcal{M} \to \mathbb{R}$ that is

1. **Symmetric**: $g_p(V_p, W_p) = g_p(W_p, V_p)$ for all vector fields V and W;
2. **Non-degenerate**: if $g_p(V_p, W_p) = 0$ for all tangent vectors $W_p \in \mathcal{T}_p\mathcal{M}$ then $V_p = 0$,

where these maps vary smoothly with p:[7] for all vector fields V and W, the function $g(V, W)$ is in $\mathcal{C}^\infty(\mathcal{M})$.

Because the maps g_p are symmetric they can be diagonalised (with no zeroes on the diagonal by non-degeneracy), giving rise to a basis-independent object called the **signature**: the collection of the signs of the diagonal. In general, a metric of the above form is called **semi-Riemannian**; however, we highlight two cases:

1. The metric is called **Riemannian** if all signs in the signature are the same;
2. In the special case where the signature has all same signs except one—for example, in standard four-dimensional physics, we have signature $(-+++)$ or $(+---)$[8]—the metric is called **Lorentzian**. This is significant because it allows us to treat $(n+1)$-dimensional spacetime—with one temporal coordinate and n spatial coordinates—as an $(n+1)$-manifold equipped with a Lorentzian metric. Latin indices (usually beginning with i, j, and k) will be used for spatial coordinates, and

[7] In the future, we will discuss relaxing this rule. For now, we continue with the ideal case of all of our maps being smooth.

[8] We will use the first convention, though the world of physics is divided in this regard.

Greek indices (μ, ν, σ and so on) will be used for spacetime coordinates. With this in mind, we have the **Minkowski metric** η, or flat metric, given in locally by

$$\eta_{\mu\nu} := -\mathrm{d}t^2 + \sum_{i=1}^{n} \left(\mathrm{d}x^i\right)^2 = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{2.2}$$

Manifolds equipped with the above metrics are called **semi-Riemannian**, **Riemannian**, and **Lorentzian manifolds**, respectively, and are written $(\mathcal{M}, g)$.[9]

A map $\psi : \mathcal{M} \to \mathcal{N}$ between semi-Riemannian manifolds $(\mathcal{M}, g)$ and $(\mathcal{N}, h)$ is called an **isometry** if it is a diffeomorphism and satisfies $\psi^* h = g$. A metric is **diffeomorphism-invariant** if this holds for $(\mathcal{N}, h) = (\mathcal{M}, g)$. The group formed by the set of all such diffeomorphisms is called the **isometry group** of $(\mathcal{M}, g)$ and is written $\mathscr{I}(\mathcal{M}, g)$, or $\mathscr{I}$ when no confusion may arise.

The metric also gives us a map[10]

$$g_p^\flat : \mathcal{T}_p\mathcal{M} \to \mathcal{T}_p^*\mathcal{M}, \quad g_p^\flat(V_p) = g(V_p, \cdot) \quad \text{for} \quad V_p \in \mathcal{T}_p\mathcal{M},$$

and its inverse

$$g_p^\sharp : \mathcal{T}_p^*\mathcal{M} \to \mathcal{T}_p\mathcal{M}, \quad g_p(g_p^\sharp(\omega_p), V_p) = \omega_p(V_p) \quad \text{for} \quad \omega_p \in \mathcal{T}_p^*\mathcal{M} \text{ and } V_p \in \mathcal{T}_p\mathcal{M},$$

which induce bundle isomorphisms $g^\flat : \mathcal{TM} \to \mathcal{T}^*\mathcal{M}$ and $g^\sharp : \mathcal{T}^*\mathcal{M} \to \mathcal{TM}$ known as **musical isomorphisms**. With the metric and its inverse written locally as g_{ij} and g^{ij} such that $g_{ij}g^{jk} = \delta_i^k$, these isomorphisms allow the metric to **raise** and **lower indices** of tensors. For a vector $V = V^i e_i$ (decomposed in a basis $(e_i)_{i=1}^n$, whose dual basis we write as $(e^i)_{i=1}^n$ in an abuse of notation) and a covector $\omega = \omega_j e^j$ we have

$$V_j := \left(g^\flat(V)\right)_j = g_{ij}V^i \quad \text{and} \quad \omega^i := \left(g^\sharp(\omega)\right)^i = g^{ij}\omega_j.$$

This process extends to general tensors, and, as hinted at before, we continually redefine new tensors when we raise and lower indices while keeping the same symbol for the tensor. For example, for a $(2,2)$-tensor τ we define a $(3,1)$-tensor (still called τ) locally by $\tau_i^{mkl} = g^{mj}\tau_{ij}^{kl}$.

Metrics allow us to measure things on our manifold, as in the following definition. (Note that the following definitions and discussion should have asterisks and slight alterations in the Lorentzian case; once we begin to study General Relativity in Chapter 4, causal structure will play an important role. As such, the following page or so should be read with the Riemannian case in mind, and revisited to make the necessary amendments after Section 4.2.)

Definition 2.1.5 (Geodesic)**.** The **length** of a curve $\gamma : [0,1] \to \mathcal{M}$ (with speed $\dot{\gamma} := \frac{\mathrm{d}}{\mathrm{d}\lambda}\gamma$) is defined to be

$$\ell(\gamma) := \int_0^1 \sqrt{\pm g_{\gamma(\lambda)}\left(\dot{\gamma}(\lambda), \dot{\gamma}(\lambda)\right)}\mathrm{d}\lambda,$$

[9] We will consider the Riemannian and Lorentzian cases in this text, though the step to general semi-Riemannian geometry is usually not too complicated.

[10] This is since $\mathcal{V} \cong \mathcal{V}^*$ for any finite-dimensional vector space $\mathcal{V}$.

where the $+$ is for a Riemannian manifold, and the $-$ is for a Lorentzian manifold. *Note that a further assumption is required in the Lorentzian case: we want the curves γ to be causal, which is described in Definition 4.2.1.* As mentioned, we will revisit this later in Chapter 4; until then, the Riemannian definitions will suffice. The Riemannian length is a functional with input γ and does not depend on the parameter λ. It is interpreted as the integral of the norm of the velocity vector along the curve, which indeed should give a distance.[11]

In the Riemannian case, this curve-length defines a metric $d_g : \mathcal{M} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$ (the **distance**) induced by the (Riemannian) metric g, given by

$$d_g(p,q) := \inf_{\gamma} \ell(\gamma), \tag{2.3}$$

where p and q are points on $\mathcal{M}$ and the infimum is taken over curves γ that start at p and end at q.[12] The **diameter** $\operatorname{diam}(\mathcal{M}, g)$ is the supremum of the possible distances:[13]

$$\operatorname{diam}(\mathcal{M}, g) := \sup \{ d_g(p,q) \mid p, q \in \mathcal{M} \}.$$

Curves γ of extremal length are called **geodesics** and solve the **geodesic equation**:

$$\ddot{\gamma}^i + \Gamma^i_{jk} \dot{\gamma}^j \dot{\gamma}^k = 0, \tag{2.4}$$

where we have defined the **Christoffel symbols** Γ^i_{jk} to be

$$\Gamma^i_{jk} := \frac{1}{2} g^{mi} \left(\partial_j g_{mk} + \partial_k g_{mj} - \partial_m g_{jk} \right), \tag{2.5}$$

which, though they have indices, are not tensors. A Riemannian manifold is called geodesically complete if any two points[14] can be connected by a geodesic, which we assume our manifolds in this text to be. Note that this notion of completeness coincides with the notion of completeness in the metric space sense via the distance function defined above.

By noting that the geodesic equation (2.4) is a second-order ordinary differential equation, we have existence and uniqueness of geodesics by the Picard-Lindelöf Theorem: given a point $p \in \mathcal{M}$ on a semi-Riemannian manifold and a vector $V_p \in \mathcal{T}_p\mathcal{M}$ there exists a unique geodesic written $\gamma_{p,V}$ defined on an open interval containing the origin such that

$$\gamma_{p,V}(0) = p \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}\lambda} \gamma_{p,V}(\lambda) \bigg|_{\lambda=0} = V.$$

This satisfies $\gamma_{p,\beta V}(\lambda) = \gamma_{p,V}(\beta\lambda)$ for all $\beta, \lambda \in \mathbb{R}$ where the above curves are defined.[15] We can now define the exponential map.

---

[11]This also aligns with the intuitive notion of calling the metric the 'line element,' written $g = \mathrm{d}s^2$ in physics, as integrals along γ are then

$$\int_\gamma \sqrt{|g|} = \int_\gamma \sqrt{|\mathrm{d}s^2|} = \int_\gamma \mathrm{d}s.$$

[12]In the Lorentzian case, we have a supremum over possible causal curves. This will be revisited in Chapter 4.

[13]Note that by the Hopf-Rinow Theorem a Riemannian manifold is (geodesically) complete if and only if it is complete as a metric space (see Chapter IV.4 of [KN63]). Furthermore, if a manifold is complete, then it is compact if and only if it has finite diameter.

[14]This needs to be sharpened to 'any two causally-separated points' in the Lorentzian case. This is guaranteed by global hyperbolicity (which we will assume in Chapter 4; see Definition 4.2.5 for details).

[15]This is because the curves $\lambda \mapsto \gamma_{p,\beta V}(\lambda)$ and $\lambda \mapsto \gamma_{p,V}(\beta\lambda)$ both satisfy the geodesic equation and have the same initial value, so are equal by uniqueness.

Definition 2.1.6 (Exponential map). At a point $p \in \mathcal{M}$, we define the **exponential map** $\exp_p$ defined on neighbourhood $\mathcal{U} \subset \mathcal{T}_p\mathcal{M}$ of the origin by its action on a tangent vector $V_p \in \mathcal{U}$ as

$$\exp_p : \mathcal{U} \to \mathcal{M}, \quad \exp_p(V_p) = \gamma_{p,V}(1).$$

This map is smooth, sometimes written as $\exp_p(\lambda V_p) = \gamma_{p,V}(\lambda)$, and often considered as $\exp : \mathcal{TM} \to \mathcal{M}$. We thus obtain a diffeomorphism from a neighbourhood $\mathcal{U}$ of the origin $0 \in \mathcal{T}_p\mathcal{M}$ to a neighbourhood $\mathcal{N}_p$ of $p \in \mathcal{M}$, called a **normal neighbourhood**. We choose coordinates $(x^i)_{i=1}^n$ of $\mathcal{T}_p\mathcal{M}$ such that their partial derivatives form an orthonormal set—these are called **normal coordinates**, or locally inertial coordinates, and satisfy the following:

1. Geodesics are straight lines: $\gamma_{p,V}(\lambda) = \lambda V_p$ is a geodesic for as long as it exists;
2. The metric is flat: $g_{ij} = \delta_{ij}$ (Riemannian case), or $g_{\mu\nu} = \eta_{\mu\nu}$ (Lorentzian case).

In particular, the second point implies that all Christoffel symbols (2.5) vanish. This useful because if we prove a coordinate-independent identity (involving tensors with fully-contracted indices, for example) in normal coordinates, then it is true in general.

The metric also allows us to define integration over our manifold $\mathcal{M}$ as follows.

Definition 2.1.7 (Integration). A **volume form** ω on a semi-Riemannian n-manifold $\mathcal{M}$ is a nowhere-vanishing n-form.[16] We normalise ω as follows:

$$\omega(\partial_1, \dots, \partial_n) = \sqrt{\pm \det g} =: \sqrt{\pm g},$$

where the positive sign is for a Riemannian manifold and the negative sign for a Lorentzian manifold, and the second equality is notational shorthand.[17] The volume form allows us to **integrate** a function $f \in \mathcal{C}^\infty(\mathcal{M})$ over $\mathcal{M}$ as

$$\int_{\mathcal{M}} f := \int_{\mathcal{M}} f\omega = \int_{\mathcal{M}} \sqrt{\pm \det g(x)} f(x) \mathrm{d}^n x.$$

We will almost always omit the volume element, save for when its presence is crucial. This definition allows us to define the **volume** of $\mathcal{M}$ as[18]

$$\mathrm{vol}(\mathcal{M}, g) := \int_{\mathcal{M}} 1.$$

We now turn to a new manner of differentiating tensors.

Definition 2.1.8 (Covariant derivative). A **covariant derivative** is an assignment of a vector field V to a map ∇_V that takes vector fields to vector fields in the following way:

1. The assignment $V \mapsto \nabla_V$ is $\mathcal{C}^\infty(\mathcal{M})$-linear:

$$\nabla_{fV}(W) = f\nabla_V W$$

   for vector fields W;

[16] This exists since we assume all of our manifolds to be orientable.
[17] Note that this is independent of coordinates.
[18] When $n = 2$ this is often called the **area** of $\mathcal{M}$.

2. It satisfies a Leibniz-type rule:

$$\nabla_V(fW) = V(f)W + f\nabla_V W,$$

for scalar functions f.

The tensor $\nabla_V W$ describes the derivative of a vector W along the direction V, and depends only upon the value of V at the point $p \in \mathcal{M}$ where all of this is taking place and the values of W in the coordinate patch in which p lives. Thus, we can make our focus local and consider the above in some coordinate patch. Suppose that for every p in this patch, $(e_i(p))_{i=1}^n$ form a basis of $\mathcal{T}_p\mathcal{M}$ (restricted to the patch in question). From this perspective we find that the covariant derivative is completely determined by the **connection coefficients** A_{ij}^k defined by

$$A_{ij}^k e_k := \nabla_{e_i}(e_j),$$

where we make the important remark that despite the index notation, the connection coefficients are not tensors. Finally, note that in parallel our previous coordinate-shorthand, we will write $\nabla_i := \nabla_{\partial_i}$. Using tensor products the covariant derivative can be uniquely extended to act on generic tensors.[19]

The covariant derivative whose connection coefficients are the Christoffel symbols Γ^i_{jk} is called the **Levi-Civita connection**, which satisfies:

1. **Metric compatibility**: $U(g(V,W)) = g(\nabla_U V, W) + g(V, \nabla_U W)$, for vector fields U, V and W;

2. **Vanishing torsion**: $\nabla_V W - \nabla_W V = [V,W]$.

It exists, is unique, and is given by the formula:

$$\begin{aligned}2g(\nabla_U V, W) &= U(g(V,W)) + V(g(W,U)) - W(g(U,V)) \\ &\quad - g(U,[V,W]) + g(V,[W,U]) + g(W,[U,V]).\end{aligned}$$

We will assume every semi-Riemannian manifold considered in this text to be equipped with the Levi-Civita connection, unless explicitly noted otherwise.

Finally, for a given metric g, the Levi-Civita connection allows us to define the **Laplacian** $\Delta_g = \Delta$ (where we drop the g unless confusion may arise) as $\Delta := g^{ij}\nabla_i\nabla_j$. It acts on any (k,l)-tensor.

We now have two ways of differentiating tensors in hand. To summarise: for a (k,l)-tensor τ and a vector field V on a semi-Riemannian manifold $(\mathcal{M}, g)$ equipped with Levi-Civita connection ∇ with corresponding Christoffel symbols Γ^i_{jk}, we have the two following differentiation rules:

1. **Covariant derivative**:

$$\nabla_\mu \tau^{\omega_1\cdots\omega_l}_{\zeta_1\cdots\zeta_k} = \partial_\mu \tau^{\omega_1\cdots\omega_l}_{\zeta_1\cdots\zeta_k} + \sum_{i=1}^{l} \Gamma^{\omega_i}_{\mu\alpha}\tau^{\omega_1\cdots\alpha\cdots\omega_l}_{\zeta_1\cdots\zeta_k} - \sum_{i=1}^{k} \Gamma^{\beta}_{\mu\zeta_i}\tau^{\omega_1\cdots\omega_l}_{\zeta_1\cdots\beta\cdots\zeta_k},$$

where α and β have replaced ω_i and ζ_i, respectively.

[19] See Section A.3 in the appendix for details.

2. **Lie derivative:**

$$\mathcal{L}_V \tau^{\omega_1\cdots\omega_l}_{\zeta_1\cdots\zeta_k} = \nabla_V \tau^{\omega_1\cdots\omega_l}_{\zeta_1\cdots\zeta_k} + \sum_{i=1}^{k} (\nabla_{\zeta_i} V^\beta) \tau^{\omega_1\cdots\omega_l}_{\zeta_1\cdots\beta\cdots\zeta_k} - \sum_{i=1}^{l} (\nabla_\alpha V^{\omega_i}) \tau^{\omega_1\cdots\alpha\cdots\omega_l}_{\zeta_1\cdots\zeta_k},$$

where, as above, α and β have replaced ω_i and ζ_i, respectively.[20]

One special case of the Lie derivative is important enough to earn its own definition.

Example 2.1.1 (Conformal Killing field)**.** On a semi-Riemannian n-manifold $(\mathcal{M}, g)$ equipped the Levi-Civita connection ∇, we call a vector field V a **conformal Killing field**, or conformal Killing vector, if for some function $\lambda \in \mathcal{C}^\infty(\mathcal{M})$ we have

$$\mathcal{L}_V g = \lambda g,$$

which is known as the **conformal Killing equation**. By tracing both sides we obtain

$$\lambda = \frac{2}{n} \mathrm{div}(V),$$

for $\mathrm{div}(V)$ the **divergence** of the vector field V given in local coordinates by $\nabla_i V^i$. Thus, our conformal Killing equation in local coordinates is

$$\nabla_i V_j + \nabla_j V_i = \frac{2}{n} g_{ij} \nabla_k V^k.$$

Note that when $\lambda = 0$, the above reduces to what is called the **Killing equation**, whose solutions are **Killing vectors**.

We close this section with several simple examples of metrics and their respective Levi-Civita connections.

Example 2.1.2 (Flat space)**.** In $(2+1)$ dimensions with coordinates (t, x, y) the Minkowski metric takes the local form

$$\eta = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It has vanishing Christoffel symbols (that is, $\Gamma^k_{ij} = 0$ for all i, j and k), meaning that covariant derivatives and partial derivatives coincide: $\partial_i = \nabla_i$. Similarly, n-dimensional Euclidean space the flat metric takes the form $g_{ij} = \delta_{ij}$, and so all of the Christoffel symbols vanish as well. The geodesic equation for these spaces takes the form $\ddot{\gamma}(t) = 0$, and we recover the expected straight lines.

In 2 dimensions we can write the flat metric in polar coordinates (ϱ, θ) as

$$g^{\text{polar}} := \mathrm{d}\varrho^2 + \varrho^2 \mathrm{d}\theta^2 = \begin{pmatrix} 1 & 0 \\ 0 & \varrho^2 \end{pmatrix}.$$

Now, using the definition of the Christoffel symbols, we find that all vanish except for

$$\Gamma^\varrho_{\theta\theta} = -\varrho \quad \text{and} \quad \Gamma^\theta_{\varrho\theta} = \Gamma^\theta_{\theta\varrho} = \frac{1}{\varrho}.$$

[20] The Lie derivative is explored in depth in any textbook on differential geometry. In essence, it measures the flow of a tensor along a vector field.

Example 2.1.3 (Hyperbolic half-plane)**.** On the **half-plane** $\mathbb{H}^2 := \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ we can define the following **hyperbolic metric**:

$$g^{\text{hyp}} := \frac{1}{y^2}\left(\mathrm{d}x^2 + \mathrm{d}y^2\right) = \begin{pmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{pmatrix}.$$

Now, computing the Christoffel symbols we find that the only non-zero ones are

$$\Gamma^x_{xy} = \Gamma^x_{yx} = \Gamma^y_{yy} = -\Gamma^y_{xx} = -\frac{1}{y}.$$

Example 2.1.4 (Round metric)**.** On the 2-sphere $\mathbb{S}^2$ we define the natural[21] **round metric** in polar coordinates (θ, ϕ) by

$$g^{\text{round}} := \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2 = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}.$$

The non-zero Christoffel symbols are

$$\Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta \quad \text{and} \quad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta.$$

Example 2.1.5 (Rotationally symmetric)**.** For a 2-dimensional Riemannian manifold (Σ, g), the metric is called **rotationally symmetric** about a point $q \in \Sigma$ if it can be written in polar coordinates (ϱ, θ) (where $\varrho = 0$ is the point q) as

$$g^{\text{rs}} := \mathrm{d}\varrho^2 + h(\varrho)^2\mathrm{d}\theta^2 = \begin{pmatrix} 1 & 0 \\ 0 & h(\varrho)^2 \end{pmatrix},$$

for $h(\varrho)$ positive on some interval $(0, A)$ (with $A > 0$) and vanishing when $\varrho = 0$ and when $\varrho = A$, since our manifold is compact and smooth.[22] Using the definition of the Christoffel symbols, we find that all vanish except for

$$\Gamma^\varrho_{\theta\theta} = -hh' \quad \text{and} \quad \Gamma^\theta_{\varrho\theta} = \Gamma^\theta_{\theta\varrho} = \frac{h'}{h},$$

where we have denoted derivative by ϱ as $(\cdot)'$.

2.2 Curvature in Various Guises

A much better understanding of curvature was one of Riemann's great insights in his work on manifolds. To motivate the definition of the Riemann tensor we take two alternate yet equivalent routes, each demonstrating how different a space is from being 'flat.'

The concept of curvature is often linked to second-order derivatives, or quadratic terms in functions. To generalise this notion to manifolds, consider the normal coordinates from Definition 2.1.6 in some neighbourhood of a point $p \in \mathcal{M}$. We wish to see how far a given Riemannian manifold $(\mathcal{M}, g)$ is from having $g_{ij} = \delta_{ij}$. In normal coordinates,

[21]Natural because it is induced by the embedding $\mathbb{S}^2 \subset \mathbb{R}^3$.

[22]Note that with $h(\varrho) = \varrho$ then we recover the result for flat 2-space in polar coordinates.

the first-order derivatives vanish, and we can define a $(4,0)$-tensor R_{iklj} as the tensor[23] about p satisfying

$$g_{ij} = \delta_{ij} + \frac{1}{3}R_{iklj}x^k x^l + \mathcal{O}(x^3), \tag{2.6}$$

where $\mathcal{O}(x^3)$ denotes higher-than-quadratic terms in x and the factor of $\frac{1}{3}$ ensures that this tensor agrees with our second definition. This $(4,0)$-tensor will become our Riemann tensor, satisfying our intuition that quadratic terms and curvature are related.

The second approach considers the following question: what is the difference in parallel-transporting[24] vector V first along a direction x^k and then along x^l versus first along x^l and then x^k? This is best described as the commutator of covariant derivatives ∇_k and ∇_l. One can think of this as travelling around a parallelogram, as depicted in Figure 2.1: first apply ∇_k then ∇_l, before subtracting ∇_l then ∇_k.

Computing this directly from the definition of the Levi-Civita connection,[25] we have

$$\begin{aligned}[\nabla_k, \nabla_l]V^i &= (\nabla_k\nabla_l - \nabla_l\nabla_k)\, V^i \\ &= \left(\partial_k\Gamma^i_{jl} - \partial_l\Gamma^i_{jk} + \Gamma^i_{km}\Gamma^m_{jl} - \Gamma^i_{lm}\Gamma^m_{jk}\right) V^j.\end{aligned}$$

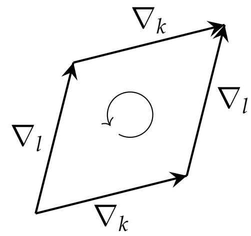

Figure 2.1: Visualisation of the commutator of covariant derivatives.

The term in parenthesis is then identified as the Riemann $(3,1)$-tensor, which has one index raised compared to its $(4,0)$-tensor counterpart described in (2.6). This derivation is often given in physics; its mathematical equivalent is given by considering the following operator $R(U,V)$,[26] which is labelled by vector fields U and V and acts on a third vector field W as follows:

$$R(U,V)W := \nabla_U\nabla_V W - \nabla_V\nabla_U W - \nabla_{[U,V]}W.$$

This collapses to the previous equation since the commutator of partial derivatives (coordinate directions) vanishes.

With the above in mind, we give our official definition of this famous tensor.

Definition 2.2.1 (Riemann tensor)**.** For a semi-Riemannian manifold $(\mathcal{M}, g)$ equipped with the Levi-Civita connection ∇ we define the **Riemann tensor**, or Riemann curvature tensor, as the $(3,1)$-tensor Riem defined by

$$\mathrm{Riem}(U,V,W,\omega) := \omega\left(\nabla_U(\nabla_V W) - \nabla_V(\nabla_U W) - \nabla_{[U,V]}W\right),$$

for vector fields U, V and W and covector ω. In local coordinates, it takes the following form (where we simplify to R for brevity):

$$R^i_{jkl} := \partial_k\Gamma^i_{jl} - \partial_l\Gamma^i_{jk} + \Gamma^i_{km}\Gamma^m_{jl} - \Gamma^i_{lm}\Gamma^m_{jk}.$$

Both of these expressions can be verified to behave tensorially. A semi-Riemannian manifold is called **flat** if its Riemann tensor vanishes.

[23] It is not immediately clear that this is a tensor; however, this is merely motivation for our eventual Definition 2.2.1, so we will (implicitly) consider it to be a result that this object is indeed a tensor.

[24] To **parallel-transport** a tensor (field) τ along a curve γ means taking a covariant derivative of the tensor in the direction of the tangent to γ: $\nabla_{\dot\gamma}\tau = 0$.

[25] Note that this derivation can be achieved with a general connection (and without coordinates, of course), with a term that includes the torsion tensor (defined in Definition A.4.1 of the appendix) appearing.

[26] This operator R is sometimes called the curvature tensor, which may lead to some confusion—hence its relegation to the footnotes.

The Riemann tensor contains a lot of information about the curvature of the manifold, yet many of its components are interrelated via various symmetries.[27] Because there are often too many components to deal with, we define the following tensors that, in some sense, decompose the Riemann tensor.

Definition 2.2.2 (Ricci and Einstein tensors[28])**.** Consider a semi-Riemannian manifold $(\mathcal{M}, g)$.

1. The **Ricci tensor** Ric (written R in coordinates, where there is no confusion with the Riemann tensor because of the number of indices) is a symmetric $(2,0)$-tensor on an n-manifold given by the following trace of the Riemann tensor:

$$R_{kl} := R^i_{kil}.$$

   If the Ricci tensor of a manifold is zero, we call the manifold **Ricci-flat**.

2. The Ricci tensor has the **Ricci scalar** (also written R), or scalar curvature, as a trace:

$$R := g^{kl} R_{kl}.$$

3. These combine in two similar ways to form **Einstein tensors**:

$$E_{kl} := R_{kl} - \frac{1}{n} R g_{kl}, \quad \text{which is used in mathematics, and}$$
$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}, \quad \text{which is used in physics.}$$

We also have the following definition of a particularly nice type of manifold.

Definition 2.2.3 (Einstein manifold)**.** A metric is called an **Einstein metric** if $\mathrm{Ric} = 2\Lambda g$ for some constant $\Lambda \in \mathbb{R}$, and a semi-Riemannian manifold equipped with such a metric is called an **Einstein manifold**.

In our dimensions of interest, we make the following remarks:

1. If $n = 2$, the Ricci tensor has a single independent component (the same as the Riemann tensor), and both Einstein tensors vanish;

2. If $n > 2$, a metric is Einstein if and only if it has vanishing (mathematical) Einstein tensor;

3. If $n = 3$, the Ricci tensor has six independent components (the same as the Riemann tensor).

From these, we can state that in 2 dimensions (as in our case of interest for Ricci flow), the Ricci scalar R contains all possible curvature information, as does the Ricci tensor Ric in 3 dimensions (as in our case of interest for General Relativity). Concretely, in 2 dimensions we have

$$R_{ijkl} = \frac{1}{2} R \left(g_{ik} g_{jl} - g_{il} g_{jk} \right), \tag{2.7}$$

[27] For more details, see Proposition A.4.3 in the appendix.

[28] See Definition A.4.3 of the appendix for yet another tensor commonly defined alongside these three, called the Weyl tensor.

while in 3 dimensions we have

$$R_{ijkl} = \frac{1}{2}R\left(g_{ik}g_{jl} - g_{il}g_{jk}\right) + E_{il}g_{jk} + E_{jk}g_{il} - E_{ij}g_{kl} - E_{kl}g_{ij}. \tag{2.8}$$

We now calculate some of these tensors to build upon our examples from the previous section.

Example 2.2.1 (Flat space #2)**.** The Minkowski, Euclidean, and polar coordinate metrics have vanishing Riemann tensor (and thus Ricci tensor and scalar, too) since the Christoffel symbols are all zero, as we expect for flat space.

Example 2.2.2 (Hyperbolic half-plane #2)**.** Our hyperbolic metric on $\mathbb{H}^2$ has non-zero (independent) Riemann tensor component, Ricci tensor and Ricci scalar:

$$R^x_{\ yxy} = -\frac{1}{y^2}, \quad R_{ij} = \begin{pmatrix} -\frac{1}{y^2} & 0 \\ 0 & -\frac{1}{y^2} \end{pmatrix} \quad \text{and} \quad R = -2.$$

This has constant negative Ricci scalar, as expected for hyperbolic space.

Example 2.2.3 (Round metric #2)**.** Our round metric on $\mathbb{S}^2$ has non-zero (independent) Riemann tensor component, Ricci tensor and Ricci scalar:

$$R^\theta_{\ \phi\theta\phi} = \sin^2\theta, \quad R_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \quad \text{and} \quad R = 2.$$

This has constant positive Ricci scalar, as one would guess for a sphere.

Example 2.2.4 (Rotationally symmetric #2)**.** Our rotationally symmetric metric has non-zero (independent) Riemann tensor component, Ricci tensor and Ricci scalar:

$$R^\theta_{\ \varrho\theta\varrho} = -\frac{h''}{h}, \quad R_{ij} = -h''\begin{pmatrix} \frac{1}{h} & 0 \\ 0 & h \end{pmatrix} \quad \text{and} \quad R = -2\frac{h''}{h}.$$

The first three examples shared something in common: constant curvature. This is a tricky concept to define, however, because the requirement that the entire Riemann tensor is constant is far too restrictive. It is also undefined since constancy could only refer to its components, which are coordinate-dependent. Instead, we define the sectional curvature.

Definition 2.2.4 (Sectional curvature)**.** At a point p on a semi-Riemannian manifold $(\mathcal{M}, g)$, we define the **sectional curvature** K_p of linearly-independent vectors $V_p, W_p \in \mathcal{T}_p\mathcal{M}$ as

$$K_p(V_p, W_p) := \frac{\text{Riem}_p(V_p, W_p, V_p, W_p)}{g_p(V_p, V_p)g_p(W_p, W_p) - g_p(V_p, W_p)^2}.$$

This can be thought of as the normalisation of the Riemann tensor by dividing by the area of a parallelogram spanned by V_p and W_p, which brings to light the requirement of them being linearly independent. With this in hand, the concept of constant curvature arises from the sectional curvature in the following way.

Definition 2.2.5 (Constant curvature)**.** A semi-Riemannian manifold has **constant curvature** if all of its sectional curvatures K_p coincide to some value called the **curvature** K.

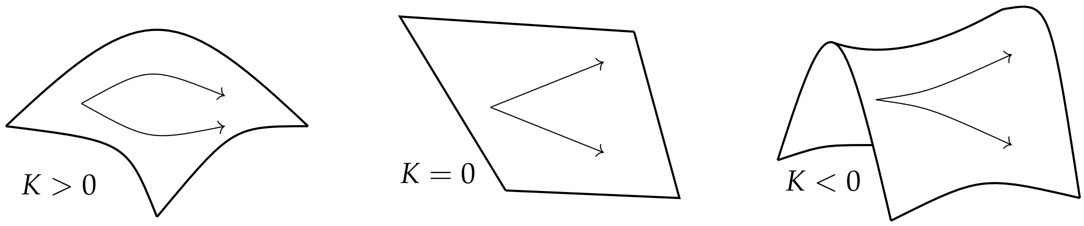

Figure 2.2: Geodesic behaviour in spaces of curvature K.

Spaces of constant curvature can be thought of in terms of their geodesic behaviour: if a space has positive curvature, then geodesics tend toward one another, whereas if a space has negative curvature, geodesics tend away from one another. In a space of zero curvature, geodesics 'remain parallel,' or preserve the same angle between them that they began with. Figure 2.2 qualitatively depicts these three scenarios.

For 2-dimensional cases, this is sometimes thought of as the sum of the interior angles of a triangle drawn on the surface: if this sum exceeds π, that (patch of the) surface has positive curvature; if it is less than π, then it has negative curvature; if it equals π (as in Euclidean space), then it is flat.

We have the following results for spaces of constant curvature.

Proposition 2.2.1. *Let $(\mathcal{M}, g)$ be a semi-Riemannian manifold of constant curvature. Then, it is Einstein and its Riemann tensor can be written as*

$$R_{ijkl} = K(g_{ik}g_{jl} - g_{il}g_{jk}),$$

for K the curvature.

Spaces of constant curvature will be of significance in our study of Ricci Flow, as we will see that a metric undergoing Ricci Flow tends to a metric of constant curvature. As previously remarked, because the Riemann tensor has only one component in 2 dimensions, the notions of constant Ricci scalar and constant curvature are equivalent and they are related by[29] $R = 2K$.

We now turn to the study of partial differential equations and their solutions, inspired by our eventual study of Ricci flow and the Einstein equations.

2.3 Partial Differential Equations

Many processes in physics—such as the flow of heat, Einstein gravitation, fluid flow, and countless others—are described using partial differential equations. Though their importance cannot be overstated, we will only skim the surface this section, giving only key definitions and examples.

Our discussions thus far have been quite geometric, dealing with generic smooth manifolds. During this section, however, we will often concentrate on patches of $\mathbb{R}^n$ for simplicity, which can be applied to n-manifolds by using charts appropriately.

Definition 2.3.1 (Elliptic, hyperbolic and parabolic partial differential equations)**.** A second-order partial differential equation in n variables $(x^i)_{i=1}^n$ for a function $u = u(x^1, \dots, x^n)$ can be (locally) written

$$\left(A^{ij}\partial_i\partial_j + \text{lower-order terms}\right) u = 0,$$

[29]In 2 dimensions, the curvature is equal to the **Gaussian curvature**, yet another notion of curvature.

where the matrix $A^{ij} = A^{ij}(x, u, \partial u)$ varies from point to point in the coordinate patch considered (and may be dependent on u itself, or its derivatives—see 'linearity' below), and the term in parentheses is often called the **differential operator**. We now have three special cases, differentiated by the eigenvalues of A^{ij}:

1. If all are non-zero and have the same sign, then the equation is **elliptic**;
2. If all are non-zero and one has the opposite sign from the rest, then the equation is **hyperbolic**;
3. If all have the same sign except one which is zero, then the equation is **parabolic**.

Note that in practice, this third case arises most often with the presence of a first-order derivative in one of the variables, which we frequently call t for time. Thus parabolic equations can be thought of as elliptic equations with one extra term: a first-order temporal derivative.

In the simple case of two variables $u = u(x, y)$ and coefficients A^{ij} independent of u and ∂u (written instead as A, B and C), we write

$$\left(A\partial_{xx} + 2B\partial_{xy} + C\partial_{yy} + \text{lower-order terms}\right) u = 0,$$

and the three above cases are the only possibilities, differentiated by their sign of the discriminant $B^2 - AC$: elliptic if negative, hyperbolic if positive and parabolic if zero.

Finally, recall that a partial differential equation for a function u is **linear** if all instances of u and derivatives of u appear as linear terms. It is **quasi-linear** if the highest-order derivatives of u appear linearly, and **non-linear** otherwise.

The descriptions above will be revisited in Section 3.5, when we prove short-time existence for Ricci flow. For now, we give canonical examples of these types of equations.

Example 2.3.1 (Poisson and Laplace equations)**.** The **Poisson equation**,

$$\Delta u = f,$$

is linear and elliptic, as is the special case when f vanishes, which is known as the **Laplace equation**. Solutions to the Laplace equation are called **harmonic functions**. The Poisson equation describes the potential distribution of a mass or charge, such as in the Newtonian gravitational equation

$$\Delta\phi = 4\pi\rho,$$

where ϕ is the **scalar gravitational potential** of a massive object with density ρ. The elliptic nature of this gravitational theory was an indication of its violation of relativity, since, as we will discuss shortly, the information at any point is felt immediately at any other point, without being slowed by a finite speed of propagation.

Example 2.3.2 (Wave equation)**.** The **wave equation**,

$$\partial_{tt} u = \Delta u, \tag{2.9}$$

is linear and hyperbolic. It describes the propagation of waves: waves along a string, sound waves, light waves, and more. Maxwell showed that the equations that bear his name are wave equations for waves produced by the electromagnetic field propagating at the speed of light, corresponding to the wave interpretation of light.

Example 2.3.3 (Heat equation). The **heat equation**,

$$\partial_t u = \Delta u, \tag{2.10}$$

is linear and parabolic. It describes the diffusion of a heat distribution throughout a domain, moving from areas of high concentration to lower concentrations (averaging the distribution), with a constant distribution being the limit at temporal infinity.

The usefulness of these simple examples lies in the fact that lots of qualitative understanding of a given equation can be grasped by comparison with the well-known behaviours of the above cases. For instance, the speed of propagation of data is infinite in the parabolic case, whereas it is finite in the hyperbolic case (the speed of the wave). In a hyperbolic equation, the smoothness (as in differentiability, or $\mathcal{C}^\infty$-smoothness) of the solution[30] depends on the smoothness of the initial data (and propagates by wave-speed), whereas solutions to elliptic equations are smooth, (almost) regardless of initial conditions.[31] Finally—and most crucially to our chapter on Ricci flow—parabolic equations tend to 'smoothen' (though 'average' might be a better term, since confusion with $\mathcal{C}^\infty$-smoothness or differentiability may occur) out the initial data, as heat diffusing to a constant temperature throughout a domain.

Elliptic and parabolic equations benefit from what are known as maximum and minimum principles, which use initial data to give bounds on the solution at other points in the domain. In essence, they state that the maximum and minimum of a function satisfying an elliptic or parabolic equation must occur on the boundary of its domain. Before presenting the only formulation of such a principle that we will use in this text, we make the following notational comments:

1. We resume x as the notation for a point to coincide with most literature and to emphasise that this takes place in (subsets of) $\mathbb{R}^n$—requiring charts to apply these results to a general manifold;

2. Contraction via the metric will often be written $\langle \cdot, \cdot \rangle$, with corresponding norm $|\cdot|$.[32]

Consider a partial differential equation for a function $u : \Omega \times [0, T] \to \mathbb{R}$ (for some $\Omega \subset \mathbb{R}^n$ open, connected and bounded, and $T > 0$) of the form

$$\partial_t u = \Delta u + F(u) + \langle \nabla u, V \rangle,$$

for a function $F(u)$ that we assume to be locally Lipschitz and $V_t = V$ a time-dependent family of vector fields. This is called a **reaction-diffusion equation**, where $F(u)$ is the **reaction term**, which works to prevent the parabolic (diffusion) operator $\Delta - \partial_t$ from 'averaging out' the function u. (It turns out that the **gradient term** $\langle \nabla u, V \rangle$ has no effect on the maximum or minimum principle.) With this reaction-diffusion equation in mind, we have the following maximum and minimum principles, which state that if u satisfies certain initial bounds and the reaction-diffusion equation, then it has (time-dependent) bounds that can be explicitly calculated with knowledge of $F(u)$.

[30] This is partial motivation for our impending review of Sobolev spaces in the pages to come.

[31] Note that this is barring very nasty circumstances.

[32] Given for a vector in the usual way via $|V|^2 = \langle V, V \rangle$. Extensions to tensors in general are straightforward.

Proposition 2.3.1 (Reaction-diffusion maximum and minimum principles)**.** *As above, consider the reaction-diffusion equation for $u : \Omega \times [0,T] \to \mathbb{R}$ (for some $T > 0$) given by*

$$\partial_t u = \Delta u + F(u) + \langle \nabla u, V \rangle, \tag{2.11}$$

where $F(u)$ some (locally Lipschitz) function and $V_t = V$ is a one-parameter family of vector fields. Then,

1. *Let $\alpha(t)$ solve the equation*

$$\partial_t \alpha = F(\alpha) \quad \text{with} \quad \alpha(0) = \alpha_0;$$

   if $u(x,t)$ satisfies[33]

$$\partial_t u \geq \Delta u + F(u) \quad \text{and} \quad u(x,0) \geq \alpha_0 \quad \text{for all} \quad x \in \Omega,$$

   then $u(x,t) \geq \alpha(t)$ for all $(x,t) \in \Omega \times [0,T]$;

2. *Let $\beta(t)$ solve the equation*

$$\partial_t \beta = F(\beta) \quad \text{with} \quad \beta(0) = \beta_0;$$

   if $u(x,t)$ satisfies

$$\partial_t u \leq \Delta u + F(u) \quad \text{and} \quad u(x,0) \leq \beta_0 \quad \text{for all} \quad x \in \Omega,$$

   then $u(x,t) \leq \beta(t)$ for all $(x,t) \in \Omega \times [0,T]$.

Combining the above gives that if for all $x \in \Omega$ we have

$$\partial_t u = \Delta u + F(u) + \langle \nabla u, V \rangle \quad \text{and} \quad \alpha_0 \leq u(x,0) \leq \beta_0,$$

then

$$\alpha(t) \leq u(x,t) \leq \beta(t) \quad \text{for all} \quad (x,t) \in \Omega \times [0,T].$$

Idea of the proof. (Since it has been proven in many places—see, for example, Theorem 4.4 of [CK04] or Theorem 3.2 of [She06]—the proof has been omitted.)

The idea is to look at

$$\partial_t (u - \alpha) \geq \Delta(u - \alpha) + \langle \nabla(u - \alpha), V \rangle + F(u) - F(\alpha),$$

and use the locally Lipschitz nature of F to take care of the reaction terms. Eventually, it can be reduced to the simpler case where F is linear, which in turn can be translated into a more standard parabolic maximum principle result where one modifies the function u by some small $\varepsilon > 0$ and inspects the first- and second-order derivatives. □

These principles will be vital tools in our study of Ricci flow. Various tensorial quantities will satisfy reaction-diffusion equations of the form (2.11), and we will use Proposition 2.3.1 to find time-dependent bounds for the evolving quantities.

One important topic within the study of partial differential equations is the space of possible solutions. What requirements must the solution functions satisfy to be reasonable solutions? In this direction, we briefly discuss what are called Sobolev spaces, which address exactly this notion. We first introduce multi-index notation, which will simplify the definitions to come.

[33] That is, $u(x,t)$ is not required to solve (**??**). This condition is less stringent.

Definition 2.3.2 (Multi-index notation)**.** Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a list of non-negative integers called a **multi-index** with **length** $|\alpha| := \sum_{i=1}^n \alpha_i$. For coordinates $(x^i)_{i=1}^n$ on an n-manifold $\mathcal{M}$, we write $D^\alpha := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$.

Now we introduce a particularly nice class of functions on our manifold, known as test functions.

Definition 2.3.3 (Test function)**.** For a manifold $\mathcal{M}$ with atlas $(U_i, \varphi_i)_i$, we let $\mathcal{C}_c^\infty(\mathcal{M})$ be the functions in $\mathcal{C}^\infty(\mathcal{M})$ with compact support. We write $\mathcal{D}(\mathcal{M})$ for the space of **test functions**, which is $\mathcal{D}(\mathcal{M}) := \mathcal{C}_c^\infty(\mathcal{M})$ as a set and is equipped with a topology given as follows. For positive integers k, we say that test functions $f_k \in \mathcal{D}(\mathcal{M})$ have limit f in $\mathcal{D}(\mathcal{M})$ if for every multi-index α, every chart (U_i, φ_i) and every $F_i \in \mathcal{C}_c^\infty(\varphi_i(U_i))$ we have[34]

$$\left\| D^\alpha \left(F_i \circ (f_k - f) \circ \varphi_i^{-1} \right) \right\|_\infty \to 0.$$

We now define distributions, which are objects that generalise functions.

Definition 2.3.4 (Distribution)**.** We write $\mathcal{D}'(\mathcal{M})$ for the space of **distributions**, which as a set is the dual of the space of test functions ($\mathcal{D}'(\mathcal{M}) := \mathcal{D}(\mathcal{M})^*$) and is equipped with the **weak topology**: distributions $u_k \in \mathcal{D}'(\mathcal{M})$ have limit $u \in \mathcal{D}'(\mathcal{M})$ if for every $f \in \mathcal{D}(\mathcal{M})$, we have $u_k(f) \to u(f)$. We will write $\langle u, f \rangle := u(f)$ to be consistent with standard conventions.[35]

We now give two simple examples of distributions.

Example 2.3.4 (Functions as distributions)**.** On $\mathbb{R}^n$, every locally integrable and measurable (though not necessarily smooth) function $\psi : \mathbb{R}^n \to \mathbb{R}$ defines a distribution $u_\psi \in \mathcal{D}'(\mathbb{R}^n)$ by (Lebesgue) integration:

$$\langle u_\psi, f \rangle := \int_{\mathbb{R}^n} \psi f \quad \text{for} \quad f \in \mathcal{D}(\mathbb{R}^n).$$

Example 2.3.5 (Dirac delta function)**.** The **Dirac delta function** δ_0 cannot be defined as a function, though it is a distribution: $\langle \delta_0, f \rangle := f(0)$, for $f \in \mathcal{D}(\mathbb{R}^n)$.

The above examples motivate the following definition of how to differentiate distributions.

Definition 2.3.5 (Weak derivative)**.** On $\mathbb{R}^n$, the **weak derivative** of a distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ is defined for a multi-index α and a test function $f \in \mathcal{D}(\mathbb{R}^n)$ by

$$\langle D^\alpha u, f \rangle := (-1)^{|\alpha|} \langle u, D^\alpha f \rangle.$$

This is inspired by partial integration, and indeed by observation, we have $D^\alpha u_\psi = u_{D^\alpha \psi}$ for functions $\psi : \mathbb{R}^n \to \mathbb{R}$.

[34]Recall that the $\mathscr{L}^\infty$**-norm** is the essential supremum:

$$\|\psi\|_\infty := \inf \left\{ C \geq 0 \,\middle|\, |\psi(x)| \leq C \text{ for almost every } x \right\}.$$

The space of functions on a manifold $\mathcal{M}$ where the above norm is finite is $\mathscr{L}^\infty(\mathcal{M})$. In our case, the supremum tending to zero should be sufficient, since our test functions are smooth.

[35]This should not be confused with contraction via the metric. We will only employ this notation until Sobolev spaces have been defined.

Finally, we can define our sought-after Sobolev spaces.

Definition 2.3.6 (Sobolev space)**.** For a manifold $\mathcal{M}$ and a positive integer s (called the **Sobolev parameter**), we define the **Sobolev space** $\mathscr{H}^s(\mathcal{M})$ as[36]

$$\mathscr{H}^s(\mathcal{M}) := \left\{ u \in \mathscr{L}^2(\mathcal{M}) \;\middle|\; \text{for all } \alpha \text{ with } |\alpha| \leq s, \text{ we have } D^\alpha u \in \mathscr{L}^2(\mathcal{M}) \right\},$$

where the previous discussion arises because we only require the derivatives $D^\alpha u$ to be weak.[37] These spaces become Hilbert spaces[38] when equipped with the inner product $\langle \cdot, \cdot \rangle_s$ given by

$$\langle u, v \rangle_s := \sum_{|\alpha| \leq s} \langle D^\alpha u, D^\alpha v \rangle_{\mathscr{L}^2(\mathcal{M})}.$$

Until now we have been assuming our tensors to be smooth—equivalent to $\mathcal{C}^\infty(\mathcal{M})$-linear maps. This is not necessary, however: in general they can be thought of as $\mathscr{H}^s(\mathcal{M})$-linear maps, which will be called $\mathscr{H}^s$-tensors, or simply tensors (with their Sobolev nature implicit).[39] Most importantly, our metrics no longer need to be required to be smooth but can instead be $\mathscr{H}^s$. For this train of thought to work, we could assume that $s > 3$ throughout, to be safe.

However, in practice, it will be much easier to consider everything to be $\mathcal{C}^\infty$-smooth. This tangent into Sobolev spaces would allow us to be much more precise in our discussions of both Ricci flow and the Einstein equations; nevertheless, other than a few mentions in the text to come, we will suppress the s from our notation and from our discourse and proceed blindly, hoping that everything behaves well under the assumption of smoothness. Typically, $\mathcal{C}^\infty$-smooth results are proved via existence in Sobolev spaces $\mathscr{H}^s$ for sufficiently large s and then sending $s \to \infty$. One of the final times we will include this s is in the following section, which defines several new infinite-dimensional spaces to join the ranks of our diffeomorphism groups, isometry groups, and the spaces defined in this past section on our list of function spaces.

2.4 Hilbert Manifolds

In the spirit of our brief mention of the space of solutions to a partial differential equation, we define the following generalisation to the definition of a manifold, which allows for the construction to be infinite-dimensional.

---

[36]Recall that for $1 \leq p < \infty$ the $\mathscr{L}^p$**-norm** is given on a semi-Riemannian manifold $\mathcal{M}$ by

$$\|\psi\|_p := \left(\int_{\mathcal{M}} |\psi|^p \right)^{\frac{1}{p}}.$$

The space of functions where the above norm is finite is $\mathscr{L}^p(\mathcal{M})$.

[37]Sobolev spaces are often defined more generally as depending on two parameters, one of which acts as our s and the other acts as the p in $\mathscr{L}^p$, which we have taken to be 2.

[38]Recall that a **Hilbert space** is a metric space with inner product that is complete with respect to the inner product. Note that $\mathscr{L}^2(\mathcal{M})$ is a Hilbert space with inner product $\langle \cdot, \cdot \rangle_2$ given by

$$\langle \psi, \phi \rangle_2 := \int_{\mathcal{M}} \psi\phi.$$

[39]We note that for our tensor operations defined in Section 2.1, we want the product of two tensors to once again be a tensor. Thankfully, a consequence of the **Sobolev Embedding Theorem** is that for an n-manifold $\mathcal{M}$, $\mathscr{H}^s(\mathcal{M})$ is algebra—that is, it is closed under multiplication—when $s > \frac{n}{2}$.

Definition 2.4.1 (Hilbert manifold)**.** A **Hilbert manifold** is a paracompact Hausdorff space such that every point has a neighbourhood that is homeomorphic to an open set in a (possibly infinite-dimensional) Hilbert space. Such a manifold can be considered differentiable by defining an atlas and differentiable transition functions whose values are taken in the Hilbert space.

Further, we can define a **Hilbert submanifold**[40] as a Hilbert manifold with a map to another Hilbert manifold that is a homeomorphism onto its image and whose push-forward is injective. We can also equip a Hilbert manifold with its inner product on each tangent space (which are each canonically isomorphic to the Hilbert space itself[41]) acting as a Riemannian metric, giving a **Hilbert-Riemannian manifold**.[42]

The precise definitions of these spaces are not crucial for our discussions, however. We simply want to think of an infinite-dimensional function space as a manifold. For details, see the introductory chapters of [Tro92].

We now state important examples of Hilbert manifolds, the second of which is crucial in our discussions to come. For reasons which will become apparent in our study of General Relativity,[43] we now deal with Riemannian (not semi-Riemannian) manifolds for the rest of this section, and we preserve the Sobolev parameter s, not requiring $\mathcal{C}^\infty$-smoothness.

Example 2.4.1 (Trivial Hilbert manifold)**.** Any Hilbert space is a Hilbert manifold whose global chart is the identity.

Example 2.4.2 (Space of metrics)**.** For a Riemannian manifold $(\mathcal{M}, g)$, we define the **space of metrics** $\mathscr{M}^s$ to be the Hilbert manifold whose points are $\mathscr{H}^s$-Riemannian metrics. As a sub-example we let

$$\mathscr{M}^s_\lambda := \{g \in \mathscr{M}^s \text{ whose Ricci scalar is constant and equal to } \lambda \in \mathbb{R}\}.$$

Because almost all of our discussion involves smooth metrics, we make the note here that $\mathscr{M}$ and $\mathscr{M}_\lambda$ are defined in an identical fashion to the spaces above but instead requiring metrics to be smooth;[44] it turns out that they are submanifolds of $\mathscr{M}^s$ and $\mathscr{M}^s_\lambda$, where their smooth structure can be established by making them **Inverse Limit Hilbert** (ILH) **manifolds**, which are not quite Hilbert manifolds.[45]

Further, these Hilbert manifolds can be equipped with the **$\mathscr{L}^2$-metric** $G^{\mathscr{L}^2}$, which is defined as follows:

1. Consider two elements ξ and Ξ of $\mathcal{T}_g\mathscr{M}^s$, the tangent space of $\mathscr{M}^s$ at a metric g;

2. Since $\mathcal{T}_g\mathscr{M}^s$ is isomorphic (in the way mentioned in Definition 2.4.1) to the space of **symmetric rank-**2 **$\mathscr{H}^s$-tensors** on $\mathcal{M}$, written $\mathscr{S}^s_2$, we can think of ξ and Ξ as elements of $\mathscr{S}^s_2$;[46]

[40] Though we have not yet stated the definition of a submanifold—they will be the principal characters of Section 4.3—this definition of a Hilbert submanifold is a logical translation of it.

[41] Recall that any vector space $\mathcal{V}$ is canonically isomorphic to its tangent space $\mathcal{T}\mathcal{V}$.

[42] One can continue this line of thought even further: a **Hilbert-Lie group** is a smooth Hilbert manifold whose group composition and inversion operations are smooth, and so on.

[43] This will be because our goal is to split spacetime into space and time, and we will use the following ideas on the 'space' part alone.

[44] These can also be seen as $\mathscr{M} := \cap_s \mathscr{M}^s$.

[45] See [Omo70], [Koi78] and [Koi79] for fundamental work in the ILH procedure.

[46] Note that by definition $\mathscr{M}^s$ is a Hilbert submanifold of $\mathscr{S}^s_2$ since it is an open subspace (a cone).

3. We then have an inner product on $\mathscr{S}_2^s$ given in local coordinates (at g) by contraction with the metric g: $\langle \xi, \Xi \rangle := g^{ij} g^{kl} \xi_{ik} \Xi_{jl}$, which is all evaluated at some $p \in \mathcal{M}$, of course;

4. Now, we define the $\mathscr{L}^2$-metric as

$$G_g^{\mathscr{L}^2}(\xi, \Xi) := \int_{\mathcal{M}} \langle \xi, \Xi \rangle \, .$$

The previous example treats the largest possible solution space to any partial differential equation for a Riemannian metric. In the case of Ricci flow, the solution is a time-dependent family of metrics $g(t)$. Thus if Ricci flow exists on some interval $[0, T]$ (for some $T > 0$), the solution is a curve $g : [0, T] \to \mathscr{M}$ on the Hilbert manifold $\mathscr{M}$. In the case of the Einstein equations (once we have made the desired 'space' and 'time' split of spacetime), the solutions form some subset of $\mathscr{M}$, denoting the possible Riemannian metrics on the spatial part of spacetime. We will study this subset in great detail in Section 4.4. We will discover that under certain circumstances it can be reduced to be finite-dimensional, rather than infinite-dimensional, as it is at first glance.

We close this introductory chapter with a discussion of the Uniformisation Theorem.

2.5 Uniformisation Theorem

The Uniformisation Theorem dates back to the nineteenth century, having been conjectured by Poincaré and Klein in 1882 [Poi82] and 1883 [Kle83], respectively. We will state two forms of the Theorem, first formulated in complex geometry, then in the language of Riemannian manifolds. We first define a 2-dimensional complex manifold, called a Riemann surface.

Definition 2.5.1 (Riemann surface)**.** A smooth oriented 2-manifold $\mathcal{M}$ whose local trivialisations $(\varphi_\alpha)_\alpha$ take values in $\mathbb{C}$ and whose transition functions are holomorphic[47] is called a **Riemann surface**. The atlas in this case is called a **complex structure** c, where

$$c := (U_\alpha, \varphi_\alpha)_\alpha \quad \text{for open} \quad U_\alpha \subset \mathcal{M} \quad \text{with} \quad \varphi_\alpha : U_\alpha \to \mathbb{C}.$$

The set of all complex structures on $\mathcal{M}$ is denoted $\mathscr{C}(\mathcal{M})$ (or simply $\mathscr{C}$), and a Riemann surface is often denoted $(\mathcal{M}, c)$.

Two Riemann surfaces $\mathcal{M}$ and $\mathcal{N}$ are **biholomorphic** if there exists a map $\psi : \mathcal{M} \to \mathcal{N}$ such that for all trivialisations $(\varphi_\alpha)_\alpha$ and $(\phi_\beta)_\beta$ on $\mathcal{M}$ and $\mathcal{N}$, respectively, the function $\phi_\beta \circ \psi \circ \varphi_\alpha^{-1}$ is holomorphic. Two Riemann surfaces are **conformal** if there exists a holomorphic map between them with non-vanishing derivative (the function is then also called **conformal**).

The Uniformisation Theorem gives a conformal classification of Riemann surfaces.

Theorem 2.1 (Uniformisation Theorem)**.** *Every simply connected Riemann surface is conformal to the complex plane* $\mathbb{C}$*, the Riemann sphere*[48] $\mathbb{S}$*, or the upper half-plane* $\mathbb{H}$*.*

We would like to translate this result into one that classifies Riemannian 2-manifolds. We now define an intermediary step between the real and complex cases.

[47] Recall that a complex-valued map is **holomorphic** if it is equal to a convergent power series; equivalently, if it is complex-differentiable at every point.

[48] Think of this as $\mathbb{C} \cup \{\infty\}$, the extended complex plane.

Definition 2.5.2 (Almost complex manifold)**.** On an n-manifold $\mathcal{M}$, an **almost complex structure** is a $(1, 1)$-tensor J such that for all $p \in \mathcal{M}$,

1. J squares to the identity: $J_p^2 = -\mathbb{1}_p$;
2. $(V_p, J_p V_p)$ is an oriented basis for $\mathcal{T}_p\mathcal{M}$ for all $V_p \in \mathcal{T}_p\mathcal{M}$.

(The second requirement is analogous to the choice of $\pm$i for the imaginary unit on $\mathbb{C}$.)

A manifold $\mathcal{M}$ with an almost complex structure J is called an **almost complex manifold** and is written $(\mathcal{M}, J)$. The space of all almost complex structures on $\mathcal{M}$ is written $\mathscr{A}(\mathcal{M})$. Though we will not prove this, $\mathscr{A}$ can be thought of as a Hilbert manifold in the same way as $\mathscr{M}$.[49]

We give a simple example that will help us to connect the space of complex structures and the space of almost complex structures.

Example 2.5.1 (On $\mathbb{R}^2$)**.** On $\mathbb{R}^2$, the simplest almost complex structure is[50]

$$\hat{J} := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

This allows us to transition between almost complex structures on $\mathbb{R}^2$ to complex structures on $\mathbb{C}$—or on patches of either, and thus on manifolds. To see this, define for any chart (U, φ) in a complex structure c the map

$$J_\varphi := (\varphi_*)_p^{-1} \circ \hat{J} \circ (\varphi_*)_p. \tag{2.12}$$

This defines an almost complex structure, as can be readily checked.

The above is also independent of the chart chosen, which leads to the following result, which describes a bijection (in 2 dimensions) between the space of complex structures $\mathscr{C}$ and the space of almost complex structures $\mathscr{A}$.

Proposition 2.5.1. *In two dimensions, there exists a bijection $\varpi : \mathscr{C} \to \mathscr{A}$ given locally by the expression for J_φ in* (2.12).[51]

Proof. See Theorem 1.1.1 of [Tro92] for details. ☐

Since our case of interest $(\mathcal{M}, g)$ is 2-dimensional and real, we can use $\hat{J}$ to connect almost complex structures to Riemannian 2-manifolds as well. We argue as follows: a Riemannian metric defines angles on the manifold, and at each point $p \in \mathcal{M}$ the tangent space $\mathcal{T}_p\mathcal{M}$ is 2-dimensional. Thus the metric induces an almost complex structure J by demanding that J rotate vectors in $\mathcal{T}_p\mathcal{M}$ counterclockwise by $\frac{\pi}{2}$.[52]

---

[49] Where once more we have omitted the Sobolev parameter s and are assuming the presence of a smooth structure in the ILH manner described in Section 2.4.

[50] This is a counterclockwise rotation by $\frac{\pi}{2}$.

[51] Note that the correspondence of almost complex structures to complex structures is only bijective in two dimensions. See [FT84] or Chapter IX.2 of Volume II of [KN63] for details. (Plus, we only defined complex structures in 2-dimensions.)

[52] We can go even further and define a **symplectic form** ω to be a closed non-degenerate 2-form. When a manifold $\mathcal{M}$ is equipped with a symplectic form, called its **symplectic structure**, it is called a **symplectic manifold** and is written $(\mathcal{M}, \omega)$. Now, a Riemannian manifold $(\mathcal{M}, g)$ equipped with an almost complex structure J (which we assume to be integrable—that is, in bijection with a complex structure; in 2 dimensions, this is always the case by Proposition 2.5.1)—sometimes called a **Hermitian manifold**—induces a symplectic structure via $\omega(\cdot, \cdot) := g(\cdot, J(\cdot))$, where the **Hermitian form** is $h := g - \mathrm{i}\omega$. This ω is then called the **Kähler structure**, so the manifold $(\mathcal{M}, g, J, \omega)$ is called **Kähler**, where any two of the three structures define the third (as $g(\cdot, \cdot) = \omega(J(\cdot), \cdot)$, etc.). The tensors g and ω are called **compatible** with J, as $g(\cdot, \cdot) = g(J(\cdot), J(\cdot))$ and $\omega(\cdot, \cdot) = \omega(J(\cdot), J(\cdot))$.

We now want to describe conformal equivalence in the language of Riemannian manifolds.

Definition 2.5.3 (Conformal)**.** Two metrics g and $\hat{g}$ on a manifold $\mathcal{M}$ are **conformal**, or conformally equivalent, if there exists a function $\lambda \in \mathcal{C}^\infty(\mathcal{M})$ such that $\hat{g} = \mathrm{e}^{-2\lambda} g$ everywhere on $\mathcal{M}$.[53]

This definition can be expanded to include the **conformal action** of the $\mathcal{C}^\infty$-smooth functions on the space of metrics $\mathscr{M}$:[54]

$$\mathscr{M} \times \mathcal{C}^\infty \to \mathscr{M}, \quad (g, \lambda) \mapsto \mathrm{e}^{-2\lambda} g. \tag{2.13}$$

This action describes the split of $\mathscr{M}$ into conformal equivalence classes.

Because we do not allow the conformal transformation to change sign, we can think of conformally equivalent metrics as simply having a different scale (albeit with the scaling factor changing from point to point). It allows us to partition Riemannian manifolds into conformal equivalence classes. With all of our tools in hand, we have the following reformulation of the Uniformisation Theorem.

Theorem 2.2 (Uniformisation Theorem, #2)**.** *Every (complete) Riemannian metric on a closed 2-manifold is conformal to a (complete) metric with constant curvature.*

One may ask where the simply-connected requirement went. In Riemannian geometry there is the following result, valid in any dimension, which classifies all spaces of constant curvature.

Theorem 2.3. *For $n \geq 2$, every simply connected Riemannian n-manifold $(\mathcal{M}, g)$ with constant (sectional) curvature and complete metric is isometric to the n-sphere S^n, flat Euclidean space $\mathbb{R}^n$, or the upper half-plane $\mathbb{H}^n$.*

This will classify all universal covering spaces[55] of Riemannian 2-manifolds, which can then be constructed by taking a quotient of the universal cover by a suitable discrete subgroup of the isometry group.

Using the notation from Section 2.4, we can state this second formulation of the Uniformisation Theorem symbolically:

$$\mathscr{M}/\mathcal{C}^\infty \cong \mathscr{M}_\lambda, \quad \text{where} \quad \lambda \in \{1, 0, -1\},$$

where the quotient is via the conformal action (2.13).

This result will come into play as a link between the studies of 2-dimensional Ricci flow and $(2+1)$-dimensional General Relativity, the topics of the two following chapters. We will show that Ricci flow can be used to prove the Uniformisation Theorem, and it will play a significant role in our inspection of General Relativity, by dividing possible universal manifolds into three case studies.

[53]This is equivalent, of course, to having a function $\lambda' \in \mathcal{C}^\infty(\mathcal{M})$ such that $\hat{g} = \mathrm{e}^{2\lambda'} g$. We will use both versions.

[54]Sometimes this equivalently written as a point-wise multiplication action on $\mathscr{M}$ by positive smooth functions. We also note that it is possible to include the Sobolev parameter s and consider the conformal action of $\mathscr{H}^s$ on $\mathscr{M}^s$ defined analogously.

[55]Recall that the **universal covering space** of a topological space is one that contains all connected covering spaces of the space, and in particular is simply connected.

Chapter 3

Ricci Flow

In 1982, Hamilton (see [Ham82]) set forth a programme whose primary goal was to resolve the Poincaré Conjecture.[1] His project centred around a (system of) partial differential equation(s) consisting of varying the Riemannian metric of a manifold $\mathcal{M}$ by some parameter—often taken to be t and representing time—in the following way:

$$\partial_t g = -2\mathrm{Ric}.$$

This equation is called the **Ricci flow equation**, or Hamilton's Ricci flow, and has solution a one-parameter family of Riemannian metrics $g(t)$. (We will refer to it most as 'Ricci flow,' with no article.) Note that to become a well-posed problem, initial data must be provided: here, and throughout the following sections, we (often implicitly) assume the requirement that the initial metric $g(0) = g_0$ be given. The idea behind Ricci flow is to 'smoothen' a metric (in the sense of curvature being constant, rather than in the $\mathcal{C}^\infty$ sense): we will see that under certain conditions, a metric undergoing Ricci flow will tend towards a metric of constant curvature.[2]

Recall from the previous chapter our definition of $\mathscr{M}$ as the space of all possible smooth Riemannian metrics on a manifold $\mathcal{M}$. The solution $g(t)$ (which we assume to exist for some interval $[0, T]$ for $T > 0$) to Ricci flow with initial metric g_0 can then be visualised as a path on $\mathscr{M}$: it is a smooth curve $g : [0, T] \to \mathscr{M}$ that begins at the point $g(0) = g_0$. Tending to a metric of constant curvature would be if this curve ended on (or tended to, as $T \to \infty$) some point $g_\infty \in \mathscr{M}$, which is a metric of constant curvature. Figure 3.1 depicts this set up as if $\mathscr{M}$ was finite-dimensional.

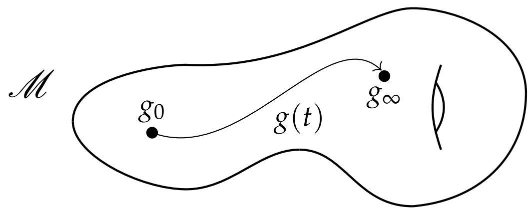

Figure 3.1: Visualisation of a solution $g(t)$ to Ricci flow with initial metric $g(0) = g_0$ on the space of metrics $\mathscr{M}$.

[1] This has since been proven and so can be stated as a Theorem.

Theorem 3.1 (Poincaré Conjecture)**.** *Every closed, simply-connected* 3*-manifold is homeomorphic to the* 3*-sphere.*

[2] The rough idea to prove the Poincaré Conjecture was then to argue that if the resultant metric was compact and of constant curvature, then it could only be (homeomorphic to) one thing: the 3-sphere!

Note that Ricci flow is invariant under diffeomorphism: if ψ is a diffeomorphism of $\mathcal{M}$ and $g(t)$ is a solution to Ricci flow with initial metric g_0, then $\psi^* g(t)$ is a solution to Ricci flow with initial metric $\psi^* g_0$. This arises from the Ricci tensor being **diffeomorphism-equivariant**:

$$\mathrm{Ric}(\psi^* g) = \psi^* \mathrm{Ric}(g).$$

This can be loosely argued by considering the diffeomorphism to be a change of chart (that is, consider $\psi^* g$ to be the metric in the new chart), and noting that since the Ricci tensor is a trace of another tensor (the Riemann tensor), it should behave in this way under a coordinate change.[3] This equivariance will also appear in General Relativity, as the Einstein equations will also be diffeomorphism-equivariant.

One final comment that will be discussed in greater detail in Section 3.5 is that Ricci flow can be loosely thought to be a parabolic equation, in the same vein as the heat equation (2.10). Though Ricci flow is not linear (and not strongly parabolic, either), this characterisation of it as a heat-type equation explains the similarity in its 'smoothening' behaviour towards the metric, just as a heat distribution averages out over time.

Overview. Section 3.1 introduces Ricci flow and the behaviour of some tensorial quantities under it, as well as defines gradient Ricci solitons. Section 3.2 applies the maximum principle to the Ricci scalar. Section 3.3 discusses mean curvature flow and compares its qualitative behaviour to that of Ricci flow. Section 3.4 addresses the (short- and long-time) existence of Ricci flow. Section 3.5 summarises the plan for proving the Uniformisation Theorem. Sections 3.6 and 3.7 prove the Uniformisation Theorem for the case of an average Ricci scalar $r \leq 0$. Sections 3.8 through 3.13 treat the case $r > 0$.

Reference Guide. A great introduction to Ricci flow is [CK04], and it forms a backbone to this chapter (particularly its Chapter 5). Lecture notes [Top06] (which is now a book) and [Lan19] as well as the thesis [She06] introduce the material clearly, too. The collection of papers on Ricci flow [CCY03] contains reams of information on the subject. The paper [CLT06] contains the final piece of the proof of the Uniformisation Theorem via Ricci flow (see Proposition 3.9.1 in this text).

3.1 Formulations of Ricci Flow

Before simplifying our discussion to our 2-dimensional case of interest, we state several examples of closely-related flow equations that are considered in the general n-dimensional case of a Riemannian manifold $(\mathcal{M}, g)$, where the time-dependence of the metric and associated tensors is withheld for notational simplicity.

Example 3.1.1 (the normalised Ricci flow)**.** We define the **average Ricci scalar** r by

$$r := \frac{1}{\mathrm{vol}(\mathcal{M}, g)} \int_{\mathcal{M}} R, \tag{3.1}$$

[3] One can show that diffeomorphisms $\psi : \mathcal{M} \to \mathcal{M}$ induce maps $*\psi^{(k,l)} : \mathcal{T}^{(k,l)}\mathcal{M} \to \mathcal{T}^{(k,l)}\mathcal{M}$, defined at each $p \in \mathcal{M}$ by

$$*\psi_p^{(k,l)}(\tau_p) = \tau^{\omega_1 \cdots \omega_l}_{\zeta_1 \cdots \zeta_k}\big|_p \left(\psi_p^{-1}\right)^* (\mathrm{d}x^{\zeta_1}) \otimes \cdots \otimes \left(\psi_p^{-1}\right)^* (\mathrm{d}x^{\zeta_k}) \otimes \left(\psi_p\right)_* (\partial_{\omega_1}) \otimes \cdots \otimes \left(\psi_p\right)_* (\partial_{\omega_l}),$$

before carrying this definition through the definitions of the Levi-Civita connection and the Riemann tensor and concluding. See Chapter 5.4 of [Lan18] for details.

where we have defined integration over Riemannian manifolds in Definition 2.1.7. This then allows us to write the **the normalised Ricci flow** as[4]

$$\partial_t g = -2\mathrm{Ric} + \frac{2}{n} r g.$$

Consider the functions

$$\phi(t) := \exp\left(\frac{2}{n}\int_0^t r(t')\mathrm{d}t'\right) \quad \text{and} \quad \tau(t) := \int_0^t \phi(t')\mathrm{d}t'.$$

If $g(t)$ is a solution to Ricci flow, then the metrics $\bar{g}(\tau) := \phi(t)g(t)$ solve the normalised Ricci flow, as can be verified by direct computation.[5] In this manner we can move between solutions of the two flows.

Example 3.1.2 (Ricci-DeTurck flow)**.** The **Ricci-DeTurck flow** has the following form:

$$\partial_t g = -2\mathrm{Ric} + \mathcal{L}_\xi g, \tag{3.2}$$

for some time-dependent vector field $\xi_t = \xi$ called the **DeTurck vector**. This flow will appear in our later discussions, in the special case where the DeTurck vector is the gradient of a scalar function.

We will see that solutions to the Ricci-DeTurck differ from solutions to Ricci flow by pulling-back by time-dependent diffeomorphisms. That is, if $g(t)$ solves the Ricci-DeTurck flow for some DeTurck vector ξ_t and ψ_t is a one-parameter family of diffeomorphisms generated by ξ_t, then the metrics $\bar{g}(t) := \psi_t^* g(t)$ solve Ricci flow.

The importance of the Ricci-DeTurck flow will become apparent in Section 3.5, when we use the fact that it is strongly parabolic—whereas the usual Ricci flow is what is called weakly parabolic—to prove that this modified flow exists for short times. The relation of solutions between the flows (via pulling-back by diffeomorphisms) mentioned above will then allow us to conclude the short-time existence result for the standard Ricci flow.

Now, we will consider 2-dimensional Riemannian manifolds (Σ, g). Using the simplified form of the Ricci tensor in 2 dimensions arising from (2.7) and adding a constant factor of the metric for reasons that will become clear shortly, we have the following formulation of our central definition.

Definition 3.1.1 (2-dimensional Ricci flow)**.** The **Ricci flow** of a closed connected Riemannian 2-manifold (Σ, g) is given for some parameter t by

$$\partial_t g = (\rho - R)g,$$

where $\rho \in \mathbb{R}$ is constant.[6]

We will also revert to index notation frequently in the upcoming sections, and our most frequent usage of Ricci flow will be in the following normalised form:

$$\partial_t g_{ij} = (r - R)g_{ij}, \tag{3.3}$$

where r is the average Ricci scalar.

[4]As is implied by its name, this has the property that the volume of the manifold remains constant throughout the flow, as we shall soon see.

[5]The only trick to this is that $\mathrm{Ric}(\bar{g}) = \mathrm{Ric}(g)$, which arises by conformal invariance of the Ricci tensor (multiplying the metric by a positive factor leaves the Ricci tensor unchanged). We will see this again next chapter when we encounter the Einstein equations.

[6]By constant, we mean that it should be constant in space, and that under this equation, it should remain constant in time. Also, the case of vanishing ρ is sometimes called the **Yamabe flow**, which does not coincide with Ricci flow in higher dimensions.

In 2 dimensions, we have the following fundamental geometric result, which we state without proof.

Theorem 3.2 (Gauss-Bonnet)**.** *For $\chi(\Sigma)$ the Euler characteristic[7] of a closed 2-dimensional Riemannian manifold (Σ, g) with Ricci scalar R, we have*

$$\int_\Sigma R = 4\pi\chi(\Sigma).$$

Recall that the definition of the average Ricci scalar r is

$$r := \frac{1}{\mathrm{vol}(\Sigma, g)} \int_\Sigma R,$$

from which it is clear that r is constant in space. Since the Euler characteristic of a manifold undergoing the normalised Ricci flow is unchanging, Theorem 3.2 gives that the time dependence of r is only (inversely) related to the time dependence of the volume of the manifold. Thus once we establish that the normalised Ricci flow preserves volume, r will be constant in time and will thus satisfy our desired qualities for a chosen ρ.

Next, we have the following result, valid in any dimension.

Proposition 3.1.1. *Let h_{ij} be a symmetric $(2,0)$-tensor. For a Riemannian metric g_{ij} on an n-manifold $\mathcal{M}$ evolving as*

$$\partial_t g_{ij} = h_{ij},$$

we have the following:

1. *The inverse metric evolves as*

$$\partial_t g^{ij} = -h^{ij};$$

2. *The Christoffel symbols evolve as*

$$\partial_t \Gamma^k_{ij} = \frac{1}{2} g^{km} \left(\nabla_i h_{jm} + \nabla_j h_{im} - \nabla_m h_{ij} \right); \tag{3.4}$$

3. *The Laplacian (whose explicit dependence on $g(t)$ we show only here) evolves as*

$$\partial_t \Delta_{g(t)} = -h_{ij} \nabla^i \nabla^j - \left(\nabla^i h_{ij} - \frac{1}{2} \nabla_j h \right) \nabla^j,$$

where h is the trace of h_{ij};

4. *The Ricci scalar evolves as*

$$\partial_t R = -\Delta h + \nabla^i \nabla^j h_{ij} - h^{ij} R_{ij};$$

[7] Recall that the **Euler characteristic** χ of a 2-dimensional topological space is the homotopy invariant defined by

$$\chi := V - E + F,$$

where V, E and F are the number of 0-cells (vertices), 1-cells (edges) and 2-cells (faces), respectively, for any CW complex homotopic to the space.

5. *The volume element evolves as*

$$\partial_t(\mathrm{d}A) = \frac{1}{2}h\mathrm{d}A,$$

though we note that we withhold this from our notation unless absolutely necessary;

6. *The volume evolves as*

$$\partial_t \mathrm{vol}(\mathcal{M}, g) = \frac{1}{2}\int_{\mathcal{M}} h. \tag{3.5}$$

The final point implies that the normalised Ricci flow indeed preserves volume, as replacing

$$h_{ij} = -2R_{ij} + \frac{2}{n} r g_{ij}$$

makes the right-hand side of (3.5) *vanish.*

Idea of the proof. (For details, see Lemmas 3.1-3.3 and 3.9 of [CK04].)

These can be proven by direct computation. The most interesting point is for the Levi-Civita connection: though Christoffel symbols are not tensors, the difference between between Christoffel symbols is tensorial and thus by the definition of the derivative, $\partial_t \Gamma^k_{ij}$ is too. By proving the equality (3.4) in normal coordinates, we can conclude that it holds in general by coordinate-independence. □

There are also similar evolution equations for the Riemann and Ricci tensors, though because of the property of 2 dimensions that all curvature information is contained within the Ricci scalar R, the above points are sufficient for our discussion.[8] Using $h_{ij} = (\rho - R)g_{ij}$ and returning to our notation of Σ being a 2-manifold, we have

$$\partial_t g^{ij} = (R-\rho)g^{ij}; \tag{3.6}$$

$$\partial_t \Gamma^k_{ij} = \frac{1}{2}\left(-\delta^k_j \nabla_i R - \delta^k_i \nabla_j R + g_{ij}\nabla^k R\right); \tag{3.7}$$

$$\partial_t \Delta = (R-\rho)\Delta; \tag{3.8}$$

$$\partial_t R = \Delta R + R(R-\rho); \tag{3.9}$$

$$\partial_t(\mathrm{d}A) = (\rho - R)\mathrm{d}A; \tag{3.10}$$

$$\partial_t \mathrm{vol}(\Sigma, g(t)) = \rho \mathrm{vol}(\Sigma, g(t)) - 4\pi\chi(\Sigma), \tag{3.11}$$

where the final point uses the Gauss-Bonnet Theorem 3.2. Using the results above, we have the following evolution equation for the average Ricci scalar:

$$\partial_t r = \partial_t \left(\frac{1}{\mathrm{vol}(\Sigma, g(t))}\int_\Sigma R\right) = -\frac{1}{\mathrm{vol}(\Sigma, g(t))^2}\left(\int_\Sigma R\right)(\rho\mathrm{vol}(\Sigma, g(t)) - 4\pi\chi(\Sigma)) = r^2 - r\rho,$$

which vanishes when the normalised Ricci flow (where $\rho = r$) is considered, as expected.

We now give three important examples of solutions to Ricci flow.

[8]Furthermore, under the assumptions of Proposition 3.1.1, we have

$$\partial_t \int_\Sigma R = \int_\Sigma h^{ij}\left(\frac{1}{2}Rg_{ij} - R_{ij}\right).$$

This vanishes for $n = 2$ (as expected by the Gauss-Bonnet Theorem 3.2), and motivates the future result that the (vacuum) Einstein equations arise from the varying of the Einstein-Hilbert action, as will be seen in Section 4.1.

Example 3.1.3 (Einstein metric)**.** Suppose (Σ, g_0) is an Einstein manifold with Ricci tensor $\mathrm{Ric}_0 = 2\Lambda g_0$, and consider the standard (vanishing ρ) case of Ricci flow with initial metric $g(0) = g_0$. The solution metric is

$$g(t) = (1 - 4\Lambda t)g_0,$$

which is valid for $t \in [0, \frac{1}{4\Lambda})$. This can be checked by differentiating the above metric with respect to time and using the result that under a constant (over the manifold) conformal transformation, the Ricci tensor is unchanged: $\mathrm{Ric}(\mathrm{e}^{-2\lambda}g) = \mathrm{Ric}(g)$ when $\lambda \in \mathbb{R}$ constant—this is why there is the restriction on the time of validity. Once the equality of the two sides of the Ricci flow equation is established, we foresee the uniqueness of the solution[9] to conclude that it is the only solution.

In 2 dimensions, the assumption of the metric being Einstein is the same as assuming the initial Ricci scalar to be the constant $R_0 = 4\Lambda$. If we consider the normalised Ricci flow in 2 dimensions (3.3), then the metric is unchanging.

On the other hand, Einstein metrics are interesting because they are fixed points of the normalised Ricci flow: if $\partial_t g = 0$, then

$$\mathrm{Ric} = \frac{1}{n} r g.$$

In 2 dimensions, this merely states that the Ricci scalar is equal to its average over the manifold and thus is constant throughout, which we already knew was a consequence of the metric being Einstein.

Example 3.1.4 (Round metric)**.** The round metric on a 2-sphere with radius $\varrho > 0$ and its corresponding Ricci scalar are given by

$$g^{\mathrm{round},\varrho} = \varrho^2 g^{\mathrm{round}} \quad \text{and} \quad R = \frac{2}{\varrho^2},$$

where $g^{\mathrm{round}} := \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2$ is the round metric on the unit sphere. If we write ϱ_0 for the initial radius and allow the radius to change in time (writing $\varrho(t)$), standard Ricci flow (with $\rho = 0$) becomes the following simple ordinary differential equation:

$$\partial_t \left(\varrho(t)^2 g^{\mathrm{round}} \right) = - \left(\frac{2}{\varrho(t)^2} \right) \left(\varrho(t)^2 g^{\mathrm{round}} \right).$$

Cancelling factors and solving for $\varrho(t)$, we find the solution

$$\varrho(t) = \sqrt{\varrho_0^2 - 2t},$$

which demonstrates that the manifold will shrink to a point in finite time as $t \to \frac{1}{2}\varrho_0^2$.[10] In the case of the normalised Ricci flow, we have $R = r$ and therefore the metric does not change in time.

[9]This will be discussed in Section 3.5.

[10]Note that this same procedure works for the n-sphere with n-dimensional Ricci flow, obtaining

$$\varrho(t) = \sqrt{\varrho_0^2 - 2(n-1)t},$$

which vanishes as $t \to \varrho_0^2/(2(n-1))$.

Example 3.1.5 (Hyperbolic half-plane)**.** The same procedure as the example above can be completed with the hyperbolic metric on $\mathbb{H}^2$ to find the equation[11]

$$\partial_t \varrho(t)^2 = 2 \quad \implies \quad \varrho(t) = \sqrt{\varrho_0^2 + 2t},$$

which blows up to infinite size as $t \to \infty$.

These two last examples motivate the following definition.

Definition 3.1.2 (Self-similar solution)**.** A solution $g(t)$ to Ricci flow (with $\rho = 0$) with initial metric g_0 is called **self similar** if it can be written

$$g(t) = \psi_t^* g_0,$$

where ψ_t is a time-dependent family of conformal diffeomorphisms (where a diffeomorphism is **conformal** if it keeps metric within its conformal equivalence class when pulling-back) with the property $\psi_0 = \mathbb{1}$. As such, a self-similar solution evolves only via diffeomorphism and conformal re-scaling.[12]

A related definition is the following.

Definition 3.1.3 (Ricci soliton)**.** Consider a Riemannian manifold (Σ, g) with Ricci scalar R and average Ricci scalar r. We call g a **Ricci soliton** if it satisfies

$$(r - R)\, g_{ij} = \nabla_i V_j + \nabla_j V_i, \tag{3.12}$$

for V a vector field.[13]

Though this is defined without mention of Ricci flow, we will often consider all elements of (3.12) to be time-dependent, with the metric g undergoing the normalised Ricci flow. In this case, the solution $g(t)$ is called a **Ricci soliton solution**. By observation, this amounts to demanding that the metric change as

$$\partial_t g = \mathcal{L}_V g,$$

where $V_t = V$ is time-dependent.

Furthermore, if V is the negative gradient of some smooth time-dependent scalar function f (that is, if $V_i = -\nabla_i f$), then g is called a **gradient Ricci soliton**. In this case, the defining equation becomes

$$(R - r) g_{ij} = 2\nabla_i \nabla_j f.$$

The function f appearing in the previous equation deserves its own definition.

---

[11]Similar to the previous example, the n-dimensional case involves including a factor of $(n-1)$.

[12]Including diffeomorphism changes here arises from our interest in solutions to Ricci flow up to diffeomorphism. This will later appear in more detail in our study of the Einstein equations (also diffeomorphism-equivariant), where the space of solutions will resemble 'metrics modulo diffeomorphisms.'

[13]In n dimensions, this equation is often written

$$-2\mathrm{Ric}(g_0) = \mathcal{L}_V g_0 + 2\lambda g_0,$$

for λ a scaling constant, which is negative, zero or positive denoting a **shrinking**, **steady** or **expanding** Ricci soliton. In our discussion, r takes the role of λ.

Definition 3.1.4 (Potential function)**.** The trace of the gradient Ricci soliton equation is

$$\Delta f = R - r.$$

Since the right-hand side integrates to zero by the definition of r, the equation is solvable for any solution of the normalised Ricci flow. We call a solution f of the above equation the **potential function**, or potential of the curvature. Note that the potential is unique up to adding a function of time alone, since on a closed surface the only harmonic functions are constants.

The above definitions of self-similar and Ricci soliton solutions correspond to one another in the following way. Consider differentiating a self-similar metric with respect to time and identifying V_t with the vector field generated by the diffeomorphisms ψ_t. We then obtain

$$\partial_t g = \partial_t \left(\psi_t^* g_0\right) = \mathcal{L}_V g.$$

We will use the above notions of Ricci solitons and potential functions frequently in our proof of the Uniformisation Theorem. The following section gives us our first taste of the promised usefulness of the maximum principle and evolution equations.

3.2 Maximum Principle Revisited

In this short section we will make a first use of our prized maximum principle, found in Proposition 2.3.1, which will be of utmost importance in the sections to come. It will allow us to find time-dependent bounds for the Ricci scalar R.

Recall that we are interested in reaction-diffusion equations (for a function u) of the form

$$\partial_t u = \Delta u + F(u) + \langle \nabla u, V \rangle,$$

where $F(u)$ is some (locally Lipschitz) function, $V_t = V$ is a one-parameter family of vector fields and $\langle \cdot, \cdot \rangle$ denotes contraction via the metric. The maximum principle allows us to ignore the Laplacian and gradient terms, focusing only on the function F to find time-dependent bounds for u if it is initially (at $t = 0$) bounded. We can use Proposition 2.3.1 on our evolution equation for the Ricci scalar R (3.9).

Similarly to Section 2.3, we will revert to using x to denote a point on our manifold, and this convention will continue throughout the rest of this chapter.

Example 3.2.1 (Ricci scalar evolution)**.** Using Proposition 2.3.1 with (3.9), we have $F(R) = R(R - \rho)$ and $V = 0$. Thus, we search for a solution $\alpha(t)$ (and another one $\beta(t)$) to

$$\partial_t \alpha = \alpha(\alpha - \rho). \tag{3.13}$$

Direct computation for non-zero ρ gives

$$\alpha(t) = \frac{\rho}{e^{C+\rho t} + 1}, \quad \text{where} \quad C \in \mathbb{R} \quad \text{is such that} \quad \alpha_0 = \frac{\rho}{e^C + 1}.$$

Rearranging to explicitly include α_0, we find

$$\alpha(t) = \frac{\alpha_0 \rho}{(\rho - \alpha_0)e^{\rho t} + \alpha_0} = \frac{\rho}{1 - \left(1 - \frac{\rho}{\alpha_0}\right) e^{\rho t}},$$

which can be checked to satisfy the ordinary differential equation (3.13). Thus, if we can find bounds α_0 and β_0 such that for all $x \in \Omega$ we have $\alpha_0 \leq R(x,t) \leq \beta_0$, then

$$\frac{\alpha_0 \rho}{(\rho - \alpha_0)e^{\rho t} + \alpha_0} \leq R(x,t) \leq \frac{\beta_0 \rho}{(\rho - \beta_0)e^{\rho t} + \beta_0} \quad \text{for all} \quad (x,t) \in \Omega \times [0,T].$$

Note that the above does not work when $\rho = 0$. However, in this case, a similar (and simpler) computation can be made to find

$$\alpha(t) = \frac{\alpha_0}{1 - \alpha_0 t},$$

which satisfies the above for $F(R) = R^2$, as can be verified.

If we now inspect the above in the context of the normalised Ricci flow, allowing ρ to be the average Ricci scalar r and explicitly replacing α with R (whose role it was playing), we have the following ordinary differential equation:

$$\partial_t R = R(R - r) \quad \text{with} \quad R(0) = R_0,$$

which has solutions:

1. If $r \neq 0$ and $R_0 \neq 0$, then
$$R(t) = \frac{r}{1 - \left(1 - \frac{r}{R_0}\right) e^{rt}};$$
2. If $r = 0$, then
$$R(t) = \frac{R_0}{1 - R_0 t};$$
3. If $R_0 = 0$ (which means that $r = 0$, since it is independent of time), then $R(t) = 0$.

Thus, we can draw the following conclusion. If $R_0 > \max\{r, 0\}$, then for

$$0 < T := \begin{cases} -\frac{1}{r} \log\left(1 - \frac{r}{R_0}\right) & \text{if } r \neq 0, \\ \frac{1}{R_0} & \text{if } r = 0, \end{cases}$$

we have $R(t) \to \infty$ as $t \to T$. Note that thankfully this divergence is only within the context of the ordinary differential equation (3.13) for R, not under Ricci flow. It is a warning sign, however, and we will have to work hard to find bounds for R—because any divergence will prevent us from having existence of our flow for all time, as we shall soon see.

Another tool we will briefly mention is the **strong maximum principle**, which we discuss here without formal formulation or proof. If $u \geq 0$ is a non-negative solution to the heat equation (2.10), then intuitively we know that $u > 0$ for any time $t > 0$ unless $u = 0$ everywhere.[14] A similar principle applies to Ricci flow: if the Ricci scalar R is initially non-negative, then the strong maximum principle applied to its evolution equation allows us to conclude that unless $R = 0$ everywhere, R will be positive for all time $t > 0$ (as long as the solution exists).

[14] Here infinite speed of propagation of information within parabolic equations is very clear.

Moving on, we will use the tools above to inspect possible bounds (or lack thereof) of the Ricci scalar. *We will be considering the normalised Ricci flow for the rest of the chapter, so the previous ρ will again be replaced by the average Ricci scalar r, which we know to be time-independent in this context.*

From the previous discussion, we know that the Ricci scalar may blow up under the ordinary differential equation (3.13), eliminating the possibility of the maximum principle granting us a satisfactory upper bound on it. However, we can use our results to derive lower bounds for the Ricci scalar. A useful piece of notation from here on out will be the following:

$$R_{\min}(t) := \inf_{x\in\Sigma} R(x,t),$$

which is always initially well-defined (written $R_{\min,0} := R_{\min}(0)$) as we always consider R_0 to be bounded.

Proposition 3.2.1. *Let $g(t)$ be a solution of the normalised Ricci flow* (3.3). *Now,*

1. *If $r < 0$, then*

$$R - r \geq \frac{r}{1 - \left(1 - \frac{r}{R_{\min,0}}\right)e^{rt}} - r \geq (R_{\min,0} - r)e^{rt};$$

2. *If $r = 0$, then*

$$R \geq \frac{R_{\min,0}}{1 - R_{\min,0}t} > -\frac{1}{t}; \text{ and}$$

3. *If $r > 0$ and $R_{\min,0} < 0$, then*

$$R \geq \frac{r}{1 - \left(1 - \frac{r}{R_{\min,0}}\right)e^{rt}} \geq R_{\min,0}e^{-rt}.$$

In particular, the right-hand side tends to 0 as $t \to \infty$ in each case, and so by taking the infimum of the left-hand sides, we find that

1. *If $r < 0$, then $R_{\min}(t) \to r$ exponentially quickly as $t \to \infty$;*

2. *If $r > 0$ and $R_{\min}(t) \geq 0$ at some point, then it remains greater than 0 for all time;*

3. *If $r > 0$ and $R_{\min,0} < 0$, then $R_{\min}(t) \to 0$ exponentially quickly.*

Proof. In each case, the middle step of the bounds is given by the previous discussion of minimum principles, and the right-most step is a further simplification for demonstrative purposes. The conclusions follow from the bounds and the strong maximum principle (for the second). □

The following section describes the qualitative behaviour of a Riemannian manifold undergoing Ricci flow, so that we have a better idea of what exactly the process entails.

3.3 Qualitative Ricci Flow Behaviour

As hinted at in previous sections, a metric of a Riemannian manifold undergoing Ricci flow will (under certain assumptions) tend to a metric of constant curvature. Because our intuitive visualisation of curvature involves spaces with peaks[15] and valleys, or protrusions and cavities, it is challenging to think of a Riemannian manifold as its own space, instead of embedded within a larger ambient one.[16] This section attempts to qualitatively explain the Ricci flow process, using a (traditional) comparison with a different 'geometric flow': mean curvature flow.

Example 3.3.1 (Mean curvature flow)**.** Consider a one-parameter family of 2-manifolds $\mathcal{M}(t)$, all embedded in some ambient space, taken to be Euclidean. Roughly, **mean curvature flow** involves the following process:

1. For each point $p \in \mathcal{M}(t)$, approximate a neighbourhood of p by a 2-sphere of radius ϱ (write this sphere $\mathbb{S}^2_{\varrho,p}$);[17]

2. Consider a vector[18] pointing (since we are embedded in Euclidean space, there is no issue with this) from p towards the centre of $\mathbb{S}^2_{\varrho,p}$, scaled to be inversely proportional to the radius ϱ;

3. Flow the point $p \in \mathcal{M}(t)$ along this vector, where the length of the vector dictates the speed of the flow, to create the manifold one time-step later.

Note that this process can be simplified further by considering the 1-dimensional analogue, **curve shortening flow**, where a curve embedded in $\mathbb{R}^2$ flows and eventually shrinks to a point, as shown in Figure 3.2.

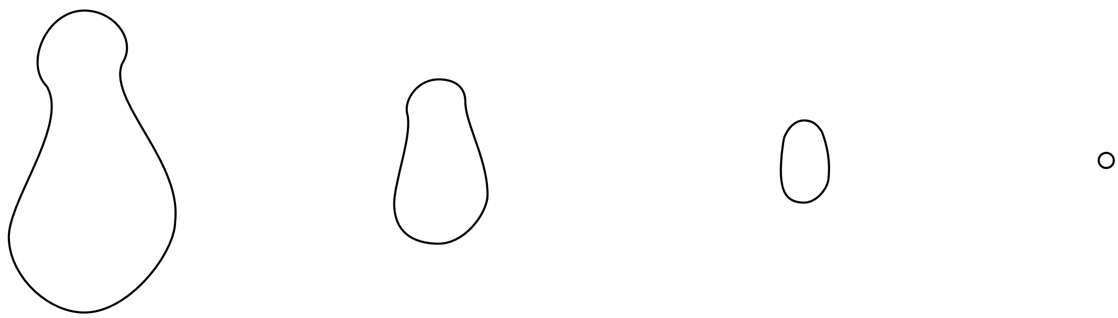

Figure 3.2: An example of curve shortening flow.

Mean curvature flow (within an ambient space) involves a flow of the differentiable manifold itself, the points of the topological space themselves moving. Peculiarities can be encountered in simple examples, such as the **neck-pinch**, described as follows. Consider a 2-manifold consisting of (something closely resembling) two copies of $\mathbb{S}^2$ connected by a cylinder $\mathbb{S}^1 \times [0, 1]$, where the radius of the cylinder is smaller than the radius of the spheres: see Figure 3.3, whose manifold resembles a dumbbell.

[15]Though not perfectly pointed (conical), as this does not satisfy the condition of being locally homeomorphic to Euclidean space.
[16]See Section 4.3 for a more precise discussion of embedded submanifolds.
[17]If a neighbourhood of p is perfectly flat, take the radius r to be very big—this is a qualitative depiction, after all.
[18]This is the **mean curvature vector**, hence the name of the flow.

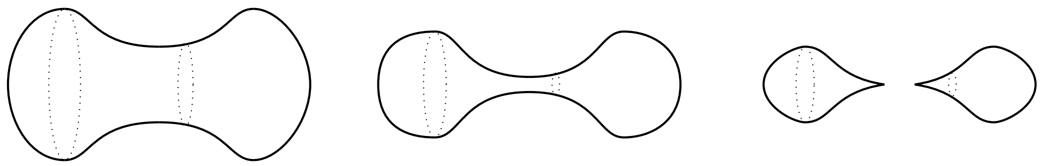

Figure 3.3: A neck-pinch on a dumbbell.

In this case, mean curvature flow dictates that the spheres will begin to shrink, as will the cylinder (along its radial directions). However, as depicted in Figure 3.3, the smaller radius of the cylinder will result in a neck-pinch: the handle of the dumbbell will shrink to a 1-dimensional line, and the 2-manifold description of the example will break down at this singularity.[19]

Despite having a single time-derivative equating second-order spatial derivatives, mean curvature flow and Ricci flow differ. One similarity lies in their tendencies to 'smoothen' manifolds to create spaces of constant curvature. However, the curvature in the case of Ricci flow is metric-induced, rather than arising from embedding in an ambient space. In Ricci flow, the manifold itself is static: the points remain stationary, the only thing changing being the metric. As such, singularities of the type described above cannot arise. However, in Ricci flow, singularities of another form do arise, in the following way.

Consider the same dumbbell set-up as in the mean curvature flow example. Ricci flow will also induce a neck-pinch, except that the stationary nature of the points of the manifold implies that instead of a neck-pinch as seen in Figure 3.3, which is easy to visualise because of our embedding in an ambient Euclidean space, it is the metric that creates the illusion[20] of the points along the neck shrinking to a point. This is done in the only ways the metric can: the volume (which is calculated using the metric; see Definition 2.1.7) shrinks to zero along the neck and geodesics that wrap around the neck have lengths (also metric-dependent; see Definition 2.1.5) that shrink to zero. This can also be described as having a shrinking injectivity radius, defined as follows.

Definition 3.3.1 (Injectivity radius)**.** Consider a Riemannian manifold $(\mathcal{M}, g)$. Then the **ball** of radius $\varrho > 0$ centred at $p \in \mathcal{M}$ is written $\mathcal{B}_{g,\varrho}(p)$ and is given by

$$\mathcal{B}_{g,\varrho}(p) := \{q \in \mathcal{M} \mid d_g(p,q) < \varrho\}, \tag{3.14}$$

[19]Note that mean curvature flow (just as in Ricci flow) has only a single time-derivative of the object undergoing the flow. Physically, this means that the points move with no acceleration, which would require a second-order temporal derivative. Thus, if a sphere with a small bulge undergoes mean curvature flow, the bulge, though shrinking, remains until the manifold becomes singular. The greater radius of the points on the bulge will guarantee them a higher velocity, but the bulge will only 'flatten' (to create a sphere) asymptotically. On the other hand, if there was an acceleration term then the larger radius of the bulge would result in the points of the bulge 'catching up' with the points on the sphere after some finite time—after this point, however, the higher speed of these points would carry them past the sphere's non-bulge points, creating a dent, which would then have a mean curvature vector pointing outward, sending (after slowing, stopping and changing direction) the points in the dent back outward, and so on, making the bulge oscillate (bulge-dent-bulge-dent-etc.) as the manifold as a whole shrinks. This oscillation reminds us of a wave, which it should, since the wave equation (2.9) indeed has a second-order temporal derivative.

[20]In some sense, if we consider our starting point to be the differentiable manifold $\mathcal{M}$ rather than the Riemannian manifold $(\mathcal{M}, g)$, then the metric g is artificially added and thus problems it encounters are not dire. However, since we consider the Riemannian manifold to be fundamental, problems faced by the metric are of utmost importance.

for $d_g(p,q)$ the distance between p and q induced by the metric g defined by (2.3). Now, at $p \in \mathcal{M}$, the **injectivity radius** $\mathrm{inj}_g(p)$ is given by

$$\mathrm{inj}_g(p) := \sup\{\varrho > 0 \mid \exp_p : \mathcal{B}_{g,\varrho}(p) \to \mathcal{M} \text{ is injective}\},$$

where the exponential map was defined in Definition 2.1.6. The injectivity radius of the manifold itself $\mathrm{inj}(\mathcal{M}, g)$ is

$$\mathrm{inj}(\mathcal{M}, g) := \inf_{p \in \mathcal{M}} \mathrm{inj}_g(p).$$

In general, spaces of positive curvature (such as spheres, dumbbells, etc.) encounter singularities of the type described above during Ricci flow, signified quantitatively by a vanishing injectivity radius. In the 2-dimensional case, however, neck-pinches do not occur, for the following reason. The handle of the dumbbell is roughly $\mathbb{S}^1 \times [0,1]$, with extremal bulbs (the 'weights') considered to be copies of $\mathbb{S}^2$. Though the $\mathbb{S}^2$ bulbs have positive curvature and shrink under Ricci flow to a point in finite time as in Example 3.1.4, the handle has no curvature (as $\mathbb{S}^1$ is flat) and therefore does not pinch off. Neck-pinches do arise in higher dimensions, since the handle is then $\mathbb{S}^{n-1} \times [0,1]$, which has positive curvature when $n \geq 3$.

In our 2-dimensional context, for the purposes of proving the Uniformisation Theorem, we wish to show that Ricci flow exists for all time. This makes the case of a shrinking 2-sphere unfortunate, which is why we will consider the normalised Ricci flow. In this case, no singularity arises: instead, the metric will tend to one of constant curvature.

We have restricted ourselves to the case of closed Riemannian 2-manifolds Σ with genus $\mathrm{g}(\Sigma) = \mathrm{g}$. The positive, zero and negative curvature cases here correspond to surfaces with genera $\mathrm{g} = 0$, $\mathrm{g} = 1$ and $\mathrm{g} > 1$, respectively. Arising from the Uniformisation Theorem (which we have yet to prove), this is because these surfaces can be constructed by a quotient of their universal covering spaces by a discrete subgroup, where the universal coverings here are the 2-sphere $\mathbb{S}^2$, flat 2-space $\mathbb{R}^2$ and hyperbolic 2-space $\mathbb{H}^2$, respectively, which have positive, zero and negative constant curvature.[21]

The following section establishes short- and long-time existence for Ricci flow by comparing it to the Ricci-DeTurck flow, which is shown to be parabolic (and thus existence and uniqueness results from parabolic theory apply to it).

3.4 Existence and Uniqueness of Solutions

This section briefly addresses the existence and uniqueness of solutions to Ricci flow. First, we will show that the flow exists and is unique for short times. Then, we will state and discuss a result on long-time existence. We will then assume long-time existence for the sections to come before encountering the bound necessary to secure it.

Our survey of short-time existence will be as brief as possible, leaving the details to be found in Chapter 5 of [Top06], Chapter 3 of [CK04] or Chapter 5 and Appendix A of [She06]. We will follow these references (and the literature at large) and prove the existence for standard (non-normalised) Ricci flow in n dimensions, and hence we write our manifold as $\mathcal{M}$. We return to our discussions of Section 2.3, though with slightly modified notation: we are considering partial differential equations of the form

$$\partial_t u = L(u) \quad \text{with} \quad u(x,0) = u_0(x), \tag{3.15}$$

[21] When we apply the Uniformisation Theorem to General Relativity we will use the classification of surfaces by genus, and consider the three situations mentioned above as case studies.

where L is a differential operator of order k and u is a time-dependent function on $\mathcal{M}$. To be precise, we will refer to Section A.2 of the appendix and consider a vector bundle $\pi : \mathcal{E} \to \mathcal{M}$, where now $u : \mathcal{M} \times [0, T) \to \mathcal{E}$ for some $T > 0$, and $L : \mathcal{C}^\infty(\mathcal{E}) \to \mathcal{C}^\infty(\mathcal{E})$. We will be interested in differential operators of the form

$$L(u) = \sum_{|\alpha| \leq k} L_\alpha \partial^\alpha u, \tag{3.16}$$

where the L_α are homomorphisms from $\mathcal{E}$ to itself and we have employed the multi-index notation introduced in Section 2.3. This notation will be used throughout our short-time existence discussion. We have the following definition.

Definition 3.4.1 (Strong parabolicity)**.** For a covector ω, the **principal symbol** of the operator L in the direction of ω is a vector bundle homomorphism $\hat{\sigma}[L](\omega) : \mathcal{E} \to \mathcal{E}$ defined by

$$\hat{\sigma}[L](\omega)(u) := \sum_{|\alpha| = k} L_\alpha(u) \prod_i \omega^{\alpha_i}.$$

Now, the partial differential equation (3.15) is called **strongly parabolic** is there exists $\delta > 0$ such that at each $p \in \mathcal{M}$, for all non-zero covectors ω and non-zero functions u, we have

$$\langle \hat{\sigma}[L](\omega)(u), u \rangle > \delta |\omega|^2 |u|^2.$$

However, the Ricci tensor is not linear and so cannot be written in the form (3.16). As such, we define its linearisation—in the same spirit as a derivative being the linearisation of a function—as follows.

Definition 3.4.2 (Linearisation)**.** The **linearisation** of an operator $L : \mathcal{C}^\infty(\mathcal{E}) \to \mathcal{C}^\infty(\mathcal{E})$ is the linear map $D[L] : \mathcal{C}^\infty(\mathcal{E}) \to \mathcal{C}^\infty(\mathcal{E})$ given by

$$D[L](u) := \left. \frac{\mathrm{d}}{\mathrm{d}t} L(u(t)) \right|_{t=0}.$$

We now call (3.15) **strongly parabolic** if its linearisation $\partial_t u = D[L](u)$ is.

The following general theorem guarantees existence and uniqueness of solutions to strongly parabolic equations.

Theorem 3.3. *A strongly parabolic equation of the form* (3.15) *has a solution on some time interval* $[0, T)$ *(for* $T > 0$*), which is unique as long as it exists.*

Proof. The proof can be found in [LSU88], a gigantic text on parabolic theory. □

We can use the above in our Ricci flow (where we resume the notation $\partial_t g_{ij} = h_{ij}$ from Proposition 3.1.1) and compute the following linearisation:

$$D[-2\mathrm{Ric}](h)_{ij} = g^{kl} \nabla_k \nabla_l h_{ij} + \nabla_i V_j + \nabla_j V_i + \text{lower-order terms in } h, \tag{3.17}$$

where the lower-order terms will not contribute to the principal symbol, and where we have defined

$$V_i := g^{kl} \left(\frac{1}{2} \nabla_i h_{kl} - \nabla_l h_{ik} \right).$$

This follows from the evolution equation of the Ricci tensor, which can be computed from the results in Proposition 3.1.1 (or is found in Chapter 3 of [CK04]).[22]

The first term in (3.17) is the Laplacian, which is fine for our desires of strong parabolicity. The Lie derivative along V is what makes Ricci flow not strongly parabolic, which induces us to turn to the Ricci-DeTurck flow from Example 3.1.2 (this is called the DeTurck trick, hence the naming of the flow—it was first done in [DeT83]). If $\bar{g}$ is a solution to the Ricci-DeTurck flow (3.2) for some DeTurck vector ξ then a similar computation (which can be found in Chapter 5 of [She06], among other places) to the one comparing solutions to the normalised Ricci flow and standard Ricci flow gives that the metric $g(t) := \varphi_t^* \bar{g}(t)$ (where the diffeomorphisms φ_t induce the DeTurck vector $\xi = \xi_t$) solves Ricci flow.[23]

Now, if we write $D[\mathrm{Ric}, \xi]$ for the linearisation of the differential operator found in the Ricci-DeTurck flow with DeTurck vector ξ, then we have

$$D[\mathrm{Ric}, \xi](h)_{ij} = g^{kl}\nabla_k \nabla_l h_{ij} + \nabla_i V_j + \nabla_j V_i - D[\nabla_i \xi_j + \nabla_j \xi_i](h) + \text{lower-order terms in } h.$$

We now can choose[24] ξ so that the terms involving V and ξ cancel. We let

$$\xi_i := -g^{kl} g_{ij} \left(\Gamma^j_{kl} - \tilde{\Gamma}^j_{kl} \right),$$

for some fixed constant connection[25] $\tilde{\nabla}$ with connection coefficients $\tilde{\Gamma}^j_{kl}$. This choice cancels the desired terms of the linearisation of Ricci-DeTurck flow, leaving its principal symbol to satisfy

$$\langle \hat{\sigma}(D[\mathrm{Ric}, \xi])(\omega)(h), h \rangle = |\omega|^2 |h|^2,$$

which satisfies the strong parabolicity criterion in Definition 3.4.1 by choosing some $\delta < 1$. Thus, by Theorem 3.3, the Ricci-DeTurck flow exists for short times. By finding the diffeomorphisms φ_t that generate ξ_t we can pull-back solutions to the Ricci-DeTurck flow to conclude that standard Ricci flow exists (and is unique by arguments found in Chapter 4.4 of [CK04]) for short times as well.

Note: the rest of the chapter will take place in a 2-dimensional context. We also introduce a crucial piece of notation, which we will employ repeatedly for the rest of the chapter: *we write $(\Sigma, g(t), g_0)$ for the one-parameter family of closed and connected 2-dimensional Riemannian manifolds that solve the (normalised) Ricci flow with initial metric $g_0 := g(0)$.*

As is common in the theory of partial differential equations, the long-time existence of solutions will be proved by *a priori* bounds. We will assume the following result, whose proof will follow once we find uniform bounds on the curvature.

Proposition 3.4.1. *The unique solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow exists for all time.*

---

[22] If we had naïvely tried to apply the above to the usual Ricci flow, we have the following problem. The principal symbol of $-2\mathrm{Ric}$ in some direction ω is then

$$\hat{\sigma}[-2\mathrm{Ric}](\omega)(h)_{ij} = g^{kl}\left(\omega_k \omega_l h_{ij} + \omega_i \omega_j h_{kl} - \omega_l \omega_i h_{jk} - \omega_l \omega_j h_{ik} \right),$$

which can be shown to not satisfy the strongly parabolic criterion by choosing $h_{ij} = \omega_i \omega_j$, as can be checked.

[23] We will perform this computation in the 2-dimensional case in the discussion preceding Proposition 3.13.2.

[24] Decisions of this form are often called gauge-fixing, particularly in physics. We will see more of this in Section 4.8.

[25] See Section A.3 of the appendix for an introduction to general connections.

Discussion of proof. (See Proposition 5.19 and Corollary 7.2 of [CK04] for details.)

As discussed in Chapter 7 of [CK04], the length of the interval on which the solution to Ricci flow exists is bounded from below by a constant that is inversely proportional to the maximum of the Riemann tensor (and thus of the Ricci scalar, in our 2-dimensional context). This means that the proof can be reduced to the validity of the following Lemma.

Lemma 3.4.1. *The unique solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow exists on a maximal time interval $0 \leq t < T \leq \infty$. Further, $T < \infty$ only if*

$$\sup_{x\in\Sigma} |\mathrm{Riem}(x,t)| \xrightarrow{t\to T} \infty,$$

where the Riemann $(4,0)$-tensor takes some unit vector V as its inputs

$$\mathrm{Riem}(V,V,V,V)(x,t) = \mathrm{Riem}(x,t),$$

within this Lemma.

We make the remark here that the blowing-up of the Riemann tensor is equivalent (in our 2-dimensional context) to the blowing up of the Ricci scalar. Thus, when we establish uniform bounds for the Ricci scalar, the above result follows and we have our desired existence for long times. We will see bounds for R in the sections to come. □

The following section will briefly outline our plan of attack for proving the Uniformisation Theorem using Ricci flow.

3.5 Strategy for Proving Uniformisation

As the following sections will prove, our comments about the parabolic nature of Ricci flow being a driving force behind its 'smoothening' of the metric so that as time tends to infinity the metric becomes one of constant curvature, is indeed true. This is done by keeping a close eye on the Ricci scalar and by showing that it converges to the average Ricci scalar, and thus is constant throughout the manifold. Various cases will be studied, depending on the initial value of the Ricci scalar, as well as on its average. The goal will be the following Theorem.

Theorem 3.4. *On a closed, connected and oriented 2-dimensional Riemannian manifold (Σ, g_0), there exists a unique solution (which exists for all time) of Riemannian metrics $g(t)$ to the normalised Ricci flow with initial metric $g(0) = g_0$ such that $g(t)$ converges uniformly as $t \to \infty$ (in any C^k norm) to a metric g_∞ of constant curvature.*

In the language of our space of metrics $\mathscr{M}$, this result claims that any point $g_0 \in \mathscr{M}$ can be (maybe asymptotically) connected via a curve $g : [0,\infty) \to \mathscr{M}$ to some point g_∞ that has constant curvature. In this way the constant curvature metrics (which, up to re-scaling, are the Hilbert manifolds $\mathscr{M}_\lambda$ with $\lambda \in \{-1,0,1\}$) form what are sometimes called **attractor basins** for the normalised Ricci flow. Because the normalised Ricci flow does not change the conformal equivalence class of the metric g_0, Theorem 3.4 indeed implies the Uniformisation Theorem (Theorem 2.2).

This section will present our agenda for the coming sections. Proposition 3.4.1 will be assumed for the time being, so that we may speak of solutions to the normalised Ricci flow without pausing to question their long-time existence. With the existence of the flow in hand, Theorem 3.4 will now be split into the four following cases, for r the average Ricci scalar:

1. $r < 0$;
2. $r = 0$;
3. $r > 0$ and the initial Ricci scalar is non-negative;
4. $r > 0$ and the initial Ricci scalar not necessarily non-negative.

If we assume $g(t)$ to be the solution to the normalised Ricci flow, then each of our cases will be approached in the following manner:

1. Establish uniform bounds for $g(t)$ depending on g_0;
2. Show that R approaches r as $t \to \infty$;
3. Show that $|\nabla^k R|$ vanishes as $t \to \infty$.

Naturally, some of these steps will require much work. It will be useful to establish evolution equations for tensorial quantities in the manner of the Ricci scalar (3.9), then apply the maximum principle to find a decaying upper bound for the quantities, which we will relate to $|R - r|$ somehow. Then, showing that these quantities decay will allow us to conclude that $R \to r$ as $t \to \infty$.

The first two cases are comparatively simple. When $r < 0$, instead of looking directly at $R - r$ we will consider the closely-related quantity[26] $R - r + |\nabla f|^2$ (where f is the potential function introduced in Definition 3.1.4). Using the maximum principle on the evolution equation of this quantity will allow us to conclude that $R \to r$ as $t \to \infty$. When $r = 0$, we will determine decaying bounds directly on R, so that $R \to 0$ as $t \to \infty$. In both cases, we will then simply need to show that all derivatives $\nabla^k R$ of the Ricci scalar vanish to obtain our desired convergence in any $\mathcal{C}^k$ norm.

The cases with $r > 0$ are much more complicated. The second, where the initial Ricci scalar can be of mixed sign, will be proved by showing that within finite time, it turns into the first case. That is, $R_{\min}$ becomes positive at some time, and thus the proof of the first case allows us to conclude upon restarting the flow at this time. To prove the first case, however, new quantities need to be defined and inspected. A closer look at the quantity[27] $\Delta \log R + R - r$ (once again closely related to $R - r$) will help us determine the Ricci scalar at one point in terms of it at an earlier point,[28] which will help us determine a lower bound for R. An inspection of the tensor[29] $\nabla^2 f - \frac{1}{2}(R - r)g$ will allow us to show that metrics with $r > 0$ (that is, metrics on spaces with vanishing genus, like $\mathbb{S}^2$) tend to become gradient Ricci solitons, which we will show always have constant curvature—thus our limiting metric does too. Finally, bounds on the injectivity radius and the diameter of the manifold will allow us to conclude $R \to r$ and $|\nabla^k R| \to 0$ in a satisfactory manner as $t \to \infty$.[30]

The following section (Section 3.6) has been isolated to allow us to get a feel for the 'evolution-equation-then-maximum-principle' strategy. Section 3.7 will then treat the first two cases ($r \leq 0$). The six following sections will prove the $r > 0$ cases, by first establishing bounds for the metric on an interval (Section 3.8); taking a closer look at gradient Ricci soliton solutions (Section 3.9); defining and exploring a quantity known as

[26]Soon to be known as the H scalar.
[27]Soon to be known as the Q scalar.
[28]This is called a Harnack inequality.
[29]Soon to be known as the M tensor.
[30]Note that our approach using entropy and Harnack inequalities is not the only one. See [AB10] for an alternate perspective, which employs what is called isoperimetric comparisons.

the entropy (Section 3.10); proving bounds for R, the diameter and the injectivity radius (Section 3.11); relating the Ricci scalar at a point to the Ricci scalar at an earlier point (Section 3.12); and concluding (Section 3.13).

3.6 Evolution Equations in Action

As shown in the beginning of this chapter, when a metric undergoes the normalised Ricci flow (3.3), its Ricci scalar obeys the following evolution equation:

$$\partial_t R = \Delta R + R(R - r). \tag{3.18}$$

We also recall that by Definition 3.1.4, a potential function f satisfies the equation

$$\Delta f = R - r. \tag{3.19}$$

From these two simple facts we will define a new scalar quantity, H, and use its evolution equation and the maximum principle to show that $R \to r$ as $t \to \infty$.

The presence of many ∇'s and Δ's inspires us to recall several useful identities that the Riemann tensor satisfies (and once traced, that the Ricci tensor and scalar satisfy), which will be used frequently in our coming derivations and proofs. By definition, we have[31]

$$(\nabla_k \nabla_l - \nabla_l \nabla_k) \nabla^i f = R^i_{jkl} \nabla^j f \quad \text{and} \quad (\nabla_k \nabla_l - \nabla_l \nabla_k) \nabla_j f = -R^i_{jkl} \nabla_i f. \tag{3.20}$$

These combine to give the first part of the following result.

Lemma 3.6.1. 1. *In 2 dimensions, we have the following identity:*

$$\nabla \Delta = \Delta \nabla - \frac{1}{2} R \nabla, \tag{3.21}$$

which holds whenever applied to a scalar function;

2. *For any tensor τ, we have*

$$\Delta |\tau|^2 = 2\langle \Delta \tau, \tau \rangle + 2|\nabla \tau|^2. \tag{3.22}$$

Proof. 1. Employing the Riemann tensor identities (3.20) and (2.7), we have

$$\nabla_i \Delta f = \nabla_i \nabla_k \nabla^k f = \nabla_k \nabla_i \nabla^k f + R^k_{lik} \nabla^l f = \nabla_k \nabla^k \nabla_i f - \frac{1}{2} R g_{li} \nabla^l f = \Delta \nabla_i f - \frac{1}{2} R \nabla_i f,$$

for some scalar function f, as desired;

2. This identity follows by the product rule. □

As previously mentioned, potential functions are unique up to adding a function of time. With this in mind, we have the following proposition, which gives the evolution equations of the potential and of its gradient as well as a bound for the potential.

[31] In the context of these identities and the subsequent Lemma, it is more proper to write d, the exterior derivative, (or, locally, ∂_i) instead of ∇_i (when applied to a scalar function). However, for notational convenience we will take advantage of their equality when applied to scalars and use ∇ and ∇_i instead.

Proposition 3.6.1. *For a potential function $f_0(x,t)$ of a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow, there exists a function $c(t)$ only dependent on time such that the potential $f := f_0 + c$ satisfies*

$$\partial_t f = \Delta f + rf. \tag{3.23}$$

This allows us to choose our potential function to satisfy (3.23). *We then have that*

1. *There exists a constant $C \geq 0$ such that*

$$|f| \leq Ce^{rt}; \tag{3.24}$$

2. *The gradient of the potential satisfies the following evolution equation:*

$$\partial_t |\nabla f|^2 = \Delta |\nabla f|^2 - 2|\nabla\nabla f|^2 + r|\nabla f|^2. \tag{3.25}$$

Proof. We differentiate the defining equation of the potential function (3.19) for f_0:

$$\partial_t(\Delta f_0) = \partial_t(R - r).$$

Now, using the evolution equations of Δ (3.8) and R (3.18), as well as the fact that r is constant, we find

$$(R-r)^2 + \Delta(\partial_t f_0) = \Delta\Delta f_0 + R(R-r),$$

which rearranges to become

$$\Delta(\partial_t f_0) = \Delta(\Delta f_0 + rf_0).$$

Now, the only harmonic functions on closed 2-manifolds are constants, so there exists a function $\gamma(t)$ only dependent on time such that

$$\partial_t f_0 = \Delta f_0 + rf_0 + \gamma.$$

By observation, if we take

$$c(t) := -e^{rt} \int_0^t e^{-rt'} \gamma(t') dt',$$

we obtain the desired function $f := f_0 + c$, since

$$\partial_t f = \partial_t f_0 + \partial_t c = \Delta f_0 + rf_0 + \gamma - e^{rt}\left(e^{-rt}\gamma\right) - re^{rt}\int_0^t e^{-rt'}\gamma(t')dt' = \Delta f_0 + rf_0 + rc,$$

as claimed. Now,

1. The constant C is found by applying the maximum principle to the evolution equation of f (3.23);
2. We use the evolution equations of g^{ij} (3.6) and f (3.23) and our favourite identities (3.21) and (3.22) to pass down each line, respectively:

$$\begin{aligned}\partial_t |\nabla f|^2 &= \partial_t\left(g^{ij}\nabla_i f \nabla_j f\right)\\ &= (R-r)|\nabla f|^2 + 2\left(\nabla_i \partial_t f\right)\nabla^i f\\ &= (R-r)|\nabla f|^2 + 2\left\langle \nabla\Delta f + r\nabla f, \nabla f\right\rangle\\ &= r|\nabla f|^2 + 2\langle \Delta\nabla f, \nabla f\rangle\\ &= \Delta|\nabla f|^2 - 2|\nabla\nabla f|^2 + r|\nabla f|^2,\end{aligned}$$

as desired. □

We now make an important definition.

Definition 3.6.1 (H scalar)**.** For f a potential function, let H be the following scalar quantity:

$$H := R - r + |\nabla f|^2. \tag{3.26}$$

This quantity is related to $R - r$, which is what we want to show vanishes in a satisfactory manner. We have added a $|\nabla f|^2$ because of the effect it will have on the evolution equation of H. By Proposition 3.6.1, $|\nabla f|^2$ behaves well in time. Thus, we hope that H will as well, so that an application of the maximum principle on its evolution equation gives us a decaying bound. If H is bounded, then $R - r$ will be as well (since $|\nabla f|^2 \geq 0$), and thus we will have our desired decaying bound on $R - r$. The following result addresses these wishes.

Proposition 3.6.2. *The quantity H defined above evolves under the normalised Ricci flow (with initial metric g_0) as*

$$\partial_t H = \Delta H - 2\left|\nabla\nabla f - \frac{1}{2}(\Delta f)g\right|^2 + rH. \tag{3.27}$$

Furthermore, there exists a constant C dependent only on g_0 such that

$$R - r \leq H \leq Ce^{rt}. \tag{3.28}$$

Proof. Using the defining equation of f (3.19) and the evolution equation of R (3.18), we find

$$\partial_t(R - r) = \Delta R + R(R - r) = \Delta(R - r) + (\Delta f)^2 + r(R - r).$$

We then combine this with our evolution equation for $|\nabla f|^2$ (3.25) to obtain

$$\begin{aligned}
\partial_t H &= \partial_t(R - r) + \partial_t|\nabla f|^2 \\
&= \Delta(R - r) + (\Delta f)^2 + r(R - r) + \Delta|\nabla f|^2 - 2|\nabla\nabla f|^2 + r|\nabla f|^2 \\
&= \Delta H - 2\left|\nabla\nabla f - \frac{1}{2}(\Delta f)g\right|^2 + rH,
\end{aligned}$$

where the final line follows from the definition of H (3.26), as well as the identity

$$\left|\nabla\nabla f - \frac{1}{2}(\Delta f)g\right|^2 = |\nabla\nabla f|^2 - \frac{1}{2}(\Delta f)^2. \tag{3.29}$$

Using the maximum principle, we obtain the desired constant C, and the lower bound follows by observation. □

As promised, the previous bounds on H give the following result regarding time-dependent upper and lower bounds of the Ricci scalar.

Proposition 3.6.3. *For a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow, there exists a constant C dependent only on g_0 such that*

1. *If $r < 0$, then*

$$r - Ce^{rt} \leq R \leq r + Ce^{rt};$$

2. *If $r = 0$, then*

$$-\frac{C}{1+Ct} \leq R \leq C;$$

3. *If $r > 0$, then*

$$-Ce^{-rt} \leq R \leq r + Ce^{rt}.$$

Proof. The upper bounds follow from (3.28), and the lower bounds from the minimum principle result Proposition 3.2.1. □

We now possess enough machinery to begin our proof of Theorem 3.4.

3.7 Non-Positive Average Ricci Scalar

This section aims to prove the cases of Theorem 3.4 where $r < 0$ and $r = 0$, respectively. Thankfully, most of the ground work was done in the previous section.

As mentioned in Section 3.5, we would like to establish uniform bounds for the metrics $g(t)$. The nice evolution equation (3.23) for our potential function f allows us to state the following result, bounding our metrics $g(t)$ in the case of a non-positive average Ricci scalar.

Proposition 3.7.1. *When $r \leq 0$ for a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow, there exists a constant $C \geq 1$ dependent only on g_0 such that for as long as the solution exists we have*

$$\frac{1}{C}g_0 \leq g(t) \leq Cg_0.$$

The metrics $g(t)$ are thus all called ***uniformly equivalent****.*

Proof. Using the definition of f (3.19) and its evolution equation (3.23), we have

$$\partial_t g = (r - R)g = (\Delta f)g = (\partial_t f - rf)g.$$

Integrating both sides in time, we obtain

$$g(x,t) = \exp\left(f(x,t) - f(x,0) - r\int_0^t f(x,t')\mathrm{d}t'\right) g(x,0).$$

Writing C' for the bound for f in (3.24), we have for any vector V,

$$\begin{aligned}\left|\log\left(\frac{g_{(x,t)}(V,V)}{g_{(x,0)}(V,V)}\right)\right| &\leq \left|f(x,t) - f(x,0) - r\int_0^t f(x,t')\mathrm{d}t'\right| \\ &\leq 2C'e^{rt} - r\frac{C'}{r}e^{rt'}\Big|_{t'=0}^{t'=t} \\ &= C'\left(e^{rt} + 1\right).\end{aligned}$$

From this we can conclude that there exists the desired C. □

Now, we consider the case $r < 0$. We know that

1. By Proposition 3.4.1, the solution $g(t)$ exists for all time;

2. By Proposition 3.7.1, the metrics $g(t)$ are uniformly equivalent;
3. By Proposition 3.6.3, there exists a constant C dependent only on the initial metric g_0 such that

$$|R - r| \leq Ce^{rt},$$

which means R is approaching r exponentially quickly.

All that remains to show is that the convergence is uniform in any C^k norm, as claimed in the following result.

Proposition 3.7.2. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r < 0$. Then for every integer $k > 0$ there exists a constant $C_k < \infty$ such that for all time $0 < t < \infty$ we have*

$$\sup_{x\in\Sigma} \left|\nabla^k R(x,t)\right|^2 \leq C_k e^{\frac{1}{2}rt}.$$

Proof. We will prove this using induction on k. The base case is implied by the following Lemma, which provides yet another evolution equation.

Lemma 3.7.1. *For any solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow, we have*

$$\partial_t |\nabla R|^2 = \Delta|\nabla R|^2 - |\nabla\nabla R|^2 + (4R - 3r)|\nabla R|^2. \tag{3.30}$$

Proof of Lemma. Using our favourite identity (3.21) and the evolution equation for R (3.18), we have

$$\partial_t(\nabla R) = \nabla(\Delta R + R(R-r)) = \Delta\nabla R + \frac{3}{2}R\nabla R - r\nabla R.$$

Using the evolution equation for g^{ij} (3.6), we then obtain

$$\begin{aligned}\partial_t|\nabla R|^2 &= \partial_t \left(g^{ij}\nabla_i R\nabla_j R\right) \\ &= (R-r)|\nabla R|^2 + 2\left\langle \Delta\nabla R + \frac{3}{2}R\nabla R - r\nabla R, \nabla R\right\rangle.\end{aligned}$$

We obtain the desired result by using R in the identity (3.22) and rearranging. □

This evolution equation allows us to conclude that the desired constant C_1 exists via

$$\begin{aligned}\partial_t|\nabla R|^2 &= \Delta|\nabla R|^2 - |\nabla\nabla R|^2 + (4R-3r)|\nabla R|^2 \\ &\leq \Delta|\nabla R|^2 - |\nabla\nabla R|^2 + \left(r + 4Ce^{rt}\right)|\nabla R|^2 \\ &\leq \Delta|\nabla R|^2 + \frac{1}{2}r|\nabla R|^2,\end{aligned}$$

for t large enough. An application of the maximum principle then gives the existence of C_1.

Now that the base case is established, the general case (assuming for $1 \leq j \leq k-1$) is obtained by using commutators of the sort

$$\left(\nabla^k\Delta - \Delta\nabla^k\right)R = \sum_{j=0}^{\lfloor k/2\rfloor} (\nabla^j R) \otimes_g (\nabla^{k-j}R),$$

where $\lfloor\cdot\rfloor$ is the floor function and $\otimes_g$ denotes a finite linear combination of contractions of tensors with respect to $g(t)$. See Proposition 5.27 of [CK04] for the full proof. □

Next, we tackle the case with vanishing average Ricci scalar. Same as before, we know that the solution exists for all time and that the metrics $g(t)$ are uniformly equivalent. Thus we have proved the case if we show that the Ricci scalar and all of its derivatives vanish as $t \to \infty$.

We recall that in the case $r = 0$, the potential function satisfies

$$\Delta f = R \quad \text{and} \quad \partial_t f = \Delta f.$$

That is, the potential satisfies the heat equation. As time nears infinity, its 'distribution' will average out so that it becomes constant in space. Thus, its Laplacian will vanish, and so will the Ricci scalar, as desired. We will keep this argument in mind as we explicitly show that R vanishes. In this direction, we want to give decaying bounds for $|\nabla f|^2$ and $R + |\nabla f|^2$, whose combination will give the desired bound for the Ricci scalar.

Proposition 3.7.3. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with Ricci scalar R and average Ricci scalar $r = 0$. Then, there exist constants $C_1, C_2 < \infty$ dependent only on g_0 such that for all time $0 \leq t < \infty$, the potential function f satisfies*

1.

$$\sup_{x \in \Sigma} |\nabla f(x,t)|^2 \leq \frac{C_1}{1+t};$$

2.

$$\sup_{x \in \Sigma} \left(|R(x,t)| + 2\,|\nabla f(x,t)|^2 \right) \leq \frac{C_2}{1+t}.$$

Consequently, the Ricci scalar R tends to 0 as $t \to \infty$.

Proof. 1. Applying the maximum principle to the evolution equation for $|\nabla f|^2$ (3.25), we have the existence of some constant C_1' dependent on g_0 such that $|\nabla f|^2 \leq C_1'$ for all time. Next, consider

$$\partial_t \left(t|\nabla f|^2 \right) = |\nabla f|^2 + t \left(\Delta |\nabla f|^2 - 2|\nabla\nabla f|^2 \right) \leq \Delta \left(t|\nabla f|^2 \right) + |\nabla f|^2. \tag{3.31}$$

Since $\Delta(f^2) = 2f\Delta f + 2|\nabla f|^2$, we have

$$\partial_t \left(f^2 \right) = 2f\Delta f = \Delta \left(f^2 \right) - 2|\nabla f|^2. \tag{3.32}$$

Combining (3.31) and (3.32) gives

$$\partial_t \left(t|\nabla f|^2 + f^2 \right) \leq \Delta \left(t|\nabla f|^2 + f^2 \right),$$

and so the maximum principle yet again gives the existence of some C_1'' (dependent on g_0) such that

$$t|\nabla f|^2 + f^2 \leq C_1''.$$

From this we deduce $|\nabla f|^2 \leq C_1''/t$ for non-zero t, and so combining this with the previous estimate (C_1') we obtain the desired C_1;

2. For convenience, we make the following definition:

$$\Xi := R + 2|\nabla f|^2.$$

Next, using the evolution equations for R (3.18) and $|\nabla f|^2$ (3.25), we have

$$\partial_t \Xi = \Delta \Xi + R^2 - 4|\nabla\nabla f|^2 \leq \Delta \Xi - R^2,$$

where we have used

$$R^2 = (\Delta f)^2 \leq 2|\nabla\nabla f|^2.$$

In the same spirit as the proof of the first point, we have

$$\partial_t (t\Xi) \leq R + 2|\nabla f|^2 + \Delta (t\Xi) - tR^2, \tag{3.33}$$

where our goal is to reduce this inequality to only the temporal derivative on the left-hand side and the Laplacian term on the right-hand side. Seemingly complicating our lives for the moment, we write

$$-tR^2 = -\frac{1}{2}tR^2 - \frac{1}{2}t\Xi^2 + 2t|\nabla f|^2 \left(R + |\nabla f|^2\right). \tag{3.34}$$

We remark that by the first point in this proposition, there exists some $c > 0$ such that $t|\nabla f|^2 \leq c$. Thus, inspecting the terms of the right-hand sides of (3.33) and (3.34) that do not contain Ξ, we have

$$\begin{aligned} -\frac{1}{2}tR^2 + R + 2|\nabla f|^2 + 2t|\nabla f|^2 \left(R + |\nabla f|^2\right) &\leq -\frac{1}{2}tR^2 + (1+2c)R + 2(1+c)|\nabla f|^2 \\ &\leq -\left(\sqrt{\frac{t}{2}}R - \frac{1}{\sqrt{2t}}(1+2c)\right)^2 + \frac{c'}{t}, \end{aligned}$$

for some $c' > 0$. Putting (3.33) and (3.34) together with the previous inequality, we obtain

$$\partial_t (t\Xi) \leq \Delta (t\Xi) - \frac{1}{2}t\Xi^2 - \left(\sqrt{\frac{t}{2}}R - \frac{1}{\sqrt{2t}}(1+2c)\right)^2 + \frac{c'}{t}.$$

From this we conclude that there is some $c'' > 0$ large enough that if $t\Xi \geq c''$, then

$$\partial_t (t\Xi) \leq \Delta (t\Xi),$$

which, by the maximum principle, implies that $\Xi = R + 2|\nabla f|^2$ has the desired decaying bound. Thus the desired result on the supremum follows, since $|\nabla f|^2$ is uniformly bounded by the proof of the first point, and R is bounded by Proposition 3.6.3. □

Next, we have the following result dealing with the derivatives of the Ricci scalar, whose proof we skip over in the name of brevity.

Proposition 3.7.4. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r = 0$. Then for every integer $k > 0$ there exists a constant $C_k < \infty$ such that for all time $0 < t < \infty$ we have*

$$\sup_{x\in\Sigma} \left|\nabla^k R(x,t)\right|^2 \leq \frac{C_k}{(1+t)^{k+2}}.$$

Idea of the proof. (See Proposition 5.33 of [CK04] for details.)
Similarly to Proposition 3.7.2, this is done with induction. □

With the above results we can conclude the desired Theorem 3.4 in the case of vanishing average Ricci scalar. In the next six sections we tackle the most difficult case: positive average Ricci scalar.

3.8 Positive Average Ricci Scalar I: Introduction

The following sections provide the proof of the final part of Theorem 3.4, where $r > 0$. New quantities will need to be defined, and we will be moving back and forth between two cases, where the initial Ricci scalar is either non-negative or is of mixed sign (depending on $x \in \Sigma$).

We state the following notation:

$$R_{\min}(t) := \inf_{x\in\Sigma} R(x,t) \quad \text{and} \quad R_{\max}(t) := \sup_{x\in\Sigma} R(x,t).$$

Now, consider the case where the initial Ricci scalar $R(x,0)$ is non-negative for all $x \in \Sigma$. For notational convenience, we will write $R(\cdot,0) \geq 0$ to signify a non-negative Ricci scalar. Note that in this case, $R_{\min,0}$ is either zero or greater than zero. By the strong maximum principle, if $R_{\min,0}$ is zero then some small time $\varepsilon > 0$ later it will be positive, unless $R_{\min,0} = R = r = 0$, in which case it is zero for all time (a case we have already treated). So, unless $R = 0$, $R_{\min}(t) > 0$ for all $t > 0$, which means that up to restarting the flow after some small time $\varepsilon > 0$, we may assume $R(\cdot,0) > 0$.

Following the strategy outlined in Section 3.5, we first wish to obtain reasonable bounds for the metric. We thus have the following result, which gives bounds for the metric on some time interval.

Lemma 3.8.1. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$ and, for some $t_0 \geq 0$ to be specified, write*

$$I := \left[t_0, t_0 + \frac{1}{2R_{\max}(t_0)}\right].$$

Then,

1. *For any $t_0 \in [0,\infty)$, the estimate*
$$g(x,t) \geq \frac{1}{\mathrm{e}} g(x,t_0),$$
holds for all $x \in \Sigma$ and for all $t \in I$;

2. *If $R(\cdot,0) \geq 0$, then for any times $0 \leq t_0 \leq t < \infty$ we have*
$$g(x,t) \leq \mathrm{e}^{r(t-t_0)} g(x,t_0);$$

3. *If $R(\cdot,0)$ changes sign, then for any times $0 \leq t_0 \leq t < \infty$ we have*
$$g(x,t) \leq \left(\mathrm{e}^{r(t-t_0)} \frac{\left(1-\frac{r}{s_0}\right) - \mathrm{e}^{-rt}}{\left(1-\frac{r}{s_0}\right) - \mathrm{e}^{-rt_0}}\right) g(x,t_0),$$

where $s(t)$ is the solution to the equation (3.13) *with initial condition*

$$s_0 := s(0) = \begin{cases} 0 & \text{if } R_{\min,0} \geq 0, \\ R_{\min,0} & \text{if } R_{\min,0} < 0. \end{cases}$$

That is,

$$s(t) = \begin{cases} 0 & \text{if } R_{\min,0} \geq 0, \\ \frac{r}{1-\left(1-\frac{r}{R_{\min,0}}\right)e^{rt}} & \text{if } R_{\min,0} < 0. \end{cases}$$

In particular, the first two points give that if $R(\cdot, 0) \geq 0$, then for any $t_0 \in [0, \infty)$ the estimate

$$\frac{1}{\mathrm{e}} g(x, t_0) \leq g(x, t) \leq \sqrt{\mathrm{e}} g(x, t_0),$$

holds for all $x \in \Sigma$ and for all $t \in I$.

Proof. We will use the following Lemma in the proof of the first point.

Lemma 3.8.2. *Under the assumptions of Lemma 3.8.1, for any $t_0 \in [0, \infty)$, the estimate*

$$R_{\max}(t) \leq 2R_{\max}(t_0),$$

holds for all $x \in \Sigma$ and for all $t \in I$.

Proof of Lemma. Since $r > 0$, $R_{\max}(t) > 0$ for all time, and so the evolution equation for R (3.18) gives that at a maximum (in space):

$$\partial_t R \leq \Delta R + R^2.$$

The solution of

$$\begin{cases} \partial_t \alpha = \alpha^2 \\ \alpha(t_0) = R_{\max}(t_0) \end{cases} \qquad \text{is} \qquad \alpha(t) = \frac{1}{R_{\max}^{-1}(t_0) + t_0 - t},$$

so by the maximum principle we have our desired estimate within I. □

Now, we consider each point in turn.

1. From the Ricci flow equation (3.3), we write

$$g(x, t) = \exp\left(\int_{t_0}^{t} \left(r - R(x, t')\right) \mathrm{d}t' \right) g(x, t_0).$$

If $t \in I$, then Lemma 3.8.2 gives

$$\int_{t_0}^{t} \left(r - R(x, t')\right) \mathrm{d}t' \geq -\int_{t_0}^{t} R(x, t')\mathrm{d}t' \geq -2\int_I R_{\max}(t_0)\mathrm{d}t' = -1,$$

from which we conclude the result;

2. The explicit form of $s(t)$ gives

$$\int_{t_0}^{t}\left(r - R(x,t')\right)\mathrm{d}t' \geq \int_{t_0}^{t}\left(r - s(t')\right)\mathrm{d}t',$$

and so if $R_{\min,0} \geq 0$, then

$$g(x,t) \leq \exp\left(\int_{t_0}^{t}\left(r - s(t')\right)\mathrm{d}t'\right) g(x,t_0) \leq \mathrm{e}^{r(t-t_0)} g(x,t_0),$$

as desired;

3. If $R_{\min,0} < 0$, then we integrate our known expression for $s(t)$:

$$-\int_{t_0}^{t} s(t')\mathrm{d}t' = \int_{t_0}^{t} \frac{r}{1-\left(1-\frac{r}{R_{\min,0}}\right)\mathrm{e}^{rt'}}\mathrm{d}t' = \log\left(\mathrm{e}^{-rt'} - \left(1 - \frac{r}{R_{\min,0}}\right)\right)\Bigg|_{t'=t_0}^{t'=t}.$$

Using $s_0 = R_{\min,0}$, we have

$$g(x,t) \leq \exp\left(\int_{t_0}^{t}\left(r - s(t')\right)\mathrm{d}t'\right) g(x,t_0) \leq \left(\mathrm{e}^{r(t-t_0)} \frac{\left(1-\frac{r}{s_0}\right) - \mathrm{e}^{-rt}}{\left(1-\frac{r}{s_0}\right) - \mathrm{e}^{-rt_0}}\right) g(x,t_0),$$

as desired.

The final conclusion arises by observation upon combining the first two points. □

The following section inspects gradient Ricci solitons, which we will see to be an intermediate step on our way to constant curvature: we will show that (a modified version of) the normalised Ricci flow with $r > 0$ tends metrics to become gradient Ricci solitons, which (as we will see shortly) have constant curvature.

3.9 Positive Average Ricci Scalar II: Gradient Ricci Solitons

It took new innovations to understand the case of positive average Ricci scalar. One of these was to consider gradient Ricci solitons. As we will see, when $r > 0$ they are always of constant curvature, and thus can be thought of as attractor basins for the normalised Ricci flow: metrics will want to tend towards them, and if the flow begins at a gradient Ricci soliton then it will remain static. In the following section, we will define a quantity called the entropy which will describe the tendency of the normalised Ricci flow with $r > 0$ to turn metrics into gradient Ricci solitons.

To begin, recall the equation defining gradient Ricci solitons for a potential f:

$$(R - r)g_{ij} = 2\nabla_i\nabla_j f. \tag{3.35}$$

By combining this with the defining equation of f (3.19), this equation is equivalent to the vanishing of the following quantity.

Definition 3.9.1 (M tensor)**.** For f a potential function, let $M^f = M$ be the following symmetric 2-tensor:[32]

$$M_{ij}^f = M_{ij} := \nabla_i\nabla_j f - \frac{1}{2}(\Delta f)g_{ij}. \tag{3.36}$$

[32] This is the trace-free part of the **Hessian** of f. We have dropped the f from our notation for succinctness. Recall too that we have encountered this before, in (3.27).

As can be verified by direct computation, it satisfies the following identity:

$$|M|^2 = |\nabla\nabla f|^2 - \frac{1}{2}(\Delta f)^2, \tag{3.37}$$

which is a restatement of (3.29).

This M tensor will allow us to encode the notion of a certain metric being a gradient Ricci soliton by simply noting when it vanishes. Eventually, we will consider the following Ricci-DeTurck flow (with DeTurck vector ∇f), first encountered in Example 3.1.2:

$$\partial_t g_{ij} = 2M_{ij} = 2\nabla_i\nabla_j f - (R-r)g_{ij} = (r-R)g_{ij} + \left(\mathcal{L}_{\nabla f} g\right)_{ij}.$$

We will call it the **gradient Ricci-DeTurck flow**. Note that its fixed points are gradient Ricci solitons, and so similarly to the right-hand side of the usual normalised Ricci flow (3.3) vanishing as $t \to \infty$ and forming a constant curvature metric, when the right-hand side of the gradient Ricci-DeTurck flow vanishes the metric will be a gradient Ricci soliton. The following result claims that gradient Ricci solitons (when $r > 0$) are of constant curvature.

Proposition 3.9.1. *All gradient Ricci solitons on closed Riemannian* 2*-manifolds with $r > 0$ have constant curvature.*[33]

Proof. Comparing the gradient Ricci soliton equation (3.35) to the conformal Killing equation

$$\nabla_i V_j + \nabla_j V_i = \lambda g_{ij},$$

we notice that ∇f is a Killing vector for $\lambda = R - r$. Letting J be the almost complex structure describing a counter-clockwise rotation about the origin by $\frac{\pi}{2}$,[34] it follows that $J(\nabla f)$ is a Killing vector:

$$\nabla_k\left(J^i_j\nabla_i f\right) + \nabla_j\left(J^i_k\nabla_i f\right) = J^i_j\nabla_k\nabla_i f + J^i_k\nabla_j\nabla_i f = \frac{1}{2}(R-r)(J_{jk} + J_{kj}) = 0,$$

by anti-symmetry. We then state the following Lemma.

Lemma 3.9.1. *If (Σ, g) is a Riemannian* 2*-manifold with non-zero Killing vector V such that V vanishes at $q \in \Sigma$, then (Σ, g) is rotationally symmetric about q.*

Proof of Lemma. Consider the one-parameter family of isometries generated by V: that is, $\psi_\lambda : \Sigma \to \Sigma$ with $\lambda \in \mathbb{R}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\psi_\lambda(p) = V\left(\psi_\lambda(p)\right) \quad \text{and} \quad \psi_0 = \mathbb{1}.$$

By construction, q is a fixed point of ψ_λ for all $\lambda \in \mathbb{R}$, and so the push-forward

$$\left(\psi_\lambda(q)\right)_* : \mathcal{T}_q\Sigma \to \mathcal{T}_q\Sigma,$$

[33] The proof of this fact often calls upon the Uniformisation Theorem—as in Proposition 5.21 of [CK04], for example, where it uses the Uniformisation-dependent Kazdan-Warner identity, or in Theorem 10.1 of [Ham88]—which we obviously want to avoid!

[34] This was introduced in Definition 2.5. Concretely, this almost complex structure is defined by

$$J(\partial_i) = J^j_i\partial_j \quad \text{where} \quad J(\partial_x) = \partial_y \quad \text{and} \quad J(\partial_y) = -\partial_x,$$

from which it is clear that $J^2 = -\mathbb{1}$, as is desired for an almost complex structure.

is an oriented isometry. Since the manifold is 2-dimensional, the oriented isometry group of $(\mathcal{T}_q\Sigma, g_q)$ is SO(2) $\cong \mathbb{S}^1$. We have two cases: either the assignment $\lambda \mapsto (\psi_\lambda(p))_*$ is the trivial homomorphism, in which case all maps $(\psi_\lambda(p))_*$ are zero maps. This is not the case, since our Killing vector is non-trivial, so $\lambda \mapsto (\psi_\lambda(p))_*$ is a non-trivial homomorphism, and thus (since the isometry group is $\mathbb{S}^1$) we can find some $\lambda_0 > 0$ such that $(\psi_{\lambda_0})_* = (\psi_0)_*$.[35]

Now we argue that isometries are uniquely defined by their push-forwards at a point. Consider an isometry ψ whose push-forward $(\psi_p)_*$ is known for some $p \in \Sigma$. To determine $\psi(q)$ we consider the geodesic $\gamma_{p,W}$ (beginning at p with speed W) that extends from p to q:[36] we have $q = \exp_p(W)$. Applying ψ to both sides, we have

$$\psi(q) = \psi(\exp_p(W)) = \exp_{\psi(p)}((\psi_p)_*(W)) = \exp_p((\psi_p)_*(W))$$

where the second equality arises from ψ being an isometry. From this we can conclude that $(\psi_{\lambda_0})_* = (\psi_0)_*$ implies $\psi_{\lambda_0} = \psi_0$. Thus we have shown that a non-trivial isometric action of $\mathbb{S}^1$ on (Σ, g), so the manifold is rotationally symmetric. □

Since every gradient on a closed surface must vanish for some point, ∇f and hence (by anti-symmetry) $J(\nabla f)$ must vanish at some $q \in \Sigma$. If ∇f is trivial, then we have constant curvature; so, we assume that it is non-trivial and hence $J(\nabla f)$ is too. We can then apply this Lemma to our context and write our metric in its rotationally symmetric form as

$$g = \mathrm{d}\varrho^2 + h(\varrho)^2\mathrm{d}\theta^2 \quad \text{for} \quad \theta \in [0, 2\pi] \quad \text{and} \quad \varrho \in [0, A],$$

for some $A > 0$, where $h(\varrho)$ is positive on $(0, A)$ and vanishing at $\varrho = 0$ and $\varrho = A$ by compactness. We recall from Examples 2.1.5 and 2.2.4 that the (non-vanishing) Christoffel symbols and Ricci scalar of this metric take the form

$$\Gamma^{\varrho}_{\theta\theta} = -hh', \quad \Gamma^{\theta}_{\varrho\theta} = \Gamma^{\theta}_{\theta\varrho} = \frac{h'}{h} \quad \text{and} \quad R = -2\frac{h''}{h}, \tag{3.38}$$

where $(\cdot)'$ denotes a derivative with respect to ϱ.

For a positive average Ricci scalar (which we take to be $r = 2$ without loss of generality), the gradient Ricci soliton equation becomes

$$\left(-2\frac{h''}{h} - 2\right) g_{ij} = 2\nabla_i\nabla_j f,$$

where we can assume by rotational symmetry that the potential function is only radially dependent: $f = f(\varrho)$. Thus the right-hand side can be expanded and the equation reduces to coupled ordinary differential equations (as g is diagonal):

$$-\frac{h''}{h} - 1 = f'' \quad \text{and} \quad -h''h - h^2 = hh'f'. \tag{3.39}$$

Dividing the second equation by h^2 and equating the right-hand sides gives

$$\frac{f''}{f'} = \frac{h'}{h},$$

[35]Think of this λ_0 as having 'gone around the circle' $\mathbb{S}^1$ to return to the original point.

[36]This exists because Σ is complete. See Chapter 2 for details (in particular Definitions 2.1.5 and 2.1.6), with further elaboration in Appendix A.

from which we obtain $f' = ah$ for some constant a. Substituting this back into the left-most soliton equation of (3.39) and multiplying by hh', we obtain

$$-h'h'' - hh' = ahh'^2. \tag{3.40}$$

Integrating (3.40) along $[0, A]$ gives

$$-\frac{1}{2}h'^2\Big|_{\varrho=0}^{\varrho=A} - \frac{1}{2}h^2\Big|_{\varrho=0}^{\varrho=A} = a\int_0^A h(\varrho)h'(\varrho)^2 \mathrm{d}\varrho. \tag{3.41}$$

We required our manifold to be smooth everywhere. In particular, at $\varrho = 0$ and $\varrho = A$, which require

$$h(0) = h(A) = 0 \quad \text{and} \quad h'(0) = -h'(A).$$

Thus the left-hand side of our integrated equation (3.41) is zero, and so we can conclude by the positivity of the integrand of the right-hand side that a must also be zero. When substituted back into $f' = ah$, we find f to be constant, and thus the right-hand side of the left-most equation in (3.39) vanishes. From this we conclude that $h'' = -h$, which can be substituted into (3.38) to find a Ricci scalar constant and equal to 2, as desired.[37] □

As hinted at this beginning of this section, we know turn to defining a quantity which will encode the fact that the normalised Ricci flow tends metrics to be gradient Ricci solitons: the entropy.

3.10 Positive Average Ricci Scalar III: Entropies

We now turn to a new quantity: the entropy, defined for a positive Ricci scalar—though it will have an equivalent in the case where $R(\cdot, 0)$ changes sign. It will be constant on gradient Ricci solitons (hence on metrics of constant curvature by Proposition 3.9.1), and in any other case it will be decreasing.

Definition 3.10.1 (Entropy). For a closed 2-dimensional Riemannian manifold (Σ, g) with positive Ricci scalar $R > 0$, we define the **entropy**, or surface entropy, N to be

$$N := \int_\Sigma R \log R. \tag{3.42}$$

Note that typically entropies (integrals of a quantity multiplied by its logarithm) have a negative sign, and are shown to increase, as in thermodynamics. Here, for no particular reason other than to follow the literature, we have the opposite sign convention. Our goal will be the following result.

Proposition 3.10.1. *For a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with Ricci scalar $R(\cdot, 0) > 0$ (and thus with average Ricci scalar $r > 0$), the entropy N is decreasing unless $R(\cdot, 0) = r$, in which case N is constant.*

It will be directly implied by the following result, which gives an explicit expression for the time-derivative of the entropy.

[37] This can also be seen in the gradient Ricci soliton equation (3.35): since the right-hand side vanishes, our metric is Einstein and thus of constant curvature by Proposition 2.2.1 and the discussion that follows it.

Proposition 3.10.2. *For a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$ and initial Ricci scalar $R(\cdot, 0) > 0$, the entropy N is decreasing:*

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -\int_\Sigma \frac{|\nabla R + R\nabla f|^2}{R} - 2\int_\Sigma |M|^2 \leq 0,$$

where f is the potential function defined by (3.19) *and the M tensor is defined by* (3.36). *In particular, the entropy is decreasing unless $g(t)$ is a gradient Ricci soliton.*

Proof. We first state a Lemma, which gives an intermediate form for the derivative of the entropy.

Lemma 3.10.1. *Under the assumptions of the proposition,*

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -\int_\Sigma \frac{|\nabla R|^2}{R} + \int_\Sigma (R - r)^2.$$

Proof. Using the evolution equations of R (3.18) and the volume element $\mathrm{d}A$ (3.10), we have

$$\partial_t(R\mathrm{d}A) = \big(\partial_t R\big)\mathrm{d}A + R\big(\partial_t(\mathrm{d}A)\big) = (\Delta R + R(R-r))\mathrm{d}A + R(r-R)\mathrm{d}A = \Delta R \mathrm{d}A.$$

Using this and the evolution equation of R (3.18) once more, we have

$$\begin{aligned}
\frac{\mathrm{d}N}{\mathrm{d}t} &= \int_\Sigma \big(\partial_t \log R\big) R\mathrm{d}A + \int_\Sigma \log R\big(\partial_t(R\mathrm{d}A)\big) \\
&= \int_\Sigma (\Delta R + R(R-r)) + \int_\Sigma \log R\ \Delta R \\
&= \int_\Sigma R(R-r) - \int_\Sigma (\nabla \log R + 1)\, R \\
&= \int_\Sigma (R-r)^2 - \int_\Sigma \frac{|\nabla R|^2}{R},
\end{aligned}$$

where in the final step we have integrated by parts (using that Σ is closed so the integral of ∇R vanishes on it) and used

$$\int_\Sigma r(R-r) = 0, \tag{3.43}$$

by definition of r (3.1). □

We now make several further computations. Using the definition of f (3.19), integrating by parts and using our favourite identity (3.21), we obtain:

$$\begin{aligned}
\int_\Sigma (R-r)^2 = \int_\Sigma (\Delta f)^2 \\
&= -\int_\Sigma \langle \nabla f, \nabla \Delta f\rangle \\
&= -\int_\Sigma \left(\langle \nabla f, \Delta \nabla f\rangle - \frac{1}{2}R|\nabla f|^2\right) \\
&= \int_\Sigma \left(|\nabla\nabla f|^2 + \frac{1}{2}R|\nabla f|^2\right).
\end{aligned}$$

Using the identity satisfied by $|M|^2$ (3.37) and the previous computation we have

$$\int_\Sigma |M|^2 = \int_\Sigma \left(|\nabla\nabla f|^2 - \frac{1}{2}(\Delta f)^2\right) = \frac{1}{2}\int_\Sigma \left((R-r)^2 - R|\nabla f|^2\right).$$

Next, consider the first integral in our desired evolution equation of N. Using the definition of f (3.19) and integrating by parts we have

$$\begin{aligned}\int_\Sigma \frac{|\nabla R + R\nabla f|^2}{R} &= \int_\Sigma \left(\frac{|\nabla R|^2}{R} + R|\nabla f|^2 + 2\langle \nabla R, \nabla f\rangle \right) \\ &= \int_\Sigma \left(\frac{|\nabla R|^2}{R} + R|\nabla f|^2 - 2R(R-r) \right) \\ &= \int_\Sigma \left(\frac{|\nabla R|^2}{R} + R|\nabla f|^2 - 2(R-r)^2 \right),\end{aligned}$$

where in the final step we have used that R and r integrate to the same result as in (3.43).

Combining these with our Lemma, we obtain

$$0 \geq -\int_\Sigma \frac{|\nabla R + R\nabla f|^2}{R} - 2\int_\Sigma |M|^2 = -\int_\Sigma \frac{|\nabla R|^2}{R} + \int_\Sigma (R-r)^2 = \frac{\mathrm{d}N}{\mathrm{d}t}, \tag{3.44}$$

as desired.

The final claim is clear when noticing the positivity of the integrands in: if both integrands vanish, then equality is satisfied. So we want to show that the integrands vanish on gradient Ricci solitons. The numerator of the first integral is zero on gradient Ricci solitons because of the following short Lemma.

Lemma 3.10.2. *Consider a potential function f. When $R > 0$, the gradient Ricci soliton equation* (3.35) *implies that* $\log R + f$ *is a constant.*

Proof of Lemma. Because the gradient Ricci soliton equation implies that M vanishes, if we consider the divergence of M when $R > 0$, we have

$$\nabla^j M_{ji} = \nabla^j \left(\nabla_j \nabla_i f - \frac{1}{2} g_{ij} \Delta f \right) = \frac{1}{2}\left(\nabla_i \Delta f + R\nabla_i f \right) = \frac{1}{2}\left(R\nabla_i f + \nabla_i R \right),$$

where we have used the definitions of f (3.19) and M (3.36) as well as our identity (3.21). If we now assume that the divergence of M is zero, then

$$R\nabla_i f + \nabla_i R = 0 \iff \nabla\left(\log R + f\right) = 0,$$

so $\log R + f$ is constant in space. □

From this Lemma, we can conclude that on gradient Ricci solitons (where the curvature is constant by Proposition 3.9.1), $R\nabla(\log R + f)$ vanishes, which is the numerator of the first integral of (3.44). The second integrand is simply $|M|^2$, which vanishes on gradient Ricci solitons by definition of M (3.36). □

The equation for the change in entropy given by Proposition 3.10.2 is interesting for several reasons. Continuing or metaphor of (normalised) Ricci flow being a heat-like equation, constant entropy would correspond to thermal equilibrium, and so gradient Ricci solitons can be thought of as equilibria. This proposition also allows us to conclude the proof of Proposition 3.10.1 as follows.

Proof of Proposition 3.10.1. That the entropy is decreasing is clear from Proposition 3.10.2, unless it is constant—whence it is a gradient Ricci soliton and thus of constant curvature, and remains so for all time. If it becomes constant at some time $t_0 \in [0, \infty)$, then the M tensor is zero at t_0, and so $g(t_0)$ is a gradient Ricci soliton. By Proposition 3.9.1, at this time the curvature is constant ($R(\cdot, t_0) = r$) and so $g(t) = g(t_0)$ for all times $t > t_0$. The key point here is that as soon as the solution becomes a gradient Ricci soliton (and thus of constant curvature), it remains so for all time. □

The previous results only hold for the initially positive Ricci scalar, since the entropy is not defined for the case where $R(\cdot,0)$ changes sign. To resolve this, consider the following. For $r > 0$ and an initial condition[38] $s_0 := s(0) < R_{\min,0}$ we recall that the ordinary differential equation[39]

$$\partial_t s = s(s-r) \quad \text{has the solution} \quad s(t) = \frac{r}{1-\left(1-\frac{r}{s_0}\right)e^{rt}}. \tag{3.45}$$

We also have the following evolution equation:

$$\partial_t\big(R-s\big) = \Delta(R-s) + (R-r+s)(R-s).$$

Since initially $R_{\min,0} > s_0$, the maximum principle applied to the above gives $R > s$ for as long as the solution exists. This motivates the following definition.

Definition 3.10.2 (Modified entropy)**.** For R and s described as above on a closed Riemannian 2-manifold (Σ, g), we define the **modified entropy**, or modified surface entropy, $\hat{N}$ to be

$$\hat{N}(g,s) := \int_\Sigma (R-s)\log(R-s). \tag{3.46}$$

Unfortunately, unlike in the $R(\cdot,0) > 0$ case, the modified entropy is not decreasing in general. Instead, we will show that it is bounded from above, as stated in the following result.

Proposition 3.10.3. *For a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$, there exists a constant $C \in \mathbb{R}$ depending only on g_0 such that the modified entropy $\hat{N}$ satisfies*

$$\hat{N}(g,s) \le C,$$

for $s(t)$ the solution (3.45).

Proof. This will be proved in several steps. First, we have the following Lemma, which gives a form of the derivative of the modified entropy.

Lemma 3.10.3. *Under the assumptions of Proposition 3.10.3, the modified entropy $\hat{N}$ satisfies the following evolution equation:*

$$\frac{d\hat{N}}{dt} = -\int_\Sigma \frac{|\nabla R + (R-s)\nabla f|^2}{R-s} - 2\int_\Sigma |M|^2 - s\int_\Sigma \left(|\nabla f|^2 + s - r - (R-s)\log(R-s)\right), \tag{3.47}$$

where f is the potential function and the M tensor is defined by (3.36).

Proof of Lemma. We will prove this in a similar fashion to our proof of Proposition 3.10.2. Writing $Y := R - s$ and using the evolution equations for R (3.18), dA (3.10) and s (3.45), the quantity $Y dA$ has the following evolution equation:

$$\partial_t\big(Y dA\big) = (\Delta Y + (R-r+s)Y)dA + Y(r-R)dA = (\Delta Y + sY)dA.$$

[38] For clarity: this initial condition is required for our desire for the quantity $R-s$ to be positive, so that it can take the role of R in our previous definitions, where $R(\cdot,0) > 0$.

[39] As is standard in the literature, we have used s instead of α. This is to remind us that s is to be used only under our current assumptions, whereas α is for any application of the maximum principle.

This allows us to write

$$\begin{aligned}
\frac{\mathrm{d}\hat{N}}{\mathrm{d}t} &= \int_\Sigma (\partial_t \log Y) Y \mathrm{d}A + \int_\Sigma \log Y \big(\partial_t (Y \mathrm{d}A)\big) \\
&= \int_\Sigma (\Delta Y + (R - r + s) Y) + \int_\Sigma \log Y (\Delta Y + sY) \\
&= \int_\Sigma (R - r + s + s \log Y) Y - \int_\Sigma (\nabla \log Y + 1) \nabla Y \\
&= \int_\Sigma (R - r + s + s \log Y) Y - \int_\Sigma \frac{|\nabla Y|^2}{Y} \\
&= \int_\Sigma (R - r + s + s \log (R - s))(R - s) - \int_\Sigma \frac{|\nabla R|^2}{R - s},
\end{aligned}$$

where we integrated by parts twice and used that Σ is closed to ignore boundary terms. We can also expand the first two terms of our desired form (3.47) as

$$-\int_\Sigma \frac{|\nabla R + (R - s)\nabla f|^2}{R - s} = -\int_\Sigma \left(\frac{|\nabla R|^2}{R - s} - 2R(R - r) + (R - s)|\nabla f|^2 \right),$$

and

$$-2 \int_\Sigma |M|^2 = \int_\Sigma \left(R |\nabla f|^2 - R(R - r) \right).$$

Combining these three equations gives the desired result. □

The only term in our evolution equation for $\hat{N}$ that is positive (and hence likely to cause problems when we attempt to bound $\hat{N}$) is the one containing $-s|\nabla f|^2$, since $s < 0$. To control this term, we have the following Lemma.

Lemma 3.10.4. *Under the assumptions of Proposition 3.10.3 (where we assume the solution $g(t)$ to exist for times $0 \leq t < T$ for some $T > 0$), there exists a constant C' depending only on g_0 such that*

$$\int_0^T \mathrm{e}^{-rt} \int_\Sigma |\nabla f|^2 \mathrm{d}A \, \mathrm{d}t \leq C',$$

for f the potential function.

Proof of Lemma. Using the evolution equations of f (3.23) and $\mathrm{d}A$ (3.10), we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \left(\mathrm{e}^{-rt} \int_\Sigma f \, \mathrm{d}A \right) &= -r\mathrm{e}^{-rt} \int_\Sigma f \, \mathrm{d}A + \mathrm{e}^{-rt} \int_\Sigma \big((\partial_t f) \mathrm{d}A + f \partial_t (\mathrm{d}A) \big) \\
&= -r\mathrm{e}^{-rt} \int_\Sigma f + \mathrm{e}^{-rt} \int_\Sigma (\Delta f + rf + f(r - R)) \\
&= \mathrm{e}^{-rt} \int_\Sigma (\Delta f - f \Delta f) \\
&= \mathrm{e}^{-rt} \int_\Sigma |\nabla f|^2,
\end{aligned}$$

where in the last line we have integrated by parts. We then integrate this along $[0, T]$ and use the bound on f from (3.24) to find

$$\int_0^T \mathrm{e}^{-rt} \int_\Sigma |\nabla f|^2 \mathrm{d}A \, \mathrm{d}t = \left(\mathrm{e}^{-rt} \int_\Sigma f \, \mathrm{d}A \right) \Bigg|_{t=0}^{t=T} \leq C',$$

as desired. □

Now, from the explicit form of $s(t)$ (3.45), there exists a $C'' > 0$ depending only on g_0 such that $-C''e^{-rt} \leq s \leq 0$. From this fact and the evolution equation for $\hat{N}$ (3.47), we have

$$\begin{aligned}\frac{d\hat{N}}{dt} &\leq -s \int_\Sigma \left(|\nabla f|^2 + s - r - (R-s)\log(R-s)\right) \\ &\leq C''e^{-rt} \int_\Sigma |\nabla f|^2 + C''e^{-rt} \left|\hat{N}\right|.\end{aligned}$$

Integrating this inequality, we obtain

$$\hat{N} \leq \int_0^T \left(C''e^{-rt} \int_\Sigma |\nabla f|^2 + C''e^{-rt} \left|\hat{N}\right| \right) dt \leq C,$$

where we have used the Lemma in the final inequality and have taken advantage of the positivity of $|\hat{N}|$. □

The entropy estimates of this past section encode the tendency of the normalised Ricci flow to make a metric become closer to a metric whose entropy is unchanging. They will also allow us to obtain uniform upper bounds for the Ricci scalar, as we shall soon see. Before turning once more to evolution equations and maximum principles, we will state bounds for the Ricci scalar, the diameter and the injectivity radius.

3.11 Positive Average Ricci Scalar IV: Key Bounds

This section will present upper bounds for the Ricci scalar and the diameter, both in the case of a positive initial Ricci scalar and in general. The section will then close with a lower bound for the injectivity radius, which we will not prove. We begin with a result that does not depend on the initial sign of R.

Lemma 3.11.1. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r \geq 0$ and Ricci scalar $|R(\cdot, 0)| \leq k$ for some constant $k > 0$. Then there exists a constant $C < \infty$ such that the estimate*

$$\sup_{x \in \Sigma} |\nabla R|(x,t) \leq C \frac{k}{\sqrt{t}},$$

holds for all times $0 < t \leq 1/(kC)$.

Proof. For convenience, define

$$\zeta := t|\nabla R|^2 + R^2.$$

We will want to use the maximum principle on its evolution equation. Using the evolution equation of R (3.18), we find

$$\partial_t \left(R^2\right) = 2R\partial_t R = 2R\left(\Delta R + R(R-r)\right) = \Delta\left(R^2\right) - 2|\nabla R|^2 + 2R^2(R-r). \tag{3.48}$$

Using the evolution equation of $|\nabla R|^2$ (3.30),

$$\partial_t \left(t|\nabla R|^2\right) = \Delta\left(t|\nabla R|^2\right) - 2t|\nabla\nabla R|^2 + \left(t(4R-3r)+1\right)|\nabla R|^2. \tag{3.49}$$

Adding (3.48) and (3.49), we have

$$\partial_t \zeta \leq \Delta\zeta + (4tR - 1)|\nabla R|^2 + 2R^3. \tag{3.50}$$

To be able to properly apply the maximum principle on this evolution equation, we will need to estimate $|R|$ on some time interval.

At a (spatial) minimum of R, we have

$$\partial_t R \geq R(R-r) \quad \Longrightarrow \quad R_{\min}(t) \geq \min\{0, R_{\min,0}\} \geq -k.$$

On the other hand, at a maximum,

$$\partial_t R \leq R^2,$$

since $R_{\max} \geq 0$. By inspection for the minimum case and by the maximum principle applied to the maximum case, we thus have

$$|R| \leq \frac{k}{1-kt} \quad \text{for} \quad t \in \left[0, \frac{1}{k}\right). \tag{3.51}$$

By noting that the maximum of the right-hand side of (3.51) on the interval $0 \leq t \leq \frac{1}{2k}$ occurs at $\frac{1}{2k}$ (since the function is increasing), we have $|R| \leq 2k$ on $0 \leq t \leq \frac{1}{2k}$. This implies $4t|R| \leq 1$ while $0 \leq t \leq \frac{1}{8k}$, so that our gradient term in (3.50) can be ignored and we can write

$$\partial_t \zeta \leq \Delta\zeta + (4tR-1)|\nabla R|^2 + 2R^3 \leq \Delta\zeta + 16k^3 \quad \text{for} \quad t \in \left[0, \frac{1}{8k}\right].$$

Finally, noting that at $t = 0$ we have $\zeta_0 = R^2 \leq k^2$, and using the maximum principle yet again, we have

$$\zeta \leq k^2 + 16k^3 t \leq k^2 + \frac{16k^3}{8k} = 3k^2 \quad \text{for} \quad t \in \left[0, \frac{1}{8k}\right].$$

Rearranging this gives the desired inequality. □

Using this, we can find an upper bound for the Ricci scalar. In the case of $R(\cdot, 0) \geq 0$, recall that (by the strong maximum principle) up to restarting the flow after some small time $\varepsilon > 0$, we may assume $R(\cdot, 0) > 0$. Then for all time, R is bounded from below by 0, and by Proposition 3.6 we have

$$R \leq Ce^{rt},$$

for a constant $C > 0$ depending only on the initial metric g_0. Before stating our upper bounds on R and on the diameter of our manifold, we state the following result (sometimes called Klingenberg's Lemma) without proof, which will be used in coming proofs.

Lemma 3.11.2 (Klingenberg). *Consider a Riemannian manifold (Σ, g) such that the Ricci curvature is bounded*[40] *by some constant $c > 0$ in the following way:*

$$\mathrm{Ric}(V, V) \leq cg(V, V) \quad \textit{for all vectors } V.$$

Then, the injectivity radius defined in Definition 3.3.1 satisfies the following bound:

$$\mathrm{inj}(\Sigma, g) \geq \min\left\{\frac{\sqrt{2}\pi}{\sqrt{c}}, \textit{half the length of the shortest closed geodesic on } \Sigma\right\}.$$

[40]The original formulation applies to n-manifolds and requires that only the sectional curvature be bounded.

Idea of the proof. (See Theorem III.2.4 of [Cha06] for details.)

Consider a unit speed geodesic $\gamma(t)$ whose endpoints p and q are conjugate to one another (that is, there exists a non-trivial Jacobi field[41] along γ that vanishes at p and q). Then Theorem II.6.3 of [Cha06] states that the earliest t where this can occur is inversely proportional to the square-root of the bound (here called c) of the (sectional) curvature.

Denote $C(p)$ to be the set of all points such that there are multiple geodesics connecting them to p (this is called the **cut locus** of p, though this definition is not quite complete—see [Cha06]). Now, if q is the closest point in $C(p)$ to p, then if q is not conjugate to p along a geodesic connecting p and q then it is the mid-point of a closed geodesic beginning and ending at p (this can be proved by contradiction). Combining these facts proves the desired result. □

We can now state our bounds for the Ricci scalar and the diameter.

Lemma 3.11.3. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with initial Ricci scalar $R(\cdot, 0) > 0$ (and thus with average Ricci scalar $r > 0$). Then,*

1. *There exists a constant $C > 1$ depending only on g_0 such that*
$$\sup_{(x,t)\in\Sigma\times[0,\infty)} R(x,t) \leq C;$$
2. *There exists a constant $C' > 0$ depending only on g_0 such that*
$$\operatorname{diam}(\Sigma, g(t)) \leq C'.$$

Proof. 1. We make the following definitions:
$$k_1 := \max_{(x,t)\in\Sigma\times[0,1]} R(x,t) \quad \text{and} \quad k(T) := \max_{(x,t)\in\Sigma\times[0,T]} R(x,t) \geq k_1,$$
where $T > 1$. Both of these are well-defined since we have time-dependent bounds for R. Our goal is to show that $k(T)$ is bounded independently of T, from which we will conclude our result. We will often write $k = k(T)$ from here onward.

First, assume $k(T) > \max\{k_1, \frac{1}{4}\}$, so $T > 1$. Now, let $(x_1, t_1) \in \Sigma \times [0, \infty)$ be a point such that $R(x_1, t_1) = k$ (that is, it maximises R) and let $t_0 := t_1 - \frac{1}{4k} > 0$.

From the proof of Lemma 3.11.1, when $|R| \leq k$ for the interval $t_0 \leq t \leq t_1$,
$$|\nabla R(x,t)| \leq \frac{2k}{\sqrt{t - t_0}} \quad \text{for all} \quad (x,t) \in \Sigma \times (t_0, t_1].$$
In particular, at time $t = t_1$, we have
$$|\nabla R(x, t_1)| \leq 4k^{\frac{3}{2}} \quad \text{for all} \quad x \in \Sigma.$$
Write $\mathcal{B}$ for the ball centred at x_1 with radius $1/\sqrt{64k}$, computed using the metric $g(t_1)$,[42] and consider some $y \in \mathcal{B}$. Connect y to x_1 using a geodesic $\gamma(t')$ with unit

[41] Loosely, a **Jacobi field** describes the behaviour of geodesics in a neighbourhood of a given geodesic. See the definition on page 78 of [Cha06] directly preceding Theorem II.5.1, or Chapter VIII.1 of [KN69].

[42] Note that by our notation of (3.14),
$$\mathcal{B} := \mathcal{B}_{g(t_1), \frac{1}{\sqrt{64k}}}(x_1),$$
which is unwieldy.

speed. Then, by the Fundamental Theorem of Calculus,

$$R(x_1,t_1) - R(y,t_1) = \int_\gamma \partial_{t'} \left(R(\gamma(t'),t_1)\right) \mathrm{d}t' \leq \int_\gamma |\nabla R(\gamma(t'),t_1)| \mathrm{d}t' \leq \frac{4k^{\frac{3}{2}}}{8k^{\frac{1}{2}}} = \frac{k}{2},$$

so $R(y,t_1) \geq \frac{k}{2}$ for all $y \in \mathcal{B}$, since $R(x_1,t_1) = k$ by definition.

By the strong maximum principle, $R(\cdot,t_1) > 0$, so the entropy N is well-defined at t_1, and by Proposition 3.10.2 we can bound it from above by some $C > 0$. Using that the minimum of $R\log R$ is $-\mathrm{e}^{-1}$, we have

$$C \geq N(g(t_1)) := \int_\Sigma R\log R \geq \int_{\mathcal{B}} R\log R - \frac{1}{\mathrm{e}}\mathrm{vol}(\Sigma, g(t_1)).$$

Inspecting the first term, we have

$$\int_{\mathcal{B}} R\log R \geq \frac{k}{2}\log\left(\frac{k}{2}\right)\mathrm{vol}(\mathcal{B}, g(t_1)),$$

so that together these inequalities imply

$$C \geq \frac{k}{2}\log\left(\frac{k}{2}\right)\mathrm{vol}(\mathcal{B}, g(t_1)) - \frac{1}{\mathrm{e}}\mathrm{vol}(\Sigma, g(t_1)) \geq c\log\left(\frac{k}{2}\right),$$

where the constant $c > 0$ exists by volume comparison results.[43] Thus, k has a uniform upper bound and therefore R is bounded as desired.

2. Consider points $(p_i)_{i=1}^m$ on Σ such that

$$d_{g(t)}(p_i,p_j) \geq \frac{2\pi}{\sqrt{R_{\max}(t)}} \quad \text{for all} \quad 1 \leq i < j \leq m.$$

By Klingenberg's Lemma 3.11.2,

$$\mathrm{inj}(\Sigma, g(t)) \geq \frac{\sqrt{2}\pi}{\sqrt{R_{\max}(t)}},$$

so that the balls centred at p_i with radius $\sqrt{2}\pi/\sqrt{R_{\max}(t)}$, computed using the metric $g(t_1)$, are pairwise disjoint—call them $\mathcal{B}_i$. The volume comparison result mentioned in the proof of the previous point gives the existence of $\varepsilon > 0$ such that

$$\mathrm{vol}(\Sigma, g(t)) \geq \sum_{i=1}^m \mathrm{vol}(\mathcal{B}_i, g(t)) \geq m\frac{\varepsilon}{R_{\max}(t)},$$

which, when combined with the upper bound $C > 0$ for R obtained in the first point of this result, gives

$$m \leq \frac{R_{\max}(t)}{\varepsilon}\mathrm{vol}(\Sigma, g(t)) \leq C\mathrm{vol}(\Sigma, g(t)),$$

so the number of points is bounded since we have a compact manifold, and thus the diameter is finite. □

[43] See Chapter III.4 of [Cha06] for details.

Note that the upper bound for the Ricci scalar obtained in this result, combined with its time-dependent upper bound in (3.28) establish that the Ricci scalar will not blow-up during the normalised Ricci flow (at least when $R(\cdot, 0) > 0$, as this is what Lemma 3.11.3 grants us). As long as we can prove that the case $R(\cdot, 0)$ of mixed sign turns into the $R > 0$ case, we therefore have our long-time existence result Proposition 3.4.1, which depended on R being uniformly bounded. Now, before moving to the case of a general initial Ricci scalar, we state the following result without a full proof.

Lemma 3.11.4. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$. Denote the maximum (over all space) Ricci scalar by $R_{\max}(t)$. Then for all time $0 \le t < \infty$, the injectivity radius satisfies*

$$\operatorname{inj}(\Sigma, g(t)) \ge \min\left\{ \operatorname{inj}(\Sigma, g_0), \min_{t' \in [0,t]} \frac{\sqrt{2}\pi}{\sqrt{R_{\max}(t')}} \right\}.$$

Idea of the proof. (See Proposition 5.65 of [CK04] for details.)

Klingenberg's Lemma 3.11.2 states that a lower bound for the injectivity radius is the smaller of half the length of the shortest closed geodesic or $\sqrt{2}\pi/\sqrt{R_{\max}(t)}$. But as long as the length of the shortest closed geodesic is less than $2\sqrt{2}\pi/\sqrt{R_{\max}(t)}$, it will be increasing in time (and thus the claim follows). This can be proved by noting the following facts:

1. A closed geodesic of length less than $2\sqrt{2}\pi/\sqrt{R_{\max}(t)}$ is stable[44] when it is the shortest closed geodesic on the manifold (see Lemma 5.69 of [CK04]);
2. The integral of the Ricci scalar R over stable closed geodesics is non-positive (see Lemma 5.70 of [CK04]);
3. For a one-parameter family of closed geodesics γ_λ, the derivative with respect to λ of the length (at $\lambda = \mathring{\lambda}$) is proportional to the integral along $\gamma_{\mathring{\lambda}}$ of $r - R$ (see Lemma 5.71 of [CK04]), which gives a lower bound of this derivative (at $\lambda = \mathring{\lambda}$) in terms of the length of $\gamma_{\mathring{\lambda}}$ (see Corollary 5.72 of [CK04]);
4. For any shortest closed geodesic at a time t, one can find a shorter closed geodesic at any time $t' < t$, and so the length is increasing (see Lemma 5.73 of [CK04]). □

Now, we turn to the case where the initial Ricci scalar is of mixed sign. Here, we have similar bounds, proved with the help of our modified entropy estimates and the injectivity bound above.

Lemma 3.11.5. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$. Then,*

1. *There exists a constant $C > 0$ depending only on g_0 such that*

$$\sup_{(x,t) \in \Sigma \times [0,\infty)} R(x,t) \le C;$$

2. *There exists a constant $C' > 0$ depending only on g_0 such that*

$$\operatorname{diam}(\Sigma, g(t)) \le C'.$$

[44] A closed geodesic is **stable** if nearby closed geodesics have greater lengths. See Definition 5.68 of [CK04] for a precise definition.

Idea of the proof. (See Lemma 5.74 and Corollary 5.75 of [CK04] for details.)

Similarly to Lemma 3.11.3, this is proved by defining $k(T)$ and showing that it is bounded independently of T. Using the modified entropy estimate found in Proposition 3.10.3, the injectivity radius bound from Lemma 3.11.4, and volume comparisons results from Chapter III.4 of [Cha06] provide the upper bound for R. The diameter bound arises from a near-identical argument to the one in Lemma 3.11.3. □

This now grants us our desired bound on the Ricci scalar R, which we know will not blow up under any circumstance during the normalised Ricci flow. Thus, our long-time existence result Proposition 3.4.1 holds.

With upper bounds for the Ricci scalar and diameter and a lower bound for the injectivity radius in hand, we turn once more to evolution equations and the maximum principle. The following section introduces what are known as Harnack inequalities, which will be useful in finding uniform (positive) lower bound for our Ricci scalar, after which we will be able to conclude our desired Uniformisation Theorem.

3.12 Positive Average Ricci Scalar V: Harnack Inequalities

The goal of this section will be to bound the Ricci scalar at some point (x_2, t_2) from below in terms of the Ricci scalar at an earlier[45] point (x_1, t_1). Inequalities of this form are called **Harnack inequalities**, and are either in differential or classical (integrated) form. In the same spirit as our entropies, we will define new quantities and modified versions of these quantities, and their evolution equations will allow us to find Harnack inequalities described above.

Definition 3.12.1 (L and Q scalars). 1. For a positive Ricci scalar $R > 0$, we define L and Q to be the following scalar quantities:

$$L := \log R \quad \text{and} \quad Q := \Delta L + R - r; \tag{3.52}$$

2. Let $\hat{L}$ and $\hat{Q}$ be the following scalar quantities

$$\hat{L} := \log (R - s) \quad \text{and} \quad \hat{Q} := \Delta \hat{L} + R - r, \tag{3.53}$$

where $s(t)$ is the solution

$$s(t) = \frac{r}{1 - \left(1 - \frac{r}{s_0}\right) e^{rt}} \quad \text{to} \quad \begin{cases} \partial_t s = s(s - r), \\ s_0 < R_{\min,0} < 0. \end{cases} \tag{3.54}$$

To motivate the definition of Q (and its modified version), we note that on gradient Ricci solitons $Q = 0$, because of the following argument. Recall from the discussion following Lemma 3.10.2 (within the proof of Proposition 3.10.2) that on gradient Ricci solitons, the quantity $\nabla R + R\nabla f = 0$ is constant. Thus the divergence of its quotient by R must vanish:

$$0 = \nabla \left(\frac{\nabla R + R \nabla f}{R} \right) = \nabla \left(\nabla \log R + \nabla f \right) = \Delta L + R - r = Q.$$

As previously mentioned, quantities that are constant on solitons are useful in deriving bounds for other quantities, often via the maximum principle applied to their evolution

[45] That is, having $t_1 < t_2$.

equations. In the case of Q and $\hat{Q}$, this procedure will directly give us what are known as Harnack inequalities for the Ricci scalar R. Note that the above motivation works only for a positive Ricci scalar; the modified versions of L and Q are motivated by the definitions of our modified entropy.

The scalars satisfy the following evolution equations.

Proposition 3.12.1. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$. Now,*

1. *If $R(\cdot, 0) > 0$, then*

$$\partial_t L = \Delta L + |\nabla L|^2 + R - r, \tag{3.55}$$

$$\partial_t Q = \Delta Q + 2\langle \nabla Q, \nabla L\rangle + 2\left|\nabla\nabla L + \frac{1}{2}(R-r)g\right|^2 + rQ, \tag{3.56}$$

where L and Q are defined in Definition 3.12.1;

2. *If $R(\cdot, 0)$ is of mixed sign, then*

$$\partial_t \hat{L} = \Delta \hat{L} + |\nabla \hat{L}|^2 + R - r + s, \tag{3.57}$$

$$\partial_t \hat{Q} = \Delta \hat{Q} + 2\langle \nabla \hat{Q}, \nabla \hat{L}\rangle + 2\left|\nabla\nabla \hat{L} + \frac{1}{2}(R-r)g\right|^2 + s\left|\nabla \hat{L}\right|^2 + (r-s)\hat{Q} + s(R-r), \tag{3.58}$$

where $\hat{L}$ and $\hat{Q}$ are defined in Definition 3.12.1.

Proof. 1. Using the evolution equation of R (3.18), we have

$$\begin{aligned}\partial_t \log R &= \frac{1}{R}(\Delta R + R(R-r)) \\ &= \left(\frac{\Delta R}{R} - \frac{|\nabla R|^2}{R^2}\right) + \frac{|\nabla R|^2}{R^2} + R - r \\ &= \Delta \log R + |\nabla \log R|^2 + R - r.\end{aligned}$$

Using the above and the evolution equation of Δ (3.8) (along with $\partial_t R = R\partial_t L$, by definition of L) we have

$$\begin{aligned}\partial_t Q &= (\partial_t \Delta)L + \Delta(\partial_t L) + R(\partial_t L) \\ &= (R-r)\Delta L + \Delta\left(\Delta L + |\nabla L|^2 + R - r\right) + R\left(\Delta L + |\nabla L|^2 + R - r\right).\end{aligned}$$

Now using the definition of Q (3.52), the identity (3.22) on the $\Delta|\nabla L|^2$ term and the identity (3.21) on the resulting $\Delta\nabla L$ term, we have

$$\begin{aligned}\partial_t Q &= \Delta Q + 2\langle \Delta\nabla L, \nabla L\rangle + 2|\nabla\nabla L|^2 + R\left(2\Delta L + |\nabla L|^2 + R - r\right) - r\Delta L \\ &= \Delta Q + 2\left\langle \nabla\Delta L + \frac{1}{2}R\nabla L, \nabla L\right\rangle + 2|\nabla\nabla L|^2 + R\left(2\Delta L + |\nabla L|^2 + R - r\right) - r\Delta L.\end{aligned}$$

Now we substitute $\Delta L = Q - (R - r)$ to obtain

$$\begin{aligned}\partial_t Q &= \Delta Q + 2\langle \nabla Q - \nabla R, \nabla L\rangle + 2|\nabla\nabla L|^2 + R\left(2\Delta L + 2|\nabla L|^2 + R - r\right) - r\Delta L \\ &= \Delta Q + 2\langle \nabla Q, \nabla L\rangle + 2|\nabla\nabla L|^2 + R(R-r) + 2R\Delta L - r\Delta L + 2R|\nabla L|^2 - 2\langle \nabla R, \nabla L\rangle.\end{aligned}$$

The final two terms cancel because of

$$\langle \nabla R, \nabla L\rangle = \langle R\nabla L, \nabla L\rangle = R|\nabla L|^2,$$

and the third and fourth terms of the desired evolution equation (3.56) appear as

$$\begin{aligned} 2\left|\nabla\nabla L + \frac{1}{2}(R-r)g\right|^2 + rQ &= 2|\nabla\nabla L|^2 + (R-r)^2 + 2(R-r)\Delta L + r(\Delta L + R - r) \\ &= 2|\nabla\nabla L|^2 + R(R-r) + 2R\Delta L - r\Delta L, \end{aligned}$$

allowing us to conclude.

2. Similar computations give the results for the modified evolution equations. See Equation (5.41) and Lemma 5.59 of [CK04] for details. □

To make use of these evolution equations we will call upon the maximum principle. From (3.56) we have

$$\partial_t Q \geq \Delta Q + 2\langle \nabla Q, \nabla L\rangle + Q^2 + rQ, \tag{3.59}$$

since for any symmetric 2-tensor τ we have

$$2|\tau|^2 = 2g^{ik}g^{jl}\tau_{ij}\tau_{kl} \geq \left(g^{ij}\tau_{ij}\right)^2 = (\mathrm{tr}\tau)^2,$$

where Q is the trace of $\nabla\nabla L + \frac{1}{2}(R-r)g$ (which plays the role of τ). Thus, to apply the maximum principle to Q, the corresponding ordinary differential equation to inspect is

$$\partial_t\alpha = \alpha^2 + r\alpha \quad \text{with} \quad \alpha_0 := \alpha(0) < -r < 0.$$

This has the solution

$$\alpha(t) = -\frac{r\alpha_0 \mathrm{e}^{rt}}{\alpha_0\mathrm{e}^{rt} - \alpha_0 - r} = -\frac{Cr\mathrm{e}^{rt}}{C\mathrm{e}^{rt} - 1} \quad \text{where} \quad C := \frac{\alpha_0}{\alpha_0 + r} > 1.$$

This grants us our first Harnack inequalities.

Proposition 3.12.2. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$ and bounded Ricci scalar R. Then,*

1. ***Differential Harnack inequality**:*

$$\partial_t \log R - |\nabla \log R|^2 \geq -\frac{Cr\mathrm{e}^{rt}}{C\mathrm{e}^{rt} - 1},$$

where $C > 1$ is a constant only depending on g_0;

2. ***Classical Harnack inequality**: there exist constants $C_1' > 1$ and $C' > 0$ only depending on g_0 such that the estimates*

$$\frac{R(x_2, t_2)}{R(x_1, t_1)} \geq \mathrm{e}^{-\frac{A}{4}}\frac{C_1'\mathrm{e}^{rt_1} - 1}{C_1'\mathrm{e}^{rt_2} - 1} \geq \mathrm{e}^{-\frac{A}{4} - C'(t_2 - t_1)}, \tag{3.60}$$

hold for all points $x_1, x_2 \in \Sigma$ and times $0 \leq t_1 < t_2$, where $A = A(x_1, x_2, t_1, t_2)$ is defined by

$$A := \inf_\gamma \int_{t_1}^{t_2} \left|\frac{\mathrm{d}\gamma}{\mathrm{d}t}\right|^2 \mathrm{d}t, \tag{3.61}$$

where the infimum is taken over all curves $\gamma : [t_1, t_2] \to \Sigma$ such that $\gamma(t_1) = x_1$ and $\gamma(t_2) = x_2$.

Proof. 1. By inspecting the definition of Q (3.52) and evolution equation for L (3.55) we obtain

$$Q = \partial_t \log R - |\nabla \log R|^2,$$

which, combined with our discussion of the maximum principle applied to the evolution inequality of Q (3.59) gives the desired inequality;

2. For points $(x_1, t_1), (x_2, t_2) \in \Sigma \times [0, \infty)$ and a curve γ as described, the Fundamental Theorem of Calculus gives

$$\begin{aligned}
\log \frac{R(x_2, t_2)}{R(x_1, t_1)} &= \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t} \big(\log R(\gamma(t), t)\big) \mathrm{d}t \\
&= \int_{t_1}^{t_2} \left(\partial_t \log R(\gamma(t), t) + \left\langle \nabla \log R, \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right\rangle \right) \mathrm{d}t \\
&\geq \int_{t_1}^{t_2} \left(|\nabla \log R|^2 - \frac{Cr\mathrm{e}^{rt}}{C\mathrm{e}^{rt} - 1} - |\nabla \log R| \left| \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right| \right) \mathrm{d}t \\
&\geq \int_{t_1}^{t_2} \left(-\frac{Cr\mathrm{e}^{rt}}{C\mathrm{e}^{rt} - 1} - \frac{1}{4} \left| \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right|^2 \right) \mathrm{d}t \\
&= -\frac{1}{4} \int_{t_1}^{t_2} \left| \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right|^2 \mathrm{d}t - \log(C\mathrm{e}^{rt_2} - 1) + \log(C\mathrm{e}^{rt_1} - 1),
\end{aligned}$$

where the differential Harnack inequality is used the first inequality. Exponentiating both sides and inserting the definition of A (3.61) gives the result. The second inequality of (3.60) is a simplification of the first. □

We now want to provide analogous result in the case where the initial Ricci scalar has arbitrary sign. Inspecting the evolution equation for $\hat{Q}$ (3.58), we note that there is a problematic negative term: $s|\nabla \hat{L}|^2$. This makes using the maximum principle trickier than before, though by considering the quantity $\hat{Q} + s\hat{L}$ we can circumvent this problem, as done in the proof of the result below.

Proposition 3.12.3. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$. Then,*

1. ***Modified differential Harnack inequality**:*

$$\partial_t \log(R - s) - |\nabla \log(R - s)|^2 - s \geq -C,$$

where $C > 0$ is a constant only depending on g_0;

2. ***Modified classical Harnack inequality**: there exists a constant $C' > 0$ only depending on g_0 such that the estimate*

$$\frac{R(x_2, t_2) - s(t_2)}{R(x_1, t_1) - s(t_1)} \geq \mathrm{e}^{-\frac{A}{4} - C'(t_2 - t_1)}, \tag{3.62}$$

holds for all points $x_1, x_2 \in \Sigma$ and times $0 \leq t_1 < t_2$, where $A = A(x_1, x_2, t_1, t_2)$ is defined as in (3.61).

Proof. 1. By inspecting the evolution equation for Q (3.56) we obtain

$$\hat{Q} = \partial_t \log(R-s) - |\nabla \log(R-s)|^2 - s.$$

Thus, if we can bound $\hat{Q}$ by below by some constant $-C < 0$, then we are done. As hinted in the discussion preceding the statement of the result, we consider the quantity $\hat{P} := \hat{Q} + s\hat{L}$. By the evolution equations of s (3.54) and $\hat{L}$ (3.57), we have

$$\partial_t \left(s\hat{L}\right) = \Delta \left(s\hat{L}\right) + s\left|\nabla\hat{L}\right|^2 + s(R-r+s) + (s-r)\left(s\hat{L}\right).$$

Now, there exists some $c > 0$ such that $\hat{L} \geq -c(1+t)$, and we can write

$$s\left|\nabla\hat{L}\right|^2 = 2\left\langle \nabla\left(s\hat{L}\right), \nabla\hat{L}\right\rangle - s\left|\nabla\hat{L}\right|^2,$$

which makes the gradient of $s\hat{L}$ explicitly appear in contraction with a vector field. This allows us to state

$$\partial_t \left(s\hat{L}\right) \geq \Delta\left(s\hat{L}\right) + 2\left\langle \nabla\left(s\hat{L}\right), \nabla\hat{L}\right\rangle - s\left|\nabla\hat{L}\right|^2 - c',$$

for some constant $c' > 0$. Using this and the evolution equation of $\hat{Q}$ (3.58), we have

$$\begin{aligned}\partial_t\hat{P} &= \partial_t\hat{Q} + \partial_t\left(s\hat{L}\right)\\ &\geq \Delta\hat{P} + 2\left\langle\nabla\hat{P}, \nabla\hat{L}\right\rangle + \hat{Q}^2 + (r-s)\hat{Q} - c'',\end{aligned}$$

for some $c'' > 0$. Now since $s\hat{L}$ is bounded, there exists some $\tilde{C} > 0$ such that

$$\hat{Q}^2 + (r-s)\hat{Q} - c'' \geq \frac{1}{2}\left(\hat{Q} + 2s\hat{L}\hat{Q}^2 + s^2\hat{L}^2 - \tilde{C}^2\right) = \frac{1}{2}\left(\hat{P}^2 - \tilde{C}^2\right).$$

Finally, an application of the maximum principle on the evolution equation of $\hat{P}$ gives some $C > 0$ such that

$$\hat{P} \geq \min\left\{\min_{x\in\Sigma}\hat{P}(x,0), -C\right\},$$

which, recalling $\hat{P} = \hat{Q} + s\hat{L}$ and that $s\hat{L}$ is bounded give the desired bound on $\hat{Q}$;

2. For points $(x_1, t_1), (x_2, t_2) \in \Sigma \times [0, \infty)$ and a curve γ as described, we have

$$\begin{aligned}\log\frac{R(x_2,t_2) - s(t_2)}{R(x_1,t_1) - s(t_1)} &= \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t}\big(\log(R(\gamma(t),t) - s(t))\big)\mathrm{d}t\\ &\geq \int_{t_1}^{t_2}\left(-C + s(t) - \frac{1}{4}\left|\frac{\mathrm{d}\gamma}{\mathrm{d}t}\right|^2\right)\mathrm{d}t,\end{aligned}$$

where the same procedure has been followed as in the proof of the non-modified case (Proposition 3.12.2). Integrating s using its definition (3.54), exponentiating both sides and inserting the definition of A (3.61) gives the result. □

We have now assembled the necessary bits and pieces to conclude our proof.

3.13 Positive Average Ricci Scalar VI: Conclusion

This section will allow us to conclude the proof of the Uniformisation Theorem using the tools devised throughout this chapter. The remaining piece was the case of the proof of Theorem 3.4 in which the average Ricci scalar was positive ($r > 0$). We will first prove the result for the case of a positive initial Ricci scalar $R(\cdot, 0) > 0$. Afterwards, we will show that in the case of an initial Ricci scalar of mixed sign, the Ricci scalar eventually becomes positive, which allows us to conclude by restarting the flow at that time and using the first case.

We first need to show that the Ricci scalar is uniformly bounded from below by a positive constant.

Proposition 3.13.1. *For a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with Ricci scalar $R(\cdot, 0) > 0$, there exists a constant $C > 0$ depending only on g_0 such that*

$$R(x,t) \geq C > 0 \quad \textit{for all} \quad (x,t) \in \Sigma \times [0, \infty).$$

Proof. From the classical Harnack inequality (3.60), we have

$$\frac{R(x_2, t_2)}{R(x_1, t_1)} \geq \mathrm{e}^{-\frac{A}{4} - C'(t_2 - t_1)}.$$

Note that if $0 \leq t \leq 1$ (choosing $t_1 = 0$ in the above, or considering Proposition 3.2.1), then $R \geq c' \mathrm{e}^{-r}$ for some $c' > 0$, so we have a bound in this case. Now, consider the case $t \geq 1$. Choose $x_1 \in \Sigma$ such that

$$r \leq R(x_1, t-1) \leq R_{\max}(t-1).$$

Then,

$$R(x,t) \geq \mathrm{e}^{-\frac{A}{4} - C} R(x_1, t-1) \geq r \mathrm{e}^{-C}\, \mathrm{e}^{-\frac{A}{4}},$$

so that it suffices to find a uniform upper bound for $A(x_1, t-1, x, t)$. From the upper bound on R from Lemma 3.11.3 and the bounds on g from Lemma 3.8.1, we have

$$\mathrm{e}^{-C} g(t-1) \leq g(t') \leq \mathrm{e}^{r} g(t-1) \quad \text{for all} \quad t' \in [t-1, t].$$

Consider a geodesic γ of constant speed joining x to x_1 parametrised by arc-length:

$$\left|\dot{\gamma}(t')\right|_{g(t)} = d_{g(t)}(x, x_1).$$

Now, we bound the desired A from above as follows:

$$A \leq \int_{t-1}^{t} \left|\gamma'(t')\right|^2_{g(t')} \mathrm{d}t' \leq \mathrm{e}^{r+C} \int_{t-1}^{t} \left|\gamma'(t')\right|^2_{g(t)} \mathrm{d}t' \leq \mathrm{e}^{r+C} \left(d_{g(t)}(x, x_1)\right)^2,$$

which is finite because the diameter is bounded. □

We will use this lower bound for the Ricci scalar to apply the maximum principle once more, this time on the evolution equation of our M tensor (which we have yet to calculate). This evolution equation will allow us to determine a decaying bound for $|M|$. In this direction, we turn to what we previously called the gradient Ricci-DeTurck flow and our discussion of gradient Ricci solitons.

By Proposition 3.9.1, we know that gradient Ricci solitons have constant curvature, and that the vanishing of the M tensor implies a solution is a gradient Ricci soliton. Thus, if we can show that M (and its derivatives, to satisfy our demand that convergence is in any $\mathcal{C}^k$ norm) vanishes as $t \to \infty$, then we can conclude that the limiting metric is of constant curvature, as desired. Recall the gradient Ricci-DeTurck flow:

$$\partial_t g_{ij} = 2M_{ij} = 2\nabla_i \nabla_j f - (R - r)g_{ij} = (r - R)g_{ij} + \left(\mathcal{L}_{\nabla f} g\right)_{ij}. \tag{3.63}$$

It has solutions that differ from our usual the normalised Ricci flow (3.3) by pulling-back by a one-parameter family of diffeomorphisms ψ_t generated by the one-parameter family of vector fields $\nabla f(t)$. To see this, consider a solution $g(t)$ to the gradient Ricci-DeTurck flow above and define

$$\bar{g}(t) := \psi_t^* g(t),$$

for ψ_t a one-parameter family of diffeomorphisms.[46] Now,

$$\begin{aligned}
\partial_t \bar{g}(t) &= \partial_\lambda \left(\psi_{t+\lambda}^* g(t+\lambda)\right)\big|_{\lambda=0} \\
&= \psi_t^* \left(\partial_t g(t)\right) + \partial_\lambda \left(\psi_{t+\lambda}^* g(t)\right)\big|_{\lambda=0} \\
&= \psi_t^* \left((r-R)g(t) + \mathcal{L}_{\nabla f(t)} g(t)\right) + \partial_\lambda \left(\left(\psi_t^{-1} \circ \psi_{t+\lambda}\right)^* \psi_t^* g(t)\right)\Big|_{\lambda=0} \\
&= (r-R)\bar{g}(t) + \psi_t^* \left(\mathcal{L}_{\nabla f(t)} g(t)\right) - \mathcal{L}_{(\psi_t^{-1})_* \nabla f(t)} \psi_t^* g(t) \\
&= (r-R)\bar{g}(t).
\end{aligned}$$

Now, since $|M|^2$ is invariant under diffeomorphism, if a bound for it can be found supposing the gradient Ricci-DeTurck flow, then the bound will also hold under the usual normalised Ricci flow. First, however, we need to calculate the evolution equation for M.

Proposition 3.13.2. *For a metric undergoing the normalised Ricci flow, the tensor M defined as* (3.36) *satisfies*

$$\partial_t M = \Delta M + (r - 2R)M; \tag{3.64}$$
$$\partial_t |M|^2 = \Delta M - 2|\nabla M|^2 - 2R|M|^2. \tag{3.65}$$

Proof. First, we compute the following combination of derivatives of the potential:

$$\begin{aligned}
\nabla_i \nabla_j \Delta f &= \nabla_i \nabla_j \nabla_k \nabla^k f \\
&= \nabla_i \nabla_k \nabla_j \nabla^k f - \nabla_i \left(R_{jl} \nabla^l f\right) \\
&= \nabla_k \nabla_i \nabla_j \nabla^k f - R^l_{ikj} \nabla_l \nabla^k f - R_{il} \nabla_j \nabla^l f - R_{jl} \nabla_i \nabla^l f - \left(\nabla_i R_{jl}\right) \nabla^l f \\
&= \Delta \nabla_i \nabla_j f - \nabla^k \left(R^l_{ikj} \nabla_l f\right) - R^l_{ikj} \nabla_l \nabla^k f - R_{il} \nabla_j \nabla^l f - R_{jl} \nabla_i \nabla^l f - \left(\nabla_i R_{jl}\right) \nabla^l f \\
&= \Delta \nabla_i \nabla_j f - \frac{1}{2}\left(\nabla_i R \nabla_j f + \nabla_j R \nabla_i f - g_{ij} \nabla^k R \nabla_k f\right) - 2R\left(\nabla_i \nabla_j f - \frac{1}{2} g_{ij} \Delta f\right),
\end{aligned}$$

where we have repeatedly used the Riemann tensor identities (2.7) and (3.20), their traces, and our favourite identity (3.21).

[46] Recall that we argued this would be true in our discussion of the short-time existence of Ricci flow in Section 3.5.

Next, using the defining equations of M (3.36) and f (3.19) as well as the evolution equations of f (3.23), Γ_{ij}^k (3.7) and R (3.18), we have

$$\begin{aligned}
\partial_t M_{ij} &= \partial_t \left(\nabla_i \nabla_j f - \frac{1}{2}(R-r) g_{ij} \right) \\
&= \nabla_i \nabla_j (\partial_t f) - \left(\partial_t \Gamma_{ij}^k \right) \nabla_k f - \frac{1}{2} g_{ij} (\partial_t R) - \frac{1}{2}(R-r)(\partial_t g_{ij}) \\
&= \nabla_i \nabla_j (\Delta f + rf) + \frac{1}{2} \left(\delta_j^k \nabla_i R + \delta_i^k \nabla_j R - g_{ij} \nabla^k R \right) \nabla_k f - \frac{1}{2} g_{ij} \left(\Delta R + R(R-r) \right) \\
&\quad + \frac{1}{2}(R-r)^2 g_{ij} \\
&= \nabla_i \nabla_j \Delta f + \frac{1}{2} \left(\nabla_i R \nabla_j f + \nabla_j R \nabla_i f - g_{ij} \nabla^k R \nabla_k f \right) - \frac{1}{2} g_{ij} \Delta R + r M_{ij}.
\end{aligned}$$

Now, we use our computation of $\nabla_i \nabla_j \Delta f$ to cancel the term in parentheses to find

$$\partial_t M_{ij} = \Delta \nabla_i \nabla_j f - 2R \left(\nabla_i \nabla_j f - \frac{1}{2} g_{ij} \Delta f \right) - \frac{1}{2} g_{ij} \Delta R + r M_{ij}.$$

Using the definitions of M (3.36) and f (3.19) once more, we obtain

$$\begin{aligned}
\partial_t M_{ij} &= \Delta \nabla_i \nabla_j f - \frac{1}{2} g_{ij} \Delta R + (r - 2R) M_{ij} \\
&= \Delta \left(\nabla_i \nabla_j f - \frac{1}{2}(R-r) g_{ij} \right) + (r-2R) M_{ij} \\
&= \Delta M + (r-2R) M,
\end{aligned}$$

as desired. Finally, using the evolution equations for g^{ij} (3.6) and M (3.64), we have

$$\begin{aligned}
\partial_t |M|^2 &= \partial_t \left(g^{ik} g^{jl} M_{ij} M_{kl} \right) \\
&= 2 \langle M, \Delta M + (r-2R) M \rangle + 2(R-r)|M|^2 \\
&= \Delta |M|^2 - 2|\nabla M|^2 - 2R|M|^2,
\end{aligned}$$

where we have used the identity (3.22) on M. □

With the evolution equation for M in hand, we can apply the maximum principle to to bound it.

Lemma 3.13.1. *Consider a solution $(\Sigma, g(t), g_0)$ to the normalised Ricci flow with average Ricci scalar $r > 0$ and the M tensor defined as* (3.36). *Then, there exists a constant $C > 0$ such that*

$$|M| \leq C e^{-rt},$$

and so the M tensor vanishes exponentially quickly as $t \to \infty$.

Proof. Proposition 3.13.1 grants us a lower bound for R, we can use the maximum principle on the evolution equation for $|M|^2$ (3.65) to find the desired bound. □

This holds for both the normalised Ricci flow (3.3) and the gradient Ricci-DeTurck flow (3.63) by diffeomorphism invariance. The following result gives bounds for the derivatives of M, for a given solution of the gradient Ricci-DeTurck flow, which we will then convert back to being bounds for derivatives of R in the usual the normalised Ricci flow context.

Lemma 3.13.2. *Consider a solution* $(\Sigma, g(t), g_0)$ *to the gradient Ricci-DeTurck flow* (3.63) *with average Ricci scalar* $r > 0$ *and the* M *tensor defined as* (3.36)*. Then, for every integer* $k > 0$ *there exist constants* $C_k, C_k' < \infty$ *depending only on* g_0 *such that*

$$\left|\nabla^k M\right|^2 \leq C_k \mathrm{e}^{-C_k' t}.$$

Proof. A similar procedure to prove this as the one used in the case of $r \leq 0$ (see, for example, Propositions 3.7.2 and 3.7.4). See Corollary 5.63 of [CK04] for details. □

With vanishing M and derivatives of M, it follows that derivatives of R are similarly bounded with constants c_k and c_k' (for integers $k > 0$). Now, by diffeomorphism invariance, these bounds for $|\nabla^k R|$ must hold for the usual normalised Ricci flow, allowing us to conclude the $R(\cdot, 0) > 0$ case of Theorem 3.4.

To finish our proof of Theorem 3.4, we need to resolve the case where $R(\cdot, 0)$ changes sign. As previously mentioned, it suffices to show that in finite time, the Ricci scalar will become positive, so that the $R(\cdot, 0) > 0$ case allows us to conclude by restarting the flow once R is positive.

Proposition 3.13.3. *Consider a solution* $(\Sigma, g(t), g_0)$ *to the normalised Ricci flow with average Ricci scalar* $r > 0$*. Then there exists* $T < \infty$ *such that*

$$\inf_{(x,t) \in \Sigma \times [T, \infty)} R(x, t) > 0.$$

Proof. We first want find a lower bound for $R - s$, where $s(t)$ is defined by (3.54). Similarly to our proof of Proposition 3.13.1, consider $t \geq 1$ and choose $x_1 \in \Sigma$ such that

$$0 < r \leq R(x_1, t-1) \leq R_{\max}(t-1).$$

Now, the modified classical Harnack inequality (3.62) gives

$$R(x,t) - s(t) \geq \mathrm{e}^{-C} \left(R(x_1, t-1) - s(t-1) \right) \mathrm{e}^{-\frac{A}{4}} \geq \mathrm{e}^{-C} r \mathrm{e}^{-\frac{A}{4}}, \tag{3.66}$$

so we will once more attempt to find an upper bound for $A = A(x_1, t-1, x, t)$. Using our bounds for the metric from Lemma 3.8.1, for any $t' \geq t - 1$ we have

$$c g(x, t-1) \leq g(x, t') \leq C g(x, t-1) \quad \text{for some constants} \quad c < C < \infty.$$

From this we have in a similar manner to Proposition 3.13.1 that A is bounded from above by the diameter, which is in turn bounded by Lemma 3.11.5. Now that $R - s$ is bounded from below, we take an infimum over $x \in \Sigma$ on the bound (3.66), so there exists a constant $\varepsilon > 0$ such that

$$R_{\min}(t) \geq \varepsilon - s(t).$$

Now, $s \to 0$ as $t \to \infty$, so we indeed have that the Ricci scalar becomes positive. □

Since the curvature now becomes positive, we can restart the normalised Ricci flow at this (finite) time, and use the conclusion from the initially positive Ricci scalar case to resolve our result Theorem 3.4 for the $r > 0$ case. Thus we have completed our proof of the Uniformisation Theorem. □

Chapter 4

General Relativity

With a thorough understanding of Ricci flow in hand, we turn to the second topic of this text: General Relativity. We will see that the $(2+1)$-dimensional case is best tackled with the Uniformisation Theorem in hand. General Relativity gives a description of classical gravity in terms of the curvature of a spacetime manifold in an elegant geometric way that its predecessor, first studied by Newton, did not address.

Newtonian gravity faced many challenges. A glaring one was revealed by Einstein's theory of Special Relativity, which postulated that no information could move faster than the speed of light.[1] Classical Newtonian gravity broke this rule by asserting that changes in position or in mass of elements in a gravitational system were instantaneously felt by all elements of the system. This meant that if the Earth doubled in mass all of a sudden, the Sun would immediately sense the weight gain—the information travelling from the Earth to the Sun faster than the speed of light! To solve this problem (and others), Einstein (and others) set out to find improved laws of gravity.

Overview. Section 4.1 introduces the Einstein equations and the Einstein-Hilbert action. Section 4.2 discusses causal assumptions on spacetime and Cauchy surfaces. Section 4.3 explores embedded Riemannian submanifolds and their extrinsic curvature. Section 4.4 defines the Teichmüller space and relates it to the Uniformisation Theorem and the space of solutions to the Einstein equations. Section 4.5 applies the ADM (or Hamiltonian) formalism to the Einstein equations, splitting spacetime into space and time. The two following sections explicitly demonstrate the reduction of the Einstein equations to dynamics on the cotangent bundle of the Teichmüller space of a Cauchy surface, first for the case of a greater-than-unit genus (Section 4.6) and then for genera of zero or unity (Section 4.7). Section 4.8 discusses the equations satisfied by the lapse and the shift, as well as the Einstein equations as coupled flow equations.

Reference Guide. A basic introduction to General Relativity (from a physics perspective) is found in [Car19], and a standard text for the $(2+1)$-dimensional case is [Car03]. Classic advanced texts are [MTW73] and [Wal10]. A mathematical text on Einstein manifolds (with some discussion of General Relativity) is found in [Bes07]. A great discussion of causal structures is found in [HE73]. A rigorous book dedicated to the Cauchy problem in General Relativity is [Rin09], and a similarly well-rounded (that is, geometry- and partial differential equation-based) approach to the Einstein equations is found in [CB08]. The spacetime split formalism is well-presented in [Gou12]. Relevant Teichmüller theory can be found in [Tro92]. The original paper performing the Hamiltonian spacetime split of the

[1] We choose our units so that the speed of light is unity, as well as Newton's gravitational constant—these are sometimes called **geometric units**.

Einstein equations is [ADM59]. The original paper detailing the reduction of the Einstein equations to the cotangent bundle of the Teichmüller space is [Mon89].

4.1 Einstein Equations

In 1915, Einstein (see [Ein15]) succeeded on his quest and wrote down the equations that bear his name, founding the theory of General Relativity.

Definition 4.1.1 (Einstein equations)**.** For a Lorentzian manifold $(\mathcal{M}, g)$, the Einstein tensor $G_{\mu\nu}$ (introduced in Definition 2.2.2) must satisfy the **Einstein equations**:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu},$$

for the **energy-momentum tensor** $T_{\mu\nu}$. (We have included the **cosmological constant**, $\Lambda \in \mathbb{R}$, though it was added in 1917; see [Ein22].) The energy-momentum tensor contains all of the information of the mass and energy distribution throughout a spacetime. Sometimes, the cosmological constant is seen as the energy of the vacuum and thus is absorbed into $T_{\mu\nu}$—we will leave Λ to the side for now and return to it later. Note too that these coupled partial differential equations require initial data to be a solvable system whose solution is the spacetime metric $g_{\mu\nu}$.

In 2 dimensions $G_{\mu\nu}$ vanishes, so the only non-trivial solution of the Einstein equations has an energy-momentum tensor equal to the vacuum energy. In other words, when $n = 2$, every metric satisfies the Einstein equations! For this reason, we assume that $n \geq 3$, eventually restricting ourselves to the $n = 3$, or $(2+1)$-dimensional, case. By tracing the Einstein equations we obtain

$$R = 2\frac{n\Lambda - 8\pi T}{n-2},$$

where $T := T^\mu_\mu$. This allows us to rewrite the Einstein equations as

$$R_{\mu\nu} = 8\pi T_{\mu\nu} + 2\frac{\Lambda - 4\pi T}{n-2} g_{\mu\nu},$$

which contains the exact same amount of information as the original formulation.[2]

Next, we note that the Einstein equations are (seemingly paradoxically) both over-determined and under-determined. Since not all possible given initial data give rise to solutions, we call the equations **over-determined**; and since for any solution g and diffeomorphism ψ of $\mathcal{M}$ we obtain another solution, $\psi^* g$, making them **under-determined**. This can be seen by noting that the Einstein tensor is diffeomorphism-equivariant: it satisfies $G(\psi^* g) = \psi^* G(g)$,[3] and we require $T_{\mu\nu}$ to behave in a similar fashion.

Finally, it can be shown that in the 'Newtonian limit' these equations reduce to the Newtonian gravitational equation (see Example 2.3.1), as one would hope.

We have not explicitly explained what shape the energy-momentum tensor takes, and will not consider it in any detail in this text. The only example we give is the following, which we will assume throughout.

[2] In fact, this reversed version (omitting the cosmological constant Λ) was how Einstein first formulated it.

[3] See the discussion at the beginning of Chapter 3 about the diffeomorphism equivariance of the Ricci tensor.

Example 4.1.1 (Vacuum energy-momentum tensor)**.** In a vacuum, the energy-momentum tensor is identically zero: $T_{\mu\nu} = 0$. The Ricci scalar then becomes

$$R = \frac{2n\Lambda}{n-2},$$

and so we obtain the **vacuum Einstein equations**:

$$R_{\mu\nu} = \frac{2\Lambda}{n-2} g_{\mu\nu}.$$

When $\Lambda = 0$, this is simply the requirement that spacetime be Ricci-flat. In our dimension of concern, $n = 3$, this becomes $R_{\mu\nu} = 2\Lambda g_{\mu\nu}$, which was the definition of an Einstein metric in Definition 2.2.

We often content ourselves with solving the vacuum equations, since solving the Einstein equations with a given energy-momentum tensor is often very challenging. Even this task is arduous, however. Thankfully, it did not take long for interesting solutions to appear, such as those describing one of the most famous predictions of Einstein's theory of General Relativity: the existence of black holes. We now state two examples of solutions to the vacuum Einstein equations.

Example 4.1.2 (Minkowski metric)**.** The Minkowski metric $\eta_{\mu\nu}$ given by (2.2) is a trivial solution to the vacuum Einstein equations with no cosmological constant.

Example 4.1.3 (Bañados-Teitelboim-Zanelli metric)**.** An example of a $(2+1)$-dimensional black hole is given by the **Bañados-Teitelboim-Zanelli metric**, or BTZ metric:

$$g^{\text{BTZ}} := -N^2 \mathrm{d}t^2 + N^{-2}\mathrm{d}r^2 + r^2(\mathrm{d}\phi + N^\phi \mathrm{d}t)^2, \tag{4.1}$$

where the **lapse** N and **shift** N^ϕ functions are defined as

$$N := \sqrt{-M + \frac{r^2}{\ell^2} + \frac{J^2}{4r^2}} \quad \text{and} \quad N^\phi := -\frac{J}{2r^2},$$

and the black hole is characterised by the **angular momentum** J and the **mass** M. It is a solution to the Einstein equations with cosmological constant $\Lambda = -\ell^{-2}$ and has Killing vectors ∂_t and ∂_ϕ (a metric with these properties is often called **stationary** and **axially symmetric**). This example is discussed at length in [Car95a].

Before turning to the geometric tools required to understand the Einstein equations and their solutions, we make the following definition that shines light on the origin of the equations themselves.

Definition 4.1.2 (Einstein-Hilbert action)**.** The action[4] that gives rise to the vacuum Einstein equations is the **Einstein-Hilbert action** $\mathcal{S}^{\text{EH}}$ given on a Lorentzian manifold $(\mathcal{M}, g)$ by

$$\mathcal{S}^{\text{EH}}(g) := \int_{\mathcal{M}} (R - 2\Lambda),$$

[4] Recall that in mechanics, the **action** of a system is the integral of the **Lagrangian**—this will be more properly recalled in Section 4.5 when we employ the Hamiltonian approach to General Relativity. Using the calculus of variations, one can vary the action to find the **Euler-Lagrange equations**, which dictate the mechanics of the system. For example, in electromagnetism, we have the **Maxwell action**

$$\mathcal{S}^{\text{M}} := -\frac{1}{4}\int_{\mathcal{M}} F_{\mu\nu}F^{\mu\nu},$$

for the electromagnetic field tensor $F_{\mu\nu}$, which gives rise to the Maxwell equations. For details on the derivation of the Euler-Lagrange equations, see Chapter 8 of [Eva10].

for Ricci scalar R and cosmological constant Λ. Note that as usual, we have omitted the volume form in our integrand: there is an implicit $\sqrt{-\det g}$ term.

We note that varying this action with respect to the metric g requires computing how three terms vary: $\sqrt{-\det g}$ (arising from integrating over $\mathcal{M}$), g^{-1} and Ric (both arising from the definition of R). This computation is lengthy and will not be done here,[5] though it turns out that the variation of the Ricci tensor only contributes a divergence-like term, which vanishes upon integration, as we assume our manifold to be closed.[6]

Before decomposing the Einstein equations and viewing them as coupled partial differential equations, we need to build up more machinery to understand the structure of spacetime. First, we will inspect the causal structure of Lorentzian manifolds and make several important definitions of properties that we will assume our universal manifold to satisfy to make the Einstein equations well-posed.

4.2 Causal Structure and Cauchy Surfaces

One may ask how different this chapter's Lorentzian context is from the Riemannian one of Chapter 3. How much of an effect can one different sign in the signature of the metric have on both geometric intuition and computations? How special is this 'time' direction? This section attempts to address these questions, revealing just how beautiful a Lorentzian manifold description of the universe is, complete with a precise notion of causality. The goal of this section is to determine the demands that are necessary and sufficient to write our spacetime manifold as $\mathcal{M} \cong \Sigma \times [0,1]$, where just as in the previous chapter Σ is a Riemannian (hence spatial) manifold, and the interval $[0,1]$ is a time interval. Effectively this is a split of spacetime into 'space' and 'time.'

On a Lorentzian manifold, the lack of positive-definiteness of the metric gives the possibility of three lengths of vectors: positive, zero and negative.

Definition 4.2.1 (Timelike, null and spacelike vectors and curves)**.** On a Lorentzian manifold $(\mathcal{M}, g)$, a non-zero vector V_p at a point $p \in \mathcal{M}$ is called

1. **Timelike** if $g_p(V_p, V_p) > 0$;
2. **Null** if $g_p(V_p, V_p) = 0$;
3. **Spacelike** if $g_p(V_p, V_p) < 0$.

A curve (on in general any subset of a manifold) is described as **timelike**, **null**, or **spacelike** if its tangent vector at every point satisfies the corresponding requirement above. A vector or curve is **causal** if it is timelike or null—named as such because while light travels on null curves and massive matter on timelike curves, moving on a spacelike curve will require a greater velocity than the speed of light, which is not allowed in relativity.

(We may now revisit Definition 2.1.5 to properly define the length of curves and the distance function in the Lorentzian context, both of which only consider causal curves.)

The above definition now brings to light the importance of the 'time' coordinate in our Lorentzian metric.

[5]See Chapter 4.3 of [Car19] for a derivation or Chapter 4 of [Bes07] for a more thorough one.

[6]See the footnote directly after Proposition 3.1.1 for some motivation of the connection between the Einstein equations and the Einstein-Hilbert action.

Definition 4.2.2 (Time-orientable)**.** If a metric admits a timelike vector field then it is called **time-orientable**, and a choice of such a field gives a time-orientation. We will always assume the existence of a **time function** $\mathtt{t} : \mathcal{M} \to \mathbb{R}$ such that

$$\nabla \mathtt{t} = g^\sharp(\mathtt{dt}),$$

is always timelike and $\mathtt{t}$ is nowhere-vanishing—this certainly satisfies the previous time-orientable condition. A **spacetime** is a connected, oriented, time-oriented Lorentzian manifold.

Frequently this time function is implicitly assumed: often authors will simply write $t = \mathtt{t}$ for the zeroth coordinate. We will not mention $\mathtt{t}$ explicitly many times, as *we assume all of our Lorentzian manifolds from now onward to be spacetimes*.

Consider a causal curve that returns to its starting point. If matter (or information) were to travel along this curve, the notions of past, present and future would be erased, since 'later' on its journey it would return to the same point in spacetime where it 'began' its journey. The notion of causality may seem delicate; nevertheless, we make the following definition.

Definition 4.2.3 (Causal)**.** A spacetime is called **causal** if it contains no closed causal curves.

Despite this, we will require a stronger property than the above for our considered spacetimes. We are interested in characterising spacetimes completely based only on data given on a subset of the manifold (as is usual in the formulation of an initial value problem of a partial differential equation). Considering information to propagate on causal curves, we define another subset of spacetime which includes all points that can be affected (or have been affected by) the points in the original subset.

Definition 4.2.4 (Future- and past-directed)**.** For a timelike vector field[7] T on a spacetime $(\mathcal{M}, g)$, a causal vector V_p (at a point $p \in \mathcal{M}$) is called **future-directed** if

$$g_p(V_p, T_p) < 0.$$

A curve is **future-directed** if each of its tangent vectors are. Similarly, a vector V_p is **past-directed** if $-V_p$ is future-directed, and a curve is **past-directed** if all of its tangent vectors are past-directed.

If there exists a future-directed causal curve from $p \in \mathcal{M}$ to $q \in \mathcal{M}$ then we write $p \leq q$. Similarly, $p \ll q$ if there exists a future-directed timelike curve from p to q.[8] We write

$$J^+(p) := \{q \in \mathcal{M} \mid p \leq q\}, \quad J^-(p) := \{q \in \mathcal{M} \mid q \leq p\} \quad \text{and} \quad J^\pm(\mathcal{U}) := \bigcup_{p \in \mathcal{U}} J^\pm(p),$$

for $\mathcal{U} \subset \mathcal{M}$ a subset of spacetime. The first two of these sets are known as the **causal future** and **causal past** of a point p and can be thought of the points that can be impacted by events at p, and those that could have impacted p, respectively (with the union of these spaces allowing for many points acting like p in the two previous characterisations). Finally, we define the **causal double-cone** $J(p, q)$ for points $p \ll q$ in $\mathcal{M}$ to be

$$J(p, q) := J^+(p) \cap J^-(q),$$

[7] This is often considered to be the gradient of a time function, though any timelike vector field suffices here.

[8] Note that these both define transitive relations.

which can be thought of as the set of all points that can both be impacted by events at p and can impact events at q (forming a diamond-like shape in the Minkowski case).

Finally, we can define a stronger notion than causality, which we will assume all of our spacetimes to have from here on out.

Definition 4.2.5 (Globally hyperbolic)**.** A spacetime $(\mathcal{M}, g)$ is **globally hyperbolic** if

1. All of the definable sets $J(p, q)$ are compact;

2. It is **strongly causal**: any neighbourhood of any point $p \in \mathcal{M}$ contains an open set $\mathcal{U}_p$ such that any causal curve with endpoints in $\mathcal{U}_p$ lies entirely within $\mathcal{U}_p$.

To see the usefulness of the above definition, we need to define a Cauchy surface, whose definition relies on the notion of extendibility of curves.

Definition 4.2.6 (Extendible)**.** On a semi-Riemannian manifold $(\mathcal{M}, g)$, a (smooth) curve $\gamma : [a, b) \to \mathcal{M}$ is

1. **Future-extendible** if it has an extension $\gamma : [a, b] \to \mathcal{M}$ and **future-inextendible** if it does not;

2. **Incomplete** if inextendible and has finite arc-length and **complete** otherwise.

Similarly, for a curve $\gamma : (a, b] \to \mathcal{M}$ we define the notions of **past-(in)extendible** and **(in)complete**.

Now, we have the following crucial definitions.

Definition 4.2.7 (Cauchy surface and developments)**.** Consider a spacetime $(\mathcal{M}, g)$ and a subset $\Sigma \subset \mathcal{M}$. Then,

1. We call Σ a **Cauchy surface** if every inextendible timelike curve in $\mathcal{M}$ intersects it exactly once;

2. The **domain of dependence**, or Cauchy development, $D^+(\Sigma)$ is the set of all points $p \in \mathcal{M}$ such that every past-directed timelike curve starting at p intersects Σ;

3. The **domain of influence** $D^-(\Sigma)$ is the set of all points $p \in \mathcal{M}$ such that every future-directed timelike curve starting at p intersects Σ.

Cauchy surfaces are exactly the subsets of spacetime on which we need initial data for the Einstein equations to be well-posed. The following important properties of the above definitions summarise their usefulness.

Proposition 4.2.1. *A spacetime $(\mathcal{M}, g)$ is globally hyperbolic if and only if it contains a Cauchy surface Σ. If this is the case, then*

1. $D^+(\Sigma) \cup D^-(\Sigma) = \mathcal{M}$;

2. *Any other Cauchy surface is diffeomorphic to Σ;*

3. *$\mathcal{M}$ can be written as a **foliation** of Cauchy surfaces Σ_t called **time-slices** (known also as **leaves** of the foliation):*

$$\mathcal{M} = \bigcup_{t \in [0,1]} \Sigma_t \quad \text{or} \quad \mathcal{M} = \bigcup_{t \in \mathbb{R}} \Sigma_t,$$

where the Σ_t are non-intersecting, and so has the topology $\mathcal{M} \cong [0, 1] \times \Sigma$, as we will assume in this text (or $\mathcal{M} \cong \mathbb{R} \times \Sigma$, which we will not consider);

4. *Every inextendible causal curve intersects* Σ*;*

5. Σ *is **achronal**:*[9] *no two points* p *and* q *in* Σ *have* $p \ll q$*.*

Proof. See Chapter 8 of [Wal10] or Chapter 4.3 of [Lan18] for details. □

The above proposition allows us to consider any globally hyperbolic spacetime to be a foliation of level sets of our time function $\mathtt{t}$:

$$\Sigma_t = \{p \in \mathcal{M} \mid \mathtt{t}(p) = t\}.$$

When these sets arise from a time function as in Definition 4.2.2, they are spacelike sets and therefore have an induced Riemannian metric. (We will make this assumption for our submanifold Σ throughout this text.) Now, if we are given initial data on a single time-slice, we expect that because of the globally hyperbolic nature of spacetime we can understand events occurring at all other points. This allows us to write the Einstein equations as an initial value problem, where the data is specified on any Cauchy surface. *For the rest of the text we assume all of our spacetimes to be globally hyperbolic.*

With an understanding of time-slices and foliations of spacetime under our belts, we turn to the mathematical subject of submanifolds. This will allow us to look at level sets of the time function as distinct manifolds embedded within spacetime.

4.3 Submanifolds and Extrinsic Curvature

Much of Einstein's ingenuity came from considering spacetime as a single manifold, effectively treating time and space on the same level, an idea dating back to Minkowski. Nevertheless, we have split our manifold $\mathcal{M} \cong \Sigma \times [0,1]$, and so hope to split the Einstein equations so that time and space are seen explicitly. To do so, we will consider slices of our manifold that will represent the spatial part of spacetime at various fixed times. This notion becomes precise when we delve into the study of embedded submanifolds.

Definition 4.3.1 (Submanifold)**.** For manifolds Σ (usually referred to as the submanifold itself) and $\mathcal{M}$ (the **ambient manifold**), a map $\psi : \Sigma \to \mathcal{M}$ defines a **submanifold** $\psi(\Sigma) \subset \mathcal{M}$ if[10]

1. ψ is a homeomorphism onto its image;

2. For all $q \in \Sigma$, the push-forward $(\psi_*)_q : \mathcal{T}_p\Sigma \to \mathcal{T}_{\psi(q)}\mathcal{M}$ is injective.

A **hypersurface** is a submanifold with a codimension[11] of 1.[12] For $(\mathcal{M}, g)$ a semi-Riemannian manifold, the **induced metric**[13] of Σ is the restriction of g to $\mathcal{T}\Sigma$.

Considering the tangent space $\mathcal{T}\Sigma$ as a subspace of $\mathcal{T}\mathcal{M}$ we encounter a new vector bundle: the normal bundle.[14]

[9] In particular, any Cauchy surface is spacelike.

[10] It is also possible (and equivalent) to define a submanifold starting from a subset of a manifold and requiring the subset to have properties such as local trivialisations.

[11] Recall that the **codimension** of a subspace is the difference between the dimension of the ambient space and the dimension of the subspace.

[12] Therefore, Cauchy surfaces are hypersurfaces.

[13] This is also called the **first fundamental form** of the submanifold.

[14] Section A.2 of the appendix contains more developed discussions of vector bundles in general.

Example 4.3.1 (Normal bundle)**.** For a semi-Riemannian manifold $(\mathcal{M}, g)$, the **normal bundle** (whose total space is written $N\Sigma$) of a submanifold $\Sigma \subset \mathcal{M}$ is composed of normal spaces $N_q\Sigma$ defined for points $q \in \Sigma$ as

$$N_q\Sigma := (\mathcal{T}_q\Sigma)^\perp,$$

where the perpendicularity is with respect to the inner product g. This allows us to decompose the tangent spaces of the ambient manifold at points $q \in \Sigma$ as

$$\mathcal{T}_q\mathcal{M} = \mathcal{T}_q\Sigma \oplus N_q\Sigma,$$

which in turn induces **tangential** and **normal projections**

$$\mathrm{proj}^\top : \mathcal{TM} \to \mathcal{T}\Sigma \quad \text{and} \quad \mathrm{proj}^\perp : \mathcal{TM} \to N\Sigma.$$

As a notational tool, we will write $V^\top := \mathrm{proj}^\top(V)$ and $V^\perp := \mathrm{proj}^\perp(V)$ for the images of the projections for some vector V.

Finally, we define a **normal vector** η to the submanifold Σ to be a vector such that

$$g(V, \eta) = 0 \quad \text{for all} \quad V \in \mathcal{TM}.$$

It is **unit** if $g(\eta, \eta) = \pm 1$ (where the positive sign is for a Riemannian manifold and the negative sign is for a Lorentzian manifold).

The above example allows us to use these projections to decompose objects that live on the tangent bundle of the ambient manifold $\mathcal{M}$. If we write $\nabla^{\mathcal{M}}$ for the Levi-Civita connection on $\mathcal{M}$ we can do the following decomposition at any point $q \in \Sigma$:

$$\nabla^{\mathcal{M}}_V W = (\nabla^{\mathcal{M}}_V W)^\top + (\nabla^{\mathcal{M}}_V W)^\perp,$$

for vectors V and W in $\mathcal{T}\Sigma$.

We then have the following definition, which helps describe the curvature of the submanifold within the ambient one.

Definition 4.3.2 (Extrinsic curvature)**.** For a hypersurface[15] Σ with unit normal η of a semi-Riemannian manifold $(\mathcal{M}, g)$ with Levi-Civita connection ∇, the **extrinsic curvature** is a $(2, 0)$-tensor κ given on vector fields V and W by[16]

$$\kappa(V, W) = g(\eta, \nabla_V W),$$

which is the coefficient[17] of the normal projection of $\nabla_V W$. It can be verified to be a tensor by it being $\mathcal{C}^\infty$-linear in each entry. The trace of this tensor is the **mean extrinsic curvature** and is written κ, which will hopefully not cause confusion: the extrinsic curvature will always have explicit vector inputs or will be written with indices.

We write $R^{\mathcal{M}}$ and $\nabla^{\mathcal{M}}$ for the Riemann $(4, 0)$-tensor and the Levi-Civita connection associated to the manifold $\mathcal{M}$ (and similarly for $\Sigma \subset \mathcal{M}$). The extrinsic curvature then satisfies the following properties.

[15]It is possible to define the extrinsic curvature for a submanifold of general codimension. For details, see Chapter VII of[KN69].

[16]Note that there is a choice of sign in this definition, and many authors have the opposite convention from this text.

[17]Note that closely related to the extrinsic curvature is the **second fundamental form** $I\!I$ (read 'two') given by $I\!I(V, W) := (\nabla^{\mathcal{M}}_V W)^\perp$.

Proposition 4.3.1. *The extrinsic curvature κ of a hypersurface $\Sigma \subset \mathcal{M}$ (with unit normal η) is symmetric and satisfies*

1. *The **Gauss formula**, or the Gauss-Weingarten equation:*

$$\nabla_V^{\mathcal{M}} W = \nabla_V^{\Sigma} W \pm \kappa(V, W)\eta,$$

where vector fields V and W in $\mathcal{T}\Sigma$ have been arbitrarily extended to $\mathcal{T}\mathcal{M}$ and $\pm$ refers to Riemannian (+) and Lorentzian (−);

2. *The **Gauss-Codazzi equations** for vector fields U, V, W and X on $\mathcal{M}$:*

$$R^{\mathcal{M}}(\eta, V, W, X) = (\nabla_W \kappa)(V, X) - (\nabla_V \kappa)(W, X), \tag{4.2}$$
$$R^{\mathcal{M}}(U, V, W, X) = R^{\Sigma}(U, V, W, X) \pm \kappa(U, X)\kappa(V, W) \mp \kappa(U, W)\kappa(V, X), \tag{4.3}$$

where in the lower equation the $\pm, \mp$ refers to Riemannian and Lorentzian, read top to bottom. In particular, if Σ is embedded in a flat manifold, the left-hand sides of the above equations will vanish.

Proof. 1. Symmetry follows from the computation:

$$\kappa(V, W) = g(\eta, \nabla_V W) = g(\eta, \nabla_W V) = \kappa(W, V),$$

by the torsion-free nature of the Levi-Civita connection;

2. The Gauss formula follows from the definition of κ;

3. The Gauss-Codazzi equations follow from several calculations, which we omit here. See Chapter 3.5 of [Gou12] or Chapter 6.3 of [Lan18] for details. □

We will translate this knowledge of hypersurfaces into the context of a Cauchy surface Σ embedded in a spacetime $\mathcal{M}$. This Cauchy surface will be of considerable interest, not only because it contains all of the topological information of $\mathcal{M}$ (since our time interval $[0, 1]$ is topologically trivial), but because it is a Riemannian 2-manifold. As such, we can apply our Uniformisation Theorem (usually in the form (2.2)) to a Cauchy surface. With this in mind, we return the space of all possible Riemannian metrics of a manifold, first discussed in Section 2.4.

4.4 Reduction to Teichmüller Space

Consider a $(2+1)$-dimensional spacetime $(\mathcal{M}, g)$ with a 2-dimensional Cauchy surface Σ that satisfies the Einstein equations. In the following section, we will split the Einstein equations so that Riemannian metric on Σ can be used to reconstruct the full spacetime metric that solves the Einstein equations. As such, we now focus on Σ and its possible Riemannian metrics, supposing that there exists a partial differential equation on Σ whose solution is a Riemannian metric on Σ. We recall that the largest possible solution space for such a set-up is the space of all Riemannian metrics $\mathscr{M}(\Sigma) = \mathscr{M}$. Since the Einstein equations do not allow for every element of $\mathscr{M}$ to be a solution (upon reconstruction of the Lorentzian metric as described above), they restrict this possible solution space. This section will inspect how else this infinite-dimensional space $\mathscr{M}$ can be reduced. Our goal is to motivate the following assertion: the space of Riemannian metrics for Σ modulo diffeomorphisms and modulo conformal equivalences (here the Uniformisation

Theorem makes its appearance) is isomorphic to the cotangent bundle of the space of non-biholomorphically equivalent complex structures on Σ (seen as a Riemann surface), which is known as the Teichmüller space of Σ. Riemannian metrics (modulo conformal equivalences and diffeomorphisms) will parametrise the Teichmüller space, and a modified version of the extrinsic curvature (which will be called the gravitational momentum π) will parametrise the cotangent space above any point of the Teichmüller space.

In this section we will often reinstate our notation of s as a Sobolev parameter, no longer requiring our functions to be $\mathcal{C}^\infty$-smooth.[18] From Chapter 2 we retrieve the following function spaces associated to a manifold Σ, where Σ is withheld from our notation for brevity:

1. $\mathscr{D}$: the group of diffeomorphisms; we will write $\mathscr{D}_0$ for the diffeomorphisms which are homotopic to the identity;[19]

2. $\mathscr{H}^s$: the Sobolev space of functions whose weak s-times derivatives are in $\mathscr{L}^2$;

3. $\mathscr{M}$: the space of Riemannian metrics; $\mathscr{M}_\lambda \subset \mathscr{M}$ denotes those with Ricci scalar $R = \lambda \in \mathbb{R}$;

4. $\mathscr{C}$: the space of complex structures (when Σ is seen as a Riemann surface);

5. $\mathscr{A}$: the space of almost complex structures (when Σ is seen as a Riemannian manifold).

We recall that a Riemann surface is a 2-dimensional complex manifold whose atlas is called a complex structure $c \in \mathscr{C}$. It turns out that diffeomorphisms are not equivalent to biholomorphisms in the complex case (unlike in the real case). However, we do have that the pull-back of a complex structure by a diffeomorphism gives rise to another Riemann surface, biholomorphic to the first via this diffeomorphism, as succinctly put in the following result, which we state without proof.

Proposition 4.4.1. *For a Riemann surface (Σ, c), a diffeomorphism $\psi : \Sigma \to \Sigma$ is a biholomorphism from (Σ, c) to $(\Sigma, \psi^* c)$, where the pull-back complex structure $\psi^* c \in \mathscr{C}$ is given by*

$$\psi^* c := (\psi^{-1}(U_\alpha), \varphi_\alpha \circ \psi)_\alpha.$$

Using this pull-back action, we can act with diffeomorphisms on our space of complex structures. This leads to the following definitions of moduli spaces, which are used to study Riemann surfaces which are not biholomorphic.

Definition 4.4.1 (Teichmüller space)**.** For a Riemann surface Σ, the **Riemann moduli space** $\mathscr{R}(\Sigma) = \mathscr{R}$ is given by

$$\mathscr{R} := \mathscr{C} / \mathscr{D},$$

where the quotient is with respect to the conformal action described in Definition 2.5.3. The **Teichmüller space** $\mathscr{T}(\Sigma) = \mathscr{T}$ is given by[20]

$$\mathscr{T} := \mathscr{C} / \mathscr{D}_0.$$

[18] For more discussion on this matter, return to Section 2.3. A great resource is [FT84].

[19] This is sometimes called the connected component of the identity.

[20] The difference between these spaces is more clear when using physics terminology: elements of $\mathscr{D}_0$ are known as 'small' diffeomorphisms, since they are formed by exponentiating and composing infinitesimal transformations (hence they lie close to the identity, which is the exponential of the trivial transformation). This induces the naming of elements of $\mathscr{D} \setminus \mathscr{D}_0$ as 'large' diffeomorphisms. An example of a large diffeo-

Teichmüller spaces will be crucial in our study of $(2+1)$-dimensional General Relativity. They satisfy the following wonderful property, discovered by Riemann himself.

Proposition 4.4.2. *The Teichmüller space of a Riemann surface* Σ *with genus* $\mathrm{g}(\Sigma)$ *has the following dimension:*

$$\dim \mathscr{T} = \begin{cases} 0 & \text{if } \mathrm{g}(\Sigma) = 0; \\ 2 & \text{if } \mathrm{g}(\Sigma) = 1; \\ 6\mathrm{g}(\Sigma) - 6 & \text{if } \mathrm{g}(\Sigma) > 1. \end{cases}$$

Proof. This is a standard result in complex geometry. See [FT84], [Dir51] or Chapter 0 of [Tro92] for details. □

Now that we have a function space of finite dimension, we need to link it back to our study of spaces of metrics. We will use our space of almost complex structures $\mathscr{A}$ as an intermediary, recalling that by Proposition 2.5.1 these are equivalent (in 2 dimensions) to complex structures via a bijection $\varpi : \mathscr{C} \to \mathscr{A}$. With this in mind, we can act on $\mathscr{A}$ at a point $p \in \Sigma$ with diffeomorphisms $f \in \mathscr{D}$ via

$$(f^* J)_p := (f_*)_p^{-1} J_{f(p)} (f_*)_p, \tag{4.4}$$

which behaves nicely, as stated in the following proposition.

Proposition 4.4.3. *The bijection* $\varpi : \mathscr{C} \to \mathscr{A}$ *from Proposition 2.5.1 is diffeomorphism-equivariant: for* $f \in \mathscr{D}$*, we have*

$$\varpi(f^* c) = f^* \varpi(c),$$

where the left-hand pull-back action is from Proposition 4.4.1 and the right-hand action is (4.4).

Proof. Let $\varphi \in c$ and $\varphi \circ f \in f^* c$ for some $f \in \mathscr{D}$. Then,

$$(\varphi \circ f)_*^{-1} \hat{J} (\varphi \circ f)_* = f_*^{-1} (\varphi_*^{-1} \hat{J} \varphi_*) f_*.$$

Now, $\varpi(f^* c)$ is the left-hand side and $f^* \varpi(c)$ is the right-hand side, so we are done. □

We next recall the discussion in Section 2.5 that described almost complex structures are being the bridge between Riemannian 2-manifolds and Riemann surfaces. Importantly, for an area element $\mathrm{d}A_g$ associated to a metric $g \in \mathscr{M}$, we can consider the following map:[21]

$$\Phi : \mathscr{M} \to \mathscr{A}, \quad \Phi(g) = -g^{-1} \mathrm{d}A_g.$$

morphism on a torus is a **Dehn twist**: the handle is cut along a closed curve (giving a cylinder), then one end is twisted by 2π before re-gluing. See Appendix A of [Car03] for details about the Dehn twist. Sometimes in the Riemannian context, $\mathscr{R}$ is called conformal superspace and $\mathscr{T}$ is quantum conformal superspace (see [FM96a], for example), which underlines the ability of the quantum context to differentiate things that the classical case cannot. These spaces represent the degrees of freedom of a system, as discussed in [Dir50] and [Dir51].

[21] This is an explicit description of the notion that metrics provide measurement of angles, and this induces the notion of rotation in a tangent space by $\frac{\pi}{2}$, which is an almost complex structure.

Locally, this takes the form $\Phi(g)^i_j = -g^{ik}\left(\mathrm{d}A_g\right)_{jk}$, which can be verified to be an almost complex structure. It can be shown that it is invariant under quotient by a positive function (that is, we can pass to the quotient $\mathscr{M}/\mathcal{C}^\infty$ by the conformal action), and that the resulting function $\Phi : \mathscr{M}/\mathcal{C}^\infty \to \mathscr{A}$ is a diffeomorphism of Hilbert manifolds.[22]

The discussion on the previous pages can be summarised in the following series of diffeomorphisms (as Hilbert manifolds):

$$\frac{\mathscr{M}/\mathcal{C}^\infty}{\mathscr{D}_0} \cong \frac{\mathscr{A}}{\mathscr{D}_0} \cong \frac{\mathscr{C}}{\mathscr{D}_0} =: \mathscr{T}, \tag{4.5}$$

where we have not included a Sobolev parameter s, which could appear in the numerators of the first two steps. This equivalence allows us to speak of the conformal classes of solution metrics to the Einstein equations (once they have been split, as was described at the beginning of this section) as elements of a Teichmüller space. By Proposition 4.4.2, this space is now finite-dimensional, which is a great reduction from the infinite-dimensional beginnings of $\mathscr{M}$.

The numerator of the left-most space in (4.5) begs to be simplified by the Uniformisation Theorem. As remarked at the end of Chapter 2, its statement allows us to write

$$\mathscr{M}/\mathcal{C}^\infty \cong \mathscr{M}_\lambda, \quad \text{where} \quad \lambda = \begin{cases} 1 & \text{if } \mathrm{g}(\Sigma) = 0; \\ 0 & \text{if } \mathrm{g}(\Sigma) = 1; \\ -1 & \text{if } \mathrm{g}(\Sigma) > 1. \end{cases} \tag{4.6}$$

Combining this expression of the Uniformisation Theorem with our knowledge of the dimension of the Teichmüller space from Proposition 4.4.2 and our equivalences (4.5) gives us the following:

$$\frac{\mathscr{M}_1}{\mathscr{D}_0} \cong \mathbb{R}^0, \quad \frac{\mathscr{M}_0}{\mathscr{D}_0} \cong \mathbb{R}^2 \quad \text{and} \quad \frac{\mathscr{M}_{-1}}{\mathscr{D}_0} \cong \mathbb{R}^{6\mathrm{g}-6}.$$

This is suggestive of the notion that the zero genus case of $(2+1)$-dimensional General Relativity is empty, because there are no dimensions to work with.[23] We will discuss this further in several sections, when we analytically inspect the three genus-dependent cases.

For the final portion of this section we will consider the $\mathrm{g}(\Sigma) > 1$ case, which will be the one that we inspect most closely in the sections to come. For reasons that will become clearer in future sections, we wish to scrutinise the quotient map $\mathscr{M}_{-1} \to \mathscr{M}_{-1}/\mathscr{D}_0$. It turns out that it comes with a fibre bundle structure. In fact, it is a principle bundle, which uses Lie group actions to affix a space to each point of a manifold, instead of vector spaces, as for vector bundles.[24]

Definition 4.4.2 (Principal bundle)**.** For a Lie group[25] $\mathcal{G}$ (called the **structure group**), a **principal bundle**, or principal $\mathcal{G}$-bundle, over a manifold $\mathcal{M}$ is an open surjective map $\pi : \mathcal{P} \to \mathcal{M}$ such that the following conditions hold:

[22]It is more correct to work with a Sobolev parameter s here, and state that $\Phi : \mathscr{M}^s/\mathscr{H}^s \to \mathscr{A}^s$ is the function in question. However, it reduces to the smooth case without too much difficulty—see [FT84] or Chapter 1 of [Tro92] for details.

[23]And thus by a footnote from Definition 4.4.1, no degrees of freedom.

[24]For details of vector bundles, see Section A.2 of the appendix.

[25]Recall that a **Lie group** is a smooth manifold $\mathcal{G}$ that is also a group: its inversion and composition maps must be smooth. Symbolically, we require that the map $\Upsilon : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ given by $\Upsilon(\chi, \upsilon) = \chi^{-1}\upsilon$ is smooth.

1. For every point $p \in \mathcal{M}$, $\mathcal{G}$ acts (on the right) freely[26] on the fibres $\mathcal{P}_p := \pi^{-1}(p)$;[27]

2. Every point $p \in \mathcal{M}$ has a neighbourhood $\mathcal{U} \subset \mathcal{M}$ with corresponding local trivialisation $\varphi_{\mathcal{U}} : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times \mathcal{G}$ such that the following diagram commutes:

$$\begin{array}{ccc} \pi^{-1}(\mathcal{U}) & \xrightarrow{\varphi_{\mathcal{U}}} & \mathcal{U} \times \mathcal{G} \\ & \searrow^{\pi} & \downarrow \mathrm{pr}_{\mathcal{U}} \\ & & \mathcal{U} \end{array}$$

   and the local trivialisations are compatible with the $\mathcal{G}$ action.

A **principal bundle morphism** is a map $\phi : \mathcal{P} \to \mathcal{Q}$ between principal $\mathcal{G}$-bundles $\pi : \mathcal{P} \to \mathcal{M}$ and $\varrho : \mathcal{Q} \to \mathcal{N}$ that commutes with the $\mathcal{G}$-action. Note that if $\mathcal{G}' \subset \mathcal{G}$ is a closed compact subgroup of $\mathcal{G}$ then the quotient map $\mathcal{G} \to \mathcal{G}/\mathcal{G}'$ is a trivial principal $\mathcal{G}'$-bundle, and (like with vector bundles) a bundle is trivial if and only if a global section can be found.[28]

This detour is useful because of the following result, which we state without proof.

Proposition 4.4.4. *The quotient map $\mathscr{M}_{-1} \to \mathscr{M}_{-1}/\mathscr{D}_0$ defines a trivial principal bundle.*

Proof. See [Mon89] or Chapter 2.5 of [Tro92]. □

The use of the above result will become apparent when we consider a global section of it (whose existence is guaranteed by triviality), which will help us to see the Einstein equations in $(2+1)$ dimensions as a Hamiltonian evolution on the cotangent bundle of a Teichmüller space. With this goal in mind, we turn to our time-space split of the Einstein equations.

4.5 ADM Decomposition

The time-space split of the Einstein equations employs the Hamiltonian formalism of mechanics, of which we give a brief reminder here.

Definition 4.5.1 (Hamiltonian mechanics)**.** On a closed n-manifold $\mathcal{M}$, the **action** $\mathcal{S}$ is given by the integral over $\mathcal{M}$ of the **Lagrangian** (density) $\mathcal{L}$:

$$\mathcal{S} = \int_{\mathcal{M}} \mathcal{L} = \int L(t)\mathrm{d}t,$$

where $L(t)$ is the Lagrangian associated to the Lagrangian density $\mathcal{L}$. This describes the **Lagrangian mechanics** of a physical system via the **principle of least action**, wherein the variation of the action is equal to zero.[29] This gives rise to the **Euler-Lagrange equations**,

[26] Recall that a group action is **free** if the only element that fixes points is the identity.

[27] Note that the fibres inherit a $\mathcal{G}$ action that is free and **transitive** (any two elements can be linked by a group element acting one of them).

[28] In physics, $\mathcal{G}$-bundles are useful in gauge theory, and a global section of a principal bundle corresponds to a global choice of gauge.

[29] Note that all of these details can be made precise using the calculus of variations, which we will not go into here. See, for example, Chapter IV of [CH53] or Chapter 8 of [Eva10].

which express the evolution of **generalised coordinates** $(q^i)_i$ in a parameter λ (often written t to signify time) as

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = \frac{\partial L}{\partial q^i}$$

where we write $\dot{q}^i := \partial_\lambda q^i$.

If we wish to change the dependence of our dynamics on the variables $(\dot{q}^i)_i$ to variables $(p_i)_i$, known as **generalised momenta**, or conjugate momenta, given by

$$p_i := \frac{\partial L}{\partial \dot{q}^i},$$

then we perform a Legendre transformation[30] on the Lagrangian to find the **Hamiltonian** H given by

$$H := \dot{q}^i p_i - L.$$

Now the dynamics of the system evolve via the **Hamilton-Jacobi equations**:

$$\dot{q}_i = \frac{\partial H}{\partial p^i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q^i},$$

which give rise to **Hamiltonian mechanics**.

Suppose a particular generalised coordinate $\mathring{q}$ only appears in the Lagrangian via a multiplicative factor:

$$L = L' + \mathring{q}\Omega$$

where L' does not depend on $\mathring{q}$, and its λ-derivative does not appear at all (implying that it has no conjugate momentum). We call such a coordinate a **Lagrange multiplier**, since it implies the existence of a **constraint**:

$$\Omega = 0,$$

which is obtained by either the Euler-Lagrange or Hamilton-Jacobi equations. Note too that this process allows one to include a desired constraint to a system, by adding a Lagrange multiplier term to the Lagrangian.

Finally, the pairs $(q^i, p_i)_i$ parametrise what is known as the **phase space**, which is a manifold in its own right. The submanifold of the phase space where the constraints (if any) are satisfied is known as the **constraint surface**.

We recall from Definition 4.1.2 that the action in the context of General Relativity is the Einstein-Hilbert action, given on a Lorentzian 3-manifold $(\mathcal{M}, g)$ by

$$\mathcal{S}^{\mathrm{EH}} = \int_{\mathcal{M}} (R - 2\Lambda)\,. \tag{4.7}$$

If we consider $\mathcal{M}$ to be a closed spacetime with a compact Cauchy surface Σ, then this integral can be split over $[0,1]$ and Σ.[31] With this in mind, we hope to decompose our

[30] A **Legendre transformation** has a precise definition which we will not explore here, merely giving the transition from Lagrangian to Hamiltonian mechanics as an example of its use. Details can be found in Chapter 3.3 of [Eva10].

[31] We have assumed our manifold to be compact; in general, this interval can have bounds t_{i} and t_{f}, where the initial and final times may be infinite in extent.

spacetime metric g and our Ricci scalar R into forms which involve the metric and Ricci scalars on Σ. Having two versions of objects, one residing on $\mathcal{M}$ and the other on Σ may cause confusion and for this reason we introduce **prescript notation**: *the pre-superscript denotes the dimensionality of an object*, whenever such confusion may arise. Thankfully, objects with explicit indices will often be safe from confusion, as those with Greek indices include the time coordinate and those with Latin indices are purely spatial.

We choose coordinates adapted to the foliation $\mathcal{M} \cong \cup_t \Sigma_t$, such that the zeroth coordinate t is the parameter of the foliation Σ_t, and local spatial coordinates (x^1, x^2) on the spacelike time-slices Σ_t so that there are natural tangent vectors $(\partial_i)_{i=1,2}$ to Σ_t and a natural 1-form dt. In this coordinate system, we can now let the lapse $N(t, x^i)$ be

$$N = \frac{1}{\sqrt{-g(\nabla t, \nabla t)}},$$

where we could have written $\mathtt{t}$ to refer back to our time function $\mathtt{t} : \mathcal{M} \to \mathbb{R}$, but we have used t for notational simplicity. In partnership to the notion of lapse, the shift vector $N^i(t, x^i)$ is defined such that $\partial_0 - N^i(t, x^i)\partial_i$ (where ∂_0 is the vector dual to the covector dt) is orthogonal to Σ_t. Using these we can write the overall spacetime metric as[32]

$${}^{(3)}g = -N^2 \mathrm{d}t^2 + {}^{(2)}g_{ij} \left(\mathrm{d}x^i + N^i \mathrm{d}t \right) \left(\mathrm{d}x^j + N^j \mathrm{d}t \right).$$

More compactly, we make the following definitions, valid in the coordinate system described above, which we will assume to hold in the text that follows.

Definition 4.5.2 (Lapse and shift)**.** For a spacetime $(\mathcal{M}, g)$ in $(2+1)$ dimensions (with a Cauchy surface Σ) and coordinates described above, we define the **lapse** and **shift** as

$$N := \frac{1}{\sqrt{-\,{}^{(3)}g^{00}}} \quad \text{and} \quad N_i := {}^{(3)}g_{0i}.$$

Note that the shift will have its index raised and lowered by ${}^{(2)}g$, since it only runs through spatial directions. These give the spacetime metric and its inverse the following forms:

$${}^{(3)}g = \begin{pmatrix} -N^2 + N^i N_i & N_i \\ N_i & {}^{(2)}g_{ij} \end{pmatrix} \quad \text{and} \quad {}^{(3)}g^{-1} = \frac{1}{N^2} \begin{pmatrix} -1 & N^i \\ N^i & {}^{(2)}g^{ij} N^2 - N^i N^j \end{pmatrix}.$$

Note that normal coordinates from Definition 2.1.6 here mean taking shift $N^i = 0$; further, **geodesic coordinates** mean taking lapse $N = 1$. Together (that is, **geodesic normal coordinates**), these simplify the relation between the metric of the spacetime and the submanifold to

$${}^{(3)}g^{\mathrm{gnc}} = -\mathrm{d}t^2 + {}^{(2)}g_{ij} \mathrm{d}x^i \mathrm{d}x^j,$$

with both the spacetime metric and its inverse being block diagonal.

The following result (whose proof we omit) allows us to neglect the shift in many of our computations, as it shows that it is always possible to choose geodesic coordinates globally.

[32] These are the same lapse and shift as appeared in our BTZ metric from Example 4.1.3.

Proposition 4.5.1. *Let $(\mathcal{M}, g)$ an n-dimensional globally hyperbolic spacetime with Cauchy surface Σ. Then it is isometric to the product $\mathbb{R} \times \Sigma$, where the metric takes the form*

$$^{(n)}g = -N^2 \mathrm{dt}^2 + {}^{(n-1)}g,$$

where $^{(n-1)}g$ is a Riemannian metric on Σ, N is non-zero, and $\mathtt{t} : \mathbb{R} \times \Sigma \to \mathbb{R}$ is the projection (and a time function).

Proof. See Theorem 1.1 of [BS05]. □

Because of this, we will not focus our attention on the lapse and the shift, knowing that in the end we will want to use geodesic normal coordinates. Our final ingredient is the extrinsic curvature: if η^μ is the normal to the Cauchy surface Σ, we recall that the extrinsic curvature $\kappa_{\mu\nu}$ (always written with indices to differentiate it from its trace κ, the mean extrinsic curvature) of Σ is given locally by

$$\kappa_{\mu\nu} = {}^{(3)}\nabla_\mu \eta_\nu. \tag{4.8}$$

This arises from Definition 4.3.2, which, in taking a contraction with the normal vector, only preserves components normal to the submanifold. We will choose a normal proportional to the zeroth component (the 'time' direction), and as such we will focus solely on the components κ_{ij}.

We now have all of the tools for the Hamiltonian treatment of General Relativity.

Definition 4.5.3 (ADM formalism)**.** Consider a $(2+1)$-dimensional spacetime $(\mathcal{M}, g)$ with lapse N, shift N^i, and Cauchy surface Σ (whose metric is $^{(2)}g$). The **ADM formalism**, named after Arnowitt, Deser and Misner,[33] uses a normal to the Cauchy surface that points in the 'time' direction: $\eta_\mu = (N, 0, 0)$. This gives the following form to the extrinsic curvature of Σ (which will be proved shortly):

$$\kappa_{ij} = -\frac{1}{2N}\left(\partial_t g_{ij} - {}^{(2)}\nabla_i N_j - {}^{(2)}\nabla_j N_i\right). \tag{4.9}$$

This decomposition splits the action integral over $\mathcal{M}$ as integrals over $[0,1]$ and Σ. Using (traces of) the Gauss-Codazzi equations (4.2), our Einstein-Hilbert action (4.7) takes the form

$$^{(3)}\mathcal{S}^{\mathrm{EH}} = \int_{[0,1]} \mathrm{d}t \int_\Sigma \mathrm{d}^2x \sqrt{^{(2)}g}N\left({}^{(2)}R - 2\Lambda + \kappa_{ij}\kappa^{ij} - \kappa^2\right), \tag{4.10}$$

where we note that indices are raised and lowered by $^{(2)}g$, as will be the custom for any object living in 2 dimensions from now on, and boundary terms have been disregarded as they do not contribute. Note that the time-derivatives of the shift and the lapse do not appear in the Lagrangian, so they do not have conjugates and are thus treated as Lagrange multipliers.

Now, the **gravitational momentum** π^{ij} conjugate[34] to g_{ij} is given by[35]

$$\pi^{ij} := \frac{\partial}{\partial(\partial_t g_{ij})}\left(\sqrt{^{(2)}g}N\left({}^{(2)}R - 2\Lambda + \kappa_{ij}\kappa^{ij} - \kappa^2\right)\right) = \sqrt{^{(2)}g}\left(\kappa^{ij} - g^{ij}\kappa\right), \tag{4.11}$$

[33] Their seminal work [ADM59] was in $(3+1)$ dimensions, though, and has been adapted accordingly.

[34] The conjugate to (k,l)-tensor appearing in the Lagrangian is a (l,k)-tensor.

[35] Note that this is actually a **tensor density**, which is defined to be a tensorial object that picks up a factor of the Jacobian during a change of coordinates. Here, the factor of $\sqrt{^{(2)}g}$ in its definition makes it as such, which arises from us using the Lagrangian instead of the Lagrangian density in its definition. However, the difference between tensors and tensor densities in our discussion is not sufficient enough to warrant lengthy investigation, so we will ignore it for now.

where only the extrinsic curvature contains terms with $\partial_t g_{ij}$ so it alone contributes. It is symmetric and we write its trace π. (Note that a choice of sign convention for the extrinsic curvature may impact the sign of the conjugate, but we are consistent in our choices in this text.)

Finally, we define the **ADM action**:

$$ {}^{(3)}\mathcal{S}^{\mathrm{ADM}} := \int_{[0,1]} \mathrm{d}t \int_{\Sigma} \mathrm{d}^2 x \left(\pi^{ij}\partial_t g_{ij} - N\mathcal{H} - N_i\mathcal{H}^i \right). \tag{4.12} $$

In the above expression we have made two important definitions: the **Hamiltonian constraint** $\mathcal{H}$ and the **momentum constraint** $\mathcal{H}^i$, which are given by

$$ \mathcal{H} := \frac{1}{\sqrt{{}^{(2)}g}} \left(\pi_{ij}\pi^{ij} - \pi^2 \right) - \sqrt{{}^{(2)}g} \left({}^{(2)}R - 2\Lambda \right) \quad \text{and} \quad \mathcal{H}^i := -2\,{}^{(2)}\nabla_j \pi^{ij}. \tag{4.13} $$

These are named as such because the variation of the shift and lapse within the action (acting as Lagrange multipliers) give rise to the Hamiltonian and momentum constraints:

$$ \mathcal{H} = 0 \quad \text{and} \quad \mathcal{H}^i = 0. \tag{4.14} $$

As usual in the Hamiltonian formalism, the pairs $(g_{ij}, \pi^{ij})_{i,j}$ parametrise the phase space, and the solutions to the above constraints will lie on the constraint surface. Finally, the form above of the ADM action gives our first glimpse of the Hamiltonian:

$$ H = \left(\partial_t g_{ij}\right) \pi^{ij} - \mathcal{L} = N\mathcal{H} + N_i\mathcal{H}^i, $$

which vanishes when the constraints are satisfied (and where are usual we use that Σ is closed to ignore boundary terms).[36]

We state the above as results, which arise from standard manipulations.

Proposition 4.5.2. *For the definitions given above, we have the following:*

1. *The extrinsic curvature from* (4.8) *does reduce to the form in* (4.9)*;*
2. *The ADM action* (4.12) *is equal to the Einstein-Hilbert action* (4.10).

Proof. 1. This proof is computational, uses the definition of the Levi-Civita connections for the metrics ${}^{(3)}g$ and ${}^{(2)}g$.[37] For $\eta_\mu = (N, 0, 0)$, the purely-spatial components of the extrinsic curvature are

$$ \begin{aligned} \kappa_{ij} &= {}^{(3)}\nabla_i \eta_j := \partial_i \eta_j - {}^{(3)}\Gamma^{\alpha}_{ij}\eta_\alpha \\ &= -\frac{1}{2} N g^{0\beta} \left(\partial_i g_{\beta j} + \partial_j g_{\beta j} - \partial_\beta g_{ij} \right) \\ &= -\frac{1}{2} N \left(-\frac{1}{N^2}\partial_i N_j + \frac{N^k}{N^2}\partial_i g_{jk} - \frac{1}{N^2}\partial_j N_i + \frac{N^k}{N^2}\partial_j g_{ik} + \frac{1}{N^2}\partial_0 g_{ij} - \frac{N^k}{N^2}\partial_k g_{ij} \right) \\ &= -\frac{1}{2N} \left(\partial_0 g_{ij} - \partial_i N_j - \partial_j N_i + N^k \left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij} \right) \right) \\ &= -\frac{1}{2N} \left(\partial_t g_{ij} - {}^{(2)}\nabla_i N_j - {}^{(2)}\nabla_j N_i \right), \end{aligned} $$

as desired.

[36] This gives rise to the statement that a closed $(2+1)$-dimensional universe has zero total energy.

[37] See Section 2.1 for details.

2. This proof will be accomplished in several steps. First, we invert the definition of π^{ij} (4.11):

$$\kappa_{ij} = \frac{1}{\sqrt{{}^{(2)}g}} \left(\pi_{ij} - g_{ij}\pi\right). \tag{4.15}$$

Next, using the forms of the extrinsic curvature (4.9) and (4.15), we have

$$\begin{aligned} \pi^{ij}\partial_t g_{ij} &= \pi^{ij}\left(2N\kappa_{ij} + {}^{(2)}\nabla_i N_j + {}^{(2)}\nabla_j N_i\right) \\ &= 2N\frac{1}{\sqrt{{}^{(2)}g}}\left(\pi^{ij}\pi_{ij} - \pi^2\right) + N_i\left(-2\nabla_j\pi^{ij}\right), \end{aligned}$$

where we have used that π^{ij} is symmetric and have integrated by parts (discarding the boundary term that appears as it does not contribute to the action), noticing the momentum constraint appearing in the right-most term. Introducing the Hamiltonian constraint (4.13) gives

$$\pi^{ij}\partial_t g_{ij} = 2N\mathcal{H} + 2N\sqrt{{}^{(2)}g}\left({}^{(2)}R - 2\Lambda\right) + N_i\mathcal{H}^i.$$

We can then write the ADM action (4.12) as

$$\begin{aligned} {}^{(3)}\mathcal{S}^{\text{ADM}} &= \int_{[0,1]} \mathrm{d}t \int_\Sigma \mathrm{d}^2x \left(\pi^{ij}\partial_t g_{ij} - N\mathcal{H} - N_i\mathcal{H}^i\right) \\ &= \int_{\mathcal{M}} \mathrm{d}^3x\, 2N\left(N\mathcal{H} + \sqrt{{}^{(2)}g}\left({}^{(2)}R - 2\Lambda\right)\right) \\ &= \int_{\mathcal{M}} \mathrm{d}^3x \sqrt{{}^{(2)}g}N\left({}^{(2)}R - 2\Lambda + \frac{1}{\det{}^{(2)}g}\left(\pi_{ij}\pi^{ij} - \pi^2\right)\right). \end{aligned}$$

Finally, using the relation between the extrinsic curvature and the gravitational momentum (4.11) once more, we find

$$\begin{aligned} {}^{(3)}\mathcal{S}^{\text{ADM}} &= \int_{\mathcal{M}} \mathrm{d}^3x \sqrt{{}^{(2)}g}N\left({}^{(2)}R - 2\Lambda + \frac{1}{\det{}^{(2)}g}\left(\pi_{ij}\pi^{ij} - \pi^2\right)\right) \\ &= \int_{\mathcal{M}} \mathrm{d}^3x \sqrt{{}^{(2)}g}N\left({}^{(2)}R - 2\Lambda + \left(\kappa_{ij} - g_{ij}\kappa\right)\left(\kappa^{ij} - g^{ij}\kappa\right) - \left(g^{ij}\left(\kappa_{ij} - g_{ij}\kappa\right)\right)^2\right) \\ &= \int_{[0,1]} \mathrm{d}t \int_\Sigma \mathrm{d}^2x \sqrt{{}^{(2)}g}N\left({}^{(2)}R - 2\Lambda + \kappa_{ij}\kappa^{ij} - \kappa^2\right) = {}^{(3)}\mathcal{S}^{\text{EH}}, \end{aligned}$$

as desired. □

The ADM form of the action is useful because the constraints are explicitly present, and thus once we assume our solution to the dynamical equations to satisfy them (that is, we assume they vanish), we will easily be able to simplify the action.

We now divide our attention between three cases, which arise from characterising the Cauchy surface Σ as having a genus of 0, 1, or greater than 1. This split arises from our Uniformisation Theorem (4.6), and much like our proof in Chapter 3, the different cases require vastly different amounts of work.

4.6 Genus Case Studies: Higher Genera Surfaces

When the cosmological constant Λ is zero in $(2+1)$ dimensions, the vacuum Einstein equations (4.1.1) imply that not only is spacetime Ricci-flat, it is (Riemann-)flat, since we know that the Riemann tensor and the Ricci tensor are related by (2.8). If $\Lambda \neq 0$, then we obtain a spacetime of constant curvature 6Λ. Physically, these facts imply that there do not exist gravitational waves in $(2+1)$ dimensions, as we require mass to create curvature in the space. Physicists would say that the theory has no local degrees of freedom, as was discussed in Section 4.4. Globally, there are no degrees of freedom (hence a vacuous problem) if the space has trivial fundamental group—that is, if it has zero genus.[38] This is discussed in depth in [Car03].

From another perspective, one could guess from our Teichmüller space exploration in Section 4.4, the case of a $(2+1)$-dimensional spacetime with a Cauchy surface Σ that has genus $g(\Sigma) = 0$ is simple. The dimension of the Teichmüller space is zero, leaving no possible non-trivial dynamics.

If the spacetime has non-trivial fundamental group, then one can think of the dimension of the Teichmüller space as counting the (finite!) number of global degrees of freedom. It turns out that in the unit genus case, the explicit form of the Hamiltonian can be found, as is done in Chapter 3.3 of [Car03] as well as in [Mon89]. Despite this, we will first focus our efforts the higher genera case ($g > 1$), before returning to the case $g(\Sigma) = 1$ in the following section to comment on its solution.

We wish to demonstrate that for spacetimes with Cauchy surfaces of genus $g(\Sigma) > 1$ the above ADM dynamics actually occur (given certain assumptions) on the cotangent bundle of the Teichmüller space of Σ, which would align with our earlier analysis. In doing so, we will find an implicitly-defined form of the Hamiltonian of the system.

First, we have the following definition of a particularly nice kind of tensor, of which we will make great use of in the coming pages.

Definition 4.6.1 (Transverse and traceless)**.** On a semi-Riemannian manifold (Σ, g) with Levi-Civita connection ∇, a rank-2 tensor τ is **transverse** if $\nabla_i \tau^{ij} = 0$ and **traceless** if $\tau^i_i = 0$. *We denote the transverse and traceless part of a tensor τ^{ij} with an over-line: $\overline{\tau^{ij}}$.*[39]

Note that transverse and traceless tensors[40] are transverse and traceless with respect to the whole conformal class of the metric with respect to which they are originally transverse and traceless. That is, if $\overline{\tau^{ij}}$ is transverse and traceless with respect to g then it is with respect to $e^{-2\lambda} g$ for any $\lambda \in C^\infty(\Sigma)$.

We now have the following Lemma, which will allow us to decompose symmetric tensors into their transverse and traceless part and two other parts.

Lemma 4.6.1. *Any symmetric rank-2 tensor τ on a closed semi-Riemannian 2-manifold (Σ, g) can be decomposed as*

$$\tau_{ij} = \overline{\tau_{ij}} + f g_{ij} + \left(\nabla_i Y_j + \nabla_j Y_i - g_{ij} \nabla_k Y^k \right), \tag{4.16}$$

*for some vector field Y and $f \in C^\infty(\Sigma)$. The term in parentheses is known as the **conformal Killing form** of Y and is traceless. Note that the decomposition is unique up to the addition of*

[38] Recall that the **fundamental group** π_1 is the group of equivalence classes (called **homotopy classes**) under homotopy of closed curves on a topological space.

[39] In the literature this is often written $\tau^{ij^{\mathrm{TT}}}$, which is cumbersome, though more descriptive.

[40] Symmetric transverse traceless tensors are called **holomorphic quadratic differentials** in the study of Riemann surfaces.

another conformal Killing vector to Y. We also have that the three terms in the decomposition are $\mathscr{L}^2$-orthogonal, and $f = \frac{1}{2}g_{ij}\tau^{ij}$.

Proof. We have $f = \frac{1}{2}g_{ij}\tau^{ij}$ by tracing both sides of (4.16), and the pair-wise $\mathscr{L}^2$-orthogonality because the first and third terms are traceless and so vanish when integrated against the second term, and their orthogonality arises from the computation

$$\int_\Sigma \overline{\tau^{ij}} \left(\nabla_i Y_j + \nabla_j Y_i - g_{ij}\nabla_k Y^k\right) = \int_\Sigma \nabla_i \left(\overline{\tau^{ij}} Y_j\right) + \nabla_j \left(\overline{\tau^{ij}} Y_i\right) - \left(\overline{\tau^{ij}} g_{ij}\right) \nabla_k Y^k = 0,$$

where we have used that the manifold has no boundary to integrate out a total derivative and that $\overline{\tau^{ij}}$ is transverse and traceless. □

Now, we have the following Lemma, which relates various objects relating to conformal metrics.

Lemma 4.6.2. *For conformal Riemannian metrics g and $\hat{g} := \mathrm{e}^{2\lambda} g$ on a 2-manifold Σ, we have the following relations (where we use hats to designate objects associated to $\hat{g}$):*

1. *The determinants* $\det g$ *and* $\det \hat{g}$ *are related by*

$$\det \hat{g} = \left(\mathrm{e}^{2\lambda}\right)^2 \det g;$$

2. *For vectors V and W, the Levi-Civita connections ∇ and $\hat{\nabla}$ are related by*

$$\hat{\nabla}_V W = \nabla_V W + V(\lambda)W + W(\lambda)V - g(V,W)\mathrm{grad}(\lambda),$$

   *where the **gradient** of a function λ is the vector field* $\mathrm{grad}(\lambda)$ *given by*

$$\mathrm{grad}(\lambda) := g^\sharp(\mathrm{d}\lambda);$$

3. *The Christoffel symbols Γ^k_{ij} and $\hat{\Gamma}^k_{ij}$ are related by*

$$\hat{\Gamma}^k_{ij} = \Gamma^k_{ij} + (\partial_i\lambda)\delta^k_j + (\partial_j\lambda)\delta^k_i - (g^{kl}\partial_l\lambda)g_{ij};$$

4. *For a vector field V, the weighted (by the square-root of the determinant of the metric) conformal Killing forms are invariant:*

$$\sqrt{\hat{g}}\left(\hat{\nabla}^i V^j + \hat{\nabla}^j V^i - \hat{g}^{ij}\hat{\nabla}_k V^k\right) = \sqrt{g}\left(\nabla^i V^j + \nabla^j V^i - g^{ij}\nabla_k V^k\right),$$

   where the left-hand side's indices are raised and lowered by $\hat{g}$ and the right-hand side's by g;

5. *The Ricci scalars R and $\hat{R}$ are related by*[41]

$$\hat{R} = \mathrm{e}^{-2\lambda}\left(-2\Delta_g\lambda + R\right).$$

[41] If $\mathrm{g}(\mathcal{M}) > 1$ then the equation

$$\Delta_g\lambda = \frac{1}{2}\left(R + \mathrm{e}^{2\lambda}\right),$$

has a unique solution and thus by the final point $\hat{g}$ can be taken to have $\hat{R} = -1$. This is an alternate way of looking at the $\mathrm{g}(\mathcal{M}) > 1$ case of the Uniformisation Theorem, and can be proved with partial differential equation methods. For details, see Chapter 10 of [Don11].

Proof. Each point can be proved by direct computation, starting from $\hat{g} := \mathrm{e}^{2\lambda} g$ and using the definitions of each quantity. We have not done it here because it is not illuminating. □

We now restrict ourselves to the case where $\mathrm{g}(\Sigma) > 1$*, and we write* $\hat{g}$ *for a Riemannian metric on* Σ *with associated Ricci scalar* $\hat{R} = -1$. Note that since $\hat{g}$ is only defined in two dimensions we will not need prescript notation.

We now decompose our gravitational momentum in the following result.

Proposition 4.6.1. *Gravitational momentum* π^{ij} *on a Cauchy surface* Σ *(with* $\mathrm{g} > 1$*) of constant mean extrinsic curvature* κ *(that is,* $\partial_i \kappa = 0$*) that solves the momentum constraint* $\mathcal{H}^i = 0$ *can be decomposed as*

$$\pi^{ij} = \overline{\pi^{ij}} - \frac{1}{2}\kappa\sqrt{{}^{(2)}g}g^{ij}. \tag{4.17}$$

Proof. We first note that the trace of π^{ij} is

$$\pi := g_{ij}\pi^{ij} = -\sqrt{{}^{(2)}g}\kappa,$$

so Lemma 4.6.1 gives

$$\pi^{ij} = \overline{\pi^{ij}} - \frac{1}{2}\kappa\sqrt{{}^{(2)}g}g^{ij} + \mathrm{e}^{-2\lambda}\sqrt{{}^{(2)}g}\left({}^{(2)}\nabla^i Y^j + {}^{(2)}\nabla^j Y^i + g^{ij}\,{}^{(2)}\nabla_l Y^l\right). \tag{4.18}$$

Here, the vector Y^i is uniquely determined as the solution to the equation

$${}^{(2)}\nabla_i\left(\pi^{ij} + \frac{1}{2}\kappa\sqrt{{}^{(2)}g}g^{ij}\right) = {}^{(2)}\nabla_i\left(\mathrm{e}^{-2\lambda}\sqrt{{}^{(2)}g}\left({}^{(2)}\nabla^i Y^j + {}^{(2)}\nabla^j Y^i - g^{ij}\,{}^{(2)}\nabla_l Y^l\right)\right).$$

Now, if we impose the momentum constraint (4.14)—namely, asking π^{ij} to be transverse—and assume that Σ is a hypersurface of constant mean extrinsic curvature ($\partial_i\kappa = 0$), then the left-hand side of the above expression vanishes. We can then use Lemma 4.6.2, which asserts the invariance of conformal Killing forms under a change to a conformal metric. We are then left with

$$\nabla_i\left(\sqrt{\hat{g}}\left(\hat{g}^{il}\hat{\nabla}_l\left(\hat{g}_{jm}Y^m\right) + \hat{\nabla}_j Y^i - \delta^i_j\hat{\nabla}_m Y^m\right)\right) = 0.$$

The unique solution of this equation is $Y^i = 0$, since a compact 2-manifold with constant negative curvature has no non-trivial conformal Killing fields. (This is implied by Theorem 4.44 of [Bes07], as mentioned in Remark 3.3 of [BK17].) Thus our decomposition (4.18) simplifies to the desired expression (4.17). □

We now assume that the momentum constraint $\mathcal{H}^i = 0$ holds, as it did in the previous proposition. Seeing the usefulness of the transverse and traceless part of the gravitational momentum, we make the following definition.

Definition 4.6.2 (Conformal momentum)**.** The **conformal momentum** $\overline{p^{ij}}$ is a completely transverse and traceless (where we will keep the over-line to emphasise this property) symmetric rank-2 tensor (density) defined by

$$\overline{p^{ij}} := \mathrm{e}^{2\lambda}\overline{\pi^{ij}}. \tag{4.19}$$

Having already imposed the momentum constraint $\mathcal{H}_i = 0$, we now turn to the Hamiltonian constraint $\mathcal{H} = 0$.

Proposition 4.6.2. *In the case* $\mathrm{g} > 1$ *and using the definition of the conformal momentum* (4.19), *the Hamiltonian constraint* (4.14) *takes the form*

$$\Delta_{\hat{g}}\lambda = \beta_+ \mathrm{e}^{2\lambda} + \beta_- \mathrm{e}^{-2\lambda} + \beta_0, \tag{4.20}$$

where we have defined the coefficients

$$\beta_+ := \frac{1}{4}\kappa^2, \quad \beta_- := -\frac{1}{2}\frac{\hat{g}_{ik}\hat{g}_{jl}\overline{p^{ij}}\,\overline{p^{kl}}}{\det\hat{g}}, \quad \textit{and} \quad \beta_0 := \frac{1}{2}\hat{R} = -\frac{1}{2}.$$

Recalling that κ *is the trace of the extrinsic curvature, the solutions are classified into the following cases.*

1. *If* $\kappa^2 = 0$*, then there are no solutions;*

2. *If* $\kappa^2 > 0$*, and* $\hat{g}_{ij}$ *and* $\overline{p^{ij}}$ *are smooth, then a unique solution* λ *exists and is smooth;*

3. *If* $\kappa^2 > 0$*, and* $\hat{g}_{ij}$ *and* $\overline{p^{ij}}$ *are in* $\mathscr{H}^s$ *and* $\mathscr{H}^{s-1}$*, respectively, then a unique solution* λ *exists and is in* $\mathscr{H}^{s+1}$.

Proof. See [Mon86] for details. □

Working backwards, our procedure above can therefore be outlined as follows.

1. Choose a Riemannian metric $\hat{g}$ on a spacelike hypersurface Σ (with $\mathrm{g}(\Sigma) > 1$ and constant non-zero mean extrinsic curvature κ) with $\hat{R} = -1$, which exists by the Uniformisation Theorem;

2. Choose a symmetric transverse and traceless (with respect to $\hat{g}$) tensor (density) $\overline{p^{ij}}$;

3. Solve (4.20) for λ;

4. Set

$$g_{ij} = \mathrm{e}^{2\lambda}\hat{g}_{ij} \quad \text{and} \quad \pi^{ij} = \mathrm{e}^{-2\lambda}\overline{p^{ij}} - \frac{1}{2}\kappa\sqrt{{}^{(2)}g}\,g^{ij}. \tag{4.21}$$

In this way we go from the space $\mathscr{M}_{-1}$ and a choice of $\overline{p^{ij}}$ and obtain a metric in $\mathscr{M}$ and its conjugate momentum π^{ij}.

One may wonder what has happened to our lapse N and our shift N^i, even though we anticipate taking simple choices of these (if allowed). We have already required the time-slices to have constant non-zero mean extrinsic curvature κ, which can be thought of as a temporal gauge fixing, specifying the lapse. As a spatial gauge we require our metric $\hat{g}$ to remain within its original global section throughout the dynamics, which specifies the shift. We will state the explicit equations that N and N^i must satisfy in Section 4.8, though for now we assume that they are being satisfied.

Our next step is to translate the above steps into terms that involve our Teichmüller space $\mathscr{T}(\Sigma)$, which is diffeomorphic to the quotient $\mathscr{M}_{-1}/\mathscr{D}_0$ (we will omit the Sobolev parameter s and focus on smooth maps only). Recall that since our principal bundle $\mathscr{M}_{-1} \to \mathscr{M}_{-1}/\mathscr{D}_0$ is trivial by Proposition 4.4.4, it has global sections diffeomorphic to

$\mathbb{R}^{6g-6}$ (for genus $g := g(\Sigma)$). We set $(q^\alpha)_{\alpha=1}^{6g-6}$ to be coordinates on this space, allowing us to express any global section as a smooth set of metrics $\hat{g}$:

$$(q^\alpha)_{\alpha=1}^{6g-6} \to \left\{ \hat{g}_{ij}(x^k, q^\alpha) \quad \text{and} \quad \hat{R}\left(\hat{g}\left(q^\alpha\right)\right) = -1 \right\},$$

where we have chosen coordinates $(x^k)_{k=1,2}$ on the hypersurface. With points on our Teichmüller space in mind, we hope to find corresponding covectors, which will parametrise the cotangent bundle.[42] In this direction, we make the following definition.

Definition 4.6.3 (Momentum components)**.** Let $\hat{g}_{ij}(x^k, \mathring{q}^\alpha)$ be a point on a global section of the trivial bundle $\mathscr{M}_{-1} \to \mathscr{M}_{-1}/\mathscr{D}_0$, and $\overline{p^{ij}}(x^k)$ a symmetric transverse and traceless tensor (density) with respect to it. We define **momentum components** $(p_\alpha)_{\alpha=1}^{6g-6}$ of $\overline{p^{ij}}(x^k)$ conjugate to $(q^\alpha)_{\alpha=1}^{6g-6}$ to be

$$p_\alpha := \int_\Sigma \left(\overline{p^{ij}}(x^k) \frac{\partial \hat{g}_{ij}}{\partial q^\alpha}(x^k, q^\beta) \right) \bigg|_{q^\alpha = \mathring{q}^\alpha}.$$

Note that we will often write $(p_\alpha)_\alpha$ and $(q^\alpha)_\alpha$ without specifying the range of α, which will always be from 1 to $6g-6$ (when $g > 1$).

We now have the following result.

Proposition 4.6.3. *The components $(p_\alpha)_{\alpha=1}^{6g-6}$ uniquely determine $\overline{p^{ij}}(x^k)$.*

Idea of a proof. (See [Mon89] for details.)

This result follows from the fact that the space of symmetric transverse and traceless rank-2 tensor densities at a point $\hat{g}_{ij}(x^k, \mathring{q}^\alpha)$ has dimension $6g-6$, as is the space of tangent vectors $\frac{\partial \hat{g}_{ij}}{\partial q^\alpha}(x^k, q^\beta)\Big|_{q^\alpha = \mathring{q}^\alpha}$. Noting that $\overline{p^{ij}}(x^k)$ integrated over Σ against some tensor field $(\mathcal{L}_V \hat{g})_{ij}$ vanishes (for some vector field V on Σ), one can argue that p_α indeed specifies $\overline{p^{ij}}(x^k)$. □

With the above result we may consider $(p_\alpha)_\alpha$ to be components of a covector of our Teichmüller space $\mathscr{T}(\Sigma) \cong \mathbb{R}^{6g-6}$ above a point $(\mathring{q}^\alpha)_\alpha$. Thus, $(q^\alpha, p_\alpha)_{\alpha=1}^{6g-6}$ are coordinates on the cotangent bundle of the Teichmüller space: $\mathcal{T}^*\mathscr{T}(\Sigma) \cong \mathbb{R}^{12g-12}$. We know that the points of $\mathcal{T}^*\mathscr{T}(\Sigma)$ label the $\mathscr{D}_0$-equivalence classes of solutions to the constraint equations (4.14) with constant non-zero mean extrinsic curvature $\kappa \neq 0$.

Translating our previous procedure to include the current terms: for a spacelike hypersurface Σ with $g(\Sigma) > 1$ and constant non-zero mean extrinsic curvature κ, choosing a point $(q^\alpha, p_\alpha)_\alpha$ on $\mathcal{T}^*\mathscr{T}(\Sigma)$ determines a Riemannian metric $\hat{g}_{ij}(x^k, q^\alpha)$ with $\hat{R} = -1$ and a symmetric transverse and traceless (with respect to $\hat{g}$) tensor (density) $\overline{p^{ij}}(x^k)$. We can then solve (4.20) for λ and set our g_{ij} and π^{ij} as in (4.21).

Finally, we wish to determine the exact (implicit) form of the Hamiltonian whose dynamics resides on $\mathcal{T}^*\mathscr{T}(\Sigma)$. Note that the constraint terms $\mathcal{H}^i$ and $\mathcal{H}$ have vanished since we assume our solution parametrised by some $(x^k, q^\alpha, p_\alpha, \kappa)$ to solve them—we

[42]This procedure is a natural in the Hamiltonian formalism of mechanics and symplectic geometry in general. We will not explore the details of this field here, simply remarking that cotangent bundles are not so foreign as one may expect.

denote this assumption with an asterisk on the action. We substitute our expressions for π^{ij} and g_{ij} into the action (4.12) to find

$$\begin{aligned} {}^{(3)}\mathcal{S}^{\mathrm{ADM}*} &= \int_{[0,1]} \mathrm{d}t \int_{\Sigma} \mathrm{d}^2x \left(\pi^{ij}\partial_t g_{ij}\right) \\ &= \int_{\mathcal{M}} \mathrm{d}^3x \left(\mathrm{e}^{-2\lambda}\overline{p^{ij}} - \frac{1}{2}\kappa\sqrt{\hat{g}}\hat{g}^{ij}\right)\partial_t\left(\mathrm{e}^{2\lambda}\hat{g}_{ij}\right) \\ &= \int_{\mathcal{M}} \mathrm{d}^3x \left(\overline{p^{ij}}\partial_t\hat{g}_{ij} - \kappa\sqrt{\hat{g}}\partial_t\mathrm{e}^{2\lambda} - \frac{1}{2}\mathrm{e}^{2\lambda}\kappa\sqrt{\hat{g}}\hat{g}^{ij}\partial_t\hat{g}_{ij}\right), \end{aligned}$$

where we have simply expanded all of the terms. Now, using $\hat{g}^{ij}\partial_t\hat{g}_{ij} = -\hat{g}_{ij}\partial_t\hat{g}^{ij}$ and absorbing factors of $\kappa\sqrt{\hat{g}}$ into the time-derivatives, we have

$$\begin{aligned} {}^{(3)}\mathcal{S}^{\mathrm{ADM}*} &= \int_{\mathcal{M}} \mathrm{d}^3x \left(\overline{p^{ij}}\partial_t\hat{g}_{ij} - \kappa\sqrt{\hat{g}}\partial_t\mathrm{e}^{2\lambda} + \frac{1}{2}\mathrm{e}^{2\lambda}\kappa\sqrt{\hat{g}}\hat{g}_{ij}\partial_t\hat{g}^{ij}\right) \\ &= \int_{\mathcal{M}} \mathrm{d}^3x \left(\overline{p^{ij}}\partial_t\hat{g}_{ij} - \partial_t\left(\mathrm{e}^{2\lambda}\kappa\sqrt{\hat{g}}\right) + \mathrm{e}^{2\lambda}\hat{g}_{ij}\partial_t\left(\frac{1}{2}\kappa\sqrt{\hat{g}}\hat{g}^{ij}\right)\right). \end{aligned}$$

Next, using our (gauge) choice that κ is constant over Σ, we find

$$ {}^{(3)}\mathcal{S}^{\mathrm{ADM}*} = \int_{\mathcal{M}} \mathrm{d}^3x \left(\overline{p^{ij}}\frac{\partial\hat{g}_{ij}}{\partial q^\alpha}\frac{\mathrm{d}q^\alpha}{\mathrm{d}t} - \partial_t\left(\mathrm{e}^{2\lambda}\kappa\sqrt{\hat{g}}\right) + \frac{\mathrm{d}\kappa}{\mathrm{d}t}\mathrm{e}^{2\lambda}\sqrt{\hat{g}} + \mathrm{e}^{2\lambda}\hat{g}_{ij}\kappa\partial_t\left(\frac{1}{2}\sqrt{\hat{g}}\hat{g}^{ij}\right)\right). \tag{4.22}$$

We claim that the final term of the right-hand side of (4.22) vanishes. To see this, note that the vector $\partial_t\hat{g}_{ij}$ is an element of the tangent space $\mathcal{T}_{\hat{g}}\mathscr{M}_{-1}$, so by Theorem 8.2 of [FT84] it can be decomposed (in an $\mathscr{L}^2$-orthogonal way) as

$$\partial_t\hat{g}_{ij} = \overline{\xi_{ij}} + (\mathscr{L}_V\hat{g})_{ij},$$

for V some vector field (which is unique) and $\overline{\xi^{ij}}$ is transverse and traceless with respect to $\hat{g}$. As argued in [Mon89], from this we find an expression for $\partial_t\left(\frac{1}{2}\sqrt{\hat{g}}\hat{g}^{ij}\right)$ that is traceless with respect to $\hat{g}$, and so indeed the final term of (4.22) vanishes.

Using this, introducing the momentum components p_α from Definition 4.6.3 and integrating a total time-derivative, (4.22) becomes

$$ {}^{(3)}\mathcal{S}^{\mathrm{ADM}*} = \int_{[0,1]} \mathrm{d}t \left(p_\alpha\frac{\mathrm{d}q^\alpha}{\mathrm{d}t} + \frac{\mathrm{d}\kappa}{\mathrm{d}t}\int_\Sigma \mathrm{d}^2x\sqrt{{}^{(2)}g}\right) - \int_\Sigma \mathrm{d}^2x\left(\kappa\sqrt{{}^{(2)}g}\right)\Bigg|_{t=0}^{t=1}. \tag{4.23}$$

We have almost reached our goal. The final term of (4.23) will vanish because the boundary terms will not contribute to the action. Also, we have already assumed our Cauchy surface to be of constant mean extrinsic curvature, but we still have the ability to choose what time-slicing to make so that (4.23) takes an even simpler form. We make the following definition of our desired time-slicing.

Definition 4.6.4 (York time)**.** The time-slicing given by **York time** $\mathtt{T}$ is defined for a hypersurface of constant non-zero mean extrinsic curvature κ by

$$\mathtt{T} := -\kappa,$$

seen as a time function on $\mathcal{M}$. It was first introduced in [YJ72].

By employing York time to fix the time coordinate $t = \mathtt{T}$, our reduced action (4.23) takes the elegant form

$$ {}^{(3)}\mathcal{S}^{\mathrm{ADM}*} = \int \mathrm{d}\mathtt{T} \left(p_\alpha \frac{\mathrm{d}q^\alpha}{\mathrm{d}\mathtt{T}} - \int_\Sigma \mathrm{d}^2x \sqrt{{}^{(2)}g} \right). \tag{4.24} $$

By inspecting this action integral we make the following definition.

Definition 4.6.5 (ADM Hamiltonian)**.** The **ADM Hamiltonian** H^{ADM} is given by

$$ H^{\mathrm{ADM}}(q^\alpha, p_\alpha, \mathtt{T}) := \int_\Sigma \mathrm{d}^2x \sqrt{{}^{(2)}g} = \int_{\Sigma_\mathtt{T}} \mathrm{d}^2x \left(\mathrm{e}^{2\lambda} \sqrt{\hat{g}} \right) (q^\alpha, p_\alpha, \mathtt{T}). $$

This is the area functional for the hypersurface $\Sigma_\mathtt{T}$,[43] where we have included the subscript as a reminder of the dependence on the York time.

This Hamiltonian is independent of the choice of global section of $\mathscr{M}_{-1} \to \mathscr{M}_{-1}/\mathscr{D}_0$, since the volume functional is independent of the choice of section, and so the Hamiltonian is as well. Though implicitly defined, this is a remarkable Hamiltonian as it informs us that the areas of the Cauchy surfaces that foliate spacetime into time-slices are what determine the Hamiltonian dynamics of the system, taking place on the cotangent bundle of the Teichmüller space $\mathcal{T}^*\mathscr{T}(\Sigma)$.

It also gives the desired form of the action as

$$ {}^{(3)}\mathcal{S}^{\mathrm{ADM}*} = \int \mathrm{d}\mathtt{T} \left(p_\alpha \frac{\mathrm{d}q^\alpha}{\mathrm{d}\mathtt{T}} - H^{\mathrm{ADM}} \right), $$

where we recognise the integrand as taking the form $L = \dot{q}^i p_i - H$.

The following section gives a brief overview of the zero and unit genus cases of the same problem addressed in this section: reducing the Einstein equations to a discussion of Teichmüller spaces.

4.7 Genus Case Studies: Zero and Unit Genus

The analysis in the previous section was done in the case $\mathrm{g} > 1$. We made the comment at the beginning of last section that the zero genus case was vacuous, because of the lack of Teichmüller dimensions. This can also be explicitly seen with the new tools we have developed. Consider the Hamiltonian constraint in the form given by Proposition 4.6.2:

$$ \Delta_{\hat{g}}\lambda = \beta_+ \mathrm{e}^{2\lambda} + \beta_- \mathrm{e}^{-2\lambda} + \beta_0. \tag{4.25} $$

In the case $\mathrm{g} = 0$, the β-coefficients take different forms than before. According to [Mon89], β_- will vanish as a direct consequence of having a Teichmüller space of zero dimension (since this is a measure of the transverse-traceless symmetric 2-tensors on a space, which vanishes on $\mathbb{S}^2$). We also discover that $\beta_0 > 0$, and when $\hat{g}$ is taken to be the round metric on $\mathbb{S}^2$ then $\beta_0 = \frac{1}{2}R = 1$ (see Example 2.2.3). With these in mind it can be shown that (4.25) has no solutions.

We now turn to the final case, where $\mathrm{g} = 1$. With our Uniformisation Theorem, we can now assume any given metric on such a space to be conformal to a flat metric $\hat{g}$. The

[43] Note that since the Hamiltonian is dependent on time, energy is not conserved by evolution of the mechanics of the system. This is because the surfaces $\Sigma_\mathtt{T}$, though diffeomorphic to one another, do not have equal area.

same arguments follow as in the last section, and one obtains that the vectors Y^i found in the decomposition of the gravitational momentum π^{ij} (4.18) are Killing vectors with respect to $\hat{g}$. These fields form a 2-dimensional space.

Another 2-dimensional space that one finds is the space of conformal momenta $\overline{p^{ij}}$ defined by (4.19), which, in the unit genus case are arbitrary transverse and traceless (with respect to $\hat{g}$) symmetric tensor densities. Now, $\beta_0 = 0$ since the Ricci scalar of $\hat{g}$ vanishes by flatness, and both β_+ and β_- are constant because of the covariantly-constant nature of $\overline{p^{ij}}$. This reduces our Hamiltonian constraint to

$$\Delta_{\hat{g}}\lambda = \beta_+ \mathrm{e}^{2\lambda} + \beta_- \mathrm{e}^{-2\lambda}. \tag{4.26}$$

If one of β_+ or β_- vanishes, then (4.26) only has a solution when the other is also zero, which implies that λ is a constant. By discussion in the Appendix of [Mon86], a unique solution in the case of β_+ and β_- both non-zero exists via

$$\mathrm{e}^{4\lambda} = -\frac{\beta_-}{\beta_+} < 0.$$

In this case, too, λ is a constant.

Consider the case where $\kappa > 0$. Instead of employing York time, we fix our time coordinate as

$$\kappa = \exp\left(\frac{t}{(2\pi)^2}\right). \tag{4.27}$$

(For simplicity, we will not introduce a new font—such as for the York time $\mathtt{T}$—instead keeping t as the time coordinate.) Along with this choice, we can choose coordinates $(q^\alpha)_{\alpha=1,2}$ and momentum components $(p_\alpha)_{\alpha=1,2}$ to parametrise the 2-dimensional Teichmüller space and its associated cotangent space. If we consider the conformal momenta $\overline{p^{ij}}$ and the metric $\hat{g}$ to be expressed in terms of these canonical variables $(q^\alpha, p_\alpha)_{\alpha=1,2}$, then the reduced action (found in (4.24) for the $\mathrm{g} > 1$ case) takes the form

$${}^{(3)}\mathcal{S}^{\mathrm{ADM}*} = \int \mathrm{d}t \left(p_\alpha \frac{\mathrm{d}q^\alpha}{\mathrm{d}t} - \left(\frac{2\hat{g}_{ik}\hat{g}_{jl}\overline{p^{ij}}\,\overline{p^{kl}}}{\det \hat{g}} \right)^{\frac{1}{2}} \right) + \text{boundary terms}.$$

The fixing of the time coordinate (4.27) ensures that the explicit Hamiltonian above is independent of time. The evolution equations induced by this Hamiltonian can be solved explicitly, as found in [Mar84].[44] In this way, $(2+1)$-dimensional General Relativity can be 'solved.'

Our final section returns to the lapse and the shift, inspecting the equations they must satisfy, before presenting the spacetime split Einstein equations, called the 'Einstein flow.'

4.8 Lapse, Shift and Einstein Flow

Much of the analysis in the previous sections ignored the specifics of the lapse N and the shift N^i. This was partly because of Proposition 4.5.1 allowing us to choose a globally vanishing shift. We had also introduced normal coordinates ($N = 1$) and geodesic coordinates ($N^i = 0$), whose choice greatly simplifies the formulation of the Einstein

[44] See [Car95b] for details, as well.

equations in the ADM formalism. However, as noted in [Mon89], geodesic normal coordinates may encounter difficulties (known as geodesic focusing), and so it is worthwhile to explore the specific equations that the lapse and the shift must satisfy. As with our genus case studies, more effort has been made to explore the case $\mathrm{g} > 1$.

Proposition 4.8.1. *For metric $\hat{g}$ with Ricci scalar $\hat{R} = -1$ on a $(2+1)$-dimensional spacetime with Cauchy surface Σ with constant non-zero mean extrinsic curvature κ and genus $\mathrm{g}(\Sigma) > 1$, the equation satisfied by the lapse N is*

$$e^{2\lambda}\frac{\partial \kappa}{\partial t} = -\Delta_{\hat{g}} N + NP, \tag{4.28}$$

where we have defined

$$P := \frac{e^{-2\lambda}}{\det \hat{g}}\left(\hat{g}_{ik}\hat{g}_{jl}\overline{p^{ij}}\,\overline{p^{kl}} + \frac{1}{2}e^{4\lambda}\kappa^2 \det \hat{g}\right).$$

It has a unique smooth solution N, which is positive on Σ if and only if $\partial_t \kappa > 0$.

Idea of a proof. (See [Mon89] for details.)

The function P is positive, so by linear elliptic partial differential equation theory (see Chapter 6 of [Eva10] for the general theory), there is a unique solution N. By the maximum principle, if $\partial_t \kappa > 0$ then $N > 0$. Thus our choice $t = \mathtt{T}$ gives a unique smooth positive solution N. □

In the case of $\mathrm{g} = 1$, fixing the time as in (4.27) gives a unique solution to the corresponding equation (4.28) for the lapse:

$$N = \frac{1}{(2\pi)^2\kappa}.$$

(See [Mon89] for details.)

Though Proposition 4.5.1 grants us the ability to choose a vanishing shift, we will still briefly discuss the equation it must satisfy under certain gauge conditions.

Definition 4.8.1 (CMCSH)**.** Consider a fixed background Riemannian metric $\tilde{g}$, denote its associated Christoffel symbols $\tilde{\Gamma}^k_{ij}$ and let a vector field V be defined locally as

$$V^k := \hat{g}^{ij}\left(\hat{\Gamma}^k_{ij} - \tilde{\Gamma}^k_{ij}\right).$$

If we require $V^k = 0$, then we have **spatial harmonic coordinates**. Often these are combined with our coordinate choice of York time: the extrinsic mean curvature κ is proportional to the time coordinate (that is, $\kappa = -t$). This is called the **constant mean curvature** coordinate choice, and together these are known as constant mean curvature and spatial harmonic coordinates, or **CMCSH**.

With the above definition in mind, we have the following equation for the shift under the choice of CMCSH coordinates.

Proposition 4.8.2. *In CMCSH coordinates and under the same assumptions as Proposition 4.8.1, the equation satisfied by the shift N^i is*

$$h^{ij}\left(\hat{\Gamma}^k_{ij} - \tilde{\Gamma}^k_{ij}\right) - \frac{1}{2}\left(\hat{\nabla}^i h^k_i + \hat{\nabla}_i h^{ik} - \hat{\nabla}^k h^i_i\right) = 0, \tag{4.29}$$

where indices are raised and lowered by $\hat{g}$ (with associated Christoffel symbols $\hat{\Gamma}^k_{ij}$) and we have defined

$$h_{ij} := -2N\kappa_{ij} + \hat{\nabla}_i N_j + \hat{\nabla}_j N_i.$$

Under (4.29) *and the equation satisfied by the lapse* (4.28)*, the CMCSH 'gauge fixing' is preserved throughout the dynamics of the system.*

Proof. See [Mon07] for details. □

Note that (4.29) does admit a choice of zero shift, as we hoped. Also, one reason for imposing the CMCSH conditions is that now the Einstein equations with these constraints form what is known as an elliptic-hyperbolic system, as the Einstein equations can be written in hyperbolic form (see the discussion at the end Section 4.1, or in completeness throughout [Rin09]), and the equations for the shift and the lapse are elliptic.

Using the decomposition from the previous sections, we are able to write the Einstein equations as coupled equations dictating the time-evolution of the metric and the extrinsic curvature of the Cauchy surface.

Definition 4.8.2 (Einstein flow)**.** For a $(2+1)$-dimensional spacetime $(\mathcal{M}, g)$ with a Cauchy surface Σ (whose induced metric, Levi-Civita connection and Ricci tensor are g_{ij}, ∇_i and ${}^{(2)}R_{ij}$) with extrinsic curvature κ_{ij} (with trace κ), lapse N and shift N^i, the Einstein equations (with zero cosmological constant here for simplicity) can be written in a form called the **Einstein flow**:

$$\begin{aligned}\partial_t g_{ij} &= -2N\kappa_{ij} + \nabla_i N_j + \nabla_j N_i;\\ \partial_t \kappa_{ij} &= N\left({}^{(2)}R_{ij} - 2\kappa_{ik}\kappa^k_j + \kappa\kappa_{ij}\right) - \nabla_i\nabla_j N + N^k\partial_k\kappa_{ij} + \kappa_{ik}\partial_j N^k + \kappa_{jk}\partial_i N^k.\end{aligned}$$

If we place ourselves in geodesic normal coordinates, where the lapse is unity and the shift vanishes, these take the simpler form:

$$\partial_t g_{ij} = -2\kappa_{ij} \quad \text{and} \quad \partial_t \kappa_{ij} = {}^{(2)}R_{ij} - 2\kappa_{ik}\kappa^k_j + \kappa\kappa_{ij},$$

where the spatial derivatives have vanished, leaving us with coupled ordinary differential equations.[45]

The Einstein flow is combined with the constraints that arise from the Gauss-Codazzi equations (4.2), namely:

$$\nabla^j \kappa_{ij} - \nabla_i \kappa = 0 \quad \text{and} \quad {}^{(2)}R + \kappa^2 - \kappa_{ij}\kappa^{ij} = 0,$$

which are simply the Hamiltonian and momentum constraints (4.14).

[45] In [Mon89], Moncrief writes that

> this choice [of geodesic normal coordinates] eliminates all spatial derivatives in the evolution equations and thus effectively reduces them to decoupled systems of ordinary differential equations along each normal geodesic. Certain particular solutions (including those obtained by taking quotients of Minkowski space by suitably chosen discrete subgroups of the Lorentz group) can in fact be globally foliated by geodesic normal slices. In general, however, one expects Gaussian normal coordinate systems to develop singularities unrelated to the natural boundaries of the space-times under study. In any case the solution of Einstein's equations through the use of Gaussian coordinates does not really realize Witten's objective to reduce these equations to a finite-dimensional Hamiltonian system defined on the cotangent bundle of Teichmüller space.

This will not be explored further, though we expect that a better understanding of the possible alternative foliation of spacetime by null hypersurfaces, as explored in Chapters 3 and 4 of [Are13], among other sources.

As we had hoped, we have the following.

Proposition 4.8.3. *The Einstein flow equations with constraints form a system of partial differential equations giving the time-evolution of the metric and the extrinsic curvature of an embedded Cauchy surface of a spacetime manifold.*

Proof. As mentioned above, the constraint equations arise from the Gauss-Codazzi equations: taking two traces and using a normal coordinate system to simplify calculations—since a tensor identity is valid in any coordinates—gives the desired result. The evolution equations for the metric and the extrinsic curvature come from the Hamilton-Jacobi equations for the ADM action; equivalently, they arise from a long computation first done in the original ADM paper [ADM59] involving manipulations of the Einstein equations themselves. Note that the evolution of the metric is seen in (4.9), but the evolution of the extrinsic curvature requires significant effort to find. □

Until now, we have not explicitly mentioned the initial value (Cauchy) problem in General Relativity. Entire textbooks—for example, see [Rin09] or [CB08]—have been written on the subject of viewing the Einstein equations as partial differential equations, and as such we will keep our discussion brief. Though they appeared in 1915, it was not until 1952 that Choquet-Bruhat (in her paper [CB52], using the covariant formalism; in [CBR83], she proved this for the time-space split formalism) proved that under certain gauge choices, the Einstein equations can be written as a quasi-linear hyperbolic system of equations, and form a well-posed initial value problem given data on a Cauchy surface (as it was later formulated). The quickest way to argue that this is the case is to note that they can be written as

$$g^{\mu\nu}(u)\partial_\mu\partial_\nu u = F(u, \partial u),$$

for $u = g_{\mu\nu}$ the solution metric and $F(u, \partial u)$ the function containing all terms that are not of highest (second) order. The left-hand side appears for similar reasons as when it does in the linearisation of Ricci flow—see (3.17)—and since the coefficients $g^{\mu\nu}(u)$ depend on u, the equations are quasi-linear. To see hyperbolicity, return to Definition 2.3.1 and note that since the coefficients $g^{\mu\nu}$ are (the inverse of) a Lorentzian metric, we will indeed get one eigenvalue of opposite sign.[46]

The hyperbolic nature of the Einstein equations allows us to draw the comparison with the canonical hyperbolic equation: the wave equation (2.9) (along the same vein as in Ricci flow when we made qualitative comparisons to the heat equation (2.10)). From this, the postulate of relativity that *information cannot propagate faster than the speed of light* appears in the guise of gravitational waves: the waves arising from the hyperbolic nature of the Einstein equations propagate at the speed of light!

[46] If the metric was Riemannian then the operator $g^{\mu\nu}(u)\partial_\mu\partial_\nu$ would be elliptic—a Laplacian!

Chapter 5

Conclusions

At the onset of this project, a broad goal of combining Ricci flow and General Relativity was established. Several possible paths were explored, and eventually this text was born. The primary objectives were to inspect the Einstein equations and Ricci flow in the same dimension, and to hope that one could learn from the other via comparisons. In the end, the text used the Uniformisation Theorem to link the 2(-spatial)-dimensional versions of the two subjects, first using the normalised Ricci flow to give a proof of the Uniformisation Theorem, before applying it to General Relativity. In the latter part, it split the discussion into three parts (divided by their conformal class) and allowed one to work in the simpler case of a metric with constant curvature, which was related to an initial given metric by a conformal transformation.

Though this is the end result, several avenues for future exploration of intersections between Ricci flow and General Relativity lie in wait. We have presented these as examples, and have kept discussion informal, presenting resources for interested readers instead of bogging down the concluding chapter with technical details.

Example 5.1 (As forming singularities)**.** Though our discussions in Chapter 3 had the conclusion that in 2 dimensions the normalised Ricci flow does not encounter singularities, the general case is vastly different. Much of the early work on Ricci flow (see [Ham82], for a start) dealt with the case of positive curvature, wherein (for $n \geq 3$) singularities form as the curvature blows up. Major steps in confronting these difficulties were made by Perelman in his proof of the Poincaré Conjecture (Theorem 3.1) in the papers [Per02], [Per03a] and [Per03b], where he showed that a method called **surgery** can be employed indefinitely to deal with singularities.

On the other hand, stemming from the work of Penrose in [Pen65] and Hawking in [Haw66], singularities in General Relativity began to be explored rigorously.[1] In $(2+1)$-dimensional General Relativity, no solutions to the Einstein equations that have curvature singularities have been found (though in higher dimensions, many such metrics exist: see any reference on General Relativity for details), though an interesting example of a $(2+1)$-dimensional black hole can be found in Example 4.1.3, and in [MR93] it is shown that an initial metric (describing 'dust') in $(2+1)$ dimensions may collapse to form the BTZ metric (4.1).

Nevertheless, the Einstein flow as presented in Section 4.8 (and its higher-dimensional counterparts) may encounter singular points, as discussed in the footnote following Definition 4.8.2. As noted in the footnote, a starting point could be to explore null foliations of

[1] Loosely, Penrose studied black hole-type singularities, while Hawking's results dealt with Big Bang-type singularities.

spacetime (an introduction to this can be found in Chapters 3 and 4 of [Are13]). Perhaps the study of singularities in General Relativity could learn from techniques developed for Ricci flow.

Example 5.2 (Space of metrics)**.** Both Ricci flow and the Einstein equations are partial differential equations whose solutions are semi-Riemannian metrics. The abstract tools of Hilbert manifolds developed in Section 2.4 allow one to see the solutions of these equations as subsets of a space of metrics $\mathscr{M}$ (possibly with a Sobolev parameter s to keep track of differentiability). We have described the solution to Ricci flow as a curve $g(t) : [0, \infty) \to \mathscr{M}$, and the solution to the Einstein equations as a subspace (quotiented by diffeomorphisms).

It could be interesting to inspect the space $\mathscr{M}$ and its properties in more detail: with an ILH structure it becomes a smooth (Hilbert) manifold, to which we can associate a (Hilbert-)Riemannian metric (see Definition 2.4.1). Though it is infinite-dimensional, what information can we glean about its topology? Or, depending on what metric we grant it, what do geodesics look like? Thus far, we have only introduced an intrinsic metric induced by the inner product associated to the underlying Hilbert space, as well as our $\mathscr{L}^2$-metric described in Example 2.4.2. Could a metric exist whose geodesics are exactly the curves $g(t)$ that are Ricci flow solutions?

On a similar vein, Einstein metrics play a role in both (normalised) Ricci flow and the Einstein equations: they are fixed points of the former (see Example 3.1.3) and solutions of the latter (see Example 4.1.1). In fact, Andersson and Moncrief explored in [AM11] to what extent Einstein metrics are 'attractors' for the Einstein flow.[2] In [IMS11], Isenberg et al. state that the size of the neighbourhood in $\mathscr{M}$ of a constant curvature metric $\hat{g}$ on which Ricci flow has $\hat{g}$ as an attractor is of interest. As promising as those two statements may seem, it is entirely possible that no information can be connected via this link, as many partial differential equations may have Einstein metrics as solutions or fixed points because of the simple definition of Einstein metrics.

The following idea arose from a conversation with Dr. Pau Figueras, and has a more physical perspective than the above examples.

Example 5.3 (Black holes in Anti-de Sitter space)**.** In 1983, Hawking and Page (see [HP83]) studied black holes at thermal equilibrium in Anti-de Sitter space (with cosmological constant $\Lambda < 0$).[3] They discovered that for temperatures above a critical temperature $T_0 = \frac{1}{2\pi}\sqrt{-\Lambda}$, there exist two possible masses for black holes at equilibrium, amounting to two real solutions of a quadratic equation for their radii. These 'large' and 'small' black holes are thermodynamically stable and unstable, respectively.

Note that mass may be defined in the ADM formalism—called the **ADM mass**—by considering asymptotically flat (globally hyperbolic) spacetimes and defining the energy E to be proportional to the integral over increasingly-large spheres $\mathbb{S}_\varrho$ (whose radii ϱ tend to infinity) of the normal projection of a difference of derivatives of the metric g on the Cauchy surface:

$$E := \frac{1}{16\pi} \lim_{\varrho \to \infty} \int_{\mathbb{S}_\varrho} \left(\partial^\mu g_{\mu\nu} - \partial_\nu g^\mu_\mu\right) \eta^\nu.$$

[2] As they note in the paper, the end result is parallel to the results presented in [AMT97] and [Mon07], a simpler version of which roughly forms the contents of our Chapter 4.

[3] This can be thought of as black holes confined to a finite-sized box, as Anti-de Sitter has natural boundary conditions similar to those imposed by a box at some finite radius. For details, see Chapter 5.2 of [HE73].

(See Chapter 11.2 of [Wal10] for the precise definition.) The important point to note is that the 'large' and 'small' black holes described above have different, well-defined masses.

In 2006, Headrick and Wiseman (see [HW06]) considered the following form of Ricci Flow

$$\partial_t g_{\mu\nu} = -2R_{\mu\nu} + 2\Lambda g_{\mu\nu}. \tag{5.1}$$

They suggested that this could be used to flow a metric describing the unstable 'small' black hole to either the 'large' black hole or to 'hot empty space.' In the first instance, the black hole simply grows in size. However, this is peculiar, since Woolgar (see [Woo08]) studied the conservation of mass along Ricci flow—though via a different formulation than (5.1)—and showed that (under some assumptions) it is conserved. The second instance—going from the 'small' black hole to 'hot empty space'—is also of interest, since going from the presence of a black hole to empty space involves a change in topology—therefore a singularity must arise in the flow.

Appendix A

Supplement to Section 2.1

The following appendix includes basic constructions of manifolds (Section A.1), vector bundles (Section A.2) and connections (Section A.3), which are abstractions on the definitions found in the main text, as well as a handful of additional definitions (Section A.4). It has been written to be self-contained, though most of its information can be read in partnership with Section 2.1 of the main text. These definitions can be found in every differential geometry textbook: for example, see [Lee13].

A.1 Manifolds

The study of physics often is described by partial differential equations. The first generalisation of the real n-dimensional space $\mathbb{R}^n$ on which these equations live is the concept of a space that locally looks like $\mathbb{R}^n$: a manifold.

Definition A.1.1 (Topological manifold)**.** An n-dimensional **manifold**, or n-manifold, denoted $\mathcal{M}$, is a paracompact Hausdorff[1] space such that every point has a neighbourhood that is homeomorphic to an open set in $\mathbb{R}^n$.

The above defines a topological n-manifold, of which we give two brief examples.

Example A.1.1 (Euclidean space)**.** Trivially, n-dimensional Euclidean space $\mathbb{R}^n$ is an n-manifold.

Example A.1.2 (Product manifold)**.** For $\mathcal{M}$ an m-manifold an $\mathcal{N}$ an n-manifold, the **product manifold** $\mathcal{M} \times \mathcal{N}$ is an $(m+n)$-manifold, where the topology is the product topology.

We are more interested in manifolds with the additional structure of differentiability, or smoothness. As such, we consider the maps between the open sets of the manifold and $\mathbb{R}^n$ more closely so as to define coordinate patches and transition functions between the patches.

Definition A.1.2 (Differentiable manifold)**.** Let $\mathcal{M}$ be an n-manifold. Then,

1. For $\mathcal{U} \subset \mathcal{M}$ an open set (called a **coordinate patch**) and $\varphi : \mathcal{U} \to \mathbb{R}^n$ an open and injective map (called a **coordinate system**, or chart), we write the **coordinates** $(x^i)_{i=1}^n$ of a point x in $\mathcal{U}$ for $x^i = \varphi^i(x)$, where $\varphi^i : \mathcal{U} \to \mathbb{R}$ are the slices of φ along the standard Euclidean basis of $\mathbb{R}^n$;

[1] A topological space is **Hausdorff** if any two points can be separated by non-intersecting open sets and **paracompact** if every covering has a refinement that is locally finite (that is, every point has a neighbourhood that only intersects a finite number of elements of the refinement). Note that a Hausdorff space is paracompact if and only if every cover admits a partition of unity subordinate to it.

2. The **transition functions** between coordinate systems $\varphi_\alpha : \mathcal{U}_\alpha \to \mathbb{R}^n$ and $\varphi_\beta : \mathcal{U}_\beta \to \mathbb{R}^n$ are the maps $\varphi_\beta \circ \varphi_\alpha^{-1}$ and $\varphi_\alpha \circ \varphi_\beta^{-1}$, which, when $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ is non-empty, are defined from the image of this intersection in $\mathbb{R}^n$ to $\mathbb{R}^n$—and thus we have a well-established notion of differentiability of such functions;

3. An **atlas** on $\mathcal{M}$ is an open covering $\mathcal{M} = \cup_\alpha \mathcal{U}_\alpha$ equipped with coordinate systems $\varphi_\alpha : \mathcal{U}_\alpha \to \mathbb{R}^n$ and is called $\mathcal{C}^k$ (where k is a positive integer or infinity) if all well-defined transition functions between its coordinate systems are $\mathcal{C}^k$. A $\mathcal{C}^k$**-structure** on $\mathcal{M}$ is an equivalence class of $\mathcal{C}^k$-atlases where two are equivalent if their union is once again a $\mathcal{C}^k$-atlas (we will always assume our atlas is maximal to this property: any atlas compatible with it will be contained in it), and a **smooth manifold** is a manifold with a $\mathcal{C}^\infty$-structure (a **smooth structure**);

4. If all transition functions in an atlas preserve orientation (that is, have positive Jacobian), then $\mathcal{M}$ is called **orientable**, and an **orientation** on $\mathcal{M}$ is an atlas satisfying this condition;

5. A manifold is called **closed** if it is compact and boundaryless, a property most of the manifolds in this text possess.

This sharpens the definition of a topological manifold to that of a differentiable manifold; in this text, all manifolds are equipped with an atlas and will be smooth and orientable unless explicitly stated otherwise. We now use this smooth structure to define smooth functions on the manifold, smooth maps between manifolds, and the isomorphisms of smooth manifolds: diffeomorphisms.

Definition A.1.3 (Smoothness)**.** Let $\mathcal{M}$ be a smooth manifold. Then,

1. A function $f : \mathcal{M} \to \mathbb{R}$ is called **smooth**, or of class $\mathcal{C}^\infty(\mathcal{M})$, if for every coordinate system φ_α in the atlas, $f \circ \varphi_\alpha^{-1}$ is smooth in the usual sense;

2. For manifolds $\mathcal{M}$ and $\mathcal{N}$, a function $\psi : \mathcal{M} \to \mathcal{N}$ is called **smooth** if for each $f \in \mathcal{C}^\infty(\mathcal{N})$, the function $f \circ \psi : \mathcal{M} \to \mathbb{R}$ is smooth. If ψ is bijective with smooth inverse then ψ is a **diffeomorphism**. If $\mathcal{N} = \mathcal{M}$, the group (with composition given by function composition) formed by the set of all such (orientation-preserving) diffeomorphisms is called the **diffeomorphism group** of $\mathcal{M}$ and is written $\mathscr{D}(\mathcal{M})$.

One example of a particularly nice manifold is one that is also a group.

Example A.1.3 (Lie group)**.** A group $\mathcal{G}$ is a **Lie group** if it is a smooth manifold and its inversion and composition maps are smooth. Symbolically, we require the smoothness of the mapping $\Upsilon : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ given by

$$\Upsilon(\chi, \upsilon) = \chi^{-1}\upsilon.$$

In the dimensions of interest to this text, we have the following result.

Proposition A.1.1. *Any homeomorphic smooth n-manifolds are diffeomorphic if $n \leq 3$.*

The above definitions allow us to understand and eventually differentiate functions that live on manifolds. An example of such a function is the temperature function on the Earth. Unfortunately, on a general manifold, it is not possible to 'subtract' nearby points to find a vector that points from one to the other, like it is in $\mathbb{R}^n$. To generalise this notion, we define the concept of a tangent vector defined at a point on the manifold[2] as follows.

[2]Note that we now transition from using x as our typical point on the manifold to p, because x will be associated to the coordinates $(x^i)_{i=1}^n$ associated to some patch around $p \in \mathcal{M}$.

Definition A.1.4 (Tangent vector)**.** At a point $p \in \mathcal{M}$, we define a **tangent vector**, or derivation, to be an $\mathbb{R}$-linear function $V_p : \mathcal{C}^\infty(\mathcal{M}) \to \mathbb{R}$ that satisfies a Leibniz-type rule:

$$V_p(fg) = V_p(f)g(p) + f(p)V_p(g) \quad \text{for} \quad f, g \in \mathcal{C}^\infty(\mathcal{M}).$$

These are named as such because the set of all tangent vectors at $p \in \mathcal{M}$ forms a vector space $\mathcal{T}_p\mathcal{M}$ known as the **tangent space**[3] at $p \in \mathcal{M}$. If we choose a tangent vector at every point on the manifold such that for all $f \in \mathcal{C}^\infty(\mathcal{M})$ the function $V_p(f)$ is in $\mathcal{C}^\infty(\mathcal{M})$ (varying points $p \in \mathcal{M}$), then we obtain the **vector field** $V : \mathcal{M} \times \mathcal{C}^\infty(\mathcal{M}) \to \mathbb{R}$.

We note that for coordinates $(x^i)_{i=1}^n$, the partial derivatives along these coordinate directions

$$\partial_i := \frac{\partial}{\partial x^i},$$

form a basis for the tangent space. These local coordinates are often used in physics and will simplify our notation once we begin to work with tensors.

The previous definition of a tangent vector does not seem to relate nearby points until the link is drawn by the following equivalent definition of a tangent vector, this time using curves[4] along the manifold.

Definition A.1.5 (Tangent vector #2)**.** Consider a curve $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$ such that $\gamma(0) = p$. Then the velocity of this curve at $p \in \mathcal{M}$ is a **tangent vector** V_p defined on a function $f \in \mathcal{C}^\infty(\mathcal{M})$ as[5]

$$V_p(f) = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(\lambda))\bigg|_{\lambda=0}.$$

This definition is useful because much of this text will involve the inspection of curves on manifolds. We will be interested in the lengths of these curves, but before we enter the realm of metric geometry (where this is possible), we take a detour into the world of vector bundles, which nicely generalise the above notions.

A.2 Vector Bundles

Consider affixing to every point of a manifold a space. This is a general construction known as a fibre bundle, the simplest of which is where the space affixed is a vector space. To make this idea more precise we make the following definition.

Definition A.2.1 (Vector bundle)**.** A k-dimensional **vector bundle** over a manifold $\mathcal{M}$ (called the **base space**) is an open surjective map $\pi : \mathcal{E} \to \mathcal{M}$ (where $\mathcal{E}$ is a manifold known as the **total space**) such that the following conditions hold:

[3] To build on the example of a Lie group from Example A.1.3: a **Lie algebra** is a vector space $\mathfrak{g}$ equipped with an antisymmetric bilinear map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ (called the **Lie bracket**) satisfying the **Jacobi identity**:

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0, \quad \text{for all} \quad u, v, w \in \mathfrak{g}.$$

Note that for a Lie group $\mathcal{G}$, the tangent space at the identity $e \in \mathcal{G}$ forms the Lie algebra associated to the Lie group $\mathcal{G}$ when the commutator of vector fields is taken as the Lie bracket, which can be checked to satisfy the Jacobi identity.

[4] Here we use the word curve as meaning a function from a real interval to a manifold, ignoring completely the troubles of other areas of mathematics with this word.

[5] Note that we will use λ instead of the traditional t as the parameter of the curve because of the important role of time in this text.

1. For every point $p \in \mathcal{M}$, the **fibre** over p, $\mathcal{E}_p := \pi^{-1}(p)$, is a k-dimensional vector space;

2. Every point $p \in \mathcal{M}$ has a neighbourhood $\mathcal{U} \subset \mathcal{M}$ with corresponding diffeomorphism $\varphi_{\mathcal{U}} : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times \mathbb{R}^k$ called a **local trivialisation** such that the following diagram commutes:

$$\begin{array}{ccc} \pi^{-1}(\mathcal{U}) & \xrightarrow{\varphi_{\mathcal{U}}} & \mathcal{U} \times \mathbb{R}^k \\ & \overset{\pi}{\searrow} & \downarrow \mathrm{pr}_{\mathcal{U}} \\ & & \mathcal{U} \end{array}$$

   where $\mathrm{pr}_{\mathcal{U}}$ is the projection: $\mathrm{pr}_{\mathcal{U}}(q, v) = q$ for $q \in \mathcal{U}$ and $v \in \mathbb{R}^k$;

3. The local trivialisations are **fibre-wise linear**: for all $q \in \mathcal{U}$, $\varphi_{\mathcal{U}}|_{\mathcal{E}_q} : \mathcal{E}_q \to \{q\} \times \mathbb{R}^k$ is $\mathbb{R}$-linear.

We call a right-inverse of π (that is, $s : \mathcal{M} \to \mathcal{E}$ such that $\pi \circ s = \mathbb{1}$) a **section**, and denote by $\Gamma(\mathcal{E})$ the set of all sections of the vector bundle $\pi : \mathcal{E} \to \mathcal{M}$. We assume our sections to be smooth, and equip $\Gamma(\mathcal{E})$ with a vector space structure by the following relation for sections $s, \varsigma \in \Gamma(\mathcal{E})$:

$$(\alpha s + \beta \varsigma)(p) = \alpha s(p) + \beta \varsigma(p),$$

where α and β are real numbers and $p \in \mathcal{M}$. If α and β are taken to be in $\mathcal{C}^\infty(\mathcal{M})$ then the above expression (with α and β evaluated at p on the right-hand side) gives $\Gamma(\mathcal{E})$ a $\mathcal{C}^\infty(\mathcal{M})$-module structure.

We define a **vector bundle morphism** to be a map $\phi : \mathcal{E} \to \mathcal{F}$ between vector bundles $\pi : \mathcal{E} \to \mathcal{M}$ and $\varrho : \mathcal{F} \to \mathcal{N}$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathcal{E} & \xrightarrow{\phi} & \mathcal{F} \\ & \overset{\pi}{\searrow} & \downarrow \varrho \\ & & \mathcal{M} \end{array}$$

If the map is also a diffeomorphism than it is an **isomorphism** of vector bundles.

Before inspecting several key examples of the above abstract definitions, we introduce two fundamental notions: the pull-back and the push-forward, which will be recurring tools in our study of differential geometry.

Definition A.2.2 (Pull-back and push-forward)**.** Let $\mathcal{M}$ and $\mathcal{N}$ be manifolds. Then,

1. For $\psi : \mathcal{M} \to \mathcal{N}$ a smooth map and $f \in \mathcal{C}^\infty(\mathcal{N})$, the **pull-back** of f by ψ is written $\psi^* f$, is given by

$$\psi^* f := \psi \circ f,$$

   and is in $\mathcal{C}^\infty(\mathcal{M})$;

2. For $\psi : \mathcal{M} \to \mathcal{N}$ a smooth map between manifolds, the **push-forward** (or differential) of ψ at a point $p \in \mathcal{M}$ is a linear map $(\psi_*)_p : \mathcal{T}_p\mathcal{M} \to \mathcal{T}_{\psi(p)}\mathcal{N}$ given by

$$(\psi_*)_p(V_p)(f) := V_p(f \circ \psi),$$

   for $V_p \in \mathcal{T}_p\mathcal{M}$ and $f \in \mathcal{C}^\infty(\mathcal{M})$.

We now have the following useful examples of vector bundles.

Example A.2.1 (Trivial bundle)**.** The manifold $\mathcal{M} \times \mathbb{R}^k$ is a k-dimensional vector bundle over $\mathcal{M}$, known as the **trivial bundle** by considering the trivial projection of the space $\mathcal{M}$. A general vector bundle is **trivial** if it is isomorphic to the trivial bundle, and an elementary result is that a bundle is trivial if and only if it has a **global section**—that is, a section defined on the whole base manifold.

A sub-example of this would be to consider the cylinder, which is a one-dimensional vector bundle (also known as a **line bundle**) over the circle $\mathbb{S}^1$.

Example A.2.2 (Dual bundle)**.** Given a vector bundle $\pi : \mathcal{E} \to \mathcal{M}$ we can form the **dual bundle**, whose base space is $\mathcal{E}^*$, has dimension equal to the dimension of $\mathcal{E}$ and whose fibres are defined as the dual vector spaces of the original bundle's fibres:

$$(\mathcal{E}^*)_p := (\mathcal{E}_p)^*.$$

The rest of the conditions of being a vector bundle are met by dualising objects of $\pi : \mathcal{E} \to \mathcal{M}$ (such as sections) in the expected way.

Example A.2.3 (Pull-back bundle)**.** For a vector bundle $\pi : \mathcal{E} \to \mathcal{N}$ and a map of manifolds $\psi : \mathcal{M} \to \mathcal{N}$, the **pull-back bundle** $\psi^*\pi : \psi^*\mathcal{E} \to \mathcal{M}$ is given fibre-wise by

$$(\psi^*\mathcal{E})_p := \mathcal{E}_{\psi(p)}.$$

The sections of this bundle are pull-backs of sections of the original bundle: if $s \in \Gamma(\mathcal{E})$ then $\psi^* s := \psi \circ s$ is in $\Gamma(\psi^*\mathcal{E})$.

Example A.2.4 (Tensor product bundle)**.** Given two vector bundles $\pi : \mathcal{E} \to \mathcal{M}$ and $\varrho : \mathcal{F} \to \mathcal{M}$ over the same base space, their **tensor product bundle** is the unique bundle $\pi \otimes \varrho : \mathcal{E} \otimes \mathcal{F} \to \mathcal{M}$ such that the fibre over any point $p \in \mathcal{M}$ is the tensor product of the fibres over that point in the original vector bundles in the usual tensor product of vector spaces sense:

$$(\mathcal{E} \otimes \mathcal{F})_p := \mathcal{E}_p \otimes \mathcal{F}_p.$$

Sections of this bundle take the form $s_{\mathcal{E}} \otimes s_{\mathcal{F}}$ for $s_{\mathcal{E}} \in \Gamma(\mathcal{E})$ and $s_{\mathcal{F}} \in \Gamma(\mathcal{F})$. While this example may seem abstract now, it will allow us to define tensors, which will be the central characters of semi-Riemannian geometry.

Example A.2.5 (Tangent bundle)**.** Over an n-manifold $\mathcal{M}$, by considering the union of all tangent spaces $\mathcal{T}_p\mathcal{M}$ (each n-dimensional) we obtain the **tangent bundle**, whose total space is written $\mathcal{TM}$ and has $2n$ dimensions. Sections of this bundle are **vector fields**, whose space is sometimes denoted $\mathfrak{X}(\mathcal{M})$, though we will stick to $\Gamma(\mathcal{TM})$. Using local coordinates we can see that the vector $\partial_i \in \mathcal{T}_p\mathcal{M}$ sends a function $f \in \mathcal{C}^\infty(\mathcal{M})$ to $\partial_i f(p)$.

Furthermore, the push-forward of a map $\psi : \mathcal{M} \to \mathcal{N}$ gives a vector bundle morphism from the tangent bundle of $\mathcal{M}$ to that of $\mathcal{N}$ as the following diagram commutes:

$$\begin{array}{ccc} \mathcal{TM} & \xrightarrow{\psi_*} & \mathcal{TN} \\ \downarrow & & \downarrow \\ \mathcal{M} & \xrightarrow{\psi} & \mathcal{N} \end{array}$$

When the map $\psi : \mathcal{M} \to \mathcal{N}$ is a diffeomorphism, we note that vector fields $V \in \Gamma(\mathcal{TM})$ can be pushed-forward to vector fields $\psi_* V \in \Gamma(\mathcal{TN})$, given at a point $q \in \mathcal{N}$ by

$$(\psi_*(V))_q := \psi_*\left(V_{\psi^{-1}(q)}\right).$$

Example A.2.6 (Cotangent bundle)**.** By considering the dual space of each tangent space, we obtain fibres

$$\mathcal{T}_p^*\mathcal{M} := (\mathcal{T}_p\mathcal{M})^*,$$

of what is known as the **cotangent bundle** over the n-manifold $\mathcal{M}$ (whose total space is written $\mathcal{T}^*\mathcal{M}$, which is also $2n$-dimensional). Sections are known as **covectors**, or 1-forms, and the space of 1-forms is often denoted $\Omega^1(\mathcal{M})$. This nomenclature is clear when considering the wedge product of tangent spaces: we can form a bundle $\wedge^k\mathcal{T}^*\mathcal{M}$ by taking the k^{th} wedge product power of the vector spaces $\mathcal{T}_p^*\mathcal{M}$ (for points $p \in \mathcal{M}$). The sections of this bundle are called **k-forms**,[6] and their natural vector space is written $\Omega^k(\mathcal{M})$.[7]

When we consider the local coordinates $(x^i)_{i=1}^n$ around a point $p \in \mathcal{M}$ and corresponding basis of $\mathcal{T}_p\mathcal{M}$ given by $(\partial_i)_{i=1}^n$ we define the dual basis $(\mathrm{d}x^i)_{i=1}^n$ of $\mathcal{T}_p^*\mathcal{M}$ as the set of n covectors satisfying

$$\mathrm{d}x^i(\partial_j) = \delta^i_j, \quad \text{where} \quad i,j \in \{1,\dots,n\}.$$

This basis allows us to define the covector $\mathrm{d}f$ at a point $p \in \mathcal{M}$, or the differential of $f \in \mathcal{C}^\infty(\mathcal{M})$ at p, as

$$\mathrm{d}f|_p := \left.\frac{\partial f}{\partial x^i}\right|_p \mathrm{d}x^i,$$

where we have finally employed the **Einstein summation convention**: *repeated indices are summed over*, which we will use throughout this text. This expression can also be seen without reference to local coordinates as

$$\mathrm{d}f(V) = V(f), \tag{A.1}$$

[6]Note that an n-manifold is orientable if and only if there exists a nowhere-vanishing n-form. This will become crucial in defining integration over manifolds.

[7]To complete the picture, we define the **exterior derivative** to be the map $\mathrm{d} : \mathcal{C}^\infty(\mathcal{M}) \to \Omega^1(\mathcal{M})$ (where smooth functions are thought of as 0-forms) that sends a function $f \in \mathcal{C}^\infty(\mathcal{M})$ to the form $\mathrm{d}f \in \Omega^1(\mathcal{M})$ described in (A.1). We then extend this map to one from k-forms to $(k+1)$-forms as being the unique map of this kind such that

1. It squares to zero: $\mathrm{d}^2 = 0$;
2. It satisfies a Leibniz-type rule:

$$\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^k \alpha \wedge \mathrm{d}\beta,$$

where α is a k-form, and we have used that when α and β are forms, so is $\alpha \wedge \beta$ (whose rank is the sum of the ranks α and β).

This allows us to write **Stokes' Theorem** for a manifold $\mathcal{M}$ with boundary $\partial\mathcal{M}$ in a neat form:

$$\int_{\partial\mathcal{M}} \omega = \int_{\mathcal{M}} \mathrm{d}\omega,$$

for any k-form ω on $\mathcal{M}$. The exterior derivative also allows us to form the **de Rham cochain complex**:

$$0 \to \mathcal{C}^\infty(\mathcal{M}) \xrightarrow{\mathrm{d}} \Omega^1(\mathcal{M}) \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} \Omega^n(\mathcal{M}) \to 0,$$

where $\mathcal{M}$ is an n-manifold. Thus we can define the **de Rham cohomology** as

$$\mathcal{H}^k_{\mathrm{dR}}(\mathcal{M}) := \frac{\ker\left(\mathrm{d} : \Omega^k(\mathcal{M}) \to \Omega^{k+1}(\mathcal{M})\right)}{\operatorname{im}\left(\mathrm{d} : \Omega^{k-1}(\mathcal{M}) \to \Omega^k(\mathcal{M})\right)}.$$

for some tangent vector $V \in \mathcal{T}_p\mathcal{M}$ where $\mathrm{d}f$ is the unique map satisfying the above.

An extension of the ideas above leads to the definition of tensors, as found in Definition 2.1.2. We now turn to the final primary subject of this appendix: connections on vector bundles.

A.3 Connections

We introduce the concept of a connection on a vector bundle, which allows us to differentiate sections (not only tensors) of any vector bundle. Crucially, we note that connections are not unique (though at least one always exists) and therefore there are many ways of taking said derivatives.

Definition A.3.1 (Connection)**.** For a vector bundle $\pi : \mathcal{E} \to \mathcal{M}$ a **connection** on $\mathcal{E}$ is a linear operator $\nabla : \Gamma(\mathcal{E}) \to \Gamma(\mathcal{T}^*\mathcal{M} \otimes \mathcal{E})$ satisfying the following Leibniz-type rule:

$$\nabla(fs) = \mathrm{d}f \otimes s + f\nabla(s),$$

for $f \in \mathcal{C}^\infty(\mathcal{M})$ and $s \in \Gamma(\mathcal{E})$. Note that the first term on the right-hand side prevents ∇ from being $\mathcal{C}^\infty(\mathcal{M})$-linear. A section $s \in \Gamma(\mathcal{E})$ is called **parallel** if $\nabla(s) = 0$.

We now give several examples of connections. Given our focus will be on connections on tangent bundles, our examples have been geared as such.

Example A.3.1 (Trivial connection)**.** For the trivial bundle $\mathrm{pr}_{\mathcal{M}} : \mathcal{M} \times \mathbb{R}^n \to \mathcal{M}$, we declare that $\nabla(s_i) = 0$ for each section $s_i(x) = (x, e_i)$, where $(e_i)_{i=1}^n$ is the canonical basis of $\mathbb{R}^n$. Since all other sections are linear combinations of $(s_i)_{i=1}^n$, we find that a general section can be written $s = f^i s_i$ for $f^i \in \mathcal{C}^\infty(\mathcal{M})$. By the Leibniz-type rule we have

$$\nabla(f^i s_i) = \mathrm{d}f^i \otimes s_i,$$

and so we note that the connection is simply the derivative of the coefficients of $s \in \Gamma(\mathcal{M} \times \mathbb{R}^n)$.

Example A.3.2 (Tensor connection)**.** For vector bundles $\pi : \mathcal{E} \to \mathcal{M}$ and $\varrho : \mathcal{F} \to \mathcal{M}$ with connections $\nabla_{\mathcal{E}}$ and $\nabla_{\mathcal{F}}$, we can define a connection $\nabla_{\mathcal{E}\otimes\mathcal{F}}$ on the tensor bundle $\pi \otimes \varrho : \mathcal{E} \otimes \mathcal{F} \to \mathcal{M}$ as

$$\nabla_{\mathcal{E}\otimes\mathcal{F}}(s_{\mathcal{E}} \otimes s_{\mathcal{F}}) \coloneqq \nabla_{\mathcal{E}}(s_{\mathcal{E}}) \otimes s_{\mathcal{F}} + s_{\mathcal{E}} \otimes \nabla_{\mathcal{F}}(s_{\mathcal{F}}),$$

for sections $s_{\mathcal{E}} \in \Gamma(\mathcal{E})$, $s_{\mathcal{F}} \in \Gamma(\mathcal{F})$ and $s_{\mathcal{E}} \otimes s_{\mathcal{F}} \in \Gamma(\mathcal{E} \otimes \mathcal{F})$. We will use this connection implicitly very often throughout this text, as our tensors will be living on various bundles but will need connections on their tensor products.

Example A.3.3 (Covariant derivative)**.** When a connection is on the tangent bundle we call it a **covariant derivative**, or linear connection. Because the connection takes vector fields to elements of $\Gamma(\mathcal{T}^*\mathcal{M} \otimes \mathcal{T}\mathcal{M})$, we can think of it as associating to every vector field $V \in \Gamma(\mathcal{T}\mathcal{M})$ a linear operator $\nabla_V : \Gamma(\mathcal{T}\mathcal{M}) \to \Gamma(\mathcal{T}\mathcal{M})$ satisfying (for $f \in \mathcal{C}^\infty(\mathcal{M})$) the following

1. The assignment $V \mapsto \nabla_V$ is $\mathcal{C}^\infty(\mathcal{M})$-linear:

   $$\nabla_{fV}(W) = f\nabla_V(W),$$

   for $W \in \Gamma(\mathcal{T}\mathcal{M})$;

2. It satisfies a Leibniz-type rule:

$$\nabla_V(fW) = V(f)W + f\nabla_V(W).$$

This describes the derivative of a vector W along the direction V, and depends only upon the value of V at the point $p \in \mathcal{M}$ where all of this is taking place and the values of W in the coordinate patch in which p lives. Thus, we can make our focus local and consider the above in some coordinate patch. Suppose that for every p in this patch, $(e_i(p))_{i=1}^n$ form a basis of $\mathcal{T}_p\mathcal{M}$ (restricted to the patch in question). From this perspective we find that the covariant derivative is completely determined by the **connection coefficients** A_{ij}^k defined by

$$A_{ij}^k e_k := \nabla_{e_i}(e_j),$$

where we make the important remark that despite the index notation, the connection coefficients are not tensors, though the covariant derivative of a tensor remains a tensor. Finally, note that in parallel our previous coordinate-shorthand, we will write

$$\nabla_i := \nabla_{\partial_i}.$$

Example A.3.4 (Dual (linear) connection)**.** For a tangent bundle $\pi : \mathcal{TM} \to \mathcal{M}$ with covariant derivative ∇ we define the **dual (linear) connection** ∇^* on the cotangent bundle as the map satisfying the following Leibniz-type rule for all vector fields $V \in \Gamma(\mathcal{TM})$:

$$V(\langle \varsigma, s\rangle) = \langle \nabla_V^*(\varsigma), s\rangle + \langle \varsigma, \nabla_V(s)\rangle,$$

where, for vector fields $s \in \Gamma(\mathcal{TM})$ and covector fields $\varsigma \in \Gamma(\mathcal{T}^*\mathcal{M})$, the natural pairing between $\mathcal{TM}$ and $\mathcal{T}^*\mathcal{M}$ has been written $\langle \varsigma, s\rangle = \varsigma(s) \in \mathcal{C}^\infty(\mathcal{M})$.

Example A.3.5 (Pull-back (linear) connection)**.** Given a tangent bundle $\pi : \mathcal{TN} \to \mathcal{N}$, a covariant derivative ∇, and a map of manifolds $\psi : \mathcal{M} \to \mathcal{N}$, we define the **pull-back (linear) connection** $\psi^*\nabla$ on the pull-back bundle $\psi^*\pi : \psi^*\mathcal{TN} \to \mathcal{M}$ by

$$(\psi^*\nabla)_V(\psi^*s) := \psi^*\left(\nabla_{\psi_*(V)}(s)\right),$$

where $V \in \Gamma(\psi^*\mathcal{TN})$ and ψ^*s a general section when s is a section.

From here we can introduce the Levi-Civita connection, as in Definition 2.1.8. To conclude this appendix we include the following extra definitions.

A.4 Additional Definitions

Definition A.4.1 (Torsion tensor)**.** The **torsion tensor** τ of a connection ∇ on a manifold $\mathcal{M}$ is a $(2,1)$-tensor defined by

$$\tau(V, W, \omega) := \omega\left(\nabla_V W - \nabla_W V - [V, W]\right),$$

where V and W are vectors and ω is a covector.

Definition A.4.2 (Symmetric space)**.** A semi-Riemannian manifold $(\mathcal{M}, g)$ is **locally symmetric** if every point $p \in \mathcal{M}$ has a normal neighbourhood $\mathcal{N}_p \subset \mathcal{M}$ and an isometry $\ell_p : \mathcal{N}_p \to \mathcal{N}_p$ (called a **geodesic reflection**) satisfying

$$\ell_p\left(\exp_p(V_p)\right) = \exp_p(-V_p) \quad \text{for all} \quad V_p \in \exp_p^{-1}(\mathcal{N}_p).$$

The manifold is **symmetric** if the above holds for $\mathcal{N}_p = \mathcal{M}$.[8]

Proposition A.4.1. *For the geodesic reflection ℓ_p defined above,*

1. *It squares to the identity: $\ell_p^2 = \mathbb{1}_{\mathcal{N}_p}$;*
2. *It satisfies both $\ell_p(p) = p$ and $(\ell_{p,*})_p = -\mathbb{1}_{T_p\mathcal{M}}$, restricted appropriately.*

Proof. The first statement is clear from the definition of ℓ_p. The second follows from noting that $p = \exp_p(0)$ and so

$$\ell_p(p) = \ell_p(\exp_p(0)) = \exp_p(-0) = p,$$

as well as from the following computation:

$$(\ell_{p,*})_p\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\exp_p(\lambda V_p)\bigg|_{\lambda=0}\right) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\ell_p\big(\exp_p(\lambda V_p)\big)\bigg|_{\lambda=0} = \frac{\mathrm{d}}{\mathrm{d}\lambda}\exp_p(-\lambda V_p)\bigg|_{\lambda=0} = -V_p,$$

where we have used $\frac{\mathrm{d}}{\mathrm{d}\lambda}\exp_p(\lambda V_p)\big|_{\lambda=0} = V_p$, by definition. □

Proposition A.4.2. *Symmetric spaces are complete. Further, the isometry group of a symmetric space acts transitively on the manifold.*

Proof. The geodesic reflection ℓ_p extends geodesics, ensuring completeness. □

Proposition A.4.3. *The Riemann tensor* Riem *(or R^i_{jkl}) on an n-dimensional semi-Riemannian manifold $(\mathcal{M}, g)$ satisfies the following (where we have lowered the upper index in the usual way for the first point):*

1. ***(Anti-)Symmetries**: $R_{ijkl} = R_{klij} = -R_{jikl} = -R_{ijlk}$;*
2. ***Bianchi identities**:*

$$R^i_{jkl} + R^i_{klj} + R^i_{ljk} = 0 \quad \textit{and} \quad \nabla_m R^i_{jkl} + \nabla_l R^i_{jmk} + \nabla_k R^i_{jlm} = 0.$$

3. *It has $\frac{1}{12}n^2(n^2-1)$ independent components—importantly, in 2 dimensions, it has only one independent component (greatly reducing computations), and in 3 dimensions it has 6;*
4. *If $(\mathcal{M}, g)$ is locally symmetric, then the Riemann tensor is parallel:[9] $\nabla(\mathrm{Riem}) = 0$.*

Definition A.4.3 (Weyl tensor)**.** Consider a semi-Riemannian manifold $(\mathcal{M}, g)$. In dimensions $n > 2$, **Weyl tensor** is defined as

$$W_{ijkl} := R_{ijkl} + \frac{2}{n-2}\left(g_{il}R_{jk} - g_{ik}R_{jl} + g_{jk}R_{il} - g_{jl}R_{ik}\right) + \frac{2R}{(n-1)(n-2)}(g_{ik}g_{jl} - g_{il}g_{jk}),$$

is thought of as the trace-free portion of the Riemann tensor.

[8] Note that if a manifold is complete, simply-connected and locally symmetric then it is symmetric. See Chapter 6 of [KN63] for details.

[9] This is sometimes given as the definition of a symmetric space.

Appendix B

(More) Accessible Overview

When we look at the surface of the Earth—ignoring the mountains and valleys, rivers and oceans—it appears to be planar. Indeed, from a close enough viewpoint, many objects and shapes resemble *flat space*: a straight line in one dimension, a plane in two, and so on. For instance, to a tiny microbe, the surface of a mug will appear to be a plane. Any smooth line drawn on a piece of paper that does not cross itself (if we forget sharp corners for the moment) appears to be straight if peered at closely enough.

Spaces that have these characteristics are called *manifolds* in mathematics. They allow mathematicians to work with functions, derivatives, and many more mathematical objects on an arbitrary space, instead of being confined to the unnatural world of graph paper. One famous example of a manifold is the fabric of the universe itself: spacetime is a four-dimensional manifold, containing one temporal dimension and three spatial dimensions, each interacting with one another. The study of spacetime and the resulting gravity is known as General Relativity, as first introduced by Einstein in the early twentieth century.

General Relativity used the mathematical language of manifolds that had been developed in the late nineteenth century and spurred further work in geometry because of its renewed pertinence to physics. In the past hundred years, General Relativity has been investigated by hundreds of physicists and mathematicians, many of whom hope to understand its intricacies in rigorous detail and to explain its behaviour in mathematical language. One exciting outcome of General Relativity is the existence of black holes and other types of singularities—points where things go wrong: densities increase to infinity, the causal structure of spacetime faces difficulties, or other peculiarities occur.

The plan for this project was to compare General Relativity with Ricci flow, a mathematical subject introduced in the 1980s that also encounters singular points. Ricci flow deals with a diffusion-type equation on a manifold. In a diffusion (or heat-like) equation, a distribution—such as temperature in a room—evolves so that it eventually becomes constant throughout space: if one part of a room is hot then eventually that heat will spread. Instead of a heat distribution, Ricci flow spreads the *curvature* of the manifold, so that as time passes it tends to become a space of constant curvature. This process can encounter singularities, such as the one depicted in Figure B.1.[1]

A mathematical result entitled the Uniformisation Theorem was chosen to unify the pillars of Ricci flow and General Relativity. The Uniformisation Theorem, which has been known to hold true since the mid-nineteenth century, classifies all two-dimensional manifolds as being related (via a relation called *conformal equivalence*) to a manifold of *constant* curvature. Manifolds of constant curvature include flat planes, perfect spheres,

[1] This appears as Figure 3.3 in Section 3.3 of the main text.

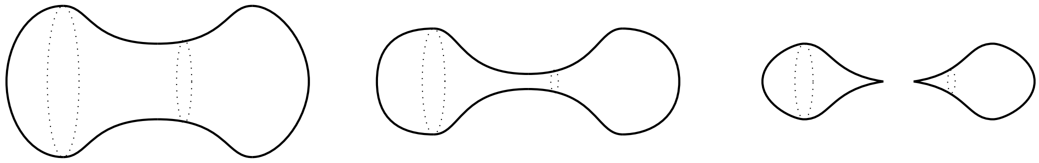

Figure B.1: An example of a singularity appearing during Ricci flow: the end result is two spheres, each having constant curvature. Since the goal of Ricci flow is to 'smoothen' the curvature of the space, this result is satisfactory; however, during Ricci flow, the manifold fractures into two pieces, which breaks down the mathematical description of the space.

and other spaces. (Because the Uniformisation Theorem deals with manifolds in two dimensions, the focus of this project was two *spatial*[2] dimensions.)

In 2006, a proof was completed that used Ricci flow to prove the Uniformisation Theorem.[3] This new proof is interesting because it gives a *constructive* proof: Ricci flow *gives* the conformal relation between the starting manifold and the resulting manifold of constant curvature. Rather than simply stating that *there exists* a connection between these manifolds (which would be a *non*-constructive proof), Ricci flow *is* that relationship.

As stated in the title, the Uniformisation Theorem is seen as a 'bridge' between Ricci flow and General Relativity because once a study of Ricci flow provides a proof, the Uniformisation Theorem is then repeatedly employed in an analysis of General Relativity. While most of the work contained herein has already been published in the mathematics literature, the hope of this project is to compile work found scattered throughout papers, textbooks, and monographs into one accessible document.

[2]The emphasis on spatial here is because our study of General Relativity is actually *three*-dimensional, since we must include one extra *temporal* dimension.

[3]The final piece of the puzzle appeared in a short paper by Chen, Lu and Tian [CLT06].

Bibliography

[AB10] B. Andrews and P. Bryan. Curvature bounds by isoperimetric comparison for normalized Ricci flow on the two-sphere. *Calculus of Variations and Partial Differential Equations*, 39(3-4):419–428, 2010.

[ADM59] R. Arnowitt, S. Deser, and C. W. Misner. Dynamical structure and definition of energy in general relativity. *Physical Review*, 116(5):1322, 1959.

[AM11] L. Andersson and V. Moncrief. Einstein spaces as attractors for the Einstein flow. *Journal of Differential Geometry*, 89(1):1–47, 2011.

[AMT97] L. Andersson, V. Moncrief, and A. J. Tromba. On the global evolution problem in 2+1 gravity. *Journal of Geometry and Physics*, 23(3-4):191–205, 1997.

[And04] M. T. Anderson. Geometrization of 3-manifolds via the Ricci flow. *Notices of the American Mathematical Society*, 51:184–193, 2004.

[Are13] S. Aretakis. Lectures on general relativity. *Lecture notes*, 2013.

[Bes07] A. L. Besse. *Einstein Manifolds*. Springer Science & Business Media, 2007.

[BK17] V. Branding and K. Kröncke. The Ricci flow with metric torsion on closed surfaces. *The Journal of Geometric Analysis*, 27(3):2098–2117, 2017.

[BS05] A. N. Bernal and M. Sánchez. Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. *Communications in Mathematical Physics*, 257(1):43–50, 2005.

[Car95a] S. Carlip. The (2+1)-dimensional black hole. *Classical and Quantum Gravity*, 12(12):2853, 1995.

[Car95b] S. Carlip. Lectures in (2+1)-dimensional gravity, 1995, gr-qc/9503024.

[Car03] S. Carlip. *Quantum Gravity in 2+1 Dimensions*, volume 50. Cambridge University Press, 2003.

[Car19] S. M. Carroll. *Spacetime and Geometry*. Cambridge University Press, 2019.

[CB52] Y. Choquet-Bruhat. Théorème d'existence pour certains systèmes d'équations aux dérivées partielles nonlinéaires. *Acta Mathematica*, 88:141–225, 1952.

[CB08] Y. Choquet-Bruhat. *General Relativity and the Einstein Equations*. OUP Oxford, 2008.

[CBR83] Y. Choquet-Bruhat and T. Ruggeri. Hyperbolicity of the 3+1 system of Einstein equations. *Communications in Mathematical Physics*, 89(2):269–275, 1983.

[CCY03] H.-D. Cao, S. C. Chu, and S.-T. Yau. *Collected Papers on Ricci Flow*. International Press, 2003.

[CH53] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience New York, 1953.

[Cha06] I. Chavel. *Riemannian Geometry: A Modern Introduction*, volume 98. Cambridge University Press, 2006.

[Cho91] B. Chow. The Ricci flow on the 2-sphere. *Journal of Differential Geometry*, 33(2):325–334, 1991.

[CK04] B. Chow and D. Knopf. *The Ricci Flow: An Introduction*, volume 1. American Mathematical Soc., 2004.

[CLT06] X. Chen, P. Lu, and G. Tian. A note on uniformization of Riemann surfaces by Ricci flow. *Proceedings of the American Mathematical Society*, pages 3391–3393, 2006.

[DeT83] D. M. DeTurck. Deforming metrics in the direction of their Ricci tensors. *Journal of Differential Geometry*, 18(1):157–162, 1983.

[Dir50] P. A. M. Dirac. Generalized Hamiltonian dynamics. *Canadian Journal of Mathematics*, 2:129–148, 1950.

[Dir51] P. A. M. Dirac. The Hamiltonian form of field dynamics. *Canadian Journal of Mathematics*, 3:1–23, 1951.

[DK20] A. Deruelle and K. Kröncke. Stability of ALE Ricci-flat manifolds under Ricci flow. *The Journal of Geometric Analysis*, pages 1–42, 2020.

[Don11] S. Donaldson. *Riemann Surfaces*. Oxford University Press, 2011.

[Ein15] A. Einstein. The field equations of gravitation. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1915:844–847, 1915.

[Ein22] A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitäts-Theorie. In *Das Relativitätsprinzip*, pages 130–139. Springer, 1922.

[Eva10] L. C. Evans. *Partial Differential Equations*, volume 19. American Mathematical Society, 2010.

[Fis04] A. E. Fischer. An introduction to conformal Ricci flow. *Classical and Quantum Gravity*, 21(3):S171, 2004.

[FK15] D. Fajman and K. Kröncke. Stable fixed points of the Einstein flow with positive cosmological constant, 2015, 1504.00687.

[FK18] D. Fajman and K. Kröncke. On the CMC-Einstein-Λ flow. *Classical and Quantum Gravity*, 35(19):195005, 2018.

[FM96a] A. E. Fischer and V. Moncrief. A method of reduction of Einstein's equations of evolution and a natural symplectic structure on the space of gravitational degrees of freedom. *General Relativity and Gravitation*, 28(2):207–219, 1996.

[FM96b] A. E. Fischer and V. Moncrief. Quantum conformal superspace. *General Relativity and Gravitation*, 28(2):221–237, 1996.

[FM96c] A. E. Fischer and V. Moncrief. The structure of quantum conformal superspace. In *Global Structure and Evolution in General Relativity*, pages 111–173. Springer, 1996.

[FT84] A. E. Fischer and A. J. Tromba. On a purely "Riemannian" proof of the structure and dimension of the unramified moduli space of a compact Riemann surface. *Mathematische Annalen*, 267(3):311–345, 1984.

[Gou12] E. Gourgoulhon. *3+1 Formalism in General Relativity: Bases of Numerical Relativity*, volume 846. Springer Science & Business Media, 2012.

[Gra13] J. Gray. *Henri Poincaré: A Scientific Biography*. Princeton University Press, 2013.

[Ham82] R. S. Hamilton. Three-manifolds with positive Ricci curvature. *Journal of Differential Geometry*, 17(2):255–306, 1982.

[Ham88] R. S. Hamilton. The Ricci flow on surfaces. In *Mathematics and General Relativity, Proceedings of the AMS-IMS-SIAM Joint Summer Research Conference in the Mathematical Sciences on Mathematics in General Relativity, Univ. of California, Santa Cruz, California, 1986*, pages 237–262. Amer. Math. Soc., 1988.

[Haw66] S. Hawking. *Properties of Expanding Universes*. PhD thesis, University of Cambridge, 1966.

[HE73] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*, volume 1. Cambridge University Press, 1973.

[HP83] S. W. Hawking and D. N. Page. Thermodynamics of black holes in anti-de Sitter space. *Communications in Mathematical Physics*, 87(4):577–588, 1983.

[HW06] M. Headrick and T. Wiseman. Ricci flow and black holes. *Classical and Quantum Gravity*, 23(23):6683, 2006.

[IMS11] J. Isenberg, R. Mazzeo, and N. Sesum. Ricci flow in two dimensions, 2011, 1103.4669.

[Kle83] F. Klein. Neue Beiträge zur Riemann'schen Functionentheorie. *Mathematische Annalen*, 21(2):141–218, 1883.

[KN63] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*, volume 1. Wiley Classics Library, 1963.

[KN69] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*, volume 2. Wiley Classics Library, 1969.

[Koi78] N. Koiso. Nondeformability of Einstein metrics. *Osaka Journal of Mathematics*, 15(2):419–433, 1978.

[Koi79] N. Koiso. A decomposition of the space of Riemannian metrics on a manifold. *Osaka Journal of Mathematics*, 16(2):423–429, 1979.

[Krö13] K. Kröncke. Stability of Einstein metrics under Ricci flow, 2013, 1312.2224.

[Krö15] K. Kröncke. On the stability of Einstein manifolds. *Annals of Global Analysis and Geometry*, 47(1):81–98, 2015.

[Lan18] K. Landsman. Foundations of general relativity. *Lecture notes*, 2018.

[Lan19] K. Landsman. Ricci flow: An introduction. *Lecture notes*, 2019.

[Lee06] J. M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Springer Science & Business Media, 2006.

[Lee13] J. M. Lee. *Introduction to Smooth Manifolds*. Springer, 2013.

[LSU88] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Uralćeva. *Linear and Quasi-Linear Equations of Parabolic Type*, volume 23. American Mathematical Soc., 1988.

[Mar84] E. Martinec. Soluble systems in quantum gravity. *Physical Review D*, 30(6):1198, 1984.

[Mon86] V. Moncrief. Reduction of Einstein's equations for vacuum space-times with spacelike U(1) isometry groups. *Annals of Physics*, 167(1):118–142, 1986.

[Mon89] V. Moncrief. Reduction of the Einstein equations in 2+1 dimensions to a Hamiltonian system over Teichmüller space. *Journal of Mathematical Physics*, 30(12):2907–2914, 1989.

[Mon07] V. Moncrief. Relativistic Teichmüller theory: A Hamilton-Jacobi approach to 2+1-dimensional Einstein gravity. *Surveys in Differential Geometry*, 12(1):203–250, 2007.

[MR93] R. B. Mann and S. F. Ross. Gravitationally collapsing dust in 2+1 dimensions. *Physical Review D*, 47(8):3319, 1993.

[MTW73] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. Macmillan, 1973.

[Omo70] H. Omori. On the group of diffeomorphisms on a compact manifold. In *Proc. Symp. Pure Appl. Math., XV, Amer. Math. Soc*, pages 167–183, 1970.

[Pen65] R. Penrose. Gravitational collapse and space-time singularities. *Physical Review Letters*, 14(3):57, 1965.

[Per02] G. Perelman. The entropy formula for the Ricci flow and its geometric applications, 2002, math/0211159.

[Per03a] G. Perelman. Finite extinction time for the solutions to the Ricci flow on certain three-manifolds, 2003, math/0307245.

[Per03b] G. Perelman. Ricci flow with surgery on three-manifolds, 2003, math/0303109.

[Poi82] H. Poincaré. Mémoire sur les fonctions Fuchsiennes. *Acta mathematica*, 1:193–294, 1882.

[Poi04] H. Poincaré. Cinquième complément à lâanalysis situs. *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 18(1):45–110, 1904.

[Rin09] H. Ringström. *The Cauchy Problem in General Relativity*, volume 6. European Mathematical Society, 2009.

[She06] N. Sheridan. Hamilton's Ricci flow. *Honours thesis*, 2006.

[Top06] P. Topping. *Lectures on the Ricci Flow*, volume 325. Cambridge University Press, 2006.

[Tro92] A. Tromba. *Teichmüller Theory in Riemannian Geometry*. Birkhäuser Bäsel, 1992.

[Wal10] R. M. Wald. *General Relativity*. University of Chicago press, 2010.

[Woo08] E. Woolgar. Some applications of Ricci flow in physics. *Canadian Journal of Physics*, 86(4):645–651, 2008.

[YJ72] J. W. York Jr. Role of conformal three-geometry in the dynamics of gravitation. *Physical Review Letters*, 28(16):1082, 1972.

Index of Terms

Index of Symbols

symbol	meaning
$\mathbb{1}$	identity
∂_i	partial derivative in the x^i direction
$(\partial_i)_{i=1}^{n}$	coordinate basis of $\mathcal{T}\mathcal{M}$
∇	Levi-Civita connection
Γ^i_{jk}	Christoffel symbols
$\gamma_{p,V}$	geodesic starting at p with velocity V
Δ_g	Laplacian of g
η	normal to a submanifold
κ_{ij}	extrinsic curvature
Λ	cosmological constant
π	circle constant
π^{ij}	gravitational momentum
ϱ	radius
Σ	2-dimensional manifold, or Cauchy surface
$\mathscr{A}$	almost complex structures
$\mathcal{B}_{g,\varrho}(p)$	ball with radius ϱ at p
$\mathscr{C}$	complex structures
$\mathcal{C}^\infty(\mathcal{M})$	smooth functions on $\mathcal{M}$
$\mathbb{C}$	complex plane
c	complex structure
$\mathscr{D}_{(0)}$	(connected component of) diffeomorphism group
$(\mathrm{d}x^i)_{i=1}^{n}$	dual coordinate basis
$d_g(p,q)$	distance between p and q
e	Euler's constant
$\exp_{(p)}$	exponential function (at p)
f	generic function, or potential function
$\mathrm{g}(\Sigma)$	genus of Σ
g	semi-Riemannian metric
$\mathscr{H}^s$	Sobolev space with parameter s
$\mathcal{H}$	Hamiltonian constraint

symbol	meaning
$\mathcal{H}^i$	momentum constraint
$\mathbb{H}^n$	hyperbolic n-space
H	H scalar
i	imaginary unit
$\mathcal{L}_V$	Lie derivative in the direction of V
L ($\hat{L}$)	(modified) L scalar
$\ell(\gamma)$	length of γ
$\mathscr{M}_{(\lambda)}$	space of Riemannian metrics (with Ricci scalar $\lambda \in \mathbb{R}$)
$\mathcal{M}$	generic manifold
$(\mathcal{M}, g)$	semi-Riemannian manifold
M	M tensor
N and N^i	lapse and shift
n	dimension of a manifold
$(p_\alpha)_\alpha$	momentum components
$\overline{p}^{ij}$	conformal momentum
Q ($\hat{Q}$)	(modified) Q scalar
$\mathbb{R}$	real numbers
Ric, R_{kl}	Ricci tensor
Riem, R^i_{jkl}	Riemann tensor
R	Ricci scalar
r	average Ricci scalar
$\mathcal{S}$	action
$\mathbb{S}^n$	n-sphere
s	Sobolev parameter
$\mathscr{T}$	Teichmüller space
$\mathcal{T}_{(p)}\mathcal{M}$	tangent space (at p)
$\mathcal{T}^*_{(p)}\mathcal{M}$	cotangent space (at p)
$\mathcal{T}^{(k,l)}\mathcal{M}$	(k,l)-tensor space
$\mathtt{T}$	York time
$\mathtt{t}$	time function
t	time coordinate
$V_{(p)}$	vector field (at p)
$\mathrm{vol}(\mathcal{M}, g)$	volume of $(\mathcal{M}, g)$
$(x^i)_{i=1}^{n}$	coordinates on a patch