Ricci Flow: A First Introduction. Lecture notes for master course by Klaas Landsman

Contents

Ricci Flow: A First Introduction

Klaas Landsman

October 21, 2019

1 Introduction

The aim of these notes (and the associated course of seven lectures) is to bring readers with elementary knowledge of Riemannian geometry and PDEs to a level where they can start reading serious books and articles about Hamilton's Ricci Flow and Perelman's proof of the Poincaré and Geometrization Conjectures based on this technique (see Literature). There will, of course, be some overlap with the introductory parts of such literature, for example with Chow & Knopf (2004), Topping (2006), and Andrews & Hopper (2011), but we aim to give more background and discuss some basic concepts and examples in a bit more detail. Our aim will have been achieved if those familiar with the material below would go and study e.g. Tao (2008) and/or Morgan and Fong (2010). However, even short of this, the material we present is interesting by itself as an example of what is called *Geometric Analysis*, a modern mathematical discipline in which the 2019 Abel Prize was awarded to Karen Uhlenbeck on March 19th, 2019.

Let (M,g) be an n-dimensional Riemannian manifold, which throughout these notes we assume to be connected, orientable, and without boundary (unless explicitly stated otherwise).

- ***Ricci Flow*** on M is a solution to the PDE (Hamilton, 1982)

$$\partial_t g(t) = -2\text{Ric}(g(t)), \tag{1.1}$$
$$g(0) = g, \tag{1.2}$$

where $\text{Ric}(g)$ is the Ricci *tensor* defined by g and we have suppressed the x-dependence of g and Ric ($x \in M$). Oe often writes Ric for $\text{Ric}(g)$, so that in local coordinates we have

$$\frac{\partial g_{ij}}{\partial t} = -2\text{Ric}_{ij}. \tag{1.3}$$

So the idea is that the metric becomes a dynamical object, hoping that it will 'flow' to some particularly desirable metric from more or less arbitrary initial data g (see below).

It is interesting to compare Ricci Flow with two other flows on a Riemannian manifold:

- ***Yamabe Flow*** is the arguably even simpler equation (now suppressing the (t) as well)

$$\partial_t g = -R(g) \cdot g, \tag{1.4}$$

i.e., $\partial_t g_{ij} = -R g_{ij}$, where R is the Ricci *scalar* derived from g;

- ***Einstein Flow*** is the more complicated pair of equations

$$\partial_t g = -2k, \tag{1.5}$$

$$\partial_t k = \mathrm{Ric} + \mathrm{Tr}\,(k)\cdot k - 2k^2, \tag{1.6}$$

where k is some symmetric 2-tensor of the same type as the metric (to be interpreted as the second fundamental form), and $\mathrm{Tr}\,(k) = g^{ij}k_{ij}$ and $k^2_{ij} = k^l_i k_{lj} = g^{lm}k_{il}k_{mj}$. The initial values (g,k) at $t=0$ are subject to the following *constraints* (where ∇ is the Levi-Civita connection):

$$R + \mathrm{Tr}\,(k)^2 - \mathrm{Tr}\,(k^2) = 0; \tag{1.7}$$

$$\nabla_j k^j_i - \nabla_i \mathrm{Tr}\,(k) = 0. \tag{1.8}$$

We will return to the Einstein equations only at the very end of the course, but note that where Ricci Flow and Yamabe Flow are *parabolic* (at least in suitable coordinates), Einstein Flow is *hyperbolic*. Nonetheless, in all cases the solution may and often will become singular after some finite time, and in this light it is useful to compare especially the first case (where thanks to Perelman one has very good control over the singularities) with the last (where despite the brilliant work of Hawking and Penrose many questions about singularities remain).

Let us now give some motivation for studying Ricci Flow and Yamabe Flow, which, as we shall prove next week, coincide in dimension 2. The ***Uniformization Theorem*** for Riemann surfaces was a notable achievement of pure mathematics in the early 20th Century;[1] it seems no accident that Poincaré gave one of the proofs of this theorem *and* is the man behind the conjecture named after him, since *conceptually* the two are closely related (although *technically* this relationship only became clear with the invention of Ricci Flow in 1982). We will avoid complex analysis, but for historical reasons we first give (more or less) the original version of this theorem.[2] In what follows, a simply connected space is also connected by definition.

Theorem 1.1 (Uniformization Theorem, version 1) *Every simply connected Riemann surface is (biholomorphically) isomorphic to exactly one of the following three possibilities:*

- *The Riemann sphere $\mathbb{S}$;*
- *The complex plane $\mathbb{C}$;*
- *The upper half plane $\mathbb{H}$ in $\mathbb{C}$ (or, equivalently, the open unit disk $\mathbb{D}$ in $\mathbb{C}$).*

It turns out that this is equivalent to a simple statement purely in the language of Riemannian geometry, where we say that two Riemannian metrics g_1 and g_2 (defined on the same manifold M) are ***conformally equivalent*** if $g_2 = C\cdot g_1$ for some strictly positive function $C \in C^\infty(M)$.

Theorem 1.2 (Uniformization Theorem, version 2) *Every complete Riemannian metric on a simply connected 2d manifold is conformally equivalent to a metric with constant curvature.*[3]

[1] For a nice and exhaustive historical survey of the uniformization theorem see de Saint-Gervais (2010).

[2] A Riemann *surface* is defined through its complex structure, whereas a Riemannian *manifold* is defined by its metric. Roughly speaking (in dimension 2), complex structures up to biholomorphic equivalence bijectively correspond to Riemannian metrics up to the equivalence relation defined by isometry and conformal equivalence.

[3] A Riemannian metric is called ***complete*** if either one (and hence both) of the following conditions is satisfied (whose equivalence follows from the fundamental Hopf-Rinow Theorem of Riemannian Geometry):

1. The underlying topological space is complete in the metric d derived from the Riemannian metric g;
2. Each geodesic defined by g is defined for all time (i.e. can be infinitely extended in both directions).

As we shall see, in $d=2$ all notions of curvature (Gauss curvature, Ricci scalar, Ricci tensor, Riemann tensor, sectional curvature) are essentially the same, so it does not matter which particular one is used here.

The reason that these two formulations of the Uniformization Theorem are equivalent is this:

Theorem 1.3 *In any dimension $n \geq 2$, every Riemannian manifold (M,g) for which M is simply connected and g is complete and has constant sectional curvature, is isometrically isomorphic to exactly one of the following possibilities:*

- *The n-dimensional sphere S^n_k in $\mathbb{R}^{n+1}$ with radius ρ (and metric inherited from $\mathbb{R}^{n+1}$);*
- *$\mathbb{R}^n$ with flat (i.e. Euclidean) metric;*
- *The upper hyperplane H^n_ρ in $\mathbb{R}^n$ with a "hyperbolic" metric parametrized by ρ.*[4]

Moreover, in cases 2 and 3 the metric can be rescaled by a positive constant to achieve $\rho = 1$.

For $n = 2$ these are the three cases in Theorem 1.1, but now seen as Riemannian geometries.[5] As such, they were of historic importance in the 19th Century in providing the first examples of non-Euclidean geometries, subsequently generalized into Riemannian geometry.

The conceptual importance of this result can hardly be overstated, since it shows that the possibility of equipping a manifold with a metric with constant curvature classifies that manifold. This is precisely the philosophy behind Ricci Flow: starting with an arbitrary (Riemannian) metric g on some manifold M, the flow drives it to a metric with constant curvature (albeit at the expense of huge complications in $d = 3$). This is seen more clearly from the induced flow of the Ricci scalar R, which, under Ricci Flow, satisfies the reaction-diffusion equation

$$\partial_t R = \Delta_g R + 2\mathrm{Tr}\,(\mathrm{Ric}^2), \tag{1.9}$$

where Δ_g is the (scalar) Laplacian defined by g. As in the heat equation, the diffusion term $\Delta_g R$ tends to spread the 'heat', that is, homogenize the curvature. The picture is relatively clean in $d = 2$, where, still assuming that M is simply connected (see below for the general case), for an arbitrary initial metric $g = g(0)$ there are, once again, our familiar three possibilities in disguise:

- The solution breaks down at some finite time $0 < T < \infty$, where $\det(g(T)) = 0$, but

  $$\tilde{g}(T) = \lim_{t \to T} \frac{g(t)}{T - t} \tag{1.10}$$

  exists. Then M must be diffeomorphic to S^2 and $\tilde{g}(T)$ is its usual ("round") metric.

- The solution exists for all times t *and* has a limit $g(\infty)$ for $t \to \infty$. In that case, M is diffeomorphic to $\mathbb{R}^2$ and the limit metric $g(\infty)$ is simply the flat one.

- The solution exists for all t *but* does not have a limit for $t \to \infty$ since $\det(g(t)) \to \infty$. Then M is again diffeomorphic to $\mathbb{R}^2$, hence to $\mathbb{H}$, and

  $$\hat{g}(\infty) = \lim_{t \to \infty} \frac{g(t)}{t} \tag{1.11}$$

  exists and equals its hyperbolic metric (perhaps up to scaling by a constant).

[4]This is *not* the metric induced by the one on $\mathbb{R}^n$; we will describe it in detail later on. The upper half-plane $\mathbb{H}^n \subset \mathbb{R}^n$ is defined by $x^n > 0$ and $(x^1, \ldots, x^{n-1}) \in \mathbb{R}^{n-1}$ arbitrary. One may also take the open unit ball in $\mathbb{R}^n$.

[5]So $\mathbb{S} = S^2$ as a space, but $\mathbb{S}$ is seen as a complex curve whereas S^2 is seen as a $2d$ Riemannian manifold, etc.

In particular, in all cases the limit metrics have constant curvature, confirming the intuition that any room (with an initial temperature gradient) eventually reaches thermal equilibrium; just replace temperature by curvature. These cases can be told apart by the ***Gauss-Bonnet Theorem***

$$\int_M R = 4\pi\chi, \tag{1.12}$$

where χ is the ***Euler characteristic*** of M seen as a topological space, and hence is necessarily an integer.[6] If $\chi > 0$ one has the first case, $\chi = 0$ gives the second, and $\chi < 0$ yields the third.

The above scenario can easily and explicitly (albeit somewhat circularly) be demonstrated if we assume that the initial metric g is a so-called ***Einstein metric***,[7] that is, satisfies

$$\mathrm{Ric}(g) = \lambda \cdot g, \tag{1.13}$$

or $\mathrm{Ric}_{ij} = \lambda g_{ij}$, for some constant $\lambda \in \mathbb{R}$. In that case, the Ricci Flow with $g(0) = g$ is given by

$$g(t) = (1 - 2\lambda t)g, \tag{1.14}$$

as follows by checking that this is *a* solution to (1.1) - (1.2), and anticipating uniqueness of solutions for short time (as we will prove later), which makes (1.14) *the* solution.[8] Once again:

- If $\lambda > 0$, then $\det(g)$ for $t = T = 1/2\lambda$, but the rescaled limit metric (1.10) exists:

$$\lim_{t\to T} \frac{g(t)}{T-t} = 2\lambda g. \tag{1.15}$$

- If $\lambda = 0$, then $g(t)$ is g for all time, including the limit $t \to \infty$.
- If $\lambda < 0$, then g exists for all time but blows up in the limit; the rescaled limit (1.11) is

$$\lim_{t\to\infty} \frac{g(t)}{t} = -2\lambda g. \tag{1.16}$$

Using some topology, Theorem 1.1 gives rise to a complete classification of compact (orientable) Riemann surfaces, which we briefly mention because a similar (but vastly more complicated) construction applies in $d = 3$ and motivates Thurston's Geometrization Conjecture:

Corollary 1.4 *Any compact Riemann surface Σ is (biholomorphically) isomorphic to $\Gamma\backslash\mathbb{U}$, where $\mathbb{U}$ is $\mathbb{S}$, $\mathbb{C}$, or $\mathbb{H}$, and Γ is a discrete subgroup of the group of biholomorphic bijections of $\mathbb{U}$ acting freely and discontinuously on $\mathbb{U}$ (i.e., no Γ-orbit has an accumulation point).*[9]

- For $\mathbb{U} = \mathbb{S}$ only the trivial group Γ is possible (this gives genus $g = 0$);
- For $\mathbb{U} = \mathbb{C}$, Γ can only be the lattice $\mathbb{Z}^2$ in $\mathbb{R}^2 \cong \mathbb{C}$, yielding the donut $\mathbb{Z}^2\backslash\mathbb{C} \cong \mathbb{T}^2$ ($g = 1$);
- For $\mathbb{U} = \mathbb{H}$, Γ must be a discrete subgroup of $PSL(2,\mathbb{R})$ (giving all genera $g > 1$).

[6] For an arbitrary triangulation of M one has $\chi = v - e + f$, where v is the number of vertices, e is the number of edges, and f is the number of faces of the triangulation. This is a topological invariant.

[7] This term originates in Einstein's equations for the gravtitational field with cosmologcal constant, and is also used in mathematics. Metrics of constant curvature are Einstein metrics, but the converse is false in $d > 3$.

[8] For *constant* $C > 0$ one has $\mathrm{Ric}(C \cdot g) = \mathrm{Ric}(g)$, to be used with $C = 1 - 2\lambda t$, where $0 \leq t < 1/2\lambda$.

[9] Equivalently, each $x \in \mathbb{U}$ has a nbhd U such that $U \cap \gamma \cdot U = \emptyset$ for all $\gamma \neq e$.

If we switch from Riemann surfaces to $2d$ Riemannian manifolds and hence start from Theorem 1.2 instead of Theorem 1.1, the corresponding version of this corollary reads as follows:[10]

Corollary 1.5 *Any metric on a compact* $2d$ *Riemannian manifold is conformally equivalent to a metric with constant curvature (and may be realized analogously to the list in Corollary 1.4).*

The ***Yamabe Problem*** is inspired by Corollary 1.5 and asks if in arbitrary dimension any complete Riemannian metric on a simply connected manifold is conformally equivalent to a metric with constant Ricci scalar.[11] The Yamabe Problem has been solved in the positive for *compact* manifolds (which are automatically complete), using the following strategy, going back to Yamabe himself (see Lee & Parker, 1987, and Bär, 2007/08, for a complete treatment): parametrize

$$\overline{g}(x) = \varphi^{p-2}(x)g(x), \tag{1.17}$$

where $p = 2n/(n-2)$ and $n = \dim(M)$ as usual (e.g. $n = 3 \Rightarrow p = 6$). Defining

$$\tilde{\Delta}_g = -4 \cdot \frac{n-1}{n-2} \cdot \Delta_g + R(g), \tag{1.18}$$

a straightforward computation gives

$$R(\overline{g}) = \varphi^{1-p}\tilde{\Delta}_g\varphi. \tag{1.19}$$

so that the Ricci scalar $R(\overline{g})$ in the conformally transformed metric is constant, say $R(\overline{g}) = \lambda$, if

$$\tilde{\Delta}_g\varphi = \lambda\varphi^{p-1}. \tag{1.20}$$

This is a complicated nonlinear *elliptic* PDE, which took about 30 years to be treated correctly.

Alternatively, one could use Yamabe Flow (1.4), albeit in the so-called *normalized form*

$$\partial_t g = -(R(g) - \langle R(g)\rangle) \cdot g, \tag{1.21}$$

where $\langle R(g)\rangle$, the average Ricci scalar, is a (g-dependent) constant function on M given by

$$\langle R(g)\rangle = \frac{\int_M R(g)}{\mathrm{Vol}(M)}, \tag{1.22}$$

taken with respect to the Riemannian volume form, so that in any coordinate system one has

$$\mathrm{Vol}(M) = \int_M d^n x \sqrt{\det(g(x))}; \tag{1.23}$$

$$\int_M R = \int_M d^n x \sqrt{\det(g(x))}R(x). \tag{1.24}$$

Eq. (1.21) may be solved from the *Ansatz*

$$g(x,t) = \varphi(x,t)^{p-2}g(x), \tag{1.25}$$

cf. (1.17), which gives

$$\partial_t(\varphi^{1-p}) = \frac{n+2}{4}(-\tilde{\Delta}_g\varphi + \langle R(g)\rangle\varphi^{p-1}). \tag{1.26}$$

This is a nonlinear *parabolic* PDE, which turns out to be easier to analyze than (1.20); the corresponding metric $g(x,t)$ converges to one with constant curvature as $t \to \infty$ (cf. Brendle, 2011)! Hence the Yamabe Problem can be solved (anew) using the Ricci Flow philosophy.

[10]Corollary 1.4 may be found in any book on Riemann surfaces, cf. G.A. Jones and D. Singerman, *Complex Functions* (CUP, 1987). For Corollary 1.5 see Vinberg (1983).

[11]Continuing footnote 3, his is the only choice among the many equivalent notions of curvature which all coincide in $d = 2$ for which there is any hope for the problem to have a solution.

The Poincaré Conjecture and Thurston's Geometrization Conjecture

The most spectacular use of Ricci Flow to date has been Perelman's proof of the Poincaré Conjecture,[12] and the ensuing proof of Thurston's Geometrization Conjecture,[13] which implies the Poincaré Conjecture. We now briefly review these conjectures (which have become theorems).

Theorem 1.6 (Poincaré Conjecture, 1904) *Any compact simply connected* 3*d manifold is diffeomorphic to the three-sphere* S^3.

It does not matter whether this is stated for topological manifolds or for smooth manifolds, since in $d = 2$ and $d = 3$ these notions coincide. The situation is as follows (Scorpan, 2005, p. vii):

- In dimension $d \leq 3$ there is exactly one smooth structure on any topological manifold;
- In $d = 4$ there are examples of uncountably many smooth structures on a given topological manifold (even on $\mathbb{R}^4$, as shown by Donaldson) and there isn't even a single known example of a topological manifold admitting only finitely many smooth structures.
- In $d > 4$ any compact topological manifold has only finitely many smooth structures.

So the Poincaré Conjecture is a purely topological problem, which makes it all the more remarkable that it was eventually solved using geometric analysis.[14] It is also true in $d = 2$, where it follows from the Uniformization Theorem.[15] It is false in any dimension $d > 3$.[16]

The first step towards proving the Poincaré Conjecture using Ricci Flow was the following:

Theorem 1.7 (Hamilton, 1982) *If some compact* 3*d manifold* M *admits a Riemannian metric with strictly positive Ricci curvature* $\mathrm{Ric}(g) > 0$ *(i.e.,* $\mathrm{Ric}_{ij}(x)X_x^i X_x^j > 0$ *for all points* $x \in M$ *and all nonzero vectors* $X_x \in T_xM$*), then* M *also has a metric with constant (sectional) curvature.*

Here and in what follows, we use the *Einstein summation convention: repeated indices in diagonal position are summed over*. Combined with Theorem 1.3, this clearly implies:

Corollary 1.8 *Any compact simply connected* 3*d manifold* M *that admits a Riemannian metric with strictly positive Ricci curvature is diffeomorphic to* S^3.

So what was lacking towards the Poincaré Conjecture is the implication: *Any compact simply connected* 3*d manifold supports a Riemannian metric with strictly positive Ricci curvature.*

This followed from the work of Perelman, which was general enough to settle the *Geometrization Conjecture* (also called *Program*) due to Thurston (1982), which we now briefly describe.[17]

---

[12] The original papers are Perelman (2002, 2003ab). For elaborations on his proof see e.g. Kleiner & Lott (2006), Morgan & Tang (2007) and Tao (2008), which is summarized in Tao (2006).

[13] See Cao & Zho (2006), Bessières et al (2010), and Morgan & Fong (2010) for detailed treatments.

[14] The relationship between topology and Riemannian geometry arguably started with the Gauss–Bonnet Theorem. Examples of topological results proved using Riemannian geometry that predate Hamilton (1982) include ***Myer's Theorem***, which states that if M is a connected n-dimensional manifold admitting a Riemannian metric g for which $\mathrm{Ric}(g) \geq (n-1) \cdot C \cdot g$, for some constant $C > 0$, then M is compact with finite $\pi_1(M)$, as well as the ***Sphere Theorem***, which states that if $\frac{1}{4} < C(x) \leq 1$ for all sectional curvatures $C(x)$, then $M \cong S^n$, see Brendle (2010) for a proof using Ricci Flow. Analogous results are described in the 'panoramic survey' Berger (2003).

[15] One may equip any smooth 2d manifold with either a complex structure or a Riemannian metric; the other two simply connected cases allowed by the Uniformization Theorem are not compact and hence only S^2 remains.

[16] However, the assumptions may be strengthened to requiring $\pi_k(M) \cong \pi_k(S^d)$ for all homotopy groups $\pi_k(M)$, in which case the corresponding conjecture is sometimes true. For example, under these stronger assumptions Smale proved the claim in the topological case for all $d > 4$, and Freedman accomplished the same in $d = 4$, so that it holds for all $d \geq 2$, whereas in the smooth case it is true for $d = 1, 2, 3, 5$, and 6 (and hence false in $d = 4$).

[17] For extensive background see Scott (1983) and, much more briefly, Anderson (2004).

> 'Thurstons great insight was to see how geometry could be used to understand the topology of three-manifolds.' (Jackson, 2006)

Thurston's ideas on (compact) three-dimensional manifolds are best understood in the light of Corollaries 1.4 and 1.5, in one dimension lower, to which there are both analogies and contrasts:

- Instead of the *three* spaces S^2, $\mathbb{R}^2$, and H^2, from which all compact $2d$ manifolds M can be produced as quotients by suitable discrete groups Γ (for example, by endowing M with a Riemannian metric and using Corollary 1.5), in $d = 3$ there are *eight* basic spaces, called

$$S^3, \quad \mathbb{R}^3, \quad H^3, \quad S^2 \times \mathbb{R}, \quad H^2 \times \mathbb{R}, \quad \widetilde{SL_2(\mathbb{R})}, \quad \text{Nil}, \quad \text{Sol}. \tag{1.27}$$

  The first three might have been expected as straightforward extensions from $d = 2$. The last three are specific three-dimensional Lie groups discussed in some detail in Appendices A and B. In fact, also $S^3 \cong SU(2)$ and $\mathbb{R}^3$ are three-dimensional Lie groups, whereas the remaining three cases, viz. H^3, $S^2 \times \mathbb{R}$, and $H^2 \times \mathbb{R}$, are at least *homogeneous spaces* of Lie groups (as will be explained later). Each of these eight spaces is endowed with a specific Riemannian metric, and as such is simply called a *geometry* by Thurston.

- It is no longer the case that any compact $2d$ manifolds M is a quotient of one of these eight spaces by some discrete group. Instead, M first has to be decomposed into "irreducible" spaces M_i,[18] from which M can be recovered (up to homeomorphism) as their connected sum over either S^2 or $\mathbb{T}^2$; see Scott (1983) or Anderson (2004)–this notion is also very well explained in Wikipedia, see `https://en.wikipedia.org/wiki/Connected_sum`.

Thurston's Geometrization Conjecture (now a theorem) may be then stated as follows:

Theorem 1.9 *Each compact three-manifold M can be obtained (up to diffeomorphism, or, equivalently, homeomorphism), as a connected sum of finitely many irreducible components M_i. Each M_i is a quotient $M_i = \Gamma\backslash\mathbb{V}$, where $\mathbb{V}$ is a Riemannian manifold in the list* (1.27) *and Γ is a discrete subgroup of the isometry group of $\mathbb{V}$ acting freely and discontinuously on $\mathbb{V}$.*

Equivalently, the decomposition of M into irreducible components M_i has finitely many terms and each M_i takes the form stated in the theorem (again, up to diffeomorphism). In fact, this is the way the theorem is proved using Ricci Flow, and it is remarkable that this technique gives both the decomposition of M and the identification of the ensuing components M_i. Roughly speaking, the singularities of the flow split M up into the M_i, on each of which the Ricci Flow then has to be manipulated so as to obtain asymptotic metrics that allow the identification of M_i. Alas, this identification is not simply made in terms of constant curvature, for in that case only the first three spaces in (1.27) would have been involved, but nonetheless it turns out that, just as in $2d$, the "majority" of possibilities arises from the hyperbolic geometry H^3. Finally, let us note that the Geometrization Conjecture implies the Poincaré Conjecture, similarly to the Unformization Theorem implying its $2d$ analogue, namely by an explicit inspection of cases and excluding all other possibilities (the proof by Perelman predated this strategy, though).

Those who wish to read about this in a leisurely way are referred to the beautiful surveys by Milnor (1982, 2003, 2014) and Anderson (2004); see also Morgan (2006) on the Poincaré Conjecture.[19] Details in Cao & Zhu (2006), Morgan & Fong (2010), Bessières et al (2010), etc.

[18]Technically, by "irreducible" we mean *both* sphere-irreducible and torus-irreducible.

[19]Supplemented by the *Georgi Perelman documentary* on YouTube (link in the references), the book Gessen (2009) about Perelman, and the article Jackson (2006) on the politics involved. Interestingly, Perelman not only refused the Fields Medal in 2006 and the Clay Millennium Prize of 1M$ in 2010, but in 1996 he had already declined a prestigious prize from the European Mathematical Society for outstanding young mathematicians.

2 Curvature in various guises and dimensions

If only to settle our notation.e start with a quick review of some material that should be familiar. In what follows, M is a connected and orientable manifold without boundary (as already agreed). Its tangent bundle is denoted by TM and the space of smooth sections of TM, i.e. the space of smooth vector fields on M, is called $\mathfrak{X}(M)$, so that $X \in \mathfrak{X}(M)$ is a map $x \mapsto X_x \in T_xM$. The space of smooth sections of the cotangent bundle T^*M is called $\mathbb{R}^n(M)$, and we define a ***tensor*** of type (k,l) as a map $T : \mathfrak{X}(M)^k \times \Omega(M)^l \to C^\infty(M)$ that is $C^\infty(M)$-linear in all entries:

$$T(f_1X_1,\dots,f_kX_k,g_1\theta^1,\dots,g_l\theta^l) = f_1\cdots f_k\cdot g_1\cdots g_l\cdot T(X_1,\dots,X_k;\theta^1,\dots,\theta^l). \tag{2.1}$$

2.1 Levi–Civita connection and Riemann tensor

A ***linear connection*** on M, or, equivalently, a ***covariant derivative*** on $\mathfrak{X}(M)$, associates to each vector field $X \in \mathfrak{X}(M)$ a linear map $\nabla_X : \mathfrak{X}(M) \to \mathfrak{X}(M)$, such that:

1. The map $X \mapsto \nabla_X$ is $\mathbb{R}$-linear as well as $C^\infty(M)$-linear, i.e.
$$\nabla_{fX}Y = f\nabla_XY \ \ (f \in C^\infty(M)); \tag{2.2}$$
2. The map $Y \mapsto \nabla_XY$ is $\mathbb{R}$-linear but not $C^\infty(M)$-linear: it satisfies the ***Leibniz rule***
$$\nabla_X(fY) = (Xf)Y + f\nabla_XY \ \ (f \in C^\infty(M)). \tag{2.3}$$

This definition also makes sense on any open $U \in \mathscr{O}(M)$, and in fact if $x \in U$, then $\nabla_XY(x)$ only depends on the value of X at x and the restriction of Y to U (this follows from (2.2) - (2.3) and the definition of a manifold). Hence we may compute covariant derivatives locally: for any local frame (e_a) for $\mathfrak{X}(M)$ on U (which consists of maps $e_a : U \to TM$, $a = 1,\dots,n$, such that at each $x \in U$ the vectors $e_a(x) \in T_xM$ form a basis of T_xM), the connection ∇ is then completely characterized by its ***connection coefficients*** ω^c_{ab}, defined (at each x) by

$$\nabla_{e_a}e_b = \omega^c_{ab}e_c. \tag{2.4}$$

Indeed, to compute ∇_XY from (2.4), we write $X = X^ae_a$ etc.,and from (2.2) - (2.3) obtain

$$\begin{aligned}\nabla_XY &= \nabla_{X^ae_a}(Y^be_b) = X^a\nabla_{e_a}(Y^be_b) = X^a(e_a(Y^b)\cdot e_b + Y^b\nabla_{e_a}e_b)\\ &= X^a(e_a(Y^c) + Y^b\omega^c_{ab})e_c.\end{aligned} \tag{2.5}$$

We write ∇_XY^a for $(\nabla_XY)^a$, so that $\nabla_XY = (\nabla_XY^a)e_a$. We therefore have

$$\nabla_XY^a = XY^a + \omega^a_{bc}X^bY^c, \tag{2.6}$$

where XY^a is the (defining) action of the vector field X on the function $Y^a \in C^\infty(U)$. In terms of the canonical coordinate basis $(e_i = \partial_i = \partial/\partial x^i)$, abbreviating $\nabla_i = \nabla_{\partial_i}$, we therefore have

$$\omega^k_{ij} = dx^k(\nabla_i\partial_j); \tag{2.7}$$

$$\nabla_XY^k = X^i(\partial_iY^k + \omega^k_{ij}Y^j); \tag{2.8}$$

$$\nabla_iY^k = \partial_iY^k + \omega^k_{ij}Y^j. \tag{2.9}$$

The ***torsion*** T_∇ of a given linear connection ∇ is a tensor of type $(2,1)$, defined by

$$T_\nabla(X,Y,\theta) = \theta(\nabla_XY - \nabla_YX - [X,Y]); \tag{2.10}$$

a simple computation shows that this expression is indeed $C^\infty(M)$-linear in each entry.

Theorem 2.1 (Levi-Civita) *Any Riemannian manifold (M, g) admits a unique linear connection ∇, called the* **Levi-Civita connection***, such that:*

1. *The torsion T_∇ associated to ∇ vanishes;*
2. *The connection ∇ and the metric g are related by*

$$X(g(Y,Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \ \ (X,Y,Z \in \mathfrak{X}(M)). \tag{2.11}$$

Proof: using torsion-freeness in the form $\nabla_X Y - \nabla_Y X = [X,Y]$, eq. (2.11) may be rewritten as

$$g(\nabla_X Y, Z) = \tfrac{1}{2}(Xg(Y,Z) + Yg(Z,X) - Zg(X,Y) - g(X,[Y,Z]) + g([X,Y],Z) + g(Y,[Z,X])), \tag{2.12}$$

often called the ***Koszul formula***. which shows both existence and uniqueness of ∇. □

The connection coefficients of ∇ are called ***Christoffel symbols***, which are usually denoted by Γ^k_{ij} $(= \omega^k_{ij})$. In a coordinate basis, where all commutators vanish, eq. (2.12) immediately gives

$$\Gamma^k_{ij} = \tfrac{1}{2} g^{kl}(\partial_j g_{li} + \partial_i g_{lj} - \partial_l g_{ij}). \tag{2.13}$$

The following map, *labeled* by $X, Y \in \mathfrak{X}(M)$, to be distinguished from its *argument* Z,

$$\Omega(X,Y) : \mathfrak{X}(M) \to \mathfrak{X}(M); \tag{2.14}$$

$$\Omega(X,Y)Z = ([\nabla_X, \nabla_Y] - \nabla_{[X,Y]})Z, \tag{2.15}$$

where $X, Y, Z \in \mathfrak{X}(M)$, and (2.15) is $C^\infty(M)$-linear in each of the three separately. Hence

$$\mathrm{Riem}(\theta, Z, X, Y) = \theta(\Omega(X,Y)Z) \tag{2.16}$$

defines a tensor of type $(3,1)$ called the ***Riemann (curvature) tensor***.[20] Equivalently,

$$\mathrm{Riem}(W, Z, X, Y) = g(W, (\Omega(X,Y)Z)). \tag{2.17}$$

makes Riem of type $(4,0)$. In coordinates, the Riemann tensor is (re)defined and expressed by

$$[\nabla_i, \nabla_j] Z^k = R^k_{lij} Z^l; \tag{2.18}$$

$$R^k_{lij} = \partial_i \Gamma^k_{lj} - \partial_j \Gamma^k_{li} + \Gamma^k_{im}\Gamma^m_{jl} - \Gamma^k_{jm}\Gamma^m_{il}; \tag{2.19}$$

$$R_{klij} = g_{km} R^m_{lij}, \tag{2.20}$$

where $R^k_{lij} \equiv \mathrm{Riem}^k_{lij}$. The ***Bianchi identities*** read, first abstractly and then in coordinates:[21]

$$\Omega(X,Y)Z + \Omega(Y,Z)X + \Omega(Z,X)Y = 0; \tag{2.21}$$

$$(\nabla_X \mathrm{Riem})(Y,Z) + (\nabla_Y \mathrm{Riem})(Z,X) + (\nabla_Z \mathrm{Riem})(X,Y) = 0; \tag{2.22}$$

$$R^k_{lij} + R^k_{ijl} + R^k_{jli} = 0; \tag{2.23}$$

$$\nabla_m R^k_{lij} + \nabla_j R^k_{lmi} + \nabla_i R^k_{ljm} = 0. \tag{2.24}$$

[20]Bernhard Georg Friedrich Riemann (1826–1866) was one of the greatest and most influential mathematicians in recent history. His *Habilitationsschrift* from 1854 entitled *Über die Hypothesen, welche der Geometrie zu Grunde liegen* is a blueprint for modern differential geometry, especially from a metric point of view. You can find it for example on `https://www.maths.tcd.ie/pub/HistMath/People/Riemann/Geom/Geom.pdf`.

[21]These identities follow from the property $\psi_t^*(\mathrm{Riem}(g)) = \mathrm{Riem}(\psi_t^* g)/dt$ under any one-paramter group of diffeomorphisms ψ_t by taking d/dt at $t = 0$, see Kazdan (1981). The contracted Bianchi's identities used in general relativity state this same property for the Ricci tensor.

2.2 Sectional curvature

Some more identities satisfied by R, which are crucial for what follows, are:

$$R_{klji} = -R_{klij}; \qquad R_{lkij} = -R_{klij}; \qquad R_{ijkl} = R_{klij}, \tag{2.25}$$

of which the first is trivial from (2.18), the second states that each map $\Omega(X,Y)$ is an isometry of T_xM, and the third is conceptually bizarre. In any case, these symmetries lead to a useful reinterpretation of the Riemann tensor. One may regard $\text{Riem}_x \in (T_x^*M)^{\otimes 4}$ as a linear map

$$\widetilde{\text{Riem}}_x : T_xM \otimes T_xM \to T_xM \otimes T_xM, \tag{2.26}$$

obtained by first noting that $\Omega_x(X,Y) \in \text{Hom}(T_xM, T_xM)$, second that for any finite-dimensional vector space V one has $\text{Hom}(V,V) \cong V^* \otimes V$ through the map $V^* \otimes V \to \text{Hom}(V,V)$ given by linear extension of $\theta \otimes v \mapsto (w \mapsto \theta(w)v)$, third that $V^* \cong V$ in the presence of an inner product on V, so that $\text{Hom}(V,V) \cong V \otimes V$. Explicitly, the map (2.26) is given by linear extension of

$$\partial_i \otimes \partial_j \mapsto g^{lm} R^k_{lij} \partial_k \partial_m, \tag{2.27}$$

so that by (2.25),[22] the map $\widetilde{\text{Riem}}_x$ in (2.26) - (2.26) restricts to a symmetric linear map

$$\widehat{\text{Riem}}_x : \wedge^2 T_xM \to \wedge^2 T_xM. \tag{2.28}$$

By (2.25), the map (2.26) is symmetric (i.e. self-adjoint) with respect to the inner product

$$\langle X_1 \otimes X_2, Y_1 \otimes Y_2 \rangle_x = g_x(X_1, Y_1) g_x(X_2, Y_2) \tag{2.29}$$

on $T_xM \otimes T_xM$, and hence $\widetilde{\text{Riem}}_x$ is entirely specified by the associated quadratic form[23]

$$\tilde{Q}_x : T_xM \otimes T_xM \mapsto \mathbb{R}; \tag{2.30}$$

$$(X,Y) \mapsto R(X,Y,X,Y), \tag{2.31}$$

Consequently, writing $X \wedge Y = \frac{1}{2}(X \otimes Y - Y \otimes X)$, also (2.28) is determined by a quadratic form

$$Q_x : \wedge^2 T_xM \to \mathbb{R}; \tag{2.32}$$

$$X \wedge Y \mapsto \langle X \wedge Y, \widehat{\text{Riem}}_x(X \wedge Y) \rangle_x = \text{Riem}_x(X,Y,X,Y). \tag{2.33}$$

It is easy to show that $X, Y \in T_xM$ are linearly independent iff $P_x(X,Y) \neq 0$, where

$$P_x(X \wedge Y) := g_x(X \wedge Y, X \wedge Y) = g_x(X,X) g_x(Y,Y) - g_x(X,Y)^2, \tag{2.34}$$

is the square of the (metric) area of the parallelogram in T_xM with sides X and Y, up to a sign. Assuming this is nonzero, the ***sectional curvature*** $C_x(X,Y)$ of the pair (X,Y) is defined by

$$C_x(X,Y) := \frac{Q_x(X \wedge Y)}{P_x(X \wedge Y)} = \frac{\text{Riem}_x(X,Y,X,Y)}{g_x(X,X) g_x(Y,Y) - g_x(X,Y)^2}. \tag{2.35}$$

The specific combination in (2.35) makes $C_x(X,Y)$ independent of the choice of X and Y within the plane (in T_xM) they span,[24] *and hence makes C_x a function of that plane only*. Thus we may as well use an orthonormal pair $X = e_a$, $Y = e_b$, so that at the end of the day we simply have

$$C_x(e_a, e_b) = \text{Riem}_x(e_a, e_b, e_a, e_b). \tag{2.36}$$

[22] Let V be a (real) vector space. Defining $\tau : V \otimes V \to V \otimes V$ by linear extension of $v \otimes w \mapsto w \otimes v$, the space $\wedge^2 V \equiv V \otimes_A V \subset V \otimes V$ is the antisymmetric part of $V \otimes V$, defined as the eigenspace of τ with eigenvalue -1.

[23] If $T : V \to V$ is linear and symmetric with respect to some inner product $\langle \cdot, \cdot \rangle$ on V, i.e., $\langle X, TY \rangle = \langle TX, Y \rangle$ for all $X, Y \in V$, then the associated quadratic form $Q : V \to \mathbb{R}$ is defined by $Q(X) = \langle X, TX \rangle$. The map T may be then be recovered from Q (and the inner product) via the polarization formula $\langle X, TY \rangle = \frac{1}{4}(Q(X+Y) - Q(X-Y))$.

[24] See Kobayashi & Nomizu, 1963, p. 200, Prop. 1.3.

2.3 Ricci tensor and Ricci scalar

The Riemann tensor contains al possible information about curvature (in Riemannian geometry, this is more or less true by definition). In particular, a key motivation for Riemann himself was:

Proposition 2.2 *A Riemannian manifold is flat iff* $\text{Riem} = 0$.

Here 'flat' means that the metric is Euclidean (i.e., $g = \text{diag}(1, \ldots, 1)$) in a suitable coordinate system (which does not imply that $M \cong \mathbb{R}^n$, think of the cylinder, which is flat). This proposition is almost trivial from left to right, but the converse implication is actually quite hard to prove.[25] Short of the Riemann tensor, there also exist weaker measures of curvature. The main actor in Ricci Flow is, of course, the ***Ricci tensor***,[26] which like the metric has type $(2,0)$; it is defined by

$$\text{Ric}(X,Y) = \sum_{a=1}^{n} \text{Riem}(e_a, X, e_a, Y), \tag{2.37}$$

where (e_a) is any orthonormal frame. Equivalently,[27] in coordinates its components are

$$\text{Ric}_{ij} \equiv R_{ij} = R^l_{ilj} = g^{kl} R_{kilj}. \tag{2.38}$$

Note that Ric is symmetric by (2.25). From Ric, we define the the ***scalar curvature*** by

$$R = \sum_{a=1}^{n} \text{Ric}(e_a, e_a) = \sum_{a,b=1}^{n} C(e_a, e_b) = g^{ij} R_{ij}, \tag{2.39}$$

where of course in the second sum the terms $a \neq b$ do not contribute and hence due to symmetry the sum just has $(n^2 - n)/2$ terms. For example, in $n = 3$ the Ricci scalar (at a point x) is the average of the sectional curvatures of the x-y, x-z, and y-z planes (within the tangent space T_xM).

Furthermore, the Ricci tensor defines two ***Einstein tensors***, most easily by their components

$$G_{ij} = R_{ij} - \tfrac{1}{2} g_{ij} R; \tag{2.40}$$

$$E_{ij} = R_{ij} - \frac{1}{n} g_{ij} R. \tag{2.41}$$

Physicists use G_{ij} because, as will be explained later, it emerges from the calculus of variations applied to the functional $g \mapsto \int_M R(g)$. Mathematicians, on the other hand, use E_{ij} because it is simply the traceless part of Ric (note that $g^{ij} E_{ij} = 0$). Moreover, to explain the name, suppose

$$\text{Ric} = \lambda g \tag{2.42}$$

for some constant $\lambda \in \mathbb{R}$, in which case we say that (M, g) is an ***Einstein manifold***, and that g is an ***Einstein metric***. Then $R = \lambda \cdot n$ is constant and $\lambda = R/n$, so that (2.42) implies $E_{ij} = 0$. In $d > 2$, also the converse is true;[28] prove this yourself from the Bianchi identity (2.24). Thus:

Proposition 2.3 *In $d > 2$, a metric satisfies* (2.42) *iff its Einstein tensor* (2.41) *vanishes.*

[25] For a relatively elementary proof (that avoids the Frobenius Theorem) see Heckman (2019), p. 46.

[26] J.R. Goodstein, *Einstein's Italian Mathematicians: Ricci, Levi-Civita, and the birth of General Relativity* (AMS, 2018) IS A joint biography of Gregorio Ricci-Curbastro (1853-1925) and Tullio Levi-Civita (1873-1941).

[27] Authors use various sign conventions for the Riemann tensor, but all Ricci tensors and scalars coincide.

[28] We will shortly see that $E_{ij} = 0$ in $d = 2$, where we know since Gauß that non-constant R is certainly possible.

The symmetries (2.25) enable one to count the number of independent components of the Riemann tensor in various dimensions n, namely $n^2(n^2-1)/12$ (check!). Therefore:

1. In $n=2$ the Riemann tensor has just one independent component R_{1212}, and also
$$g^{-1} = \begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix} = \frac{1}{\det(g)} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix}, \tag{2.43}$$
so that the Ricci tensor $R_{ij} = g^{kl}R_{kilj}$ must equal $R_{ij} = g_{ij}R_{1212}/\det(g)$, whence
$$R_{ij} = \tfrac{1}{2} g_{ij} R, \tag{2.44}$$
and $R_{1212} = \frac{1}{2}\det(g)\cdot R = \det(g)\cdot K$, where the Gaussian curvature K is given (either as a definition or as a theorem[29]) by $K = C(\partial_1, \partial_2) = R_{1212}/\det(g)$, cf. (2.35), so that $K = \frac{1}{2}R$.

2. In $n=3$ the Riemann tensor has 6 independent components, as does the Ricci tensor! So these two must carry the same information.[30] This can be understood as follows from linear algebra. If V has an inner product, any symmetric bilinear map $T: V\otimes V \to \mathbb{R}$ is equivalent to a self-adjoint linear map $\tilde{T}: V \to V$ via $T(v\otimes w) = \langle v, \tilde{T}w\rangle$. In particular the Ricci tensor $\mathrm{Ric}_x : T_xM \otimes T_xM \to \mathbb{R}$ at a point $x \in M$ is equivalent to a linear map
$$\widetilde{\mathrm{Ric}_x} : T_xM \to T_xM; \tag{2.45}$$
$$g_x(X, \widetilde{\mathrm{Ric}_x}Y) = \mathrm{Ric}_x(X,Y). \tag{2.46}$$
In $n=3$ (only!),[31] one has $\wedge^2 T_xM \cong T_x^*M \cong T_xM$, which isomorphism also makes linear maps $\wedge^2 T_xM \to \wedge^2 T_xM$ and $T_xM \to T_xM$ equivalent. As a case in point, if the Ricci tensor as in (2.45) is diagonalized by an orthonormal basis (e_1, e_2, e_3) of T_xM with eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$, then the Riemann tensor as in (2.28) is diagonal with respect to the basis $(e_1\wedge e_2, e_2 \wedge e_3, e_3 \wedge e_1)$ of $\wedge^2 T_xM$ with corresponding eigenvalues
$$(\lambda_1+\lambda_2-\lambda_3, \lambda_2+\lambda_3-\lambda_1, \lambda_1-\lambda_2+\lambda_3).$$
Trivially, the Ricci scalar is then given by the trace of either $\widetilde{\mathrm{Ric}_x}$ or $\hat{R}_x$, i.e.
$$R_x = \lambda_1+\lambda_2+\lambda_3. \tag{2.47}$$

3. In $n=4$ (the case of interest to physics) the Riemann tensor has 20 independent components, whereas the Ricci tensor only has 10. The geometric information in the Riemann tensor that is not passed on to the Ricci tensor is contained in the ***Weyl tensor***, defined by
$$W_{klij} = R_{klij} + (g_{k[j}R_{i]l} + g_{l[i}R_{j]k}) + \tfrac{1}{3}(R\cdot g_{k[i}g_{j]l}), \tag{2.48}$$
where $[\cdots]$ denotes *antisymmetrization* in the enclosed indices, much as $(\cdots)$ denotes *symmetrization* in the enclosed indices (for example, $g_{k[j}R_{i]l} = g_{kj}R_{il} - R_{ki}g_{jl}$). The symmetries of W, namely $W_{(kl)ji} = W_{lk(ij)} = 0$ and $W_{ijkl} = W_{klij}$, then also give W 10 independent components. According to Roger Penrose, W is the key to the Universe!

[29] See e.g. Heckman (2019), Theorem 3.14, p. 49.

[30] One way to see the equivalence of the Riemann and Ricci tensors is through the Einstein tensor (2.41): in terms of the notation $(P\odot Q)_{ijkl} := P_{il}Q_{jk} + P_{jk}Q_{il} - P_{ik}Q_{jl} - P_{jl}Q_{ik}$, one finds $\mathrm{Riem} = \frac{1}{4}R(g\odot g) + E\odot g$.

[31] If V is a $3d$ vector space with basis (e_1,e_2,e_3), then every element of $\wedge^3 V$ is proportional to $\omega = e_1\wedge e_2 \wedge e_3$ and hence $A \in \wedge^2 V$ defines $\hat{A}\in V^*$ by $A\wedge v = \hat{A}(v)\omega$. Since $\wedge^2 V$ has a basis $(e_1\wedge e_2, e_2\wedge e_3, e_3\wedge e_1)$, counting dimensions shows that $A\mapsto \hat{A}$ gives a bijection $\wedge^2 V \cong V^*$. If ε_{ijk} is the totally antisymmetric (Levi-Civita) symbol, $v = v^i e_i \in V$, $A = A^{ij}e_i\wedge e_j \in \wedge^2 V$, and $\hat{A} = \hat{A}_i\theta^i \in \wedge^2 V$ with $\theta^i(e_j) = \delta^i_j$, then $\hat{A}_i = \varepsilon_{ijk}A^{jk}$, or $\hat{A}_1 = A^{23}$, $\hat{A}_2 = -A^{13}$, and $\hat{A}_3 = A^{12}$. Dually, $V \cong (\wedge^2 V)^* = \wedge^2 V^*$ under $v \mapsto \hat{v}$, where $\hat{v}_{ij} = \varepsilon_{ijk}v^k$ (this applies e.g. to the magnetic field).

3 Spaces of constant curvature

We now turn the the Riemannian manifolds that are of greatest interest to Ricci Flow:[32]

Definition 3.1 *We say that a Riemannian manifold (M,g) has* **constant curvature** *if all sectional curvatures $C_x(X,Y)$ coincide (where $x \in M$ and $X,Y \in T_xM$).*

In $d = 2$ this simply means that the scalar curvature is constant, and the condition becomes increasingly stringent in higher dimension, as T_xM contains an increasing number of plane whose sectional curvatures is to be constant. This exactly compensates for the larger variety of possible Riemannian manifolds in high dimension, since Theorem 1.3 gives exactly the same classification in $d = 2$ and $d > 2$. An interesting difference between these cases, however, is:[33]

Proposition 3.2 *Let $\dim(M) \geq 3$. If for all $x \in M$ the sectional curvatures $C_x(X,Y)$ are independent of X and Y, then they are also independent of x, i.e. (M,g) has constant curvature.*

We sketch the proof and leave it as an exercise to fill in the details. Let $C_x(X,Y) = k(x)$ for all $X,Y \in T_xM$ and some $k \in C^\infty(M)$. In terms of the auxiliary tensor S of type $(4,0)$ defined by

$$S_x(V,W,X,Y) = g_x(V,X)g_x(W,Y) - g_x(V,Y)g_x(W,X), \tag{3.1}$$

eq. (2.35) then gives $\text{Riem}_x(X,Y,X,Y) = k(x)S(X,Y,X,Y)$, but since the Riemann tensor is completely defined by its sectional curvatures, this implies $\text{Riem}_x = k(x)S$. Taking the covariant derivative with respect to an arbitrary vector-field $U \in \mathfrak{X}(M)$ gives $\nabla_U \text{Riem} = (Uk) \cdot S$, since $\nabla_U S = 0$ by definition of the Levi-Civita connection (which gives $\nabla_U g = 0$). Eq. (2.22) gives

$$\begin{aligned}(Uk)\cdot(g(Z,Y)X - g(Z,X)Y) + (Xk)\cdot(g(Z,U)Y - g(Z,Y)U)&\\ + (Yk)\cdot(g(Z,X)U - g(Z,U)X) = 0,&\end{aligned} \tag{3.2}$$

for arbitrary $X,Y,Z \in \mathfrak{X}(M)$. Now in $d \geq 3$ we may take $Z = U$ to be unit vectors and (X,Y,Z) mutually perpendicular, so that (3.2) yields $(Xk)\cdot Y - (Yk)\cdot X = 0$. Since this is true for all $X \perp Y$, it follows that $Xk = Yk = 0$ and hence k is constant. □

This proof gives a useful formula for the Riemann tensor in case of constant curvature, viz.

$$R_{ijkl} = k(g_{ik}g_{jl} - g_{il}g_{jk}), \tag{3.3}$$

where k is the common value of all sectional curvatures, called *the* curvature of (M,g). It is easy to show from (3.3) that a constant curvature metric must be Einstein (exercise). We have already seen that an Einstein metric has constant *scalar* curvature. In $d = 3$, this can be boosted to constant *sectional* curvature (exercise), so that in our dimension of interest, we have:

Proposition 3.3 *In $d = 3$, a metric has constant curvature iff it is an Einstein metric.*

We now turn to the classification of (complete) Riemannian manifolds with constant curvature. This given by Theorem 1.3 in the simply connected case, from which the general case follows analogously to Corollary 1.4, that is, by taking quotients with respect to certain discrete isometry groups.[34] We start with an elementary description of the case $n = 2$, which will not be used in the sequel, but is useful for developing intuition and for its historical value. This is followed by a restatement Theorem 1.3 in terms of group theory, on the basis of which it will also be proved.

[32] Those with ***constant scalar curvature***, in which the Ricci scalar R is constant, are of less interest.

[33] Recall that we assume that M is connected.

[34] We will treat some easy cases; the general case is a subject by itself. See Vinberg (1993) and Wolf (2010).

- Let S^2_ρ be the two-sphere with radius $\rho > 0$, with metric induced from the flat one on $\mathbb{R}^3$,

$$S^2_\rho = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 = \rho^2\}. \tag{3.4}$$

Riemann's point of view was to use the stereographic projection from the north pole onto the $z=0$ plane in $\mathbb{R}^3$, which is given by the formulae

$$(x,y,z) \mapsto (u,v,0); \qquad u = \frac{\rho x}{\rho - z}; \qquad v = \frac{\rho y}{\rho - z}, \tag{3.5}$$

where $(x,y,z) \neq (0,0,\rho) \equiv N$ (which point is sent to infinity) so that we obtain a diffeomorphism from $S^2_\rho \backslash N$ to $\mathbb{R}^2$. A straightforward computation then gives the metric

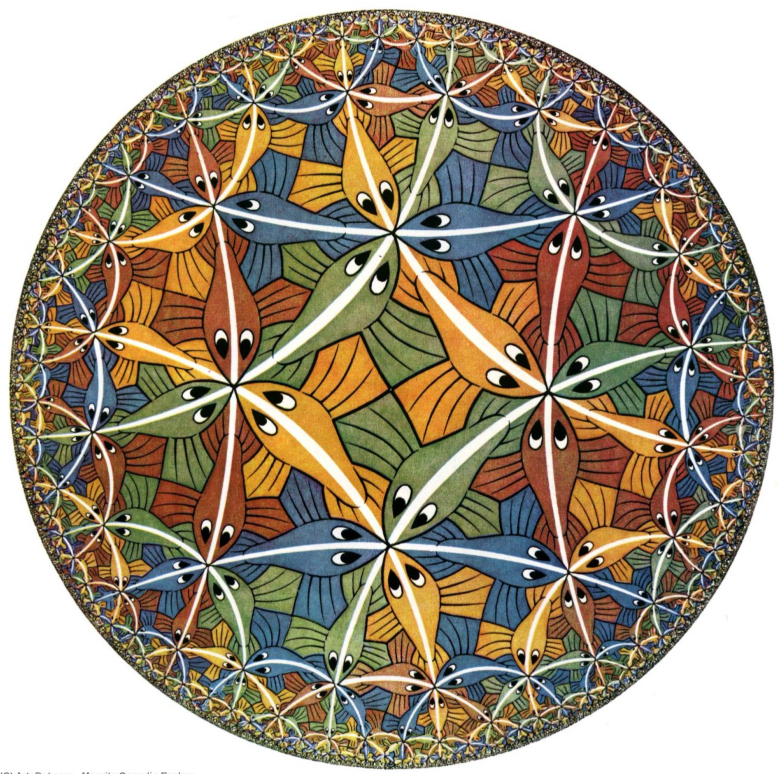
$$ds^2 = 4 \cdot \frac{du^2+dv^2}{1+k(u^2+v^2)}, \tag{3.6}$$

where $k = 1/\rho^2 > 0$ is the value of the constant curvature that appears in (3.3).

- Riemann then identified the metric carrying hyperbolic geometry as the open disc

$$D^2_\rho = \{(u,v) \in \mathbb{R}^2 \mid u^2+v^2 < \rho\}, \tag{3.7}$$

equipped with formally the same metric (3.6), but this time with $k = -1/\rho^2 < 0$.



Circle Limit III by M.C. Escher, showing geodesics of the hyperbolic metric on the disc

In 1906, Hilbert proved that it is impossible to isometrically embed D^2_ρ with its hyperbolic metric in $\mathbb{R}^3$, equipped with its usual (Euclidean) metric $g = \mathrm{diag}(-1,1,1)$. However, an embedded model of hyperbolic geometry is given by the relativistic ***mass shell*** in $\mathbb{R}^3$,

$$M^2_\rho = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2-z^2 = -\rho^2, z>0\}, \tag{3.8}$$

where $\mathbb{R}^3$ *is equipped with the Minkowski metric* $\eta = \mathrm{diag}(-1,1,1)$.

3.1 Homogeneous manifolds

These models have group-theoretical descriptions, which play a fundamental role in their classification. In preparation, we quote the following basic technical result without proof:[35]

Proposition 3.4 *Let G be a Lie group and $K \subset G$ a closed subgroup.*[36] *Then K is itself a Lie group and there exists a smooth structure on the* **homogeneous space** *G/K such that:*

1. $\dim(G/K) = \dim(G) = \dim(K)$*;*
2. *The canonical projection $G \to G/K$, $\gamma \mapsto \gamma K$, is smooth;*
3. *The canonical G-action $G \times (G/K) \to G/K$, $(\gamma_1, \gamma_2 K) \mapsto (\gamma_1\gamma_2)K$, is smooth.*

We write such group actions as $\gamma_1(\gamma_2 K) = (\gamma_1\gamma_2)K$. It is clear that G acts transitively on G/K (for any $x \in G/K$ and $y \in G/K$ there is $\gamma \in G$ such that $y = \gamma x$). Without loss of generality,[37] we may also assume that G acts effectively on G/K (if $\gamma x = x$ for all $x \in G/K$, then $x = e$).

The following isomorphism will be very useful in the proof of Theorem 1.3:

$$T_K(G/K) \cong \mathfrak{g}/\mathfrak{k}, \tag{3.9}$$

where $\mathfrak{g}$ and $\mathfrak{k}$ are the Lie algebras of the Lie groups G and K, respectively. To see this, let us consider a more general situation, where a Lie group G acts smoothly on a manifold M, that is, $\varphi : G \times M \to M$ is a smooth G-action on M. We will write $\varphi_\gamma(x)$ (or simply $\gamma \cdot x$) for $\varphi(\gamma, x)$, so that each map $\varphi_\gamma : M \to M$ is a diffeomorphism. For each $A \in \mathfrak{g}$ we define a map

$$A_M : C^\infty(M) \to C^\infty(M); \tag{3.10}$$

$$A_M f(x) = \frac{d}{dt} f(e^{tA} \cdot x)_{|t=0}. \tag{3.11}$$

This defines a derivation on $C^\infty(M)$ and hence a vector field on M, so that $A_M \in \mathfrak{X}(M)$, and we have a map $A \mapsto A_M$ from $\mathfrak{g}$ to $\mathfrak{X}(M)$. It can be shown that our map has good properties:[38]

Proposition 3.5 *The map $A \mapsto A_M$ is linear and for all $A, B \in \mathfrak{g}$ satisfies*

$$[A_M, B_M] = -[A, B]_M. \tag{3.12}$$

In other words, our map is an anti-homomorphisms of Lie algebras (with respect to the usual commutator bracket of vector fields). Clearly, at any $x \in M$ we obtain a map $\mathfrak{g} \to T_x M$ by regarding $A_M(x)$ as an element of $T_x M$. In the case at hand, where $M = G/K$, we take $x = K$ and note that the map $\mathfrak{g} \mapsto T_K(G/K)$ has kernel $\mathfrak{k}$; indeed, if $A \in \mathfrak{k}$, then $\exp(tA) \in K$ by definition of a Lie algebra (see Appendix B), but $kK = K$ for any $k \in K$, whence the right-hand side of (3.11) vanishes. This proves that $\mathfrak{k}$ lies in the kernel. Conversely, $\gamma K = K$ iff $\gamma \in K$, and $k \in K$ lies near the identity iff $k = \exp(tA)$ for some $A \in \mathfrak{k}$. Surjectivity of $\mathfrak{g} \mapsto T_K(G/K)$ follows from a dimension count based on Proposition 3.4.1, so that finally (3.9) follows from linear algebra.

[35] See e.g. Goodman & Wallach (2009),Theorem D.2.10, or Kobayashi & Nomizu (1963), Proposition 4.2.

[36] Anticipating the fact that K will be compact in our applications, we use K instead of the traditional letter H.

[37] If G does not act effectively on G/K, take the largest normal subgroup $K_0 \subset K$ that is also normal in G, and define $G^* = G/K_0$ and $K^* = K/K_0$. Then $G/K \cong G^*/K^*$ and G^* acts effectively on G^*/K^*. An example where this is necessary occurs if $K \subset Z(G)$, in which case all of K acts trivially on G/K. Although the isometry group of a Riemannian manifold always acts effectively on M, we still need some care, as in the case S^3 in §3.4.

[38] See e.g. J.E. Marsden & T.S. Ratiu, *Introduction to Mechanics and Symmetry*, Second edition (Springer, 1999).

The isomorphism (3.9) gets more body of we combine it with the residual K-action on $T_K(G/K)$. For any diffeomorphism φ of a manifold M, the derivative φ'_x maps T_xM linearly to $T_{\varphi(x)}M$ (and the pullback φ^*_x linearly maps $T^*_{\varphi(x)}M$ to T^*_xM).[39] If $\varphi(x)=x$, then $\varphi'_x \in \mathrm{Hom}(T_xM)$. If the diffeomorphisms φ come from a G-action on M, we may define the ***stabilizer*** of x by

$$G_x = \{\gamma \in G \mid \gamma \cdot x = x\}. \tag{3.13}$$

The maps $\varphi'_\gamma : T_xM \to T_xM$, where $\gamma \in G_x$, then combine into a representation

$$\pi_x : G_x \to GL(T_xM); \tag{3.14}$$

$$\gamma \mapsto \varphi'_\gamma, \tag{3.15}$$

called the ***isotropy representation*** of G_x in T_xM (here $GL(T_xM)$ consists of all invertible linear maps from T_xM to T_xM). This applies in particular to $M = G/K$ and $x = K$, so that we obtain

$$\pi_K : K \to GL(T_K(G/K)); \tag{3.16}$$

$$k \mapsto \varphi'_k. \tag{3.17}$$

We will now explicitly find π_K under the isomorphism (3.9). We know that any group G acts on itself by the ***adjoint action*** $\mathrm{Ad}_\gamma(\delta) = \gamma\delta\gamma^{-1}$. If G is a Lie group,[40] this action defines a representation Ad' of G on its Lie algebra $\mathfrak{g}$, defined by $\mathrm{Ad}'_\gamma(X) = \gamma X \gamma^{-1}$. This action may, of course, be restricted to $K \subset G$, and it is easy to see that this restriction quotients to $\mathfrak{g}/\mathfrak{k}$. In our application to spaces with constant curvature, $\mathfrak{g}$ will have a canonical decomposition

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}, \tag{3.18}$$

where (trivially) not only $\mathfrak{k}$, but also $\mathfrak{p}$ is invariant under Ad'_k for any $k \in K$ (if K is connected, this is equivalent to $[\mathfrak{k}, \mathfrak{p}] \subseteq \mathfrak{p}$). In that case, we may evidently replace the isomorphism (3.9) by

$$T_K(G/K) \cong \mathfrak{p}, \tag{3.19}$$

Proposition 3.6 1. *Under the isomorphism* (3.9)*, the isotropy representation* (3.16) - (3.17) *of K on $T_K(G/K)$ is mapped to the adjoint action of K on $\mathfrak{g}/\mathfrak{k}$ (still denoted by* Ad'*), i.e.,*

$$\pi_K(k)[A] = [\mathrm{Ad}'_k(A)], \tag{3.20}$$

where $A \in \mathfrak{g}$ and $[A] \in \mathfrak{g}/\mathfrak{k}$, seen as an element of $T_K(G/K)$ via the isomorphism (3.9).

2. *Consequently, under the isomorphism* (3.19)*, assuming that $\mathfrak{p}$ is* $\mathrm{Ad}'(K)$*-invariant, the same isotropy representation of K is mapped to the adjoint action of K on $\mathfrak{p}$.*

Indeed, for any $A \in \mathfrak{g}$, $k \in K$, and $f \in C^\infty(G/K)$ we have, cf. (3.11) and (3.17),

$$(\pi_K(k)A_{G/K})f(K) = \frac{d}{dt} f(ke^{tA} \cdot K)_{|t=0} = \frac{d}{dt} f(ke^{tA}k^{-1} \cdot K)_{|t=0}$$
$$= \frac{d}{dt} f(e^{tkAk^{-1}} \cdot K)_{|t=0} = (\mathrm{Ad}'_k A)_{G/K} f(K). \qquad \square$$

---

[39] If we regard $X_x \in T_xM$ as a point derivation, i.e., as a map $X_x : C^\infty(M) \to \mathbb{R}$ that satisfies $X_x(fg) = X_x(f)g(x) + f(x)X_x(g)$, then $(\varphi'_x X_x) f(\varphi(x)) = X_x(f \circ \varphi)(x)$, and dually $(\varphi^*_x \theta_{\varphi(x)})(X_x) = \theta_{\varphi(x)}(\varphi'_x X)$ for $\theta_{\varphi(x)} \in T^*_{\varphi(x)}M$.

[40] It follows from our definition of a Lie algebra in Appendix B that Ad' is well defined as well as linear.

Homogeneous spaces arise if a Lie group G acts smoothly and transitively on a manifold M (in which case M is called a ***homogeneous G-space***). Then $M \cong G/K$ with $K = G_{x_0}$ (i.e. the stability group of some fixed $x_0 \in M$), under the diffeomorphism $M \to G/K$, $x \mapsto \gamma K$, where $\gamma \in G$ satisfies $\gamma x_0 = x$; the inverse map $G/K \to M$ is $\gamma K \mapsto \gamma x_0$ (both maps are independent of the choice of $\gamma \in \gamma K$), and this identification $M \leftrightarrow G/K$ is G-equivariant. Thus our trio is:

$$S^2 \cong SO(3)/SO(2); \tag{3.21}$$

$$\mathbb{R}^2 \cong E(2)_0/SO(2); \tag{3.22}$$

$$M^2 \cong SO(2,1)_0/SO(2), \tag{3.23}$$

where the Lie groups in question are defined in Appendix A and we denote the component of G that contains the identity by G_0. For simplicity we put $S^2 = S^2_1$ and $M^2 = M^2_1$; the story for general $\rho > 0$ is similar. To verify (3.21), let $SO(3)$ act on S^2 by restricting its defining action on $\mathbb{R}^3$, and take $x_0 \in S^2$ to be the north pole, in which case the $SO(2)$ in (3.21) consists of rotations around the z-axis. For (3.22), let $E(2)_0$ act on $\mathbb{R}^2$ in the defining representation and take $x_0 = (0,0)$, so that the $SO(2)$ in (3.22) is rotations of the plane. Finally, (3.23) is obtained from $x_0 = (0,0,1)$ in (3.8), so that also here the $SO(2)$ consists of rotations around the z-axis.

Writing (3.21) - (3.23) generically as $\mathbb{V} \cong G/K$, where $K = SO(2)$ is the same in all cases, the $\mathrm{Ad}'(K)$-invariant decomposition (3.18) applies to each G in the list. In all three cases we have $\mathfrak{g} \cong \mathbb{R}^3$, $\mathfrak{k} \cong \mathbb{R}$, and $\mathfrak{p} \cong \mathbb{R}^2$, and it follows from (B.8), (B.19), and (B.10), respectively, that

the $SO(2)$-action on $\mathfrak{p}$ is just its defining action on $\mathbb{R}^2$.

It is worth spelling this out in some detail, explaining a few useful techniques along the way. Let $u : G \to GL(V)$ be a representation of a Lie group G on a *finite-dimensional* vector space V. Then the following is unproblematic: for $A \in \mathfrak{g}$ we define a (linear) map $du(A) : V \to V$ by

$$du(A)v = \frac{d}{dt} u\left(e^{tA}\right) v_{|t=0}. \tag{3.24}$$

This construction gives a linear map $du : \mathfrak{g} \to \mathrm{Hom}(V)$, which satisfies

$$[du(A), du(B)] = du([A,B]); \tag{3.25}$$

$$e^{du(A)} = u\left(e^A\right). \tag{3.26}$$

In particular, if G is connected, then u can be recovered from du via (3.26).[41]

For example, the adjoint representation $\mathrm{Ad}' : G \to GL(\mathfrak{g})$ defines a Lie algebra homomorphism $\mathrm{ad} : \mathfrak{g} \to \mathrm{Hom}(\mathfrak{g})$,[42] where $\mathrm{ad} \equiv dAd'$, given by

$$\mathrm{ad}(A)B = [A,B]. \tag{3.27}$$

For $G = SO(3)$, the commutation relations (B.8) show that $\mathrm{ad}(e_3)e_1 = e_2$ and $\mathrm{ad}(e_3)e_2 = -e_1$, where e_3 is the generator of the subgroup $SO(2)$ of $SO(3)$ that consists of rotations around the z-axis. This means that as a matrix relative to the basis (e_1, e_2) of $\mathbb{R}^2$, the restriction of the linear map $\mathrm{ad}(e_3) : \mathfrak{so}(3) \to \mathfrak{so}(3)$ to $\mathfrak{p} = \mathrm{span}(e_1, e_2) \cong \mathbb{R}^2$ (which restriction is well defined, as the above relations show) is just the usual generator of $\mathfrak{so}(2)$, see (B.18), which is obtained from the defining action id of $G = O(2)$ on $V = \mathbb{R}^2$ by the procedure (3.24). By exponentiation, we then conclude that the corresponding Ad-action of $SO(2)$ on $\mathfrak{p}$ is the defining action, too. It follows from (B.19) and (B.10) that the same is true for the cases (3.22) and (3.23), respectively.

[41] If G is simply connected, this gives an equivalence between finite-dimensional Lie group and Lie algebra representations. This fails for infinite-dimensional representations, where (3.24) leads to unbounded operators.

[42] Each map $\mathrm{ad}(A)$ is even a derivation of $\mathfrak{g}$ as a Lie algebra, as follows from the Jacobi identity.

3.2 Homogeneous Riemannian manifolds and symmetric spaces

So far, this was an exercise in Lie group theory and differential geometry; we now bring in a metric. The relationship between homogeneous spaces and Riemannian geometry is twofold:

1. Given $M = G/K$, one may study possible G-invariant Riemannian metrics g on M.
2. Given (M,g), one may find out if the isometry group of g possibly acts transitively on M.

In general, a metric g on M is *invariant* under a diffeomorphism φ of M if

$$\varphi^* g = g \qquad \Leftrightarrow \qquad g_{\varphi(x)}(\varphi'_x(X), \varphi'_x(Y)) = g_x(X,Y) \;\; \forall x \in M, X, Y \in T_x M. \tag{3.28}$$

The set of all such diffeomorphisms φ is the ***isometry group*** of (M,g), denoted by $\mathrm{Iso}(M,g)$. If M is a G-space, we say that g is G-invariant if $\varphi_\gamma^* g = g$ for all $\gamma \in G$. If this is the case, then $G \subseteq \mathrm{Iso}(M,g)$ by definition (typically without equality),[43] If in addition G acts transitively on M, we say that (M,g) is a ***homogeneous Riemannian manifold***, so that $M \cong G/K$.

We return to the second point in the next section. The first is settled as follows:[44]

Proposition 3.7 *1. There is a bijective correspondence between G-invariant metrics on G/K and $\mathrm{Ad}'(K)$-invariant inner products on $\mathfrak{g}/\mathfrak{k}$, and hence, if* (3.18) *applies, on* $\mathfrak{p}$.

2. There is a unique G-invariant metric on G/K (up to scaling by a positive constant) iff the $\mathrm{Ad}'(K)$-action on $\mathfrak{g}/\mathfrak{k}$ (or, if applicable, $\mathfrak{p}$) is irreducible.

Proof. To prove the first claim, just use (3.9) or (3.19): any inner product on $\mathfrak{g}/\mathfrak{k}$ or $\mathfrak{p}$ defines a metric g on $T_K(G/K)$, which the G-action then pushes to any other point. Invariance under G clearly requires $\varphi_k^* g_K = g_K$ for any $k \in K$, so that Proposition 3.6 shows that $\mathrm{Ad}'(K)$-invariance of the inner product is necessary. It is a simple exercise to show that it is also sufficient.

The second claim follows from two facts. First, any inner product on a vector space V can be written in terms of some fiducial inner product $\langle \cdot, \cdot, \rangle$ and a positive operator $A \in \mathrm{Hom}(V)$ (where positive means: $\langle v, Av \rangle \geq 0$ for all $v \in V$). This follows by regarding an inner product as a special quadratic form. Second, *Schur's Lemma* from group theory states that the $\mathrm{Ad}'(K)$-action on $\mathfrak{g}/\mathfrak{k}$ (or $\mathfrak{p}$) is irreducible iff any operator commuting with all operators Ad'_k ($k \in K$) is necessarily a multiple of the unit operator id_V. So this leaves $A = \lambda \cdot \mathrm{id}_V$, for some $\lambda > 0$. □

Proposition 3.8 *For any Riemannian manifold (M,g) and $G \subseteq \mathrm{Iso}(M,g)$, the isotropy representation $\pi_x(G_x)$ defined in* (3.14) - (3.15) *is injective and has closed range.*

If M is complete, any isometry φ of M is determined by its tangent map φ'_x at some fixed $x \in M$: to find $\varphi(y)$, by Hopf–Rinow there is a geodesic γ from x to y, so that $y = \exp_x(Y)$ for some $Y \in T_x M$, and if φ is an isometry, then $\varphi(\exp_x(Y)) = \exp_x(\varphi'_x(Y))$.[45] Injectivity of π_x then immediately follows from (3.15). For closedness see Cheeger & Ebin, 1975, p. 61 and 37.

[43]If M is orientable, as we assume, then $\mathrm{Iso}(M,g)$ has at least two components (and for spaces of constant curvature it has exactly two): one is the identity component $\mathrm{Iso}(M,g)_0$, and the other is obtained from this by composition with an orientation-preserving isometry. For example, $\mathrm{Iso}(S^2,g)$ (with round metric g) is $O(3)$, with $O(3)_0 = SO(3)$ and an orientation-preserving isometry is given by $\mathrm{diag}(-1,-1,-1)$. Thus $S^2 \cong SO(3)/SO(2) \cong O(3)/O(2)$. Next, $\mathrm{Iso}(\mathbb{R}^2,g)$ (with flat metric g) is $E(2)$ and $\mathbb{R}^2 \cong E(2)/O(2) \cong E(2)_0/SO(2)$. In the third case, $\mathrm{Iso}(M^2,g)$ (with hyperbolic metric g) consists of the group $O(2,1)^+$ of all elements of $O(2,1)$ with $\lambda_{00} > 0$, see (A.5), and then $\mathrm{Iso}(M^2,g)_0 = SO(2,1)^+ = SO(2,1)_0$, so that $M^2 \cong O(2,1)^+/O(2) \cong SO(2,1)^+/SO(2)$.

[44]See e.g. Prop. 3.16 in Cheeger & Ebin (1975), or corresponding results in Vinberg (1993), Part I, Chapter 1.

[45]Here $\exp_x(X) = \gamma_X(1)$, where γ_X is the unique geodesic satisfying $\gamma_X(0) = x$ and $\dot{\gamma}_X(0) = X$.

Corollary 3.9 *If G/K carries a G-invariant Riemannian metric, then K is compact.*

Proof. If $O(V)$ is the subgroup of $GL(V)$ that preserves a given inner product on V, then any orthonormal basis of V will produce an isomorphism $V \cong \mathbb{R}^n$ and hence $O(V) \cong O(n)$, which is compact. Thus $O(\mathfrak{g}/\mathfrak{k})$ and $O(\mathfrak{p})$ are compact (cf. Proposition 3.7). The previous proposition makes K isomorphic to a closed subgroup of $O(\mathfrak{g}/\mathfrak{k})$ or $O(\mathfrak{p})$, so that K is compact. □

It is clear from these results that, up to scaling, our three spaces (3.21) - (3.23) have unique G-invariant metrics (where $G = SO(3), E(2)_0$, and $SO(2,1)_0$).[46] This uniqueness is true for all spaces of constant curvature, as should be clear from the following restatement of Theorem 1.3:

Theorem 3.10 *If $n = \dim(M) \geq 2$, M is simply connected, and g is complete and has constant curvature, then up to a rescaling of the metric by a positive constant, (M,g) is isometrically isomorphic to G/K with its unique G-invariant metric, where $K = SO(n)$ and G is one of:*

- $G = SO(n+1)$, *so that* $M \cong SO(n+1)/SO(n) \cong S^n$ *with round metric,*[47] *and* $k = 1$;
- $G = E(n)_0 = SO(n) \ltimes \mathbb{R}^n$, *whence* $M \cong E(n)_0/SO(n) \cong \mathbb{R}^n$ *with flat metric, and* $k = 0$;
- $G = SO(n,1)$, *where* $M \cong SO(n,1)/SO(n) \cong M^n$ *with hyperbolic metric,*[48] *and* $k = -1$.

Our proof of this theorem is based on the concept of a *symmetric (Riemannian) space*.[49]

Definition 3.11 1. *A Riemannian manifold (M,g) is* **locally symmetric** *if each $x \in M$ has a normal neighbourhood U_x and an isometry $l_x : U_x \to U_x$ with the following properties:*[50]

$$l_x(x) = x \qquad\qquad (l_x)'_x = -\mathrm{id}_{T_xM}, \tag{3.29}$$

or, equivalently:[51]

$$l_x(\exp_x(X)) = \exp_x(-X), \;\; X \in \exp_x^{-1}(U_x) \subset T_xM. \tag{3.30}$$

2. *A Riemannian manifold (M,g) is* **symmetric** *if the above holds for $U_x = M$ ($\forall x \in M$).*

Eq. (3.30), and hence also (3.29), gives $l_x^2 = \mathrm{id}_{U_x}$. Eq. (3.30) easily implies (3.29), and the converse implication follows from the fact that, as already mentioned, a (local) isometry φ is determined by its tangent φ'_x at an arbitrary point $x \in M$ (for which, in case that $\varphi = l_x$, we take x). Furthermore, as an example of the idea that local constructions in Riemannian geometry can often be made global if (M,g) is complete and simply connected,[52] we mention that if (M,g) has these properties (as it does in Theorem 3.10) and is locally symmetric, then it is symmetric.

[46]In fact these metrics are invariant under the full (disconnected) isometry groups $O(3)$, $E(2)$, and $SO(2,1)$.

[47]This is the metric induced by the embedding $S^n \subset \mathbb{R}^{n+1}$ with flat metric.

[48]Here $M^n = \{x \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots + x_n^2 - x_0^2 = -1, x_0 > 0\}$, where (as in physics) we label $x = (x_0, x_1, \ldots, x_n)$, with metric induced by the embedding $M^n \subset \mathbb{R}^{n+1}$ with Minkowski metric $\mathrm{diag}(-1,1,\ldots,1)$. This space may also be realized as the open unit ball in $\mathbb{R}^n$ or as the upper half plane in $\mathbb{R}^n$, but in those realizations the metric looks more artificial (at least to the author). See e.g. Vinberg (1993) for an extensive treatment of hyperbolic geometry.

[49]A full treatise is Helgason (1978). See also Kobayashi & Nomizu (1969), Ch. IX, and Joos (2002), Ch. 5.

[50]A nbhd U_x of x is ***normal*** if there exists a star-shaped nbhd $\mathscr{U}_x$ of $0 \in T_xM$ such that $\exp_x : \mathscr{U}_x \to U_x$ is a diffeomorphism (where a subset $V \subset W$ of a vector space is called ***star-shaped*** if $v \in V$ implies $tv \in V$ for all $t \in [-1,1]$). In that case, the tangent vector $X \in T_xM$ in (3.30) should of course be restricted to $\mathscr{U}_x$.

[51]Therefore, l_x is often called a ***geodesic reflection***, generalizing reflection symmetry in Euclidean geometry.

[52]This ideology is the subject of Chapter VI of Kobayaski & Nomizu, 1969; see especially Corollary 7.9.

3.3 Proof of the classification theorem

Our proof of Theorem 3.10 consists of three steps, of which we state the first two as a lemma:

Lemma 3.12 *1. If (M,g) has constant curvature, then it is locally symmetric. Consequently, if (M,g) is simply connected, complete, and has constant curvature, then it is symmetric.*

2. If (M,g) is symmetric, then it is a homogeneous Riemannian manifold.

Therefore, among Riemannian manifolds we have the following implications:

constant curvature (= homogeneous and isotropic) $\Rightarrow$ symmetric $\Rightarrow$ homogeneous.

This lemma reduces the classification problem to a straightforward problem in Lie groups and Lie algebras; reinserting the constant curvature condition then leads to the list in Theorem 3.10.

We first note that the proof of Proposition 3.2 shows that if (M,g) has constant curvature, then $\nabla\text{Riem} = 0$. Lemma 3.12.1 then follows from a key fact about locally symmetric spaces:

Lemma 3.13 *A space (M,g) is locally symmetric iff $\nabla\text{Riem} = 0$.*

The easy implication is "$\Rightarrow$", which is a simple exercise. Unfortunately, we need the converse implication, whose proof rests on the following technique (going back to Cartan) for extending pointwise isometries to local isometries. Suppose $x,y \in M$ and $F : T_xM \to T_yM$ is a linear isomorphism. Let U_x and U_y be normal nbhds of x and y, respectively, so that we obtain a map

$$f : U_x \to U_y; \tag{3.31}$$

$$f = \exp_y \circ F \circ \exp_x^{-1}. \tag{3.32}$$

The point (for us) is that if F preserves both the metric (i.e. $g_y(F(X),F(Y)) = g_x(X,Y)$ for all $X,Y \in T_xM$) and the Riemann tensor (i.e., $\text{Riem}_y(F(W),F(Z),F(X),F(Y)) = \text{Riem}_x(W,Z,X,Y)$ for all $X,Y,W,Z \in T_xM$), and in addition $\nabla\text{Riem} = 0$, then f is an isometry.[53] Now take $x = y$ and $F = -\text{id}_{T_xM}$. This clearly satisfies the assumptions, simply because both g and Riem have even rank (namely 2 and 4, respectively). The ensuing map f is our local isometry l_x. □

Part 2 of Lemma 3.12 has two components. One is purely technical and stated without proof:[54]

Proposition 3.14 *The isometry group* $\text{Iso}(M,g)$ *of any Riemannian manifold is a (finite-dimensional) Lie group in the compact-open topology.*

Proposition 3.15 *The isometry group* $\text{Iso}(M,g)$ *of a symmetric space acts transitively on M.*

This relies on the fact that a symmetric space is complete; we leave this as an exercise, since in Theorem 3.10 we assume completeness. The Hopf–Rinow Theorem then states, among other things, that any two points y,z of M may be connected by a geodesic γ. So let $y = \gamma(0)$ and $z = \gamma(T)$. Then $y = l_x$ for $x = \gamma(T/2)(y)$, and we recall that l_x is an isometry. □

[53] See Kobayaski & Nomizu (1963), Theorem 7.4, which in turn goes back to the ***Cartan–Ambrose–Hicks Theorem*** of Riemannian geometry. This states that f is a (local) isometry iff F preserves g and Riem, and for all $Y \in T_xM$ such that $\exp_x(Y) \in U_x$ one has $\text{Riem}_{\exp_x(F(Y))}(P_Y(U),P_Y(V),P_Y(W),P_Y(X)) = \text{Riem}_{\exp_x(Y)}(U,V,W,X)$ for all $U,V,W,X \in T_{\exp_x(Y)}$, where $P_Y : T_{\exp_x(Y)}M \to T_xM \to T_{\exp_x(F(Y))}M$ is the composition of parallel transport along the geodesics γ_Y (traversed backward) and $\gamma_{F(Y)}$. This condition is automatically satisfied when $\nabla\text{Riem} = 0$.

[54] See e.g. Helgason (1978), §IV.2. The compact-open topology on a space of maps $F : X \to Y$ is generated by open sets of the form $\mathbb{C}_{K,U} = \{F \mid F(K) \subset U\}$, where K is compact in X and U is open in Y.

It can be shown for any Riemannian space that if $\mathrm{Iso}(M,g)$ acts transitively on M, then already its identity component $\mathrm{Iso}(M,g)_0$ does so.[55] For symmetric spaces this is an easy exercise. We conclude that under the assumptions of Theorem 3.10 we have $M \cong G/K$, with $G = \mathrm{Iso}(M,g)_0$ and $K = G_{x_0}$ for some $x_0 \in M$. By Proposition 3.7, the given G-invariant (constant curvature) metric g on M is entirely determined by some suitable inner product $\langle \cdot,\cdot \rangle$ on $\mathfrak{g}/\mathfrak{k}$, and by Proposition 3.6 the K-action on $T_{x_0}M$ is mapped to the $\mathrm{Ad}'(K)$-action on $\mathfrak{g}/\mathfrak{k}$ (which by implication preserves $\langle \cdot,\cdot \rangle$). By Proposition 3.8 the representation Ad' is injective on K. Therefore, if we choose an orthonormal basis of $\mathfrak{g}/\mathfrak{k}$ with respect to $\langle \cdot,\cdot \rangle$, and hence obtain an identification $\mathfrak{g}/\mathfrak{k} \cong \mathbb{R}^n$, we may also identify $K \cong \mathrm{Ad}'(K)$ with a certain subgroup of $SO(n)$.

Lemma 3.16 *If, in the situation just described, (M,g) has constant curvature, then $K = SO(n)$.*

This follows by the argument in the proof of Lemma 3.13, which is applicable because constant curvature implies $\mathrm{Riem} = k \cdot S$, see the proof of Proposition (3.2) and especially eq. (3.1). By definition, any element $R \in SO(n)$ preserves the inner product, and hence, the metric, and hence, by the above formula, the Riemann tensor. Thus R comes from an isometry, i.e. $R \in K$. □

Therefore, we now know that $M \cong G/K$ as a homogeneous Riemannian manifold, where

$$G = \mathrm{Iso}(M,g)_0; \tag{3.33}$$

$$K = SO(n). \tag{3.34}$$

Since $SO(n)$ acts irreducibly on $\mathbb{R}^n$, so that $\mathrm{Ad}'(K)$ is irreducible on $\mathfrak{g}/\mathfrak{k}$, there is exactly one possible metric g on M (up to scaling by a positive constant), cf. Proposition 3.7. We now transfer the involutions l_x on M to G. Since for all $x \in M$ and $\gamma \in \mathrm{Iso}(M,g)$ one has (exercise)

$$\gamma l_x \gamma^{-1} = l_{\gamma x}, \tag{3.35}$$

it is sufficient to consider a single $l_{x_0} : M \to M$, where $x_0 \in M$ is arbitrary. For (3.33), define

$$l : G \to G; \tag{3.36}$$

$$\gamma \mapsto l_{x_0} \gamma l_{x_0}. \tag{3.37}$$

Using (3.35) and the definition of the maps l_x, it is easy to show that l has the properties

$$l \neq \mathrm{id}_G; \qquad l^2 = \mathrm{id}_G; \qquad l(\gamma\delta) = l(\gamma) l(\delta). \tag{3.38}$$

We defined l by (3.37) for (3.33) - (3.34), in which context (3.38) follow from the definition. Conversely, for any Lie group G one may start with a nontrivial smooth involutive automorphism (3.36), i.e. a map (3.36) satisfying (3.38), called a ***Cartan involution*** on G, and define

$$K = G^l \equiv \{\gamma \in G \mid l(\gamma) = \gamma\} \tag{3.39}$$

as the fixed-point set of l. Then construct a family $(l_x)_{x \in G/K}$ of diffeomorphisms of G/K by

$$l_K(\gamma K) = l(\gamma) K; \tag{3.40}$$

$$l_{\gamma K}(x) = \gamma \cdot l_K(\gamma^{-1} \cdot x). \tag{3.41}$$

[55] The following analysis could have been performed with $\mathrm{Iso}(M,g)$, too, in which case $O(n)$ replaces $SO(n)$.

If K is connected (as is the case in our application), these procedures are equivalent,[56] so that one may start either with a symmetric Riemannian manifold (M,g) or with the corresponding group-theoretical data (G,l). It is striking how many interesting Riemannian manifolds are obtained in this way (including all those on which serious harmonic analysis is possible).[57]

Three important classes of examples, which are also very easy to state, are:[58]

$$G = SL_n(\mathbb{R}), \qquad l(A) = (A^T)^{-1}, \qquad K = SO(n); \tag{3.42}$$

$$G = SO(n+m), \qquad l(A) = \Gamma(A^T)^{-1}\Gamma, \qquad K = SO(n)\times SO(m); \tag{3.43}$$

$$G = SO(n,m)^+, \qquad l(A) = (A^T)^{-1}, \qquad K = SO(n)\times SO(m). \tag{3.44}$$

where $\Gamma = \mathrm{diag}(1, \overset{n\,\text{times}}{\ldots\ldots}, 1, -1, \overset{m\,\text{times}}{\ldots\ldots}, -1)$. The first line is irrelevant to Theorem 3.10 (expect for $n = 2$, for which it yields the same quotient as the third line for $n = 2$, $m = 1$), but the second line produces S^n for $m = 1$, which very choice in the third line gives the hyperbolic spaces M^n.

The classification problem is much easier at the level of Lie algebras, to which we now turn. Since $l : G \to G$ is smooth, it has a derivative $l' : \mathfrak{g} \to \mathfrak{g}$, defined by, cf. (B.2),

$$l'(A) = \frac{d}{dt} l\left(e^{tA}\right)_{t=0}. \tag{3.45}$$

As in (3.26), this map satisfies $\exp(l'(A)) = l(\exp(A))$. From this, and $l^2 = \mathrm{id}_G$, we compute

$$l' \circ l'(A) = \frac{d}{dt} l\left(e^{tl'(A)}\right)_{t=0} = \frac{d}{dt} l\left(l(\exp(tA))\right)_{t=0} = \frac{d}{dt}\left(e^{tA}\right)_{t=0} = A, \tag{3.46}$$

so that $(l')^2 = \mathrm{id}_{\mathfrak{g}}$. We therefore have our promised canonical decomposition (3.18), in which $\mathfrak{k}$ and $\mathfrak{p}$ are the eigenspaces of l' with eigenvalue 1 and -1, respectively. Furthermore, it follows from the last entry in (3.38) that l' is a Lie algebra automorphism, i.e., l' is linear and (exercise)

$$l'([A,B]) = [l'(A), l'(B)]. \tag{3.47}$$

This implies the following properties (of which the first one is trivial since $K \subset G$ is a subgroup):

$$[\mathfrak{k},\mathfrak{k}] \subset \mathfrak{k}; \qquad [\mathfrak{k},\mathfrak{p}] \subset \mathfrak{p}; \qquad [\mathfrak{p},\mathfrak{p}] \subset \mathfrak{k}. \tag{3.48}$$

We now return to our classification problem (i.e., the proof of Theorem 3.10), so that Proposition 3.7 and 3.33 - (3.34) apply, as well as the remarks preceding Lemma 3.16. Consequently, $\mathfrak{p} \cong \mathbb{R}^n$ and the $\mathrm{Ad}'(K)$-action on $\mathfrak{p}$ is the defining action of $SO(n)$ on $\mathbb{R}^n$. By (3.27), the derivative of the $\mathrm{Ad}'(K)$-action is the $\mathrm{ad}(\mathfrak{k})$-action, that is, for $A \in \mathfrak{k} = \mathfrak{so}(\mathfrak{n})$ and $v \in \mathfrak{p}$ we have

$$[A,v] = A \cdot v, \tag{3.49}$$

where $A \cdot v$ is the derivative of the defining action of $SO(n)$, see (3.24). Since the Lie bracket $[A,B]$ for $A,B \in \mathfrak{k}$ is also known (because $\mathfrak{k} = \mathfrak{so}(\mathfrak{n})$), all we need to find out to determine $\mathfrak{g}$ as a Lie algebra (and hence, by Lie's Third Theorem in Appendix C, to determine G as a Lie group) is the commutator $[u,v] \in \mathfrak{k}$ of $u,v \in \mathfrak{p}$ (identified with $\mathbb{R}^n$). To start, for $n = 2$ we must have

$$[u,v]w = \rho(\langle u,w\rangle v - \langle v,w\rangle u), \tag{3.50}$$

[56] If K is disconnected, then $G_0^l \subseteq K \subseteq G^l$. See Helgason (1978) or even wikipedia, `symmetric space`.

[57] See Jost (2002) and Terras (2016) for examples and Helgason (1978) and wikipedia for a classification.

[58] In Cartan's classification these correspond to types AI (noncompact), BDI (compact), and BDI (noncompact), respectively. See Helgason (1978), Table V, p. 518 or Terras (2016), Table 2.1, p. 357 for a summary.

for some constant $\rho \in \mathbb{R}$. This is easily shown by taking the usual basis (e_1, e_2) of $\mathbb{R}^2$ as well as the generator $j_3 \equiv e_3$ of $\mathfrak{so}(2)$, see (B.18). The only possibility is $[e_1, e_2] = \rho e_3$, which by linear extension gives (3.50). Rescaling of the metric by a positive constant then leads to the three possibilities $\rho = 1, 0, -1$. This leads to the following three possibilities:

$$\rho = 1: \qquad [e_1, e_2] = e_3; \qquad [e_3, e_1] = e_2; \qquad [e_3, e_2] = -e_1; \tag{3.51}$$

$$\rho = 0: \qquad [e_1, e_2] = 0; \qquad [e_3, e_1] = e_2; \qquad [e_3, e_2] = -e_1; \tag{3.52}$$

$$\rho = -1: \qquad [e_1, e_2] = -e_3; \qquad [e_3, e_1] = e_2; \qquad [e_3, e_2] = -e_1. \tag{3.53}$$

These are the Lie algebras of $SO(3)$, $E(2)$, and $SO(2,1)$, respectively, see (B.8), (B.19), and (B.10). It follows from Theorem C.1 that G (assumed connected) must be:

$$\rho = 1: \qquad G = SO(3); \tag{3.54}$$

$$\rho = 0: \qquad G = E(2)_0; \tag{3.55}$$

$$\rho = -1: \qquad G = SO(2,1)^+, \tag{3.56}$$

or any other Lie group differing only in its topology (as made precise by Theorem C.1). However, the candidates just listed are the only ones whose quotients with $SO(2)$ give simply connected spaces (as may be checked case by case). This proves Theorem 3.10 for $n = 2$.

The general proof just needs one additional argument (plus some Lie algebra bookkeeping). Namely, (3.50) holds in any dimension! To see this, we recall that the adjoint action of $K = SO(n)$ on $\mathfrak{g}$ consists of Lie algebra automorphisms (indeed this is true for all of G). Hence

$$[Ru, Rv] = \mathrm{Ad}(R)([u, v]) = R[u, v]R^{-1}, \tag{3.57}$$

for any $R \in SO(n)$ and $u, v \in \mathfrak{p} \cong \mathbb{R}^n$, with $[u, v] \in \mathfrak{so}(\mathfrak{n})$. If $n > 2$, we may take three mutually orthogonal vectors u, v, w and take R to be the reflection in the (hyper)plane orthogonal to w.[59] Then $Ru = u$, $Rv = v$, and $R^{-1}w = Rw = -w$ by construction, so that (3.57) gives

$$[u, v]w = -R([u, v]w). \tag{3.58}$$

By definition of R (which implies that $Rx = -x$ is only true if x is a multiple of w), this implies that $[u, v]w$ is a multiple of w, which is impossible for (infinitesimal) rotations unless $[u, v]w = 0$. Therefore, $[u, v]$ maps any vector orthogonal to u and v to zero, which yields (3.50) for any n.

The finishing touch is to note that the covariance property (3.57) has not only delivered the conclusion just given, but also implies that the constant ρ in (3.50) is independent of the u-v plane (since one can rotate any plane into any other plane). This means that the Lie algebra $\mathfrak{g}$ is now entirely known, and it is a straightforward exercise (which we spare the reader) to find the right basis of $\mathfrak{g}$ for the three cases $\rho = 1, 0, -1$ and hence reproduce the known Lie algebras of $SO(n+1)$, $E(n)$, and $SO(n,1)$, again with respect to a suitable basis.[60] The final identification

$$\rho = 1: \qquad G = SO(n); \tag{3.59}$$

$$\rho = 0: \qquad G = E(n)_0; \tag{3.60}$$

$$\rho = -1: \qquad G = SO(n,1)^+, \tag{3.61}$$

also uses practically the same (topological) case by case arguments as for $n = 2$. □

[59] We are cheating, since this reflection does not lie in $SO(n)$ but in the other component of $O(n)$, where $\det(R) = -1$. However, the entire argument can be carried through using $G = \mathrm{Iso}(M, g)$ and $K = O(n)$, cf. footnote 55.

[60] This basis is provided by the root space decomposition of semi-simple Lie algebras and requires more background than is offered in our appendices. Helgason (1978) is a complete reference for this background.

3.4 The Thurston classification

We close this chapter with the result of Thurston's classification in $d = 3$, which generalizes the case $n = 3$ of Theorem 3.10. This classification lists (up to isometry) all simply connected homogeneous Riemannian 3-manifolds $M = G/K$ (where $K \subseteq SO(3)$) for which:

1. the metric on M has a maximal isometry group;
2. G has a discrete subgroup Γ acting freely and discontinuously on M with $\Gamma\backslash M$ compact.

Since the aim of the classification is to produce all compact 3-manifolds, the second condition should be clear; the difference with $d = 2$ lies in the fact that constant curvature is replaced by the more general requirement of homogeneity (and it is only clear with hindsight that this was the right point to stop, rather than, say, the less general symmetric spaces or some more general class of Riemannian manifolds than homogeneous ones). We now rearrange the solution (1.27) according to K, which makes the special case of constant curvature stand out as the first line:

$$\begin{aligned}
&K = SO(3): \quad && M = S^3; && M = \mathbb{R}^3; && M = M^3; && \\
&K = SO(2): \quad && M = S^2 \times \mathbb{R}; && M = M^2 \times \mathbb{R}; && M = \widetilde{SL_2(\mathbb{R})}; && M = \mathrm{Nil}; \\
&K = \{e\}: \quad && M = \mathrm{Sol}. &&&&&&
\end{aligned} \tag{3.62}$$

By Proposition 3.7, only the spaces in the top line have unique G-invariant Riemannian metrics (up to scaling); for the others there is a choice, which is made by enforcing condition 1. Of the eight spaces listed in (3.62), the following five are simply connected $3d$ Lie groups:

$$S^3 \cong SU(2), \qquad \mathbb{R}^3, \qquad \widetilde{SL_2(\mathbb{R})}, \qquad \mathrm{Nil} = \mathrm{Heis}_3(\mathbb{R}), \qquad \mathrm{Sol} = E(1,1)_0. \tag{3.63}$$

are $3d$ Lie groups, see Appendix A for the definition of these groups.[61] On a different note, all spaces except S^3 and $S^2 \times \mathbb{R}$ are diffeomorphic (but not isometric!) to $\mathbb{R}^3$. We now give brief descriptions of the metrics on the spaces (3.62) and their associated isometry groups.

1. $M = S^3$ with round metric (i.e. obtained from the embedding $S^3 \hookrightarrow \mathbb{R}^4$). We already know that $S^3 \cong SO(4)/SO(3)$, with unique $SO(4)$-invariant metric (up to scaling), but there is another group-theoretic description coming from an "accident" in $d = 3$, namely
$$S^3 \cong SU(2); \tag{3.64}$$
see (C.1). Now any group G carries a transitive $G \times G$ action defined by $(x,y) \cdot z = xzy^{-1}$), whose stabilizer at the identity $e \in G$ is $K = G_d = \{(\gamma,\gamma) \mid \gamma \in G\}$, so that $K \cong G$.

[61] These are precisely the $3d$ simply connected *unimodular* Lie groups, which had previously been classified by Milnor (1976). A Lie group (and more generally a locally compact group) is called ***unimodular*** if it has a nonzero measure that is both left- and right-invariant under the G-action (any locally compact group has a left-invariant measure called the ***Haar measure***, so what is at stake is whether this measure is also right-invariant). It can be shown that a Lie group is unimodular iff $\mathrm{Tr}(\mathrm{ad}(A)) = 0$ for all $A \in \mathfrak{g}$, so that, using Theorem C.1, the list (3.63) may be proved by classifying all $3d$ unimodular Lie algebras. This is done in Milnor (1976), whose basic lemma states that any $3d$ unimodular Lie algebra has a basis (n_1, n_2, n_3) with Lie bracket (B.26) in our Appendix B. This leads to the table at the end of Appendix B. Note that $E(2)$ occurs in this table but not in (3.63); this is because $E(2) \cong \mathbb{R}^3$ as manifolds and the Riemannian metric on $\mathbb{R}^3$ used in Thurston's classification does not use the group structure of $E(2)$. The so-called ***Bianchi classification*** predates and generalizes this, since it classifies all $3d$ Lie algebras. The relevance of the unimodularity condition for Thurston's program is that G has a discrete subgroup Γ whose left-action on G has a compact quotient iff G is unimodular.

On any (real) Lie algebra $\mathfrak{g}$, the ***Cartan–Killing*** form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is defined by

$$B(X,Y) = \tfrac{1}{2}\mathrm{Tr}\,(\mathrm{ad}_X \mathrm{ad}_Y). \tag{3.65}$$

This bilinear form is always $\mathrm{Ad}(G)$-invariant, is nondegerenate iff $\mathfrak{g}$ is semi-simple (i.e., the direct sum of simple Lie algebras), and is negative definite iff G is compact. Therefore, if G is a compact semi-simple Lie group, like $SU(2)$, we have an $\mathrm{Ad}(G)$-invariant inner product on $\mathfrak{g}$, given by *minus* B. As a case in point, it follows from (B.13), (B.14), and (B.7) that under the identification $\mathfrak{su}(2) \cong \mathbb{R}^3$ we obtain $\mathrm{ad}(h_i) = e_i$, $i = 1,2,3$, and since $\mathrm{Tr}\,(e_i e_j) = -2\delta_{ij}$, i.e., $B(h_i h_j) = \delta_{ij}$, minus B is just the usual inner product on $\mathbb{R}^3$.

Let G be a compact semi-simple Lie group. Under the identification (3.64), we put the inner product $-B$ on $T_eG \cong \mathfrak{g}$, and move this to any other point by the left G-action. This gives a $G \times G$-invariant metric g on G. Hence for $G = SU(2)$ it is tempting to write

$$S^3 \cong (SU(2) \times SU(2))/SU(2), \tag{3.66}$$

which is true, but for the analysis of §3.1 to apply, one must assume that the G-action on M is effective. This is not the case for the $SU(2) \times SU(2)$ action on itself (and hence on S^3), and more generally, is not the case whenever G has a nontrivial center $Z(G)$, which for $SU(2)$ consists of $Z(SU(2)) = \{1_2, -1_2\}$. So we should rewrite (3.66) as

$$S^3 \cong ((SU(2) \times SU(2))/\mathbb{Z}_2)/(SU(2)/\mathbb{Z}_2) \cong ((SU(2) \times SU(2))/\mathbb{Z}_2)/SO(3), \tag{3.67}$$

where the first $\mathbb{Z}_2$ is the subgroup of $SU(2) \times SU(2)$ consisting of $(1_2, 1_2)$ and $(-1_2, -1_2)$, and the second one is the subgroup $\{1_2, -1_2\}$ of $SU(2)$ just mentioned. We also used

$$SO(3) \cong SU(2)/\mathbb{Z}_2, \tag{3.68}$$

and to make the story complete we mention a third (related) "accident" in $d = 3$, namely

$$SO(4) \cong (SU(2) \times SU(2))/\mathbb{Z}_2, \tag{3.69}$$

which is most easily proved at the Lie algebra level and then noting that $SO(4)$ is doubly connected whereas $SU(2) \times SU(2)$ is simply connected, because S^3 is, cf. (3.64). Hence (3.67) recovers our original description $S^3 \cong SO(4)/SO(3)$, and under all these isomorphisms, the $SU(2) \times SU(2)$-invariant metric g on S^3 that is ultimately given by (minus) the Cartan–Killing form is the same as the round metric, whose (connected) isometry group $\mathrm{Iso}(S^3, g)_0$ may be written as either side of (3.69), depending on the choice of either $G = SO(4)$ or $G = (SU(2) \times SU(2))/\mathbb{Z}_2$, respectively, in writing $S^3 \cong G/SO(3)$.

2. $M = \mathbb{R}^3$ with flat metric, $M = M^3$ with hyperbolic metric, $M = S^2 \times \mathbb{R}$ with product metric (of the round metric on S^2 and the flat metric on $\mathbb{R}$), and $M = M^2 \times \mathbb{R}$ with product metric (of the hyperbolic metric on M^2 and the flat metric on $\mathbb{R}$) are easy and have:

$$\mathrm{Iso}(\mathbb{R}^3, g)_0 = E(3)_0; \tag{3.70}$$

$$\mathrm{Iso}(M^3, g)_0 = SO(3,1)^+; \tag{3.71}$$

$$\mathrm{Iso}(S^2 \times \mathbb{R}, g)_0 = SO(3) \times \mathbb{R}; \tag{3.72}$$

$$\mathrm{Iso}(M^2 \times \mathbb{R}, g)_0 = SO(2,1) \times \mathbb{R}; \tag{3.73}$$

3. The remaining three (group) cases are more complicated, see Scott (1983) for the explicit metrics. The isometry groups of the metrics on $\widetilde{SL_2(\mathbb{R})}$ and $M = \mathrm{Nil}$ are (by construction) certain extensions of these groups by $SO(2)$, and finally $\mathrm{Iso}(\mathrm{Sol}, g)_0$ is simply Sol itself.

4 Ricci Flow

At last, we now return to the Ricci Flow equation (1.1) or (1.3). Much of the analysis is done through derived equations for the various curvature tensors, of which the one for the curvature scalar R , viz. (1.9), even has a clearer intuitive meaning than the Ricci Flow equation itself. We first state and derive these equations, and then return to the Ricci Flow equation itself, proving existence (al least for short time) and uniqueness (this order is pedagogical rather than logical).

4.1 Derived equations

The following theorem states the flow of the curvature tensors and some other relevant objects.

Theorem 4.1 *If the metric g satisfies the Ricci Flow equation* (1.1)*, then the associated curvature quantities, i.e. the scalar curvature and the Ricci and Riemann tensors, satisfy:*

$$\partial_t R = \Delta_g R + 2R^{ij}R_{ij}; \tag{4.1}$$

$$\partial_t R_{ij} = \Delta_g R_{ij} - 2R_i^k R_{kj} + 2R_{ikjl}R^{kl}; \tag{4.2}$$

$$\begin{aligned}\partial_t R_{ijkl} = {} & \Delta_g R_{ijkl} + R_l^m R_{ijmk} - R_k^m R_{ijml} - R_j^m R_{imkl} - R_i^m R_{mjkl} \\ & + 2(B_{ijkl} - B_{ijlk} + B_{ikjl} - B_{iljk}),\end{aligned} \tag{4.3}$$

where $\Delta_g = g^{ij}\nabla_i\nabla_j$ is the covariant Laplacian and $B_{ijkl} = R^m_{ijn}R^n_{lkm}$. Moreover, if M is compact,

$$\partial_t \mathrm{Vol}(M) = -\int_M R; \tag{4.4}$$

$$\partial_t \int_M R = \int_M (-R^2 + 2R_{ij}R^{ij}); \tag{4.5}$$

$$\partial_t \langle R \rangle = \langle R \rangle^2 - \langle R^2 \rangle + 2\langle R_{ij}R^{ij} \rangle. \tag{4.6}$$

See also (1.22) - (1.24). In any coordinate system, $\int_M f$ means $\int_M d^n x \sqrt{\det(g(x))} f(x)$, for say $f \in C(M)$.[62] One sees that Ricci Flow is a reaction-diffusion equation, with a diffusion (or 'heat') term $\Delta_g \cdots$ and, typically nonlinear, reaction terms involving higher curvature tensors, expect for Riem, where the buck stops. The Ricci Flow equation (1.1) for the metric is of the same type, but it will take us some effort to show this (see the next section). Also, one sees from (4.4) that *positive/negative* curvature *decreases/increases* the volume under Ricci Flow.[63]

[62] We recall our standing assumption that M is *orientable*. Recall that this means that there exists an atlas $(U_\alpha, \varphi_\alpha)$, where the U_α form an open cover of M and each $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$ is injective and open, whose transition functions $\varphi_\beta \circ \varphi_\alpha^{-1} : V_{\alpha\beta} \to \mathbb{R}^n$, where $V_{\alpha\beta} = \varphi_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$) all have positive Jacobian. An *orientation* of an orientable manifold, then, is an atlas satisfying this condition. It can be shown that M is orientable iff it admits a nowhere vanishing n-form $\omega \in \Omega^n(M)$; one then only accepts charts φ whose coordinates $(x^1, \ldots, x^n)$ satisfy $\omega(\partial_1, \ldots, \partial_n) > 0$. In the presence of a metric there is a canonical normalization of ω, given by the condition

$$\omega(\partial_1, \ldots, \partial_n) = \sqrt{\det(g)} \qquad \Leftrightarrow \qquad \omega_x = \sqrt{\det(g(x))}\, dx^1 \wedge \cdots \wedge dx^n.$$

This condition is well defined, since ω keeps this form under coordinate transformations: indeed, by elementary calculus one has $\sqrt{\det(g(x_\beta))} = J_{\alpha\beta}^{-1}\sqrt{\det(g(x_\alpha))}$, where $J_{\alpha\beta} = \det|\partial x^i_\beta / \partial x^j_\alpha|$ is the Jacobian of the coordinate transformation from x_α to x_β). For any reasonable $f : M \to \mathbb{R}$ one then has $\int_M f \equiv \int_M f\omega = \int_M d^n x \sqrt{\det(g(x))} f(x)$.

[63] In particular, in $d = 2$, by (2.44) we have $\partial_t R = \Delta_g R + R^2$ and $\partial_t \int_M R = 0$, as follows also from the Gauss–Bonnet Theorem (1.12). This also turns (4.4) into $\partial_t \mathrm{Vol}(M) = -4\pi\chi$, and hence $\mathrm{Vol}(M)(t) = \mathrm{Vol}(M)(0) - 4\pi\chi t$. It follows that if $\chi > 0$, then $\mathrm{Vol}(M)$ becomes zero in finite time, whereas $\chi < 0$ gives eternal expansion.

Proof. We will not prove all equations above, but illustrate the method and leave horrific cases like (4.3) as exercises to the reader. It is instructive to first consider general flows

$$\partial_t g_{ij}(t,x) = h_{ij}(t,x), \tag{4.7}$$

and then put $h = -2Ric$ or $h_{ij} = -2R_{ij}$ at the end of the computation. The simplest result is

$$\partial_t g^{ij} = -g^{ik}g^{jl}h_{kl} = -h^{ij}, \tag{4.8}$$

which follows from the defining property $g^{ij}g_{jk} = \delta^i_k$, and also illustrates the convention that indices are raised and lowered with $g(t)$, as usual (we often omit the t-dependence). Next,

$$\partial_t \Gamma^k_{ij} = \tfrac{1}{2}g^{kl}(\nabla_i h_{jl} + \nabla_j h_{il} - \nabla_l h_{ij}), \tag{4.9}$$

where the Christoffel symbol Γ was defined in (2.13). This may be proved either by direct computation, or the following instructive trick, which can often be used.

First note that although the coefficients Γ^k_{ij} (for fixed $g(t)$) do not form the components of a tensor, their derivatives $\partial_t \Gamma^k_{ij}$ do. This is because *the difference between two connections is a tensor*. Let ∇ and $\tilde{\nabla}$ be connections on TM. Then the difference $(\nabla_X - \tilde{\nabla}_X)Y$ is $C^\infty(M)$-linear in $Y \in \mathfrak{X}(M)$, unlike $\nabla_X Y$ and $\tilde{\nabla}_X Y$, since the spoiler $(Xf)Y$ in the Leibniz rule (2.3) drops out. For example, let ∇ be the Levi-Civita connection for a given metric g and let $\tilde{\nabla}$ be the one for some other metric $\tilde{g}$. We then have a *tensor* $\hat{\Gamma}$, defined by $\hat{\Gamma}(X,Y,\theta) = \theta(\nabla_X Y - \tilde{\nabla}_X Y)$, whose connection coefficients are $\Gamma^k_{ij} - \tilde{\Gamma}^k_{ij}$, cf. (2.7). In particular, take $g = g(t+s)$ and $\tilde{g} = g(t)$. Since $\partial_t \Gamma^k_{ij}(g) = \lim_{s\to 0}(\Gamma^k_{ij}(g) - \Gamma^k_{ij}(\tilde{g}))/s$, the $\partial_t \Gamma^k_{ij}$ form the components of a tensor $\partial_t \Gamma$.

This is useful, because if σ and τ are tensors of the same type, say $(1,1)$, then $\sigma = \tau$ is true iff for each $x \in M$ one has $\sigma^j_i(x) = \tau^j_i(x)$ in just *one* specific coordinate system (x^i) defined on some nbhd U of x, which system may even depend on x. For in that case we have $\sigma_x(\partial_i, dx^j) = \tau_x(\partial_i, dx^j)$, and so, by $C^\infty(M)$-linearity of σ and τ in its arguments, $\sigma(X,\theta) = \tau(X,\theta)$, where we write $X = X^i\partial_i$ and $\theta = \theta_j dx^j$ as usual, for some $X^i \in C^\infty(U)$ and $\theta_j \in C^\infty(U)$. Similarly for tensors of any type (k,l). It therefore suffices to verify (4.9) in *geodesic normal coordinates*,[64] where at $x = x_0$ we have $\nabla = \partial$, and hence (4.9) is very easy to check.

Similarly, we compute $\partial_t R_{ij}$ from its definition (2.38) with (2.19), viz.

$$R_{ij} = \partial_k \Gamma^k_{ij} - \partial_j \Gamma^k_{ik} + \Gamma^k_{kl}\Gamma^l_{ij} - \Gamma^l_{jk}\Gamma^k_{il}. \tag{4.10}$$

Since only the first two terms contain second-order derivaties of g_{ij}, in computing $\partial_t R_{ij}$ in GNC only these are nonzero, and can be computed from (4.9). This gives the tensor equality

$$\partial_t R_{ij} = \tfrac{1}{2}(\nabla_k \nabla_i h^k_j + \nabla_k \nabla_j h^k_i - \nabla_i \nabla_j h^k_k - \nabla^k \nabla_k h_{ij}), \tag{4.11}$$

in which the substitution $h = -2\text{Ric}$ and by some rearrangements involving (2.18) give (4.2). From this and (4.8), which gives $\partial_t g^{ij} = 2R^{ij}$, one easily finds (4.1).
With more work, the technique we used for (4.2) also yields (4.3). For (4.4) - (4.6) we use

$$\frac{\partial \det(g)}{\partial g_{ij}} = g^{ij}\det(g), \tag{4.12}$$

[64] Geodesic normal coordinates (GNC) are defined on a normal nbhd U_{x_0} of some fixed $x_0 \in M$, relative to an orthonormal basis (e_i) of $T_{x_0}M$: the coordinates of $x = \exp_{x_0}(X) \in U_{x_0}$, where $X = X^i e_i$, are just $x^i = X^i$. It is a simple exercise to show that in these coordinates $x^i_0 = 0$, $g_{ij}(x_0) = \delta_{ij}$, and $\partial_k g_{ij}(x_0) = 0$, so that also $\Gamma^k_{ij}(x_0) = 0$.

which follows from straightforward linear algebra (exercise).[65] This relation implies

$$\partial_t \sqrt{\det(g)} = \frac{\partial \sqrt{\det(g)}}{\partial g_{ij}} h_{ij} = \frac{1}{2\sqrt{\det(g)}} \frac{\partial g}{\partial g_{ij}} h_{ij} = \tfrac{1}{2}\sqrt{\det(g)}\, g^{ij} h_{ij} = \tfrac{1}{2}\sqrt{\det(g)} h, \tag{4.13}$$

from which (4.4) is immediate. To prove (4.5), we show that under (4.7) we have

$$\partial_t \int_M R = \int_M h^{ij}(\tfrac{1}{2} g_{ij} R - R_{ij}), \tag{4.14}$$

where we recall that in coordinates we integrate with respect to the measure $d^n x \sqrt{\det(g(x))}$. Consequently, we have to compute $\partial_t(\sqrt{\det(g)} g^{ij} R_{ij})$, which is the sum of three terms:

$$(\partial_t \sqrt{\det(g)}) g^{ij} R_{ij} = \sqrt{\det(g)}\, h^{ij} \cdot \tfrac{1}{2} g_{ij} R; \tag{4.15}$$

$$\sqrt{\det(g)} (\partial_t g^{ij}) R_{ij} = -\sqrt{\det(g)}\, h^{ij} \cdot R_{ij}; \tag{4.16}$$

$$\sqrt{\det(g)} g^{ij} (\partial_t R_{ij}) = \sqrt{\det(g)}\, \nabla_i X^i, \tag{4.17}$$

where $X^i = \nabla_j h^{ij} - \nabla^i h^j_j$ is a vector field on M whose explicit form does not matter, since

$$\int_M \nabla_i X^i = 0 \tag{4.18}$$

for any vector field on M (recalling that M is compact without boundary).[66] This follows from

$$\sqrt{\det(g)}\, \nabla_i X^i = \partial_i(\sqrt{\det(g)}\, X^i), \tag{4.19}$$

and Gauss's Theorem (calculus) or Stokes's Theorem (geometry). Eq. (4.19) follows from

$$\partial_i \sqrt{\det(g)} = \sqrt{\det(g)}\, \Gamma^j_{ij}, \tag{4.20}$$

whose proof is an exercise.[67] Thus we are finally in a position to compute:

$$\partial_t \int_M R = \partial_t \int_M d^n x \sqrt{\det(g)} g^{ij} R_{ij} = \int_M d^n x \sqrt{\det(g)} h^{ij} (\tfrac{1}{2} g_{ij} R - R_{ij}), \tag{4.21}$$

which is (4.14); putting $h_{ij} = -2R_{ij}$ as usual then gives (4.5). Eq. (4.6) immediately follows from (4.4) and (4.5). Using these equations, it is easy to show that the ***normalized Ricci Flow***

$$\frac{\partial g_{ij}}{\partial t} = -2R_{ij} + \frac{2}{n} g_{ij} \langle R(g) \rangle \tag{4.22}$$

preserves volume; that is, if $g(t)$ solves (4.22), then $\mathrm{Vol}(M)$, which through its dependence on $\sqrt{\det(g(t))}$ is a function of t, is independent of t. Normalized Ricci Flow is often better behaved than Ricci Flow and we will return to it often. One sees at once that its fixed points ($\partial_t g = 0$) must be Einstein metrics, and in $d = 3$ therefore have constant curvature (cf. Proposition 3.3).

---

[65] *HInt:*, $\partial \det(g)/\partial g_{ij} = m^{ij}$, i.e. the minor = cofactor of g_{ij}, and $g^{ij} = m^{ji}/\det(g)$.

[66] Continuing footnote 62, a more abstract proof of (4.18) is as follows. Eq. (4.19) is a coordinate version of the geometric formula $\mathscr{L}_X \omega = \omega \nabla \cdot X$. ***Cartan's formula*** for the Lie derivative of exterior forms states that $\mathscr{L}_X = d i_X + i_X d$, where $X \in \mathfrak{X}(M)$, i.e., for any p-form $\alpha \in \Omega^p(M)$, $p > 0$, we have $\mathscr{L}_X \alpha = d(i_X \alpha) + i(d\alpha)$, where $d : \Omega^p(M) \to \Omega^{p+1}(M)$ is the ***exterior derivative*** (defined in coordinates by $(d\alpha)_{\mu_1 \cdots \mu_{p+1}} = \partial_{\mu_1} \alpha_{\mu_2 \cdots \mu_{p+1}}$) and $i_X : \Omega^p(M) \to \Omega^{p-1}(M)$ is the ***insertion operation***, defined in coordinates by $(i_X \alpha)_{\mu_2 \cdots \mu_p} = X^{\mu_1} \alpha_{\mu_1 \mu_2 \cdots \mu_p}$. Since $\omega \in \Omega^n(M)$ we must have $d\omega = 0$, so that Cartan's formula gives $\mathscr{L}_X \omega = d(i_X \omega)$, and hence, with the first equation in this footnote, $\omega \nabla \cdot X = d(i_X \omega)$. The geometric version of Stokes's Theorem states that $\int_M d\alpha = \int_{\partial M} \alpha$, for any $\alpha \in \Omega^n(M)$, so that $\int_M \omega \nabla \cdot X = \int_{\partial M} i_X \omega$. If M has no boundary, this gives (4.18).

[67] *Answer.* Since the first term in (2.13) cancels the last if $j = k$, we have $\Gamma^j_{ij} = \frac{1}{2} g^{jk} \partial_i g_{jk}$. Diagonalizing the symmetric invertible matrix (g_{jk}), yielding nonzero eigenvalues $(\lambda_1, \ldots, \lambda_n)$ and realizing that (g^{jk}) is its inverse gives $g^{jk} \partial_i g_{jk} = \frac{\partial_i \lambda_1}{\lambda_1} + \cdots + \frac{\partial_i \lambda_n}{\lambda_n}$, But also $2 \frac{\partial_i \sqrt{\det(g)}}{\sqrt{\det(g)}} = g^{-1} \partial_i g = \frac{\partial_i (\lambda_1 \cdots \lambda_n)}{\lambda_1 \cdots \lambda_n} = \frac{\partial_i \lambda_1}{\lambda_1} + \cdots + \frac{\partial_i \lambda_n}{\lambda_n}$.

4.2 Ricci Flow as a parabolic PDE

In contrast to the derived equations (4.1) - (4.3) for the curvature tensors, the original Ricci Flow equation (1.1) for the metric itself looks quite opaque. We will now show that in suitable coordinates (1.1) is a mildly nonlinear ("quasilinear") version of the heat equation; later, this will also be the key to the proof of (short-time) existence and uniqueness of its solutions.[68]

The simplest approach is to introduce ***harmonic ccordinates*** x^i, which by definition satisfy

$$\Delta_g x^i = 0, \tag{4.23}$$

where $\Delta_g = g^{kl}\nabla_k\nabla_l$ as before, and the index i is not treated as a vector index: the x^i are seen as n functions of given coordinates $(y^1,\ldots,y^n)$, in terms of which the metric is given. Since

$$\Delta_g f = g^{kl}\nabla_k\nabla_l f = g^{kl}(\partial_k\partial_l - \Gamma^j_{kl}\partial_j)f \equiv (\Delta - \Gamma^j\partial_j)f, \tag{4.24}$$

for any function $f \in C^\infty(M)$, or locally $f \in C^\infty(U)$ on a chart domain $U \subset M$, where

$$\Delta = g^{kl}\partial_k\partial_l; \tag{4.25}$$

$$\Gamma^j = g^{kl}\Gamma^j_{kl}, \tag{4.26}$$

eq. (4.23) is a nonlinear elliptic PDE for the n functions x^i. For this, local existence and uniqueness can be proved,[69] subject to initial conditions such as $x^i(p) = 0$ and $\partial x^i/\partial y^j(p) = \delta^i_j$ for some $p \in U$, where $U \subset M$ is the patch on which the original y-coordinates are defined. Having switched to harmonic coordinates, we of course have $\partial_j x^i = \delta^i_j$, so that (4.24) reads

$$\Delta_g x^i = -\Gamma^i, \tag{4.27}$$

and (4.23) is therefore equivalent to the condition

$$\Gamma^i = 0. \tag{4.28}$$

The point of using harmonic ccordinates is that in *arbitrary* coordinates one has (exercise)

$$-2R_{ij} = \Delta_g g - g_{ik}\partial_j\Gamma^k - g_{jk}\partial_i\Gamma^k + O(g,\partial g), \tag{4.29}$$

where the remainder $O(g,\partial g)$ contains only first or no derivatives of the metric g; note that (4.10) has 24 terms! The right-hand side starts like a heat equation, but the next two terms, which also contain second-order derivatives of the metric, ruin its parabolic nature. However, in view of (4.28), *in harmonic ccordinates* the Ricci Flow equation (1.1) reads

$$\partial_t g_{ij} = \Delta g_{ij} + O'(g,\partial g), \tag{4.30}$$

where again the remainder only contains derivatives of lower order than two; it differs from $O(g,\partial g)$ in (4.29) by $\Delta_g g - \Delta g$, which contains no second-order derivatives.

[68] What we are now going to explain is called the ***DeTurck trick*** in the literature, after DeTurck (1983). This method originates in the work of Y. Choquet-Bruhat on the initial value problem in General Relativity (GR), which in terms was motivated by harmonic maps in Riemannian geometry, see e.g. Jost (2002), Chapter 8. Despite a clear formal analogy, the situation in GR is quite different, however: applying a diffeomeorphism that fixes the initial data to a solution of the Einstein equations gives *another* solution, which is therefore underdetermined, and hence the Choquet-Bruhat trick picks a convenient solution within its diffeomorphism class. For Ricci Flow, the DeTurck trick maps a solution to a convenient *non* solution, which has to be transformed back so as to find the solution.

[69] See DeTurck & Kazdan (1981), which relies on standard elliptic regularity theorems.

Our gain is that, unlike (1.1), eq. (4.30) is a ***quasilinear second-order parabolic PDE***:

- It is *quasilinear* because the highest (in this case, second) order derivates of the unknown (i.e. g_{ij}) occur linearly (but don't be fooled: even the truncated equation is highly nonlinear, since the coefficients g^{kl} in $\Delta_0 = g^{kl}\partial_k\partial_l$ contain g_{ij});
- It is *second order* because the highest derivatives are order two;
- It is *parabolic* because it takes the form

$$\partial_t u = a^{kl}(x,t,u,\partial u)\partial_k\partial_l u + O(u,\partial u), \tag{4.31}$$

  where for any fixed v (in the class of functions where one looks for solutions) one has

$$a(t,x,v,\partial v) \geq \lambda \cdot 1_n, \tag{4.32}$$

  for some $\lambda > 0$, where $a = (a^{kl})$ is seen as an $n \times n$ matrix and 1_n is the $n \times n$ unit matrix.

This is very good, and if we can solve it, the metric in arbitrary coordinates is obtained by undoing the transformation from these coordinates to harmonic coordinates. The transformation to harmonic coordinates is time-dependent, as is its undoing, since it is done with respect to the metric $g(t)$ defining Δ_g in (4.23). This is no problem, but it *is* sometimes problematic that the method is restricted to a single coordinate patch, with no easy transition to other patches (this is because of the ugly property that the i in x^i is not treated as a vector index, like the i in Γ^i).

To remedy this, one may start from a fixed background metric $\hat{g}$ on M, and replace (4.28) by

$$\tilde{\Gamma}^i \equiv g^{kl}\tilde{\Gamma}^i_{jk} = 0, \qquad\qquad (\tilde{\Gamma}^i_{jk} = \Gamma^i_{kl} - \hat{\Gamma}^i_{kl}), \tag{4.33}$$

where $\hat{\Gamma}^i_{kl}$ are the Christoffel symbols for the metric $\hat{g}$. The crucial difference compared to (4.27) - (4.28) is that the $\tilde{\Gamma}^i_{jk}$ are the components of a tensor (cf. §4.1), so that (4.33) is coordinate-independent. It is, therefore, not a condition on the *coordinates* but a condition on the *metric* (called a *gauge* in the physics literature).[70] Without any conditions, eq. (4.30) is replaced by

$$\begin{aligned} -2R_{ij} &= \Delta_g g - g_{ik}\partial_j\tilde{\Gamma}^k - g_{jk}\partial_i\tilde{\Gamma}^k + \tilde{O}(g,\partial g) \\ &= \Delta_g g - \nabla_i\tilde{\Gamma}_j - \nabla_j\tilde{\Gamma}_i + \tilde{O}'(g,\partial g), \end{aligned} \tag{4.34}$$

where $\tilde{\Gamma}_i = g_{ij}\tilde{\Gamma}^j$ (recall that $\tilde{\Gamma}$ is a vector). This follows from (4.29), since $\partial_j\tilde{\Gamma}^k$ and $\partial_i\tilde{\Gamma}^k$ in the first line involve second derivatives of $\hat{g}$ but not of g, and the terms $\nabla_i\tilde{\Gamma}_j$ and $\nabla_j\tilde{\Gamma}_i$ in the second line (to be used shortly) differ from those just mentioned in the first line by first derivatives of g only. Therefore, for metrics $\overline{g}$ satisfying (4.33), the Ricci Flow equation (1.1) becomes

$$\partial_t\overline{g} = \Delta_{\overline{g}}\overline{g} + \tilde{O}'(\overline{g},\partial\overline{g}), \tag{4.35}$$

with $\overline{g}(0) = g_0$. Of course, as in (4.30), one may also replace the covariant Laplacian Δ_g by the coordinate Laplacian Δ, since the two differ by first derivatives of the metric, which enter the $O(g,\partial g)$ terms. Either way, one again obtains a quasilinear second-order parabolic PDE.

Compared to the original use of harmonic coordinates, coordinate dependence has now been traded for background metric (i.e. $\hat{g}$) dependence; if in the second approach one makes a choice of coordinates and of $\hat{g}$ such that $\hat{g}$ is the flat (Euclidean) metric, the first approach is recovered.

[70] A gauge condition selects a unique metric $\overline{g}$ from the set of metrics ψ^*g, where ψ ranges among all diffeomorphisms of M. Cf. Choquet-Bruhat (2009), §VI.7.4.

To complete this coordinate-free approach, we proceed to show how a metric $\overline{g}$ that solves (4.35) can be transformed into a metric g that solves (1.1) with the same initial condition.[71]

To this end, let $\overline{g}(t)$ solve (4.35) and define $g(t)$ by

$$g(t) = \psi_t^* \overline{g}(t), \tag{4.36}$$

where ψ_t is a time-dependent diffeomorphism to be constructed in such a way that $g(t)$ solves (1.1), with the same initial condition. First, eq. (4.36) implies, using some differential geometry,

$$\partial_t g_{ij}(t) = \psi_t^*(\partial_t \overline{g}_{ij}(t) + \mathscr{L}_{\xi(t)} \overline{g}_{ij}(t)) = \psi_t^*(\partial_t \overline{g}_{ij}(t) + \overline{\nabla}_i \xi_j + \overline{\nabla}_j \xi_i), \tag{4.37}$$

where $\overline{\nabla}$ is the Levi-Civita connection for the metric $\overline{g}(t)$, and $\mathscr{L}_{\xi(t)}$ is the Lie derivative with respect to the time-dependent vector field

$$\xi(t) = d\psi_t/dt, \tag{4.38}$$

again to be chosen so that $g(t)$ solves (1.1), which we rewrite in terms of $\overline{g}(t)$ via (4.36). Thus

$$\partial_t g(t) = -2\mathrm{Ric}(g(t)) = -2\mathrm{Ric}(\psi_t^* \overline{g}) = -2\psi_t^* \mathrm{Ric}(\overline{g}) \equiv -2\psi_t^* \overline{\mathrm{Ric}}. \tag{4.39}$$

Comparing (4.37) and (4.39), we see that $g(t)$ solves (1.1) iff the metric $\overline{g}(t)$ solves

$$\partial_t \overline{g}_{ij} = -2\overline{R}_{ij} - \overline{\nabla}_i \xi_j - \overline{\nabla}_j \xi_i. \tag{4.40}$$

To relate this to (4.35), we use the identity (4.34) for $\overline{g}$ instead of g, so that $R_{ij} = \mathrm{Ric}(g)_{ij}$ is replaced by $\overline{R}_{ij} = \mathrm{Ric}(\overline{g})_{ij}$ and $\tilde{\Gamma}_i$ is replaced by $\tilde{\Gamma}'_i = \overline{\Gamma}^i - \hat{\Gamma}^i$. Thus (4.34) rewrites (4.40) as

$$\partial_t \overline{g}_{ij} = \Delta_g g_{ij} - \overline{\nabla}_i (\tilde{\Gamma}'_j + \xi_j) - \overline{\nabla}_j (\tilde{\Gamma}'_i + \xi_i) + \tilde{O}'(g, \partial g), \tag{4.41}$$

Clearly, eqs. (4.35) and (4.41) coincide (as they must for this whole approach to make sense) iff

$$\xi_i = -\tilde{\Gamma}'_i, \tag{4.42}$$

from which ψ_t is determined by solving (4.38) with initial condition $\psi_0 = \mathrm{id}_M$. Thus the Ricci Flow $g(t)$ may be found, in principle, by first solving (4.35) and then following the above steps.

The point is not to carry out any of this in practice; it is rather that the quasilinear second-order parabolic PDE (4.35) is of a type for which short-time existence and uniqueness of the solution are known,[72] whereas (1.1) is a weakly parabolic PDE, for which such results were, initially, not available.[73] In conclusion, short-time existence and uniqueness of the solution of (1.1) now follow either locally from the corresponding results in harmonic coordinates, or globally from the same results for $\overline{g}(t)$, from which $g(t)$ can be constructed by (4.36).

---

[71] The subsequent argument is taken from Topping (2006), §5.2 and Brendle (2010), Chapter 2. An alternative way to complete the argument is to show that any metric can be brought into a metric satisfying the gauge condition (4.33) through a suitable coordinate transformation. This is done through the introduction of *generalized harmonic coordinates*, which by definition satisfy $\Delta_g x^i = g^{kl} \hat{\Gamma}^i_{jk} \partial_k x^j \partial_l x^k$. The relationship between these different ways of completing the argument lies in the fact that a diffeomorphism may at least locally be described as a change of coordinates: if (U, φ) is a chart, consisting of open set $U \subset M$ and coordinates $\varphi :\to \mathbb{R}^n$, and $\psi : U \to U$ is a local diffeomorphism, then $(U, \tilde{\varphi} = \varphi \circ \psi)$ is a new chart and hence $\tilde{\varphi} :\to \mathbb{R}^n$ are new coordinates on U. And *vice versa*.

[72] Though there is still at least a pedagogical gap in the literature, which we shall try to close later in these notes.

[73] Hence Hamilton's (1982)'s proof of this without DeTurck's trick was heroic, but very complicated.

4.3 Normalized Ricci Flow

We now show that existence and uniqueness results for Ricci Flow transfer to ***normalized Ricci Flow***, defined by (4.22). As in the literature we write r or $r(g)$ for $\langle R(g)\rangle$, so that (4.22) reads

$$\partial_t g_{ij} = \frac{2}{n} r(g) g_{ij} - 2R_{ij}, \tag{4.43}$$

or, in $d = 2$, as already stated in (1.21) in connection with Yamabe flow,

$$\partial_t g_{ij} = (r(g) - R(g)) \cdot g_{ij}. \tag{4.44}$$

We have already mentioned that, as an immediate consequence of (4.4) - (4.6):

Proposition 4.2 *Normalized Ricci Flow preserves volume (evolving under Ricci Flow).*

Another, even more immediate fact is that:

Proposition 4.3 *The stationary points of normalized Ricci Flow are Einstein metrics.*[74]

Indeed, one obtains $\lambda = r(g)/n$ in (1.13). Yet another perspective on (4.43) is given by:

Proposition 4.4 *Assume M compact. For any $t \in [0,T]$ for which Ricci Flow exists, define*

$$C(t) = \left(\frac{\mathrm{Vol}(M)(0)}{\mathrm{Vol}(M)(t)}\right)^{2/n}; \tag{4.45}$$

$$\tau(t) = \int_0^t ds\, C(s), \tag{4.46}$$

so that $d\tau/dt = C(t)$. Then the rescaled and reparametrized metric

$$\hat{g}(\tau) = C(t(\tau))g(t(\tau)) \tag{4.47}$$

satisfies normalized Ricci Flow (in τ) iff $g(t)$ satisfies Ricci Flow (in t).

Proof. We write $\tilde{g}(t) = C(t)g(t)$, so that $\hat{g}(\tau) = \tilde{g}(t(\tau))$. The key step is the computation

$$\frac{dC(t)}{dt} = \frac{2}{n} r(g(t))C(t) = \frac{2}{n} r(\tilde{g}(t))C(t)^2, \tag{4.48}$$

which follows from (4.4) - (4.6) and the fact that under x-invariant conformal rescalings of the metric the scalar curvature scales as $R(Cg) = C^{-1}R(g)$.[75] In particular, $R(\tilde{g}) = C^{-1}R(g)$. Then:

$$\begin{aligned}\frac{d\tilde{g}(t)}{dt} &= \frac{d}{dt}(C(t)g(t)) = \frac{2}{n} r(\tilde{g}(t))C(t)^2 g(t) + C(t)\frac{dg(t)}{dt} \\ &= C(t)\left(\frac{2}{n} r(\tilde{g}(t))\tilde{g}(t) - 2\mathrm{Ric}(\tilde{g}(t))\right),\end{aligned} \tag{4.49}$$

where in the last step we used (1.1) as well as the property $\mathrm{Ric}(g(t)) = \mathrm{Ric}(\tilde{g}(t))$. Finally,

$$\frac{d\hat{g}(\tau)}{d\tau} = \frac{dt}{d\tau}\frac{d\tilde{g}(t)}{dt} = \frac{2}{n} r(\tilde{g}(t))\tilde{g}(t) - 2\mathrm{Ric}(\tilde{g}(t)) = \frac{2}{n} r(\hat{g}(\tau))\hat{g}(\tau) - 2\mathrm{Ric}(\hat{g}(\tau)). \qquad \square$$

[74] This suggests that general relativity emerges from some diffusion process ruled by normalized Ricci Flow.

[75] Since $\mathrm{Ric}(Cg) = \mathrm{Ric}(g)$, $R = g^{ij}R_{ij}$, and g^{ij} is the inverse to g_{ij}, the scalar curvature R picks up C^{-1}.

4.4 Ricci solitons

Ricci solitons form a special and historically important class of solutions of Ricci Flow. The motivation comes from Einstein metrics as initial data, i.e., $\mathrm{Ric}(g) = \lambda g$, cf. (1.13), in which case (1.1) can be solved immediately by (1.14). More generally, suppose that $g = g(0)$ satisfies

$$\mathrm{Ric}(g) = \lambda g - \tfrac{1}{2}\mathscr{L}_Y g, \tag{4.50}$$

where Y is some vector field on M. This time, the solution to (1.1) - (1.2) is

$$g(t) = (1-2\lambda t)\psi_t^* g, \tag{4.51}$$

where ψ_t is the flow of the time-dependent vector field

$$X(t) = \frac{Y}{1-2\lambda t}. \tag{4.52}$$

This is even simpler in $d = 2$, where $R_{ij} = \frac{1}{2}g_{ij}R$, and hence (4.50) becomes

$$\mathscr{L}_Y g = (2\lambda - R)g. \tag{4.53}$$

An example of a metric satisfying this equation is in $M = \mathbb{R}^2$, with

$$g(x,y) = \frac{dx^2+dy^2}{1+x^2+y^2} = \frac{dr^2+r^2d\theta^2}{1+r^2}, \tag{4.54}$$

where $x = r\cos\theta$ and $y = r\sin\theta$ as usual. Simple computations (exercise) show that:[76]

$$\mathrm{Ric}(g)(x,y) = \frac{2g}{1+x^2+y^2} = 2\frac{dx^2+dy^2}{(1+x^2+y^2)^2}; \qquad R(g)(x,y) = \frac{4}{1+x^2+y^2}; \tag{4.55}$$

$$Y(x,y) = -2(x\partial_x + y\partial_y), \qquad Y(r,\theta) = -2r\partial_r; \tag{4.56}$$

$$\mathscr{L}_Y g = -R(g)g, \qquad \lambda = 0, \tag{4.57}$$

which of course is consistent with (2.44). Hence $X(t) = Y$ is t-independent, and its flow ψ_t is given by solving $d\psi_t(x,y)/dt = Y(\psi_t(x,y))$ with initial condition $\psi_0(x,y) = (x,y)$. This gives

$$\psi_t(x,y) = (e^{-2t}x, e^{-2t}y); \qquad \psi_t(r,\theta) = (e^{-2t}r, \theta). \tag{4.58}$$

Finally, the Ricci Flow of the metric (4.54) is given by (4.51) with $\lambda = 0$, i.e. $g(t) = \psi_t^* g$, viz.

$$g(t,x,y) = \frac{dx^2+dy^2}{e^{4t}+x^2+y^2} \qquad g(t,r,\theta) = \frac{dr^2+r^2d\theta^2}{e^{4t}+r^2}. \tag{4.59}$$

This Ricci soliton is called Hamilton's ***cigar***. It exists for all $t \in \mathbb{R}$; for $t \to -\infty$ it converges to a singular metric that blows up at $(0,0)$, and for $t \to \infty$ it becomes singular in being zero.

[76] One should use the following formulae for the Lie derivative of the metric w.r.t. Y and its pullback under ψ:

$$\mathscr{L}_Y g_{ij} = Y^k\partial_k g_{ij} + (\partial_i Y^k)g_{jk} + (\partial_j X^k)g_{ik};$$

$$\psi^* g_{ij}(x) = \left(\frac{\partial\psi^k}{\partial x^i}\frac{\partial\psi^l}{\partial x^j}g_{kl}\right)(\psi(x)).$$

5 Maximum and minimum principles

The maximum and minimum principles originate in elliptic PDEs of the kind

$$Pu = 0; \tag{5.1}$$

$$P = a^{ij}(x)\partial_i\partial_j + b^i(x)\partial_j, \tag{5.2}$$

where $u : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ is bounded, open, and connected, and the matrix $a(x)$ is strictly positive at each $x \in \Omega$.[77] The *weak* max and min 'principles' (which are theorems) state that

$$Pu \geq 0 \qquad \Rightarrow \qquad \max_{x\in\Omega} u(x) \leq \max_{x\in\partial\Omega} u(x); \tag{5.3}$$

$$Pu \leq 0 \qquad \Rightarrow \qquad \min_{x\in\partial\Omega} u(x) \leq \min_{x\in\Omega} u(x). \tag{5.4}$$

In particular, both the weak minimum and the maximum principles apply to the PDE (5.1), so that any solution of (5.1) assumes both its minimum and its maximum on the boundary.[78] We only prove (5.3), from which (5.4) follows by changing u to $-u$ and reverting inequalities.

We know from calculus that the condition for u to assume a local maximum at $x_0 \in \Omega$ is

$$\partial_i u(x_0) = 0; \tag{5.5}$$

$$\partial_i\partial_j u(x_0) \leq 0, \tag{5.6}$$

where the second condition regards $\partial_i\partial_j u(x_0) \equiv D$ as a matrix and the inequality means $D \leq 0$. The proof would be easy if the assumption had been $Pu > 0$ instead of $Pu \geq 0$, since, writing $A = a(x_0)$, the former implies $\mathrm{Tr}\,(AD) > 0$ at, contradicting the ellipticity assumption $A > 0$ and the condition $D \leq 0$ for a local maximum. Hence (5.6) must fail and so u cannot have any local maximum within Ω. Since it must take a maximum somewhere on the compact space $\Omega \cup \partial\Omega$, (5.3) follows. However, our starting point is $Pu \geq 0$ rather than $Pu > 0$. Take $\varepsilon > 0$ and define

$$u_\varepsilon(x) = u(x) + \varepsilon f(x), \tag{5.7}$$

where f is a function such that $Pf(x) > 0$ for all $x \in \Omega$.[79] Therefore, if u satisfies $Pu \geq 0$, then by linearity, $Pu_\varepsilon = Pu + \varepsilon Pf > 0$, and hence by the above argument, (5.3) holds for u_ε. Since $\lim_{\varepsilon\to 0} u_\varepsilon(x) = u(x)$ pointwise and (5.3) is an inequality, it also holds for $u = u_0$.

Similar principles apply to parabolic PDEs. We keep (5.2) with $a >$ and modify (5.1) to

$$Pu = \partial_t u. \tag{5.8}$$

The ***parabolic boundary*** $\mathscr{P}\Omega$ of $\Omega \times (0,T)$, where $\Omega \subset \mathbb{R}^n$ as above and $T > 0$, is defined by

$$\mathscr{P}\Omega = (\partial\Omega \times [0,T]) \cup (\Omega \times \{0\}); \tag{5.9}$$

here $[0,T]$ is the time interval on which we study solutions of (5.8). It lacks the part $\Omega \times \{T\}$ that would make it the boundary of $\Omega \times (0,T)$, but one has the following max-min 'principle':

---

[77] That is, $a(x) \geq \lambda(x) \cdot 1_n$ for some $\lambda(x) > 0$, where $c \geq 0$ for some $c \in M_n(\mathbb{C})$ iff $c^{ij} v_i v_j \geq 0$ for all $v \in \mathbb{R}^n$.

[78] Though we will not need this, it is worth stating that *strong* maximum or minimum principle states that we have strict inequalities in (5.3) and (5.4) unless u is constant. Every PDE book discusses these things, for example Renardy & Rogers (2004).

[79] For example, $f(x) = \exp(cx_1)$ gives $Pf(x) = (c^2 a^{11}(x) + cb_1(x))\exp(cx_1)$. Since $a^{11}(x) > 0$ by ellipticity of P, this choice of f indeed gives $Pf(x) > 0$ for sufficiently large $c > 0$.

Proposition 5.1 *If $\partial_t u \leq Pu$ ($\partial_t u \geq Pu$), then u assumes its maximum (minimum) on $\mathscr{P}\Omega$. Hence if $\partial_t u = Pu$, then u assumes both its minimum and its maximum on $\mathscr{P}\Omega$.*

If u is to have a maximum at $(t_0, x_0) \in \Omega \times (0,T)$, then the spatial condition (5.6) of course still stands, but because of the boundaries at $t = 0$ and $t = T$, in time the conditions are:

$$\partial_t u(0,x_0) \leq 0; \tag{5.10}$$
$$\partial_t u(t_0,x_0) = 0, \ \forall t_0 \in (0,T); \tag{5.11}$$
$$\partial_t u(T,x_0) \geq 0. \tag{5.12}$$

If $t_0 \in (0,T]$, then we obtain a contradiction: if $\partial_t u < Pu$, then (5.6) yields $\partial_t u(t_0,x_0) < 0$, which excludes maxima inside $\Omega \times (0,T)$ as well as on $\Omega \times \{T\}$. If $\partial_t u \leq Pu$, a simple ε-argument uses $u_\varepsilon(t,x) = u(t,x) + \varepsilon f(t)$, as in (5.7), this time with $f(t,x) = \exp(-t)$, so that

$$\partial_t u_\varepsilon = \partial_t u - \varepsilon \exp(-t) \leq Pu - \varepsilon \exp(-t) = Pu_\varepsilon - \varepsilon \exp(-t) < Pu_\varepsilon,$$

hence $\partial_t u_\varepsilon < Pu_\varepsilon$, and one finishes the argument exactly as in the elliptic case. For the minimum principle the inequalities in (5.6), (5.10), and (5.12) are reverted, but the argument is the same.

Without proof, we mention that for geometric PDEs this argument also works for manifolds, including the case where Ω has no boundary (e.g. when it is compact, or rather closed). Then u assumes both its minimum and its maximum on $\Omega \times \{0\}$ and Proposition 5.1 comes down to

$$u(0,x) \leq u(t,x) \leq u(0,x), \tag{5.13}$$

so that $u(t,x) = u(0,x)$ for all t. Fortunately, this dull result applies neither to Ricci Flow itself (where $P = \Delta_g$ is such that the matrix $a^{ij} = g^{ij}$ depends on the unknown u, viz. the g_{ij}), nor to its derived equations (see §4.1), which are more general nonlinear reaction-diffusion equations

$$\partial_t u = Pu + F(u), \tag{5.14}$$

and/or associated inequalities

$$\partial_t u \geq Pu + F(u); \tag{5.15}$$
$$\partial_t u \leq Pu + F(u). \tag{5.16}$$

Our paradigmatic example will be the elliptic geometric differential operator

$$Pu = \Delta_g + g(X, \nabla u), \tag{5.17}$$

where X is some vector field on M (which is often zero); as will be clear from its proof, the fundamental Theorem 5.2 below even applies to the case where u is some curvature invariant (seen as a complicated function of the metric g, like the differential operator Δ_g, which exceeds the setting of Proposition 5.1, where P is independent of u). The simplest cases are

$$\partial_t R = \Delta_g R + R^2 \ \ (n = 2); \tag{5.18}$$
$$\partial_t R \geq \Delta_g R + \frac{2}{n} R^2 \ \ (n \geq 2). \tag{5.19}$$

Eq. (5.18) is eq. (4.1) in $2d$, cf. (2.44), whereas (5.19) follows from (4.1) and the inequality

$$R_{ij}R^{ij} = E_{ij}E^{ij} + R^2/n \geq R^2/n, \tag{5.20}$$

where E_{ij} is the Einstein tensor (2.41), so that $R_{ij} = E_{ij} + g_{ij}R/n$, and $E_{ij}E^{ij} = \mathrm{Tr}\,(E^2) \geq 0$.

Theorem 5.2 • Minimum principle*: Let $\alpha(t)$ solve $\partial_t\alpha = F(\alpha)$ with $\alpha(0) = \alpha_0$. If* (5.15) *holds with $u(0,x) \geq \alpha_0$ for all $x \in \Omega$, then $u(t,x) \geq \alpha(t)$ for all $x \in \Omega$ and $t \in [0,T]$.*

- Maximum principle*: Let $\beta(t)$ solve $\partial_t\beta = F(\beta)$ with $\beta(0) = \beta_0$. If* (5.16) *holds with $u(0,x) \leq \beta_0$ for all $x \in \Omega$, then $u(t,x) \leq \beta(t)$ for all $x \in \Omega$ and $t \in [0,T]$.*

- *In particular, if $u(t,x)$ solves* (5.14)*, if $\alpha_0 \leq \beta_0$, and if $u(0,x)$ for all $x \in \Omega$ satisfies*

$$\alpha_0 \leq u(0,x) \leq \beta_0, \tag{5.21}$$

then the solution $u(t,\cdot)$ satisfies, for all $x \in \Omega$ and all $t \in [0,T]$,

$$\alpha(t) \leq u(t,x) \leq \beta(t). \tag{5.22}$$

Once again, here $[0,T]$ is some time interval where all solutions α, β, and u exist. Of course, the earlier min and max principles follows from this by taking $F = 0$, for in that case, $\alpha(t)$ and $\beta(t)$ are constant and the boring eq. (5.13) is recovered. Before proving this theorem, let us apply it to (5.19), from which the $2d$ case (5.18) will also be obvious. For (5.19) we evidently obtain a minimum principle. The function $\alpha(t)$ satisfies $\partial_t\alpha = 2\alpha^2/n$, and hence,

$$\alpha(t) = \frac{\alpha_0}{1 - 2\alpha_0 t/n}. \tag{5.23}$$

The minimum principle therefore gives

$$\forall_{x\in M} R(t,x) \geq \alpha_0 \qquad \Rightarrow \qquad \forall_{x\in M} R(t,x) \geq \frac{\alpha_0}{1 - 2\alpha_0 t/n}. \tag{5.24}$$

In particular, if $\alpha_0 > 0$ (positive curvature), then $R(t,x)$ will blow up at $t = T$, $T \leq n/2\alpha_0$.

Proof of Theorem 5.2. We just prove the maximum principle, from which the minimum principle follows by reversing all inequalities. As in the easier cases above, we initially assume

$$\partial_t u < Pu + F(u), \tag{5.25}$$

and later move from $<$ to $\leq$ with an ε-argument. We claim that if also $u(0,x) < \beta_0$ for all x, then $u(t,x) < \beta(t)$ for all x. Proof by contradiction: suppose there is $x_0 \in m$ and $t_0 \in (0,T]$ where $u(t_0,x_0) = \beta(t_0)$. We may assume that t_0 is the earliest time where this is the case, so that $u(t,x) \leq \beta(t)$ for all $t \in [0,t_0]$, and for (small) $h > 0$ we have $u(t_0 - h, x_0) < \beta(t_0 - h)$. Hence

$$u(t_0,x_0) > u(t_0 - h, x_0) - \beta(t_0 - h) + \beta(t_0), \tag{5.26}$$

in which the limit $h \to 0$ gives, on our *reductio ad absurdum* assumption $u(t_0,x_0) = \beta(t_0)$,

$$\partial_t u(t_0,x_0) \geq \partial_t\beta(t_0) = F(\beta(t_0)) = F(u(t_0,x_0)). \tag{5.27}$$

We are now back to the simpler cases treated earlier: at a maximum we have (5.5) - (5.6), so that $Pu(t_0,x_0) \leq 0$, from which (5.25) leads to $\partial_t u(t_0,x_0) < F(u(t_0,x_0))$, which contradicts (5.27). We leave the straightforward ε-argument that finishes the proof as an exercise. □

Let us give the simplest application of the minimum principle to (normalized) Ricci Flow that is actually used. In $d = 2$ (and M compact) we have (5.18) under Ricci Flow, and

$$\partial_t R = \Delta_g R + R(R - r) \tag{5.28}$$

under normalized Ricci Flow. Note in $d = 2$, by the Gauss–Bonnet Theorem (which makes $\int_M R$ constant and equal to the topological invariant $4\pi\chi$) and the time-independence of $\mathrm{Vol}(M)$ under Ricci Flow in any dimension (exercise) the average scalar curvature $r(g)$ is constant (in time) and hence equal to its initial value where $\mathrm{Vol}(M)$ is computed from $g(0)$, i.e.,

$$r = \frac{4\pi\chi}{\mathrm{Vol}(M)}. \tag{5.29}$$

The Uniformization Theorem (version 2, Theorem 1.2) therefore follows if we can prove that:[80]

1. The solution $g(t)$ of (5.28) exists for all t and has a limit as $t \to \infty$;
2. The curvature scalar $R(x,t)$ becomes constant in this limit (and converges to r).

This program is relatively easy to carry out for $\chi < 0$ and hence $r < 0$, in which case the bound

$$|R(x,t) - r| \leq Ce^{rt} \tag{5.30}$$

immediately gives the result.[81] This bound (which is true for any χ) consists of two parts,

$$R(x,t) \leq r + Ce^{rt}; \qquad\qquad R(x,t) \geq r - Ce^{rt}, \tag{5.31}$$

where $C > 0$, of which the last one, which for $r < 0$ takes the simpler form

$$R(x,t) \geq -Ce^{rt}, \tag{5.32}$$

will now be proved. By (5.28) and the obvious property $\Delta_g r = 0$, the quantity $\rho = R - r$ satisfies

$$\partial_t \rho = \Delta_g \rho + \rho^2 + r\rho \geq \Delta_g \rho + r\rho, \tag{5.33}$$

so that the minimum principle can be applied. Thus we solve $\partial_t \alpha(t) = r\alpha(t)$ by

$$\alpha(t) = \alpha_0 \exp(rt), \tag{5.34}$$

and just note that because M is compact, we have $\rho(0) \geq -C$ for some $C > 0$. The minimum principle then immediately gives (5.32). Alas, the first bound in (5.31) is far more difficult.[82]
Finally, Theorem 5.2 can be extended to tensorial quantities, for which Hamilton proved:[83]

Theorem 5.3 *Let u be a tensor satisfying* (5.14)*, where $u(t,x) \in V_x = (T_x^*M)^{\otimes k} \otimes (T_xM)^{\otimes l}$, and suppose $V_x \cong V$ canonically. Let the V-valued function $v(t)$ solve $\partial_t v = F(v)$, and let $K \subset V$ be convex. If $u(0,x) \in K$ for all $x \in M$, then $u(t,x) \in K(t)$ for all x, where $K(t) = \{v(t) \mid v(0) \in K\}$.*

Theorem 5.2 is then the special case where $V = \mathbb{R}$ and $K = [\alpha_0, \beta_0]$ is an interval. This general result can often be avoided, since estimates for the curvature tensors can usually be derived from estimates for associated scalars, such as R, $R_{ij}R^{ij}$, or $R_{ijkl}R^{ijkl}$, etc., which in turn give information about the underlying metric. However, it will be useful in the next chapter.

[80] In $d = 2$ any solution $g(t)$ to (5.28) is conformal to $g(0)$. This follows from the fact that the conformal *Ansatz* (1.25) with $p = 3$ leads to *a* solution to (5.28), which, by uniqueness, must then be *the* solution.

[81] This relies on the nontrivial fact that if $R(x,t)$ is uniformly bounded, then the solution $g(t)$ exists for all t.

[82] This is true even for $\chi < 0$, but the case $\chi > 0$ is exceedingly difficult. See Chow & Knopf (2004), Chapter 5.

[83] We just state this for the case that TM and hence all other tensor bundles are trivial, as is the case in $d = 3$. For good discussions see Brendle (2010), Chapter 5, Topping (2006), Chapter 9, Bennett & Chow (2004), Chapter 4.

6 Before Perelman: Hamilton's Theorem

Historically, an important step that gave confidence in the idea that Ricci Flow could lie at the basis of a proof of the Poincaré and Geometrization Conjectures is ***Hamilton's Theorem***:[84]

Theorem 6.1 *Let M be a compact 3-manifold that admits a Riemannian metric g for which* $\mathrm{Ric}(g) > 0$. *Then M also admits a metric with positive constant curvature.*

Here the assumption $\mathrm{Ric}(g) > 0$, called *strictly positive Ricci curvature*, means that there exists $\varepsilon > 0$ for which $\mathrm{Ric} \geq \varepsilon g$, which in turns means that for each $x \in M$ and $X_x \in T_xM$ we have $\mathrm{Ric}_x(X_x, X_x) \geq \varepsilon g_x(X_x, X_x)$. If M is also simply connected, then $M \cong S^3$ by Theorem 3.10 (since the other two cases are not compact), so the Poincaré Conjecture follows if we can prove that any compact simply connected 3-manifold admits a Riemannian metric g for which $\mathrm{Ric}(g) > 0$ (which Hamilton was unable to prove; it followed two decades later from the work of Perelman).

Unfortunately, it is beyond the scope of these notes to prove Theorem 6.1, but we will play around it in a way that hopefully gives some insight into the Ricci Flow program.

6.1 Ricci Flow of $SU(2)$-invariant metrics on S^3

Starting with $M = S^3$ is somewhat circular in so far as applications of Ricci Flow to the Poincaré Conjecture are concerned, especially if we start with a metric that already has a high degree of symmetry, but it is instructive all the same. Our goal is to show that under normalized Ricci Flow any left-invariant metric on S^3 converges to the "round" (constant curvature) metric.[85]

The round metric on S^3 has a group-theoretic description, which is explained in §3.4 and is based on the identification (3.64) of S^3 with the Lie group $SU(2)$, which is given in detail by (C.1). In sum, the round metric on $G = SU(2)$ comes from (minus) the Cartan–Killing B form of $\mathfrak{g} \cong T_eG$, which is translated to all of G by the left-action of G on itself, and which is also right-invariant because B is $\mathrm{Ad}(G)$-invariant. Moreover, it is the unique $G \times G$-invariant metric on G (up to rescaling by a positive constant, as usual). In the case at hand, we have $\mathfrak{g} \cong \mathbb{R}^3$ and under this isomorphism $-B$ is just the usual inner product on $\mathbb{R}^3$. However, *any* inner product $\langle\, , \rangle$ on $\mathfrak{g}$ gives a left-invariant (but typically not right-invariant) metric on G by left translation (i.e. $g_\gamma = L^*_{\gamma^{-1}} g_e$, where $L_\gamma \delta = \gamma\delta$), and it is such a more general metric on S^3 that we start with.

Theorem 6.2 *Under normalized Ricci Flow any left-invariant metric on $SU(2)$ converges to a (positive) multiple of the round metric as $t \to \infty$, and hence to a metric with constant curvature.*

All computations are based on the following lemma (Milnor, 1976), which applies to $\mathfrak{su}(2)$:

Lemma 6.3 *If $\mathfrak{g}$ is a 3d Lie algebra such that* $\mathrm{Tr}(\mathrm{ad}(X)) = 0$ *for each $X \in \mathfrak{g}$ (that is, if $\mathfrak{g}$ is* unimodular*), then $\mathfrak{g}$ has a basis (e_1, e_2, e_3) with respect to which the Lie bracket is given by*

$$[e_1, e_2] = \lambda_3 e_3; \qquad [e_2, e_3] = \lambda_1 e_1; \qquad [e_3, e_1] = \lambda_2 e_2 \qquad (\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}), \tag{6.1}$$

and this basis may be chosen to be orthonormal with respect to any given inner product on $\mathfrak{g}$.

[84] This theorem is due to Hamilton (1982), but his proof was simplified later. See Chow & Knopf (2004, Chapter 6), Brendle (2010, Chapters 5 and 6), and Sheridan (2006, Chapter 7) for complete and up-to-date treatments.

[85] See also Isenberg & Jackson (1992), Chow & Knopf (2004, Chapter 1), and Sheridan (2006, Chapter 7).

For later use, we note that the case where $G = SU(2)$ and $\lambda_1 = \lambda_2 = \lambda_3 = \lambda > 0$ corresponds to the round metric on S^3 (up to rescaling; for simplicity we assume $\lambda = 1$, and otherwise one rescales the basis vectors). For in that case, the relations (6.1) reproduce the Lie algebra of $SO(3)$, see (B.7) - (B.8). Using the basis (e_1, e_2, e_3) to identify $\mathfrak{g} = \mathfrak{su}(2) = \mathfrak{so}(3)$ with $\mathbb{R}^3$, the adjoint action of $SO(3)$ on its Lie algebra $\mathfrak{so}(3)$ is the defining action on $\mathbb{R}^3$ and the given inner product on this Lie algebra is the standard one on $\mathbb{R}^3$. This is invariant under the defining = adjoint action of $SO(3)$, and hence, by the general theory of §3.4, one obtains a left- and right invariant metric on $SU(2) \cong S^3$, and this metric is unique (up to scaling) by Proposition 3.7.

Proof. The proof is based on the isomorphism $\wedge^2\mathfrak{g} \cong \mathfrak{g}$ in $d = 3$, which turns the Lie bracket $[\,,\,] : \wedge^2\mathfrak{g} \to \mathfrak{g}$ into a map $\Lambda : \mathfrak{g} \to \mathfrak{g}$. In some initial basis (T_1, T_2, T_3) of $\mathfrak{g}$ one may take

$$[T_2, T_3] = \Lambda(T_1); \qquad [T_3, T_1] = \Lambda(T_2); \qquad [T_1, T_2] = \Lambda(T_3). \tag{6.2}$$

If the structure constants of $\mathfrak{g}$ in the given basis are defined by $[T_a, T_b] = C^c_{ab} T_c$, then

$$\Lambda = \begin{pmatrix} C^1_{23} & C^1_{31} & C^1_{12} \\ C^2_{23} & C^2_{31} & C^2_{12} \\ C^3_{23} & C^3_{31} & C^3_{12} \end{pmatrix}, \tag{6.3}$$

We now impose the unimodularity condition $\mathrm{Tr}\,(\mathrm{ad}(X)) = 0$ for each $X \in \mathfrak{g}$, with $\mathrm{ad}(T_a)^c_b = C^c_{ab}$ in the given basis. This gives $C^1_{a1} + C^2_{a2} + C^3_{a3} = 0$ for $a = 1, 2, 3$, hence

$$C^2_{12} = -C^3_{13} = C^3_{31}; \qquad C^1_{21} = -C^3_{23} = C^3_{32}; \qquad C^1_{31} = -C^2_{23} = C^2_{32}, \tag{6.4}$$

which gives $\Lambda^T = \Lambda$. Thus Λ is symmetric and hence diagonalizable. If (e_1, e_2, e_3) is a basis in which $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$, then (6.1) follows at once. Finally, if the initial basis (T_1, T_2, T_3) was orthogonal with respect to the given inner product, then, by the spectral theorem for hermitian matrices (which in this case are real, and hence symmetric), so is the final basis. □

The basis (e_1, e_2, e_3) of $\mathfrak{g} = T_eG$ defines a basis of T_xG at each $x \in G$ by left translation on G, and, by definition of a left-invariant metric, remains orthonormal. This means that, up to isometry, the eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ contain all information about the metric and describe it.

Proposition 6.4 *Let $g(t)$ solve (normalized) Ricci Flow with initial condition $g(0) = g$. Then any isometry of g is also an isometry of $g(t)$, i.e. the isometry group of $g(t)$ contains that of g.*[86]

Proof. For Ricci Flow this follows from the covariance property $\psi^*\mathrm{Ric}(g) = \mathrm{Ric}(\psi^*(g))$ under diffeomorphisms ψ, which implies that if $g(t)$ solves (1.1) with initial condition g, then $\psi^*g(t)$ solves (1.1) with initial condition ψ^*g. Therefore, if ψ is an isometry of g, i.e. $\psi^*g = g$, then both $\psi^*g(t)$ and $g(t)$ solve (1.1) with initial condition g. But since we have shown that solutions to (1.1) - (1.2) are unique, we must have $\psi^*g(t) = g(t)$. The normalized case is an exercise. □

As a significant application of this proposition (but an *intermezzo* in our story), we state:

Corollary 6.5 *A bi-invariant (i.e. left- and right-invariant) Riemannian metric on a Lie group is an Einstein metric (so in particular, in $d = 3$ it has constant curvature by Proposition 3.3).*

The proof is an exercise. This result applies, for example, to the round metric on S^3.

[86]Under boundedness assumptions on the curvature, the isometry groups of g and $g(t)$ even coincide. See B. Kotschwar, Backwards uniqueness for the Ricci flow, IMRN 21, 4064–4097 (2010) and `arXiv:0906.4920`.

Returning to our main story, Proposition 6.4 implies that if g is a left-invariant metric on g, then so is $g(t)$, and hence we can still use the basis (e_1, e_2, e_3) with associated eigenvalues $(\lambda_1, \lambda_2 \lambda_3)$. Because of the time-dependence of g, both the e_a and the λ_a now become time-dependent as well. It will simplify the computations if we replace $(\lambda_1, \lambda_2, \lambda_3)$ by (A, B, C) via

$$A = \lambda_1^2 \lambda_2 \lambda_3; \qquad B = \lambda_1 \lambda_2^2 \lambda_3; \qquad C = \lambda_1 \lambda_2 \lambda_3^2; \tag{6.5}$$

$$\lambda_1 = \frac{A}{\sqrt{ABC}}; \qquad \lambda_2 = \frac{B}{\sqrt{ABC}}; \qquad \lambda_3 = \frac{C}{\sqrt{ABC}}, \tag{6.6}$$

where for $SU(2)$ we assume all $\lambda_1, \lambda_2, \lambda_3 > 0$ (see the table at the end of Appendix B), and hence $A, B, C > 0$ also.[87] In terms of the variables (A, B, C), the Lie brackets (6.1) become

$$[e_1, e_2] = \frac{C}{\sqrt{ABC}} e_3; \qquad [e_2, e_3] = \frac{A}{\sqrt{ABC}} e_1; \qquad [e_3, e_1] = \frac{B}{\sqrt{ABC}} e_2, \tag{6.7}$$

which of course remains time-dependent. To remove this, we define a new basis (f_1, f_2, f_3) by

$$f_1 = \sqrt{A} e_1; \qquad f_2 = \sqrt{B} e_2; \qquad f_3 = \sqrt{C} e_3, \tag{6.8}$$

whose Lie bracket is time-independent, and time-dependence is transferred to the metric, viz.

$$[f_1, f_2] = f_3; \qquad [f_2, f_3] = f_1; \qquad [f_3, f_1] = f_2; \tag{6.9}$$

$$g(f_1, f_1) = A; \qquad g(f_2, f_2) = B; \qquad g(f_3, f_3) = C; \qquad g(f_a, f_b) = 0 \ (a \neq b). \tag{6.10}$$

Lemma 6.6 *The Ricci tensor is diagonal in the orthonormal basis* (f_1, f_2, f_3), *with entries*

$$R_{11} = \frac{A^2 - (B-C)^2}{2BC}; \qquad R_{22} = \frac{B^2 - (A-C)^2}{2AC}; \qquad R_{33} = \frac{C^2 - (A-B)^2}{2AB}. \tag{6.11}$$

Proof. This is a lengthy but straightforward computation, in which first $\nabla_{f_a} f_b$ is computed by finding its coefficients $g(f_c, \nabla_{f_a} f_b)$ in the (f_1, f_2, f_3) basis from the Lie brackets in (6.9) and the Koszul formula (2.12), in which just the commutators survive, so that we have

$$g(\nabla_{f_a} f_b, f_c) = -\tfrac{1}{2}(g(f_a, [f_b, f_c]) - g([f_a, f_b], f_c) - g(f_b, [f_c, f_a])). \tag{6.12}$$

Using (6.9), this makes $g(\nabla_{f_a} f_b, f_c)$ nonzero iff a, b, c are all different, in which case $\nabla_{f_a} f_b$ is therefore proportional to f_c. Using this fact in (2.17) easily gives the relations

$$R_{3132} = R_{2321} = R_{1312} = 0, \tag{6.13}$$

where $R_{abcd} = \mathrm{Riem}(f_a, f_b, f_c, f_d)$, plus similar results for the components of the Riemann tensor obtained by permutations, using (2.25). For the Ricci tensor (2.38) one then finds

$$R_{12} = R^1_{112} + R^2_{122} + R^3_{132} = A^{-1} R_{1112} + B^{-1} R_{2122} + C^{-1} R_{3132} = 0, \tag{6.14}$$

since $g^{ij} = \delta^{ij}$ in the e-basis and the only nontrivial part R_{3132} vanishes because of (6.13). Similarly, $R_{23} = 0$ and $R_{13} = 0$. Eq. (6.11) follows from a simple computation (exercise). □

For the curvature scalar we have $R = A^{-1} R_{11} + B^{-1} R_{22} + C^{-1} R_{33}$, so that (6.11) gives

$$R = \frac{-\frac{1}{2}(A^2 + B^2 + C^2) + AB + AC + BC}{ABC}. \tag{6.15}$$

[87] For the other case one obviously needs to introduce some signs in (6.5) - (6.6), since always $A, B, C > 0$.

Since R is independent of x, it equals $r(g)$. Eq. (4.22) then takes the form, for $a = 1,2,3$,

$$\partial_t g(f_a, f_a) = -2(\mathrm{Ric}(f_a, f_a) - \tfrac{1}{3} g(f_a, f_a) R), \tag{6.16}$$

where $\partial_t g(f_1, f_1) = dA/dt$, $\partial_t g(f_2, f_2) = dB/dt$, and $\partial_t g(f_3, f_3) = dC/dt$, so that we obtain

$$\frac{dA}{dt} = \tfrac{2}{3}\left(\frac{-2A^2 + (B-C)^2 + A(B+C)}{BC}\right); \tag{6.17}$$

$$\frac{dB}{dt} = \tfrac{2}{3}\left(\frac{-2B^2 + (A-C)^2 + B(A+C)}{AC}\right); \tag{6.18}$$

$$\frac{dC}{dt} = \tfrac{2}{3}\left(\frac{-2C^2 + (A-B)^2 + C(A+B)}{AB}\right), \tag{6.19}$$

where $A = A(t)$, $B = B(t)$, $C = C(t)$. It follows from these equations that $d(ABC)/dt = 0$, which reflects the invariance of $\mathrm{Vol}(M)$ under normalized Ricci Flow.[88] We may therefore set $ABC = 1$. Since the round metric has $\lambda_1 = \lambda_2 = \lambda_3 = 1$, our aim is to show that $\lambda_a - \lambda_b \to 0$, i.e., $|A-B| \to 0$, $|B-C| \to 0$, and $|A-C| \to 0$. To this end we rewrite (6.17) - (6.19) as

$$\frac{d}{dt}(A-B) = -\tfrac{2}{3}(A-B)\cdot(2B^2 + 2AB + (A-C)(A+B) + (A^2 - C^2)); \tag{6.20}$$

$$\frac{d}{dt}(B-C) = -\tfrac{2}{3}(B-C)\cdot(2A^2 + 2AC + (B-C)(B+A) + (B^2 - C^2)); \tag{6.21}$$

$$\frac{d}{dt}(A-C) = -\tfrac{2}{3}(A-C)\cdot(2C^2 + 2AC + (A-B)(A+C) + (A^2 - B^2)); \tag{6.22}$$

These equations are symmetric in A, B, and C, and so without loss of generality we may assume that $A(0) \geq B(0) \geq C(0) > 0$. Then (6.20) shows that $A(t) \geq B(t)$ for all t (for which the solution is defined), for if $A(t_0) = B(t_0)$ at some $t_0 \geq 0$, then all derivatives of $A(t) - B(t)$ vanish at t_0 and hence $A(t) = B(t)$ for all t. Similarly, eq. (6.21) gives $B(t) \geq C(t)$. If we rewrite (6.19) as

$$\frac{d}{dt}C = \tfrac{2}{3}C\cdot(C(A+B-2C) + (A-B)^2), \tag{6.23}$$

the previous two inequalities give $(C(A+B-2C) + (A-B)^2) \geq 0$ at any t, whence $dC/dt \geq 0$. Hence $C(t) \geq C(0) > 0$. Knowing, then, that $A(t) \geq B(t) \geq C(t) > 0$, in (6.22) we may estimate

$$(2C^2 + 2AC + (A-B)(A+C) + (A^2 - B^2)) \geq 2C(t)^2 \geq 2C(0)^2 > 0, \tag{6.24}$$

so that (6.22) gives $d(A-C)/dt \leq -\tfrac{4}{3}C(0)^2(A-C)$, and hence

$$|A(t) - C(t)| \leq |A(0) - C(0)| \cdot e^{-\frac{4}{3}C(0)^2 t}. \tag{6.25}$$

Since $B(t)$ lies between $A(t)$ and $C(t)$, or equally well using (6.21) and (6.22), eq. (6.25) gives

$$\lim_{t\to\infty} |A(t) - C(t)| = \lim_{t\to\infty} |B(t) - C(t)| = \lim_{t\to\infty} |A(t) - B(t)| = 0, \tag{6.26}$$

provided the solution exist for all t. But this is the case, for it follows from the constancy of $A(t)B(t)C(t)$ that the solutions $A(t)$, $B(t)$, and $C(t)$ are bounded, so that (6.26) implies that $A(t)$, $B(t)$, and $C(t)$ must each converge to the same positive constant. Returning to (6.6), this means that $\lambda_a(t) \to \lambda$ for $a = 1,2,3$, where $\lambda > 0$. As explained after Lemma 6.3, this means that normalized Ricci Flow converges to the round (= constant curvature) metric on S^3 (this is true even if the initial metric does not satisfy the assumptions of Hamilton's Theorem 6.1). □

[88] This can also be shown directly, since in suitable coordinates adapted to the frame (f_1, f_2, f_3), the quantity ABC is the determinant of the metric, which gives the volume of M.

6.2 Ricci Flow on general compact 3-manifolds

In general, there is no group action to transport some suitable basis of T_xM to other points. Nonetheless, through a trick it is still possible to mimic some of the analysis in the previous section.[89] First, since since the Ricci tensor, seen, for each $x \in M$, as a map $\widetilde{\mathrm{Ric}}_x : T_xM \to T_xM$ as in (2.45), is symmetric with respect to the inner product on T_xM provided by the metric, at $t = 0$ one can find a frame $(e_a(x))$ that diagonalizes both g_x and $\widetilde{\mathrm{Ric}}_x$. Find such a frame, where in coordinates one has $e_a(x) = e_a^i(x)\partial_i$. Next, we time-evolve the frame by solving

$$\frac{d}{dt}e_a^i(x,t) = R_j^i(x,t)e_a^j(x,t); \qquad e_a^i(0,x) = e_a^i(x), \tag{6.27}$$

where $R_j^i(x,t) = g^{ik}(x,t)R_{jk}(x,t)$ and $R_{jk}(x,t)$ and $g^{ik}(x,t)$ evolve under Ricci Flow $g_{ij}(x,t)$. It is easy to show (exercise) that the frame $(e_a(x,t))$ is still orthonormal for $g(x,t)$, that is,

$$g(t)_x(e_a(x,t), e_b(x,t)) = \delta_{ab}. \tag{6.28}$$

Furthermore, in $d = 3$ the evolution equation (4.2) for R_{ij} under Ricci Flow closes into

$$\partial_t R_{ij} = \Delta_g R_{ij} + 3RR_{ij} - 6R_i^k R_{kj} + (2R_{kl}R^{kl} - R^2)g_{ij}, \tag{6.29}$$

where $R_{ij} = R(\partial_i, \partial_j)$, to be distinguished from $R_{ab} = R(e_a, e_b)$. Moreover, if we define

$$R_{ab}(x,t) = e_a^i(x,t)e_b^j(x,t)R_{ij}(x,t), \tag{6.30}$$

then (6.29) becomes

$$\partial_t R_{ab} = \Delta_g R_{ab} + 3RR_{ab} - 4R_a^c R_{cb} + (2R_{cd}R^{cd} - R^2)\delta_{ab}, \tag{6.31}$$

where in fact $R_a^c = R_{ac}$ and $R^{cd} = R_{cd}$, since we work in an orthonormal frame. This equation has the feature that if $a \neq b$, then the right-hand side only contains diagonal terms that multiply off-diagonal terms, and hence if $R_{ab} = 0$ at $t = 0$ for $a \neq b$, then it remains zero at any time.[90] Hence we just have three coupled PDEs for the diagonal components R_{11}, R_{22}, and R_{33}, namely

$$\partial_t R_{11} = \Delta_g R_{11} + R_{11}(R_{22} + R_{33}) + (R_{22} - R_{33})^2; \tag{6.32}$$
$$\partial_t R_{22} = \Delta_g R_{22} + R_{22}(R_{11} + R_{33}) + (R_{11} - R_{33})^2; \tag{6.33}$$
$$\partial_t R_{33} = \Delta_g R_{33} + R_{33}(R_{11} + R_{22}) + (R_{11} - R_{22})^2. \tag{6.34}$$

Since these equations are coupled, we now need the minimum-maximum principle in the general form of Theorem 5.2, where in this case $V = \mathbb{R}^3$. This causes no special difficulties. In principle on should now introduce functions $(\alpha_1(t), \alpha_2(t), \alpha_3(t))$ and $(\beta_1(t), \beta_2(t), \beta_3(t))$, but is turns out to be more convenient to work with the variables

$$\rho_1 = R_{22} + R_{33} - R_{11}, \qquad \rho_2 = R_{11} + R_{33} - R_{22}, \qquad \rho_3 = R_{11} + R_{22} - R_{33}; \tag{6.35}$$
$$\lambda_1 = \alpha_2 + \alpha_3 - \alpha_1, \qquad \lambda_2 = \alpha_1 + \alpha_3 - \alpha_2, \qquad \lambda_3 = \alpha_1 + \alpha_2 - \alpha_3; \tag{6.36}$$
$$\alpha_1 = \tfrac{1}{2}(\lambda_2 + \lambda_3); \qquad \alpha_2 = \tfrac{1}{2}(\lambda_1 + \lambda_3); \qquad \alpha_3 = \tfrac{1}{2}(\lambda_1 + \lambda_2). \tag{6.37}$$

[89] The following procedure is a special case of what is called *Uhlenbeck's trick* in the Ricci Flow literature.

[90] This admittedly only shows that there is *a* solution with this feature, but uniqueness for given initial conditions shows that *the* solution also has $R_{ab}(t) = 0$ for $a \neq b$ whenever this holds at $t = 0$.

The ODEs for the minimum-maximum principle for the PDEs (6.32) - (6.34) are then given by

$$d_t\lambda_1 = \lambda_1^2 + \lambda_2\lambda_3; \qquad d_t\lambda_2 = \lambda_2^2 + \lambda_1\lambda_3; \qquad d_t\lambda_3 = \lambda_3^2 + \lambda_1\lambda_2, \tag{6.38}$$

with $d_t = d/dt$. These equations may be rewritten in the more useful form

$$\frac{d}{dt}(\lambda_1 - \lambda_2) = (\lambda_1 - \lambda_2)(\lambda_1 + \lambda_2 - \lambda_3); \tag{6.39}$$

$$\frac{d}{dt}(\lambda_2 - \lambda_3) = (\lambda_2 - \lambda_3)(\lambda_2 + \lambda_3 - \lambda_1); \tag{6.40}$$

$$\frac{d}{dt}(\lambda_1 - \lambda_3) = (\lambda_1 - \lambda_3)(\lambda_1 + \lambda_3 - \lambda_2). \tag{6.41}$$

So far, this analysis applies to any compact 3-manifold M. We now assume that M satisfies the hypothesis of Theorem 6.1, in which case it makes sense, in applying the minimum or maximum prinicple, to assume that $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$ for the initial metric $g = g(0)$. Much as in the previous section, one can now easily show that if $\lambda_1(0) \geq \lambda_2(0) \geq \lambda_3(0) > 0$ to begin with, then $\lambda_1(t) \geq \lambda_2(t) \geq \lambda_3(t) > 0$ for all t where the solutions are defined (exercise). If we then return to (6.38), we obtain $d_t\lambda_a \geq \lambda_a^2$ for $a = 1,2,3$, from which it follows that each $\lambda_a(t)$ diverges in finite time, cf. the analysis leading from (5.19) to (5.24). However, *normalized* Ricci Flow remains finite and converges to a metric with constant curvature. This is still very difficult (20 pages) to prove, but one gets a feeling why this is true if one allows the following shortcut:[91]

Lemma 6.7 *The projection of the flow $(\lambda_1(t), \lambda_2(t), \lambda_3(t))$ onto the unit sphere S^2 in $\mathbb{R}^3$ gives a system of ODEs that is asymptotically equivalent to that for normalized Ricci Flow.*

This means that we introduce new variables (x,y,z) by

$$x = \frac{\lambda_1}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \qquad y = \frac{\lambda_2}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \qquad z = \frac{\lambda_3}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \tag{6.42}$$

which indeed lie on S^2, and let them evolve via the $\lambda_a(t)$. This gives the following flow on S^2:

$$\frac{dx}{dt} = x^2 + yz - x(x^3 + y^3 + z^3 + 3xyz); \tag{6.43}$$

$$\frac{dy}{dt} = y^2 + xz - y(x^3 + y^3 + z^3 + 3xyz); \tag{6.44}$$

$$\frac{dz}{dt} = z^2 + xy - z(x^3 + y^3 + z^3 + 3xyz), \tag{6.45}$$

which, on the notation $f(x,y,x) = x(y^2+z^2) + y(x^2+z^2) + z(x^2+y^2) + 3xyz$, is equivalent to

$$\frac{d}{dt}(x-y) = -(x-y)(2z + f(x,y,z)); \tag{6.46}$$

$$\frac{d}{dt}(x-z) = -(x-z)(2y + f(x,y,z)); \tag{6.47}$$

$$\frac{d}{dt}(y-z) = -(y-z)(2x + f(x,y,z)). \tag{6.48}$$

Within the region where $x \geq y \geq z > 0$, this flow exists for all t and, by the same arguments as in the previous section, converges to its fixed point $(1,1,1)/\sqrt{3}$. Using both the minimum and the maximum principles, it follows that the components R_{11}, R_{22}, and R_{33} also converge to the same constant value, which in turn implies that the underlying metric converges to a metric with constant curvature. Given Lemma 6.7, this proves Theorem 6.1. □

[91] Unfortunately, as far as we can see this lemma can only be proved with hindsight from Theorem 6.1.

7 Existence of solutions of quaslinear parabolic PDEs

In this chapter we sketch a proof of existence (for short times) and uniqueness of solutions of PDEs of the kind we have studied so far, albeit just for a single variable.[92] Thus we are interested in second-order quaslinear parabolic PDEs equations of the type (5.14), with P given by (5.2), but we allow the coefficients a^{ij} in the second-order operator P to depend on both u and its *first* derivatives ∂u, the vector field b may depend on u, and the function F may depend on ∂u as well as on u; the key restriction is that the second-order terms $\partial_i \partial_j u$ still occur linearly. Indeed, this is the property that makes the equation *quasilinear*: the highest derivatives occur linearly. It is this feature that makes linear methods available for the first step of the proof, followed by a fixed-point argument (in Banach space) to deal with the nonlinear case.

7.1 Existence and uniqueness: linear case

Generalizing (5.2), we first consider linear elliptic operators taking the form

$$P = a^{ij}(x,t)\partial_i \partial_j + b^i(x,t)\partial_j + c(x,t), \tag{7.1}$$

as well as at functions $f(x,t)$, with associated parabolic PDE $\partial_t u = Pu + f$. We rewrite this as

$$\dot{u} = A(t)u + f(t), \tag{7.2}$$

where $A(t) = P$, but the point is to regard $A(t)$ as a parametrized family of *bounded* linear maps

$$A(t) : V \to V^*, \tag{7.3}$$

where $t \in [0,T]$ for some $T > 0$ (we will see that we can take $T > 0$ arbitrary), and V is some Banach space such that $f(t) \in V^*$. For applications to (7.1) we will take

$$V = H^1(M) \qquad \Leftrightarrow \qquad V^* = H^{-1}(M), \tag{7.4}$$

but in the general method the choice is not prescribed. The key assumptions on $A(t)$ are:

1. The map $t \mapsto A(t)$ is continuous from $[0,T]$ to the space $B(V,V^*)$ of bounded linear maps from V to V^* (with the norm topology, where $\|A\|_{B(V,V^*)} = \sup\{\|Av\|_{V^*}, v \in V, \|v\|_V = 1\}$).

2. There is a constant $C > 0$ such that for any $u \in V$ one has a so-called *coercivity estimate*

$$-\langle u, A(t)u\rangle \geq C \cdot \|u\|_V^2. \tag{7.5}$$

It is a highly nontrivial fact, which we state without proof, that elliptic operators like (7.1) satisfy this estimate,[93] provided we choose (7.4) and take the coefficients a^{ij}, b^i, and c in (7.1) to be sufficiently regular (e.g. smooth and bounded) so as to satisfy condition 1 above.

[92] The extension to the Ricci Flow equation (1.1), seems quite nontrivial, but is usually dealt with in the literature by stating that DeTurck's trick (see §4.2) reduces it to a standard situation. Dan Knopf, the co-author of the fundamental textbook Chow & Knopf (2004), admitted in an email that this is indeed a weak point in the literature on Ricci Flow. Our treatment is a mix of Renardy & Rogers (2004), §11.1, for the linear part, and Lieberman (1998), Chapter VIII, for the non linear part, but in order to make these treatments compatible the functional-analytic setting for Lieberman's use of the Schauder Fixed Point Theorem, on which his proof is based, had to be changed from Hölder spaces to Sobolev spaces, admittedly at the expense of an extra assumption, namely (7.14), that makes the proof below less general than desirable (but sufficient for our purposes).

[93] This relies on the Gårding inequality, see Renardy & Rogers (2004), Example 11.5 in §11.1.3 and §9.2.3. One may need to change $A(t)$ to $A(t) - c \cdot 1_H$ (and $u(t)$ to $u(t)\exp(-ct)$) to satisfy (7.5), e.g. for the Laplacian on $\mathbb{R}^n$.

As our last assumption, we assume V and V^* fit into a *Gelfand triple* (cf. Appendix D), viz.

$$V \subset H \subset V^*; \tag{7.6}$$

for our choice (7.4) in the context of (7.1), this is the case for $H = L^2(M)$, cf. (D.19).

Theorem 7.1 *On the assumptions just stated, for initial value $u_0 \in H$ and $f(t) \in L^2([0,T],V^*)$, eq.* (7.2) *has a unique solution u that lies in $L^2([0,T],V)$, in $H^1([0,T],V^*)$, and in $C([0,T],H)$.*

Proof. The proof starts from the following *a priori* estimate:[94] any solution to (7.2) satisfies

$$\|u\|_{L^2([0,T],V)} \leq C_2(\|f\|_{L^2([0,T],V^*)} + \|u_0\|_H), \tag{7.7}$$

where $C_2 > 0$ depends on the family $A(t)$ over $[0,T]$ but not on u. This follows by integrating the PDE (7.2) from 0 to T and writing $\langle u(t), \dot{u}(t)\rangle = \frac{1}{2}\partial_t \|u(t)\|_H^2$, which gives

$$\tfrac{1}{2}(\|u(T)\|_H^2 - \|u_0\|_H^2) - \int_0^T dt\, \langle u(t), A(t)u(t)\rangle = \int_0^T dt\, \langle f(t), u(t)\rangle, \tag{7.8}$$

where both brackets denote the V-V^* pairing.[95] Subsequently, estimate (we omit the t)

$$\langle f,u\rangle \leq |\langle f,u\rangle| \leq \|f\|_{V^*}\|u\|_V \leq \tfrac{1}{2}\left(\varepsilon\|u\|_V^2 + \frac{1}{\varepsilon}\|f\|_{V^*}^2\right), \tag{7.9}$$

for any $\varepsilon > 0$, which follows from the inequality $\varepsilon^{-1}(a - \varepsilon b)^2 \geq 0$ for any $\varepsilon > 0$ and $a, b \in \mathbb{R}$, form which $ab \leq \frac{1}{2}(\varepsilon b^2 + \varepsilon^{-1}a^2)$. A smart choice of ε (exercise) then yields (7.7).

Since (7.2) is linear, this estimate immediately proves *uniqueness* of a solution, if any. It also provides the basis of *existence* through what is called a *Galerkin method*,[96] meaning that some PDE is approximated by a finite number of ODEs, whose solution then converges to a solution of the PDE. In the case at hand, this idea is executed as follows. Take some orthonormal basis (e_n) of H that lies in V and, for any $N < \infty$, consider the system of N ODEs given by

$$\langle e_n, \dot{u}_N\rangle_H = \langle e_n, A(t)u_N(t)\rangle_H + \langle e_n, f(t)\rangle_H, \ (n = 1, \ldots, N). \tag{7.10}$$

This system can be solved for $u_N(t) \in P_N H$, where $P_N = \sum_{n=1}^N |e_n\rangle\langle e_n|$ is the projection on the linear space of the first N basis vectors e_n, with initial condition $u_N(0) = P_N u_0$. Then u_N satisfies (7.7), and since $\|P_N u_0\| \leq \|u_0\|$, we have $\|u_N\|_{L^2([0,T],V)} \leq C_3$. Now a bounded sequence in a Hilbert space has a weakly convergent subsequence, and so the sequence (u_N) has a limit point $u \in H$. It is an exercise to show that $u = \lim_N u_N$ solves the original PDE (7.2) with initial condition u_0, upon which (7.7) shows both uniqueness (which was not clear from the previous step) and the property $u \in L^2([0,T],V)$. Furthermore, since $A(t) \in B(V,V^*)$, the PDE (7.2) shows that $\dot{u} \in L^2([0,T],V^*)$ and hence $u \in H^1([0,T],V^*)$. By Sobolev embedding for $s = n = 1$ and $k = 0$ (see Theorem D.1) the last point also gives $u \in C([0,T],V^*)$, from which the step to $u \in C([0,T],H)$ is unfortunately still quite difficult and technical (and will be omitted). □

Finally, under sharper assumptions on $A(t)$ and $f(t)$, improved regularity of $u(t)$ in time follows from the PDE (7.2) itself: if $\dot{A}(t)$ and $\dot{f}(t)$ satisfy the same assumptions as $A(t)$ and $f(t)$, then $v = \dot{u}$ satisfies $\dot{v} = Av + g$ with $g = \dot{A}u + \dot{f}$, and hence $\dot{u} = v \in C([0,T],H)$. If both $A(t)$ and $f(t)$ are C^∞, iterating this argument also makes u smooth in t. In cases like (7.1), also smoothness in x can be achieved if $A(t)$ and $f(t)$ satisfy our assumptions for $V = H^k(M)$ with arbitrary $m \in \mathbb{N}$. This drives u into $C^\infty([0,T], C^\infty(M)) = C^\infty([0,T] \times M)$ and makes the solution classical.[97]

[94] *A priori* estimates are typically given before the existence of a solution has been proved.

[95] If $u \in V$, then also $u \in H$ and $u \in V^*$, but there is no ambiguity since $\langle u,u\rangle$ as given by the V-V^* pairing coincides with $\langle u,u\rangle_H$ because the pairing is given through the inner product in H.

[96] For details in the context that follows see Renardy & Rogers (2004), §11.1.2 and Evans (2010), §7.1.2.

[97] This argument is given in detail by Evans (2010), §7.1.3, study it as an exercise!

7.2 Existence for short times: quasilinear case

We now move to the nonlinear PDE $\partial_t u = Pu$, where, as opposed to (7.1), we now have

$$P = a^{ij}(x,t,u,\partial u)\partial_i\partial_j + b^i(x,t,u)\partial_j + c(x,t,u), \tag{7.11}$$

and also f may depend on u and ∂u, as well as on x and t, as before.[98] The idea of the existence proof is to initially replace u in the coefficients a^{ij}, b^i, and c as well as in $f(x,t,u,\partial u)$ by an arbitrary function v (in the class of functions where we search for solutions), so as to obtain a linear PDE to which Theorem 7.1 applies. This gives existence and uniqueness of solutions for the wrong equation, but a clever application of the ***Schauder fixed-point theorem*** (invoking Rellich's Theorem D.1 (d)) then gives the same result for (7.11). This theorem reads as follows:

Theorem 7.2 *If $K \subset B$ is a* compact convex *subset of a Banach space B and $\varphi : K \to K$ is continuous, then φ has a fixed point.*

This generalizes Brouwer's fixed point theorem to certain infinite-dimensional spaces. In order to apply it, we first agree to denote P in (7.11) by P_u, and likewise f by f_u, so that our PDE is

$$\partial_t u = P_u u + f_u. \tag{7.12}$$

We label the version of P where its arguments u and ∂u are replaced by v and ∂v as P_v, and likewise f_v means f with u and ∂u replaced by v and ∂v. In particular, $P_u = P$ and $f_u = f$. Then

$$\partial_t w = P_v w + f_v \tag{7.13}$$

is a linear version of (7.11), differing from the *linearization* of (7.11), which plays no role here.

We then need to find a Banach space B and a compact convex subset $K \subset B$ such that the map $\varphi(v) = w$, where w solves (7.13) with initial condition u_0, continuously maps K to K (this map is well defined by Theorem 7.1). If u is a fixed point of this map, then $\partial_t u = P_u u$ and hence u solves (7.12) with the right initial condition u_0, since this is the same for (7.12) and (7.13). We now assume that f has the following property: for each $\alpha > 0$ there is $C(\alpha) > 0$ such that:[99]

$$\|v\|_{L^2([0,T],V)} \leq \alpha \qquad \Rightarrow \qquad \|f_v\|_{L^2([0,T],V^*)} \leq C(\alpha)T. \tag{7.14}$$

If this is the case, then our *a priori* estimate (7.7) gives, for solutions u of (7.13),

$$\|u\|_{L^2([0,T],V)} \leq C_2(C(\alpha)T + \|u_0\|_H). \tag{7.15}$$

Now take $\alpha = C_2(1 + \|u_0\|_H)$ and $T = 1/C(\alpha)$. The assumption (7.14) then implies

$$\|v\|_{L^2([0,T],V)} \leq \alpha \qquad \Rightarrow \qquad \|u\|_{L^2([0,T],V)} \leq \alpha. \tag{7.16}$$

What is still lacking for the use of Theorem 7.2 is continuity of φ, which (since φ is nonlinear) does not follow from (7.16) and should be proved on a case by case basis for each given quasilinear elliptic operator P, and compactness of the closed ball in $L^2([0,T],V)$ with radius α. However, taking $V = H^1(M)$ as before, at the expense of moving up the regularity assumptions on u_0 and f by some arbitrary $\delta > 0$ (in a sense that will be clear immediately), we can perform the argument in the closed α-ball in $H^\delta([0,T],H^{1+\delta}(M))$ but now taken in the norm of $L^2([0,T],H^1(M))$, in which, for any $\delta > 0$, this ball is compact by Theorem D.1 (d).[100]

This gives existence of solutions; unlike in the linear case, uniqueness has to be proved separately, for example from a minimum and maximum principle (exercise–tongue in cheek!).

[98] In view of this, we might as well omit the b and c terms in P, as they can be absorbed in this more general f.

[99] This is only known to be true for Hölder spaces instead of Sobolev spaces, see Lieberman (2005), Chapter VIII. Lieberman forgets to prove continuity of φ in his Theorem 8.2, though he needs it just as much as we do!

[100] Combined with the fact that $H_1 \otimes H_2 \hookrightarrow H_3 \otimes H_4$ is compact iff $H_1 \hookrightarrow H_3$ and $H_2 \hookrightarrow H_4$ are compact.

A Lie groups

We only need *linear* Lie groups, which by definition are closed subgroups of $GL_n(\mathbb{K})$ for $\mathbb{K}=\mathbb{R}$ or $\mathbb{K}=\mathbb{C}$, i.e. the group of invertible $n\times n$ matrices with entries in $\mathbb{K}$, with group multiplication simply given by matrix multiplication.[101] As a first and hopefully well-known example, $SO(3)$ is the subgroup of $GL_3(\mathbb{R})$ consisting of matrices R that satisfy

$$R^T R = 1_3; \tag{A.1}$$

$$\det(R) = 1. \tag{A.2}$$

More generally, for some given $\Gamma \in GL_n(\mathbb{K})$, the matrices $g \in GL_n(\mathbb{K})$ that for all x, y satisfy

$$\langle gx, \Gamma gy\rangle = \langle x, \Gamma y\rangle, \tag{A.3}$$

or, in other words, leave the bilinear (or sequilinear) form $\langle x,y\rangle_\Gamma = \langle x, \Gamma y\rangle$ invariant (where $\langle\cdot,\cdot\rangle$) is the usual inner product on $\mathbb{K}^n$), form a linear Lie group G_Γ. In other words,

$$G_\Gamma = \{g \in GL_n(\mathbb{K}) \mid g^*\Gamma g = \Gamma\}. \tag{A.4}$$

For $\mathbb{K}=\mathbb{R}$, $n=3$, and $\Gamma = 1_3$ we obtain $G_\Gamma = O(3)$, which has has two components: the one containing the identity is $SO(3) \equiv O(3)_+$, singled out by the condition $\det(R)=1$, whereas the other component $O(3)_-$ consists of those elements $R \in O(3)$ with $\det(R) = -1$. Note that $SO(3)$ is connected but not *simply* connected. Furthermore, $O(3)$ and $SO(3)$ are *compact* in the topology inherited from $M_3(\mathbb{R}) \cong \mathbb{R}^9$: this follows from the Heine–Borel theorem. Indeed, $R^T R = 1_3$ implies $\|R\| = 1$, so that $SO(3)$ must be a bounded subset of $\mathbb{R}^9$, which is also closed, since the condition (A.3) is closed by continuity of matrix multiplication, as is $\det(R) = 1$. Compactness also follows from the following parametrization of $SO(3)$, with $\alpha, \beta, \gamma \in [0, 2\pi]$:

$$R^z_\gamma = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}, R^y_\beta = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}, R^x_\alpha = \begin{pmatrix} 1 & & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Staying in $n=3$ and $\mathbb{K}=\mathbb{R}$ for the moment, instead of $\Gamma = 1_3$ we may take $\Gamma = \mathrm{diag}(-1,1,1)$. Then $G_\Gamma \equiv O(2,1)$ is called the ***Lorentz group*** (in space-time dimension 3). It has *four* components, singled out by the four combinations of the two independent conditions

$$\det(\lambda) = \pm 1; \qquad\qquad \pm\lambda_{00} > 0; \tag{A.5}$$

for an indefinite Γ like this it is customary to label the entries λ_{ij} by $i, j = 0, 1, 2$ instead of $1, 2, 3$. In particular, the identity component $O(2,1)_0$ satisfies $\det(\lambda) = 1$ and $\lambda_{00} > 0$.[102] Consequently, even the subgroup $SO(2,1) = \{\lambda \in O(2,1) \mid \det(\lambda) = 1\}$ has *two* components. Another important difference with $SO(3)$ is that $SO(2,1)$ is *non-compact*. This follows, for example, from the following parametrization of $O(2,1)_0$, where $\alpha \in [0, 2\pi]$ and $\beta, \gamma \in \mathbb{R}$:

$$B^x_\gamma = \begin{pmatrix} \cosh\gamma & \sinh\gamma & 0 \\ \sinh\gamma & \cosh\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}, B^y_\beta = \begin{pmatrix} \cosh\beta & 0 & \sinh\beta \\ 0 & 1 & 0 \\ \sinh\beta & 0 & \cosh\beta \end{pmatrix}, R_\alpha = \begin{pmatrix} 1 & & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

[101] Lie groups are not necessarily closed in $M_n(\mathbb{K})$, since invertibility of matrices is an open condition (we call a condition *open* if its solution set is open, and *closed* if its solution set is closed). For example, the sequence $g_n = 1_n/n$ in $GL_n(\mathbb{R})$ converges to zero, so the limit is not in $GL_n(\mathbb{R})$. The topology used may either be the usual one or $\mathbb{K}^{n^2}$ or the matrix norm topology; these are equivalent.

[102] This follows from the fact that any matrix $\lambda \in O(2,1)$ satisfies $\lambda_{00}^2 - \sum_{k=1} 23\lambda_{k0}^2 = 1$, so that $|\lambda_{00}| \geq 1$, and from the fact that $\mathrm{sgn}(\lambda_{00})$ and $\det(\lambda)$ are continuous functions on $O(2,1)$.

From these, one obtains the matrices λ with $\det(\lambda) = 1$ and $\lambda_{00} < 0$ by multiplication with $\mathrm{diag}(-1,-1,1)$, those with $\det(\lambda) = -1$ and $\lambda_{00} > 0$ by multiplication with $\mathrm{diag}(1,-1,1)$, and finally, those with $\det(\lambda) = -1$ and $\lambda_{00} < 0$ by multiplication with $\mathrm{diag}(-1,1,1)$.

The additive (and hence abelian) groups $\mathbb{R}^n$ also fall under our definition of linear Lie groups, since one may identify $a \in \mathbb{R}^n$ with the $2n \times 2n$-matrix

$$a \equiv \begin{pmatrix} 1_n & \mathrm{diag}(a) \\ 0 & 1_n \end{pmatrix}, \tag{A.6}$$

where $\mathrm{diag}(a)$ is the diaginal $n \times n$ matrix with entries $(a_1, \dots, a_n)$ on the diagonal. Indeed, matrix multiplication reproduces addition. On the other hand, we simply put the n-torus $G = \mathbb{T}^n = U(1)^n$ (where $\mathbb{T} \equiv \{z \in \mathbb{C} : |z| = 1\}$), which is the compact sister of $\mathbb{R}^n$, with multiplication as group operation, into the diagonal of $GL_n(\mathbb{C})$.

An intermediate case (between abelian and non-abelian) is the 3*d* ***Heisenberg group***

$$\mathrm{Nil} = \mathrm{Heis}_3(\mathbb{R}) \subset GL_3(\mathbb{R}), \tag{A.7}$$

called Nil by Thurston since it is nilpotent, which consists of all real matrices

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, \quad (x,y,z) \in \mathbb{R}^3. \tag{A.8}$$

Another three-dimensional Lie groups of interest is $SL_2(\mathbb{R})$, which is the (non-compact) subgroup of $GL_2(\mathbb{R})$ consisting of all matrices with unit determinant, which is doubly connected, along with its universal cover $\widetilde{SL_2(\mathbb{R})}$ (which is no longer a Linear Lie group, however). Finally,

$$E(2) = O(2) \ltimes \mathbb{R}^2; \tag{A.9}$$

$$E(1,1) = O(1,1) \ltimes \mathbb{R}^2, \tag{A.10}$$

are the ***Euclidean group*** and the ***Poincaré group*** in dimension 2 and $1+1$, where $O(1,1)$ is defined as the subgroup of $GL_2(\mathbb{R})$ that leaves $\Gamma = \mathrm{diag}(-1,1)$ invariant, with the additive group $\mathbb{R}^2$, on which $O(1,1)$ acts in its defining representation.[103] The identity component of $E(2)$ is

$$\mathrm{Sol} = E(1,1)_0 = O(1,1)_0 \ltimes \mathbb{R}^2, \tag{A.11}$$

where Sol is Thurston's name for it since it is solvable. Here $O(1,1)_0$ is the identity component of $O(1,1)$ (which has four connected components, like $O(2,1)$), consisting of all matrices

$$\begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix}, \quad \alpha \in \mathbb{R}.$$

Note that Nil, Sol, and $\widetilde{SL_2(\mathbb{R})}$ are all diffeomorphic to $\mathbb{R}^3$, but have different group structures. All this can be generalized to any dimension in an obvious way; we have focused on some three-dimensional examples, since these play a role in Thurston's Geometrization Conjecture.

[103] Let some group L act linearly on a vector space V. Then the operation $(\lambda, v) \cdot (\lambda', v') = (\lambda\lambda', v + \lambda \cdot v')$, with inverse $(\lambda, v)^{-1} = (\lambda^{-1}, -\lambda^{-1} \cdot v)$, turns $L \ltimes V$ into a group, called the ***semi-direct product*** of L and V. If $L \subset GL_n(\mathbb{K})$ is a linear Lie group and $V = \mathbb{K}^n$, then $L \ltimes V$ is a linear Lie group in $GL_{2n}(\mathbb{K})$, realized by the matrices $\begin{pmatrix} L & \mathbf{v} \\ 0 & 1_n \end{pmatrix}$, where $\mathbf{v} \in GL_n(\mathbb{K})$ is the matrix with $v \in V$ in every column.

B Lie algebras

Abstractly, a ***Lie algebra*** over $\mathbb{K}$ is a vector space over $\mathbb{K}$ equipped with an antisymmetric bilinear map $[\cdot,\cdot] : A \times A \to A$ (i.e., $[a,b] = -[b,a]$) that satisfies the ***Jacobi identity***

$$[a,[b,c]] + [c,[a,b]] + [b,[c,a]] = 0 \ \ (a,b,c \in A). \tag{B.1}$$

Concretely, the Lie algebra of a linear Lie group $G \subset GL_n(\mathbb{K})$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, is

$$\mathfrak{g} = \{A \in M_n(\mathbb{K}) \mid e^{tA} \in G \forall t \in \mathbb{R}\}, \tag{B.2}$$

where the exponential map $\exp : \mathfrak{g} \to G$ is just given by its usual (norm-convergent) power series. Even if $\mathbb{K} = \mathbb{C}$, we often regard $\mathfrak{g}$ as a *real* vector space. Eq. (B.2) defines a Lie algebra in the above abstract sense, where the Lie bracket is simply defined as the commutator

$$[A,B] = AB - BA. \tag{B.3}$$

It is a nontrivial fact that this concrete Lie algebra is also an abstract one, notably that $\mathfrak{g}$ is a vector space and that the bracket (B.3) indeed maps $\mathfrak{g} \times \mathfrak{g}$ to $\mathfrak{g}$ (on the other hand, the Jacobi identity is easily verified). The former property follows from the ***Lie product formula***

$$e^{A+B} = \lim_{m\to\infty} \left(e^{A/m}e^{B/m}\right)^m, \tag{B.4}$$

combined with the axiom that G be closed in $GL_n(\mathbb{K})$, whereas the latter property derives from

$$[A,B] = \frac{d}{dt} e^{tA} B e^{-tA}, \tag{B.5}$$

combined with a lemma about matrices showing that if $g \in G$ and $A \in \mathfrak{g}$, then $gAg^{-1} \in \mathfrak{g}$ (which in turn follows from the definition of the exponential, implying $\exp(gAg^{-1}) = g\exp(A)g^{-1}$.)

If $G = G_\Gamma$ is defined by (A.4), then its Lie algebra is

$$\mathfrak{g}_\Gamma = \{A \in M_n(\mathbb{K}) \mid A^*\Gamma = -\Gamma A\}. \tag{B.6}$$

For example, taking $\Gamma = \text{diag}(1,1,1)$, the Lie algebra $\mathfrak{so}(3)$ of $SO(3)$ consists of all real 3×3 matrices X that satisfy $X^T = -X$. As a vector space $\mathfrak{so}(3) \cong \mathbb{R}^3$, since $\mathfrak{so}(3)$ has a basis

$$e_1 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad e_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}; \tag{B.7}$$

whose linear span gives all 3×3 real antisymmeric matrices.[104] A vector space isomorphism $\mathbb{R} \xrightarrow{\cong} \mathfrak{so}(3)$ is then given by $(x,y,z) \mapsto xe_1 + ye_2 + ze_3$. The commutators of these elements are

$$[e_1,e_2] = e_3; \qquad [e_2,e_3] = e_1; \qquad [e_3,e_1] = e_2, \tag{B.8}$$

and by linearity these determine the Lie bracket of arbitrary elements of $\mathfrak{so}(3)$.

[104] This is not the usual basis (J_1,J_2,J_3) of $\mathfrak{so}(3)$, which is defined by $J_1 = e_3$, $J_2 = -e_2$, $J_3 = J_1$.

Similarly, according to (B.6) the Lie algebra of $SO(2,1)$ consists of all real 3×3 matrices A that satisfy $A^T\mathrm{diag}(-1,1,1) = -\mathrm{diag}(-1,1,1)A$. There are good reasons for taking the basis

$$f_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad f_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad f_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{B.9}$$

with commutation relations

$$[f_1, f_2] = -f_3; \qquad [f_2, f_3] = f_1; \qquad [f_3, f_1] = f_2. \tag{B.10}$$

For $SL(2,\mathbb{R})$, whose Lie algebra consist of all real 2×2 *traceless* matrices, we may take

$$g_1 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad g_2 = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad g_3 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{B.11}$$

with commutation relations

$$[g_1, g_2] = -g_3; \qquad [g_2, g_3] = g_1; \qquad [g_3, g_1] = g_2. \tag{B.12}$$

It follows that the Lie algebras of $SO(2,1)$ and $SL_2(\mathbb{R})$ are isomorphic, namely by linear extension of the map $f_i \mapsto g_i$, $i = 1,2,3$. Nonetheless, the groups in question are not isomorphic.

Similarly, consider the Lie group $SU(2)$, defined as the set of unitary 2×2 complex matrices with unit determinant. Its Lie algebra then consists of all traceless matrices $A \in M_2(\mathbb{C})$ (this is the "S" in $SU(2)$) for which $A^* = -A$ (which is the "U"). If we take the following basis

$$h_1 = \tfrac{1}{2}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \qquad h_2 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad h_3 = \tfrac{1}{2}\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, \tag{B.13}$$

of the vector space of such matrices, with commutation relations

$$[h_1, h_2] = h_3; \qquad [h_2, h_3] = h_1; \qquad [h_3, h_1] = h_2, \tag{B.14}$$

we see that the Lie algebras of $SO(3)$ and $SU(2)$ are isomorphic by linear extension of the map $e_i \mapsto h_i$, $i = 1,2,3$, although, once again, the groups are not isomorphic. See the next section.

The next interesting three-dimensional case is the Euclidean group (A.9), which has two connected components, of which the identity component is

$$E(2)_0 = SO(2) \ltimes \mathbb{R}^2. \tag{B.15}$$

To find its Lie algebra, we note that in general the Lie algebra $\mathfrak{g}$ of a semidirect product $L \ltimes \mathbb{R}^n$ is $\mathfrak{l} \oplus \mathbb{R}^n$ as a vector space, with commutators for $A, B \in \mathfrak{l}$ and $v, w \in V$ given by

$$[(A,v),(B,w)] = ([A,B], Aw - Bv). \tag{B.16}$$

Since $SO(2)$ consists of all matrices

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \ \alpha \in [0, 2\pi], \tag{B.17}$$

we make take the basis

$$j_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad j_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad j_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{B.18}$$

the former forming a basis of $\mathbb{R}^2$, and find the commutation relations from (B.16) to be

$$[j_1, j_2] = 0; \qquad [j_2, j_3] = j_1; \qquad [j_3, j_1] = j_2. \tag{B.19}$$

For $E(1,1)$, on the other hand, we take

$$k_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad k_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad k_3 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \tag{B.20}$$

to obtain

$$[k_1, k_2] = 0; \qquad [k_2, k_3] = k_1; \qquad [k_3, k_1] = -k_2. \tag{B.21}$$

For the Lie algebra of the Heisenberg group, which consists of all matrices

$$\begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix},$$

again with $(x,y,z) \in \mathbb{R}^3$, we take the natural basis

$$l_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad l_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad l_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \tag{B.22}$$

with commutation relations

$$[l_1, l_2] = 0; \qquad [l_2, l_3] = l_1; \qquad [l_3, l_1] = 0. \tag{B.23}$$

The last three-dimensional Lie algebra of interest is simply $\mathbb{R}^3$, with the usual basis

$$m_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad m_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad m_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{B.24}$$

and commutation relations

$$[m_1, m_2] = 0; \qquad [m_2, m_3] = 0; \qquad [m_3, m_1] = 0. \tag{B.25}$$

Due to our particular choices of a basis, we can summarize all cases so far by the relations

$$[n_1, n_2] = \lambda_3 n_3; \qquad [n_2, n_3] = \lambda_1 n_1; \qquad [n_3, n_1] = \lambda_2 n_2, \tag{B.26}$$

where the λ_i are equal to ± 1 or 0 and are given by the following table (Milnor, 1976):[105]

$(\lambda_1, \lambda_2, \lambda_3)$	group
$(1,1,1)$	$SO(3)$ or $SU(2)$
$(1,1,-1)$	$SO(2,1)$ or $SL_2(\mathbb{R})$
$(1,1,0)$	$E(2)$
$(1,-1,0)$	$E(1,1)$
$(1,0,0)$	$\mathrm{Heis}_3(\mathbb{R})$
$(0,0,0)$	$\mathbb{R}^3$

[105] This table gets more body in connection with Lemma 6.3. If we refrain from the requirement of orthonormality of the basis, we are free to rescale the λ_i by recsaling one or more basis vectors. This gives the table.

C Lie's Third Theorem

We saw how (linear) Lie groups lead to Lie algebras. We now briefly discuss the converse passage from Lie algebras to Lie groups. Given any Lie algebra $\mathfrak{g} \subset M_n(\mathbb{K})$ realized in matrix form, one may define the group G_0 generated within $GL_n(\mathbb{K})$ by all ***one-paramteter subgroups*** $t \mapsto \exp(tA)$, where $A \in \mathfrak{g}$ and $t \in \mathbb{R}$. This group is connected by construction, but it is not the most general Lie group with Lie algebra (isomorphic to) $\mathfrak{g}$. For example, let us take $\mathfrak{g} = \mathbb{R}$ with zero Lie bracket. Suppose we realize this inside $M_1(\mathbb{C}) = \mathbb{C}$ as the set of all iA, $A \in \mathbb{R}$. Exponentiation gives $G_0 = \mathbb{T} \equiv \{z \in \mathbb{C} : |z| = 1\}$, with multiplication as the group operation; this one-dimensional Lie group is the torus (and analogously for $\mathbb{R}^n$, giving the n-torus $\mathbb{T}^n$). Now, however, realize $\mathfrak{g} = \mathbb{R}$ as the set of all matrices $\begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix}$ within $GL_2(\mathbb{R})$, where again $A \in \mathbb{R}$. This time, exponentiation gives, perhaps surprisingly, $\exp\begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & A \\ 0 & 1 \end{pmatrix}$. Since $\begin{pmatrix} 1 & A \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & A+B \\ 0 & 1 \end{pmatrix}$, this time, from the same Lie algebra we obtain the Lie group $G_0 = \mathbb{R}$ (as an additive group), which is simply connected (and similar for $\mathbb{R}^n$).

Other examples we have already encountered are $SO(3)$ and $SU(2)$, which have isomorphic Lie algebras but are not isomorphic as Lie groups, and similarly $SO(2,1)$ and $SL_2(\mathbb{R})$ or $\widetilde{SL_2}(\mathbb{R})$. So, depending on its specific realization, the "same" Lie algebra may give rise to very different Lie groups. ***Lie's Third Theorem*** (proved by É. Cartan) settles the matter completely:[106]

Theorem C.1 *Let $\mathfrak{g}$ be a Lie algebra. There exists a simply connected Lie group $\tilde{G}$, unique up to isomorphism, such that the Lie algebra of $\tilde{G}$ is $\mathfrak{g}$ (and any Lie group isomorophic to $\tilde{G}$ has a Lie algebra ismomorphic to $\mathfrak{g}$). Furthermore, if G is a connected Lie group with Lie algebra isomorphic to $\mathfrak{g}$, then $G \cong \tilde{G}/D$, where D is a discrete normal subgroup of the center of $\tilde{G}$.*

In the first example above, i.e. $\mathfrak{g} = \mathbb{R}$, we have $\tilde{G} = \mathbb{R}$ and $\mathbb{T} \cong \mathbb{R}/\mathbb{Z}$ (so $D = \mathbb{Z}$). In the second, $SU(2)$ is simply connected: this follows from the fact that it consist of all matrices of the form

$$\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix},\ \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1. \tag{C.1}$$

This implies that $SU(2) \cong S^3$, whose simple connectedness we take for granted. By Lie's Third Theorem, we must therefore have $SO(3) \cong SU(2)/D$, where D is a discrete subgroup of the center of $SU(2)$. The center of $SU(2)$ is $\{1_2, -1_2\}$, and to prove that D equals the center it only remains to prove that $SU(2)$ and $SO(3)$ are not isomorphic; this follows by finding a non-contractible loop in $SO(3)$ (almost any rotation around 2π will do), whence $\pi_1(SO(3)) = \mathbb{Z}_2$.

A more difficult analysis (Carter et al, 1995, pp. 56–58) shows that $SL_2(\mathbb{R})$ is homeomorphic to an open sausage $S^1 \times D$, where S^1 is the circle and D is the open disk in $\mathbb{R}^2$, so that $\pi_1(SL_2(\mathbb{R})) = \mathbb{Z}_2$ and its universal cover $\widetilde{SL_2(\mathbb{R})}$ is homeomorphic to an infinite open cylinder $\mathbb{R} \times D$, which of course is homeomorphic to $\mathbb{R}^3$. Finally, $SO(2,1)_0$ is homeomorphic to $S^1 \times \mathbb{R}^2$ and hence $\pi_1(SO(2,1)_0) = \mathbb{Z}$; it is a quotient of $\widetilde{SL_2(\mathbb{R})}$ by $\mathbb{Z}$, so at the end of the day we have

$$SO(3) \cong SU(2)/\mathbb{Z}_2; \qquad SL_2(\mathbb{R}) \cong \widetilde{SL_2(\mathbb{R})}/\mathbb{Z}_2; \qquad SO(2,1)_0 \cong \widetilde{SL_2(\mathbb{R})}/\mathbb{Z}. \tag{C.2}$$

[106] We say that two Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are isomorphic, written $\mathfrak{g}_1 \cong \mathfrak{g}_2$, if there is a vector space isomorphism $\varphi : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that $[\varphi(A), \varphi(B)]_{\mathfrak{g}_2} = \varphi([A,B]_{\mathfrak{g}_1})$ for all $A, B \in \mathfrak{g}_1$.

D Distributions and Sobolev spaces on manifolds

This is an introduction to the topics in the title, containing just what we need for chapter 7.[107]

1. ***Notation.*** Let $n > 0$ and $x \in \mathbb{R}^n$. It will convenient to write $x = (x_1, \dots, x_n)$ rather than our usual $(x^1, \dots, x^n)$. Let $\alpha = (\alpha_1, \dots \alpha_n)$, with $\alpha_i \in \mathbb{N}$ (where $0 \in \mathbb{N}$), and write

$$|\alpha| = \sum_{i=1}^{n} \alpha_i; \quad x^\alpha = x_1^{\alpha^1} \cdots x_n^{\alpha^n}; \tag{D.1}$$

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} \equiv \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} \equiv \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}. \tag{D.2}$$

2. ***Test functions.*** Let $\mathscr{D}(\mathbb{R}^n)$ be $C_c^\infty(\mathbb{R}^n)$ as a set, equipped with the topology in which $f_\lambda \to f$ iff there is a compact set $K \subset \mathbb{R}^n$ such that $\mathrm{supp}(f_\lambda) \subseteq K$ for all λ as well as $\mathrm{supp}(f) \subseteq K$, and for all multi-indices α one has $\|D^\alpha(f_\lambda - f)\|_\infty \to 0$.

   This may be generalized to manifolds M, as follows: for some given atlas (U_i, φ_i), where the U_i form an open cover of M and each $\varphi_i : U_i \to \mathbb{R}^n$ is injective and open (so that U_i is homeomorphic to $\varphi_i(U_i) \equiv V_i \subset \mathbb{R}^n$), we say that $f_\lambda \to f$ in $\mathscr{D}(M) = C_c^\infty(M)$ iff for each $\psi_i \in C_c^\infty(U_i)$ one has $\|D^\alpha(\psi_i(f_\lambda - f) \circ \varphi_i^{-1})\|_\infty \to 0$, which is defined on $\mathbb{R}^n$. This turns out to be independent of the choice of the atlas (within the equivalence class defining the smooth structure on M). Elements of $\mathscr{D}(\mathbb{R}^n)$ or $\mathscr{D}(M)$ are called ***test functions***.

   The ***rapidly decreasing (test) functions*** $\mathscr{S}(\mathbb{R}^n)$ consists of those $f \in C^\infty(\mathbb{R}^n)$ for which the function $x \mapsto x^\alpha D^\beta f$ is bounded for all multi-indices α and β. One often writes

$$\langle x \rangle = (1 + \|x\|^2)^{1/2}, \tag{D.3}$$

   and uses $x \mapsto \langle x \rangle^\alpha D^\beta f$, which of course gives the same space. The topology on $\mathscr{S}(\mathbb{R}^n)$ is such that $\varphi_\lambda \to \varphi$ iff for all $l, m \in \mathbb{N}$ and multi-indices α and β with $|\alpha| \leq l$ and $|\beta| \leq m$,

$$\|x^\alpha D^\beta(\varphi_\lambda - \varphi)\|_\infty \to 0. \tag{D.4}$$

3. A ***distribution*** on $\mathbb{R}^n$ (or M) is a continuous linear map $u : \mathscr{D}(\mathbb{R}^n) \to \mathbb{C}$ (or $u : \mathscr{D}(M) \to \mathbb{C}$). The space $\mathscr{D}'(\mathbb{R}^n)$ of distributions on $\mathbb{R}^n$ carries the *weak topology*, in which $u_\lambda \to u$ iff $u_\lambda(f) \to u(f)$ for each $f \in \mathscr{D}(\mathbb{R}^n)$. Similarly for $\mathscr{D}'(M)$ and for the space $\mathscr{S}'(\mathbb{R}^n)$ of ***tempered distributions*** on $\mathbb{R}^n$. In this topology, $\mathscr{D}(\mathbb{R}^n)$ is dense in $\mathscr{D}'(\mathbb{R}^n)$, where $u \in \mathscr{D}(\mathbb{R}^n)$ defines $u \in \mathscr{D}'(\mathbb{R}^n)$ through the L^2 inner product $u(f) = \langle \overline{u}, f \rangle_{L^2(\mathbb{R}^n)}$, and similarly for $\mathscr{S}(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n)$. Adding a middle man gives so-called ***Gelfand triples***

$$\mathscr{D}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \subset \mathscr{D}'(\mathbb{R}^n); \tag{D.5}$$

$$\mathscr{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n), \tag{D.6}$$

   in which each embedding is continuous and dense. Likewise for $\mathscr{D}(M)$, provided we equip M with a measure that in coordinates has the same null sets as Lebesgue measure. For example, any (background) Riemannian metric on M provides such a measure, viz. $d\mu(x) = d^n x \sqrt{\det(g(x)}$, see §4.1.[108] Also in that case we obtain a Gelfand triple

$$\mathscr{D}(M) \subset L^2(M) \subset \mathscr{D}'(M). \tag{D.7}$$

[107] For details see Hörmander (1990), §6.3, Taylor (1996), §4.3, Grubb (2009), §8.2, Aubin (1998), Chapter 2, or E. van den Ban & M. Crainic (2013), Chapter 2.

[108] Hörmander's definition of a distribution on M coincides with the one above if we choose such a measure.

4. ***Weak derivatives***. It will be convenient from now on to write $\langle u, f\rangle$ for $u(f)$. For each multi-index α, the ***weak derivative*** $D^\alpha u$ of $u \in \mathscr{D}'(\mathbb{R}^n)$ is defined by

$$\langle D^\alpha u, f\rangle = (-1)^{|\alpha|}\langle u, D^\alpha f\rangle. \tag{D.8}$$

This definition comes from the fake formula $\langle u, f\rangle = \int_{\mathbb{R}^n} d^n x\, u(x) f(x)$, which on repeated partial integration gives (D.8). Any linear partial differential operator may therefore be regarded as a map $L : \mathscr{D}'(\mathbb{R}^n) \to \mathscr{D}'(\mathbb{R}^n)$, with adjoint $L^* : \mathscr{D}(\mathbb{R}^n) \to \mathscr{D}(\mathbb{R}^n)$, i.e.,

$$\langle Lu, f\rangle = \langle u, L^* f\rangle. \tag{D.9}$$

For example, if $L = D^\alpha$, then $L^* = (-1)^{|\alpha|} D^\alpha$. The derivatives in Lu are called ***weak***, those in $L^* f$ being ***classical***. Similarly, a solution $u \in \mathscr{D}'(\mathbb{R}^n)$ of a linear PDE $Lu = F$ (with initial conditions), i.e. $\langle Lu, f\rangle = \langle u, L^* f\rangle$ for all $f \in \mathscr{D}(\mathbb{R}^n)$, is called ***weak***.

5. ***Sobolev spaces.*** For any $s \in \mathbb{N}$, based on (D.5), define the ***Sobolev space***

$$H^s(\mathbb{R}^n) = \{u \in L^2(\mathbb{R}^n) \mid D^\alpha u \in L^2(\mathbb{R}^n) \forall \alpha : |\alpha| \leq s\}, \tag{D.10}$$

where accordingly the derivatives inherent in D^α are weak. Clearly, $H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n)$, but it can be shown that all $H^s(\mathbb{R}^n)$ are Hilbert spaces with respect to the inner product

$$\langle u, v\rangle_s = \sum_{|\alpha| \leq s} \langle D^\alpha u, D^\alpha v\rangle_{L^2}. \tag{D.11}$$

A different perspective on Sobolev spaces comes from the ***Fourier transform***

$$\hat{f}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} d^n x\, f(x) e^{-i\xi x}; \tag{D.12}$$

$$\check{f}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} d^n \xi\, f(\xi) e^{i\xi x}, \tag{D.13}$$

which make sense as Lebesgue integrals for $f \in L^1(\mathbb{R}^n)$. If one also has $\hat{f} \in L^1(\mathbb{R}^n)$, then

$$\check{\hat{f}} = f. \tag{D.14}$$

The scope of these formulae may be extended in at least three different ways:

(a) Eq. (D.12) yields a unitary isomorphism $L^2(\mathbb{R}^n) \stackrel{\cong}{\to} L^2(\mathbb{R}^n)$ of Hilbert spaces.

(b) The Fourier transform also defines a linear homeomorphism $\mathscr{S}(\mathbb{R}^n) \stackrel{\cong}{\to} \mathscr{S}(\mathbb{R}^n)$.

(c) Defining $\hat{f}$ for $f \in \mathscr{S}'(\mathbb{R}^n)$ by $\langle \hat{f}, \varphi\rangle = \langle f, \check{\varphi}\rangle$, the Fourier transform (D.12) even defines a linear homeomorphism $\mathscr{S}'(\mathbb{R}^n) \stackrel{\cong}{\to} \mathscr{S}'(\mathbb{R}^n)$ of tempered distributions.

Using the notation (D.3) for ξ, we may then (re)define, *for any* $s \in \mathbb{R}$,

$$H^s(\mathbb{R}^n) = \{u \in \mathscr{S}'(\mathbb{R}^n) \mid \xi \mapsto \langle \xi\rangle^s \hat{u}(\xi) \in L^2(\mathbb{R}^n)\}, \tag{D.15}$$

with inner product

$$\langle u, v\rangle_s = \int_{\mathbb{R}^n} d^n \xi\, \langle \xi\rangle^{2s} \overline{\hat{u}}(\xi) \hat{v}(\xi) = \int_{\mathbb{R}^n} d^n \xi\, (1 + \|\xi\|^2)^s \overline{\hat{u}}(\xi) \hat{v}(\xi) \tag{D.16}$$

For $s \in \mathbb{N}$ this reproduces (D.10) as a vector space. Though the inner products (D.11) and (D.16) are different (one has to specify which one is used), they induce equivalent norms.

Sobolev spaces can also be defined on manifolds: for $u \in \mathscr{D}'(M)$ de define $u \in H^2(M)$ iff for each chart (U_i, φ_i) and $\chi_i \in C_c^\infty(V_i)$, where $V_i = \varphi_i(U_i) \subset \mathbb{R}^n$, the distribution $u \circ \varphi_i^{-1}\chi_i$ on $\mathscr{D}(\mathbb{R}^n)$, defined on $f \in \mathscr{D}(\mathbb{R}^n)$ by $\langle u \circ \varphi_i^{-1}\chi_i, f\rangle = \langle u, (\chi_i f) \circ \varphi_i\rangle$, is in $H^s(\mathbb{R}^n)$.

Theorem D.1 *Let M be a compact manifold (with a complete Riemannian metric) or $\mathbb{R}^n$.*

(a) For each $s \in \mathbb{R}$ the space $\mathscr{D}(M)$ is dense in $H^2(M)$.

(b) For each $s \in \mathbb{R}$ we have an isometric (Banach space) isomorphism

$$H^s(M)^* \cong H^{-s}(M), \tag{D.17}$$

understood in the following way:[109] *any continuous functional $\Lambda \in H^s(M)^*$ corresponds linearly, bijectively, and isometrically to $f \in H^{-s}(M)$ via*

$$\Lambda(u) = \langle \overline{f}, u\rangle_{L^2(M)}. \tag{D.18}$$

*(c) If $s > \frac{1}{2}n + k$, then $H^s(M) \subset C_b^k(M)$, where the embedding is continuous with respect to the norm $\|u\|_{m,\infty} = \sum_{|\alpha|\leq m} \|D^\alpha u\|_\infty$ on $C^k(M)$ (*Sobolev embedding theorem*).*

*(d) For any $s \in \mathbb{R}$ and $\delta > 0$ the injection $H^{s+\delta}(M) \hookrightarrow H^s(M)$ is compact (*Rellich*).*

(e) For $s > 0$ we have our fourth Gelfand triple

$$H^s(M) \subset L^2(M) \subset H^{-s}(M). \tag{D.19}$$

6. For PDE's, for fixed $T > 0$ one considers Banach spaces like $C([0,T], H^s(M))$, with norm

$$\|u\|_\infty = \sup_{t\in[0,T]} \|u(t)\|_s, \tag{D.20}$$

or $C^1([0,T], H^s(M))$ with analogous norm, or $L^p([0,T], H^s(M))$, $1 \leq p < \infty$, normed by

$$\|u\|_p = \left(\int_0^T dt\, (\|u(t)\|_s)^p\right)^{1/p}, \tag{D.21}$$

or $L^\infty([0,T], H^s(M))$, with norm

$$\|u\|_\infty = \operatorname{ess\,sup}_{t\in[0,T]} \|u(t)\|_s. \tag{D.22}$$

Here we define $L^p([0,T], H^s(M))$, $1 \leq p < \infty$, as the completion of $C([0,T], H^s(M))$ in the norm (D.21), and also (avoiding Banach space-valued measurable functions), *define* $L^\infty([0,T], H^s(M))$ as the (Banach) dual of $L^1([0,T], H^{-s}(M))$, in that we identify $f \in L^\infty([0,T], H^s(M))$ with the functional $\Lambda_f \in (L^1([0,T], H^{-s}(M)))^*$ given by

$$\Lambda_f(g) = \int_0^T dt\, \langle f(t), g(t)\rangle_{L^2(M)}. \tag{D.23}$$

[109] Also, $H^s(M)^* \cong H^s(M)$ through its own inner product; the pairing in (D.18) is through the L^2 inner product.

Literature

Books on Ricci Flow

B. Andrews and C. Hopper (2011), *The Ricci Flow in Riemannian Geometry* (Springer).

S. Brendle (2010), *Ricci Flow and the Sphere Theorem* (AMS).

B. Chow and D. Knopf (2004), *Ricci Flow: An Introduction* (AMS).

M. Gessen (2009), *Perfect Rigor: A Genius and the Mathematical Breakthrough of the Century* (HMH).

J.W. Morgan and F. Fong (2010), *Ricci Flow and Geometrization of 3-Manifolds* (AMS).

J. W. Morgan and G. Tian (2007), *Ricci Flow and the Poincaré Conjecture* (AMS).

Other books

T. Aubin (1998), *Some Nonlinear Problems in Riemannian Geometry* (Springer).

M. Berger (2003), *A Panoramic View of Riemannian Geometry* (Springer).

R. Carter, G. Segal, and I. MacDonald (1995), *Lectures on Lie Groups and Lie Algebras* (LMS).

J. Cheeger and D.G. Ebin (1975), *Comparison Theorems in Riemannian Geometry* (North-Holland).

Y. Choquet-Bruhat (2009), *General Relativity and the Einstein Equations* (OUP).

L.C. Evans (2010), *Partial Differential Equations*, Second Edition (AMS).

R. Goodman & N.R. Wallach (2009), *Symmetry, Representations, and Invariants* (Springer).

G. Grubb (2009), *Distributions and Operators* (Springer).

S. Helgason (1978), *Differential Geometry, Lie Groups, and Symmetric Spaces* (Academic Press).

L. Hörmander (1990), *The Analysis of Linear Partial Differential Operators I*, Second Ed. (Springer),

J. Jost (2002), *Riemannian Geometry and Geometric Analysis*, Third edition (Springer).

S. Kobayashi and K. Nomizu (1963), *Foundations of Differential Geometry, Vol. I* (Wiley).

S. Kobayashi and K. Nomizu (1969), *Foundations of Differential Geometry, Vol. II* (Wiley).

G.M. Lieberman (2005), *Second Order Parabolic Differential Equations* (World Scientific).

M. Renardy and R.C. Rogers (2004), *An Introduction to Partial Differential Equations* (Springer).

H.P. de Saint-Gervais (2010), *Uniformization of Riemann Surfaces* (EMS).

A. Scorpan (2005), *The Wild World of 4-Manifolds* (AMS).

M.E. Taylor (1996), *Partial Differential Equations I* (Springer).

A. Terras (2016), *Harmonic Analysis on Symmetric Spaces: Higher Rank Spaces, Positive Definite Matrix Space and Generalizations* (Springer).

E.B. Vinberg, ed. (1993), *Geometry II: Spaces of Constant Curvature* (Springer).

J.A. Wolf (2010), *Spaces of Constant Curvature*, 6th Edition (AMS Chelsea Publishing).

E. Zeidler (1995), *Applied Functional Analysis: Applications to Mathematical Physics* (Springer).

Articles

S. Brendle (2011), Evolution equations in Riemannian geometry, *Japan J. Math.* 6, 45–61.

H.D. Cao and X.P. Zhu (2006), A Complete Proof of the Poincaré and Geometrization Conjectures–Application of the Hamilton-Perelman theory of the Ricci flow, *Asian J. Math.* 10, 165–492.

D.M. DeTurck (1983), Deforming metrics in the direction of their Ricci tensors, *J. Diff. Geometry* 18, 157–162.

D.M. DeTurck and J.L. Kazdan (1981), Some regularity theorems in Riemannian geometry, *Annales scientifiques de l' É.N.S. (4^{e} série)* 14, 249–260.

R.S. Hamilton (1982), Three-manifolds with positive Ricci curvature, *J. Diff. Geometry* 17, 255–306.

J. Isenberg and J. Jackson (1992), Ricci flow of locally homogeneous geometries on closed manifolds, *J. Diff. Geometry* 35, 723–741.

J.M. Lee and T.H. Parker (1987), The Yamabe Problem, *Bull. American Math. Soc.* 17, 37–91.

J. Milnor (1976), Curvatures of left invariant metrics on Lie groups, *Adv. Math.* 21, 293–329.

J. Milnor (1982), Hyperbolic geometry: The first 150 years, *Bull. American Mathematical Society* (NS) 6, 9–24.

J. Milnor (2003), Towards the Poincaré Conjecture and the classification of 3-manifolds, *Notices of the AMS* 50, 1226–1233.

J. Milnor (2014), Topology through the centuries: Low dimensional manifolds (Abel Prize lecture), *Bull. American Mathematical Society* (NS) 52, 545–584.

P. Scott (1983), The Geometries of 3-manifolds, *Bull. London Math. Soc.* 15, 401–487.

W.P. Thurston (1982), Three-dimensional manifolds, Kleinian groups and hyperbolic geometry, *Bull. American Mathematical Society* (New Series) 6, 357–381.

Online resources

M.T. Anderson (2004), Geometrization of 3-manifolds via the Ricci Flow, *Notices of the AMS* 51, 184–193, `http://quantumuniverse.eu/Tom/fea-anderson.pdf`.

R. Bamler (2014), *Ricci Flow: Lecture Notes*, `http://www.mit.edu/~cmad/Papers/RicciFlowNotes.pdf`.

E. van den Ban & M. Crainic (2013), *Analysis on Manifolds: Lecture Notes for the 2012/2013 mastermath course*, `http://www.staff.science.uu.nl/∼crain101/AS-2013/chapters-1-2-3.pdf`.

C. Bär (2007/08), *Geometrische Analysis*, `https://www.math.uni-potsdam.de/fileadmin/user_upload/Prof-Geometrie/Dokumente/Lehre/Lehrmaterialien/skript-GeomAna.pdf`.

L. Bessières, G. Besson, M. Boileau, S. Maillot, and J. Porti (2010), *Geometrization of 3-manifolds*, `http://mat.uab.cat/~porti/book.pdf`.

G. Heckman (2019), *Introduction to Riemannian Geometry*, `http://www.math.ru.nl/~heckman/DiffGeom.pdf`.

J.L. Kazdan, Another proof of Bianchi's identity in Riemannian geometry, *Proc. AMS* 81, 341–342 (1981).

B. Kleiner and J. Lott (2006), *Notes on Perelman's papers*, `arXiv:math/0605667`.

K. Landsman (2018), *Foundations of General Relativity* (preliminary draft), `https://www.math.ru.nl/~landsman/FGRfinal.pdf` .

J. W. Morgan (2006), *The Poincaré Conjecture*, `http://www.icm2006.org/proceedings/Vol_I/30.pdf`.

J. W. Morgan and G. Tian (2007), *Ricci Flow and the Poincaré Conjecture*, `arxiv:0607607`.

M. Mukherjee (2012), *Ricci flows and assorted applications of PDE to geometry*, M.Sc. Thesis, `http://mukherjee.web.unc.edu/files/2014/09/masters-paper-cont-511.pdf`.

G. Perelman (2002), The entropy formula for the Ricci flow and its geometric applications, `arXiv:math/0211159`.

G. Perelman (2003a), Ricci flow with surgery on three-manifolds, `arXiv:math/0303109`.

G. Perelman (2003b), Finite extinction time for the solutions to the Ricci flow on certain three-manifolds, arXiv:math/0307245.

N. Sheridan (2006), *Hamilton's Ricci Flow*, M.Sc. Thesis, https://web.math.princeton.edu/~nsher/ricciflow.pdf.

I. Slegers (2015), *Ricci Flow on Surfaces*, B.Sc. Thesis, https://dspace.library.uu.nl/handle/1874/318805.

B. Stetler (2011), *The Ricci flow on surfaces and the uniformization theorem*, M.Sc. Thesis, http://mathematics.stanford.edu/wp-content/uploads/2013/08/Stetler-Honors-Thesis-2011.pdf.

P. Topping (2006), *Lectures on the Ricci Flow*, http://homepages.warwick.ac.uk/~maseq/topping_RF_mar06.pdf.

T. Tao (2008), https://terrytao.wordpress.com/category/teaching/285g-poincare-conjecture/.

YouTube

Richard Hamilton, *The Poincaré Conjecture* (ICM, 2006),[110] https://www.youtube.com/watch?v=fymCXcIt20g.

John W. Morgan, *Ricci flow and its applications to 3-manifold topology*,[111] https://www.youtube.com/watch?v=ho9RsgfehVg.

Georgi Perelman documentary,[112] https://www.youtube.com/watch?v=Ng1W2KUHI2s.

Perelman: De documentaire,[113] https://www.youtube.com/watch?v=RQjtcxR-mqo.

[110] Very clear plenary talk at the ICM 2006 by the inventor of Ricci Flow.

[111] Excellent first introduction to the subject, steadily interrupted however by an annoying Russian smart ass.

[112] Excellent documentary in Russian, switch subtitles on!

[113] Pathetic attempt by three Dutch film students to get hold of Perelman in St Petersburg, without any knowledge of either Russian or mathematics, not even bringing an interpreter. They turn vice into virtue by presenting their failure as some sort of a road movie, which even seems to have achieved cult status in some circles (in Dutch).