Compact Surfaces of Least Total Curvature

Alexander Polden, Sonderforschungsbereich 382, University of Tübingen
February 10th, 1997

The Uniformisation Theorem: *every metric on a compact surface is conformally equivalent to one of constant scalar curvature.* In this work, we show that the uniform metrics arise as the solutions to a natural variational problem for the total curvature of a surface.

This mirrors the situation in dimensions three and above. The Uniformisation Theorem finds a perfect analogue in higher dimensions in Yamabe's theorem — simply substitute 'n-manifold' for the word 'surface' in the above statement. In this case, the constant curvature metrics are precisely the critical points for $\int R\,d\mu$ with respect to fixed volume and conformal structure.

In two dimensions, however, $\int R\,d\mu$ is a topological constant, and the notion of minimal total curvature becomes a meaningless one.

At least, it becomes meaningless unless we consider a different curvature integral. In this work, we measure the total curvature more abstractly, as some $\int F(R)\,d\mu$. This definition is likely to be scale-dependent, so when we speak of surfaces of least total curvature, this means minimisers in the class of surfaces having a certain fixed volume. We state natural conditions on F which ensure that the set of critical points is ideally small: the stationary surfaces are precisely those with constant curvature, and these are all absolute minimisers in their topologies.

In fact, things turn out even better. Even minimising within the more restrictive framework of a conformal class (that is to say, minimising with respect only to conformal variations of the metric), the outcome is the same: only the uniform metrics are minimal, or even stationary.

The second part of this work is concerned with descent flows for $\int R^2\, d\mu$ on the 2-sphere. These are governed by parabolic equations of fourth and sixth order.

Flows which diminish $\int R^2\, d\mu$ are of obvious geometric interest. They are of interest on other counts too: as for instance in the work of Chruściel [Chruściel 1990] on Robinson-Trautman spacetimes. These may be split into a foliation whose leaves form the solution to a flow of this type, and Chruściel, by solving this flow in complete detail, provides a very sharp picture of the nature of such a spacetime.

Chruściel's result is entirely clear-cut, but it applies only to one very specific flow problem. The analysis pivots on a miraculous monotonicity property discovered by Singleton [Singleton 1989]. In this work, we produce techniques appropriate to an entire class of flows.

The foremost tool in the analysis of parabolic equations is the maximum principle. In geometric problems of this sort, it is almost the *only* tool. But the maximum principle begins and ends with second-order equations; it is of no help to us here.

Nor does the Sobolev space theory provide an immediate replacement. The 'constants' in the standard integral inequalities are no longer constant: they depend on the surface, and the surface is continuously changing.

The favourite example of singularity formation in a geometric evolution is the 'neck-pinching' phenomenon of the mean-curvature and Ricci flows. On a 2-sphere, the constant in the Sobolev inequality has the character of an isoperimetric ratio (see, for instance, [Cheeger 1970]). In the event of a neck-pinch, it really would degenerate entirely.

The existence theory therefore splits into two parts. The first of these is controlling the Sobolev constant. This is accomplished through estimates for the functional defined by Berger in his investigation of the uniformisation problem [Berger 1971]. This control legitimises the various integral inequalities. The second part is the derivation of integral estimates for the curvature using the Sobolev space interpolation estimates of Gagliardo and Nirenberg. This step parallels the analysis of the corresponding one-dimensional problem [Polden 1995], and leads directly to long-time existence and convergence results.

This work includes results from the author's doctoral thesis at the University of Tübingen.

1. The Variational Problem

Consider now a compact, orientable surface M^2 with metric g and corresponding measure $d\mu$. Let R be its curvature, and $\mathcal{E}$ its total curvature in the sense above:

$$\mathcal{E}(g) = \int_{M^2} F(R)\, d\mu.$$

In general, we shall make assumptions about F only when they are needed. From the outset, however, F will be even, smooth and strictly convex.

We shall also refer to $\mathcal{E}$ as the *energy* of the surface. $\mathcal{E}$ is not scaling-invariant. In the natural case where F is non-negative and coercive, the energy of a surface may be diminished as far as one pleases simply by blowing the metric up uniformly. Assume, therefore, that the volume is fixed: $V(g) = 1$. This allows us to apply Jensen's inequality:

$$\mathcal{E}(g) = \int_{M^2} F(R)\, d\mu \geq F\left(\int_{M^2} R\, d\mu\right) = F(4\pi\chi(M^2)),$$

with equality precisely when R is constant: the uniform metrics are absolute minimisers for $\mathcal{E}$ in their topologies.

The model here is of course the case $F(x) = x^2$:

$$\mathcal{R}(g) = \int_{M^2} R^2 \, d\mu.$$

Because the value of $\int R \, d\mu$ is legislated by the Gauß-Bonnet theorem, fixing the volume (a) fixes in turn the average, r, of the curvature and (b) shows that, for any real number σ, the associated integral

$$\mathcal{R}_\sigma = \int_{M^2} (R - \sigma)^2 \; d\mu$$

differs from $\mathcal{R}$ by addition of a constant. For the purpose of minimisation, then, the two energies are equivalent. In particular, one might equally well consider the integral

$$\mathcal{R}^* = \int_{M^2} (R - r)^2 \; d\mu,$$

and the outcome will be the same. In intuitive terms, the flattest closed surfaces with respect to $\mathcal{R}$ are at once also the roundest, not to say the most hyperbolic.

Now we compute the variational equations. Let g be a *conformal* variation of the metric g_0 — that is, consider a one–parameter family of metrics $g(\bullet, t)$ varying according to the equation

$$\frac{\partial}{\partial t} g_{ij} = \eta(x,t) \cdot g_{ij}. \tag{1}$$

The metrics $g(\bullet, t)$ can be expressed relative to g_0 (or, for that matter, relative to any fixed metric in the conformal class) by way of a conformal factor u:

$$g(x,t) = e^{2u(x,t)} \cdot g_0(x); \tag{2}$$

and in turn, the geometric attributes of $g(\bullet, t)$ — measure, volume, curvature and Laplace operator — can be related back to those of g_0:

$$d\mu = e^{2u} d\mu_0, \qquad V = \int_{M^2} e^{2u} d\mu_0$$

$$R = e^{-2u} \left(R_0 - 2\Delta_0 u \right) \qquad \Delta = e^{-2u} \Delta_0.$$

From (1), we compute the variation of the conformal factor:

$$\frac{\partial}{\partial t} g_{ij} \;=\; \frac{\partial}{\partial t} \left(e^{2u} \left(g_0 \right)_{ij} \right) \;=\; 2 \frac{\partial u}{\partial t} \, e^{2u} \left(g_0 \right)_{ij} \;=\; 2 \frac{\partial u}{\partial t} g_{ij},$$

but $\partial g / \partial t = \eta g$, and so

$$\frac{\partial u}{\partial t} = \frac{1}{2} \eta. \tag{3}$$

Knowing this much, it is then an easy matter to derive the variations in the measure and curvature, using the formulas above:

Lemma 1.1. *From (1), we derive the following variational equations*

$$\frac{\partial}{\partial t} g^{ij} = -\eta \cdot g^{ij} \qquad \frac{\partial}{\partial t} d\mu = \eta \, d\mu \qquad \frac{\partial V}{\partial t} = \int_{M^2} \eta \, d\mu$$

$$\frac{\partial R}{\partial t} = -\Delta\eta - R\eta \qquad \frac{\partial \mathcal{E}}{\partial t} = -\int_{M^2} \omega_F \eta \, d\mu$$

in which $\omega_F = \Delta f(R) + Rf(R) - F(R), f = F'$*; more to this, for any function* $\phi = \phi(x,t)$*,*

$$\frac{\partial}{\partial t}\Delta\phi = \Delta\frac{\partial\phi}{\partial t} - \eta\Delta\phi \qquad \text{and} \qquad \frac{\partial}{\partial t}|\nabla\phi|^2 = 2\left\langle \nabla\phi, \ \nabla\frac{\partial\phi}{\partial t}\right\rangle - \eta\,|\nabla\phi|^2 .$$

It follows at once that a metric which is stationary for $\mathcal{E}$ with respect to *all* conformal variations must satisfy

$$\omega_F = \Delta f + Rf - F = 0.$$

In the case $F(x) = x^2$, this is equivalent to $\Delta R + \frac{1}{2}R^2 = 0$, and the only solutions are the flat tori; but this just reflects the scale-dependence of the energy. If we impose in addition a volume constraint, the Euler equation becomes

$$\Delta f + Rf - F = \text{ constant } = \langle Rf - F\rangle , \tag{4}$$

in which $\langle Rf - F\rangle$ denotes the average value of $Rf - F$, and this now catches all of the uniform metrics. In the following section, we shall derive conditions on F which ensure that there are no other solutions beside them.

Classification of the critical points

In this section, we prove the following clear-cut theorem:

THEOREM 1.2. *If* $M^2 \neq S^2$*, then the only solutions to (4) are the metrics of constant curvature. If* $M^2 = S^2$*, then any solution to (4) must be rotationally symmetric.*

The sharper result on the sphere requires further assumptions on F:

THEOREM 1.3. *If* f *and* $1/f$ *are convex on* $\mathbb{R}_+$*, one of them strictly so, then the only solutions to (4) on* S^2 *are the metrics of constant curvature.*

These conditions cover the model case, $F(x) = x^2$; more generally, they cover $F(x) = x^{2n}$, for any $n \geq 1$.

The cornerstone of the proof is the following lemma:

LEMMA 1.4. *On a compact surface* M^2*,*

$$\omega_F \equiv \text{constant} \qquad \text{if and only if} \qquad \overset{\circ}{\nabla^2} f \equiv 0,$$

where $\overset{\circ}{\nabla^2} f$ *denotes the traceless part of the second covariant derivatives.*

PROOF. Naturally, the proof is simply computational. Assume first that $\overset{\circ}{\nabla^2} f \equiv 0$. Differentiating one further time, this becomes

$$\nabla_k\nabla_i\nabla_j f - \frac{1}{2}\nabla_k\Delta f \cdot g_{ij} = 0.$$

Now commute ∇_k and ∇_i; this introduces a curvature term, and we have:

$$\nabla_i \nabla_k \nabla_j f + \mathrm{Riem}\,{}_{kij}{}^m \cdot \nabla_m f - \frac{1}{2} \nabla_k \Delta f \cdot g_{ij} = 0.$$

Tracing with g_{kj}, this simplifies to

$$\frac{1}{2} \nabla_i \Delta f + \mathrm{Ric}\,{}_i^m \cdot \nabla_m f = 0;$$

but in two dimensions, $\mathrm{Ric} = \frac{1}{2} R \cdot g$, and so

$$\frac{1}{2} \left(\nabla_i \Delta f + R \nabla_i f \right) = 0.$$

However, $\nabla_i (Rf - F)$ is simply $R \nabla_i f$; we therefore have $\nabla_i \omega_F = 0$. Conversely, suppose ω_F is constant. Then, multiplying by Δf and integrating,

$$\int_{M^2} (\Delta f)^2 - R \, |\nabla f|^2 \, d\mu = 0. \tag{5}$$

By the Bochner identity, $\int (\Delta f)^2 \, d\mu = \int |\nabla^2 f|^2 + \frac{1}{2} R |\nabla f|^2 \, d\mu$; combining this with the previous equation in such a way as to eliminate the terms in $R|\nabla f|^2$, this becomes

$$\int_{M^2} 2 \left| \nabla^2 f \right|^2 - (\Delta f)^2 \, d\mu = 0,$$

but the left hand side is precisely $2 \int |\overset{\circ}{\nabla^2} f|^2 \, d\mu$, which settles the lemma. $\square$

The L^2–gradient of $\mathcal{E}$ with respect to *all* variations in the metric is a 2–tensor, M_{ij}. For a variation h in g, the corresponding change in $\mathcal{E}$ is simply $\int \langle M, \; h \rangle \; d\mu$. In the special case $h = \eta \cdot g$, this becomes $\int \eta \, \mathrm{tr}\, M \; d\mu$, and so (by comparison with Lemma 1.1), the trace of M has to be -2ω.

M can be split as the direct sum

$$M = \frac{1}{2} \, \mathrm{tr}\, M \cdot g \oplus \overset{\circ}{M}.$$

The trace of M is just the conformal gradient; the traceless part $\overset{\circ}{M}$ is precisely $2 \, \overset{\circ}{\nabla^2} f$. In intuitive terms, the lemma states that the complete gradient of $\mathcal{E}$ vanishes if either of its components does so. In turn, the solutions to the variational problem are the same, whether one minimises in an entire topology, or within the more restrictive scope of a conformal class.

The condition $\overset{\circ}{\nabla^2} f \equiv 0$ is geometrically an intuitive one. It says that the second derivatives of f are uniformly distributed in all directions at every point of M^2. The first consequence of this is that f fulfils a simple ODE along geodesics:

LEMMA 1.5. *Along any geodesic in M^2, f satisfies the equation*

$$\frac{d^2 f}{ds^2} = \frac{1}{2} \left(\gamma - (Rf - F) \right). \tag{6}$$

with γ here and in the future denoting the average value of $Rf - F$.

PROOF. Let $F : I \to M^2$ be a geodesic with arc-length parameter s and unit tangent vector T. Then,

$$\frac{d^2 f}{ds^2} = \nabla_T \nabla_T f,$$

and now, using the conditions $\overset{\circ}{\nabla}{}^2 f \equiv 0$ and $\omega_F \equiv \gamma - (Rf - R)$,

$$\frac{d^2 f}{ds^2} = \frac{1}{2}\Delta f \cdot \langle T,\ T\rangle = \frac{1}{2}\left(\gamma - (Rf - F)\right),$$

which is the claim of the lemma. $\square$

Let us at this point make two simple observations. First, f is critical wherever R is critical. And second, the right hand side of equation (6) is determined completely by f.

Consider now a system of polar co-ordinates (r, θ) centred at ξ, a critical point for R. Along the radial geodesics, f is determined by (6), together with the same boundary condition in every direction: $\nabla_r f|_{r=0} = 0, f(0) = f(\xi)$. In the local scale, this means that f, and hence R, are constant on geodesic circles close to ξ.

More generally, the curvature is the same at any two points equally distant from the same critical point. Indeed, the curvature is the same at two points which are equally distant from two *distinct* critical points if it is the same at the critical points themselves. This is worth singling out as a lemma of its own:

LEMMA 1.6. *if ξ_1 and ξ_2 are critical points for R with $R(\xi_1) = R(\xi_2)$, and if $\mathrm{dist}(x_1, \xi_1) = \mathrm{dist}(x_2, \xi_2)$, where the distances are measured along* any *geodesics, not necessarily minimising ones, then $R(x_1) = R(x_2)$.*

Notation: let Z denote the critical set of R; let Z_{max} and Z_{min} denote the subsets of Z on which R is maximal and minimal.

LEMMA 1.7. *Either $Z = M^2$ or Z is finite.*

PROOF. If Z is infinite, then it must accumulate at some point P. Z is closed, so P is itself a point of Z.

In this case, choose $Q \in Z$ such that $\delta = \mathrm{dist}(Q, P) < \frac{1}{4}\iota$, where ι is the injectivity radius of (M^2, g), and let ξ be any point whatever inside $B_\delta(P)$. Let C be the geodesic circle $S_\delta(P)$, and C' the geodesic circle centred at Q and passing through ξ. R is constant on C and C', by the previous lemma.

C and C' are topologically genuinely S^1s — that follows from the restriction $\delta < \frac{1}{4}\iota$ — and they cut through one another. Hence, $R(\xi) = R(Q)$. ξ was arbitrary; thus, R is constant on $B_\delta(P)$.

It follows (by the previous lemma once again) that R is constant on a disk of radius δ about any other critical point P' for which $R(P') = R(P)$. But this means the set $Z \cap \{x \in M^2 \,|\, R(x) = R(P)\}$ is open; it is obviously closed, and so it must be all of M^2. $\square$

So assume from here onwards that R is non–constant, and that the critical set Z is discrete.

LEMMA 1.8. *Let $P \in Z_{max}$ and $Q \in Z_{min}$. Then, the distance between P and Q is maximal: for any $x \in M^2$, $\mathrm{dist}(x, P) \leq \mathrm{dist}(P, Q)$ and $\mathrm{dist}(x, Q) \leq \mathrm{dist}(P, Q)$.*

PROOF. Suppose there is a point S, more distant from P than Q is. In this case, the set $\Gamma = \{x \in M^2 \,|\, \mathrm{dist}(x, P) = \mathrm{dist}(Q, P)\}$ cuts M^2 into two open subsets. In particular, Γ is infinite.

R is constant on Γ, and therefore minimal. But this means that Z is an infinite set, which implies constant curvature, by the previous lemma. This shows that Q is as far removed as possible from P; likewise P is as far removed as possible from Q. $\square$

Lemma 1.9. *R has precisely two critical points. The geodesics starting from one of these points all meet up again at the other, and they are all of the same length.*

Proof. Let $P \in Z_{max}$ and $Q \in Z_{min}$, and let $d = \operatorname{dist}(P, Q)$. Suppose that there is some other critical point, call it S, and consider the geodesic triangle PQS.

Let $d_1 = \operatorname{dist}(P, S)$ and $d_2 = \operatorname{dist}(Q, S)$; by the previous lemma, $d_1 \leq d$ and $d_2 \leq d$.

If $d_1 = d_2$, then by Lemma 1.6, $R(P) = R(Q)$, that is, $\max R = \min R$ and so R is constant. So assume $d_1 < d_2$. In this case, there is another maximum P' on the geodesic connecting Q and S. But this is now closer to Q than P itself, and this contradicts the previous lemma. $d_2 < d_1$ is impossible for the same reasons.

This rules out the presence of other critical points; S must be either P itself or Q. To prove the final claim of the lemma, note simply that every geodesic leaving P must reach a minimum at a distance of d, but this can only be Q. $\square$

Proof of Theorem 1.2. The last lemma implies that if R is nonconstant, then M^2 is topologically S^2. $\square$

Indeed, the last lemma implies more. It implies that M^2 can be covered, give or take one pole, by a single chart of polar co-ordinates centered at one of the two critical points.

To extend the arguments above to spheres, we shall have to consider the equation more precisely. To do this, we work in the global polar co-ordinate system (r, θ) just mentioned; r is arclength along the radial geodesics. In these co-ordinates, the metric has the form:

$$g_{ij} = dr \otimes dr + \phi^2 d\theta \otimes d\theta$$

where $\phi(0, \theta) = 0$, $\nabla_r \phi(0, \theta) = 1$ and $\nabla_r^2 \phi = -\frac{1}{2} R\phi$. Since R depends only on r, this equation shows the same is true of ϕ, and hence that the surface is rotationally symmetric.

Now, at last, we use the precise form of the variational equation:

Lemma 1.10. *There is a constant λ such that $\phi(r) = \lambda \nabla_r f$.*

Proof. Compute: using the conditions $\nabla_r^2 f = \frac{1}{2}(\gamma - (Rf - F))$ and $\nabla_r^2 \phi = -\frac{1}{2} R\phi$,

$$\nabla_r \left(\nabla_r \phi \, \nabla_r f - \phi \, \nabla_r^2 f \right) = 0.$$

Thus, $\nabla_r \phi \, \nabla_r f - \phi \, \nabla_r^2 f$ is constant on M^2; in fact, it vanishes, because $\phi(0, \theta) = 0$ and $\nabla_r f(0, \theta) = 0$. But this means that $\nabla_r (\nabla_r f / \phi) = 0$, which proves the lemma. $\square$

Proof of Theorem 1.3. We have already shown that if R is nonconstant, then M^2 is a rotationally symmetric S^2 with the metric and the curvature linked by the previous lemma. Now we show that this too leads to constant curvature under the convexity assumptions of Theorem 1.3.

Let $f_{\max}$ and $f_{\min}$ be the minimum and maximum values of f; assume the radial geodesics are parametrised by arc-length from 0 to L. Then, the volume of (M^2, g) is given by

$$V = \int_{M^2} d\mu = 2\pi \int_0^L \phi \, dr = 2\pi\lambda \int_0^L \nabla_r f \, dr = 2\pi\lambda (f_{\max} - f_{\min}).$$

Equally,

$$\int_{M^2} f\,d\mu = 2\pi\int_0^L f\phi\,dr = 2\pi\lambda\int_0^L f\nabla_r f\,dr = \pi\lambda(f_{\max}^2 - f_{\min}^2).$$

Comparing these two, we see that the average of f is simply $\langle f\rangle = \frac{1}{2}(f_{\max}+f_{\min})$. Moreover,

$$\begin{aligned}\int_{M^2} Rf\,d\mu &= 2\pi\int_0^L Rf\phi\,dr = -4\pi\int_0^L f\nabla_r^2\phi\,dr \\ &= -4\pi\left((f\nabla_r\phi)|_0^L - \int_0^L \nabla_r\phi\,\nabla_r f\,dr\right) \\ &= 4\pi(f_{\max}+f_{\min}) + 4\pi\int_0^L \phi\,\nabla_r\phi\,dr\end{aligned}$$

but the last integral is obviously zero, because $\phi(0) = \phi(L) = 0$. Combining these three integrals with Gauß-Bonnet,

$$V\cdot\int_{M^2} Rf\,d\mu = \int_{M^2} R\,d\mu\cdot\int_{M^2} f\,d\mu.$$

The convexity conditions, together with Jensen's inequality, allow us to conclude from this last equation that R is constant. Set $h(x) = xf^{-1}(x)$, so that $Rf(R) = h(f(R))$. h is convex — this follows by tedious computation from the evenness of F and the convexity of $1/f$ for positive x. So, assuming once more that $V = 1$, writing $f_+ = \max\{f, 0\}$, and applying Jensen's inequality,

$$\int R\int f_+ \geq \int R\int f = \int Rf \geq \int Rf_+ = \int h(f_+) \geq h\left(\int f_+\right) = \int f_+\cdot f^{-1}\left(\int f_+\right).$$

It follows that

$$f_+\left(\int_{M^2} R\,d\mu\right) \geq \int_{M^2} f_+(R)d\mu,$$

but f_+ is a convex function, so this must be an equality. The same is therefore true of the Jensen step in the preceding equation: $\int h(f_+)\,d\mu = h(\int f_+\,d\mu)$. But one or other of the functions f_+ and h must be strictly convex for positive x, by assumption, and equality therefore implies that R is constant. □

2. Descent flows for $\int R^2\,d\mu$

In the previous section, we showed that the variational problem for $\mathcal{R}$ has ideally few solutions. Since it is well known that there is a metric of constant curvature in every conformal class of metrics on a compact surface, in some sense, the case is closed; the problem is solved completely. But it is still of considerable interest to study the associated steepest descent flows.

Starting from a given initial metric, we shall consider the deformation process

$$\frac{\partial}{\partial t}g_{ij} = (\omega - \langle\omega\rangle)\,g_{ij}, \tag{7}$$

where for the remainder of this work, ω is the speed $\Delta R + \frac{1}{2}R^2$. This is the flow of most rapid descent amongst all conformal, volume-preserving variations of the metric. In parallel, we shall study the behaviour of a metric under the flow

$$\frac{\partial}{\partial t}g_{ij} = -\Delta\omega\cdot g_{ij}. \tag{8}$$

This flow of *sixth* order also fixes the volume whilst decreasing $\int R^2\,d\mu$; in fact, it is the flow of steepest descent with respect to the L^2 norm of the *potential* of the speed.

THEOREM 2.1. *Any smooth initial metric g_0 on M^2 may be extended to a time-dependent metric g_t, $0 \leq t \leq \epsilon$ which solves equation (7). Similarly, one can always find short-time solutions to (8).*

PROOF. see [Polden 1996]. Chapter 2, Theorem 2.5.1. □

The regularity theory for these equations resembles that for the corresponding one-dimensional problem (see [Polden 1996], Theorem 3.2.2.1). We have all the same ingredients: an a priori bound for the L^2-norm of curvature and a parabolic equation in which the leading term is a power of the laplacian. In contrast to the one-dimensional case, however, stating Sobolev and Poincaré inequalities which hold uniformly for all the metrics g_t is a very difficult matter.

Let g_t be an evolving metric, and let $g^\sharp$ be some fixed representative metric from the same conformal class. Let u be the conformal factor relating g_t back to $g^\sharp$: $g_t = \exp 2u(x,t) \cdot g^\sharp$. The evolution of u is given by

$$\frac{\partial u}{\partial t} = \frac{1}{2}\left(\omega - \langle\omega\rangle\right) \quad \text{(fourth order)} \quad \text{or} \quad \frac{\partial u}{\partial t} = -\frac{1}{2}\Delta\omega \quad \text{(sixth order)}.$$

The first and most difficult step in the analysis is to obtain a sup bound for u — that is, to show that the metrics g_t remain uniformly equivalent for all time. This in turn bounds the constant in the Sobolev inequality.

Expressed in terms of the background metric, the bound for the total squared curvature reads

$$\int_{M^2} e^{-2u}\left(R^\sharp - 2\Delta^\sharp u\right)^2 \, d\mu^\sharp \leq \mathcal{R}_0. \tag{9}$$

From this we obtain the following regularity result:

LEMMA 2.2. *If $||e^u||_{L^{2+\epsilon}(d\mu^\sharp)} \leq A$ in some time interval $0 \leq t \leq T$, then*

$$\sup_{M^2 \times [0,T)} |u| \leq C(A).$$

PROOF. The result is essentially the same as a theorem of Chruściel; this proof is very much simpler, conceptually and technically. For any $\phi \in C^\infty(M^2)$,

$$\int_{M^2} e^{-u}\big(R^\sharp - 2\Delta^\sharp u\big)\,\phi\, d\mu^\sharp \leq \mathcal{R}_0^{\frac{1}{2}}\, ||\phi||_{L^2(d\mu^\sharp)}^{\frac{1}{2}}.$$

Now we conduct a Moser-style iteration over the powers of e^u. Set $\phi = e^{(2a+1)u}$:

$$\int_{M^2} \left(e^{2au}R^\sharp - 2e^{2au}\Delta^\sharp u\right)\, d\mu^\sharp \leq \mathcal{R}_0^{\frac{1}{2}}\, ||e^u||_{L^{4a+2}(d\mu^\sharp)}^{2a+1}.$$

The first term on the left is easily subsumed into the right hand side; integrating the second by parts gives

$$\frac{4}{a}\int_{M^2} \left|\nabla^\sharp e^{au}\right|^2 \, d\mu^\sharp \leq C(\mathcal{R}_0)\, ||e^u||_{L^{4a+2}(d\mu^\sharp)}^{2a+1}.$$

The left hand side controls any L^p norm of e^u; the Sobolev inequality gives, for any p,

$$||e^u||_{L^{2ap}(d\mu^\sharp)}^{2a} \leq aC(p,\mathcal{R}_0)\, ||e^u||_{L^{4a+2}(d\mu^\sharp)}^{2a+1}$$

and we may iterate this to a sup bound, provided the right hand side is controlled for *some* initial $a > 0$ as in the hypothesis. That gives a one-sided bound; to make it two-sided, note that once $\sup u$ is controlled, the factor e^{-2u} in (9) is bounded from below, and thus we have a uniform $W^{2,2}(d\mu^\sharp)$ estimate for u. From there, the two-sided bound follows directly from the Moser-Harnack inequality. $\square$

The L^2-norm of e^u is simply $\sqrt{V}$, which is of course fixed during our evolutions; the above result demands control on the $L^{2+\epsilon}$-norm. These two can not be reconciled.

The standard sphere in E^3 is mapped diffeomorphically into itself by conformal inversion in any other sphere to which it is orthogonal. This gives a new metric on the sphere which is isometric to the old one; the geometric invariants V and $\mathcal{R}$ are unaffected.

The inversion process swaps the inside of the auxiliary sphere with the outside. If this sphere is very small, that means a very small portion of the standard S^2 swaps places with a very large one; the conformal factor must be very large on the small piece, and very small on the large piece. Letting the radius of the auxiliary sphere approach zero, $u \to \pm\infty$; it certainly doesn't remain bounded.

So we absolutely have to derive an $L^{2+\epsilon}$ estimate for the conformal factor. To do this, we shall consider the energy introduced by Berger [Berger 1971] in his work on the uniformisation problem. It should be no surprise that this should make an appearance; it is the natural two-dimensional counterpart to the Yamabe functional.

For $u \in W^{1,2}(M^2)$, the Berger functional is

$$\mathcal{I}(u) = \int_{M^2} \left|\nabla^\sharp u\right|^2 + R^\sharp u \, d\mu^\sharp.$$

Minimisers for the Berger functional are precisely those conformal factors u for which the metric $\exp 2u \cdot g^\sharp$ has constant curvature.

That statement was made without any consideration of the choice of background metric $g^\sharp$. In fact, the Berger functional has the following transformation property: suppose $g^\sharp = \exp 2v \cdot g^\flat$, and let $\mathcal{I}^\sharp$ and $\mathcal{I}^\flat$ denote the functionals determined by the two metrics. Then $\mathcal{I}^\sharp(u) + \mathcal{I}^\flat(v) = \mathcal{I}^\flat(u+v)$, for any u.

It is probably more natural to define the Berger functional on metrics conformal to $g^\sharp$: if $g = \exp 2u \cdot g^\sharp$, then $\mathcal{I}^\sharp(g) = \mathcal{I}^\sharp(u)$. Expressed in this way, $\mathcal{I}^\sharp(g)$ is minimal when g has constant curvature, and the transformation law becomes: $\mathcal{I}^\flat(g) = \mathcal{I}^\sharp(g) + \mathcal{I}^\flat(g^\sharp)$. So, while this definition still depends upon the choice of background metric, the dependence is slight; the background metric determines $\mathcal{I}$ up to a constant additive factor. In particular, if one now takes the derivative of $\mathcal{I}(g_t)$ over a one-parameter family of metrics, this additive constant disappears, and the result really does depend only on g_t.

Lemma 2.3. *For any choice $g^\sharp$ of background metric, the evolution of the Berger functional is given by*

$$\frac{d\mathcal{I}^\sharp}{dt} = \frac{1}{2}\int_{M^2} R\left(\omega - \langle\omega\rangle\right) d\mu \quad \text{(fourth order)} \quad \text{or} \quad \frac{d\mathcal{I}^\sharp}{dt} = \frac{1}{2}\int_{M^2} -R\Delta\omega \, d\mu \quad \text{(sixth order)}.$$

Proof. Let η be the speed; then $\partial u/\partial t = \frac{1}{2}\eta$, and by simple computation,

$$\begin{aligned}\frac{d}{dt}\int_{M^2} \left|\nabla^\sharp u\right|^2 + R^\sharp u \, d\mu^\sharp &= \int_{M^2} \left\langle \nabla^\sharp u, \ \nabla^\sharp \eta\right\rangle + \frac{1}{2}R^\sharp \eta \, d\mu^\sharp \\ &= \int_{M^2} \eta\left(\frac{1}{2}R^\sharp - \Delta^\sharp u\right) d\mu^\sharp;\end{aligned}$$

the lemma follows now substituting $R = e^{-2u}(R^\sharp - 2\Delta^\sharp u)$ and $d\mu = e^{2u} d\mu^\sharp$. $\square$

Corollary 2.4. *For any choice of $g^\sharp$, we have the bounds*

$$\mathcal{I}^\sharp(g_t) \le \mathcal{I}^\sharp(g_0) + \frac{1}{4}\sqrt{(\mathcal{R}_0^2 - \mathcal{R}_t^2)\,t} \qquad \text{(fourth order)} \quad \text{and}$$

$$\mathcal{I}^\sharp(g_t) \le \mathcal{I}^\sharp(g_0) \qquad \text{(sixth order)}.$$

Proof. The fourth-order bound first:

$$\begin{aligned}\frac{d\mathcal{I}^\sharp}{dt} &= \frac{1}{2}\int_{M^2} R\,(\omega - \langle\omega\rangle)\,d\mu \\ &\le \frac{1}{2}\left(\int_{M^2} R^2\,d\mu\right)^{\frac{1}{2}}\left(\int_{M^2}(\omega - \langle\omega\rangle)^2\,d\mu\right)^{\frac{1}{2}} \\ &= \frac{1}{2}\mathcal{R}^{\frac{1}{2}}\left(-\frac{d\mathcal{R}}{dt}\right)^{\frac{1}{2}} = \frac{1}{4}\left(-\frac{d}{dt}\mathcal{R}^2\right)^{\frac{1}{2}}\end{aligned}$$

and now integrating over time, and applying the Hölder inequality, we have the estimate for $\mathcal{I}^\sharp$ in the fourth-order case. The sixth-order flow behaves even better: in this case,

$$\begin{aligned}\frac{d\mathcal{I}^\sharp}{dt} &= -\int_{M^2}\Delta R\left(\Delta R + \frac{1}{2}R^2\right)\,d\mu \\ &= -\int_{M^2}(\Delta R)^2 - R\,|\nabla R|^2\,d\mu,\end{aligned}$$

but by the Bochner identity, as in Lemma 1.4, this is simply $-\int |\overset{\circ}{\nabla^2} R|^2\,d\mu$, and so $\mathcal{I}^\sharp$ decreases monotonically. □

So far, the background metric $g^\sharp$ has been left unchosen. Suppose that M^2 is topologically the sphere, and let $g^\sharp$ be $g^\natural$, the standard metric on the sphere.

The bound on the Berger functional, more or less, estimates the $W^{1,2}$ norm of u. We now make use of the exponential Sobolev embedding this entails to derive estimates for e^u. Unfortunately, the result requires a strong assumption of symmetry.

Lemma 2.5. *If g_0 is antipodally symmetric — that is to say, if the conformal factor u for g_0 relative to $g^\natural$ is antipodally symmetric — then there is a $\delta > 0$ such that*

$$||e^u||_{L^{2+\delta}(d\mu^\natural)} \le C\exp\left(\frac{1}{4\pi}\mathcal{I}^\natural(g_0) + \frac{1}{16\pi}\sqrt{(\mathcal{R}_0^2 - \mathcal{R}_t^2)\,t}\right) \qquad \text{(fourth order)} \quad \text{and}$$

$$||e^u||_{L^{2+\delta}(d\mu^\natural)} \le C\exp\frac{1}{4\pi}\mathcal{I}^\natural(g_0) \qquad \text{(sixth order)}.$$

Proof. The estimates follow from the Moser exceptional Sobolev inequality on the sphere (see, for instance, [Aubin 1982]): if $\int_{S^2} f\,d\mu^\natural = 0$, then

$$\int_{S^2}\exp f\,d\mu^\natural \le C\exp\left(\frac{1}{16\pi}\int_{S^2}|\nabla^\natural f|^2\,d\mu^\natural\right).$$

The constant $1/16\pi$ is optimal. Setting $f = 2\,(u - \langle u\rangle)$, we have

$$\int_{S^2} \exp 2u \, d\mu^\natural \le C \exp\left(\frac{1}{4\pi}\left(\int_{S^2} |\nabla^\natural u|^2 + 8\pi u \, d\mu^\natural\right)\right) = C \exp \frac{1}{4\pi} \mathcal{I}^\natural(g_t)$$

and this bounds L^2-norm of e^u, by Corollary 2.4. This tells us nothing new; the L^2-norm is simply the volume, which is fixed. But with the assumption of symmetry, the constant $1/16\pi$ can be improved to $1/(32\pi - \epsilon)$ – see [Aubin 1982] again — and computing as above, we derive a bound for the $L^{2+\delta}$ norm of e^u, for sufficiently small $\delta > 0$. □

COROLLARY 2.6. *Under the fourth-order flow, the conformal factor remains bounded for finite times. Under the sixth-order flow, it remains even uniformly bounded.*

PROOF. Combine the previous lemma with Lemma 2.2. □

COROLLARY 2.7. *The constant in the Sobolev inequality remains bounded — for finite times in the fourth-order case, and uniformly for all time, in the sixth-order case.*

PROOF. The conformal factor controls the Sobolev constant. □

Curvature bounds and long-time existence

Let g_t be a symmetric metric on the 2-sphere evolving in accordance with one of our two flow equations. In this section, we shall derive pointwise estimates for the curvature of g_t and its derivatives, and thereby prove the solution may be continued for all times.

From the previous section, we know that the constant in the Sobolev inequality remains bounded in any finite interval of time, and in the sixth-order case, uniformly bounded.

THEOREM 2.8. *Let g_t be a symmetric solution to (7) defined on the 2-sphere. For any T, there exist constants $M_n(T)$ such that*

$$\sup_{0\le t\le T} ||\nabla^n R||_{L^2(d\mu_t)} \le M_n.$$

These depend only on T and the initial metric g_0.

PROOF. The proof hinges on the Sobolev-space interpolation estimates of Gagliardo and Nirenberg. In two space dimensions, these read:

LEMMA 2.9. *If the function f belongs to $W^{m,p} \cap L^q(M^2, g)$, and if $\int_{M^2} f \, d\mu = 0$, then, for $0 \le j < m$,*

$$\left|\left|\nabla^j f\right|\right|_{L^r} \le C \, ||\nabla^m f||^a_{L^p} \, ||f||^{1-a}_{L^q}, \tag{10}$$

whenever a, defined by the equation

$$\frac{1}{p} = \frac{j}{2} + a\left(\frac{1}{r} - \frac{m}{2}\right) + \frac{1-a}{q}$$

lies between j/m and 1.

PROOF. See [Aubin 1982]. The proof is an iteration combining the Hölder and Sobolev inequalities with the the estimate

$$||\nabla f||^2_{L^s(M^2)} \le (\sqrt{n} + |s-2|) \, ||f||_{L^w(M^2)} \left|\left|\nabla^2 f\right|\right|_{L^v(M^2)},$$

in which $2s^{-1} = w^{-1} + v^{-1}$, and the pointwise inequality

$$\left|\nabla \left|\nabla^l f\right|\right| \leq \left|\nabla^{l+1} f\right|.$$

Since the last two inequalities are independent of the metric g, and the constant in the Hölder inequality is simply a power of the volume, the constant C here depends only on the Sobolev quotient of g. So when it comes to an *evolving* metric g_t, this inequality continues to hold uniformly in t so long as the Sobolev constant can be held in check. □

PROOF OF THE $W^{n,2}$ ESTIMATES. The evolution of the $W^{n,2}$-seminorm of curvature is given by

$$\begin{aligned}\frac{d}{dt}\int_{S^2} |\nabla^n R|^2 \, d\mu &= \frac{d}{dt}\int_{S^2} g^{i_1 j_1} g^{i_2 j_2} \ldots g^{i_n j_n} \cdot \nabla^n_{i_1 i_2 \cdots i_n} R \nabla^n_{j_1 j_2 \cdots j_n} R \, d\mu \\ &= \int_{S^2} -2 \left\langle \nabla^n R,\ \nabla^n \Delta^2 R \right\rangle + \text{tensors of lesser order}\, d\mu.\end{aligned}$$

Now commuting derivatives, $\nabla^n \Delta^2 R = \Delta^2 \nabla^n R + \text{curvature terms} = \nabla_i \nabla_j \nabla_i \nabla_j \nabla^n R + \text{curvature}$, and so

$$\frac{d}{dt}\int_{S^2} |\nabla^n R|^2 \, d\mu = \int_{S^2} -2\left|\nabla^{n+2} R\right|^2 + \text{terms of lesser order}\ d\mu. \tag{11}$$

It is neither necessary nor desirable to compute the error terms explicitly. Instead, we simply split them into two groups: those (a) which derive from the principal part of the speed, $\Delta R + \frac{1}{2}R^2$; and those (b) which are the by-products of the volume correcting term $-\langle\omega\rangle g$. In other words, we divide the error terms by their scaling properties — those in the first group scale like the leading term in (11); those in the second scale slightly better.

Consider first the terms in class (a). Some of these contain the top derivative $\nabla^{n+2} R$ and some do not; but using Young's inequality, the 'do's' may be interpolated between sufficiently large multiples of the 'don'ts' and a fraction of the leading term. The error terms may therefore be written as a sum $\sum_j \alpha_j T_j$, in which each T_j has the form

$$T_j = \prod_{i \leq n+1} \left(\nabla^i R\right)^{a_{ji}},$$

where the product simply means some contraction of the various derivatives, and where $\sum_i (2+i) a_{ji} = 2n+8$: all these terms scale like the $-(2n+8)$th power of distance.

By the Hölder inequality,

$$\left|\int_{S^2} T_j \, d\mu\right| \leq \prod_{i \leq n+1} \left|\left|\nabla^i R\right|\right|^{a_{ji}}_{L^{\beta_i a_{ji}}},$$

for any choice of the β_is with $\sum \beta_i^{-1} = 1$. Now estimate each factor in the product using the interpolation inequalities:

$$\left|\int_{S^2} T_j \, d\mu\right| \leq C(S)\left(\left|\left|\nabla^{n+2} R\right|\right|^{\gamma_0 a_{j0}}_{L^2} ||R||^{(1-\gamma_0)a_{j0}}_{L^2} + 8\pi\right) \prod_{i=1}^{n+1} \left(\left|\left|\nabla^{n+2} R\right|\right|^{\gamma_i a_{ji}}_{L^2} ||R||^{(1-\gamma_i)a_{ji}}_{L^2}\right)$$

where the dependence of C on the constant S from the Sobolev inequality has been made explicit; and in this two-dimensional problem, γ_i is given by

$$\gamma_i = \frac{1}{n+2}\left(i + 1 - \frac{2}{\beta_i a_{ji}}\right).$$

Note that the term $i = 0$ was handled separately, because R does not have average value zero. The extra 8π in the first factor above arises when one centrifuges out the mean value with the triangle inequality.

It follows then from the monotonicity of $\int R^2\,d\mu$ that

$$\left|\int_{S^2} T_j\,d\mu\right| \leq C'(S)\left(\left|\left|\nabla^{n+2}R\right|\right|_{L^2}^{\sum_{i=0}^{n+1}\gamma_i a_{ji}} + \left|\left|\nabla^{n+2}R\right|\right|_{L^2}^{\sum_{i=1}^{n+1}\gamma_i a_{ji}}\right)$$

and the smaller power may be interpolated away with Young's inequality, at the cost of introducing a constant term:

$$\left|\int_{S^2} T_j\,d\mu\right| \leq C''(S)\left|\left|\nabla^{n+2}R\right|\right|_{L^2}^{\sum_{i=0}^{n+1}\gamma_i a_{ji}} + Q_j.$$

for some sufficiently large constant Q_j; it remains only to show that the combined exponent in the above equation is less than two. This follows from the scaling condition:

$$\sum_{i=0}^{n+1}\gamma_i a_{ji} = \frac{1}{n+2}\sum_{i=0}^{n+1}\left(i+1-\frac{2}{\beta_i a_{ji}}\right)a_{ji} = \frac{1}{n+2}\left(2n+6-\sum_{i=0}^{n+1}a_{ji}\right), \tag{12}$$

and the remaining sum is easily estimated from below:

$$\sum_{i=0}^{n+1}a_{ji} > \sum_{i=0}^{n+1}\frac{i+2}{n+2}\,a_{ji} = \frac{2n+8}{n+2} > 2;$$

but the a_{ji} are whole numbers; if the sum is larger than two, then it must be at least three; returning then to (4.4.2):

$$\sum_{i=0}^{n+1}\gamma_i a_{ji} \leq \frac{2n+3}{n+2} < 2.$$

So we may apply Young's inequality again, to deduce that

$$\left|\int_{S^2} T_j\,d\mu\right| \leq \epsilon_j\left|\left|\nabla^{n+2}R\right|\right|_{L^2}^2 + Q_j(S)$$

where Q_j is chosen sufficiently large. Thus, each T_j may be estimated by an arbitrarily small fraction of the leading term in (11) and a correspondingly large constant.

This handles the error terms in class (a); those in class (b) may be dealt with in the same way. The argument above changes only in the scaling condition — this becomes $\sum_i (i+2)a_{ij} = 2n+4$ — with the ensuing change in equation (12) above. So, revisiting (11),

$$\begin{aligned}\frac{d}{dt}\int_{S^2}\left|\nabla^n R\right|^2\,d\mu &\leq -\int_{S^2}\left|\nabla^{n+2}R\right|^2\,d\mu + Q(S)\\ &\leq -\delta(S)\int_{S^2}\left|\nabla^n R\right|^2\,d\mu + Q(S);\end{aligned}$$

and as long as the Sobolev constant remains bounded, this in turn bounds the $W^{n,2}$-seminorm of the curvature. □

COROLLARY 2.10. *There exist* $\tilde{M}_n(T)$ *such that* $\sup\left|\nabla^n R\right| \leq \tilde{M}_n$ *for* $0 \leq t \leq T$.

PROOF. Sobolev. □

The sixth-order flow behaves just as well. Fundamentally, throwing more laplacians onto the speed does no harm to this argument: the energy continues to fall monotonically; the Liouville property ensures that on a compact manifold, this does not give rise to any more critical points; and the higher derivatives serve only to strengthen the Sobolev inequality. Of course, the second monotonicity property for this particular flow is an added bonus.

THEOREM 2.11. *Let g_t be a symmetric solution to (8) defined on $S^2 \times [0,T)$. Then there are constants M_n, independent of T, such that*

$$\sup_{S^2 \times [0,T)} |\nabla^n R| \leq M_n.$$

These are determined completely by g_0.

PROOF. As above. The differences are:

$$\frac{d}{dt} \int_{S^2} |\nabla^n R|^2 \, d\mu = \int_{S^2} -2 \left|\nabla^{n+3} R\right|^2 + \text{terms of lesser order} \, d\mu$$

where the scaling of the error terms is like the $-(2n+10)^{\text{th}}$ power of distance. Writing the typical error term as $T_j = \prod_i \left(\nabla^i R\right)^{a_{ji}}$ with $\sum_i (2+i) a_{ji} = 2n+10$,

$$\left| \int_{S^2} T_j \, d\mu \right| \leq C \left|\left|\nabla^{n+2} R\right|\right|_{L^2}^{\sum_{i=0}^{n+2} \gamma_i a_{ji}} + Q_j,$$

and the total exponent here is

$$\sum_{i=0}^{n+2} \gamma_i a_{ji} \leq \frac{2n+5}{n+3} < 2,$$

which means T_j may be interpolated between a fraction of the leading term and a constant. As before, this leads to a bound for the $W^{n,2}$-seminorm of R; because the Sobolev constant is uniformly bounded, this estimate is independent of time. □

With that, we have the following existence theorems:

THEOREM 2.12. *Let g_0 be a smooth, symmetric metric on the 2-sphere. Then g_0 extends to a long-time solution to the equation*

$$\frac{\partial}{\partial t} g_{ij} = (\omega - \langle \omega \rangle) \cdot g_{ij}.$$

THEOREM 2.13. *With g_0 as above, one can find a long-time solution to the equation*

$$\frac{\partial}{\partial t} g_{ij} = -\Delta \omega \cdot g_{ij}$$

with uniformly bounded curvature; moreover, the metrics g_t are all uniformly equivalent and converge along subsequences to the round metric on S^2.

PROOF. This follows at once from the smoothness estimates above. In the sixth-order case, the metrics converge along subsequences, and in view of the equation for $\mathcal{R}$, we must have $\nabla \omega = 0$ for any of the limit metrics; these must therefore be solutions to the stationary equation. But by Theorem 1.2, the only stationary solutions are the metrics of constant curvature. □

COROLLARY 2.14. *Any metric g_0 on the projective plane can be extended to a long-time solution to either of the two descent flows above; in the sixth-order case, the metrics g_t converge along subsequences to the round metric on P^2.*

PROOF. g_0 lifts to a symmetric metric on the sphere which then extends to a solution to either descent flow; this solution remains symmetric, and can therefore be pulled back to P^2. □

The flow equations are meaningfully defined on P^2. There is a subtle point to note here, though it is not a terribly important one; the analysis was conducted by way of integral estimates; but to define integrals over P^2, one has to pass to the double-cover. However, the flow is defined solely in terms of local information; it really does make sense on P^2.

References

[Aubin 1982] Thierry Aubin, *Nonlinear Analysis on Manifolds. Monge-Ampère Equations,* Springer-Verlag, New York.

[Berger 1971] M. Berger, *On Riemannian Structures of prescribed Gaussian Curvature for Compact 2-Surfaces,* J. Diff. Geom. **5** 325-332.

[Cheeger 1970] J. Cheeger, *A lower bound for the smallest eigenvalue of the Laplacian* Problems in Analysis, *Symposium in Honor of S. Bochner,* 195-199. Princeton University Press, Princeton.

[Chruściel 1990] P. T. Chruściel, *On Robinson-Trautman Space-Times,* Centre for Mathematical Analysis Research Report CMA-R23-90, The Australian National University, Canberra.

[Polden 1995] A. Polden, *Closed Curves of Least Total Curvature,* Preprint Nr. 13, Sonderforschungsbereich 382, Universität Tübingen.

[Polden 1996] A. Polden, *Curves and Surfaces of Least Total Curvature and Fourth-Order Flows*, Doktorarbeit, Universität Tübingen.

[Singleton 1989] D. Singleton, Doctoral Thesis, Monash University, Melbourne, Australia.