A FORMAL RIEMANNIAN STRUCTURE ON CONFORMAL CLASSES AND UNIQUENESS FOR THE $\sigma_2$-YAMABE PROBLEM

MATTHEW GURSKY AND JEFFREY STREETS

Abstract. We define a new formal Riemannian metric on a conformal classes of four-manifolds in the context of the $\sigma_2$-Yamabe problem. Exploiting this new variational structure we show that solutions are unique unless the manifold is conformally equivalent to the round sphere.

1. Introduction

1.1. Background. In [20], we defined a formal Riemannian metric on the space of conformal metrics on surfaces of positive (or negative) Gauss curvature. Our goal in this paper is to show that one can extend this definition to conformal classes of metrics on four-manifolds, and to explore the geometric properties of this metric and their applications. The definition we give can be extended to higher (even) dimensions, but this will be pursued in a subsequent article since there are technical issues that do not arise in two or four dimensions [21].

In addition to verifying the formal properties of this metric we prove a remarkable geometric consequence: namely, solutions of the $\sigma_2$-Yamabe problem – whose existence follows from our positivity assumption and [8] – are unique, unless the manifold is conformally equivalent to the sphere. This is a surprising departure from the classical (or $\sigma_1$-)Yamabe problem, where explicit examples of non-uniqueness are known (see Remarks 1.6 and 1.7 below). Thus, positive conformal classes on four-manifolds have a unique conformal representative whose $\sigma_2$-curvature is constant; moreover the value of this constant (after normalizing the volume) can be expressed in terms of the Euler characteristic and the $L^2$-norm of the Weyl tensor (see the introduction of [7]). We also remark that this representative has positive Ricci curvature.

To give a more detailed description it will be helpful to return to the setting of surfaces. Let $(M, g_0)$ be a compact Riemannian surface with positive Gauss curvature $K_0 > 0$, and let $[g_0]$ denote the conformal class of $g_0$. Define

\[ C^+ = \{ g_u = e^{2u} g_0 \in [g_0] : K_u = K_{g_u} > 0 \}. \]

Formally, the tangent space to $[g_0]$ at any metric $g_u \in [g_0]$ is given by $C^\infty(M)$. For $\phi, \psi \in C^\infty(M) \cong T_u([g_0])$ we define

\[ \langle \phi, \psi \rangle_u = \int_M \phi \psi K_u dA_u, \]

where $K_u$ is the Gauss curvature and $dA_u$ is the area form of $g_u$.

The definition in (1.2) is inspired by the Mabuchi-Semmes-Donaldson [27, 32, 14] metric of Kähler geometry, wherein a formal Riemann metric is put on a Kähler class by imposing on the tangent space to a given Kähler potential the $L^2$ metric with respect to the associated Kähler metric. As observed in [27], this metric enjoys many nice formal properties, for instance nonpositive sectional curvature. Moreover, it has a profound relationship to natural functionals in Kähler geometry such as the Mabuchi $K$-energy and the Calabi energy, as well as their gradient flow, the Calabi flow.

In [20] we established a number of analogous properties for the metric defined by (1.2). For example, $C^+$ endowed with the metric in (1.2) has non-positive curvature in the sense of Alexandrov.
We also showed that the normalized Liouville energy \( F : W^{1,2} \to \mathbb{R} \), defined by
\[
F[u] = \int_M |\nabla u|^2 dA_0 + 2 \int_M K_0 u dA_0 - \left( \int_M K_0 dA_0 \right) \log \left( \int_M e^{2u} dA_0 \right),
\]
is \textit{geodesically convex}. Recall that critical points of \( F \), which are precisely the conformal metrics of constant Gauss curvature, are minimizers and unique up to Möbius transformation. Many of these global geometric properties are based on existence and partial regularity results for geodesics in \( C^* \) (see Section 4 of [20] for precise statements).

In this paper we study a natural generalization of the inner product (1.5). For an \( n \)-dimensional Riemannian manifold \((n \geq 3)\), we denote the Schouten tensor by
\[
A = \frac{1}{(n-2)} \left( Ric - \frac{1}{2(n-1)} R g \right),
\]
where \( Ric \) is the Ricci tensor and \( R \) is the scalar curvature. Let \( \sigma_k(g^{-1}A) \) denote the \( k^{th} \)-symmetric function of the eigenvalues of the (1,1) tensor obtained by raising an index of \( A \); i.e.,
\[
A^j_i = g^{jk} A_{ik}.
\]
The quantity \( \sigma_k(g^{-1}A) \) is called the \( \sigma_k \)-\textit{curvature} or the \( k \)-\textit{scalar curvature}. For example,
\[
\sigma_1(g^{-1}A) = \frac{R}{2(n-1)}.
\]
For \( 1 \leq k \leq n \), we write \( A = A_g \in \Gamma^+_k \) if \( \sigma_j(g^{-1}A) > 0 \) on \( M^n \) for all \( 1 \leq j \leq k \). By (1.4), we have \( A_g \in \Gamma^+_1 \) if \( g \) has positive scalar curvature, while \( A_g \in \Gamma^+_n \) if the Schouten tensor of \( g \) is positive definite.

We will be interested in the case where \( n = 4 \) and \( k = 2 \). To this end, let \((M^4, g_0)\) be a compact Riemannian four-manifold such that \( A_{g_0} \in \Gamma^+_2 \). Given \( u \in C^\infty(M) \), let \( A_u \) denote the Schouten tensor of the conformal metric \( g_u = e^{-2u} g_0 \). We will say that \( u \) is \textit{admissible} if \( A_u \in \Gamma^+_2 \). Let
\[
C^+ = C^+(g_0) = \{ g_u \in [g_0] \mid A_u \in \Gamma^+_2 \}.
\]
By a result of Guan-Viaclovsky, [18], if \( g_u \in C^+ \) then \( g_u \) has positive Ricci curvature. As noted above, the tangent space to \( C^+ \) at any point is given by \( C^\infty(M) \). Thus, in analogy with (1.5) we define for \( \phi, \psi \in C^\infty(M) \)
\[
\langle \phi, \psi \rangle_u = \int_M \phi \psi \sigma_2(g_u^{-1}A_u) dV_u.
\]

\textbf{Remark 1.1.} To simplify the notation we will write \( \sigma_2(A) \) instead of \( \sigma_2(g^{-1}A) \). Since we will be working with conformal metrics, we will also need to distinguish between \( g^{-1}A_u \) and \( g_u^{-1}A_u \); i.e., whether we are using \( g \) or \( g_u \) to raise an index. Therefore, we will adopt the usual convention that \( \sigma_2(A_u) = \sigma_2(g^{-1}A_u) \), but write \( \sigma_2(g_u^{-1}A_u) \) when we are using \( g_u \) to raise an index. Note that
\[
\sigma_2(g_u^{-1}A_u) = e^{4u} \sigma_2(A_u).
\]
In particular,
\[
\sigma_2(g_u^{-1}A_u) dV_u = \sigma_2(A_u) dV.
\]

\textbf{Remark 1.2.} There is a sharp characterization of conformal classes for which \( C^+ \) is non-empty. In view of the conformal invariance of the integral
\[
\sigma := \int \sigma_2(g^{-1}A_g) dV_g,
\]
a necessary condition for \([g]\) to admit a metric \( g_u \in [g] \) with \( A_u \in \Gamma^+_2 \) is the positivity of the Yamabe invariant and the positivity of \( \sigma \). In [7] these conditions were shown to be sufficient. Thus we have an exact parallel with the case of two dimensions, since a conformal class of metrics on a surface admits a metric of positive Gauss curvature if and only if the total Gauss curvature is positive.
1.2. Formal metric properties. We begin by establishing in §3 some fundamental formal properties of the metric defined in (1.5). We first introduce a formal path derivative which can be regarded as the Levi-Civita connection associated to the metric. Using this we compute the curvature tensor, and furthermore show that the curvature is nonpositive:

**Theorem 1.3.** Given $(M^4, g)$ a compact Riemannian manifold, with $A_g \in \Gamma^+_2$. Then (1.5) defines a metric with nonnegative sectional curvature on $\mathcal{C}^\ast$.

Next, we derive the geodesic equation. Formal calculations derived using either the path derivative or variations of the length functional yield that a one-parameter family of conformal factors is a geodesic if and only if

$$u_t - \frac{1}{\sigma_2(A_u)} (T_1(A_u), \nabla u_t \otimes \nabla u_t) = 0,$$

where $T_1$ is the Newton transform and $\langle \cdot, \cdot \rangle$ denotes the inner product on tensor bundles induced by $g$ (the background metric). This is a degenerate fully nonlinear equation, which is related to a $\sigma_2$-type problem for the spacetime Hessian of $u$, in direct analogy to the $(n+1)$-dimensional degenerate Monge-Ampere interpretation of the Mabuchi geodesic equation in Kähler geometry. We also show that one parameter families of conformal transformations are automatically geodesics (Proposition 3.12). This is again in analogy with the fact that one parameter families of biholomorphisms generate families of Kähler potentials which are Mabuchi geodesics.

In the Kähler setting, the Mabuchi metric and its geodesics are intimately related to Mabuchi’s $K$-energy functional. This is a “relative functional” defined via path integration of a closed 1-form on a Kähler class. It was shown in [26, 27] that this functional is geodesically convex, leading to the conjecture that extremal Kähler metrics are unique up to biholomorphism in a fixed Kähler class. Confirming this conjecture requires extensive existence and regularity results for the geodesic equation. An initial theory of $C^{1,1}$ was developed in [10, 6, 2], and eventually a more refined regularity theory was developed and the conjecture finally confirmed in [11].

In our setting there is a natural analogue of Mabuchi’s functional. For surfaces it is given by the Liouville energy, or regularized determinant (1.3). In four dimensions this functional was written down by Chang-Yang in [9] (although it appears implicitly in [7]):

$$F[u] = \int \left\{ 2\Delta u |\nabla u|^2 - |\nabla u|^4 - 2Ric(\nabla u, \nabla u) + R|\nabla u|^2 - 8u\sigma_2(A_g) \right\} dV$$

$$- 2 \left( \int \sigma_2(A_g) dV \right) \log \left( \int e^{-4u} dV \right).$$

After this, Brendle-Viaclovsky [5] give a path-integration derivation of this functional which makes clearer the analogy between it and the Mabuchi functional in Kähler geometry. We will not need the precise formula, only the fact that it provides a conformal primitive for $\sigma_2(A)$; i.e., if $u_s$ is a path with $\frac{d}{ds} u_s |_{s=0} = u'$, then

$$\frac{d}{ds} F[u_s] |_{s=0} = \int u'[ - \sigma_2(g_u^{-1} A_u) + \sigma_g ] dV_u.$$

Consequently, $u$ is a critical point of $F$ if and only if $g_u = e^{-2u} g$ is a solution of the $\sigma_2$-Yamabe problem:

$$\sigma_2(g_u^{-1} A_u) \equiv \text{const.}$$

In four dimensions the existence of solutions to (1.10) in conformal classes with $\mathcal{C}^\ast \neq \emptyset$ was first proved by Chang-Gursky-Yang [8] (for surveys on solving the $\sigma_k$-Yamabe problem for general $2 \leq k \leq n$ see [37] and [33]). In particular, if $\mathcal{C}^\ast ([g])$ is non-empty, then $[g]$ always admits a critical point of $F$. Our next result gives us deeper insight into the variational structure of $F$:
Theorem 1.4. The functional $F$ in (1.8) is geodesically convex.

The proof of this theorem requires the use of a sharp curvature-weighted Poincaré inequality due to Andrews [1]. In fact, it follows from Andrews’ inequality that $F$ is strictly convex, up to to one-parameter families of conformal automorphisms on the round sphere. This sharp characterization naturally leads one to conjecture that critical points of $F$ are unique, except in the case of the sphere. We are able to confirm this surprising fact:

Theorem 1.5. Let $(M^4, g)$ be a compact Riemannian manifold such that $C^\ast([g]) \neq \emptyset$.

1. If $(M^4, g)$ is not conformal to $(S^4, g_{S^4})$, then there exists a unique solution to the $\sigma_2$-Yamabe problem in $[g]$.
2. In $[g_{S^4}]$, all solutions to the $\sigma_2$-Yamabe problem are round metrics.

Remark 1.6. This uniqueness property is in stark contrast to the Yamabe problem, in which generic conformal classes admit arbitrarily many distinct solutions (see [29]). In dimensions $n \geq 25$ the solution space may even be non-compact [3, 4].

Remark 1.7. Explicit examples of non-uniqueness for the Yamabe problem were constructed by Schoen in [31], in which he constructed Delaunay-type solutions on $S^{n-1} \times S^1$. By lifting to the universal cover $S^{n-1} \times \mathbb{R}$ and imposing symmetry, he reduced the Yamabe equation to an ODE and studied the phase portrait. Interestingly, Viaclovsky [36] carried out a similar construction for solutions of the $\sigma_k$-Yamabe problem when $k < n/2$. However, once $k \geq n/2$ the construction fails, since the admissibility condition implies the Ricci curvature of any solution would have to be positive, and $S^{n-1} \times S^1$ does not admit a metric with positive Ricci curvature.

The proof of Theorem 1.5 consists of two main phases. First we develop a weak existence/regularity theory for the geodesic equation (1.7). In general for degenerate Monge-Ampere equations one typically expects at best $C^{1,1}$ control, and indeed this is verified in the Kähler setting by Chen (with complements due to Blocki) [10, 2]. Where Mabuchi geodesics can be interpreted as solutions of a degenerate complex Monge-Ampere equation, our geodesics are solutions to a degenerate $\sigma_2$-equation (Proposition 4.1), and so one at best again expects $C^{1,1}$ regularity. However, due to some technical issues arising from the presence of first order terms in the Schouten tensor, we are not able to establish such estimates. Rather we are forced to regularize the equation by rendering the right hand side positive (which is a standard trick), but also perturbing the coefficients on the time direction term, to further break the nondegeneracy. This leads to full $C^\infty$ regularity, but only the $C^{1}$-estimates persist as the regularization parameters go to zero.

Given this, one cannot directly rigorously establish properties of $F$ related to the geodesic convexity. Nonetheless we are able to improve the regularity of an approximate geodesic connecting any two solutions to the $\sigma_2$-problem by smoothing via the parabolic flow introduced by Guan-Wang [19]. In particular we are able to take a sequence of approximate geodesics connecting two critical points for $F$, smooth them for a short time with this flow, and then show that this process yields a path of critical points for $F$, although not necessarily a geodesic. Combining this with arguments using the geodesic convexity shows that the existence of this path implies that the critical points are all round metrics on $S^4$, finishing the proof.

1.3. Outline. In §2 we establish notation and record some basic properties of the Schouten tensor and of elementary symmetric polynomials. Next in §3 we establish the basic properties of the $\sigma_2$-metric defined in (1.5). In particular we prove Theorem 1.3 and establish the geodesic convexity of the $F$ functional. Then in §4 we develop estimates for approximate solutions to the geodesic equation, leading to a weak existence theory. In §5 we show a short-time smoothing result which we will use to improve the regularity of approximate geodesics connecting any two critical points of the $F$-functional. We combine these two main technical tools in §6 to establish Theorem 1.5.
2. Background

In this section we establish our notation and some basic formulas. Although we are primarily interested in four dimensions, we will state most of the standard results for symmetric functions we will need for general $n$ and $k$.

2.1. The Schouten tensor. Given a Riemannian manifold $(M^n, g)$ let $A$ denote the Schouten tensor of $g$. Given a conformal metric $g_u = e^{-2u} g$, the tensor $A$ transforms according to

\begin{equation}
A_u = A + \nabla^2 u + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g.
\end{equation}

Let $g_u = e^{-2u(t)} g$ be a 1-parameter family of conformal metrics. Then using formula (2.1) it follows that

\begin{equation}
\frac{\partial}{\partial t} (g_u^{-1} A_u)^{ij} = 2 \left( \frac{\partial u}{\partial t} \right) (g_u^{-1} A_u)^{ij} + \left( \nabla_u^2 \frac{\partial u}{\partial t} \right)_{ij},
\end{equation}

where the Hessian is with respect to $g_u$. A direct calculation \cite{30} yields

\begin{equation}
\frac{\partial}{\partial t} \sigma_k(g_u^{-1} A_u) = (T_{k-1}(g_u^{-1} A_u), \nabla_u^2 \frac{\partial u}{\partial t})_{g_u} + 2k \frac{\partial u}{\partial t} \sigma_k(g_u^{-1} A_u),
\end{equation}

where $T_{k-1}$ is the Newton transform. Since the Newton transform is a $(1,1)$-tensor, for the pairing in (2.3) we lower an index of $T_{k-1}(g_u^{-1} A_u)$ and view it as a $(0,2)$-tensor, and use the inner product induced by $g_u$. For example, if $n = 4$ and $k = 2$,

\begin{equation}
T_1(g_u A_u) = -A_u + \sigma_1(g_u^{-1} A_u) g_u.
\end{equation}

Combining (2.3) with the variation of the volume form yields

\begin{equation}
\frac{\partial}{\partial t} \left[ \sigma_k(g_u^{-1} A_u) dV_u \right] = (T_{k-1}(g_u^{-1} A_u), \nabla_u^2 \frac{\partial u}{\partial t})_{g_u} dV_u + (n - 2k) \frac{\partial u}{\partial t} \sigma_k(g_u^{-1} A_u) dV_u.
\end{equation}

A key property we will use throughout is the following:

**Lemma 2.1.** If $k = 2$ or if the manifold is locally conformally flat, then $T_{k-1}(g^{-1} A)$ is divergence-free.

**Remark 2.2.** This was proved in \cite{38}. The essential idea also appears in \cite{30}, where the Schouten tensor is replaced with the second fundamental form of a hypersurface of a space of constant curvature. In both cases one needs that the tensor is Codazzi; i.e.,

\[ \nabla_k A_{ij} = \nabla_j A_{ik}. \]

Note that the conformal invariance of the integral

\[ \sigma = \int_M \sigma_2(g_u^{-1} A_u) dV_u \]

follows from the variational formula (2.3) and Lemma 2.1. We denote the average value

\[ \bar{\sigma} = \sigma V_u^{-1}. \]

2.2. Properties of elementary symmetric polynomials. We record some lemmas concerning elementary symmetric polynomials and Newton transforms. To begin we record basic facts which are well-known from Garding’s theory of hyperbolic polynomials \cite{10}. We use these to derive some further properties of generalized Newton transforms required for our estimates of the geodesic equation. First, given $A \in \Gamma_k^n$ we let $\sigma_k(A)$ denote the $k$-th elementary polynomial in the eigenvalues of $A$. Moreover, given $A_1, \ldots, A_k$ we define the generalized Newton transformation by

\[ [T_k]_{ij} (A_1, \ldots, A_k) := \frac{1}{k!} \delta_{j_1, \ldots, j_k} (A_1)_{i_1 j_1} \ldots (A_k)_{i_k j_k}. \]
where here $\delta$ denotes the generalized Kronecker delta function. Moreover we set

$$\Sigma_k(A_1, \ldots, A_k) = \frac{1}{(k-1)!} f_{j_1, \ldots, j_k}^{i_1, \ldots, i_k} (A_1)_{i_1 j_1} \cdots (A_k)_{i_k j_k}.$$ 

**Lemma 2.3.** One has

1. Given $A_1, \ldots, A_k \in \Gamma_k^+$, then $[T_k]_{ij}(A_1, \ldots, A_k) > 0.$
2. Given $A_1, \ldots, A_k \in \Gamma_k^+$, then $\Sigma(A_1, \ldots, A_k) > 0.$
3. If $A - B \in \Gamma_k^+$ and $A_2, \ldots, A_k \in \Gamma_k^+$ then $\Sigma(B, A_2, \ldots, A_k) < \Sigma(A, A_2, \ldots, A_k)$.

**Lemma 2.4.** Given $A, B \in \Gamma_k^+$, $A < B$, one has $T_{k-1}(A) < T_{k-1}(B)$.

**Proof.** From Lemma 2.3 for $A_i \in \Gamma_k$ one has $T_k(A_1, \ldots, A_k) > 0$. Now consider $M_t = A + t(B - A)$. Since $B - A$ is positive definite certainly it lies in $\Gamma_k^+$. It follows that

$$\frac{d}{dt} T_k(M_t) = \frac{d}{dt} [T_k](M_t, \ldots, M_t)$$
$$= \sum_{j=1}^k [T_k](M_t, \ldots, B - A, \ldots, M_t)$$
$$\geq 0.$$ 

The result follows. \qed

**Lemma 2.5.** Given $A$ a symmetric matrix and $X$ a vector, one has for $k \geq 1$,

$$\langle T_k(A - X \otimes X), X \otimes X \rangle = \langle T_k(A), X \otimes X \rangle,$$

$$\sigma_k(A - X \otimes X) = \sigma_k(A) - \langle T_{k-1}(A), X \otimes X \rangle.$$ 

**Proof.** If we express the matrix $B_t = A - tX \otimes X$ in a basis where $X$ is the first basis vector, it is clear that the function

$$f(t) = \sigma_k(B_t)$$

is a linear function of $t$. It follows that its time derivative is constant, hence

$$C = f'(t) = -\langle T_{k-1}(A - tX \otimes X), X \otimes X \rangle.$$ 

Hence

$$\langle T_{k-1}(A), X \otimes X \rangle = -f'(0) = -f'(1) = \langle T_{k-1}(A - X \otimes X), X \otimes X \rangle.$$ 

Moreover, this shows that

$$\sigma_k(A - X \otimes X) = f(1) = f(0) + \int_0^1 f'(s)ds = \sigma_k(A) - \langle T_{k-1}(A), X \otimes X \rangle.$$ 

\qed

**Lemma 2.6.** Given $A, B \in \text{Sym}^2(\mathbb{R}^4)$, $A, B \in \Gamma_2^+$ one has

$$\langle T_1(B), A \rangle^2 \geq 4\sigma_2(A)\sigma_2(B).$$
Riemannian metric defined for $g$ we define the directional derivative along the path $u$ as required.

Proof. We compute that

$$\frac{\sigma_1(A)}{\sigma_1(B)} \langle T_1(B), A \rangle = - \frac{\sigma_1(A)}{\sigma_1(B)} \langle B, A \rangle + \sigma_1(A)^2$$

$$\geq - \frac{1}{2} \left( \frac{\sigma_1(A)}{\sigma_1(B)} \right)^2 |B|^2 - \frac{1}{2} |A|^2 + \left[ \frac{\sigma_1(A)}{\sigma_1(B)} \right]^2$$

$$= - \frac{1}{2} \sigma_1(A)^2 \left( \frac{|B|^2 - \sigma_1(B)^2 + \sigma_1(B)^2}{\sigma_1(B)^2} \right) + \sigma_2(A) + \frac{1}{2} \sigma_1(A)^2$$

$$= \frac{\sigma_1(A)^2}{\sigma_1(B)^2} \sigma_2(B) + \sigma_2(A).$$

Rearranging this and applying Cauchy-Schwarz yields

$$\sigma_2(A) \leq \frac{\sigma_1(A)}{\sigma_1(B)} \langle T_1(B), A \rangle - \frac{\sigma_1(A)^2}{\sigma_1(B)^2} \sigma_2(B)$$

$$\leq \frac{1}{4\sigma_2(B)} \langle T_1(B), A \rangle^2,$$

as required. \hfill \Box

3. The $\sigma_2$-metric

In this section we define the $\sigma_2$-metric and establish fundamental properties of this metric concerning connections, torsion, curvature and distance. We end by showing the crucial geodesic convexity property of the functional $F$ of Chang-Yang.

3.1. Metric, connection, and curvature. As in the Introduction, let

$$C^+ = C^+([g]) = \{g_u = e^{-2u}g : A_u \in \Gamma_2^+ \}.$$

Definition 3.1. Let $(M^4, g)$ be a compact Riemannian four-manifold. The $\sigma_k$-metric is the formal Riemannian metric defined for $g_u \in C^+([g]) = C^+, \alpha, \beta \in T_u\mathcal{C}^+ \simeq C^\infty(M)$ via

$$\langle \alpha, \beta \rangle_u = \frac{1}{\sigma} \int_M \alpha \beta \sigma_2(g_u^{-1}A_u) dV_u.$$

Moreover, given $u_t$ a path in $\mathcal{C}^+$ and $\alpha_t$ a one-parameter family of tangent vectors with $\alpha_t \in T_{u_t}\mathcal{C}^+$, we define the directional derivative along the path $u_t$ by

$$\frac{D}{dt} \alpha := \alpha_t - \sigma_2(g_u^{-1}A_u)^{-1} \langle T_1(g_u^{-1}A_u), \nabla \alpha \otimes \nabla u_t \rangle_{g_u}$$

$$= \alpha_{t-} - \sigma_2(A_u)^{-1} \langle T_1(A_u), \nabla \alpha \otimes \nabla u_t \rangle,$$

where we have used (1.6), and the convention that $T_1(g^{-1}A_u) = T_1(A_u)$.

Lemma 3.2. The connection defined by (3.1) is metric compatible and torsion free.

Proof. First we check metric compatibility. We compute, using (2.5) and Lemma 2.1

$$\frac{d}{dt} \langle \alpha_t, \beta_t \rangle_{u_t} = \frac{d}{dt} \int_M \alpha_t \beta \sigma_2(g_u^{-1}A_u) dV_u$$

$$= \langle \dot{\alpha}, \beta \rangle + \langle \alpha, \dot{\beta} \rangle + \int_M \alpha_t \beta \left( T_1(g_u^{-1}A_u), \nabla_u^2 \frac{\partial u}{\partial t} \right) dV_u$$

$$= \langle \dot{\alpha}, \beta \rangle + \langle \alpha, \dot{\beta} \rangle - \int_M \left( T_1(g_u^{-1}A_u), (\alpha \nabla \beta + \beta \nabla \alpha) \otimes \nabla_u \frac{\partial u}{\partial t} \right) dV_u$$

$$= \left( \frac{D}{dt} \alpha, \beta \right) + \left( \alpha, \frac{D}{dt} \beta \right).$$
Next, to compute the torsion, let $u_{s,t}$ be a two parameter family of conformal factors. Then
\[
\frac{D}{\partial s} \frac{\partial u}{\partial t} - \frac{D}{\partial t} \frac{\partial u}{\partial s} = \frac{\partial^2 u}{\partial s \partial t} - \sigma_2(g_u^{-1} A_u)^{-1} \left( T_1(g_u^{-1} A_u), \nabla \frac{\partial u}{\partial s} \otimes \nabla \frac{\partial u}{\partial t} \right)_{u} \\
- \frac{\partial^2 u}{\partial s \partial t} + \sigma_2(g_u^{-1} A_u)^{-1} \left( T_1(g_u^{-1} A_u), \nabla \frac{\partial u}{\partial t} \otimes \nabla \frac{\partial u}{\partial s} \right)_{u} \\
= 0.
\]

The lemma follows.  \hfill \square

Next we compute the sectional curvature, and conclude that it is non-positive. We first record an integral identity in Lemma 3.3 and a certain general quadratic inequality in Lemma 3.4. We then obtain the curvature inequality by exploiting these identities.

**Lemma 3.3.** If $\phi, \psi \in C^\infty(M)$, then
\[
\int \left\{ \nabla^2 \phi(\nabla \psi, \nabla \psi) - \Delta \phi |\nabla \psi|^2 - \nabla^2 \psi(\nabla \psi, \nabla \phi) + \Delta \psi(\nabla \psi, \nabla \phi) \right\} \phi dV \\
= \int \left\{ - |(\nabla \phi, \nabla \psi)|^2 + |\nabla \phi|^2 |\nabla \psi|^2 \right\} dV.
\]

**Proof.** Consider the vector field
\[
X_i = (\nabla \phi, \nabla \psi)^{\perp}_i \phi - |\nabla \psi|^2 \nabla_i \phi.
\]

Taking the divergence gives
\[
\delta X = \nabla_i X_i \\
= \nabla^2 \phi(\nabla \psi, \nabla \psi) + \nabla^2 \psi(\nabla \phi, \nabla \psi) + \Delta \psi(\nabla \phi, \nabla \psi) \\
- 2\nabla^2 \psi(\nabla \psi, \nabla \phi) - \Delta \phi |\nabla \psi|^2 \\
= \nabla^2 \phi(\nabla \psi, \nabla \psi) - \Delta \phi |\nabla \psi|^2 - \nabla^2 \psi(\nabla \psi, \nabla \phi) + \Delta \psi(\nabla \psi, \nabla \phi).
\]

Therefore,
\[
I = \int \left\{ \nabla^2 \phi(\nabla \psi, \nabla \psi) - \Delta \phi |\nabla \psi|^2 - \nabla^2 \psi(\nabla \psi, \nabla \phi) + \Delta \psi(\nabla \psi, \nabla \phi) \right\} \phi dV \\
= \int (\delta X) \phi dV.
\]

On the other hand, integrating by parts gives
\[
I = \int (\delta X) \phi dV \\
= - \int (X, \nabla \phi) dV \\
= \int \left\{ - |(\nabla \phi, \nabla \psi)|^2 + |\nabla \phi|^2 |\nabla \psi|^2 \right\} dV,
\]

as claimed.  \hfill \square

**Lemma 3.4.** Let $T_1 = T_1(A)$ denote the first Newton transformation of the symmetric linear map $A : V \to V$, where $V$ is a real inner product space of dimension four. Assume $A \in \Gamma^2_+$. Then for all $X, Y \in V$,
\[-T_1(X, X)T_1(Y, Y) + T_1(X, Y)^2 + \sigma_2(A)[|X|^2|Y|^2 - (X, Y)^2] \leq 0.\]

**Proof.** Choose an orthonormal basis for $V$ which diagonalizes $T_1$, and let $\{\lambda_1, \ldots, \lambda_4\}$ denote the eigenvalues of $T_1$. Note by our assumption on $A$ we know that $\lambda_i \geq 0$ for each $i$. With respect to
this orthonormal basis write \( X = (x_1, \ldots, x_4) \) and \( Y = (y_1, \ldots, y_4) \). Then expanding and collecting terms we get

\[
-T_1(X, X)T_1(Y, Y) + T_1(X, Y)^2 = -\left( \lambda_1 x_1^2 + \cdots + \lambda_4 x_4^2 \big\{ \lambda_1 y_1^2 + \cdots + \lambda_4 y_4^2 \big\} + \left\{ \lambda_1 x_1 y_1 + \cdots \lambda_4 x_4 y_4 \right\}^2 \\
= -\lambda_1 \lambda_2 (x_1^2 y_2^2 + x_2^2 y_1^2 - 2x_1 x_2 y_1 y_2) - \lambda_1 \lambda_3 (x_3^2 y_3^2 + x_4^2 y_4^2 - 2x_3 x_4 y_3 y_4) \\
- \cdots - \lambda_3 \lambda_4 (x_3^2 y_4^2 + x_4^2 y_3^2 - 2x_3 x_4 y_3 y_4).
\]

Next, let \( Z = X \wedge Y \), whose components are

\[
z_{ij} = x_i y_j - x_j y_i.
\]

In terms of \( Z \), we can rewrite the above as

\[
-T_1(X, X)T_1(Y, Y) + T_1(X, Y)^2 = -\lambda_1 \lambda_2 z_{12}^2 - \lambda_1 \lambda_3 z_{13}^2 - \cdots - \lambda_3 \lambda_4 z_{34}^2.
\]

At the same time,

\[
|X|^2|Y|^2 - (X, Y)^2 = \frac{1}{2} |Z|^2 = z_{12}^2 + z_{13}^2 + \cdots + z_{34}^2.
\]

Therefore,

\[
-T_1(X, X)T_1(Y, Y) + T_1(X, Y)^2 + \sigma_2(A)\left[ |X|^2|Y|^2 - (X, Y)^2 \right] = -\lambda_1 \lambda_2 z_{12}^2 - \lambda_1 \lambda_3 z_{13}^2 - \cdots - \lambda_3 \lambda_4 z_{34}^2 + \sigma_2(A)\left[ z_{12}^2 + z_{13}^2 + \cdots + z_{34}^2 \right].
\]

We need to express \( \sigma_2(A) \) in terms of the eigenvalues of \( T_1 \). Since

\[
T_1 = -A + \sigma_1(A) \cdot I,
\]

taking the trace it follows that

\[
\lambda_1 + \cdots + \lambda_4 = 3\sigma_1(A).
\]

Also, taking the norm-squared in (3.3),

\[
|T_1|^2 = |A|^2 + 2\sigma_1(A)^2.
\]

Therefore,

\[
\sigma_2(A) = \frac{1}{3} ( - \lambda_1^2 - \cdots - \lambda_4^2 + \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \cdots \lambda_3 \lambda_4 ).
\]

Substituting this into (3.2),

\[
-T_1(X, X)T_1(Y, Y) + T_1(X, Y)^2 + \sigma_2(A)\left[ |X|^2|Y|^2 - (X, Y)^2 \right] = -\lambda_1 \lambda_2 z_{12}^2 - \lambda_1 \lambda_3 z_{13}^2 - \cdots - \lambda_3 \lambda_4 z_{34}^2 \\
+ \frac{1}{3} \left( - \lambda_1^2 - \cdots - \lambda_4^2 + \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \cdots \lambda_3 \lambda_4 \right) \left[ z_{12}^2 + z_{13}^2 + \cdots + z_{34}^2 \right] \\
= \frac{1}{3} \left( - \lambda_1^2 - \cdots - \lambda_4^2 - 2\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \cdots \lambda_3 \lambda_4 \right) z_{12}^2 \\
+ \frac{1}{3} \left( - \lambda_1^2 - \cdots - \lambda_4^2 + \lambda_1 \lambda_2 - 2\lambda_1 \lambda_3 + \lambda_1 \lambda_4 + \cdots \lambda_3 \lambda_4 \right) z_{13}^2 \\
+ \cdots + \frac{1}{3} \left( - \lambda_1^2 - \cdots - \lambda_4^2 + \lambda_1 \lambda_2 + \cdots + \lambda_2 \lambda_4 - 2\lambda_3 \lambda_4 \right) z_{34}^2.
\]
We claim that the coefficients of the $z_j^2$-terms are all non-positive. To see this, consider the first one:

$$-\lambda_2^2 - \cdots - \lambda_3^2 - 2\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4$$

$$= -(\lambda_1 + \lambda_2)^2 - \lambda_3^2 - \lambda_4^2 + (\lambda_1 + \lambda_2)\lambda_3 + (\lambda_1 + \lambda_2)\lambda_4 + \lambda_3\lambda_4$$

(3.5)

$$\leq -(\lambda_1 + \lambda_2)^2 - \lambda_3^2 - \lambda_4^2 + \frac{1}{2}(\lambda_1 + \lambda_2)^2 + \frac{1}{2}\lambda_3^2 + \frac{1}{2}(\lambda_1 + \lambda_2)^2 + \frac{1}{2}\lambda_4^2$$

$$+ \frac{1}{2}\lambda_3^2 + \frac{1}{2}\lambda_4^2$$

$$= 0.$$

□

Finally we prove the required curvature inequality, which is a more precise statement of Theorem 1.3.

**Theorem 3.5.** Let $(M^4, g)$ be a compact Riemannian manifold such that $A_g \in \Gamma_2^+$. Given $u \in \Gamma_2^+$ and $\phi, \psi \in T_u\Gamma_2^+$ we have

$$K(\phi, \psi) = \int \frac{1}{\sigma_2(g^{-1}_u A_u)} \left\{ - \left( T_1(g^{-1}_u A_u), \nabla\phi \otimes \nabla\phi \right) T_1(g^{-1}_u A_u), \nabla\psi \otimes \nabla\psi \right\}$$

$$+ \left( T_1(g^{-1}_u A_u), \nabla\phi \otimes \nabla\psi \right)^2 + \sigma_2(g^{-1}_u A_u) |\nabla\phi|^2 |\nabla\psi|^2 - \sigma_2(g^{-1}_u A_u) |(\nabla\phi, \nabla\psi)|^2 \right\} dV_u$$

$$\leq 0,$$

where the inner products are with respect to $g_u$

**Proof.** Let $u(s, t)$ be a 2-parameter family of conformal factors, and $\alpha = \alpha(s, t) \in T_u(s, t)\mathcal{C}^+$. Using the formula for the directional derivative in (3.1), we have

$$\frac{D}{ds} \frac{D}{dt} \alpha = \frac{\partial}{\partial s} \left( \frac{D}{dt} \alpha \right) - \frac{1}{\sigma_2(g^{-1}_u A_u)} \left( T_1(g^{-1}_u A_u), \nabla \left( \frac{D}{dt} \alpha \right) \otimes \nabla \left( \frac{\partial}{\partial s} \right) \right)$$

$$= \frac{\partial}{\partial s} \left\{ \frac{\partial \alpha}{\partial t} - \frac{1}{\sigma_2(g^{-1}_u A_u)} \left( T_1(g^{-1}_u A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial}{\partial s} \right) \right) \right\}$$

$$- \frac{1}{\sigma_2(g^{-1}_u A_u)} \left( T_1(g^{-1}_u A_u), \nabla \left( \frac{D}{dt} \alpha \right) \otimes \nabla \left( \frac{\partial}{\partial s} \right) \right)$$

(3.6)

$$= \frac{\partial^2 \alpha}{\partial s \partial t} + \frac{1}{\sigma_2(g^{-1}_u A_u)} \left( T_1(g^{-1}_u A_u), \nabla^2 \left( \frac{\partial}{\partial s} \right) \right) \left( T_1(g^{-1}_u A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial}{\partial s} \right) \right)$$

$$- \frac{1}{\sigma_2(g^{-1}_u A_u)} \left( \frac{\partial}{\partial s} \right) \left( T_1(g^{-1}_u A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial}{\partial t} \right) \right)$$

$$- \frac{1}{\sigma_2(g^{-1}_u A_u)} \left( T_1(g^{-1}_u A_u), \nabla \left( \frac{\partial}{\partial s} \alpha \right) \otimes \nabla \left( \frac{\partial}{\partial s} \right) \right) + \nabla \alpha \otimes \nabla \left( \frac{\partial^2 u}{\partial s \partial t} \right)$$

$$- \frac{1}{\sigma_2(g^{-1}_u A_u)} \left( T_1(g^{-1}_u A_u), \nabla \left( \frac{D}{dt} \alpha \right) \otimes \nabla \left( \frac{\partial}{\partial s} \right) \right).$$

In the above, we have used the fact that the inner product on symmetric 2-tensors satisfies

$$\frac{\partial}{\partial s} \langle \cdot, \cdot \rangle_u = 4 \frac{\partial u}{\partial s} \langle \cdot, \cdot \rangle_u.$$
For the last term in (3.6),
\[- \frac{1}{\sigma_2(g_u^{-1}A_u)} \{ T_1(g_u^{-1}A_u), \nabla \left( \frac{D}{Dt}(\alpha) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right) \}_u = \]
\[- \frac{1}{\sigma_2(g_u^{-1}A_u)} \{ T_1(g_u^{-1}A_u), \nabla \left\{ \frac{\partial \alpha}{\partial t} - \frac{1}{\sigma_2(g_u^{-1}A_u)} \{ T_1(g_u^{-1}A_u), \nabla \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \}_u \right\} \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \]
\[- \frac{1}{\sigma_2(g_u^{-1}A_u)} \{ T_1(g_u^{-1}A_u), \nabla \left( \frac{\partial \alpha}{\partial t} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \]
\[+ \frac{1}{\sigma_2(g_u^{-1}A_u)} \{ T_1(g_u^{-1}A_u), \nabla \left\{ \frac{1}{\sigma_2(g_u^{-1}A_u)} \{ T_1(g_u^{-1}A_u), \nabla \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \}_u \right\} \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \].

By (2.4) and (2.2),
\[\frac{\partial}{\partial s} T_1(g_u^{-1}A_u) = \frac{\partial}{\partial s} \left\{ - A_u + \sigma_1(g_u^{-1}A_u) g_u \right\} = - \nabla_u^2 \left( \frac{\partial u}{\partial s} \right) + \Delta_u \left( \frac{\partial u}{\partial s} \right) g_u.
\]

Substituting this into (3.6), we get
\[
\frac{D}{Ds} \frac{D}{Dt} \alpha = \frac{\partial^2 \alpha}{\partial s \partial t} + \frac{1}{\sigma_2(g_u^{-1}A_u)} \left\{ \frac{1}{\sigma_2(g_u^{-1}A_u)} \{ T_1(g_u^{-1}A_u), \nabla^2 \left( \frac{\partial u}{\partial s} \right) \}_u \{ T_1(g_u^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \}_u \right\} \\
- \{ T_1(g_u^{-1}A_u), \nabla \left( \frac{\partial \alpha}{\partial t} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \left\{ T_1(g_u^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \}_u \right\} \\
- \{ T_1(g_u^{-1}A_u), \nabla \left( \frac{\partial \alpha}{\partial t} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \left\{ T_1(g_u^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \}_u \right\} \\
+ \{ T_1(g_u^{-1}A_u), \nabla \left( \frac{\partial \alpha}{\partial t} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \left\{ T_1(g_u^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \}_u \right\} \right\}.
\]

Next, we rearrange the terms into two groups: those symmetric in $s, t$, and those that are not:
\[
\frac{D}{Ds} \frac{D}{Dt} \alpha = \frac{\partial^2 \alpha}{\partial s \partial t} + \frac{1}{\sigma_2(g_u^{-1}A_u)} \left\{ - \{ T_1(g_u^{-1}A_u), \nabla \left( \frac{\partial \alpha}{\partial t} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u - \{ T_1(g_u^{-1}A_u), \nabla \left( \frac{\partial \alpha}{\partial t} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \right\} \\
+ \{ T_1(g_u^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial^2 u}{\partial s \partial t} \right) \}_u \} \\
+ \frac{1}{\sigma_2(g_u^{-1}A_u)} \left\{ \frac{1}{\sigma_2(g_u^{-1}A_u)} \{ T_1(g_u^{-1}A_u), \nabla^2 \left( \frac{\partial u}{\partial s} \right) \}_u \{ T_1(g_u^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \}_u \right\} \\
+ \{ \nabla_u^2 \left( \frac{\partial u}{\partial s} \right) - \Delta_u \left( \frac{\partial u}{\partial s} \right) g_u \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \\
+ \{ T_1(g_u^{-1}A_u), \nabla \left( \frac{\partial \alpha}{\partial t} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \}_u \left\{ T_1(g_u^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \}_u \right\} \right\}.
\]
Therefore,

(3.7) 

\[
\left( \frac{D}{\partial s} \frac{D}{\partial t} - \frac{D}{\partial t} \frac{D}{\partial s} \right) \alpha = \frac{1}{\sigma_2(g_+^{-1}A_u)} \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla^2 \left( \frac{\partial u}{\partial t} \right))_u \right) T_1(g_+^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right)_u \\
- \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla^2 \left( \frac{\partial u}{\partial t} \right))_u \right) T_1(g_+^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial s} \right)_u \\
+ \left( \nabla_u^2 \left( \frac{\partial u}{\partial s} \right) - \Delta_u \left( \frac{\partial u}{\partial s} \right) g_u, \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \right) \right)_u - \left( \nabla_u^2 \left( \frac{\partial u}{\partial s} \right) - \Delta_u \left( \frac{\partial u}{\partial t} \right) g_u, \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \right) \right)_u \\
+ \left( T_1(g_+^{-1}A_u), \nabla \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial s} \right)) \right)_u \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \\
- \left( T_1(g_+^{-1}A_u), \nabla \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \alpha \otimes \nabla \left( \frac{\partial u}{\partial s} \right)) \right)_u \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right). 
\]

To compute the sectional curvature of the plane spanned by \( \left( \frac{\partial u}{\partial s}, \frac{\partial u}{\partial t} \right) \), we take \( \alpha = \frac{\partial u}{\partial s} \) in the formula above, then take the inner product with \( \frac{\partial u}{\partial s} \):

\[
\left( \left( \frac{D}{\partial s} \frac{D}{\partial t} - \frac{D}{\partial t} \frac{D}{\partial s} \right) \frac{\partial u}{\partial s}, \frac{\partial u}{\partial s} \right)_u = \\
\int \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla^2 \left( \frac{\partial u}{\partial s} \right))_u \right) T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \\
- \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla^2 \left( \frac{\partial u}{\partial s} \right))_u \right) T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \\
+ \left( \nabla_u^2 \left( \frac{\partial u}{\partial s} \right) - \Delta_u \left( \frac{\partial u}{\partial s} \right) g_u, \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u - \left( \nabla_u^2 \left( \frac{\partial u}{\partial s} \right) - \Delta_u \left( \frac{\partial u}{\partial s} \right) g_u, \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \\
+ \left( T_1(g_+^{-1}A_u), \nabla \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right) \right)_u \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \\
- \left( T_1(g_+^{-1}A_u), \nabla \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right) \right)_u \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right) dV_u. 
\]

Consider the last two lines above. Integrating by parts and using the fact that \( T_1(g_+^{-1}A_u) \) is divergence-free, we get

\[
\int \left( \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right) \right)_u \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u dV_u \\
- \left( T_1(g_+^{-1}A_u), \nabla \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right) \right)_u \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u dV_u \\
= \int \left( \left( \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right) T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \\
- \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right) T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \\
+ \frac{1}{\sigma_2(g_+^{-1}A_u)} (T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u \right) T_1(g_+^{-1}A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial s} \right) \right)_u dV_u. 
\]
Substituting this into $\text{(3.7)}$ we find that the the first two lines there cancel, and we arrive at
\[
\left(\frac{D}{\partial s} \frac{D}{\partial t} - \frac{D}{\partial t} \frac{D}{\partial s}\right) \frac{\partial u}{\partial t} u = \\
\int \left\{ \left( \nabla^2_u (\frac{\partial u}{\partial t}) - \Delta_u (\frac{\partial u}{\partial t}) \right) g_u, \nabla (\frac{\partial u}{\partial t}) \otimes \nabla (\frac{\partial u}{\partial t}) \right\}_u \\
- \left\{ \nabla^2_u (\frac{\partial u}{\partial t}) - \Delta_u (\frac{\partial u}{\partial s}) g_u, \nabla (\frac{\partial u}{\partial t}) \otimes \nabla (\frac{\partial u}{\partial s}) \right\}_u \frac{\partial u}{\partial s} dV_u \\
+ \int \sigma_2(g_u^{-1} A_u) \left\{ - \langle T_1(g_u^{-1} A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \rangle_u \langle T_1(g_u^{-1} A_u), \nabla \left( \frac{\partial u}{\partial t} \right) \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \rangle_u \\
+ \langle T_1(g_u^{-1} A_u), \nabla \left( \frac{\partial u}{\partial s} \right) \otimes \nabla \left( \frac{\partial u}{\partial t} \right) \rangle_u^2 \\
+ \sigma_2(g_u^{-1} A_u) \left| \nabla \frac{\partial u}{\partial s} \right|_u^2 \left| \nabla \frac{\partial u}{\partial t} \right|_u^2 - \sigma_2(g_u^{-1} A_u) \langle \nabla \frac{\partial u}{\partial s}, \nabla \frac{\partial u}{\partial t} \rangle_u \right\} dV_u \\
\leq 0,
\]
as required. \qed

**Remark 3.6.** The Mabuchi metric turns out to be formally an infinite dimensional symmetric space, evidenced by the sectional curvatures admitting an interpretation as the square norm of the Poisson bracket of the two tangent vector functions. There does not seem to be such an interpretation in this setting.

### 3.2. Formal metric space structure.

In this subsection we observe some fundamental properties of lengths of curves and distances in the $\sigma_2$-metric.

**Definition 3.7.** Given a path $u : [a, b] \to \mathcal{C}^+$, the length of $u$ is
\[
\mathcal{L}(u) := \int_a^b \sqrt{\alpha^2 + \beta^2} \, dt = \int_a^b \left( \int_M \left( \frac{\partial u}{\partial t} \right)^2 \sigma_2(g_u^{-1} A_u) dV_u \right)^{\frac{1}{2}} dt.
\]

A curve is a geodesic if it is a critical point for $L$.

**Lemma 3.8.** A curve $u_t \in \mathcal{C}^+$ is a geodesic if and only if
\[
u_t - \frac{1}{\sigma_2(A_u)} \langle T_1(A_u), \nabla u_t \otimes \nabla u_t \rangle = 0.
\]

**Proof.** Formally, by Lemma 3.2 the connection is indeed the Riemannian connection and so a curve is a geodesic if and only if
\[
0 = \frac{D}{\partial t} \frac{D}{\partial t} u_t = u_t - \frac{1}{\sigma_2(A_u)} \langle T_1(A_u), \nabla u_t \otimes \nabla u_t \rangle.
\]

This can also be derived by directly taking the first variation of the length functional. \qed

**Remark 3.9.** We observe a canonical isometric splitting of $T_u \mathcal{C}^+$ with respect to the $\sigma_k$ metric. In particular, the real line $\mathbb{R} \subset T_u \mathcal{C}^+$ given by constant functions is orthogonal to
\[
T_u^0 \mathcal{C}^+ := \left\{ \alpha \mid \int_M \alpha \sigma_2(g_u^{-1} A_u) dV_u = 0 \right\}.
\]
In the next lemma we show two basic properties of geodesics, namely that they preserve this isometric splitting, and are automatically parameterized with constant speed.

**Lemma 3.10.** Let \( u_t \) be a solution to (3.3). Then
\[
\frac{d}{dt} \int_M u_t \sigma_2 (g_u^{-1} A_u) dV_u = 0,
\]
\[
\frac{d}{dt} \int_M u_t^2 \sigma_2 (g_u^{-1} A_u) dV_u = 0.
\]

**Proof.** Differentiating and using (2.5),
\[
\frac{d}{dt} \int_M u_t \sigma_2 (g_u^{-1} A_u) dV_u = \int_M \left( u_{tt} \sigma_2 (g_u^{-1} A_u) + u_t \left( T_1 (g_u^{-1} A_u), \nabla^2 u_t \right)_u \right) dV_u
\]
\[
= \int_M \left( u_{tt} - \sigma_2 (g_u^{-1} A_u)^{-1} \left( T_1 (g_u^{-1} A_u), \nabla u_t \otimes \nabla u_t \right)_u \right) \sigma_2 (g_u^{-1} A_u) dV_u
\]
\[
= 0.
\]
Next
\[
\frac{d}{dt} \int_M u_t^2 \sigma_2 (g_u^{-1} A_u) dV_u = \int_M \left[ 2 \sigma_2 (g_u^{-1} A_u) u_{tt} u_t + u_t^2 \left( T_1 (g_u^{-1} A_u), \nabla^2 u_t \right)_u \right] dV_u
\]
\[
= 2 \int_M \sigma_2 (g_u^{-1} A_u) u_t \left[ u_{tt} - \frac{1}{\sigma_2 (g_u^{-1} A_u)} \left( T_1 (g_u^{-1} A_u), \nabla u_t \otimes \nabla u_t \right)_u \right] dV_u
\]
\[
= 0.
\]

**Proposition 3.11.** Given \( u_0, u_1 \in C^\infty(M) \) and \( u_t : [0, 1] \to C^+ \) a geodesic, one has
\[
\mathcal{L}(u) \geq \sigma^{-\frac{1}{2}} \max \left\{ \int_{u_1 > u_0} (u_1 - u_0) \sigma_2 (g_u^{-1} A_u) dV_u, \int_{u_0 > u_1} (u_0 - u_1) \sigma_2 (g_{u_0}^{-1} A_{u_0}) dV_{u_0} \right\}.
\]

**Proof.** Observe that the geodesic equation implies \( u_{tt} \geq 0 \), and so we obtain the pointwise inequality
\[
u_t (0) \leq u_1 - u_0 \leq u_t (1).
\]
Thus using Hölder’s inequality we have
\[
E(1) = \left( \int_M u_t^2 \sigma_2 (g_u^{-1} A_u) dV_u \right)^{\frac{1}{2}}
\]
\[
\geq \sigma^{-\frac{1}{2}} \int_M |u_t| \sigma_2 (g_u^{-1} A_u) dV_u
\]
\[
\geq \sigma^{-\frac{1}{2}} \int_{u_1 > u_0} (u_1 - u_0) \sigma_2 (g_u^{-1} A_u) dV_u.
\]
A similar argument yields
\[
E(0) \geq \sigma^{-\frac{1}{2}} \int_{u_0 > u_1} (u_0 - u_1) \sigma_2 (g_{u_0}^{-1} A_{u_0}) dV_{u_0}.
\]
Since geodesics are automatically constant speed by Lemma 3.10, the result follows.

### 3.3 Geodesics and the conformal group of the sphere

As in the two-dimensional case, we will show that the 1-parameter family of transformations that generate the conformal group of the sphere are geodesics. In anticipation of our forthcoming article on the higher-dimensional case we will prove a more general result.

Let \( (S^n, g_0) \) denote the round sphere. Using stereographic projection \( \sigma : S^n \setminus \{ N \} \to \mathbb{R}^n \), where \( N \in S^n \) denotes the north pole, one can define a one-parameter of conformal maps of \( S^n \) by conjugating the dilation map \( \delta_\alpha : x \mapsto \alpha^{-1} x \) on \( \mathbb{R}^n \) with \( \sigma \):
\[
\varphi_\alpha = \sigma^{-1} \circ \delta_\alpha \circ \sigma : S^n \to S^n.
\]
UNIQUENESS FOR THE $\sigma_2$-YAMABE PROBLEM

Taking $\alpha(t) = e^{\lambda t}$, where $\lambda$ is a fixed real number, we can define the path of conformal metrics

$$g(t) = e^{-2u(t)} g_0 = \phi_0^* g_0 = \left[ \frac{2\alpha(t)}{(1 + \xi) + \alpha(t)^2(1 - \xi)} \right]^2,$$

where $\xi = x^{n+1}$ is the $(n + 1)$-coordinate function; i.e., $N = (0, \ldots, 0, 1)$ (see [24]).

**Proposition 3.12.** If $k = n/2$, the path $g(t) = e^{-2u(t)} g_0 : (-\infty, +\infty) \to C^*$ satisfies

$$u_{tt} - \frac{1}{\sigma_k(A_u)} \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle = 0.$$

In particular, when $n = 4$ this path defines a geodesic.

**Proof.** By (3.9),

$$u = u(t) = -\log 2\alpha + \log [(1 + \xi) + \alpha^2(1 - \xi)].$$

This yields

$$u_t = -\frac{\dot{\alpha}}{\alpha} + \frac{2\alpha \dot{\alpha} (1 - \xi)}{(1 + \xi) + \alpha^2(1 - \xi)},$$

and hence

$$u_{tt} = -\frac{\alpha_{tt}}{\alpha} + \left( \frac{\alpha_{t}}{\alpha} \right)^2 + \frac{(1 + \xi) + \alpha^2(1 - \xi)}{(1 + \xi) + \alpha^2(1 - \xi)} \left[ 2\alpha \alpha_{tt} + 2\alpha^2(1 - \xi) - 4\alpha^2 \alpha_t^2 (1 - \xi)^2 \right].$$

Since $\alpha(t) = e^{\lambda t}$, we have

$$u_{tt} = 4\lambda^2 e^{2\lambda t} \frac{1 - \xi^2}{[(1 + \xi) + \alpha^2(1 - \xi)]^2}.$$  \hspace{1cm} (3.11)

Also,

$$\nabla u_t = -\frac{2\alpha \alpha_t \nabla \xi}{(1 + \xi) + \alpha^2(1 - \xi)} - \frac{2\alpha \alpha_t (1 - \xi)}{[(1 + \xi) + \alpha^2(1 - \xi)]^2} \left[ (1 - \alpha^2) \nabla \xi \right]$$

$$= \frac{-2\alpha \alpha_t \nabla \xi}{[(1 + \xi) + \alpha^2(1 - \xi)]^2} \left[ (1 + \xi) + \alpha^2 (1 - \xi) + (1 - \xi)(1 - \alpha^2) \right]$$

$$= \frac{-4\alpha \alpha_t \nabla \xi}{[(1 + \xi) + \alpha^2(1 - \xi)]^2}.$$  \hspace{1cm} (3.12)

On $S^n$, the Schouten tensor is a multiple of the identity; in fact $A(g_0) = \frac{1}{2} g_0$. Therefore, using standard identities for the symmetric functions,

$$\frac{1}{\sigma_k(g(t)^{-1} A_{g(t)})} T_1(g(t)^{-1} A_{g(t)}) = \frac{2k}{n} g(t) = g(t),$$

since $k = n/2$. Thus

$$\frac{1}{\sigma_k(g(t)^{-1} A_{g(t)})} (T_{k-1}(g(t)^{-1} A_{g(t)}), \nabla u_t \otimes \nabla u_t) = 4\lambda^2 e^{2\lambda t} \frac{\left| \nabla \xi \right|^2}{[(1 + \xi) + \alpha^2(1 - \xi)]^2}.$$  \hspace{1cm} (3.12)

Since $|\nabla \xi|^2 = 1 - \xi^2$, comparing (3.11) and (3.12) we see that $u$ satisfies (3.10).  \hspace{1cm} $\square$
Remark 3.13. We do not expect conformal vector fields on general backgrounds to generate nontrivial geodesics, and thus nonuniqueness of solutions. It follows from a result of Lelong-Ferrand/Obata [23, 28] that if \((M^n, g)\) is not conformally equivalent to the round sphere, then any conformal Killing field is a Killing field for a conformally related metric. Expressed with respect to this background metric, pullback by a family of isometries will result in no change on the level of conformal factors.

3.4. The \(F\)-functional and geodesic convexity. We now derive the geodesic convexity of the \(F\)-functional of Chang-Yang. The crucial input is a sharp curvature-weighted Poincaré inequality due to Andrews:

**Proposition 3.14.** (Andrews [1], cf. [12] pg. 517) Let \((M^n, g)\) be a closed Riemannian manifold with positive Ricci curvature. Given \(\phi \in C^\infty(M)\) such that \(\int_M \phi dV = 0\), then

\[
\frac{n}{n-1} \int_M \phi^2 dV \leq \int_M (\text{Ric}^{-1})^{ij} \nabla_i \phi \nabla_j \phi dV,
\]

with equality if and only if \(\phi \equiv 0\) or \((M^n, g)\) is isometric to the round sphere.

The convexity of \(F\) will follow from a weaker form of this inequality:

**Corollary 3.15.** Let \((M^4, g)\) be a closed Riemannian manifold such that \(A_g \in \Gamma^+_2\). Given \(\phi \in C^\infty(M)\) such that \(\int_M \phi dV = 0\), then

\[
\int_M \frac{1}{\sigma_2(A_g)} T_1(A_g)^{ij} \nabla_i \phi \nabla_j \phi dV_g \geq 4 \int_M \phi^2 dV_g - \left( \frac{4}{\int_M dV_g} \right) \left( \int_M \phi dV_g \right)^2,
\]

with equality if and only if \(\phi \equiv 0\) or \((M^4, g)\) is isometric to the round sphere.

**Proof.** We assume \(\int_M \phi dV_g = 0\). By Andrews’ Poincaré inequality we have

\[
\frac{4}{3} \int_M \phi^2 dV_g \leq \int_M (\text{Ric}^{-1})^{ij} \nabla_i \phi \nabla_j \phi dV_g.
\]

To show the claim it suffices to show that

\[
3 \text{Ric}^{-1}(X, X) \leq \frac{1}{\sigma_2(A)} T_1(X, X).
\]

Since \(\text{Ric}\) and \(T_1(A)\) commute, it suffices to show that \(\text{Ric} \circ T_1 \geq 3 \sigma_2(A) g\). Since \(\text{Ric} = 2A + \sigma_1(A) g\), this is equivalent to

\[
-2A \circ A + \sigma_1(A) A + \sigma_1(A)^2 g \geq 3 \sigma_2(A) g.
\]

Now let \(Z = A - \frac{1}{4} \sigma_1(A) g\), then we can rewrite this as

\[
-2Z^2 + \frac{9}{8} \sigma_1(A)^2 g \geq 3 \sigma_2(A) g.
\]

Now, a Lagrange multiplier argument shows that

\[
Z \circ Z \leq \frac{3}{4} |Z|^2 g.
\]

Thus

\[
-2Z^2 + \frac{9}{8} \sigma_1(A)^2 g \geq -\frac{3}{2} |Z|^2 g + \frac{9}{8} \sigma_1(A)^2 g = 3 \sigma_2(A) g.
\]

\[ \square \]

**Proposition 3.16.** The functional \(F\) is geodesically convex.

Proof. It follows from [9] that for a path of conformal metrics \( u = u(t) \),
\[
\frac{d}{dt} F[u] = \int_M u_t \left[ -\sigma_2(g_u^{-1} A_u) + \sigma \right] dV_u.
\]
(3.13)

Assuming the path is a geodesic, then differentiating again and using Lemma [3.10] we have
\[
\frac{d^2}{dt^2} F[u] = \frac{d}{dt} \int_M u_t \left[ -\sigma_2(g_u^{-1} A_u) + \sigma \right] dV_u
\]
\[
= \sigma \frac{d}{dt} \int_M u_t V_u^{-1} dV_u
\]
\[
= \sigma \int_M \left[ u_t V_u^{-1} + V_u^{-2} u_t \left( \int_M 4 u_t dV_u \right) - 4 V_u^{-1} u_t^2 \right] dV_u
\]
\[
= \sigma V_u^{-1} \left[ \int_M \frac{1}{\sigma_2(g_u^{-1} A_u)} (T_1(g_u^{-1} A_u), \nabla u_t \otimes \nabla u_t) dV_u \right.
\]
\[
- 4 \left( \int_M u_t^2 dV_u - V_u^{-1} \left( \int_M u_t dV_u \right)^2 \right) \]
\[
\geq 0,
\]
where the last line follows from Corollary [3.15]. \( \square \)

4. Estimates of the Geodesic Equation

In this section we establish several fundamental properties of the geodesic equation (3.8). Once again, for future reference we will consider a more general equation which reduces to (3.8) when \( n = 4 \) and \( k = 2 \):
\[
u_{tt} = \frac{1}{\sigma_k(A_u)} (T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t).
\]

To begin, we define a certain regularization of this equation. In particular let
\[
\Phi(u) := u_{tt} \sigma_k(A_u) - \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle.
\]

Furthermore, let
\[
\Phi_\epsilon(u) = (1 + \epsilon) u_{tt} \sigma_k(A_u) - \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle.
\]

We will fix two parameters \( \epsilon, s \), and study a priori estimates for
\[
\Phi_\epsilon(u(\cdot, s)) = sf.
\]

To obtain estimates though we will simply fix a function \( f \in C^\infty(M \times [0, 1]) \) and study the equation
\[
G_\epsilon^f(u) = \Phi_\epsilon(u) - f = 0. \quad (\star_{\epsilon,f}).
\]

As remarked on above, in the setting of Mabuchi geodesics, as observed by Semmes [32] if one complexifies the time direction the equation admits an interpretation as a certain modification of the tensor \( A \) will show up naturally in the linearized operator. Let
\[
E = E^\epsilon_u = (1 + \epsilon) u_{tt} A_u - \nabla u_t \otimes \nabla u_t.
\]

Proposition 4.1. \( u \in C^2 \) satisfies \((\star_{\epsilon,f})\) if and only if
\[
[(1 + \epsilon) u_{tt}]^{1-k} \sigma_k(E^\epsilon_u) = f.
\]
Proof. Using Lemma 2.5 and homogeneity properties of elementary symmetric polynomials we compute

\[
\sigma_k(E_u^e) = \sigma_k((1 + \epsilon)u_{tt}A_u - \nabla u_t \otimes \nabla u_t) \\
= \sigma_k((1 + \epsilon)u_{tt}A_u) - \langle T_{k-1}((1 + \epsilon)u_{tt}A_u), \nabla u_t \otimes \nabla u_t \rangle \\
= \left[(1 + \epsilon)u_{tt}\right]^{k-1} \left[(1 + \epsilon)u_{tt}\sigma_k(A_u) - \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle\right].
\]

The proposition follows. \qed

We will say that a solution \( u \) of \((\ast_{\epsilon,f})\) is admissible if \( E_u^e \in \Gamma_k^r \). As we will see below, \((\ast_{\epsilon,f})\) is elliptic for admissible solutions.

**Lemma 4.2.** Let \( u = u(s, \cdot) \in C^\infty(M \times [0, 1]) \) be a one-parameter family of smooth functions such that \( \frac{d}{ds}u(s, \cdot)\big|_{s=0} = v \). Then

\[
\frac{d}{ds}u_{tt}^{1-k}\sigma_k(E_u)\bigg|_{s=0} = \mathcal{L}(v),
\]

where

\[
\mathcal{L}(v) = (1 + \epsilon)^{k-1}u_{tt}^{1-k}f v_{tt} \\
+ u_{tt}^{1-k}\left(T_{k-1}(E_u^e), (1 + \epsilon)u_{tt}\left(\nabla^2 v + \nabla v \otimes \nabla u + \nabla u \otimes \nabla v - \langle \nabla v, \nabla u \rangle g\right) \\
- \nabla v_t \otimes \nabla u_t - \nabla u_t \otimes \nabla v_t + u_{tt}^{1-k}v_{tt}\nabla u_t \otimes \nabla u_t \right).
\]

Proof. We compute

\[
\frac{d}{ds}u_{tt}^{1-k}\sigma_k(E_u) \\
= (1 - k)u_{tt}^{1-k}\sigma_k(E_u) v_{tt} + u_{tt}^{1-k}\left(T_{k-1}(E_u^e), \frac{d}{ds}E_u^e\right) \\
= (1 - k)u_{tt}^{1-k}\sigma_k(E_u) v_{tt} \\
+ u_{tt}^{1-k}\left(T_{k-1}(E_u^e), (1 + \epsilon)v_{tt}A_u + (1 + \epsilon)u_{tt}\frac{d}{ds}A_u - \nabla v_t \otimes \nabla u_t - \nabla u_t \otimes \nabla v_t \right).
\]

The second term can be simplified using Lemma 2.5 to

\[
(1 + \epsilon)u_{tt}^{1-k}\left(T_{k-1}(E_u^e), v_{tt}A_u\right) \\
= v_{tt}(1 + \epsilon)u_{tt}^{1-k}\left[1 + \epsilon^{-1}\left(T_{k-1}(E_u^e), E_u^e + \nabla u_t \otimes \nabla u_t\right)\right] \\
= v_{tt}u_{tt}^{1-k}[k\sigma_k(E_u^e) + \langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle] \\
= k\nabla v_{tt}u_{tt}^{1-k}\sigma_k(E_u^e) + v_{tt}u_{tt}^{1-k}(1 + \epsilon)^{k-1}\langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle \\
= v_{tt}u_{tt}^{1-k}\sigma_k(E_u^e) + v_{tt}\left[(1 + \epsilon)^k\sigma_k(A_u) - f(1 + \epsilon)^{k-1}u_{tt}^{1-k}\right] \\
= v_{tt}\left[(1 + \epsilon)^k\sigma_k(A_u) + (1 + \epsilon)^k\sigma_k(A_u)\right].
\]

Hence the overall term involving \( v_{tt} \) in (4.3) is \( v_{tt}(1 + \epsilon)^k\sigma_k(A_u) \). However we can furthermore express, again using the geodesic equation and Lemma 2.5, that

\[
(1 + \epsilon)^k\sigma_k(A_u) = (1 + \epsilon)^{k-1}u_{tt}^{1-k}f + (1 + \epsilon)^{k-1}u_{tt}^{1-k}\langle T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t \rangle \\
= (1 + \epsilon)^{k-1}u_{tt}^{1-k}f + u_{tt}^{1-k}\langle T_{k-1}(E_u^e), \nabla u_t \otimes \nabla u_t \rangle.
\]

Likewise we simplify the third term of (4.3) as

\[
(1 + \epsilon)^{k-1}\left(T_{k-1}(E_u^e), (1 + \epsilon)v_{tt}\left(\nabla^2 v + \nabla v \otimes \nabla u + \nabla u \otimes \nabla v - \langle \nabla v, \nabla u \rangle g\right) \right). 
\]

Collecting these calculations yields the result. \qed
Lemma 4.3. Given \( f \geq 0 \), equation \((\ast_{\epsilon,f})\) for admissible \( u \) is strictly elliptic for \( \epsilon > 0 \), and weakly elliptic for \( \epsilon = 0 \).

Proof. We compute the principal symbol of \( L \). We will ignore the first term of \((4.2)\), which has weakly positive symbol. Now fix a vector \( V = (\lambda, X) \in T[0,1] \times TM \). It follows from \((4.2)\) that the principal symbol of \( L \) acts via

\[
L(V, V) = u_{tt}^{1-k} \left( T_{k-1}(E'_{\epsilon}), (1 + \epsilon)u_{tt}X \otimes \nabla u_t \otimes (\lambda X) - (\lambda X) \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t(\lambda^2) \right).
\]

It follows from the Cachy-Schwarz inequality that for any \( \rho > 0 \), as an inequality of matrices one has

\[
-\lambda X \otimes \nabla u_t - \lambda \nabla u_t \otimes \lambda X \leq \rho X \otimes X + \rho^{-1} \lambda^2 \nabla u_t \otimes \nabla u_t
\]

Applying this inequality with \( \rho = (1 + \frac{\epsilon}{2})u_{tt} \) yields

\[
(1 + \epsilon)u_{tt}X \otimes X - \nabla u_t \otimes (\lambda X) - (\lambda X) \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t(\lambda^2) \geq \frac{\epsilon}{2} u_{tt}X \otimes X + \frac{\epsilon}{2} u_{tt}^{-1} \lambda^2.
\]

Since \( u \) is admissible, we have \( T_{k-1}(E'_{\epsilon}) > 0 \), and the result follows. \( \square \)

4.1. \( C^0 \) estimate. To prove a \( C^0 \)-estimate we begin with two technical lemmas:

Lemma 4.4. Suppose \( \phi = \phi(t) \). Then

\[
L\phi = \phi_{tt}(1 + \epsilon)^k \sigma_k(A_u).
\]

Proof. We directly compute using \((4.2)\), Lemma \((2.5)\) and the geodesic equation that

\[
L\phi = \phi_{tt} \left\{ (1 + \epsilon)^{-k-1} u_{tt}^{-1} f + u_{tt}^{-1} \left( T_{k-1}(E'_{\epsilon}), u_{tt}^{-1} \nabla u_t \otimes \nabla u_t \right) \right\}
\]

\[
= \phi_{tt} \left\{ (1 + \epsilon)^{-k-1} u_{tt}^{-1} f + u_{tt}^{-1} \left( T_{k-1}(E'_{\epsilon}), u_{tt}^{-1} \nabla u_t \otimes \nabla u_t \right) \right\}
\]

\[
= \phi_{tt} (1 + \epsilon)^{-k-1} \left\{ u_{tt}^{-1} f + u_{tt}^{-1} \left( T_{k-1}(E'_{\epsilon}), u_{tt} \otimes \nabla u_t \right) \right\}
\]

\[
= \phi_{tt} (1 + \epsilon)^k \sigma_k(A_u).
\]

\( \square \)

Lemma 4.5. Let \( u \) be an admissible solution to \((\ast_{\epsilon,f})\). Then

\[
Lu = (k + 1)(1 + \epsilon)^{-k-1} f + (1 + \epsilon)u_{tt}^{-k-1} \left( T_{k-1}(E'_{\epsilon}), -A + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \right).
\]

Proof. To begin we directly compute using \((4.2)\) that

\[
Lu = (1 + \epsilon)^{-k-1} f + u_{tt}^{-1-k} \left( T_{k-1}(E'_{\epsilon}), (1 + \epsilon)u_{tt} \left( \nabla^2 u + 2 \nabla u \otimes \nabla u - |\nabla u|^2 g \right) - \nabla u_t \otimes \nabla u_t \right).
\]

For the second term we simplify

\[
(1 + \epsilon)u_{tt}^{-k-1} \left( T_{k-1}(E'_{\epsilon}), \nabla^2 u \right)
\]

\[
= (1 + \epsilon)u_{tt}^{-k-1} \left( T_{k-1}(E'_{\epsilon}), A_u - A \nabla u \otimes \nabla u + \frac{1}{2} |\nabla u|^2 g \right)
\]

\[
= u_{tt}^{-k} \left( T_{k-1}(E'_{\epsilon}), u_{tt}^{-1} [E_u + \nabla u_t \otimes \nabla u_t] \right)
\]

\[
+ (1 + \epsilon)u_{tt}^{-k} \left( T_{k-1}(E'_{\epsilon}), -A - \nabla u \otimes \nabla u + \frac{1}{2} |\nabla u|^2 g \right)
\]

\[
= ku_{tt}^{-k} \left( T_{k-1}(E'_{\epsilon}), \nabla u_t \otimes \nabla u_t \right)
\]

\[
+ (1 + \epsilon)u_{tt}^{-k} \left( T_{k-1}(E'_{\epsilon}), -A - \nabla u \otimes \nabla u + \frac{1}{2} |\nabla u|^2 g \right)
\]

\[
= k(1 + \epsilon)^{-k-1} f + u_{tt}^{-1-k} \left( T_{k-1}(E'_{\epsilon}), \nabla u_t \otimes \nabla u_t \right)
\]

\[
+ (1 + \epsilon)u_{tt}^{-k} \left( T_{k-1}(E'_{\epsilon}), -A - \nabla u \otimes \nabla u + \frac{1}{2} |\nabla u|^2 g \right).
\]
Combining these calculations yields the result.

**Proposition 4.6.** Let \( u \) be an admissible solution to \((\star_{\epsilon,f})\). Then
\[
\sup_{M \times [0,1]} |u| \leq C(u|_{M \times [0,1], \max f}).
\]

**Proof.** We first observe that an admissible solution to \((\star_{\epsilon,f})\) satisfies \( u_{tt} \geq 0 \), and hence by convexity one has \( \sup_{M \times [0,1]} u \leq \sup_{M \times [0,1]} u \). To obtain the lower bound, fix a constant \( \Lambda \) and let
\[
\Psi = u + \Lambda t(1 - t).
\]
Observe that at an interior spacetime minimum of \( \Psi \) one has
\[
\nabla u, \quad \nabla^2 u > 0.
\]
Using this and Lemma 4.5 yields, at such a spacetime minimum,
\[
\mathcal{L}\Psi = (k + 1)(1 + \epsilon)^{k-1} f - (1 + \epsilon)u_{tt}^{2-k} (T_{k-1}(E_u), A) - 2\Lambda \left[(1 + \epsilon)^k u_{tt}^{1-k} (T_{k-1}(E_u), u_{tt}^{1} \nabla u_{tt} \otimes \nabla u_{tt})\right].
\]
Next we claim
\[
\Psi_{tt} \nabla^2 \Psi - \nabla \Psi_t \otimes \nabla \Psi_t \geq 0.
\]
Since we are at a minimum for \( \Psi \), \( \Psi_{tt} \nabla^2 \Psi \) is a positive semidefinite matrix. The expression above is thus the difference between a positive semidefinite matrix and a negative definite rank 1 matrix. The lemma follows if we establish positivity in the nondegenerate direction of the rank 1 matrix we subtracted, i.e. \( \nabla \Psi_t \). In particular it then suffices to show
\[
\Psi_{tt} \nabla^k \nabla \Psi_t \nabla \Psi_t - |\nabla \Psi_t|^4 \geq 0.
\]
To establish this we use that \( \Psi \) is actually a spacetime minimum. This implies that the spacetime Hessian is positive semidefinite. Testing this condition against the vector \(-\sqrt{\Psi_{tt} \nabla \Psi_t \otimes \frac{\nabla \Psi_t}{\Psi^{tt}}} \frac{\partial}{\partial t}\) yields
\[
0 \leq \Psi_{tt} \nabla^k \nabla \Psi_t \nabla \Psi_t - 2 |\nabla \Psi_t|^4 + |\nabla \Psi_t|^4,
\]
as required. However, using the explicit form of \( \Psi \) we see that this implies
\[
(u_{tt} - \Lambda) \nabla^2 u - \nabla u_t \otimes \nabla u_t \geq 0,
\]
which since \( \nabla^2 u > 0 \) implies
\[
u_{tt} \nabla^2 u - \nabla u_t \otimes \nabla u_t \geq 0.
\]
Hence \( E_u \geq u_{tt} A \), and then we obtain using Lemma 2.3 that
\[
u_{tt}^{2-k} (T_{k-1}(E_u), A) = u_{tt}^{2-k} \Sigma(E_u, \ldots, E_u, A) \geq u_{tt}^{2-k} \Sigma(u_{tt} A, \ldots, u_{tt} A, A) = u_{tt} \sigma_k(A) \geq 0.
\]
We can also simplify
\[
u_{tt}^{1-k} (T_{k-1}(E_u), u_{tt}^{1} \nabla u_t \otimes \nabla u_t) = (1 + \epsilon)^{k-1} u_{tt}^{1} (T_{k-1}(A_u), \nabla u_t \otimes \nabla u_t) = (1 + \epsilon)^{k-1} u_{tt}^{-1} f + (1 + \epsilon)^{k-1} \sigma_k(A_u)\]
Combining these observations yields, at the interior minimum,
\[ \mathcal{L}\Psi \leq (k + 1)(1 + \varepsilon)^{k-1}f - 2\Lambda(1 + \varepsilon)^{k-1}\sigma_k(A_u) \]
\[ \leq (k + 1)(1 + \varepsilon)^{k-1}f - 2\Lambda(1 + \varepsilon)^{k-1}\sigma_k(A) \]
\[ \leq C f - 2\delta \Lambda, \]
for some constants \(C\) and \(\delta\) depending only on the background data and maximum of \(f\). Choosing \(\Lambda\) sufficiently large with respect to these constants yields \(\mathcal{L}\Psi < 0\). Hence \(\Psi\) cannot have an interior minimum, and the result follows. \(\square\)

**Remark 4.7.** In the following estimates, all bounds on solutions be understood to depend on
\[ \max_M \left\{ f + \frac{|f_t|}{f} + \frac{\nabla |f|}{f} + \frac{|f_{tt}|}{f} + \frac{\nabla^2 |f|}{f} \right\}, \]
but this dependence will be suppressed to simplify the exposition.

### 4.2. \(C^1\) estimates.

**Proposition 4.8.** Given \(u\) an admissible solution to \((\ast_{e,f})\), one has
\[ \sup_{M \times [0,1]} |u_t| \leq C. \]

**Proof.** First we observe that, since \(u_t \geq 0\), it follows that there is a constant such that \(u_t(0) \leq C\) by direct integration. Now fix constants \(\Lambda_1, \Lambda_2\) and consider
\[ \Phi(x,t) = u(x,t) - u(x,0) - \Lambda_1 t^2 + \Lambda_2 t, \]
where \(\Lambda_1\) is chosen large below, and \(\Lambda_2\) is chosen still larger so that \(\Phi(x,1) \geq 0\). First note using \(4.2\) that
\[ \mathcal{L}u_0 = u_{tt}^{1-k} \left( T_{k-1}(E_u^t), (1 + \varepsilon)u_{tt} \left( \nabla^2 u_0 + \nabla u_0 \otimes \nabla u + \nabla u \otimes \nabla u_0 - \langle \nabla u_0, \nabla u \rangle g \right) \right). \]

Combining this with Lemmas 4.4 and 4.5 we obtain
\[ \mathcal{L}\Phi = \mathcal{L}u - \mathcal{L}u_0 - \Lambda_1\mathcal{L}^2 \]
\[ = (1 + \varepsilon)u_{tt}^{2-k} \left( T_{k-1}(E_u), -A - \nabla^2 u_0 + \nabla u_0 \otimes \nabla u - 2\nabla u \otimes \nabla u_0 - \frac{1}{2} |\nabla u|^2 g + \langle \nabla u_0, \nabla u \rangle g \right) \]
\[ (k + 1)(1 + \varepsilon)^{k-1}f - 2\Lambda_1(1 + \varepsilon)^{k-1}u_{tt}^{1-k} f - 2\Lambda_1 u_{tt}^{1-k} \left( T_{k-1}(E_u^t), \nabla u_t \otimes \nabla u_t \right). \]

Also we have \(\nabla u = \nabla u_0\) at the minimum, so we can simplify to
\[ \mathcal{L}\Phi = -u_{tt}^{2-k} \left( T_{k-1}(E), A + \nabla^2 u_0 + \nabla u_0 \otimes \nabla u_0 - \frac{1}{2} |\nabla u_0|^2 g \right) \]
\[ + (k + 1)(1 + \varepsilon)^{k-1}f - 2\Lambda_1(1 + \varepsilon)^k\sigma_k(A_u) \]
\[ = -u_{tt}^{2-k} \left( T_{k-1}(E), A_{u_0} \right) + (k + 1)(1 + \varepsilon)^{k-1}f - 2\Lambda_1(1 + \varepsilon)^k\sigma_k(A_u). \]

At a spacetime minimum for \(\Phi\) we have \(\nabla^2 (u - u_0) \geq 0\), and hence
\[ 0 \leq \Phi_{tt} \nabla^2 \Phi - \nabla \Phi_t \otimes \nabla \Phi_t \]
\[ = (u_{tt} - 2\Lambda_1)\nabla^2 (u - u_0) - \nabla u_t \otimes \nabla u_t \]
\[ \leq u_{tt} \nabla^2 (u - u_0) - \nabla u_t \otimes \nabla u_t. \]
Using this yields

\[ E_u = [(1 + \epsilon)u_t A_u - \nabla u_t \otimes \nabla u_t] \]

\[ = [(1 + \epsilon)u_t (A + \nabla^2 u + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g) - \nabla u_t \otimes \nabla u_t] \]

\[ \geq [(1 + \epsilon)u_t (A + \nabla^2 u_0 + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u_0|^2 g)] \]

\[ = [(1 + \epsilon)u_t (A + \nabla^2 u_0 + \nabla u_0 \otimes \nabla u_0 - \frac{1}{2} |\nabla u_0|^2 g)] \].

It follows from Lemma 4.9 that

\[ \langle T_{k-1}(E), A \rangle \geq 0. \]

A similar calculation shows that at the minimum point under consideration we have

\[ \sigma_k(A_u) \geq \sigma_k(A_{u_0}). \]

Putting these estimates together yields

\[ \mathcal{L} \Phi \leq (k + 1)(1 + \epsilon)^{k-1} f - 2 \Lambda_1(1 + \epsilon)^k \sigma_k(A_{u_0}). \]

If we choose \( \Lambda_1 \) sufficiently large with respect to the positive lower bound for \( \sigma_k(A_{u_0}) \) and the maximum of \( f \) we obtain \( \mathcal{L} \Phi < 0 \), and hence \( \Phi \) cannot have an interior minimum. Thus it follows that \( \Phi_t(x, 0) \geq 0 \) for all \( x \), and thus the lower bound for \( u_t(0) \) follows. A very similar estimate yields a two sided bound for \( u_t(1) \). Since \( u_{tt} \geq 0 \) everywhere we have a two sided bound for \( u_t \) everywhere.

We next proceed to obtain the interior spatial gradient estimate. To do this we need two preliminary calculations.

**Lemma 4.9.** Let \( u \) be an admissible solution to \( (*_{\epsilon, f}) \). Then

\[ \mathcal{L} e^{-\lambda u} \geq - \lambda e^{-\lambda u} \mathcal{L} u + \frac{1}{2} \lambda^2 e^{-\lambda u} u_{tt}^{-k} \langle T_{k-1}(E_u^0), \nabla u \otimes \nabla u \rangle - C \lambda^2 e^{-\lambda u} \sigma_k(A_u) u_t^2. \]

**Proof.** To begin we directly compute using (4.2) that

\[ \mathcal{L} e^{-\lambda u} = (1 + \epsilon)^{k-1} u_{tt}^{-1} f(e^{-\lambda u})_{tt} \]

\[ + u_{tt}^{1-k} \langle T_{k-1}(E_u^0), (1 + \epsilon)u_{tt} (\nabla^2 e^{-\lambda u} + \nabla e^{-\lambda u} \otimes \nabla u + \nabla u \otimes \nabla e^{-\lambda u} - (\nabla e^{-\lambda u}, \nabla u)) \]

\[ - \nabla (e^{-\lambda u})_t \otimes \nabla u_t - u_{tt}^{1-k} u_t \otimes \nabla u_t (e^{-\lambda u})_{tt} \]

\[ = - \lambda e^{-\lambda u} \mathcal{L} u + (1 + \epsilon)^{k-1} u_{tt}^{-1} f \lambda^2 e^{-\lambda u} u_t^2 \]

\[ + \lambda^2 e^{-\lambda u} u_{tt}^{1-k} \langle T_{k-1}(E_u^0), (1 + \epsilon)u_{tt} \nabla u \otimes \nabla u - u_t \nabla u \otimes \nabla u - u_t \nabla u \otimes \nabla u - u_{tt}^{1-k} u_t^2 \nabla u_t \otimes \nabla u_t \rangle. \]

Next we observe using the Cauchy-Schwarz inequality and equation \( (*_{\epsilon, f}) \) that

\[ u_{tt}^{1-k} \langle T_{k-1}(E_u^0), (1 + \epsilon)u_{tt} \nabla u \otimes \nabla u - u_t \nabla u \otimes \nabla u - u_t \nabla u \otimes \nabla u - u_{tt}^{1-k} u_t^2 \nabla u_t \otimes \nabla u_t \rangle \]

\[ = \sigma_k(A_u) u_t^2 - 2 u_{tt}^{1-k} \langle T_{k-1}(E_u^0), \nabla u_t \otimes \nabla u \rangle + u_{tt}^{2-k} \langle T_{k-1}(E_u^0), \nabla u \otimes \nabla u \rangle \]

\[ \geq - C \sigma_k(A_u) u_t^2 + \frac{1}{2} u_{tt}^{2-k} \langle T_{k-1}(E_u^0), \nabla u \otimes \nabla u \rangle. \]

Combining these calculations yields the result. \qed

**Lemma 4.10.** Given \( u \) an admissible solution to \( (*_{\epsilon, f}) \), one has

\[ \mathcal{L} u_t^2 = 2 u_t f_t + (1 + \epsilon)^{k-1} f u_{tt} + 2 \epsilon u_{tt}^{2-k} T_{k-1}(E) \mathcal{J} \nabla_j u_t \nabla_k u_t. \]
Proof. It follows directly from the definition of $\mathcal{L}$ that $\mathcal{L}u_t = f_t$. It follows that
\[ \mathcal{L}u_t^2 = 2u_t \mathcal{L}u_t + 2(1 + \epsilon)^{k-1} f u_t \]
\[ + 2u_t^{1-k} T_{k-1}(E)^{jk} \left\{ (1 + \epsilon) u_t \nabla_j u_t \nabla_k u_t - 2 \nabla_j u_t \nabla_k u_t u_t + u_t^{1-j} \nabla_j u_t \nabla_k u_t |\nabla u_t|^2 \right\} \]
\[ = 2u_t f_t + 2(1 + \epsilon)^{1-k} f u_t + 2u_t^{2-k} T_{k-1}(E)^{jk} \nabla_j u_t \nabla_k u_t, \]
as required. \qed

Lemma 4.11. Given $u$ an admissible solution to $(\ast_{\epsilon,f})$, one has
\[ \mathcal{L} |\nabla u|^2 = 2u_t^{1-k} T_{k-1}(E)^{jk} \left\{ (1 + \epsilon) u_t \nabla_j u_t \nabla_i \nabla_k u - 2 \nabla_i \nabla_j u \nabla_k u_t \nabla_i u_t + u_t^{1-j} \nabla_j u_t \nabla_k u_t |\nabla u_t|^2 \right\} \]
\[ + 2(1 + \epsilon)^{1-k} u_t^{-1} f |\nabla u_t|^2 + 2 \left( \nabla f, \nabla u \right) - 2(1 + \epsilon) u_t^{2-k} \left\{ T_{k-1}(E), \nabla^i u \nabla^i A + R_{ijk}^l \nabla^i u \nabla^i u \right\}. \]

Proof. To begin we take the gradient of the geodesic equation to yield
\[ \nabla_i f = \nabla_i \left[ u_t^{1-k} \sigma_k(E^t_u) \right] \]
\[ = (1 - k) u_t^{1-k} \nabla_i u_t \sigma_k(E^t_u) + u_t^{1-k} \left( T_{k-1}(E^t_u), \nabla^i E^t_u \right) \]
\[ = (1 - k) u_t^{1-k} \nabla_i u_t \sigma_k(E^t_u) \]
\[ + u_t^{1-k} \left( T_{k-1}(E^t_u), (1 + \epsilon) \nabla_i u_t A_u + (1 + \epsilon) u_t \nabla_i A_u - \nabla_i \nabla u_t \nabla u_t - \nabla u_t \nabla_i u \right). \]

A calculation similar to (4.4) shows that
\[ (1 - k) u_t^{1-k} \nabla_i u_t \sigma_k(E^t_u) + u_t^{1-k} \left( T_{k-1}(E^t_u), (1 + \epsilon) \nabla_i u_t A_u \right) \]
\[ = (1 + \epsilon)^{1-k} u_t^{-1} f \nabla_i u_t + u_t^{1-k} \left( T_{k-1}(E), \nabla u_t \nabla u_t \nabla u_t \right) \]
Next we simplify
\[ \nabla_i(A_u)_{jk} = \nabla_i \left[ A_{jk} + \nabla_j \nabla_k u + \nabla_j u \nabla_k u - \frac{1}{2} |\nabla u|^2 g_{jk} \right] \]
\[ = \nabla_i A_{jk} + \nabla_i \nabla_j \nabla_k u + \nabla_i \nabla_j u \nabla_k u + \nabla_j u \nabla_i \nabla_k u - \frac{1}{2} \nabla_i |\nabla u|^2 g_{jk} \]
\[ = \nabla_i A_{jk} + \nabla_i \nabla_j \nabla_k u + R^l_{ijk} \nabla_l u + \nabla_i \nabla_j u \nabla_k u + \nabla_j u \nabla_i \nabla_k u - \frac{1}{2} \nabla_i |\nabla u|^2 g_{jk}. \]

Hence we obtain the identity
\[ (4.5) \quad \mathcal{L} \nabla_i u = \nabla_i f - (1 + \epsilon) u_t^{2-k} T_{k-1}(E)^{jk} \left\{ \nabla_i A_{jk} + R^l_{ijk} \nabla_l u \right\}. \]

On the other hand using (4.2) we have
\[ \mathcal{L} |\nabla u|^2 = 2 \left( \mathcal{L} \nabla u, \nabla u \right) + 2(1 + \epsilon)^{k-1} u_t^{-1} f |\nabla u_t|^2 \]
\[ + 2u_t^{1-k} T_{k-1}(E)^{jk} \left\{ (1 + \epsilon) u_t \nabla_j u_t \nabla_i \nabla_k u - 2 \nabla_i \nabla_j u \nabla_k u_t \nabla_i u_t + u_t^{1-j} \nabla_j u_t \nabla_k u_t |\nabla u_t|^2 \right\} \]
\[ = 2u_t^{1-k} T_{k-1}(E)^{jk} \left\{ (1 + \epsilon) u_t \nabla_j u_t \nabla_i \nabla_k u - 2 \nabla_i \nabla_j u \nabla_k u_t \nabla_i u_t + u_t^{1-j} \nabla_j u_t \nabla_k u_t |\nabla u_t|^2 \right\} \]
\[ + 2(1 + \epsilon)^{1-k} u_t^{1-k} f |\nabla u_t|^2 + 2 \left( \nabla f, \nabla u \right) - 2(1 + \epsilon) u_t^{2-k} \left\{ T_{k-1}(E), \nabla^i u \nabla^i A + R_{ijk}^l \nabla^i u \nabla^i u \right\}, \]
as required. \qed

Proposition 4.12. Given $u$ an admissible solution to $(\ast_{\epsilon,f})$, one has
\[ \sup_{M \times [0,1]} |\nabla u|^2 \leq C. \]

Proof. Without loss of generality we can assume $u < 0$. Choose $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ and let
\[ \Phi = |\nabla u|^2 + \lambda_1 u_t^2 + e^{-\lambda_2 u} + \lambda_3 t(t-1). \]
Lemmas 4.4, 4.9, 4.10, 4.11 show that

\[
\mathcal{L} \Phi \geq \mathcal{L} |\nabla u|^2 + 2\lambda_1 [f_t u_t + (1 + \epsilon)^k f u_{tt} + \epsilon u_{tt}^{2-k} T_{k-1}(E)^{jk} \nabla_j u_t \nabla_k u_t]
\]

\[-\lambda_2 \mathcal{L} \epsilon^{-\lambda_2 u} + \frac{\lambda_2^2}{2} \epsilon^{-\lambda_2 u} u_{tt}^{2-k} \left( T_{k-1}(E_u), \nabla u \otimes \nabla u \right) - C \lambda_2^2 \epsilon^{-\lambda_2 u} \sigma_k(A_u) u_t^2 + \lambda_3 \sigma_k(A_u)
\]

\[
\geq 2 \langle \nabla f, \nabla u \rangle + 2 \sigma_k(A_u) |\nabla u_t|^2 + 2 u_{tt}^{2-k} \left( T_{k-1}(E_u), \nabla_j \nabla_i u \nabla_k \nabla_l u \right)
\]

\[-4 u_{tt}^{1-k} \left( T_{k-1}(E_u)_{ij}, \nabla_i u_t \nabla_k u_t \nabla_j \nabla_k u \right) - 2 u_{tt}^{2-k} \left( T_{k-1}(E_u)_{jk}, \nabla_i A_{jk} + R_{ijk}^l \nabla_l u \right)
\]

\[-C \lambda_1 f_t + \lambda_1 f u_{tt} \]

\[-\lambda_2 \epsilon^{-\lambda_2 u} \left[ f - u_{tt}^{2-k} \left( T_{k-1}(E_u), A \right) + u_{tt}^{2-k} \left( T_{k-1}(E_u), \nabla u \otimes \nabla u \right) - \frac{1}{2} \text{tr} T_{k-1}(E_u) |\nabla u|^2 \right]
\]

\[+ \frac{\lambda^2}{2} \epsilon^{-\lambda_2 u} u_{tt}^{2-k} \left( T_{k-1}(E_u), \nabla u \otimes \nabla u \right) - C \lambda_2^2 \epsilon^{-\lambda_2 u} \sigma_k(A_u) u_t^2 + \lambda_3 \sigma_k(A_u).\]

First we observe that, using the Cauchy-Schwarz inequality and Lemma 2.5

\[4 u_{tt}^{1-k} \left( T_{k-1}(E_u)_{ij}, \nabla_i u_t \nabla_k u_t \nabla_j \nabla_k u \right)
\]

\[= 4 u_{tt}^{1-k} \left[ \left( T_{k-1}(E_u) \frac{\nabla \cdot \nabla u}{\nabla \cdot u} T_{k-1}(E_u) \frac{\nabla \cdot \nabla u}{\nabla \cdot u} \right) u_t \right]
\]

\[\leq 2 u_{tt}^{2-k} \left( T_{k-1}(E_u), \nabla^2 u \cdot \nabla^2 u \right) + 2 u_{tt}^{2-k} \left( T_{k-1}(E_u), \nabla u_t \otimes \nabla u t \right) |\nabla u_t|^2
\]

\[= 2 u_{tt}^{2-k} \left( T_{k-1}(E_u), \nabla^2 u \cdot \nabla^2 u \right) + 2 u_{tt}^{1-k} \left( T_{k-1}(A_u), \nabla u_t \otimes \nabla u t \right) |\nabla u_t|^2
\]

\[= 2 u_{tt}^{2-k} \left( T_{k-1}(E_u), \nabla^2 u \cdot \nabla^2 u \right) + 2 |\sigma_k(A_u) - f u_{tt}^{-1} | |\nabla u_t|^2.
\]

Observe the preliminary inequality

\[u_{tt}^{2-k} \text{tr} T_{k-1}(E_u) = u_{tt}^{2-k} \sigma_k(E_u)
\]

\[\geq u_{tt}^{2-k} \left[ \sigma_k(E_u) \frac{k+1}{k} \right]
\]

\[= u_{tt}^{2-k} \left[ f u_{tt}^{k-1} \right] \frac{k+1}{k}
\]

\[= f^{\frac{k+1}{k}} u_{tt}^{2-k+\frac{k^2}{2}k+1} \]

\[= f^{\frac{k+1}{k}} u_{tt}^{\frac{1}{2}}.
\]

Next observe the estimate

\[\langle \nabla f, \nabla u \rangle \leq C f u_{tt}^{\frac{1}{2}} + C f u_{tt}^{\frac{1}{2}} |\nabla u|^2
\]

\[\leq C f u_{tt}^{-1} + C f u_{tt} + C f u_{tt}^{\frac{1}{2}} |\nabla u|^2
\]

\[\leq C f u_{tt}^{-1} + C f u_{tt} + C f^{\frac{k+1}{k}} u_{tt}^{\frac{1}{2}} |\nabla u|^2.
\]

Next observe that

\[-2 u_{tt}^{2-k} \left( T_{k-1}(E_u)_{jk}, \nabla_i A_{jk} + R_{ijk}^l \nabla_l u \right) \geq - C u_{tt}^{2-k} \text{tr} T_{k-1}(E_u) \left[ 1 + |\nabla u|^2 \right].\]
Combining these preliminary observations and using Proposition 4.8 yields

\[
L \Phi \geq (\lambda_3 - C) f u_{tt}^{-1} + (\lambda_1 - C) f u_{tt} + \left( \frac{3}{4} e^{-\lambda_2 u} - C \right) f^{k-1} u_{tt}^{1/2} |\nabla u|^2 - C \lambda_1 f
+ e^{-\lambda_2 u} u_{tt}^{2-k} \left[ \frac{1}{2} \lambda_3^2 - \lambda_2 \right] (T_{k-1}(E_u), \nabla u \otimes \nabla u)
+ u_{tt}^{2-k} \operatorname{tr} T_{k-1}(E_u) \left[ -C - C |\nabla u|^2 + \frac{\lambda_2}{4} e^{-\lambda_2 u} |\nabla u|^2 \right]
+ \sigma_k(A_u) \left[ \lambda_3 - C \lambda_2^2 \right]
\geq \frac{\lambda_3}{2} f u_{tt}^{-1} + \frac{\lambda_1}{2} f u_{tt} - C \lambda_1 f + u_{tt}^{2-k} \operatorname{tr} T_{k-1}(E_u) \left[ -C + |\nabla u|^2 \right]
\geq u_{tt}^{2-k} \operatorname{tr} T_{k-1}(E_u) \left[ -C + |\nabla u|^2 \right],
\]

where the second inequality follows by choosing \( \lambda_1, \lambda_2 \) large with respect to universal constants and noting that \( e^{-\lambda_2 u} > 1 \) for every choice of \( \lambda_2 \), and then choosing \( \lambda_3 \) large with respect to these choices. The third inequality follows by choosing \( \lambda_3 \) large with respect to \( \lambda_1 \). Using the previously establishing a priori estimates for \( u \) and \( u_t \), at a sufficiently large maximum of \( \Phi \) we will have \( |\nabla u|^2 \geq C \), and hence we see that \( L \Phi > 0 \) at a sufficiently large maximum, a contradiction. The a priori estimate for \( |\nabla u|^2 \) follows. \( \square \)

4.3. \( C^2 \) estimates.

**Lemma 4.13.** Given \( u \) an admissible solution of \( (\ast_e, f) \) we have

\[
\mathcal{L} u_{tt} = -k f^k u_{tt} (1 + \epsilon)^k \mathcal{F}^{ij,kl} [(E_u)_t]_{ij} [(E_u)_t]_{kl}
+ u_{tt}^{-1} \left( T_{k-1}(E), 2(1 + \epsilon)u_{tt}^{-2} u_{tt}^{k-1} \nabla u_t \otimes \nabla u_t - 4u_{tt}^{-1} u_{tt} \nabla u_t \otimes \nabla u_t + 2 \nabla u_t \otimes \nabla u_t \right)
+ (1 + \epsilon) u_{tt}^{k-1} f^{k-1} (f^1_1)_{tt} + 2(k - 1)(1 + \epsilon)^{k-1} u_{tt}^{-1} f^{k-1} (f^1_1)_{tt} - 2(1 + \epsilon)^{k-1} u_{tt}^{-1} u_{tt} f_t
+ (1 + \epsilon) k^{-1} k^1 u_{tt}^{-2} u_{tt}^2 f_t.
\]

**Proof.** First we compute using (4.2) that

\[
\mathcal{L} u_{tt} = (1 + \epsilon)^{k-1} u_{tt}^{-1} f u_{tttt}
\]

(4.6)

\[
+ u_{tt}^{-1} \left( T_{k-1}(E), (1 + \epsilon)^{k-1} u_{tt} \left( \nabla^2 u_{tt} + \nabla u_{tt} \otimes \nabla u + \nabla u \otimes \nabla u_{tt} - \nabla u \nabla u_{tt} \otimes \nabla u + \nabla u_{tt} \nabla u_{tt} \right) \right)
\]

To simplify notation we adopt the following (standard) conventions: for an \( n \times n \) symmetric matrix \( r = r_{ij} \) we denote

\[
\mathcal{F}(r) = \sigma_k(r)^{1/k},
\]

and derivatives of \( \mathcal{F} \) with respect to the entries of \( r \) by

\[
\frac{\partial}{\partial r_{pq}} \mathcal{F}(r) = \mathcal{F}(r)^{pq},
\]

\[
\frac{\partial^2}{\partial r_{pq} \partial r_{rs}} \mathcal{F}(r) = \mathcal{F}(r)^{pq,rs}.
\]

We next need to differentiate the equation, which we can rewrite as

\[
c_e f^k u_{tt}^{k-1} = \sigma_k(E_u)^{1/k} = \mathcal{F}(E_u),
\]
where \( c_\epsilon = (1 + \epsilon)^{\frac{k-1}{k}} \). Differentiating this yields

\[
c_\epsilon \left( f^\frac{k}{t} \right)_{tt}^k + c_\epsilon \frac{k-1}{k} f^\frac{k}{kt} u_{tt}^k = \mathcal{F}^{ij} \left[ \frac{\partial}{\partial t} E_u \right]_{ij} = \frac{1}{k} \sigma_k (E_u)^{\frac{k-1}{k}} \left( T_{k-1}(E_u), (E_u)_t \right).
\]

Differentiating again yields

\[
\mathcal{F}^{ij} \left[ (E_u)_tt \right]_{ij} + \mathcal{F}^{ij, kl} \left[ (E_u)_t \right]_{ij} \left[ (E_u)_t \right]_{kl}
= c_\epsilon \left[ \left( f^\frac{k}{t} \left)_{tt} \frac{k-1}{k} + 2 \frac{k-1}{k} \left( f^\frac{k}{kt} u_{tt}^k - \frac{1}{k} (k-1) f^\frac{k}{k} u_{tt}^2 \right)
+ \frac{k-1}{k} f^\frac{k}{k} u_{tt}^2 \right] \right].
\]

(4.7)

Next we want to get an explicit formula for \((E_u)_{tt}\), which we build up to in stages. We first observe the preliminary computation

\[
(1 + \epsilon)(A_u)_t = \left[ u_{tt}^{-1} E_u + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t \right]_t
= -u_{tt}^{-2} u_{ttt} E_u + u_{tt}^{-1} (E_u)_t - u_{tt}^{-2} u_{ttt} \nabla u_t \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t
\]

Next we compute that

\[
\left[ (E_u)_t \right] = (1 + \epsilon) u_{ttt} A_u + (1 + \epsilon) u_{ttt} (A_u)_t - \nabla u_{ttt} \otimes \nabla u_t - \nabla u_t \otimes \nabla u_t
= (1 + \epsilon) u_{ttt} A_u + (1 + \epsilon) u_{ttt} (A_u)_t
\]

\[
+ (1 + \epsilon) u_{ttt} \left[ \nabla^2 u_t + \nabla u_t \otimes \nabla u_t + 2 \nabla u_t \otimes \nabla u_t + \nabla u_t \otimes \nabla u_t - |\nabla u_t|^2 g + \langle \nabla u_t, \nabla u_t \rangle g \right]
- \nabla u_{ttt} \otimes \nabla u_t - \nabla u_t \otimes \nabla u_t
\]

\[
= (1 + \epsilon) u_{ttt} A_u
+ 2 u_{ttt} \left[ -u_{tt}^{-2} u_{ttt} E_u + u_{tt}^{-1} (E_u)_t - u_{tt}^{-2} u_{ttt} \nabla u_t \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t \right]
+ (1 + \epsilon) u_{ttt} \left[ \nabla^2 u_t + \nabla u_t \otimes \nabla u_t + 2 \nabla u_t \otimes \nabla u_t + \nabla u_t \otimes \nabla u_t - |\nabla u_t|^2 g - \langle \nabla u_t, \nabla u_t \rangle g \right]
- \nabla u_{ttt} \otimes \nabla u_t - \nabla u_t \otimes \nabla u_t.
\]

Next we have, using (4.3),

\[
\left[ (E_u)_{tt} \right] = (1 + \epsilon) u_{ttt} A_u + 2(1 + \epsilon) u_{ttt} (A_u)_t + (1 + \epsilon) u_{ttt} (A_u)_{tt}
- \nabla u_{ttt} \otimes \nabla u_t - 2 \nabla u_{ttt} \otimes \nabla u_t - \nabla u_t \otimes \nabla u_t
\]

\[
= (1 + \epsilon) u_{ttt} A_u + 2(1 + \epsilon) u_{ttt} (A_u)_t
+ \left[ (1 + \epsilon) u_{ttt} \left[ \nabla^2 u_t + \nabla u_t \otimes \nabla u_t + 2 \nabla u_t \otimes \nabla u_t + \nabla u_t \otimes \nabla u_t - |\nabla u_t|^2 g + \langle \nabla u_t, \nabla u_t \rangle g \right]
- \nabla u_{ttt} \otimes \nabla u_t - 2 \nabla u_{ttt} \otimes \nabla u_t - \nabla u_t \otimes \nabla u_t \right]
\]

\[
= (1 + \epsilon) u_{ttt} A_u
+ 2 u_{ttt} \left[ -u_{tt}^{-2} u_{ttt} E_u + u_{tt}^{-1} (E_u)_t - u_{tt}^{-2} u_{ttt} \nabla u_t \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t + u_{tt}^{-1} \nabla u_t \otimes \nabla u_t \right]
+ (1 + \epsilon) u_{ttt} \left[ \nabla^2 u_t + \nabla u_t \otimes \nabla u_t + 2 \nabla u_t \otimes \nabla u_t + \nabla u_t \otimes \nabla u_t - |\nabla u_t|^2 g - \langle \nabla u_t, \nabla u_t \rangle g \right]
- \nabla u_{ttt} \otimes \nabla u_t - 2 \nabla u_{ttt} \otimes \nabla u_t - \nabla u_t \otimes \nabla u_t.
\]

Hence

\[
k \sigma_k (E_u)^{\frac{k-1}{k}} u_{tt}^{1-k} \mathcal{F}^{ij} \left[ (E_u)_t \right]_{ij}
= u_{tt}^{1-k} \{ T_{k-1}(E_u), (1 + \epsilon) u_{tttt} A_u - 2 u_{tt}^{-2} u_{tttt} E_u + 2 u_{tt}^{-1} u_{tttt} (E_u)_t - 2 u_{tt}^{-2} u_{tttt} \nabla u_t \otimes \nabla u_t
+ 4 u_{tt}^{-1} u_{tttt} \nabla u_t \otimes \nabla u_t + (1 + \epsilon) \{ u_{tt}^{-2} \nabla u_t + 2 u_{tt} \nabla u_t \otimes \nabla u_t + 2 u_{tt} \nabla u_t \otimes \nabla u_t
- u_{tt}^{-2} \nabla u_t \otimes \nabla u_t \} \} - 2 \nabla u_{ttt} \otimes \nabla u_t - 2 \nabla u_{ttt} \otimes \nabla u_t \}
= \sum_{i=1}^{12} A_i.
Comparing against (4.6) yields
\[
\mathcal{L}u_{tt} = A_1 + A_6 + A_7 + A_{10} + A_{11}
+ u_{tttt}u_{tt}^{-k}\left(\frac{T_{k-1}(E), -(1+\epsilon)A_{tt} + u_{tt}^{-1}\nabla u_t \otimes \nabla u_t \right) + (1+\epsilon)^{k-1}u_{tt}^{-1}f u_{tttt}
= A_1 + A_6 + A_7 + A_{10} + A_{11}
+ u_{tttt}\left[u_{tt}^{-k}\left(\frac{T_{k-1}(E), -E \right) + (1+\epsilon)^{k-1}u_{tt}^{-1}f \right]
= A_1 + A_6 + A_7 + A_{10} + A_{11} + u_{tttt}\left[k u_{tt}^{-k}\sigma_k(E) + (1+\epsilon)^{k-1}u_{tt}^{-1}f \right]
= A_1 + A_6 + A_7 + A_{10} + A_{11} + f (1+\epsilon)^{k-1}(1-k)u_{tt}^{-1}u_{tttt}.
\]

Hence we obtain
\[
\mathcal{L}u_{tt} = k \sigma_k(E_u) \frac{k-1}{k} u_{tt}^{-1-k} f^{ij} \left[ (E_u)_{1i} - u_{tt}^{-1-k} \left(\frac{T_{k-1}(E_u), -2u_{tt}^{-2}u_{tttt} E_u + 2u_{tt}^{-1}u_{tttt}(E_u) \right)
- 2(1+\epsilon)u_{tt}^{-2}u_{tttt} \nabla u_t \otimes \nabla u_t + 4u_{tt}^{-1}u_{tttt} \nabla u_t \otimes \nabla u_t + 2u_{tt} \nabla u_t \otimes \nabla u_t
- (1+\epsilon)u_{tt} |\nabla u_t|^2 g - 2\nabla u_{tt} \otimes \nabla u_t \right] + f (1+\epsilon)^{k-1}(1-k)u_{tt}^{-1}u_{tttt}
= k \sigma_k(E_u) \frac{k-1}{k} u_{tt}^{-1-k} \left[ \mathcal{L}^{ij, kl} \left[ (E_u)_{1i} \left[ (E_u)_{1i} \right] \right] \right]
+ (1+\epsilon)^{k-1}(1-k)u_{tt}^{-1}u_{tttt}
= \sum_{i=1}^{13} A_i.
\]

We now clean up some of the lower order terms. In particular we express
\[
k \sigma_k(E) = k \left[f u_{tt}^{-k} (1+\epsilon)^{k-1} \right] \frac{k-1}{k} u_{tt}^{-1-k} = k \left[f \frac{k-1}{k} u_{tt}^{-k} (1+\epsilon)^{\frac{(k-1)^2}{k}}.\right.
\]

Then observe
\[
A_2 = \left(k \sigma_k(E) \frac{k-1}{k} u_{tt}^{-1-k} \right) \left((1+\epsilon) \frac{k-1}{k} \left(f \frac{1}{k} \right)_{tt} u_{tt}^{-k-1} \right)
= \left(k f \frac{k-1}{k} u_{tt}^{-k} (1+\epsilon)^{\frac{(k-1)^2}{k}} \right) \left((1+\epsilon) \frac{k-1}{k} \left(f \frac{1}{k} \right)_{tt} u_{tt}^{-k-1} \right)
= \left(1+\epsilon\right)^{k-1} k f \frac{k-1}{k} \left(f \frac{1}{k} \right)_{tt}
\]
Next
\[
A_3 = \left(k \sigma_k(E) \frac{k-1}{k} u_{tt}^{-k} \right) \left((1+\epsilon) \frac{k-1}{k} \frac{2k-1}{k} \left(f \frac{1}{k} \right)_{tt} u_{tt}^{-k-1} \right)
= \left(k f \frac{k-1}{k} u_{tt}^{-k} \left(1+\epsilon\right)^{\frac{(k-1)^2}{k}} \right) \left((1+\epsilon) \frac{k-1}{k} \frac{2k-1}{k} \left(f \frac{1}{k} \right)_{tt} u_{tt}^{-k-1} \right)
= 2(k-1)(1+\epsilon)^{k-1} u_{tt}^{-1} \frac{k-1}{k} \left(f \frac{1}{k} \right)_{tt} u_{tttt}
\]
Next
\[
A_4 = \left(k f \frac{k-1}{k} u_{tt}^{-k} \left(1+\epsilon\right)^{\frac{(k-1)^2}{k}} \right) \left(-\left(1+\epsilon\right) \frac{k-1}{k} \frac{k-1}{k} \left(f \frac{1}{k} \right)_{tt} u_{tt}^{-k-1} \right)
= -\left(1+\epsilon\right)^{k-1} \left(k-1\right) u_{tt}^{-2} u_{tttt}.
\]
Next note that
\[ A_5 = k \sigma_k(E_u) \frac{k-1}{k} u_{tt}^{1-k} C_e \frac{k-1}{k} f^{1/k} u_{ttt} = (k-1)(1+\epsilon)^{k-1} f u_{tt}^{-1} u_{ttt} = -A_{13}. \]
Also observe
\[ A_6 = u_{tt}^{1-k} \left( T_{k-1}(E), 2u_{tt}^{-2} u_{ttt} E_u \right) = 2k u_{tt}^{1-k} u_{ttt} \sigma_k(E) = 2k u_{tt}^{1-k} f u_{ttt}^{k-1} (1+\epsilon)^{k-1} = 2k(1+\epsilon)^{k-1} u_{tt}^{-2} u_{ttt} f. \]
Lastly
\[ A_7 = -2u_{tt}^{1-k} \left( T_{k-1}(E), u_{tt}^{-1} u_{ttt}(E_u)_t \right) = -2u_{tt}^{1-k} u_{ttt} \left[ \sigma_k(E) \right]_t = -2(1+\epsilon)^{k-1} u_{tt}^{-2} u_{ttt} \left[ f u_{ttt}^{k-1} \right]_t = -2(1+\epsilon)^{k-1} u_{tt}^{-1} u_{ttt} \left[ f_t + (k-1) f u_{tt}^{-1} u_{ttt} \right]. \]
Inserting these simplifications into (4.8) yields the result. \[\square\]

**Proposition 4.14.** Given \( u \) an admissible solution to \((\ast, f)\), one has
\[ \sup_{M \times [0,1]} u_{tt} \leq C\epsilon^{-1}. \]

**Proof.** Let’s begin with a preliminary estimate for \( L u_{tt} \). Returning to Lemma 4.13 and considering the terms in order, one first observes by convexity of \( F \) that
\[ -k f \frac{k}{k+1} u_{tt} (1+\epsilon)^{k} \mathcal{F}^{ij,kl} [(E_u)_i]_j [(E_u)_t]_k \geq 0. \]
Also, by an application of the Cauchy Schwarz inequality one has the matrix inequality
\[ 2u_{tt}^{-2} u_{ttt}^2 \nabla u_t \otimes \nabla u_t - 4u_{tt}^{-1} u_{ttt} \nabla u_t \otimes \nabla u_t + 2\nabla u_{tt} \otimes \nabla u_{tt} \geq 0. \]
Also, since \( u \) is an admissible solution we have
\[ u_{tt}^{1-k} \left( T_{k-1}(E), u_{tt} \left| \nabla u_t \right|^2 g \right) = u_{tt}^{2-k} \left| \nabla u_t \right|^2 \text{tr} T_{k-1}(E) \geq 0. \]
Also we observe
\[ (1+\epsilon)^{k-1} k f \frac{k+1}{k} (f^{1/k})_t \leq C f \frac{k-1}{k} \left[ f^{1/k-1} f_t + f^{1-2} f_t^2 \right] \leq C f. \]
Next
\[ 2(k-1)(1+\epsilon)^{k-1} u_{tt}^{-1} f \frac{k}{k+1} (f^{1/k})_t u_{ttt} \leq C f \frac{k-1}{k} (f^{1/k-1} f_t) u_{tt}^{-1} u_{ttt} \leq C f u_{tt}^{-1} u_{ttt} \leq C \delta^{-1} f + C \delta f u_{tt}^{-2} u_{ttt}^2. \]
Also
\[ -2(1+\epsilon)^{k-1} u_{tt}^{-1} u_{ttt} f_t \leq C f u_{tt}^{-1} u_{ttt} \leq C \delta^{-1} f + C \delta f u_{tt}^{-2} u_{ttt}^2. \]
Combining these estimates and choosing \( \delta \) sufficiently small leads to the preliminary estimate (4.10)
\[ L u_{tt} \geq -2u_{tt}^{2-k} \left( T_{k-1}(E), \nabla u_t \otimes \nabla u_t \right) - C f. \]
Differentiating again yields
\[ \mathcal{L} u_t^2 \geq -C f + 2 f u_{tt} + 2 \epsilon u_{tt}^{2-k} (T_{k-1}(E), \nabla u_t \otimes \nabla u_t). \]

Now fix constants \( \lambda_i \) and let \( \Phi = u_{tt} + \lambda_1^{-1} u_t^2 + \lambda_2 (t-1). \)

Choosing \( \lambda_1 \geq 1 \), combining Lemma 4.14 with (4.10) and (4.11) yields
\[
\mathcal{L} \Phi \geq 2u_t^{2-k} (T_{k-1}(E), (\lambda_1 - 1)\nabla u_t \otimes \nabla u_t) - f (C + C\lambda_1^{-1}) + 2\lambda_1^{-1} f u_{tt} + \lambda_2 f u_t^{-1} \geq f \left[ (2\lambda_1^{-1} - \delta (C + C\lambda_1^{-1})) u_{tt} + (\lambda_2 - \delta^{-1} (C + C\lambda_1^{-1})) \right].
\]

If we now choose \( \delta \) small above with respect to universal constants and then choose \( \lambda_2 \) large with respect to \( \delta \) we conclude
\[ \mathcal{L} \Phi > 0, \]
and hence \( \Phi \) cannot have an interior maximum. The proposition follows. \( \square \)

**Lemma 4.15.** Given \( u \) an admissible solution of \((\ast_{\epsilon, f})\) we have
\[
\mathcal{L}(\Delta u) = -k \sigma_k(E) \frac{k-1}{k} u_{tt}^{1-k} f^{(ij),(kl)} \nabla_p(E_u)_{ij} \nabla_p(E_u)_{kl} + u_t^{1-k} (T_{k-1}(E), (1 + \epsilon) u_t u_t \nabla^2 u + \nabla u \otimes \nabla u + \nabla u \otimes \nabla u - (\nabla \Delta u, \nabla u) g - \nabla \Delta u_t \otimes \nabla u_t - \nabla u_t \otimes \nabla \Delta u_t + u_t^{1-k} \nabla u_t \otimes \nabla u_{tt}).
\]

Next we differentiate the equation, which we rewrite as
\[ c_{\epsilon} f^1 u_{tt}^{-1/2} = \sigma_k(E_u)^{1/2} =: \mathcal{F}(E_u). \]

Differentiating yields
\[ c_{\epsilon} \nabla_p(f^{1/2}) u_{tt}^{1/2} + c_{\epsilon} \left( \frac{k-1}{k} \right) f^{1/2} u_{tt}^{-1/2} \nabla_p u_{tt} = \mathcal{F}^{ij} \nabla_p(E_u)_{ij}. \]

Differentiating again yields
\[
\mathcal{F}^{ij} (\Delta E_u)_{ij} + \mathcal{F}^{(ij),(kl)} \nabla_p(E_u)_{ij} \nabla_w(E_u)_{kl} = c_{\epsilon} f^{1/2} u_{tt}^{-1/2} \nabla_p u_{tt}^{1/2} + \left( \frac{k-1}{k} \right) f^{1/2} u_{tt}^{-1/2} \nabla u_{tt}^{1/2} + \left( \frac{k-1}{k} \right) f^{1/2} u_{tt}^{-1/2} \Delta u_{tt}.
\]

Next we have
\[
\nabla_p(E_u)_{ij} = \nabla_p [(1 + \epsilon) u_{tt} (A_u)_{ij} - \nabla_i u_t \nabla_j u_t] = (1 + \epsilon) \nabla_p u_{tt} (A_u)_{ij} + (1 + \epsilon) u_{tt} \nabla_p (A_u)_{ij} - \nabla_p \nabla_i u_t \nabla_j u_t - \nabla_i u_t \nabla_p \nabla_j u_t.
\]
Differentiating again and commuting derivatives yields
\[
(\Delta E_u)_{ij} = (1 + \epsilon)\Delta u_{tt}(A_u)_{ij} + 2(1 + \epsilon)\nabla_p u_{tt}\nabla_p (A_u)_{ij} + (1 + \epsilon)u_{tt}\Delta (A_u)_{ij} \\
- \nabla_i\Delta u_t\nabla_j u_t - \nabla_i u_t \nabla_j \Delta u_t - 2\nabla_i \nabla_p u_{tt}\nabla_j \nabla_p u_t \\
- R_{ijp}\nabla_p u_t \nabla_j u_t - R_{jip}\nabla_p u_t \nabla_i u_t.
\]

Differentiating the equation for the Schouten tensor yields
\[
\nabla_p (A_u)_{ij} = \nabla_p A_{ij} + \nabla_p \nabla_i u\nabla_j u + \nabla_i \nabla_p u\nabla_j u + \nabla_i u \nabla_j \nabla_p u - \frac{1}{2} \nabla_p |\nabla u|^2 g.
\]
This implies
\[
\Delta (A_u)_{ij} = \Delta A_{ij} + \nabla_i \nabla_j \Delta u + \nabla_i \Delta u \nabla_j u + \nabla_i u \nabla_j \Delta u \\
+ 2\nabla_i \nabla_p u\nabla_j \nabla_p u - |\nabla^2 u|^2 g_{ij} - (\nabla u, \nabla \Delta u) g_{ij} + O(|\nabla^2 u| + |\nabla u|^2 + 1).
\]

On the other hand it is also useful to express
\[
(1 + \epsilon)\nabla_p (A_u)_{ij} = \nabla_p \left[u_{tt}^{-1}(E_u)_{ij} + u_{tt}^{-1}\nabla_i u_t \nabla_j u_t \right] \\
= u_{tt}^{-1}\nabla_p (E_u)_{ij} - u_{tt}^{-2}(E_u)_{ij}\nabla_p u_{tt} - u_{tt}^{-2}\nabla_p u_{tt}\nabla_i u_t \nabla_j u_t \\
+ u_{tt}^{-1}\nabla_i \nabla_p u_t \nabla_j u_t + u_{tt}^{-1}\nabla_i u_t \nabla_j \nabla_p u_t.
\]
Combining the above calculations yields
\[
\Delta (E_u)_{ij} = (1 + \epsilon)\Delta u_{tt}(A_u)_{ij} + 2\nabla_p u_{tt} \left[u_{tt}^{-1}\nabla_p (E_u)_{ij} - u_{tt}^{-2}(E_u)_{ij}\nabla_p u_{tt} - u_{tt}^{-2}\nabla_p u_{tt}\nabla_i u_t \nabla_j u_t \right] \\
+ u_{tt}^{-1}\nabla_i \nabla_p u_t \nabla_j u_t + u_{tt}^{-1}\nabla_i u_t \nabla_j \nabla_p u_t \\
+ (1 + \epsilon)u_{tt} \left[\nabla_i \nabla_j \Delta u + \nabla_i \Delta u \nabla_j u + \nabla_i u \nabla_j \Delta u \right] \\
+ 2\nabla_i \nabla_p u\nabla_j \nabla_p u - |\nabla^2 u|^2 g_{ij} - (\nabla u, \nabla \Delta u) g_{ij} + O(|\nabla^2 u| + |\nabla u|^2 + 1) \\
- \nabla_i \Delta u_t \nabla_j u_t - \nabla_i u_t \nabla_j \Delta u_t - 2\nabla_i \nabla_p u_t \nabla_j \nabla_p u_t.
\]
Thus
\[
k\sigma_k(E)\frac{k-1}{k}u_{tt}^{-k}\mathcal{F}^{ij}(\Delta E_u)_{ij} \\
= u_{tt}^{-k} \left(T_{k-1}(E), (1 + \epsilon)\Delta u_{tt}(A_u)_{ij} + 2u_{tt}^{-1}\nabla_p u_{tt}\nabla_p (E_u)_{ij} - 2u_{tt}^{-2}|\nabla u_{tt}|^2 (E_u)_{ij} - 2u_{tt}^{-2}|\nabla u_{tt}|^2 \nabla_i u_t \nabla_j u_t \\
+ 4u_{tt}^{-1}\nabla_p u_{tt}\nabla_i \nabla_p u_t \nabla_j u_t + (1 + \epsilon)u_{tt} \nabla_i \nabla_j \Delta u + 2(1 + \epsilon)u_{tt} \nabla_i \nabla_j \Delta u \\
+ 2(1 + \epsilon)u_{tt} \nabla_i \nabla_p u\nabla_j \nabla_p u - (1 + \epsilon)u_{tt} |\nabla^2 u|^2 g_{ij} - (1 + \epsilon)u_{tt} (\nabla u, \nabla \Delta u) g_{ij} \\
- 2\nabla_i \Delta u_t \nabla_j u_t - 2\nabla_i \nabla_p u_t \nabla_j \nabla_p u_t + u_{tt} O(|\nabla^2 u| + |\nabla u|^2 + 1) \right) \\
= \sum_{i=1}^{13} A_{ij}.
\]
Comparing this against (4.12) yields

\[ \mathcal{L}(\Delta u) = A_1 + A_6 + A_7 + A_{10} + A_{11} + u_t^{1-k} \Delta u_{tt} (T_{k-1}(E), -E) + (1 + \epsilon) A_u + \nabla u_t \otimes \nabla u_t + (1 + \epsilon) k^{-1} u_{tt}^{-1} f \Delta u_{tt} \]

\[ = A_1 + A_6 + A_7 + A_{10} + A_{11} + (1 + \epsilon) k^{-1} u_{tt}^{-1} f \Delta u_{tt} \]

Hence, collecting these calculations yields

\[ \mathcal{L}(\Delta u) = \frac{k \sigma_k(E) k^{1-k} u_t^{1-k}}{u_{tt}^{1-k}} \mathcal{F}(\Delta E_u)_{ij} \]

\[ - u_t^{1-k} \left( T_{k-1}(E), 2 u_{tt}^2 \nabla_p u_{tt} \nabla_p (E_u)_{ij} - 2 u_{tt}^2 |\nabla u_{tt}|^2 (E_u)_{ij} - 2 u_{tt}^2 |\nabla u_{tt}|^2 \nabla_i u_t \otimes \nabla_j u_t \right) \]

\[ + u_t^{1-k} \nabla_p u_{tt} \nabla_p u_{tt} \nabla_j u_t + 2(1 + \epsilon) u_{tt} \nabla_i \nabla_p u_{tt} \nabla_j u_t - (1 + \epsilon) u_{tt} \nabla_j^2 u_t \]

\[ - 2 \nabla_i \nabla_p u_{tt} \nabla_j u_t + u_{tt} \mathcal{O}(\nabla^2 u_t + |\nabla u_t|^2 + 1) \]

\[ + (1 - k)(1 + \epsilon) k^{-1} u_{tt}^{-1} f \Delta u_{tt} \]

\[ = -k \sigma_k(E) f + c_k \sigma_k(E) \frac{k^{1-k} u_t^{1-k}}{u_{tt}^{1-k}} \left[ (f + \frac{1}{k}) u_{tt}^{\frac{1}{k}} + 2 \left( \frac{k - 1}{k} \right) \left( \nabla(f + \frac{1}{k}), \nabla u_t \right) u_{tt}^{\frac{1}{k}} \right] \]

\[ + \frac{1}{k} \left( \frac{k - 1}{k} \right) f + \frac{1}{k} u_{tt}^{\frac{1}{k}} \Delta u_{tt} \]

\[ + u_t^{1-k} \left( T_{k-1}(E), -2 u_{tt}^2 \nabla_p u_{tt} \nabla_p (E_u)_{ij} + 2 u_{tt}^2 |\nabla u_{tt}|^2 (E_u)_{ij} + 2 u_{tt}^2 |\nabla u_{tt}|^2 \nabla_i u_t \otimes \nabla_j u_t \right) \]

\[ - 4 u_{tt}^2 \nabla_p u_{tt} \nabla_p u_{tt} \nabla_j u_t + 2 \nabla_i \nabla_p u_{tt} \nabla_j u_t - 2(1 + \epsilon) u_{tt} \nabla_i \nabla_p u_{tt} \nabla_j u_t \]

\[ + (1 + \epsilon) u_{tt} |\nabla^2 u_t|^2 g_{ij} + u_{tt} \mathcal{O}(\nabla^2 u_t + |\nabla u_t|^2 + 1) \]

\[ + (1 - k)(1 + \epsilon) k^{-1} u_{tt}^{-1} f \Delta u_{tt} \]

\[ = \sum_{i=1}^{14} A_i. \]

Now we simplify

\[ A_2 = \left( k \sigma_k(E) \frac{k^{1-k} u_t^{1-k}}{u_{tt}^{1-k}} \right) \left[ c_k \frac{k^{1-k}}{u_{tt}^{1-k}} \Delta(f + \frac{1}{k}) \right] \]

\[ = \left( k f \frac{k^{1-k} u_t^{1-k}}{u_{tt}^{1-k}} (1 + \epsilon)^{\frac{(k-1)^2}{k}} \left( 1 + \epsilon \right) k^{-1} u_{tt}^{\frac{1}{k}} \Delta(f + \frac{1}{k}) \right) \]

\[ = k(1 + \epsilon) k^{-1} f \frac{k^{1-k} u_t^{1-k}}{u_{tt}^{1-k}} \Delta(f + \frac{1}{k}). \]

Next

\[ A_3 = \left( k \sigma_k(E) \frac{k^{1-k} u_t^{1-k}}{u_{tt}^{1-k}} \right) \left( 2 c_k \left( \frac{k - 1}{k} \right) \left( \nabla(f + \frac{1}{k}), \nabla u_{tt} \right) u_{tt}^{\frac{1}{k}} \right) \]

\[ = \left( k f \frac{k^{1-k} u_t^{1-k}}{u_{tt}^{1-k}} (1 + \epsilon)^{\frac{(k-1)^2}{k}} \left( 2(1 + \epsilon)^{\frac{k-1}{k}} \frac{k - 1}{k} \right) \left( \nabla(f + \frac{1}{k}), \nabla u_{tt} \right) u_{tt}^{\frac{1}{k}} \right) \]

\[ = 2(1 + \epsilon)^{k-1}(k-1) f \frac{k^{1-k} u_t^{1-k}}{u_{tt}^{1-k}} \left( \nabla(f + \frac{1}{k}), \nabla u_{tt} \right) \]

\[ = (1 + \epsilon)^{k-1} \left( 2 - \frac{2}{k} \right) u_{tt}^{\frac{1}{k}} \left( \nabla f, \nabla u_{tt} \right). \]
Next
\[ A_4 = - \left( k \sigma_k(E)^{k-1} u_{tt}^{1-k} \right) \left( c_e \frac{1}{k} \left( \frac{k-1}{k} \right) f^{k} u_{tt}^{1-k} \right) \left( \nabla u_{tt} \right)^2 \]
\[ = - \left( k f^{k} u_{tt}^{1-k} (1 + \epsilon)^{k-1} \right) \left( \left( 1 + \epsilon \right)^{k-1} \frac{1}{k} \left( \frac{k-1}{k} \right) f^{k} u_{tt}^{1-k} \right) \left( \nabla u_{tt} \right)^2 \]
\[ = -(1 + \epsilon)^{k-1} \left( \frac{k-1}{k} \right) f u_{tt}^{2} \left( \nabla u_{tt} \right)^2. \]

Next
\[ A_5 = \left( k \sigma_k(E)^{k-1} u_{tt}^{1-k} \right) \left( c_e \frac{1}{k} \left( \frac{k-1}{k} \right) f^{k} u_{tt}^{1-k} \Delta u_{tt} \right) \]
\[ = \left( k f^{k} u_{tt}^{1-k} (1 + \epsilon)^{k-1} \right) \left( \left( 1 + \epsilon \right)^{k-1} \frac{1}{k} \left( \frac{k-1}{k} \right) f^{k} u_{tt}^{1-k} \right) \Delta u_{tt} \]
\[ = (k - 1)(1 + \epsilon)^{k-1} f u_{tt}^{1-k} \Delta u_{tt} \]
\[ = - A_{14}. \]

Next
\[ A_6 = - 2 u_{tt}^{1-k} \nabla \cdot \left( T_{k-1}(E), u_{tt}^{-1} \nabla (E u) \right) \]
\[ = - 2 u_{tt}^{1-k} \nabla \cdot \sigma_k(E) \]
\[ = - 2(1 + \epsilon)^{k-1} u_{tt}^{1-k} \nabla \cdot \sigma_k(E) \]
\[ = - 2(1 + \epsilon)^{k-1} u_{tt}^{1-k} \nabla \cdot \sigma_k(E) \]
\[ = - 2(1 + \epsilon)^{k-1} u_{tt}^{1-k} \nabla \cdot \sigma_k(E) \]
\[ = -(1 + \epsilon)^{k-1} f u_{tt}^{2} \left( \nabla u_{tt} \right)^2. \]

Lastly
\[ A_7 = 2 u_{tt}^{1-k} u_{tt}^{-2} \left( T_{k-1}(E), E \right) \]
\[ = 2 k u_{tt}^{1-k} u_{tt}^{-2} \left( \nabla u_{tt} \right)^2 \]
\[ = 2 k (1 + \epsilon)^{k-1} f u_{tt}^{-2} \left( \nabla u_{tt} \right)^2. \]

Collecting these simplifications yields the result.

\[ \square \]

**Proposition 4.16.** Given \( u \) an admissible solution to \((*,f)\), one has
\[ \sup_{M \times [0,1]} \Delta u \leq C \epsilon^{-1}. \]

*Proof.* We begin with a preliminary estimate for \( \mathcal{L} \Delta u \). Returning to Lemma 4.15 and considering the terms in order, one first observes by convexity of \( F \) that
\[ - k f^{k-1} u_{tt}^{1-k} (1 + \epsilon)^{k-1} F^{ij,kl} \left[ \nabla_p (E u) \right]_{ij} \left[ \nabla_p (E u) \right]_{kl} \geq 0. \]

Also, by an application of the Cauchy Schwarz inequality one has the matrix inequality
\[ 2 u_{tt}^{2} \left( \nabla u_{tt} \right)^{2} \nabla_i u_t \nabla_j u_t - 4 u_{tt}^{1} \nabla_p u_t \nabla_i \nabla_p u_t \nabla_j u_t + 2 \nabla_i \nabla_p u_t \nabla_j \nabla_p u_t \geq 0. \]

Also we observe
\[ (1 + \epsilon)^{k-1} k f^{k-1} \Delta (f^{k}) \leq C f^{k-1} \left[ f^{k-1} \Delta f + f^{k-2} \left( \nabla f \right)^{2} \right] \leq C f. \]

Next
\[ - 2 k (1 + \epsilon)^{k-1} u_{tt}^{1-k} \left( \nabla f, \nabla u_{tt} \right) \leq C f u_{tt}^{-2} \left( \nabla u_{tt} \right)^{2} \]
\[ \leq C \delta^{-1} f + C \delta u_{tt}^{-2} \left( \nabla u_{tt} \right)^{2}. \]
Combining these estimates and choosing \( \delta \) sufficiently small leads to the preliminary estimate
\[
\mathcal{L} \Delta u \geq -2(1 + \epsilon) u_{tt}^{2-k} \langle T_{k-1}(E), \nabla_i \nabla_j u \nabla_p u \rangle
\]
\[
+ u_{tt}^{2-k} \left\{ T_{k-1}(E), \frac{\nabla^2 u}{g} + O\left( |\nabla^2 u| + |u|^2 + 1 \right) \right\} - C f.
\]
(4.15)

Similar considerations applied to Lemma 4.11 yield
\[
\mathcal{L} |\nabla u|^2 \geq 2\epsilon u_{tt}^{2-k} T_{k-1}(E) \frac{\nabla^2 u}{g} - C f - u_{tt}^{2-k} \langle T_{k-1}(E), O(1) \rangle.
\]
(4.16)

Now fix a constant \( \lambda \in \mathbb{R} \) and consider
\[
\Phi = \Delta u + \epsilon^{-1} \left[ (1 + \epsilon) |\nabla u|^2 + u_t^2 + \lambda t (t - 1) \right].
\]
Combining Lemma 4.4 with lines (4.11), (4.15), and (4.16) yields
\[
\mathcal{L} \Phi \geq u_{tt}^{2-k} \left\{ T_{k-1}(E), \frac{\nabla^2 u}{g} + O\left( |\nabla^2 u| + |u|^2 + 1 \right) + \epsilon^{-1} O(1) \right\}
\]
\[
- C \epsilon^{-1} f + 2\epsilon^{-1} f u_{tt} + \lambda \epsilon^{-1} f u_{tt}^{-1}.
\]
First we observe that at a sufficiently large maximum of \( \Phi \), the existing a priori estimates imply that \( \Delta u \) is also large. In particular, at a maximum for \( \Phi \) where \( |\nabla^2 u| \geq C \epsilon^{-\frac{1}{2}} \) we obtain
\[
|\nabla^2 u| g + O\left( |\nabla^2 u| + |u|^2 + 1 \right) + \epsilon^{-1} O(1) \geq \frac{1}{2} |\nabla^2 u|^2 g,
\]
and hence since \( u \) is an admissible solution we have
\[
u_{tt}^{2-k} \left\{ T_{k-1}(E), \frac{\nabla^2 u}{g} + O\left( |\nabla^2 u| + |u|^2 + 1 \right) + \epsilon^{-1} O(1) \right\} \geq \frac{1}{2} u_{tt}^{2-k} |\nabla^2 u|^2 \text{tr} T_{k-1}(E) \geq 0.
\]
But then we can estimate
\[
C \epsilon^{-1} f \leq \epsilon^{-1} f u_{tt} + C \epsilon^{-1} f u_{tt}^{-1}.
\]
hence choosing \( \epsilon \) sufficiently large we obtain, at a sufficiently large maximum for \( \Phi \) which satisfies \( \Delta u \geq C \epsilon^{-\frac{1}{2}} \), one has
\[
\mathcal{L} \Phi > 0,
\]
a contradiction. The a priori estimate for \( \Delta u \) follows directly.

4.4. Boundary estimates. By Proposition 4.8 we already have the boundary estimate
\[
\sup_{M \times (0,1)} \left[ |u| + |u_t| + |\nabla u| \right] \leq C.
\]

In this section we prove boundary estimates for second order derivatives:

**Proposition 4.17.** Given \( u \) an admissible solution to \((\ast_{\epsilon, f})\), one has
\[
\sup_{M \times (0,1)} \left[ |u_{tt}| + |\nabla u_t| + |\nabla^2 u| \right] \leq C.
\]

**Proof.** A bound for \( |\nabla^2 u| \) is immediate. If we can prove a bound for the ‘mixed’ term \( |\nabla u_t| \), then restricting the equation for \( u \) to \( t = 0 \) we have
\[
(1 + \epsilon) u_{tt} \langle \cdot, 0 \rangle \sigma_k(A_u(\cdot, 0)) = \langle T_{k-1}(A_u(\cdot, 0)), \nabla u_t (\cdot, 0) \otimes \nabla u_t (\cdot, 0) \rangle + f
\]
\[
\leq C_1 \left[ 1 + |\nabla u_0|^2 + |\nabla^2 u_0| \right] |\nabla u_t (\cdot, 0)|^2 + C_2.
\]
Since \( u_0 \) is admissible,
\[
\sigma_k(A_u(\cdot, 0)) = \sigma_k(A_{u_0}) \geq \delta_0 > 0,
\]
and it follows that
\[
\sup_M u_{tt} (\cdot, 0) \leq C_0 (1 + \sup_M |\nabla u_t (\cdot, 0)|^2),
\]
Therefore, hence if 0 divides \( \lambda \) since 0.

We conclude that \( \Phi \leq \lambda \), \( t, C \)

\[ \frac{\partial}{\partial x^i}(u(t) - u_0(x)) + \left[ e^{\lambda (u_0 - u)} - e^{\lambda t} \right] + \Lambda t(t - 1), \]

where \( \lambda, \Lambda \) and \( \Psi \) are constants yet to be determined. By making an appropriate choice of these constants, we claim that \( \Psi \) attains a non-positive maximum on the boundary of of \( M \times [0, \tau] \). Assuming for the moment this is true, let us see how a bound for \( \nabla u_t \) follows.

Choose a point \( x_0 \in M \), and a unit tangent vector \( X \in T_{x_0}M \). Let \( \{x^i\} \) be a local coordinate system with \( X = \frac{\partial}{\partial x^i} \) at \( x_0 \). Then

\[ \frac{\partial}{\partial x^i}(u(t) - u_0(x)) + \left[ e^{\lambda (u_0 - u)} - e^{\lambda t} \right] + \Lambda t(t - 1) \]

\[ \leq |\nabla (u - u_0)(x, t)| + \left[ e^{\lambda (u_0 - u)} - e^{\lambda t} \right] + \Lambda t(t - 1) \]

\[ \leq 0. \]

Therefore,

\[ 0 \geq \lim_{t \to 0^+} \frac{1}{t} \left\{ \frac{\partial}{\partial x^i}(u(t) - u_0(x)) + \left[ e^{\lambda (u_0 - u)} - e^{\lambda t} \right] + \Lambda t(t - 1) \right\} \]

\[ = \frac{\partial}{\partial x^i}u_t(x_0, 0) + \frac{1}{t} \left[ e^{\lambda (u_0 - u)} - e^{\lambda t} \right] + \Lambda (t - 1). \]

Since \( u_t \) is bounded, an upper bound on \( \frac{\partial}{\partial x^i}u_t \) follows. Since \( X = \frac{\partial}{\partial x^i} \) was arbitrary, we obtain a bound on \( |\nabla u_t(x_0, 0)| \).

To see that such a choice of \( \lambda, \Lambda, \Psi \) and \( \tau \) are possible, we first note that

\[ \Psi(x, 0) = 0. \]

Since \( |\nabla u| \) is bounded,

\[ \Psi(x, \tau) = |\nabla u(x, \tau) - \nabla u_0(x)| + \left[ e^{\lambda (u_0(x) - u(x, \tau)) + \Psi} \right] + \Lambda \tau(\tau - 1) \]

\[ \leq C_1 + \left| e^{\lambda (u_0(x) - u(x, \Psi)) + \Psi} \right| + \Lambda \tau(\tau - 1). \]

Since \( |u_t| \) is also bounded,

\[ \left| e^{\lambda (u_0(x) - u(x, \tau)) + \Psi} \right| \leq C_2 e^{C_2 \lambda \tau + \Psi}, \]

hence if \( 0 < \tau < 1/2 \),

\[ \Psi(x, \tau) \leq C_1 + C_2 \tau e^{C_2 \lambda \tau + \Psi} - \Lambda \tau(1 - \tau) \]

\[ \leq C_1 + (C_2 \lambda e^{C_2 \lambda \tau} - \Lambda/2) \tau. \]

Therefore, if \( \Lambda \) is chosen large enough (depending on \( \tau, C_1, C_2, \lambda, \) and \( \Psi \)), then

\[ \Psi(x, \tau) \leq 0. \]

We conclude that \( \Psi \leq 0 \) on \( \partial(M \times [0, \tau]) \).

Assume the maximum of \( \Psi \) is attained at a point \( (x_0, t_0) \) which is interior (i.e., \( 0 < t_0 < \tau \)). Let

\[ \eta = \frac{\nabla u(x, 0)(x_0, t_0)}{|\nabla u(x, 0)(x_0, t_0)|}. \]

We can extend \( \eta \) locally via parallel transport along radial geodesics based at \( x_0 \). By construction,

\[ \nabla \eta(x_0) = 0, \]

\[ |\nabla^2 \eta(x_0)| \leq C(g). \]
By using a cut-off function, we can assume $\eta$ is globally defined and satisfies

$$|\eta| \leq 1,$$

with $|\eta| = 1$ in a neighborhood of $x_0$.

Define

$$H = \eta^\alpha \nabla_\alpha (u - u_0) + \left[ e^{\lambda(u_0 - u + T)} - e^{\lambda T} \right] + \Lambda t(t - 1).$$

Since $|\eta| \leq 1$,

$$H(x, t) \leq \Psi(x, t),$$

and the max of $H$ is attained at $(x_0, t_0)$. Therefore,

$$\mathcal{L}H(x_0, t_0) \leq 0.$$

To compute $\mathcal{L}H(x_0, t_0)$, let $\phi = \eta^\alpha \nabla_\alpha (u - u_0)$. Using (4.17), at $(x_0, t_0)$ we have

$$\phi_t = \eta^\alpha \nabla_\alpha u_t,$$

$$\phi_{tt} = \eta^\alpha \nabla_\alpha u_{tt},$$

$$\nabla_k \phi_t = \eta^\alpha \nabla_k \nabla_\alpha u_t.$$

Also at $(x_0, t_0)$,

$$\nabla_k \phi = \eta^\alpha \nabla_k \nabla_\alpha (u - u_0) = \eta^\alpha \nabla_k \nabla_\alpha u + O(1),$$

$$\nabla_k \nabla_\ell \phi = \nabla_k \nabla_\ell \eta^\alpha \nabla_\alpha (u - u_0) + \eta^\alpha \nabla_k \nabla_\ell \nabla_\alpha (u - u_0)$$

$$= \eta^\alpha \nabla_k \nabla_\ell \nabla_\alpha u + O(1).$$

Therefore, by the formula in (4.2), at $(x_0, t_0)$ we have

$$\mathcal{L} \phi = (1 + \epsilon)^{k-1} u_{tt}^{-1} f \eta^\alpha \nabla_\alpha u_{tt} + u_{tt}^{-k} T_{k-1}(E_u^\epsilon)_{k\ell} \left\{ (1 + \epsilon) u_{tt} \left[ \eta^\alpha \nabla_k \nabla_\ell \nabla_\alpha u ight. ight.$$

$$+ \eta^\alpha \nabla_k \nabla_\alpha u \nabla_\ell u + \eta^\alpha \nabla_k u \nabla_\ell \nabla_\alpha u - (\eta^\alpha \nabla_m \nabla_\alpha v \nabla_m u) g_{\ell \ell} + O(1) g_{k\ell} \bigg] \left. - \eta^\alpha \nabla_k \nabla_\alpha u_t \nabla_\ell u_t - \eta^\alpha \nabla_k u_t \nabla_\ell \nabla_\alpha u_t + \left[ \frac{\eta^\alpha \nabla_\alpha u_{tt}}{u_{tt}} \nabla_k u_t \nabla_\ell u_t \right] \right\}$$

$$\geq \eta^\alpha \mathcal{L} \nabla_\alpha u - C u_{tt}^{2-k} \text{tr} T_{k-1}(E_u^\epsilon).$$

Using the identity (4.5), we conclude

$$\mathcal{L} \phi \geq \langle \nabla f, \eta \rangle - C u_{tt}^{2-k} \text{tr} T_{k-1}(E_u^\epsilon)$$

$$\geq -C f - C u_{tt}^{2-k} \text{tr} T_{k-1}(E_u^\epsilon),$$

where the constants depend on $\max_M |\nabla f|/f$. 


Next, we use Lemma 4.5 to calculate

\[(4.18)\]
\[
\mathcal{L}(u - u_0) = (k + 1)(1 + \epsilon)^{k-1} f + (1 + \epsilon)u_{tt}^{2-k} \left( T_{k-1}(E_u^c), -A + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \right) - (1 + \epsilon)u_{tt}^{2-k} \left( T_{k-1}(E_u^c), \nabla^2 u_0 + \nabla u_0 \otimes \nabla u + \nabla u \otimes \nabla u_0 - (\nabla u_0, \nabla u)g \right) = (k + 1)(1 + \epsilon)^{k-1} f - (1 + \epsilon)u_{tt}^{2-k} \left( T_{k-1}(E_u^c), A + \nabla^2 u_0 \right) + (1 + \epsilon)u_{tt}^{2-k} \left( T_{k-1}(E_u^c), \nabla u \otimes \nabla u - \frac{1}{2} \text{tr} T_{k-1}(E_u^c) |\nabla u|^2 \right) - (1 + \epsilon)u_{tt}^{2-k} \left[ 2 \left( T_{k-1}(E_u^c), \nabla u \otimes \nabla u_0 - \text{tr} T_{k-1}(E_u^c) (\nabla u, \nabla u_0) \right) \right] = (k + 1)(1 + \epsilon)^{k-1} f + (1 + \epsilon)u_{tt}^{2-k} \left[ - (T_{k-1}(E_u^c), A_{u_0}) + (T_{k-1}(E_u^c), \nabla (u - u_0) \otimes \nabla (u - u_0)) \right] - \frac{1}{2} \text{tr} T_{k-1}(E_u^c) |\nabla (u - u_0)|^2 \right].
\]

Taking \( v = e^{\lambda(u_0 - u + \Upsilon)} - e^{\lambda \Upsilon} \) in Lemma 4.2, we also have

\[
\mathcal{L} \left( e^{\lambda(u_0 - u + \Upsilon)} - e^{\lambda \Upsilon} \right) = e^{\lambda(u_0 - u + \Upsilon)} \left\{ (1 + \epsilon)^{k-1} f u_{tt}^{-1} \left[ - \lambda u_{tt} + \lambda^2 u_t^2 \right] + u_{tt}^{2-k} (T_{k-1}(E_u^c), (1 + \epsilon)u_{tt} \left[ \lambda \nabla^2 (u_0 - u) + \lambda^2 \nabla (u_0 - u) \otimes \nabla (u_0 - u) + \lambda \nabla u_0 \otimes \nabla u + \lambda \nabla u \otimes \nabla (u_0 - u) - \lambda (\nabla (u_0 - u), \nabla u)g \right] + \lambda \nabla u_0 \otimes \nabla u + \lambda \nabla u_0 \otimes \nabla (u_0 - u) \right) \right\}
\]

We can estimate the term in braces as follows:

\[
(1 + \epsilon)^{k-1} f \frac{u_t^2}{u_{tt}} + u_{tt}^{2-k} \left( T_{k-1}(E_u^c), (1 + \epsilon) \nabla (u - u_0) \otimes \nabla (u - u_0) \right) + \frac{u_t}{u_{tt}} \nabla (u_0 - u) \otimes \nabla u + \frac{u_t}{u_{tt}} \nabla u_0 \otimes \nabla (u_0 - u) + \frac{u_t^2}{u_{tt}} \nabla u_0 \otimes \nabla u_0 \right) \right] \geq (1 + \epsilon)^{k-1} f \frac{u_t^2}{u_{tt}} + u_{tt}^{2-k} \left( T_{k-1}(E_u^c), \frac{1 + \epsilon}{2} \nabla (u - u_0) \otimes \nabla (u - u_0) - \frac{u_t^2}{u_{tt}} \nabla u_0 \otimes \nabla u_0 \right) \]

Using Lemma 2.5 and the regularized equation, the final (negative) term above can be rewritten:

\[ u_{tt}^{2-k} \{ T_{k-1} \left( E_u^c \right), - \frac{u_t^2}{u_{tt}} \nabla u_t \otimes \nabla v_t \} = -u_{tt}^{-k} u_t^2 \{ T_{k-1} \left( (1 + \epsilon) u_{tt} A_u - \nabla u_t \otimes \nabla v_t \right), \nabla u_t \otimes \nabla v_t \} \]

\[ = -u_{tt}^{-k} u_t^2 \{ T_{k-1} \left( (1 + \epsilon) u_{tt} A_u \right), \nabla u_t \otimes \nabla v_t \} \]

\[ = -(1 + \epsilon) u_{tt}^{-k} u_t^2 \{ T_{k-1} \left( A_u \right), \nabla u_t \otimes \nabla v_t \} \]

\[ = -(1 + \epsilon) u_{tt}^{k-1} u_t^2 \left\{ (1 + \epsilon) u_{tt} \sigma_k(A_u) - f \right\} \]

\[ = -(1 + \epsilon) u_{tt}^{k} \sigma_k(A_u) + (1 + \epsilon)^{-1} f \frac{u_t^2}{u_{tt}}. \]

Therefore,

\[
\begin{align*}
\mathcal{L} \left( e^{\lambda(\psi_0 - \psi_T)} - e^{\lambda T} \right) & \geq -\lambda e^{\lambda(\psi_0 - \psi_T)} \mathcal{L}(u - \psi_0) + \lambda^2 e^{\lambda(\psi_0 - \psi_T)} \left\{ 2(1 + \epsilon)^{k-1} f \frac{u_t^2}{u_{tt}} \right. \\
& \quad - \left. (1 + \epsilon)^k u_t^2 \sigma_k(A_u) + u_{tt}^{2-k} \left\{ T_{k-1}(E_u^c), \frac{(1 + \epsilon)}{2} \nabla(u - \psi_0) \otimes \nabla(u - \psi_0) \right\} \right). 
\end{align*}
\]

(4.19)

Also, by (4.18),

\[
\begin{align*}
-\lambda \mathcal{L}(u - \psi_0) &= -\lambda (k + 1)(1 + \epsilon)^{-1} f + \lambda (1 + \epsilon) u_{tt}^{2-k} \left\{ T_{k-1}(E_u^c), A_u \right\} \\
& \quad + u_{tt}^{2-k} \left\{ T_{k-1}(E_u^c), -\lambda(1 + \epsilon) \nabla(u - \psi_0) \otimes \nabla(u - \psi_0) \right\} + \frac{1}{2}(1 + \epsilon) \lambda u_{tt}^{2-k} \text{tr} T_{k-1}(E_u^c) |\nabla(u - \psi_0)|^2. 
\end{align*}
\]

(4.20)

Combining (4.19) and (4.20), we get

\[
\begin{align*}
\mathcal{L} \left( e^{\lambda(\psi_0 - \psi_T)} - e^{\lambda T} \right) & \geq e^{\lambda(\psi_0 - \psi_T)} \left\{ -\lambda (k + 1)(1 + \epsilon)^{-1} f + 2\lambda^2 (1 + \epsilon)^{-1} f \frac{u_t^2}{u_{tt}} - \lambda^2 (1 + \epsilon)^k u_t^2 \sigma_k(A_u) \\
& \quad + u_{tt}^{2-k} \left\{ T_{k-1}(E_u^c), (1 + \epsilon) \frac{1}{2} \lambda^2 - \lambda \right\} \nabla(u - \psi_0) \otimes \nabla(u - \psi_0) \right) + \frac{1}{2}(1 + \epsilon) \lambda u_{tt}^{2-k} \text{tr} T_{k-1}(E_u^c) |\nabla(u - \psi_0)|^2 + \lambda(1 + \epsilon) u_{tt}^{2-k} \left\{ T_{k-1}(E_u^c), A_u \right\} \right). 
\end{align*}
\]

(4.21)

Next, using Lemma 4.4, we have

\[
\mathcal{L}(\Lambda t(1 - t)) = 2\Lambda (1 + \epsilon)^k \sigma_k(A_u). 
\]

Combining the above, we conclude that at an interior maximum of \( H \),

\[
\mathcal{L}H \geq -Cf - Cu_{tt}^{2-k} \text{tr} T_{k-1}(E_u^c) + 2\Lambda (1 + \epsilon)^k \sigma_k(A_u) \\
+ e^{\lambda(\psi_0 - \psi_T)} \left\{ -\lambda(1 + \epsilon)^{-1} f + 2\lambda^2 (1 + \epsilon)^{-1} f \frac{u_t^2}{u_{tt}} - \lambda^2 (1 + \epsilon)^k u_t^2 \sigma_k(A_u) \\
+ u_{tt}^{2-k} \left\{ T_{k-1}(E_u^c), (1 + \epsilon) \frac{1}{2} \lambda^2 - \lambda \right\} \nabla(u - \psi_0) \otimes \nabla(u - \psi_0) \right) + \frac{1}{2}(1 + \epsilon) \lambda u_{tt}^{2-k} \text{tr} T_{k-1}(E_u^c) |\nabla(u - \psi_0)|^2 + \lambda(1 + \epsilon) u_{tt}^{2-k} \left\{ T_{k-1}(E_u^c), A_u \right\} \right). 
\]

Now note that since the cone \( \Gamma_k^c \) is open and \( M \) is compact there exists \( \delta > 0 \) depending only on \( u_0 \) so that \( A_u_0 - \delta g \in \Gamma_k^c \). It follows from Lemma 2.3 that

\[
\delta \text{ tr} T_{k-1}(E_u^c) = \Sigma(E_u^c, \ldots, E_u^c, \delta g) < \Sigma(E_u^c, \ldots, E_u^c, A_{u_0}) = \left( T_{k-1}(E_u^c), A_{u_0} \right). 
\]

Therefore, if \( \lambda >> 2 \), we have

\[
\begin{align*}
\mathcal{L}H & \geq \left\{ -C - \lambda(1 + \epsilon)^{-1} e^{\lambda(\psi_0 - \psi_T)} \right\} f \\
& \quad + \left\{ 2\Lambda (1 + \epsilon)^k - \lambda^2 (1 + \epsilon)^k u_t^2 e^{\lambda(\psi_0 - \psi_T)} \right\} \sigma_k(A_u) + \left\{ -C + \lambda(1 + \epsilon) \delta \right\} u_{tt}^{2-k} \text{tr} T_{k-1}(E_u^c). 
\end{align*}
\]
Observe that by choosing \( \lambda = \lambda(\delta) \) large enough, we can assume the last term in (4.21) is bounded below by
\[
\frac{\lambda}{2} \delta u_{tt}^{2-k} \tr T_{k-1}(E_u^\epsilon).
\]
(4.22)

By the Newton-Maclaurin inequality,
\[
u_{tt}^{2-k} \tr T_{k-1}(E_u^\epsilon) = (k-1)u_{tt}^{2-k} \sigma_{k-1}(E_u^\epsilon)
\geq (k-1)u_{tt}^{2-k} \sigma_k(E_u^\epsilon)^{\frac{k-1}{k}}
= (k-1)\frac{k-1}{k} u_{tt}^{\frac{k}{k}}
\geq C f u_{tt}^{\frac{1}{k}}.
\]

Combining this with (4.22) and substituting into (4.21), we get
\[
\mathcal{L}H \geq \left\{ -C - \lambda(k + 1)(1 + \epsilon)k^{-1}e^{\lambda(u_0 - u + \Upsilon)} + C\lambda \delta u_{tt}^{\frac{1}{k}} \right\} f
+ \{2\lambda(1 + \epsilon)^k - \lambda^2(1 + \epsilon)^k u_{tt}^{2} e^{\lambda(u_0 - u + \Upsilon)}\} \sigma_k(A_u).
\]

Let us fix the constant \( \Upsilon \) so that
\[
0 \leq u_0 - u + \Upsilon \leq C,
\]
then
\[
\mathcal{L}H \geq \left\{ -C - C\lambda(k + 1) + C\lambda \delta u_{tt}^{\frac{1}{k}} \right\} f + \{2\lambda(1 + \epsilon)^k - C\lambda^2 u_{tt}^{2}\} \sigma_k(A_u).
\]

Next, we assume \( \Lambda = \Lambda(\lambda, \max u_t^2) \) is chosen large enough so that the coefficient of the second term above is
\[
2\lambda(1 + \epsilon)^k - C\lambda^2 u_{tt}^{2} \geq \frac{1}{2} \lambda^2.
\]

By the regularized equation,
\[
\sigma_k(A_u) \geq \frac{f}{(1 + \epsilon)u_{tt}}.
\]

Therefore,
\[
\mathcal{L}H \geq \left\{ -C - C\lambda(k + 1) + C\lambda \delta u_{tt}^{\frac{1}{k}} + \frac{1}{2(1 + \epsilon)} \lambda^2 u_{tt}^{-1} \right\} f.
\]

If \( u_{tt} > C(\delta) \) is large then the left-hand side is positive, which would be a contradiction at an interior maximum. On the other hand, if \( u_{tt} \) is small then as long as \( \lambda \) is chosen large enough, the last term in the braces will dominate and once again we conclude \( \mathcal{L}H > 0 \). It follows that \( H \) attains its maximum on the boundary, as claimed.

\[ \square \]

4.5. Existence of approximate and regularizable geodesics. In this subsection we use the a priori estimates of the previous subsections to establish the existence of weak geodesics.

**Theorem 4.18.** Given \( u_0, u_1 \in \Gamma^+_t \), there exists \( f \in C^\infty(M \times [0, 1]) \) with \( f > 0 \) and a smooth solution \( u(x, t, s, \epsilon) : M \times [0, 1] \times [0, 1] \times (0, \epsilon_0] \to \mathbb{R} \) of \( G^\epsilon_{sf}(u_\epsilon) = 0 \) such that

1. For each \( \epsilon \in (0, \epsilon_0] \), \( u_\epsilon = u(\cdot, \cdot, \cdot, \epsilon) \) satisfies

\[
u_\epsilon(x, 0, s) = u_0(x), \quad u_\epsilon(x, 1, s) = u_1(x).
\]

2. There is a constant \( C > 0 \), independent of \( \epsilon \), such that

\[
|u_\epsilon| + |\nabla u_\epsilon| + |(u_\epsilon)_t| + \epsilon \left( |\nabla^2 u_\epsilon| + |(u_\epsilon)_t| + |(u_\epsilon)_{tt}| \right) \leq C.
\]
Proof. As the argument follows standard lines we provide only a sketch. Fix some $0 < \epsilon_0 < 1$, then choose an arbitrary $0 < \epsilon < \epsilon_0$. First we observe that it follows from (39 Proposition 3) that the path $u_t := tu_1 + (1-t)u_0$ lies in $\Gamma_k^\epsilon$. Moreover, there exists some constant $\Lambda$ for which $w_t := u_t + \Lambda t(1-t)$ satisfies $E_u^\epsilon \in \Gamma_k^\epsilon$. Let $f := \Phi_{\epsilon}(w)$, and set

$$I = \{ s \in [0,1] : \exists u \in C^{4,\alpha} \cap \Gamma_k^\epsilon, \ u \text{ solves } (\ast_{\epsilon,s,f}) \}.$$

By construction, $1 \in I$.

To verify that $I$ is open, it suffices to study the linearized equation; i.e., given $\psi \in C^\infty(M \times [0,1])$, we need to solve for some $s \in I$ then equation

$$\mathcal{L}_{u_s(\cdot,\cdot,s)} \varphi = \psi$$

with $\varphi$ satisfying Dirichlet boundary conditions. The solvability of this linear problem follows from [17], Theorem 6.13.

We claim that $I$ is closed: let \{ $u_i = u_{s_i}$ \} be a sequence of admissible solutions with $s_i \geq s_0$. The preceding \textit{a priori} estimates imply there is a constant $C$ (independent of $\epsilon$) such that

$$|u_i| + |\nabla u_i| + |(u_i)_t| + \epsilon \left( |(u_i)_{tt}| + |(\nabla u_i)_t| + |\nabla^2 u_i| \right) \leq C.$$

To obtain higher order regularity, we need to verify the concavity of the operator. Observe that the equation can be rewritten as

$$\sigma_k^\frac{1}{k} \left( \frac{1}{1-k} E_u^\epsilon \right) = f_k^\frac{1}{k}.$$

Since $\sigma_k^\frac{1}{k}$ is a concave operator, the equation is convex, and so by Evans-Krylov [15] 22 we conclude there is a constant $C = C(\epsilon, f)$ such that

$$\| u_i \|_{C^{2,\alpha}} \leq C.$$

Applying the Schauder estimates we obtain bounds on derivatives of all orders, and it follows that the set $I$ is closed. Since $I$ is open, closed, and non-empty, it follows that $I = [0,1]$. The theorem follows.

**Definition 4.19.** Given $u_0,u_1 \in \Gamma_k^\epsilon$, we say a one parameter family of $C^{1,1}$ functions $u_s(x,t) : M \times [0,1] \to \mathbb{R}$ is an $\epsilon$-geodesic from $u_0$ to $u_1$ if

$$u_s(x,0) = u_0(x), \quad u_s(x,1,s) = u_1(x), \quad G_{s_0}(u_s) = 0.$$

We furthermore will say that it is a regularizable $\epsilon$-geodesic if there exists $f_0 \in C^\infty(M \times [0,1])$ with $f_0 > 0$ and a smooth function $u(x,t,s) : M \times [0,1] \times [0,1] \times \mathbb{R}$ with the following properties:

(i) For each $s \in [0,1]$, $u(\cdot,\cdot,s)$ satisfies

$$u(x,0,s) = u_0(x), \quad u(x,1,s) = u_1(x), \quad G_{s_0}(u) = 0.$$

(ii) There is a constant $C > 0$, independent of $\epsilon$, such that

$$|u_s| + |\nabla u_s| + |(u_s)_t| + \epsilon \left( |\nabla^2 u_s| + |\nabla (u_s)_t| + |(u_s)_{tt}| \right) \leq C.$$

(iii) One has that $u(x,t,s) \to u(x,t)$ in the weak $C^{1,1}$ topology as $s \to 0$.

**Definition 4.20.** Given $u_0,u_1 \in \Gamma_k^\epsilon$, we say a one parameter family of $C^1$ functions $u(x,t)$ is a regularizable geodesic from $u_0$ to $u_1$ if there exists $f_0 \in C^\infty(M \times [0,1])$ with $f_0 > 0$ and a smooth function $u(x,t,s,\epsilon) : M \times [0,1] \times [0,1] \times [0,\epsilon_0] \to \mathbb{R}$ with the following properties:

(i) For each $\epsilon \in [0,\epsilon_0)$ $u_{\epsilon}(x,0) = u(\cdot,\cdot,\cdot,\epsilon)$ satisfies

$$u_{\epsilon}(x,0,s) = u_0(x), \quad u_{\epsilon}(x,1,s) = u_1(x), \quad G_{s_0}(u_{\epsilon}) = 0.$$
(ii) There is a constant $C > 0$, independent of $\epsilon$, such that

$$|u_\epsilon| + |\nabla u_\epsilon| + |(u_\epsilon)_t| + \epsilon \left\{ |\nabla^2 u_\epsilon| + |\nabla(u_\epsilon)_t| + |(u_\epsilon)_{tt}| \right\} \leq C.$$ 

(iii) For each $0 < \alpha < 1$, $u_\epsilon \to u$ in $C^{0,\alpha}$ as $\epsilon, s \to 0$.

We can now show existence and uniqueness of a regularizable geodesic connecting any two points in $\Gamma^+$. The key issue for uniqueness is a comparison lemma.

**Lemma 4.21.** Suppose $u, \tilde{u} \in C^\infty$ are admissible and satisfy

$$G^\epsilon_{f_1}(u) = 0,$$

$$G^\epsilon_{f_2}(\tilde{u}) = 0,$$

where $f_1 \leq f_2$. Assume further that on the boundary,

$$u(x,0) = \tilde{u}(x,0),$$

$$u(x,1) = \tilde{u}(x,1).$$

Then on $M \times [0,1]$,

$$u(x,t) \geq \tilde{u}(x,t).$$

We remark here also that the Lemma 4.21 can be used to exhibit uniqueness for solutions of the equation $G^\epsilon_0(u) = 0$.

**Corollary 4.22.** Given $u_0, u_1 \in \Gamma^+_k$, there exists a unique $\epsilon$-geodesic from $u_0$ to $u_1$.

**Proof.** Let $u(x,t,\epsilon)$ and $f$ be the data guaranteed by Theorem 4.18. Due to the a priori estimates, by Arzela-Ascoli there exists a $C^{1,1}$ limit as $s \to 0$. By definition this is an $\epsilon$-geodesic. Now suppose $\tilde{u}$ is another regularizable geodesic connecting $u_0$ to $u_1$, with regularization $\tilde{u}(x,t,\epsilon)$ and auxiliary function $\tilde{f}$. Fixing some $\delta > 0$, for sufficiently small $\epsilon > 0$ Lemma 4.21 implies that $u(x,t,\epsilon) \geq \tilde{u}(x,t,\delta)$. Since the convergence is in $C^{0,\alpha}$, sending $\epsilon \to 0$ yields $u(x,t) \geq \tilde{u}(x,t,\delta)$. We can now send $\delta \to 0$ to obtain $u(x,t) \geq \tilde{u}(x,t)$. Since the roles of $u$ and $\tilde{u}$ are interchangeable in that argument, it follows that $u(x,t) = \tilde{u}(x,t)$. \hfill \Box

**Corollary 4.23.** Given $u_0, u_1 \in \Gamma^+_k$, there exists a unique regularizable geodesic from $u_0$ to $u_1$.

**Proof.** Let $u(x,t,\epsilon)$ and $f$ be the data guaranteed by Theorem 4.18. Due to the a priori estimates, by Arzela-Ascoli there exists a $C^{0,\alpha}$ limit as both $\epsilon \to 0$ and $s \to 0$. By definition this is a regularizable geodesic. Now suppose $\tilde{u}$ is another regularizable geodesic connecting $u_0$ to $u_1$, with regularization $\tilde{u}(x,t,\epsilon)$ and auxiliary function $\tilde{f}$. Fixing some $\delta > 0$, for sufficiently small $\epsilon > 0$ Lemma 4.21 implies that $u(x,t,\epsilon) \geq \tilde{u}(x,t,\delta)$. Since the convergence is in $C^{0,\alpha}$, sending $\epsilon \to 0$ yields $u(x,t) \geq \tilde{u}(x,t,\delta)$. We can now send $\delta \to 0$ to obtain $u(x,t) \geq \tilde{u}(x,t)$. Since the roles of $u$ and $\tilde{u}$ are interchangeable in that argument, it follows that $u(x,t) = \tilde{u}(x,t)$. \hfill \Box

5. **Smoothing via Guan-Wang flow**

In this section we develop a sharper picture (Theorem 5.12) of the short-time smoothing properties of a parabolic flow introduced by Guan-Wang in [19]. This is used in the proof of Theorem 1.5 to smooth the approximate geodesics so that we can take strong limits to obtain a curve of critical points for $F$ connecting any two given critical points.

In first subsection we will derive a series of formulas for the evolution of various quantities. Since we will be quoting some of the formulas from the previous section, we will state these formulas for general dimensions. In the second subsection, where we derive some short-time estimates, we will specialize to the case $n = 4$ and $k = 2$. 
First, we recall the definition of the flow introduced in [19]:

\begin{equation}
\frac{\partial}{\partial t} u = \log \sigma_k(g_u^{-1}A_u) - V_u^{-1} \int_M \log \sigma_k(g_u^{-1}A_u) dV_g.
\end{equation}

(5.1)

For technical simplicity we will instead study an unnormalized flow

\begin{equation}
\frac{\partial}{\partial t} u = \log \sigma_k(g_u^{-1}A_u) = \log \sigma_k(A_u) + 2k u.
\end{equation}

(5.2)

As we will be able to control the size of \(u\) along this flow, the renormalizing term will only change \(u\) by a controlled constant, and have no effect on the estimates.

5.1. Evolution equations. We remark that when the dimension \(n \geq 4\), Guan-Wang assumed the manifold was locally conformally flat. For the evolutionary formulas we are interested in this assumption will not be necessary.

**Definition 5.1.** Given \(u\) an admissible solution to (5.2), define

\[
Lf = \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), \nabla^2 f + \nabla u \otimes \nabla f + \nabla f \otimes \nabla u - (\nabla u, \nabla f) g \right),
\]

\[
H = \frac{\partial}{\partial t} - L.
\]

where the derivatives and inner products are with respect to \(g\) (the fixed background metric).

**Lemma 5.2.** Let \(u\) be a solution to (5.2). Then

\[
Hu = \log \sigma_k(A_u) - k + \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), A - \nabla u \otimes \nabla u + \frac{1}{2} |\nabla u|^2 g \right) + 2k u.
\]

**Proof.** We directly compute

\[
Lu = \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), \nabla^2 u + 2 \nabla u \otimes \nabla u - |\nabla u|^2 g \right)
\]

\[
= \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), A_u - A + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \right)
\]

\[
= k + \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), - A + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \right).
\]

Combining this with (5.2) yields the result. \(\square\)

**Lemma 5.3.** Let \(u\) be a solution to (5.2) and \(\lambda \in \mathbb{R}\). Then

\[
He^{\lambda u} = \lambda e^{\lambda u} \left[ \log \sigma_k(A_u) + 2k u - k + \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), A - (1 + \lambda) \nabla u \otimes \nabla u + \frac{1}{2} |\nabla u|^2 g \right) \right].
\]

**Proof.** Note

\[
\frac{\partial}{\partial t} (e^{\lambda u}) = \lambda e^{\lambda u} \left( \log \sigma_k(A_u) + 2k u \right).
\]

Also,

\[
Le^{\lambda u} = \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), \lambda e^{\lambda u} \nabla^2 u + \lambda^2 e^{\lambda u} \nabla u \otimes \nabla u + 2 \lambda e^{\lambda u} \nabla u \otimes \nabla u - \lambda e^{\lambda u} |\nabla u|^2 g \right)
\]

\[
= \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), \lambda e^{\lambda u} \left[ A_u - A \right. \nabla u \otimes \nabla u + \frac{1}{2} |\nabla u|^2 g \left. \right] + \lambda (\lambda + 2) e^{\lambda u} \nabla u \otimes \nabla u - \lambda e^{\lambda u} |\nabla u|^2 g \right)
\]

\[
= \lambda e^{\lambda u} \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), A_u - A + (\lambda + 1) \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \right)
\]

\[
= \lambda e^{\lambda u} \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), - A + (\lambda + 1) \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \right) + \lambda k e^{\lambda u}.
\]
Therefore,
\[ H e^{\lambda u} = \frac{\partial}{\partial t}(e^{\lambda u}) - L e^{\lambda u} \]
\[ = \lambda e^{\lambda u} \left[ \log \sigma_k(A_u) + 2ku - k + \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), A - (1 + \lambda) \nabla u \otimes \nabla u + \frac{1}{2} |\nabla u|^2 g \right) \right]. \]

**Lemma 5.4.** Given \( u \) a solution to (5.2), one has
\[ H |\nabla u|^2 = 2 \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \left\{ - \nabla_i \nabla_p u \nabla_i \nabla_q u + O(|\nabla u|^2 + 1) \right\} + 4k |\nabla u|^2. \]

**Proof.** We compute
\[ \frac{\partial}{\partial t} \nabla_i u = \nabla_i \log \sigma_k(A_u) + 2k \nabla_i u \]
\[ = \sigma_k(A_u)^{-1} \left( T_{k-1}(A_u), \nabla_i A_u \right) + 2k \nabla_i u \]
\[ = \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \left\{ \nabla_i A_p q + \nabla_i \nabla_p \nabla_q u + 2 \nabla_i \nabla_p u \nabla_q u - \nabla_i \nabla_j u \nabla_j u g_{pq} \right\} + 2k \nabla_i u \]
\[ = \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \left\{ \nabla_p \nabla_q \nabla_i u + 2 \nabla_i \nabla_p u \nabla_q u - \nabla_i \nabla_j u \nabla_j u g_{pq} + (\nabla A + Rm * \nabla u)_{ipq} \right\} + 2k \nabla_i u, \]

hence
\[ \frac{\partial}{\partial t} |\nabla u|^2 = 2 \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \left\{ \nabla_p \nabla_q \nabla_i u \nabla_i u + 2 \nabla_i \nabla_p u \nabla_q u \nabla_i u - \nabla_i \nabla_j u \nabla_j u \nabla_i u g_{pq} \right\} + \left\{ (\nabla A + Rm * \nabla u) \ast \nabla u \right\}_{pq} + 4k |\nabla u|^2. \]

Also,
\[ L |\nabla u|^2 = 2 \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \left\{ \nabla_p \nabla_i u \nabla_q \nabla_i u + \nabla_p \nabla_q \nabla_i u \nabla_i u + 2 \nabla_i \nabla_p u \nabla_q u \nabla_i u - \nabla_i \nabla_j u \nabla_j u \nabla_i u g_{pq} \right\}. \]

It follows that
\[ \frac{\partial}{\partial t} |\nabla u|^2 = L |\nabla u|^2 + 2 \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \left\{ - \nabla_p \nabla_i u \nabla_q \nabla_i u + \left\{ (\nabla A + Rm * \nabla u) \ast \nabla u \right\}_{pq} + 4k |\nabla u|^2, \right\}
\]

which implies the result. \( \square \)

**Corollary 5.5.** Given \( u \) a solution to (5.2), one has
\[ (5.3) \quad H(e^{-4kt} |\nabla u|^2) = 2 e^{-4kt} \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \left\{ - \nabla_i \nabla_p u \nabla_i \nabla_q u + O(|\nabla u|^2 + 1) \right\}. \]

For the following lemma, for an \( n \times n \) symmetric matrix \( r = r_{ij} \) we denote
\[ \mathcal{F}(r) = \log \sigma_k(r), \]
and derivatives of \( \mathcal{F} \) with respect to the entries of \( r \) by
\[ \frac{\partial}{\partial r_{pq}} \mathcal{F}(r) = \mathcal{F}(r)^{pq}, \]
\[ \frac{\partial^2}{\partial r_{pq} \partial r_{rs}} \mathcal{F}(r) = \mathcal{F}(r)^{pq,rs}. \]

**Lemma 5.6.** Given \( u \) a solution to (5.2), one has
\[ H \Delta u = \mathcal{F}^{pq,rs} \nabla_i (A_u)^{pq} \nabla_i (A_u)^{rs} \]
\[ + \sigma_k(A_u)^{-1} \left\{ T_{k-1}(A_u)_{ij}, 2 \nabla_i \nabla_p u \nabla_j \nabla_p u - |\nabla^2 u|^2 g_{ij} + O(|\nabla^2 u| + |\nabla u|^2 + 1) \right\}. \]
Proof. We compute
\[ \Delta \log \sigma_k(A_u) = \nabla_t [\mathcal{F}^{pq} \nabla_i (A_u)_{pq}] + \mathcal{F}^{pq,rs} \nabla_i (A_u)_{pq} \nabla_i (A_u)_{rs} + \mathcal{F}^{pq} \Delta (A_u)_{pq}. \]
Combining this with our prior calculation of \( \Delta A_u \) (4.13) yields
\[ \frac{\partial}{\partial t} \Delta u = \Delta \log \sigma_k(A_u) + 2k \Delta u \]
\[ = \mathcal{F}^{pq,rs} \nabla_i (A_u)_{pq} \nabla_i (A_u)_{rs} + \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} (\Delta A_u)_{pq} + 2k \Delta u \]
\[ = \mathcal{F}^{pq,rs} \nabla_i (A_u)_{pq} \nabla_i (A_u)_{rs} + \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \{ \nabla_p \nabla_q (\Delta u) + \nabla_p \Delta u \nabla_q u + \nabla_p u \nabla_q \Delta u \}
+ 2 \nabla_p \nabla_q \nabla_r u - |\nabla^2 u|^2 g_{pq} - (\nabla u, \nabla \Delta u) g_{pq} + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1) \]
\[ = L(\Delta u) + \mathcal{F}^{pq,rs} \nabla_i (A_u)_{pq} \nabla_i (A_u)_{rs}
+ \sigma_k(A_u)^{-1} T_{k-1}(A_u)^{pq} \{ 2 \nabla_p \nabla_q \nabla_r u - |\nabla^2 u|^2 g_{pq} + \mathcal{O}(|\nabla^2 u| + |\nabla u|^2 + 1) \}, \]
and the result follows. \( \square \)

5.2. Estimates. In this section we specialize to the case \( n = 4 \) and \( k = 2 \), and use the evolutionary formulas from the preceding subsection to derive some short-time smoothing estimates.

Lemma 5.7. Given \( u \) a solution to (5.1), one has
\[ \frac{d}{dt} F[u] \leq 0. \]

Proof. This is immediate from the flow equation (5.1) and the formula (3.13). \( \square \)

Proposition 5.8. Let \( u \) be a solution to (5.1) with initial value \( u(\cdot, 0) = u_0 \), where \( u_0 \) is admissible. Then there are constants \( C_1 = C_1(g), \epsilon = \epsilon(|u_0|_{C^0}) \), such that \( u \) exists for all \( 0 \leq t \leq \epsilon \), and
\[ |u|_{C^0} \leq C_1 (1 + |u_0|_{C^0}). \]
for all \( 0 \leq t \leq \epsilon \).

Proof. At a maximum for \( u \), one has \( A_u \leq A \), and hence \( \sigma_2(A_u) \leq \sigma_2(A) < C \). By (5.2),
\[ \frac{d}{dt} \max u \leq C + 4 \max u. \]
Integrating this inequality we get an upper bound for \( u \). Applying a similar argument at a minimum of \( u \), we obtain a lower bound. \( \square \)

Proposition 5.9. Given \( u \) as in the previous proposition, there exists constants \( C_1 \) and \( \epsilon \) depending on \( |u_0|_{C^0} \) such that for all \( 0 \leq t \leq \epsilon \leq 1 \), one has
\[ |\nabla u|_{C^0} \leq C_1 (|u_0|, |\nabla u_0|). \]

Proof. Let
\[ \Phi = e^{-st} |\nabla u|^2 + \Lambda e^{-2u} - \mu, \]
where \( \Lambda, \mu > 0 \) will be specified later. Combining Corollary 5.5 and Lemma 5.3, and using the fact that at a maximum of \( \Phi \) we have \( H \Phi \geq 0 \), it follows
\[ 0 \leq H \Phi = 2 \sigma_2(A_u)^{-1} T_1(A_u)^{pq} \{ -e^{-st} \nabla_i \nabla_p u \nabla_i \nabla_q u + e^{-st} \mathcal{O}(1 + |\nabla u|^2) - \Lambda e^{-2u} [\nabla_p u \nabla_q u + \frac{1}{2} |\nabla u|^2 g_{pq}] \}
- 2 \Lambda e^{-2u} [\log \sigma_2(A_u) + 4u - 2 + \sigma_2(A_u)^{-1} (T_1(A_u), A)] - \mu \]
\[ = I_1 + I_2 - \mu. \]
We can estimate the terms in braces in \( I_1 \) by
\[ -e^{-st} \nabla_i \nabla_p u \nabla_i \nabla_q u + e^{-st} \mathcal{O}(1 + |\nabla u|^2) - \Lambda e^{-2u} [\nabla_p u \nabla_q u + \frac{1}{2} |\nabla u|^2 g_{pq}] \leq \{ C + (C - \frac{\Lambda}{2} e^{-2u}) |\nabla u|^2 \} g_{pq}. \]
By Proposition 5.8 for $0 \leq t \leq \epsilon \leq 1$ we have a uniform bound on $|u|$ depending only on the initial data, hence if $\Lambda > 1$ is chosen large enough,

$$C + (C - \frac{\Lambda}{2}e^{-2u})|\nabla u|^2 \leq C - |\nabla u|^2.$$ 

If $|\nabla u|$ remains uniformly bounded we have nothing to prove, so we may assume that at the maximum of $\Phi$ the gradient of $u$ is large, hence at a maximum of $\Phi$ we have

$$I_1 \leq 0.$$ 

To estimate $I_2$, we first consider the case where $\sigma_2(A_u) \geq 1$. Then $\log \sigma_2(A_u) \geq 0$ and the remaining terms in brackets are either bounded or non-negative, hence

$$I_2 - \mu \leq C(\Lambda, \max |u|) - \mu \leq 0,$$

if $\mu$ is chosen large enough. On the other hand, using Lemma 2.6 we see that

$$\sigma_2(A_u)^{-1}(T_1(A_u), A) \geq \sigma_2(A_u)^{-1} \sigma_2(A_u)^{-1} \sigma_2(A_u)^{-1} \sigma_2(A_u)^{-1} = \frac{\sigma_2(A)^2}{\sigma_2(A_u)^{2n}}.$$ 

It follows there is a small constant $\delta = \delta(\sigma_2(A))$ such that if $0 < \sigma_2(A_u) \leq \delta$, then

$$\log \sigma_2(A_u) + \sigma_2(A_u)^{-1}(T_1(A_u), A) \geq 0.$$ 

Then arguing as we did in the case where $\sigma_2(A_u) \geq 1$, we can choose $\mu$ large enough to achieve (5.3) again. Finally, in the intermediate range $\delta \leq \sigma_2(A_u) \leq 1$, all the terms in the brackets in $I_2$ are bounded are non-positive, and we again conclude that (5.4) holds once $\mu$ is chosen large enough. It follows that $H\Phi \leq 0$, and the result follows from the maximum principle.

**Proposition 5.10.** Suppose $u$ is a solution to (5.2) with $n = 4$ on $[0, T], T \leq 1$, such that

$$\sup_{M \times [0, T]} |u| \leq N.$$ 

There exists a constant $C = C(\Lambda)$ such that for all $t \in [0, T]$, one has

$$t |\log \sigma_2(A_u)| \leq C.$$ 

**Proof.** First, note that

$$\frac{\partial}{\partial t} \log \sigma_2(A_u) = \sigma_2(A_u)^{-1} \langle T_1(A_u), \frac{\partial}{\partial t} A_u \rangle$$

$$= \sigma_2(A_u)^{-1} \left\{ T_1(A_u), \nabla^2 \log \sigma_2(A_u) + \nabla u \otimes \nabla \log \sigma_2(A_u) + \nabla \log \sigma_2(A_u) \otimes \nabla u - \langle \nabla u, \nabla \log \sigma_2(A_u) \rangle g + 4 \nabla^2 u + 8 \nabla u \otimes \nabla u - 4 |\nabla u|^2 g \right\}$$

$$= L(\log \sigma_2(A_u)) + 4 \sigma_2(A_u)^{-1} \langle T_1(A_u), \nabla^2 u + 2 \nabla u \otimes \nabla u - |\nabla u|^2 g \rangle$$

$$= L(\log \sigma_2(A_u)) + 4 \sigma_2(A_u)^{-1} \langle T_1(A_u), A_u - A + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \rangle$$

$$= L(\log \sigma_2(A_u)) + 8 + 4 \sigma_2(A_u)^{-1} \langle T_1(A_u), -A + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \rangle,$$

hence

$$H(\log \sigma_2(A_u)) = 8 + 4 \sigma_2(A_u)^{-1} \langle T_1(A_u), -A + \nabla u \otimes \nabla u - \frac{1}{2} |\nabla u|^2 g \rangle.$$ 

Set

$$\Phi := t \log \sigma_2(A_u) + \Lambda e^{-2u} - \mu t.$$
We will show that by choosing \( \Lambda, \mu \gg 1 \) sufficiently large (depending on \( N \)), \( H\Phi \leq 0 \). This will give an upper bound on \( \Phi \) depending only on the initial \( C^0 \)-norm of \( u \).

To begin, we combine (5.6) with (5.3) to get

\[
H\Phi = -\mu + 8t + 4\Lambda(1 - 2u)e^{-2u} + (1 - 2\Lambda e^{-2u}) \log \sigma_2(A_u)
\]

(5.7)

\[
+ \sigma_2(A_u)^{-1}(T_1(A_u), - (4t + 2\Lambda e^{-2u}) A + (4t - 2\Lambda e^{-2u}) \nabla u \otimes \nabla u - (2t + \Lambda e^{-2u})|\nabla u|^2 g).
\]

By choosing \( \Lambda \) large enough (depending on the constant \( N \) in (5.5)) we may assume the coefficient of the log-term

\[
1 - 2\Lambda e^{-2u} \leq -1.
\]

(5.8)

For \( t \) small (depending on \( N \) and \( \Lambda \)) the coefficients of the gradient terms in (5.7) are also non-positive, so we have

\[
H\Phi \leq -\mu + 8t + 4\Lambda(1 - 2u)e^{-2u} + (1 - 2\Lambda e^{-2u}) \log \sigma_2(A_u)
\]

(5.9)

\[
- (4t + 2\Lambda e^{-2u}) \sigma_2(A_u)^{-1}(T_1(A_u), A).
\]

If \( \mu \gg 1 \) is chosen large enough, the first three terms on the RHS of (5.7) can be bounded above by \(-\mu/2\), and we conclude

\[
H\Phi \leq -\mu/2 + (1 - 2\Lambda e^{-2u}) \log \sigma_2(A_u) - (4t + 2\Lambda e^{-2u}) \sigma_2(A_u)^{-1}(T_1(A_u), A).
\]

(5.10)

By Lemma 2.6 we have

\[
\sigma_2(A_u)^{-1}(T_1(A_u), A) \geq \sigma_2(A_u)^{-1}\left[4\sigma_2(A_u)^\frac{1}{2}\sigma_2(A)^\frac{1}{2}\right] \geq \delta \sigma_2(A_u)^{-\frac{1}{2}} > 0,
\]

hence

\[
-(4t + 2\Lambda e^{-2u}) \sigma_2(A_u)^{-1}(T_1(A_u), A) \leq -C_1 \sigma_2(A_u)^{-\frac{1}{2}}.
\]

(5.11)

If \( \sigma_2(A_u) \geq 1 \), it follows from (5.8), (5.10), and (5.11) that \( H\Phi \leq 0 \). On the other hand, if \( \sigma_2(A_u) < 1 \), then

\[
H\Phi \leq -\mu/2 - \log \sigma_2(A_u) - C_1 \sigma_2(A_u)^{-\frac{1}{2}},
\]

and by choosing \( \mu \gg 1 \) large enough (depending only on \( C_1 \)) once again we have \( H\Phi \leq 0 \).

To obtain a lower bound for \( \log \sigma_2(A_u) \), we consider

\[
\tilde{\Phi} := -t \log \sigma_2(A_u) + \Lambda e^{-2u} - \mu t,
\]

and apply a similar argument. We will omit the details. \( \Box \)

**Proposition 5.11.** Suppose \( u \) is a solution to (5.2) with \( n = 4 \) on \([0, T], T \leq 1\), such that

\[
\sup_{M \times [0, T]} \left\{ |\nabla u|^2 + |u| \right\} \leq A.
\]

There exists a constant \( C = C(A) \) such that for all \( t \in [0, T] \), one has

\[
t \Delta u \leq C.
\]

**Proof.** Let

\[
\Phi = t \Delta u + |\nabla u|^2,
\]

where \( \Lambda \gg 1 \) will be chosen later. A direct calculation using Lemmas 5.4 and 5.6 and some elementary estimates yields

\[
H\Phi = \Delta u + t Fpqrs \nabla_i(A_u)_{pq} \nabla_i(A_u)_{rs}
\]

(5.12)

\[
+ \sigma_2(A_u)^{-1} T_1(A_u)_{pq} \left[ 2(t - 1) \nabla_i \nabla_p u \nabla_i \nabla_q u - t |\nabla^2 u|^2 g_{pq} + O(t |\nabla^2 u| + |\nabla u|^2 + 1) \right].
\]
If \( \Phi \) attains a large space-time maximum, say \( \Phi \geq B \geq 2A \), then
\[
t \Delta u \geq B - A \geq \frac{1}{2} B,
\]
hence
\[
t \left| \nabla^2 u \right|^2 \geq \frac{B^2}{16t}.
\]
Therefore, if \( t \leq 1 \), the terms in braces in (5.12) can be estimated as
\[
2 (t - 1) \nabla_i \nabla_p u \nabla_i \nabla_q u - t \left| \nabla^2 u \right|^2 g_{pq} + O(t \left| \nabla^2 u \right| + \left| \nabla u \right|^2 + 1) \leq \left\{ - t \left| \nabla^2 u \right|^2 + C t \left| \nabla^2 u \right| + C(A) \right\} g_{pq}
\]
\[
\leq \left\{ - \frac{t}{2} \left| \nabla^2 u \right|^2 + C' \right\} g_{pq}
\]
\[
\leq \left\{ - \frac{B^2}{32t} + C' \right\} g_{pq}
\]
\[
\leq 0,
\]
if \( B \) is large enough. Thus we conclude \( H \Phi < 0 \) at a sufficiently large maximum, proving the result. \( \square \)

**Theorem 5.12.** Let \((M^4, g)\) be a compact Riemannian manifold such that \( g \in \Gamma^+_2 \). Given \( u_0 \in \Gamma^+_2 \), there exists \( \epsilon = \epsilon(\left| u_0 \right|, \left| \nabla u_0 \right|) \) and \( C = C(\left| u_0 \right|, \left| \nabla u_0 \right|) \) such that the solution to (5.1) with initial condition \( u_0 \) exists on \([0, \epsilon]\) and moreover satisfies
\[
(5.13) \quad -C \leq \Delta u \leq \frac{C}{t}, \quad -C \leq t \log \sigma_2(A_u) \leq C.
\]
Furthermore, choosing \( l \in \mathbb{N}, 0 < \alpha < 1 \) there exists \( C_2 = C(\left| u_0 \right|, \left| \nabla u_0 \right|, l, \alpha) \) such that
\[
\left| u_{t} \right|_{C^{l; \alpha}} \leq C.
\]

**Proof.** The equation (5.2) is strictly parabolic for \( u_0 \in \Gamma^+_2 \), and so there exists a solution on some small time interval \([0, \eta]\). By Propositions 5.8 and 5.9, as long as the solution exists there is a uniform upper bound on \( \left| u \right|_{C^1} \) on \([0, \epsilon]\) where \( \epsilon \) depends only on \( \left| u_0 \right|_{C^1} \). The estimates of (5.13) follow from Propositions 6.11 and 5.10. Given these it follows that equation (5.2) is uniformly parabolic on \([0, \epsilon]\), and hence by the Evans-Krylov estimates \(15 \), \(22 \) there is a uniform \( C^{2; \alpha} \) estimate for \( u \) on \([0, \epsilon]\). Schauder estimates now imply that for any \( l, \alpha \) there are uniform \( C^{l; \alpha} \) bounds on \( u \) on \([0, \epsilon]\), which in particular proves that the solution actually exists for this whole time interval as well. Given these estimates, one relates the solution to (5.2) to the solution to (5.1) by adding a time dependent constant to \( u \) which fixed the volume to be \( V_{u_0} \). Since \( u \) is a priori bounded and this has no effect on any of the derivative estimates the result follows. \( \square \)

6. **Uniqueness of Solutions to \( \sigma_2 \)-Yamabe Problem**

In this section we combine the previous results to establish Theorem 1.5. As described in the introduction, the proof consists of a few main steps. In particular, we use Theorem 4.18 to connect any two critical points for \( F \) by an \( \epsilon \)-geodesic. Applying the geodesic convexity of \( F \) we obtain that the curve must consist of near-minimizers for \( F \). We then smooth this approximate geodesic via Theorem 5.12. Taking the limit as \( \epsilon \to 0 \) of these smoothed paths yields a nontrivial one-parameter family of minimizers of \( F \). Using our knowledge of the geodesic convexity of \( F \) we can show that this can only happen if the background conformal class is \([g_{S^4}]\), and the endpoints of the path are round metrics. Note that, unlike the Kähler setting, we are unable to show that the approximate geodesics converge directly to a nontrivial smooth geodesic due to the lack of stronger regularity results for the geodesics.
Lemma 6.1. Given $u_0, u_1$ two admissible critical points of $F$, one has $F[u_0] = F[u_1]$, and $F[u] \geq F[u_0]$ for all admissible $u$. Moreover, given $f$ and $u = u(x, t, s, \epsilon)$ the approximate geodesics given by Theorem 4.18, one has for any $t \in [0, 1],$

$$\lim_{s, \epsilon \to 0} F(u(\cdot, t, s, \epsilon)) = F[u_0].$$

Proof. Fix $f$, and let $u = u(x, t, s, \epsilon)$ be the approximate geodesics guaranteed by Theorem 4.18 connecting $u_0$ and $u_1$. To begin we repeat the calculation of Proposition 3.16 for these paths. Fix some $s, \epsilon$ and compute:

$$\frac{d^2}{dt^2} F[u] = \frac{d}{dt} \int_M u_t \left[ -\sigma_2(g_u^{-1}A_u) + \sigma \right] dV_u$$

$$= -\int_M u_t \sigma_2(g_u^{-1}A_u) + u_t \left( T_1(g_u^{-1}A_u), \nabla^2 u_t \right) dV_u$$

$$+ \sigma \int_M \left[ u_t V_u^{-1} + V_u^{-2} u_t \left( \int_M 4u_t dV_u \right) - 4V_u^{-1} u_t^2 \right] dV_u$$

$$= \int_M \left[ \epsilon u_t \sigma_2(g_u^{-1}A_u) - sf \right] dV_u + \sigma V_u^{-1} \int_M \left[ \frac{1}{\sigma_2(g_u^{-1}A_u)} sf - \epsilon u_t \right] dV_u$$

$$+ \sigma V_u^{-1} \int_M \left[ T_1(g_u^{-1}A_u), \nabla u_t \otimes \nabla u_t \right] - 4 \left( \int_M u_t^2 dV_u - V_u^{-1} \left( \int_M u_t dV_u \right)^2 \right) dV_u.$$

Applying Corollary 3.15 to the above equation yields

$$\frac{d^2}{dt^2} F \geq -\int_M sf dV_u - \sigma V_u^{-1} \epsilon \int_M u_t.$$

Now let us estimate using the uniform $C^1$ estimate

$$\int_0^1 \int_M u_t dV_u = \int_0^1 \left[ \frac{\partial}{\partial t} \int_M u_t dV_u - \int_M 4u_t^2 \right] dt$$

$$= \int_M u_t dV_u \bigg|_{t=1}^{t=1} - \int_0^1 \int_M 4u_t^2 dV_u dt$$

$$\leq C.$$

Hence, integrating the inequality (6.1) and using that $u_0$ is a critical point yields

$$\frac{d}{dt} F[u](t) = \frac{d}{dt} F[u](t) - \frac{d}{dt} F[u](0) = \int_0^t \frac{d^2}{dt^2} F dt \geq -C(s + \epsilon).$$

Integrating this in time and sending $s, \epsilon \to 0$ yields

$$F[u_1] \geq F[u_0].$$

But since the roles of $u_0$ and $u_1$ are interchangeable, we obtain $F[u_0] = F[u_1].$ \hfill \square

Lemma 6.2. Fix $(M^4, g)$ with $A_g \in T^2_+$, and suppose $u \in C^\infty(M)$ is an admissible critical point of $F$. Then either $u$ is an isolated critical point for $F$ or $(M, g_u)$ is isometric to $(S^4, g_{S^4})$.

Proof. Suppose $u$ is not an isolated critical point, so that there exists a sequence of admissible conformal factors \{u_i\}, $u_i \neq u$, converging in $C^\infty$ to $u$, normalized so that $\int_M (u - u_i) dV_u = 0$. We aim to use the convexity properties to show that the minimum eigenvalue of the linear operator

$$L(\phi) = -\left( T_1(g_u^{-1}A_u), \nabla^2 \phi \right)_{g_u} - 4\sigma \phi$$

is zero. Since $u$ satisfies $\sigma_2(A_u) \equiv \sigma$ and has unit volume, this lowest eigenvalue is characterized variationally as

$$\lambda_1 = \inf_{\{\phi : \int_M \phi dV_u = 0\}} \frac{\sigma}{\int_M} \left( \sigma_2(A_u)^{-1} \left( T_1(A_u), \nabla \phi \otimes \nabla \phi \right) - 4\phi^2 \right) dV_u.$$
We next evaluate this at (1.9) and (2.5), yielding contradiction from this setup. First we make a second variation calculation along this path using $w$ consists of admissible functions. Note that $\lambda$.

It follows from Corollary [3.15] that $\lambda_1 \geq 0$, with equality if and only if $(M^4, g_u)$ is isometric to $(S^4, g_{S^4})$. We suppose that $\lambda_1 > 0$ and derive a contradiction.

Fix a sufficiently large $i$ so that the path

$$w(x, t) = (1 - t)u + tu_i$$

consists of admissible functions. Note that $w_{ti} = 0$, and by construction $\frac{dF(w(\cdot,t))}{dt}(0) = \frac{dF(w(\cdot,t))}{dt}(1) = 0$. It follows that for any $i$ there exists $t_i \in [0,1]$ such that $\frac{d^2F(w(\cdot,t))}{dt^2}(t_i) = 0$. We aim to derive a contradiction from this setup. First we make a second variation calculation along this path using (1.9) and (2.5), yielding

$$\frac{d^2}{dt^2} F[w(\cdot,t)] = \frac{d}{dt} \int_M w_t \left( -\sigma_2(g_w^{-1}A_w) + \overline{\sigma} \right) dV_w$$

We next evaluate this at $t_i$. Using that $w^i := w(\cdot,t_i)$ converges to $u$ as $i \to \infty$ yields

$$0 = \int_M \left[ \langle T_1(g_w^{-1}A_w), \nabla w_t \otimes \nabla w_t \rangle - n\overline{\sigma}w_t^2 \right] dV_w.$$  

We now evaluate this at $t_i$. Using that $w^i := w(\cdot,t_i)$ converges to $u$ as $i \to \infty$ yields

$$0 = \int_M \left[ \langle T_1(g_w^{-1}A_w), \nabla w_t \otimes \nabla w_t \rangle - n\overline{\sigma}w_t^2 \right] dV_w.$$  

We now evaluate this at $t_i$. Using that $w^i := w(\cdot,t_i)$ converges to $u$ as $i \to \infty$ yields

$$0 = \int_M \left[ \langle T_1(g_w^{-1}A_w), \nabla w_t \otimes \nabla w_t \rangle - n\overline{\sigma}w_t^2 \right] dV_w.$$  

If $\lambda_1 > 0$ then for sufficiently large $i$ this implies that $w_t = u_i - u = 0$, a contradiction. It follows that $\lambda_1 = 0$, and hence by Corollary [3.15] $(M^4, g_u)$ is isometric to $(S^4, g_{S^4})$. \qed
Proof of Theorem 7.3 See Figure 8 for a schematic outline of the argument. Suppose there exist two distinct solutions \( u_0 \) and \( u_1 \) to the \( \sigma_2 \)-Yamabe problem. Let \( u(x, t, s, \epsilon) \) be the family of approximate geodesics connecting \( u_0 \) to \( u_1 \) guaranteed by Theorem 4.18. Noting the a priori estimates on \( |u|_{C^0} \) and \( |\nabla u|_{C^0} \) are independent of \( s, \epsilon \) we have by Theorem 5.12 that the solution to the flow equation (6.1) with initial condition \( u(\cdot, t, s, \epsilon) \) exists on some time interval \([0, \eta]\), and moreover the solution at time \( \eta \), call it \( v(x, t, s, \epsilon) \) has uniform \( C^{k,\alpha} \) estimates independent of \( s, \epsilon \) and stays uniformly in the interior of \( \Gamma^* \), in the sense that \( T_1(g^{-1}_wA_w) \) has uniform upper and lower bounds. Due to these estimates we can obtain one-parameter family of smooth functions \( v(x, t) = \lim_{s, \epsilon \to 0} v(x, t, s, \epsilon) \), which is continuous in \( t \). Moreover, by Lemmas 5.7 and 6.1 we see that \( F(v(\cdot, t)) = F[u_0] \). It follows that \( v(\cdot, t) \) is a nontrivial path of critical points for \( F \) through \( u_0 \), and hence by Lemma 6.2 we conclude that \( (M^4, g_u) \) is isometric to \( (S^4, g_{S^4}) \).

References

[1] B. Andrews, unpublished.
[2] Z. Blocki, On geodesics in the space of Kähler metrics, Advanced Lectures in Mathematics 21, p. 3-20, International Press, 2012.
[3] S. Brendle, Blow-up phenomena for the Yamabe equation J. Am. Math. Soc. 21, 951-979 (2008).
[4] S. Brendle, F.C. Marques, Blow-up phenomena for the Yamabe equation II J. Diff. Geom. 81, 225-250 (2009).
[5] S. Brendle, J. Viaclovsky, A variational characterization for \( \sigma_2 \), Calc. Var. 20, 399-402 (2004).
[6] E. Calabi, X.X. Chen, The space of Kähler metrics II, J. Diff. Geom. 61 (2002), 173-193.
[7] S.Y.A. Chang, M. J. Gursky, P. Yang, An equation of Monge-Ampère type in conformal geometry, and four-manifolds of positive Ricci curvature, Ann. of Math. (2) 155 (2002), no. 3, 709-787.
[8] S.Y.A. Chang, M. J. Gursky, P. Yang, An a priori estimate for a fully nonlinear equation on four-manifolds. Dedicated to the memory of Thomas H. Wolff, J. Anal. Math. 87 (2002), 151-186.
[9] S.Y.A. Chang, P. Yang, The inequality of Moser and Trudinger and applications to conformal geometry, Dedicated to the memory of Jorgen K. Moser Comm. Pure Appl. Math. 56 (2003), no. 8, 11351150.
[10] X.X. Chen, The space of Kähler metrics, J. Diff. Geom. 56 (2000), 189-234.
[11] X.X. Chen, G. Tian, Geometry of Kähler metrics and foliations by holomorphic discs Publ. Math. de L’IHES, Vol. 107, No. 1, 1-107.
[12] B. Chow, P. Lu, L. Ni, Hamilton’s Ricci flow, Lectures in Contemporary Mathematics, Science Press, Beijing.
[13] S.K. Donaldson, Conjectures in Kähler geometry, Strings and geometry, 71-78, Clay Math. Proc., 3, Amer. Math. Soc., Providence, RI 2004.
[14] S. K. Donaldson, Symmetric spaces, Kähler geometry and Hamiltonian dynamics, in Northern California Symplectic Geometry Seminar, Amer. Math. Soc. Transl. Ser. 2, 196, Amer. Math. Soc., Providence, 1999, 13-33.
[15] L. C. Evans, Classical solutions of fully nonlinear, convex, second-order elliptic equations, Comm. Pure Appl. Math. 35 (1982), 333-363.
[16] L. Garding, An inequality for hyperbolic polynomials J. Math. Mech. 8 1959 957-965.
[17] D. Gilbarg, N. Trudinger, Elliptic partial differential equations of second order, Springer, 1998.
[18] P. Guan, J. A. Viaclovsky, G. Wang, Some properties of the Schouten tensor and applications to conformal geometry, Trans. Amer. Math. Soc. 355 (2003), no. 3, 925933.
[19] P. Guan, G. Wang, A fully nonlinear conformal flow on locally conformally flat manifolds, J. Reine und Angew. Math. 557 (2003), 219-238.
[20] M. J. Gursky, J. Streets, A formal Riemannian structure on conformal classes and the inverse Gauss curvature flow, preprint.
[21] M. J. Gursky, J. Streets. Variational structure of the \( v_3 \)-Yamabe problem, preprint.
[22] N. V. Krylov, Boundedly inhomogeneous elliptic and parabolic equations in a domain, Izv. Akad. Nauk SSSR Ser. Mat. 47 (1983), 75-108.
[23] J. Lelong-Ferrand, Transformations conforme et quasi conforme des variés riemannienes; application la démonstration d’une conjecture de A. Lichnerowicz. C. R. Acad. Sci. Paris Sr. A-B 269 1969
[24] J. Lee, T. Parker, The Yamabe problem, Bull. AMS Vol 17, No. 1, (1987).
[25] A. Li, Y. Li, On some conformally invariant fully nonlinear equations, Comm. Pure Appl. Math. 56 (2003), no. 10, 14161464.
[26] T. Mabuchi, K-energy maps integrating Futaki invariants, Tohoku Math. Journ. 38 (1986), 575-593.
[27] T. Mabuchi, Some symplectic geometry on compact Kähler manifolds (1), Osaka J. Math 24 (1987), 227-252.
[28] M. Obata, Certain conditions for a Riemannian manifold to be isometric with a sphere, J. Math. Soc. Japan 14 (1962), 333-340.
[29] D. Pollack, *Nonuniqueness and high energy solutions for a conformally invariant scalar equation*, Comm. Anal. Geom. 1 (1993), no. 3-4, 347-414.

[30] R. Reilly, *Variational properties of functions of the mean curvatures for hypersurfaces in space forms*, J. Diff. Geom. 8 (1973) 465-477.

[31] R. Schoen, *Variational theory for the total scalar curvature functional for Riemannian metrics and related topics*, Topics in calculus of variations (Montecatini Terme, 1987), 1201-154, Lecture Notes in Math., 1365, Springer, Berlin, 1989.

[32] S. Semmes, *Complex Monge-Ampere equations and symplectic manifolds*, Amer. J. Math. 114 (1992), 495-550.

[33] W. Sheng, N.S. Trudinger, X.-J. Wang, *The k-Yamabe problem. Surveys in differential geometry*, Vol. XVII, 427-457, Surv. Differ. Geom., 17, Int. Press, Boston, MA, 2012.

[34] N. S. Trudinger, X.-J. Wang, *The intermediate case of the Yamabe problem for higher order curvatures*, Int. Math. Res. Not. IMRN 2010, no. 13, 2437-2458.

[35] Y. Tashiro, *Complete Riemannian manifolds and some vector fields*, Trans. Amer. Math. Soc 117 (1965), 251-275.

[36] J. A. Viaclovsky, *Conformally invariant Monge-Ampere equations: global solutions*, Trans. Amer. Math. Soc. 352 (2000), no. 9, 4371-4379.

[37] J. A. Viaclovsky, *Conformal geometry and fully nonlinear equations. Inspired by S. S. Chern*, 435-460, Nankai Tracts Math., 11, World Sci. Publ., Hackensack, NJ, 2006.

[38] J.A. Viaclovsky, *Conformal geometry, contact geometry, and the calculus of variations*, Duke. Math. J., Volume 101, No. 2 (2000), 283-316.

[39] J.A. Viaclovsky, *Estimates and Existence Results for some Fully Nonlinear Elliptic Equations on Riemannian Manifolds*, Comm. Anal. and Geometry 10 (2002), no. 4, 815-846.

Department of Mathematics University of Notre Dame, Notre Dame, IN 46556

*E-mail address:* mgursky@nd.edu

Department of Mathematics, University of California, Irvine, CA 92617

*E-mail address:* jstreets@uci.edu