SCALAR CURVATURE AND THE EXISTENCE OF GEOMETRIC STRUCTURES ON 3-MANIFOLDS, I.

MICHAEL T. ANDERSON

ABSTRACT. This paper analyses the convergence and degeneration of sequences of metrics on a 3-manifold, and relations of such with Thurston's geometrization conjecture. The sequences are minimizing sequences for a certain (optimal) scalar curvature-type functional and their degeneration is related to the sphere and torus decompositions of the 3-manifold under certain conditions.

0. INTRODUCTION.

This paper and its sequel are concerned with the limiting behavior of minimizing sequences for certain curvature integrals on the space of metrics on a 3-manifold M, and the relations of such behavior with the geometrization conjecture of Thurston [37]. First, recall the statement of Thurston's conjecture, in the case of closed, oriented 3-manifolds.

Geometrization Conjecture (Thurston).

Let M be a closed, oriented 3-manifold. Then M admits a canonical decomposition into domains, each of which carries a canonical geometric structure.

A geometric structure on a 3-manifold is a complete, locally homogeneous Riemannian metric. There are exactly eight such structures, namely the three constant curvature geometries, together with two further product geometries, $H^2\times\mathbb{R}$, $S^2\times\mathbb{R}$ and three twisted product geometries, $SL(2,\mathbb{R})$, Nil and Sol, c.f. [37], [33].

The decomposition of M is along certain essential 2-spheres and essential tori embedded in M. Recall that a 3-manifold M is *irreducible* if every embedded 2-sphere S^2 in M bounds a 3-ball $B^3 \subset M$.

The sphere decomposition [22], [24] states that M may be decomposed as a union of irreducible 3-manifolds, in that M is diffeomorphic to a connected sum of closed, oriented 3-manifolds,

$$M = (M_1\#M_2\#\cdots\#M_p)\#(N_1\#N_2\#\cdots\#N_q)\#(\#_1^r(S^2\times S^1)), \tag{0.1}$$

where each factor is irreducible, with the exception of $S^2\times S^1$. The manifolds M_i are defined to be those factors with infinite fundamental group, while N_j are the factors with finite π_1. Elementary 3-manifold topology implies that each manifold M_i is a closed, 3-dimensional $K(\pi,1)$, while each N_j has universal cover given by a homotopy 3-sphere.

Thus, the sphere decomposition allows one to understand the topology (and geometry) of 3-manifolds, in terms of the irreducible factors M_i or N_j, since the topology and geometry of the standard factor $S^2\times S^1$ can be viewed as clear.

Partially supported by NSF Grants DMS 9505744 and 9802722.

2000 *Mathematics Subject Classification.* Primary 58J60, 57M50; Secondary 58E11, 53C20.

In the present paper, we are only concerned with the structure of the M_i factors, i.e. irreducible 3-manifolds with infinite fundamental group. The torus decomposition will be discussed later, c.f. Conjecture I and §2.

The geometrization conjecture can be considered as a question concerning the existence of a "best possible" metric on a given 3-manifold. This is a classical, albeit very difficult, question in Riemannian geometry. As mentioned above, we approach this existence question by seeking the minimum of a natural functional on the space of smooth metrics on M.

Let $\mathbb{M}$ denote the space of all smooth Riemannian metrics on M and consider the L^2 norm of the scalar curvature as a functional on $\mathbb{M}$. This needs to be weighted with the appropriate power of the volume in order to obtain a scale-invariant functional. Thus, consider

$$\mathcal{S}^2 : \mathbb{M} \to \mathbb{R}, \quad \mathcal{S}^2(g) = \big(v^{1/3} \int_M s_g^2 dV_g\big)^{1/2}, \tag{0.2}$$

where s_g is the scalar curvature of the metric g, dV_g is the volume form associated with g, and v is the volume of (M, g). As explained in detail in [5, §1], c.f. also Remark 3.2 below, it is technically advantageous to consider a slightly weaker functional than $\mathcal{S}^2$, namely

$$\mathcal{S}_-^2 : \mathbb{M} \to \mathbb{R}, \quad \mathcal{S}_-^2(g) = \big(v^{1/3} \int_M (s_g^-)^2 dV_g\big)^{1/2}, \tag{0.3}$$

where $s^- = \min(s, 0)$ is the non-positive part of s. If M carries a metric of positive scalar curvature, then it also carries scalar-flat metrics, so that $\inf \mathcal{S}_-^2 = \inf \mathcal{S}^2 = 0$. In this case, there is an infinite dimensional family of minimizers of $\mathcal{S}_-^2$, so that there is no close relation between the geometry of such metrics and the topology of M.

On the other hand, let the Sigma constant $\sigma(M)$ be the supremum of the scalar curvatures of unit volume Yamabe metrics on M, c.f. [1], [30]. This is a topological invariant of the 3-manifold M, which should be thought of as an analogue of the Euler characteristic for surfaces. The set of closed 3-manifolds divides naturally into three topological classes, according to whether $\sigma(M)$ is negative, zero, or positive.

By the resolution to the Yamabe problem [29], $\sigma(M) \leq 0$ if and only if M carries no metrics of positive scalar curvature. It is not difficult to see, c.f. [5, Prop. 3.1], that if $\sigma(M) \leq 0$, then

$$\inf_{\mathbb{M}} \mathcal{S}_-^2 = \inf_{\mathbb{M}} \mathcal{S}^2 = |\sigma(M)|. \tag{0.4}$$

Thus, for this paper we are exclusively interested in closed 3-manifolds satisfying

$$\sigma(M) \leq 0. \tag{0.5}$$

We assume (0.5) holds throughout the paper. By a well-known result of Gromov-Lawson [15, Thm. 8.1], a closed 3-manifold M satisfies $\sigma(M) \leq 0$ if M has at least one non-empty factor M_i with infinite fundamental group in the sphere decomposition (0.1). Thus,

$$M_i \neq \emptyset, \quad \text{for some} \quad i \quad \Rightarrow \sigma(M) \leq 0. \tag{0.6}$$

Further, any metric g_o on M realizing $\inf \mathcal{S}_-^2$, (or $\inf \mathcal{S}^2$), is necessarily an Einstein metric, and thus of constant sectional curvature, c.f. §2.

Of course, an arbitrary closed, oriented 3-manifold does not admit an Einstein metric; this is the case for instance if M has a non-trivial sphere decomposition (0.1). On the other hand, if M is irreducible, we conjecture that there exist minimizing sequences $\{g_i\}$ for $\mathcal{S}_-^2$ which effectively implement the geometrization of M, in the sense that one can deduce the torus decomposition of M, and the geometrization of each canonical domain in M, from the limiting geometric behavior of $\{g_i\}$. This is expressed in the following Conjectures, (equivalent to Conjectures I and II of [1], in the context of maximizing sequences of Yamabe metrics on M in place of minimizing sequences of $\mathcal{S}_-^2$).

Conjecture I. *Let M be a closed, oriented, irreducible 3-manifold, with*

$$\sigma(M) < 0.$$

Then there is a finite collection $\mathcal{T}$ of disjoint, embedded, incompressible tori $T_i^2 \subset M$, which separate M into a union of two types of manifolds:

$$M \setminus \dot{\cup}\, T_i^2 = \dot{\cup}\, H_j \cup \dot{\cup}\, G_k. \tag{0.7}$$

Each H_j is a complete, connected hyperbolic manifold, of finite volume. The collection of boundary components of $\cup H_j$, i.e. the canonical tori in the hyperbolic cusps of $\{H_j\}$, forms exactly the collection $\mathcal{T}$. Each G_k is a connected graph manifold with toral boundary components, and the union of such boundary components again gives $\mathcal{T}$. This decomposition of M is unique up to isotopy of M.

Let $vol_{-1}H_j$ denote the volume of H_j in the hyperbolic metric. Then the Sigma constant of M is given by

$$|\sigma(M)| = (6\sum_j vol_{-1}H_j)^{2/3}. \tag{0.8}$$

In particular, if M is atoroidal, i.e. M contains no $\mathbb{Z} \oplus \mathbb{Z} \subset \pi_1(M)$, then M admits a hyperbolic structure which realizes the Sigma constant, i.e.

$$|\sigma(M)| = (6 vol_{-1} M)^{2/3}. \tag{0.9}$$

It is well-known that the condition that M be irreducible and atoroidal is also necessary for M to admit a hyperbolic metric.

Conjecture II. *Let M be a closed, oriented, irreducible 3-manifold, with*

$$\sigma(M) = 0.$$

Then M is a graph manifold satisfying

$$|\pi_1(M)| = \infty.$$

The Sigma constant $\sigma(M)$ is realized by a smooth metric in $\mathbb{M}_1$ if and only if M is a flat 3-manifold.

A *graph manifold* is a union of S^1 bundles over surfaces, or equivalently a union of Seifert fibered spaces, glued together by toral automorphisms along toral boundary components, c.f. [38,39]. This is also exactly the class of 3-manifolds which admit an F-structure, in fact a polarized F-structure, in the sense of Cheeger-Gromov [7,8], c.f. also [13, App.2].

The geometrization of graph manifolds is relatively straightforward and well understood. Briefly, any irreducible graph manifold admits a further decomposition into domains, with toral boundary components, each of which admits a geometric structure modelled on one of the Seifert fibered geometries or it admits a Sol geometry. This decomposition and geometrization of the graph manifolds is obtained naturally from the proofs of the results below, c.f. §2.3 for further discussion.

Topologically, these two conjectures imply that a closed, oriented, irreducible 3-manifold with $\sigma(M) \leq 0$ is a union of hyperbolic manifolds and graph manifolds glued together along incompressible toral boundary components. The union of the graph manifolds in M is also called the *characteristic subvariety* of M. The case $\sigma(M) = 0$ implies the absence of hyperbolic components, while the case $\sigma(M) < 0$ implies their existence.

Geometrically, Conjectures I and II are meant to describe the limiting behavior of a suitable minimizing sequence $\{g_i\}$ for $\mathcal{S}^2_-$. Thus, in I, conjecturally the sequence $\{g_i\}$ converges on the manifolds H_j in (0.7) to a complete, finite volume, constant curvature metric, (with scalar curvature $\sigma(M)$), and collapses the graph manifold components G_k in (0.7) along S^1 or T^2 fibers. In II, conjecturally the sequence $\{g_i\}$ fully collapses M along S^1 or T^2 fibers, so that there is no limiting metric in general. In these cases of collapse, although the metrics are not converging, their degeneration implies the existence of a well defined topological structure, (a graph manifold structure), incompressibly embedded in M, which essentially describes how the degeneration is occuring. In effect, the limiting behavior of the sequence $\{g_i\}$ implements, or performs, the geometrization of M.

Conjectures I and II, together with (0.6), are easily seen to imply the geometrization conjecture for closed, oriented irreducible 3-manifolds with infinite fundamental group, c.f. [1, §4], or also Remark 2.11 and §2.3.

In this paper, we prove Conjectures I and II in case either one of two additional topological assumptions hold on M. Some further background is needed to explain these assumptions. For a given $\varepsilon > 0$ small, consider the scale-invariant perturbation of $\mathcal{S}^2_-$ given by

$$I_\varepsilon^- = \varepsilon v^{1/3} \int |z|^2 dV + \big(v^{1/3} \int (s^-)^2 dV\big)^{1/2}. \tag{0.10}$$

The functional I_ε^- is a perturbation of $\mathcal{S}^2_-$ in the direction of the L^2 norm of the trace-free Ricci curvature z. While the existence of minimizers of $\mathcal{S}^2_-$ is difficult to prove - this is the essential geometric content of Conjectures I and II - it is not so difficult to prove the existence of minimizers g_ε of I_ε^-. This is done in [2,5], where the following results are proved, (c.f. also Theorem 1.1 below for more details). For any closed oriented 3-manifold M, (not necessarily irreducible), there exists a maximal domain Ω_ε and a $C^{2,\alpha}$ complete Riemannian metric g_ε on Ω_ε, for which the pair $(\Omega_\varepsilon, g_\varepsilon)$ realizes $\inf_{\mathbb{M}_1} I_\varepsilon^-$, in that

$$I_\varepsilon^-(g_\varepsilon) = \varepsilon \int_{\Omega_\varepsilon} |z_{g_\varepsilon}|^2 + \big(\int_{\Omega_\varepsilon} (s_{g_\varepsilon}^-)^2 \big)^{1/2} = \inf_{\mathbb{M}_1} I_\varepsilon^-, \tag{0.11}$$

and

$$vol_{g_\varepsilon} \Omega_\varepsilon = 1. \tag{0.12}$$

The domain Ω_ε weakly embeds in M, in the sense that any compact domain with smooth boundary embeds as such a domain in M. There is an exhaustion of Ω_ε by compact domains K_j with each ∂K_j given by a finite collection of smooth tori, such that the complement $M \backslash K_j$ is a graph manifold in M and so admits an F-structure $\mathcal{F}_j$.

The domain Ω_ε is empty exactly when the closed 3-manifold M itself is a graph manifold. If M is not a graph manifold, then $\Omega_\varepsilon \neq \emptyset$, and one may have $\Omega_\varepsilon = M$, or Ω_ε only weakly embedded in M.

It is proved in [5, (3.3)-(3.4)], that if g_ε is a minimizer for I_ε^- as above, then as $\varepsilon \to 0$,

$$\mathcal{S}^2_-(g_\varepsilon) \to |\sigma(M)|, \tag{0.13}$$

and

$$\varepsilon \int |z_{g_\varepsilon}|^2 dV_{g_\varepsilon} \to 0. \tag{0.14}$$

Thus the family $(\Omega_\varepsilon, g_\varepsilon)$ as $\varepsilon \to 0$, may be considered as a specific minimizing family for the functional $\mathcal{S}^2_-$. This family is optimal in that it has the least amount of curvature in L^2 in the following sense: given any $(\Omega_\varepsilon, g_\varepsilon)$ as above, for any smooth unit volume metric $\bar{g}$ on M for which

$$\mathcal{S}^2_-(\bar{g}) \leq \mathcal{S}^2_-(g_\varepsilon),$$

one has

$$\int_M |z_{\bar{g}}|^2 dV_{\bar{g}} \geq \int_{\Omega_\varepsilon} |z_{g_\varepsilon}|^2 dV_{g_\varepsilon}.$$

A pair $(\Omega_\varepsilon, g_\varepsilon)$ as above is called a minimizing pair for I_ε^-. We point out that it is not known, (without resolution of Conjectures I and II), if a minimizing pair is unique. Thus neither the metrics g_ε, nor the topological type of the domains Ω_ε, are known to be unique, in the sense that they depend only on the topology of M. The only exception to this is when M is a graph manifold, when, as mentioned above, both Ω_ε, and so g_ε, are empty. On the other hand, the collection of all minimizing pairs $(\Omega_\varepsilon, g_\varepsilon), \varepsilon > 0$, depends only on the topology of M.

With this background, we now are able to state the first topological assumption on M.

Definition 0.1. *Let M be a closed, oriented 3-manifold. Then M is* tame *if there exists a sequence $\varepsilon = \varepsilon_i \to 0$, and a sequence of minimizing pairs $(\Omega_\varepsilon, g_\varepsilon)$ for $I_\varepsilon{}^-$ on M such that*

$$\int_{\Omega_\varepsilon} |z_\varepsilon|^2 dV_{g_\varepsilon} \leq \Lambda, \tag{0.15}$$

for some $\Lambda < \infty$.

It is clear that the condition that M is tame is a topological condition, i.e. depends only on the smooth, and hence topological, structure of M. In fact, from the definition of g_ε, M is tame if and only if there exists a sequence $\{g_i\}$ of smooth metrics on M, which form a minimizing sequence for $\mathcal{S}_-^2$, and which have uniformly bounded z-curvature in L^2.

Theorem 0.2. *Let M be a closed oriented 3-manifold with $\sigma(M) \leq 0$. If M is tame, then Conjectures I and II hold for M.*

We mention that Theorem 0.2 remains valid even if the tame manifold M is not irreducible. Of course the hyperbolic part $\cup H_j$ of M in (0.7) is irreducible, but the graph manifold part $\cup G_k$ of M, and hence M itself, could be reducible. In particular, a tame 3-manifold need not be irreducible, c.f. §1.

A proof of the analogue of Theorem 0.2 for maximizing sequences of Yamabe metrics was outlined in [1], without full details however. The proof given here may be adapted to such sequences of Yamabe metrics without difficulty. Theorem 0.2 is also analogous to a recent result of Hamilton [16] on the long-time behavior of the Ricci flow assuming a uniform L^∞ bound on the curvature.

Next we turn to the more interesting and difficult case where M is not tame, so that an arbitrary minimizing sequence for $\mathcal{S}_-^2$ has curvature diverging to infinity in L^2. In particular, this is the case for any sequence of minimizing pairs $(\Omega_\varepsilon, g_\varepsilon)$, with $\varepsilon = \varepsilon_i$ any sequence converging to 0. The metrics g_ε are thus degenerating and one would like to relate this degeneration to the topology of M.

In this case, it is proved in [5, Thm. B], c.f. also Theorem 1.3 below, that for any sequence $\varepsilon_i \to 0$ and sequence of minimizing pairs $(\Omega_\varepsilon, g_\varepsilon), \varepsilon = \varepsilon_i$, there exist base points $y_\varepsilon \in \Omega_\varepsilon$ and scale factors $\rho(y_\varepsilon) \to 0$ as $\varepsilon \to 0$, such that the rescaled or blow-up metrics

$$g'_\varepsilon = \rho(y_\varepsilon)^{-2} \cdot g_\varepsilon \tag{0.16}$$

based at $\{y_\varepsilon\}$ have a subsequence converging to a complete non-flat Riemannian manifold (N, g', y) of non-negative scalar curvature and uniformly bounded curvature. The limit (N, g') minimizes the L^2 norm of z over all comparison metrics $\bar{g}$ of non-negative scalar curvature on N such that $vol_{\bar{g}}K \leq vol_g K$ and $\bar{g}|_{N\setminus K} = g|_{N\setminus K}$, for some arbitrary compact set $K \subset N$. The base points $y_\varepsilon \in \Omega_\varepsilon$ are preferred in that the curvature of g_ε is (locally) maximal near y_ε.

The structure of these complete metrics (N, g, y) models the small scale geometry of the degeneration of $(\Omega_\varepsilon, g_\varepsilon)$, near the base points y_ε where the curvature blows up. The structure of the limit (N, g') is discussed in more detail in §1. We point out that the base points y_ε, as well as the choice of subsequence above may not be unique. Thus it is possible apriori that there are many distinct blow-up limits (N, g').

Definition 0.3. *Let M be a closed oriented 3-manifold. Then M is* spherically tame *if there exists a complete non-flat blow-up limit (N, g', y) as above which has an* asymptotically flat end *$E \subset N$ in the following sense: there is a diffeomorphism $F\colon \mathbb{R}^3 \setminus B \to E$, for some compact ball $B \subset \mathbb{R}^3$,* *such that the metric g' has the form*

$$g'_{ij} = (1 + \frac{2m}{r})\delta_{ij} + h, \tag{0.17}$$

in the chart F, where $h = O(r^{-2}), |D^p h| = O(r^{-2+p}), p = 1,2$, $r(x) = |x|$. The parameter m is the mass of the end E and is assumed to be positive.

Again this condition depends only on the smooth structure of the manifold M. Thus the class of non-tame 3-manifolds, (closed and oriented), divides into two classes, namely those which are spherically tame, and those which are not. Of course, this condition is expressed in terms of analysis and geometry, and so its relation with standard topological concepts may not be immediately clear. In this respect, we prove the following result in §3.

Theorem 0.4. *Let M be a closed oriented 3-manifold with $\sigma(M) \leq 0$, which is not tame but is spherically tame. Then M is reducible, i.e. M contains an essential embedded 2-sphere.*

In fact we prove that the "natural" 2-sphere S^2 in the asymptotically flat end $E \subset N$ is an essential 2-sphere in M. The method of proof is by a cut and paste or comparison argument. If this S^2 bounds a 3-ball B^3 in M, we construct a specific metric glueing of such a B^3 onto S^2 so as to decrease the value of I_ε^- a definite amount below that of $I_\varepsilon^-(g_\varepsilon)$, for ε sufficiently small, which contradicts the minimizing property of g_ε. An analogous cut and paste argument is used in the proof of Theorem 0.2 to prove the tori $\mathcal{T}$ in (0.7) are incompressible.

The fact that it is possible to carry out such cut and paste arguments on spheres and tori is a strong reason for preferring the functional I_ε^- over other candidates.

Theorem 0.4 relates the geometry and analysis of the metrics g_ε with the sphere decomposition (0.1). In particular, it implies that if M is irreducible and spherically tame, then M must be tame. Hence, via Theorems 0.2 and 0.4, Conjectures I and II are reduced to the question of whether a closed oriented 3-manifold is automatically spherically tame, i.e. to the following:

Sphere conjecture. *Any closed oriented 3-manifold M which is not tame is spherically tame.*

The Sphere conjecture will be the focus of the sequel paper. This conjecture asserts that there exists a blow-up limit (N, g', y) of a non-tame minimizing sequence $(\Omega_\varepsilon, g_\varepsilon), \varepsilon = \varepsilon_i \to 0$, which has an asymptotically flat end. More loosely speaking, it claims that one can detect 2-spheres S^2 in M, (embedded in almost flat regions), from the degenerating geometry of a suitable minimizing sequence for $\mathcal{S}^2_-$.

The existence of such an end might seem to be a strong condition. In fact, asymptotically flat ends are the most natural end structure for complete blow-up limits, and no situations are known where such a metric might have a non-asymptotically flat end. In this regard, [5, Thm.C], c.f. also Theorem 1.4 below, gives a relatively simple characterization of blow-up limits (N, g') which have a finite number of ends, each of which is asymptotically flat. To prepare for work to follow in the sequel, in §4 we extend this result by characterizing those blow-up limits (N, g') which have at least one asymptotically flat end, c.f. Proposition 4.1. Further, in Theorem 4.2 the arguments proving Theorem 0.4 are shown to extend to limits (N, g') which have a suitable almost flat annulus in place of a full asymptotically flat end.

Briefly, the paper is organized as follows; more detailed remarks on the contents are given at the beginning of each section or subsection. Following discussion of some necessary background material in §1, the proof of Theorem 0.2 is given in §2. Also, in §2.3, we discuss the geometrization of graph manifolds, as obtained from the limiting behavior of minimizing sequences for $\mathcal{S}^2_-$, c.f. Theorem 2.14. The proof of Theorem 0.3 is given in §3, while §4 collects some results and techniques on the structure of blow-up limits (N, g') needed for the sequel paper.

While this paper relies to a certain extent on results from earlier papers [2]-[5], we have made the paper at least logically almost self-contained in that the results needed from these papers are summarized in §1.

1. Background Material.

Throughout the paper, we assume knowledge of the Cheeger-Gromov theory of convergence and collapse of Riemannian manifolds, [12], [7,8], as well as the L^2 Cheeger-Gromov theory on 3-manifolds, [2, §3], c.f. also [5, §2] for a summary.

The following geometric quantities will be used frequently.

Definition 1.0. *Let (N, g) be a complete Riemannian manifold, and $x \in N$.*

(i). *The L^2 curvature radius at x is the largest radius $\rho(x)$ such that for any geodesic ball $B_y(s) \subset B_x(\rho(x)), s \leq \rho(x)$, one has*

$$\frac{s^4}{volB_y(s)}\int_{B_y(s)}|r|^2 dV \leq c_o; \tag{1.1}$$

where r is the Ricci curvature and c_o is a fixed small positive parameter. Throughout the paper, we set $c_o = 10^{-3}$ for convenience.

(ii). *The volume radius $\nu(x)$ at x is given by*

$$\nu(x) = \sup\{r : \frac{vol(B_y(s))}{s^3} \geq \mu, \forall B_y(s) \subset B_x(r), s \leq r\}, \tag{1.2}$$

where again μ is a fixed small parameter, which measures the degree of the volume collapse near x. As above, to be concrete, we assume $\mu = 10^{-1}$ throughout the paper.

There are similar definitions for the $L^{k,p}$ curvature radius, $k \geq 0$, $1 < p < \infty$, where the L^2 norm of r in (1.1) is replaced by its $L^{k,p}$ norm, and the power s^4 is replaced by the power making the expression analogous to (1.1) scale-invariant. It is important to note that the $L^{k,p}$ curvature radius is continuous under convergence in the strong $L^{k,p}$ topology, c.f. [2, §3] and further references therein.

A domain Ω is *weakly embedded* in a closed 3-manifold M if every subdomain $K \subset \Omega$ with compact closure and with smooth boundary in Ω may be smoothly embedded as such a domain in M; in this case, we write

$$\Omega \subset\subset M. \tag{1.3}$$

A domain is defined to be an open 3-manifold, not necessarily connected.

The Euler-Lagrange equations $\nabla I_\varepsilon{}^-(g) = 0$ for the functional $I_\varepsilon{}^-$ from (0.11) at a unit volume metric g are the following system of elliptic PDE, c.f. [5, §3]:

$$\varepsilon\nabla\mathcal{Z}^2 + L^*\tau + \phi\cdot g = 0, \tag{1.4}$$

$$2\Delta(\tau + \frac{\varepsilon s}{12}) + \frac{1}{4}s\tau = -\frac{1}{2}\varepsilon|z|^2 + 3c. \tag{1.5}$$

Here L^* is the L^2 adjoint of the linearization of the scalar curvature, given by

$$L^*u = D^2u - \Delta u\cdot g - u\cdot r, \tag{1.6}$$

where D^2 is the Hessian and $\Delta = trD^2$ is the Laplacian w.r.t. the metric. The term $\nabla\mathcal{Z}^2$ in (1.4) is the gradient of $\mathcal{Z}^2 = \int|z|^2dV$, and is of the form

$$\nabla\mathcal{Z}^2 = D^*Dz + \frac{1}{3}D^2s - 2\overset{\circ}{R}\circ z + \frac{1}{2}(|z|^2 - \frac{1}{3}\Delta s)\cdot g,$$

c.f. [6,§4H]. The lower order terms in (1.4)-(1.5) are given by $\phi = -\frac{1}{4}s\tau + c$, $c = \frac{1}{12\sigma}\int (s^-)^2 + \frac{\varepsilon}{6}\int |z|^2$, s is the scalar curvature of g and

$$\tau = \frac{s^-}{\sigma}, \tag{1.7}$$

with $\sigma = (\int (s^-)^2)^{1/2}$. Note that τ is non-positive and has L^2 norm equal to 1.

The following result from [5,Thm.3.8] proves the existence and basic geometric properties of minimizers of the functional ${I_\varepsilon}^-$.

Theorem 1.1. (Geometric Decomposition for ${I_\varepsilon}^-$). *Suppose $\sigma(M) \leq 0$. For any $\varepsilon > 0$, there is a complete, $L^{3,p} \cap C^{2,\alpha}$ Riemannian metric g_ε, defined on a domain $\Omega_\varepsilon \subset\subset M$, which realizes $\inf_{\mathbb{M}_1} {I_\varepsilon}^-$, in the sense that*

$$ {I_\varepsilon}^-(g_\varepsilon) = \varepsilon \int_{\Omega_\varepsilon} |z_{g_\varepsilon}|^2 + \Big(\int_{\Omega_\varepsilon} (s_{g_\varepsilon}^-)^2 \Big)^{1/2} = \inf_{\mathbb{M}_1} {I_\varepsilon}^-, \tag{1.8}$$

and

$$ vol_{g_\varepsilon}\Omega_\varepsilon = 1. \tag{1.9}$$

The metric g_ε weakly satisfies the Euler-Lagrange equations (1.4)-(1.5) and is C^∞smooth, in fact real-analytic, away from the locus where $s = 0$.

Further, the curvature R of g_ε is uniformly bounded on Ω_ε and Ω_ε consists of a finite number $Q = Q(\varepsilon, M)$ of components. There is an exhaustion of Ω_ε by compact domains K_j with ∂K_j given by a finite collection of smooth tori, such that the complement $M \setminus K_j$ is a graph manifold embedded in M.

As in the Introduction, $(\Omega_\varepsilon, g_\varepsilon)$ is called a minimizing pair for ${I_\varepsilon}^-$. The domain Ω_ε is empty if and only if M itself is a graph manifold. In fact,

$$\inf {I_\varepsilon}^- = 0, \quad \text{for all} \ \ \varepsilon > 0, \tag{1.10}$$

if and only if M is a graph manifold, c.f. [13, App.2] or [5, §3.1]. Of course, this implies inf $\mathcal{S}_-^2 = 0$ also for graph manifolds. (The converse is the main content of Conjecture II). Graph manifolds can be characterized geometrically as exactly the class of 3-manifolds which admit a volume collapse with bounded curvature, i.e. there exists a sequence of metrics $\{g_k\}$ on M such that $|R_{g_k}| \leq 1$ everywhere, while $vol_{g_k}M \to 0$ as $k \to \infty$. Equivalently, they admit an F-structure, in fact a polarized F-structure, in the sense of Cheeger-Gromov [7,8]. Briefly, a 3-manifold admits an F-structure if there is a partition of M into domains, each of which admits either an S^1 or a T^2 action, for which the dimension of the orbits is positive everywhere. The F-structure is polarized if the group actions are locally free.

It is interesting and important to note that the class of graph manifolds is closed under connected sums, c.f. [39, p.91], or also [34, Lemma 4], so that graph manifolds may well be reducible. Via (1.10), this shows in particular that tame 3-manifolds need not be irreducible, c.f. also §2.3 for further discussion.

Theorem 0.1 gives a geometric decomposition of M w.r.t. the functional ${I_\varepsilon}^-$, in that M is the union of Ω_ε, or more precisely $K_j \subset \Omega_\varepsilon$, and its graph manifold complement in M, with $(\Omega_\varepsilon, g_\varepsilon)$ a solution to a natural geometric variational problem. However, it is very unlikely that this geometric decomposition corresponds to the geometric decomposition given in Conjecture I on arbitrary, in particular on reducible 3-manifolds. In situations where M is irreducible and Conjectures I and II are known to hold, then these two decompositions do coincide, c.f. Remark 2.10.

On the other hand, without such knowledge, it is unknown if Ω_ε even has finite topological type, or if the number of components of Ω_ε remains uniformly bounded on a sequence $\varepsilon = \varepsilon_i \to 0$. As stated in the Introduction, it is also unknown if $(\Omega_\varepsilon, g_\varepsilon)$ is unique, for $\varepsilon > 0$ fixed.

The behavior of the potential function $\tau = \tau_\varepsilon$ from (1.4)-(1.5) plays an important role throughout the paper. Let $T = T_\varepsilon = \sup_{\Omega_\varepsilon} |\tau_\varepsilon| = -\inf_{\Omega_\varepsilon} \tau_\varepsilon$. Then $T \geq 1$, since the L^2 norm of τ_ε over Ω_ε equals 1. By [5,(3.38)], one has the bound

$$1 \leq T \leq (1 + \frac{2\varepsilon}{\sigma}\mathcal{Z}^2(g_\varepsilon))^{1/2}. \tag{1.11}$$

The following result from [5, Thms. 3.10-3.11] summarizes some of the properties of τ_ε.

Proposition 1.2. *If $\sigma(M) < 0$, then the function τ_ε satisfies*

$$\inf_{\Omega_\varepsilon} \tau_\varepsilon \to -1, \tag{1.12}$$

and, for any $p < \infty$,

$$\int_{\Omega_\varepsilon} |\tau_\varepsilon + 1|^p dV_{g_\varepsilon} \to 0, \tag{1.13}$$

as $\varepsilon \to 0$. If $\sigma(M) = 0$, then

$$\inf_{\Omega_\varepsilon} s_\varepsilon \to 0, \text{ as } \varepsilon \to 0. \tag{1.14}$$

Further, in both cases,

$$\int_{\Omega_\varepsilon} |\nabla\tau_\varepsilon|^2 dV_{g_\varepsilon} \to 0. \tag{1.15}$$

We next summarize below two of the main results, namely Theorems B,C, of [5] concerning the structure of the blow-up limits (N, g) discussed in §0. In analogy to the Euler-Lagrange equations (1.4)-(1.5), the $\mathcal{Z}_c^2$ equations are an elliptic system of equations in a metric g, given by

$$\alpha\nabla\mathcal{Z}^2 + L^*\tau = 0, \tag{1.16}$$

$$\Delta(\tau + \frac{\alpha}{12}s) = -\frac{\alpha}{4}|z|^2, \tag{1.17}$$

for some constant $\alpha > 0$. As previously in (1.4)-(1.5), the function τ is viewed as a potential function.

Theorem 1.3. (Structure of Blow-up Limits). *Suppose $\sigma(M) \leq 0$, and let $(\Omega_\varepsilon, g_\varepsilon), \varepsilon = \varepsilon_i \to 0$, be a sequence of minimizing pairs for $I_\varepsilon^{\ -}$. Suppose the sequence $\{(\Omega_\varepsilon, g_\varepsilon)\}$ degenerates, in the sense that*

$$\int_{\Omega_\varepsilon} |z_{g_\varepsilon}|^2 dV_{g_\varepsilon} \to \infty, \quad \text{as} \quad \varepsilon \to 0. \tag{1.18}$$

Then there exist points $\{y_\varepsilon\} \in (\Omega_\varepsilon, g_\varepsilon)$, with $\rho(y_\varepsilon) \to 0$, such that the blow-up metrics

$$g'_\varepsilon = \rho(y_\varepsilon)^{-2} \cdot g_\varepsilon, \tag{1.19}$$

based at y_ε, have a subsequence converging in the strong $L^{2,2}$ topology to a limit (N, g', y). The limit (N, g') is a complete, non-flat Riemannian manifold, with uniformly bounded curvature, and non-negative scalar curvature s.

Further, (N, g') minimizes the L^2 norm of the curvature z over all metrics $\bar{g}$ of non-negative scalar curvature satisfying $vol_{\bar{g}}K \leq vol_{g'}K$ and $\bar{g}|_{N\setminus K} = g'|_{N\setminus K}$, for some compact set $K \subset N$. The metric g' is $C^{2,\beta} \cap L^{3,p}$ smooth, for any $\beta < 1$ and $p < \infty$, and is an $L^{3,p}$ weak solution of the $\mathcal{Z}_c^2$ equations (1.16)-(1.17).

The potential function τ is non-positive, and τ and s are locally Lipschitz functions on N with disjoint supports, in the sense that

$$s \cdot \tau \equiv 0.$$

The metric g' is C^∞ smooth, in fact real-analytic, and the convergence to the limit is C^∞ smooth, uniformly on compact subsets, in any region where $\tau < 0$ or $s > 0$ on the limit. The complete *manifold N is weakly embedded in Ω_ε, in the sense that any smooth compact domain $K \subset N$ embeds in Ω_ε, provided ε is sufficiently small, depending on K. Consequently, N weakly embeds in M.*

The blow-up limits (N, g') model the small-scale degeneration of the sequence $(\Omega_\varepsilon, g_\varepsilon)$ in neighborhoods of the base points y_ε. Theorem 1.3 implies in particular that the metrics g'_ε in (1.19) do not collapse, in the sense of Cheeger-Gromov, near the base points y_ε. The potential function τ in (1.16)-(1.17) is a limit of the functions $\tau_\varepsilon, \varepsilon = \varepsilon_i$, in (1.7), or suitable renormalizations thereof.

Theorem 1.4. (Asymptotically Flat Ends). *Let (N, g, τ) be a complete non-flat $\mathcal{Z}_c^2$ solution, i.e a solution of (1.16)-(1.17). Suppose there exists a compact set $K \subset N$ and a constant $\omega_o < 0$ such that the potential function*

$$\omega = \tau + \frac{\alpha}{12}s : N \to \mathbb{R},$$

in (1.17) satisfies

$$\omega \leq \omega_o < 0, \tag{1.20}$$

on $N \setminus K$, and that the level sets of ω in N are compact.

Then N is an open 3-manifold of the form

$$N = P\#(\#_1^q \mathbb{R}^3), \tag{1.21}$$

where P is a closed 3-manifold, (possibly empty), admitting a metric of positive scalar curvature, i.e. $\sigma(P) > 0$, and $1 \leq q < \infty$.

Each end $E = E_k$, $1 \leq k \leq q$, of N is asymptotically flat in the sense of Definition 0.3, and the potential ω has the expansion

$$\omega = \omega_E + \frac{m_E|\omega_E|}{r} + O(r^{-2}), \tag{1.22}$$

where $\omega_E < 0$ is a constant depending on (E, g) and $m_E > 0$ is the mass of the end E.

Theorem 1.4 characterizes the complete $\mathcal{Z}_c^2$ solutions for which all ends are asymptotically flat, (except possibly when the potential ω goes to 0 at infinity in some end).

The simplest example of a metric satisfying the conclusions of Theorem 1.4 is the Schwarzschild metric, (on the space-like hypersurface),

$$g_S = (1 - \frac{2m}{r})^{-1}dr^2 + r^2 ds_{S^2}^2, \tag{1.23}$$

defined on $[2m, \infty) \times S^2$, and isometrically doubled across the horizon $\Sigma = \{r = 2m\}$; observe that Σ is a totally geodesic 2-sphere, of constant curvature $(2m)^{-2}$. The metric g_S is asymtotically flat at each end. We refer to [3, Prop. 5.1] for the exact form of the potential τ for this metric.

Note that if blow-ups $\{g'_\varepsilon\}$ of $\{g_\varepsilon\}$ converge to the Schwarzschild metric, then the metrics $\{g_\varepsilon\}$ themselves are collapsing or crushing the core S^2 in g_S to a point. Similarly, this occurs for the 2-spheres in any asymptotically flat end of a blow-up limit. If such S^2's are essential, the metrics $\{g_\varepsilon\}$ are, in effect, then performing or carrying out a process analogous to the sphere decomposition (0.1) of M; c.f. the proof of Theorem 2.14 for a concrete illustration of this.

We close this background section with a discussion of the regularity of the metrics $(\Omega_\varepsilon, g_\varepsilon)$ as $\varepsilon \to 0$ when M is tame. Thus, the L^2 bound on the curvature (0.15) may not apriori control the L^∞ norm of the curvature or even the L^2 curvature radius ρ of $(\Omega_\varepsilon, g_\varepsilon)$ from (1.1). As an example, consider the following family of 2-dimensional metrics suggested by Gallot [10]: for any $\varepsilon > 0$,

$$h_\varepsilon = dr^2 + (\varepsilon + r)^{2k} d\theta^2,$$

on $(-1,1) \times S^1$. For $k > 3$, the metric h_ε satisfies the bound (0.15), with $\Lambda = k^2(k-1)^2/(k-3)$. However, the curvature blows up in L^∞ and the L^2 curvature radius goes to 0, in that

$$\rho_\varepsilon(0,\theta) \to 0 \text{ as } \varepsilon \to 0,$$

for any $\theta \in S^1$. The metrics h_ε collapse at $r = 0$ as $\varepsilon \to 0$, in the sense that the volume radius goes to 0, (but do not collapse at any $r \neq 0$). Observe that this transition from non-collapse to collapse takes place within bounded distance, so that the limit metric is incomplete. Of course there are similar metrics in dimensions ≥ 3.

However, for the "special" metrics g_ε, this phenomenon does not occur, as shown in the following Lemma.

Lemma 1.5. *For any minimizing pair $(\Omega_\varepsilon, g_\varepsilon)$, there is a constant $\rho_o = \rho_o(\Lambda) > 0$ such that if (0.15) holds, then, for all $x_\varepsilon \in (\Omega_\varepsilon, g_\varepsilon)$,*

$$\rho(x_\varepsilon) \geq \rho_o. \tag{1.24}$$

Proof: The proof is by contradiction, so suppose (0.15) holds, but (1.24) does not, on some sequence $\varepsilon = \varepsilon_j \to 0$. Then by [5, Thm.4.5], (a slightly stronger version of Theorem 1.3 above), there is a sequence of base points $y_\varepsilon \in \Omega_\varepsilon$, for some sequence of minimizing pairs $(\Omega_\varepsilon, g_\varepsilon), \varepsilon = \varepsilon_j \to 0$, with

$$\rho(y_\varepsilon) \to 0, \tag{1.25}$$

such that the blow-up metrics $g'_\varepsilon = \rho(y_\varepsilon)^{-2} \cdot g_\varepsilon$ based at y_ε, have a subsequence converging to a complete, non-flat limit (N, g') having the properties in Theorem 1.3. We point out again that this result implies in particular that the metrics g'_ε do not collapse near y_ε in this situation. The convergence is in the strong $L^{2,2}$, (in fact $C^{2,\alpha}$), topology on compact subsets of N, (c.f (1.26) below).

Now by the scaling properties of curvature and volume, one has

$$\int_{\Omega_\varepsilon} |z_{g'_\varepsilon}|^2 dV_{g'_\varepsilon} = \rho(y_\varepsilon) \cdot \int_{\Omega_\varepsilon} |z_{g_\varepsilon}|^2 dV_{g_\varepsilon} \to 0,$$

by (0.15) and (1.25). It follows that the limit (N, g') has $z_{g'} = 0$, and hence (N, g') is flat, a contradiction. ■

With regard to higher order regularity, by [5, Thm.4.2/Rmk.4.3] there is a constant $d_o > 0$, independent of ε, such that

$$\rho_\varepsilon^{1,p}(x) \geq d_o \cdot \rho_\varepsilon(x), \tag{1.26}$$

for all $x \in \Omega_\varepsilon$; here d_o depends only on p and $\rho^{1,p}$ is the $L^{1,p}$ curvature radius. This means the metric is controlled locally in $L^{3,p}$, in harmonic coordinates, in any region where it is controlled in $L^{2,2}$, (away from the boundary). By Sobolev embedding, $L^{3,p} \subset C^{2,\alpha}$, for any $\alpha = \alpha(p) < 1$. In particular, (1.26) implies that the curvature $|r|$ satisfies

$$|r_\varepsilon|(x) \leq K \cdot \rho_\varepsilon(x)^{-2}, \tag{1.27}$$

where K depends only on the choice of c_o in (1.1). Hence for tame 3-manifolds, there is a constant $\lambda = \lambda(\Lambda) < \infty$ such that $|r_\varepsilon| \leq \lambda$ on some, (in fact any), sequence of minimizing pairs $(\Omega_\varepsilon, g_\varepsilon)$, with $\varepsilon = \varepsilon_i \to 0$. Further, again by [5, Rmk.4.3], $|\nabla^k r_\varepsilon|$ is then bounded as $\varepsilon \to 0$, for any $k < \infty$, in regions where the potential $\tau = \tau_\varepsilon < 0$ or where $s_{g_\varepsilon} > 0$. Of course these estimates come from the fact that g_ε satisfies the elliptic system (1.4)-(1.5); the main point is then to show that the estimates are independent of ε.

2. Geometrization of Tame 3-Manifolds.

In this section, we prove Theorem 0.2. Thus, throughout this section, it is assumed that M is a tame, closed oriented 3-manifold with $\sigma(M) \leq 0$, so that there is a sequence $\varepsilon = \varepsilon_i \to 0$, and a sequence of minimizing pairs $(\Omega_\varepsilon, g_\varepsilon)$, $\varepsilon = \varepsilon_i$, $\Omega_\varepsilon \subset\subset M$, such that the L^2 curvature is uniformly bounded, i.e.

$$\int_{\Omega_\varepsilon} |z_\varepsilon|^2 dV_{g_\varepsilon} \leq \Lambda, \tag{2.1}$$

for some $\Lambda < \infty$. It follows from Lemma 1.5 that

$$\rho_\varepsilon(x) \geq \rho_o, \tag{2.2}$$

for some $\rho_o = \rho(\Lambda) > 0$.

In §2.1, we prove various convergence and collapse results needed for the proof of Theorem 0.2; based on these, the proof will be completed in §2.2. A detailed discussion of the geometrization of graph manifolds from this point of view is then given in §2.3.

§2.1. We apply the L^2, (or in view of (2.2) and (1.27), the L^∞), Cheeger-Gromov theory as described in [2, §3] to the sequence $(\Omega_\varepsilon, g_\varepsilon)$, with $\varepsilon = \varepsilon_i \to 0$. There are exactly three possible behaviors for the sequence $(\Omega_\varepsilon, g_\varepsilon)$; namely a subsequence either converges, collapses or forms cusps. These cases are distinguished by the possible behaviors of the volume radius ν_ε as $\varepsilon \to 0$, in that either:

- $\nu_\varepsilon(x) \geq \nu_o, \forall x \in \Omega_\varepsilon$, for some $\nu_o > 0$.
- $\nu_\varepsilon(x) \to 0, \forall x \in \Omega_\varepsilon$, as $\varepsilon \to 0$.
- There exist points x_ε and y_ε in Ω_ε such that
$$\nu_\varepsilon(x_\varepsilon) \geq \nu_o \text{ and } \nu_\varepsilon(y_\varepsilon) \to 0, \text{ as } \varepsilon \to 0.$$

We treat each of these cases separately.

Proposition 2.1. (Convergence). *Let $(\Omega_\varepsilon, g_\varepsilon)$ be a sequence of minimizing pairs satisfying (2.1), for some sequence $\varepsilon = \varepsilon_i \to 0$. Suppose that there is a uniform constant $\nu_o > 0$ such that*

$$\nu_\varepsilon(x) \geq \nu_o > 0, \tag{2.3}$$

for all $x \in (\Omega_\varepsilon, g_\varepsilon)$. Then

$$\Omega_\varepsilon = M,$$

for all $\varepsilon > 0$ small, and a subsequence of (M, g_ε) converges to a constant curvature metric g_o on M, of scalar curvature $\sigma(M)/6$ and volume 1.

Proof: By Theorem 1.1, if $\Omega_\varepsilon \subset\subset M$ but $\Omega \neq M$, then g_ε collapses, (for ε fixed), with bounded curvature everywhere at infinity in $(\Omega_\varepsilon, g_\varepsilon)$. This implies that $\nu_\varepsilon(x) \to 0$, as x diverges to infinity in Ω_ε, contradicting (2.3). Thus (2.3) implies that $\Omega_\varepsilon = M$, and g_ε is globally defined on M.

Note also that there is a uniform bound on the diameter of (M, g_ε), i.e.

$$diam_{g_\varepsilon} M \leq D, \tag{2.4}$$

where D depends only on ν_o. This follows since $vol_{g_\varepsilon} M = 1$, and by (2.3), there is a uniform lower bound on $vol_{g_\varepsilon} B_x(\nu_o)$, for any $x \in M$. Hence, there is a uniform bound on the number of disjoint geodesic balls of radius ν_o in (M, g_ε).

Since the curvature of (M, g_ε) is uniformly bounded in L^2, it follows by the L^2 Cheeger-Gromov theory that a subsequence converges, (modulo diffeomorphisms of M), in the weak $L^{2,2}$ topology to a limit $L^{2,2}$ metric g_o defined on M, (c.f. [2,Thm.3.7]). In fact, by (2.2) and (1.26), the curvature of g_ε is uniformly bounded in $L^{1,p}$, for any fixed $p < \infty$, and hence the convergence is actually in the weak $L^{3,p}$ topology. By the compactness of the embedding $C^{2,\alpha} \subset L^{3,p}$, it follows that the convergence of g_ε to g_o is in the $C^{2,\alpha}$ topology. In particular, the limit metric g_o is at least $L^{3,p}$ smooth.

The limit metric g_o is a weak solution to the limit equations

$$L^*\tau - (\frac{1}{4}s\tau + \frac{1}{12}\sigma(M)) \cdot g = 0, \tag{2.5}$$

$$2\Delta\tau + \frac{1}{4}s\tau = \frac{1}{4}|\sigma(M)|. \tag{2.6}$$

obtained by taking the limit as $\varepsilon_i \to 0$ in (1.4)-(1.5), c.f. also (0.13)-(0.14). The equations (2.5)-(2.6) are just the Euler-Lagrange equations for a critical point of the functional $\mathcal{S}^2_-$.

By definition, i.e. from (1.7), on (M, g_ε) one has

$$\int_M \tau_\varepsilon^2 dV_\varepsilon = 1. \tag{2.7}$$

Since the local and global geometry of (M, g_ε) is uniformly bounded, standard elliptic theory applied to the trace equation (1.5), c.f. [11, §8], implies that τ_ε is uniformly bounded in $L^{1,p}, p < \infty$. Hence, the limit function τ is in $L^{1,p}$, and is non-positive since τ_ε is. In particular, (2.7) also holds for the limit function τ.

Suppose first that $\sigma(M) < 0$. We may apply the minimum principle to the trace equation (2.6) in a neighborhood of a point q realizing the minimal value of τ. Since $\tau \leq 0$ and the L^2 average of τ over M is 1, $\tau(q) \leq -1$. Since then $(s\tau)(q) \geq |\sigma(M)|$, the minimum principle implies that

$$\tau \equiv -1, s \equiv \sigma(M). \tag{2.8}$$

Alternately, (2.8) can be derived from (1.13) and (1.15). From (2.5) and the definition of L^* in (1.6), it follows that

$$-r + \frac{1}{3}\sigma(M) = 0,$$

and hence

$$z = 0. \tag{2.9}$$

Thus, (M, g_o) is a smooth unit volume metric, of constant sectional curvature $\sigma(M)/6$. The convergence to the limit is then in the C^∞ topology, by the remarks following (1.27).

Next suppose $\sigma(M) = 0$. From (1.14), one sees that the limit scalar curvature s is non-negative. Since the limit function τ is non-positive, it follows that $s \cdot \tau \equiv 0$ on the limit. Hence the trace equation (2.6) gives $\Delta\tau = 0$ on (M, g_o), so that again by the minimum principle, τ is constant. Hence, by (2.7), we must have $\tau \equiv -1, s = \sigma(M) = 0$, and (2.8) holds in this situation also. Thus (2.5) again implies that (M, g_o) is a flat 3-manifold of unit volume.

■

In case $\sigma(M) < 0$, the metric g_o is unique, (up to isometry), by the Mostow rigidity theorem [25]. Of course, if $\sigma(M) = 0$, i.e. g_o is a flat metric, one cannot expect uniqueness, since the moduli space of flat metrics on M may well be non-trivial.

Remark 2.2. It follows that under the assumption (2.3), $\inf \mathcal{S}^2_-$ is realized by a constant curvature metric g_o on M. Clearly ${I_\varepsilon}^-(g) \geq \mathcal{S}^2_-(g)$, for all $g \in \mathbb{M}$ and $\varepsilon \geq 0$, with equality if and only if $z_g = 0$. In particular,

$$ {I_\varepsilon}^-(g_o) = I_o^-(g_o) = \mathcal{S}^2_-(g_o) = |\sigma(M)|. \tag{2.10}$$

In case $\sigma(M) < 0$, it follows from (0.4) and Mostow rigidity that the constant curvature metric g_o uniquely realizes $\inf {I_\varepsilon}^-$, for all $\varepsilon \geq 0$, among metrics on M. Thus, the "family" of metrics g_ε satisying (2.3) is constant, i.e.

$$g_\varepsilon = g_o, \tag{2.11}$$

for all $\varepsilon \geq 0$. More precisely, the isometry class of the metrics g_ε is constant, in the sense that for any $\varepsilon > 0$, there is a diffeomorphism ψ_ε of M such that $\psi_\varepsilon^* g_\varepsilon = g_o$. This means of course that the sequence $(\Omega_\varepsilon, g_\varepsilon)$ in Proposition 2.1 is in fact a constant sequence, modulo diffeomorphisms. Note however that this discussion does not automatically preclude the existence of other minimizers $(\Omega_\varepsilon, g_\varepsilon)$ of I_ε^- with $\Omega_\varepsilon \subset\subset M$, for which (2.3) does not hold.

Similarly, if $\sigma(M) = 0$, the metrics g_ε satisfying (2.3) are all flat metrics on M, for all $\varepsilon \geq 0$. Since the moduli space $\mathcal{M}_F$ of flat metrics on M may be non-trivial, the metrics g_ε may not be unique. However, the volume assumption (2.3) implies that the metrics g_ε remain within a compact set of the moduli space $\mathcal{M}_F$.

Proposition 2.3. (Collapse). *Let $(\Omega_\varepsilon, g_\varepsilon)$ be a sequence of minimizing pairs satisfying (2.1), for some sequence $\varepsilon = \varepsilon_i \to 0$. Suppose that*

$$\nu_\varepsilon(x) \to 0, \tag{2.12}$$

for all $x \in (\Omega_\varepsilon, g_\varepsilon)$. Then M is a graph manifold with

$$\sigma(M) = 0.$$

Proof: Under the assumption (2.12), $(\Omega_\varepsilon, g_\varepsilon)$ becomes arbitrarily thin at every point. Recall from §1 that the complement of a sufficiently large compact set K_ε in Ω_ε admits a polarized F-structure $\mathcal{F}_\infty = \mathcal{F}_\infty(\varepsilon)$, i.e. a graph manifold structure, and that this structure extends to the rest of M, so that $V_\varepsilon = M \setminus K_\varepsilon$ admits a polarized F-structure, also called $\mathcal{F}_\infty$.

Now from Lemma 1.5, (1.27) and the Cheeger-Gromov theory [7,8], there is a $\delta_1 = \delta_1(\Lambda)$ such that if

$$\nu_\varepsilon(x) \leq \delta_1, \tag{2.13}$$

for all $x \in (\Omega_\varepsilon, g_\varepsilon)$, then Ω_ε admits a polarized F-structure $\mathcal{F}_\varepsilon$. It is clear that the F-structures $\mathcal{F}_\infty$ and $\mathcal{F}_\varepsilon$ are compatible on their intersections, and thus define a global F-structure $\mathcal{F}$ on M.

Thus, (2.12), or even the weaker condition (2.13), implies that M itself is a graph manifold. We have already noted in (1.10) that in this case, for any $\varepsilon > 0$,

$$\inf I_\varepsilon^- = \inf \mathcal{S}_-^2 = \sigma(M) = 0, \tag{2.14}$$

which proves the result. In fact, as noted in §0, any minimizing sequence for I_ε^-, with $\varepsilon > 0$ fixed, collapses M along an F-structure, (or possibly a sequence of F-structures), unless M is a flat manifold. In particular, the estimate (2.12) implies then that either

$$\Omega_\varepsilon = \emptyset, \quad \forall \varepsilon > 0, \tag{2.15}$$

or $M = \Omega_\varepsilon$ and g_ε is a collapsing sequence of flat metrics on M.

∎

Remark 2.4. We point out that this discussion shows that any closed graph manifold G necessarily satisfies $\sigma(G) \geq 0$. Hence if $\sigma(M) < 0$, then M cannot be a graph manifold.

Proposition 2.5. (Cusps I). *Suppose $\sigma(M) < 0$, and let $(\Omega_\varepsilon, g_\varepsilon)$ be a sequence of minimizing pairs satisfying (2.1), for some sequence $\varepsilon = \varepsilon_i \to 0$. Suppose that there exist x_ε and y_ε in Ω_ε such that*

$$\nu_\varepsilon(x_\varepsilon) \geq \nu_o \quad \text{and} \quad \nu_\varepsilon(y_\varepsilon) \to 0, \quad \text{as} \quad \varepsilon \to 0, \tag{2.16}$$

for some $\nu_o > 0$. Then there is a subsequence, also denoted $\{\varepsilon\}$, of $\{\varepsilon_i\}$, such that

$$\Omega_\varepsilon = \Omega$$

is independent of ε, and is given by a finite union of complete, non-compact hyperbolic manifolds of finite volume. There is an embedding

$$\Omega \subset M,$$

for which the complement $M \setminus \Omega$ is a graph manifold. The metrics g_ε converge to the complete constant curvature metric g_o, of scalar curvature $\sigma(M)/6$, and

$$vol_{g_o}\Omega = 1,$$

while collapsing the graph manifold $M \setminus \Omega$ to a lower dimensional space.

Proof: Suppose the sequence $\{(\Omega_\varepsilon, g_\varepsilon)\}, \varepsilon = \varepsilon_i$, has basepoints $\{x_\varepsilon\}$ and $\{y_\varepsilon\}$ satisfying (2.16), and consider the pointed sequence $\{(\Omega_\varepsilon, g_\varepsilon, x_\varepsilon)\}$. As in the proof of Proposition 2.1, from Lemma 1.5 and (1.27), the curvature of $(\Omega_\varepsilon, g_\varepsilon)$ is uniformly bounded, independent of ε. It follows from the L^∞ Cheeger-Gromov theory, (c.f. [2,§2] and [3,Rmk.5.5]), that there are diffeomorphisms ψ_ε of Ω_ε, and a maximal connected open domain $\Omega_o \subset\subset M$ such that a subsequence of $(\psi_\varepsilon)^* g_\varepsilon, \varepsilon = \varepsilon_i$, converges weakly in the $L^{2,p}$ topology based at x_ε, and uniformly on compact subsets, to a limit $L^{2,p}$ metric g_o on Ω_o, with

$$0 < vol_{g_o}\Omega_o \leq 1. \tag{2.17}$$

By (1.26), the convergence is actually in the weak $L^{3,p}$ and strong $C^{2,\alpha}$ topologies, and the limit metric g_o is $L^{3,p}$. The curvature of (Ω_o, g_o) is uniformly bounded in L^∞, and since $(\Omega_\varepsilon, g_\varepsilon)$ is complete for all ε, so is (Ω_o, g_o). As in Proposition 2.1, the metric g_o is a (weak) solution of the Euler-Lagrange equations (2.5)-(2.6) for $\mathcal{S}_-^2$.

Since $\sigma(M) < 0$, it follows from (1.13) that $\tau_\varepsilon \to -1$ almost everywhere on Ω_ε as $\varepsilon \to 0$. Hence the limit function $\tau = \lim \tau_\varepsilon$ satisfies

$$\tau \equiv -1, s \equiv \sigma(M), \tag{2.18}$$

as in (2.8). The same reasoning as in (2.8)-(2.9) shows that

$$z_{g_o} = 0, \tag{2.19}$$

so that (Ω_o, g_o) is a complete, non-compact manifold of constant sectional curvature $\sigma(M)/6$ and volume at most 1.

We now fix the subsequence of $g_\varepsilon, \varepsilon = \varepsilon_i$ above, but do not change the notation for the subsequence. Suppose there is another sequence of points $\{x_\varepsilon^1\}$ in Ω_ε, with

$$\nu_\varepsilon(x_\varepsilon^1) \geq \nu_o, \tag{2.20}$$

for some constant $\nu_o > 0$, and $\{x_\varepsilon^1\}$ does not (sub)-converge to a point in Ω_o. Then one may repeat the above process to obtain another smooth maximal open domain $\Omega_1 \subset\subset M$, with smooth complete metric g_o, which again is of constant curvature, and scalar curvature $s = \sigma(M) < 0$.

This process may be repeated as many times as necessary until there are no sequences left satisfying (2.20), and not subconverging to a previously defined domain. There is clearly a uniform upper bound Q on the number of domains $\Omega_j, j = 0, 1, ..., Q$, obtained in this way, since $vol\Omega_j \geq 10^{-1}\nu_o^3$ and the total volume of $\Omega = \cup\Omega_j$ is at most 1. In fact, the bound Q is independent of the choice of ν_o in (2.20) or (2.16), and depends only on the 3-manifold M. For if (U, g_o) is any complete, constant curvature, open 3-manifold, with scalar curvature $\sigma(M)$, then

$$vol_{g_o}U \geq v_o > 0, \tag{2.21}$$

where v_o depends only on an upper bound for $|\sigma(M)|; v_o$ is essentially the Margulis constant, c.f. [36] or [37]. Thus, choosing ν_o above small, i.e. $\nu_o << v_o$, it follows that there is a fixed bound $Q = Q(M)$ on the number of domains Ω_j obtained by the process above.

Let

$$\Omega = \bigcup_o^Q \Omega_j. \tag{2.22}$$

Then (Ω, g_o) is a collection of a finite number of connected, noncompact, complete Riemannian manifolds, each with constant curvature metric g_o with $s_{g_o} = \sigma(M) < 0$. In particular, each component Ω_j is of finite topological type, having a finite number of ends, i.e. cusps, each diffeomorphic to $T^2 \times \mathbb{R}^+$. It follows that not only is each component Ω_j weakly embedded in M, but Ω_j is actually embedded in M as an open domain. Consequently, there is an embedding $\Omega \subset M$. The complement $M \setminus \Omega$ is a compact manifold with boundary consisting of a finite number of tori, which has a sequence of F-structures $\mathcal{F}_\varepsilon$ along which a subsequence of the metrics $g_\varepsilon, \varepsilon = \varepsilon_i$, collapse $M \setminus \Omega$ to a lower dimensional space. In particular, $M \setminus \Omega$ is a graph manifold with toral boundary.

By the construction of Ω,

$$vol_{g_o}\Omega \leq 1. \tag{2.23}$$

We claim that

$$vol_{g_o}\Omega = 1. \tag{2.24}$$

The proof of this is exactly the same as the proof of (1.9), given in [2, Thm. 5.7]. Briefly, suppose $vol_{g_o}\Omega < 1$, so that $vol_{g_o}\Omega \leq 1 - \mu$, for some $\mu > 0$. Using the structure of graph manifolds, for any $\delta > 0$, one may construct a metric $\bar{g} = \bar{g}_\delta$ on M, agreeing with g_o on a prescribed sufficiently large compact set $K \subset \Omega$, such that $vol_{\bar{g}} M \setminus K \leq \delta$ and $|\bar{s}| \leq C$ on $M \setminus K$, where C is independent of K and δ. Thus, choosing δ much smaller than μ implies that

$$v(\bar{g})^{1/3}\int_M (s_{\bar{g}}^-)^2 dV_{\bar{g}} < v(g_o)^{1/3}\int_\Omega (s_{g_o}^-)^2 dV_{g_o} = \sigma(M)^2,$$

which contradicts (0.4). Finally, the arguments proving (2.11) in Remark 2.2 also prove that $\Omega = \Omega_\varepsilon$ and $g_o = g_\varepsilon$ in this situation, by means of the Mostow-Prasad rigidity theorem [27]. This completes the proof.

∎

Remark 2.6. As in Remark 2.2, one sees that $\inf I_\varepsilon{}^- = \inf \mathcal{S}_-^2 = |\sigma(M)|$ is realized by the union of the complete constant negative curvature metrics g_o on $\Omega \subset M$. Apriori however, as before, it may be possible that (Ω, g_o) is not unique; this will be discussed further in §2.2.

We now turn to the analogue of Proposition 2.5 in case $\sigma(M) = 0$.

Proposition 2.7. (Cusps II). *Suppose $\sigma(M) = 0$ and let $(\Omega_\varepsilon, g_\varepsilon)$ be a sequence of minimizing pairs satisfying (2.1), for some sequence $\varepsilon = \varepsilon_i \to 0$. Then $(\Omega_\varepsilon, g_\varepsilon)$ cannot form cusps, i.e. either there exists $\nu_o > 0$ such that*

$$\nu_\varepsilon(x) \geq \nu_o > 0, \quad \text{for all} \quad x \in \Omega_\varepsilon,$$

or

$$\nu_\varepsilon(x) \to 0, \quad \text{for all} \quad x \in \Omega_\varepsilon.$$

Proof: The proof is by contradiction and so we assume $(\Omega_\varepsilon, g_\varepsilon)$ does form cusps as $\varepsilon = \varepsilon_i \to 0$, i.e. there are base points $x_\varepsilon, y_\varepsilon$ satisfying (2.16). The same arguments as in the proof of Proposition 2.5 prove that a subsequence of $\{(\Omega_\varepsilon, g_\varepsilon, x_\varepsilon)\}$ converges to a complete connected maximal limit (Ω_o, g_o, x) with

$$\int_{\Omega_o} |z|^2 dV_o \leq \Lambda, \quad vol_{g_o}\Omega_o \leq 1. \tag{2.25}$$

By (1.14), $s_{g_o} \geq 0$ everywhere and by (1.27), g_o has uniformly bounded curvature. Since (Ω_o, g_o) is of finite volume, it collapses everywhere along an F-structure at infinity. In particular, a neighborhood of infinity of Ω_o is a graph manifold. Further, by continuity, (Ω_o, g_o) minimizes the L^2 norm of z among all compact perturbations of g_o with non-negative scalar curvature and with volume at most that of g_o.

As in the proof of Proposition 2.5, this process may be repeated for any other sequence of base points $\{x_\varepsilon^j\}$ satisfying (2.16) giving rise to disjoint cusps Ω_j in the limit. Each such Ω_j satisfies (2.25) with the same Λ. Again however this process terminates after a finite number Q of steps. This is because if j is sufficiently large, then Ω_j satisfies (2.25) and $vol\Omega_j \leq \delta_1$, where $\delta_1 = \delta_1(j)$ may be made arbitrarily small by choosing j sufficiently large. However, such a manifold is a graph manifold by (2.13). This implies that the minimizing pairs $(\Omega_\varepsilon, g_\varepsilon), \varepsilon = \varepsilon_i$, will have collapsed Ω_j for ε sufficiently small, i.e. in effect $\Omega_j = \emptyset$, as in (2.15). Thus, as in (2.22), the maximal domain $\Omega = \cup\Omega_j$ consists of a finite number of connected components. For similar reasons, the same arguments proving (2.24), proves that

$$vol_{g_o}\Omega = 1. \tag{2.26}$$

As in Proposition 2.1, the limit metric g_o and limit potential $\tau = \lim \tau_\varepsilon$ are a solution of the Euler-Lagrange equations (2.5)-(2.6) with $s \geq 0$. Since the region where $s > 0$ is disjoint from the region where $\tau < 0$, $s \cdot \tau \equiv 0$, and hence the limit function τ is harmonic on (Ω, g_o). By (2.7), τ is in $L^2(\Omega, g_o)$. This is easily seen to imply, by a standard cutoff argument, that τ is constant on each component Ω_j of Ω; (this can also be seen from the estimate (1.15)).

If the limit potential $\tau \neq 0$ on some Ω_j, then as in (2.8)-(2.9), again one has $z = 0$, and hence the complete limit (Ω_j, g_o) is flat. However, any complete non-compact flat manifold must have infinite volume, contradicting (2.26). Thus, such cusps cannot form.

It follows we must have $\tau \equiv 0$ on the limit (Ω, g_o). Although this situation is very special and unlikely to occur, it will take some further arguments to rule it out. (In case Ω_j is irreducible, for some j, these arguments can be bypassed, c.f. Remark 2.8 below).

Recall that the L^2 norm of τ_ε on $(\Omega_\varepsilon, g_\varepsilon)$ is 1 by (2.7). Hence in this situation, all of τ_ε is concentrating at infinity in Ω_ε as $\varepsilon \to 0$, and so by (2.26), concentrating on a set of measure converging to 0. In particular, τ_ε must be unbounded in this region, so that $T_\varepsilon = \sup|\tau_\varepsilon| \to \infty$, as $\varepsilon \to 0$. By (1.11), this forces

$$\frac{\varepsilon}{\sigma} \to \infty, \quad \text{as } \varepsilon \to 0, \tag{2.27}$$

where $\sigma = \sigma(g_\varepsilon) = \mathcal{S}_-^2(g_\varepsilon)$. It is then natural to consider the renormalized functional

$$\frac{1}{\varepsilon}I_\varepsilon{}^- = \mathcal{Z}^2 + \frac{1}{\varepsilon}\mathcal{S}_-^2 \tag{2.28}$$

on $(\Omega_\varepsilon, g_\varepsilon)$, together with the renormalized Euler-Lagrange equations (1.4)-(1.5), i.e.

$$\nabla\mathcal{Z}^2 + L^*\bar{\tau} + \bar{\phi} \cdot g = 0, \tag{2.29}$$

$$2\Delta(\bar{\tau} + \frac{s}{12}) + \frac{1}{4}s\bar{\tau} = -\frac{1}{2}|z|^2 + 3\bar{c}, \tag{2.30}$$

where $\bar{\tau} = \tau/\varepsilon, \bar{c} = c/\varepsilon$ and $\bar{\phi} = \phi/\varepsilon$. By (2.27) and (2.26), $\bar{c}$ converges to $\frac{1}{6}\int_\Omega |z|^2 dV_o \leq \Lambda$ as $\varepsilon \to 0$, and by the regularity following (1.27), $\nabla\mathcal{Z}^2$, $|z|^2$ and s are uniformly bounded as $\varepsilon \to 0$.

We divide the discussion now into three cases, according to the behavior of $\bar{\tau} = \bar{\tau}_\varepsilon$ as $\varepsilon \to 0$, on a sequence of points $y_\varepsilon \in \Omega_\varepsilon$ converging to a limit point $y \in \Omega$. These behaviors are either $\bar{\tau}(y_\varepsilon) \to -\infty$, $\bar{\tau}(y_\varepsilon)$ remains bounded, or $\bar{\tau}$ tends to 0 in neighborhoods of y_ε as $\varepsilon \to 0$. (This discussion closely resembles, at least formally, that in [5, §4.1]). Let Ω_y be the component of Ω containing y.

Case (i). Suppose $\bar{\tau}(y_\varepsilon) \to -\infty$ as $\varepsilon \to 0$. In this case, divide the equations (2.29)-(2.30) further by $|\bar{\tau}(y_\varepsilon)|$. Let $\widetilde{\tau} = \widetilde{\tau}_\varepsilon = \bar{\tau}/|\bar{\tau}(y_\varepsilon)|$. Since the right hand side of (2.30) is uniformly bounded, elliptic regularity, c.f. [11, Ch.8] implies that $\widetilde{\tau}_\varepsilon$ is uniformly bounded at points a bounded distance to the base points y_ε. Hence, a subsequence converges to a limit function $\widetilde{\tau}$. The limit (renormalized) equations on (Ω_y, g_o) are then

$$L^*\widetilde{\tau} = 0, \quad \Delta\widetilde{\tau} = 0, \tag{2.31}$$

i.e. the static vacuum Einstein equations, (c.f. [4]). Since $\widetilde{\tau} \leq 0$, the maximum principle implies that $\widetilde{\tau} < 0$ everywhere. Hence (Ω_y, g_o, y) is a complete solution to the static vacuum equations, with non-vanishing potential. By [4, Thm.3.2], it follows that (Ω_y, g_o) is flat, which gives a contradiction as before.

Case (ii). Suppose $\bar{\tau}(y_\varepsilon)$ is bounded as $\varepsilon \to 0$. The limit equations then take the form

$$\nabla\mathcal{Z}^2 + L^*\bar{\tau} + \bar{c}\cdot g = 0, \tag{2.32}$$

$$2\Delta(\bar{\tau} + \frac{s}{12}) = -\frac{1}{2}|z|^2 + 3\bar{c}, \tag{2.33}$$

with $\bar{c} = \mathcal{Z}^2(g_o)/6$.

Now on the one hand, the metric g_ε is a critical point of the functional $\frac{1}{\varepsilon}I_\varepsilon{}^-$. On the other hand, the second term in (2.28) converges to 0 on g_ε, by (2.27). Thus let η be a positive cutoff function on Ω_y, with $\eta \equiv 1$ on $B_y(R)$, $\eta \equiv 0$ on $\Omega_y \setminus B_y(2R)$, with $|d\eta| \leq c/R$, and consider for instance the variation of g_ε given by

$$g_{\varepsilon,t} = g_\varepsilon - t\eta r_\varepsilon.$$

Note that since r_ε is uniformly bounded as $\varepsilon \to 0$, the metrics $g_{\varepsilon,t}$ are well defined for all $t \leq t_o = t_o(R)$, independent of ε.

A straightforward computation shows that $\frac{d}{dt}\mathcal{S}_-^2(g_{\varepsilon,t}) \leq 0$, for R sufficiently large. (Alternately, if $\frac{d}{dt}\mathcal{S}_-^2(g_{\varepsilon,t})$ were positive, replace $g_{\varepsilon,t}$ by $g_\varepsilon + t\eta r$ in the ensuing argument). Hence

$$0 \leq \frac{1}{\varepsilon}\mathcal{S}_-^2(g_{\varepsilon,t}) \leq \frac{1}{\varepsilon}\mathcal{S}_-^2(g_\varepsilon),$$

for all t small, and so both terms converge to 0 as $\varepsilon \to 0$ by (2.27). Since the metrics converge smoothly to the limit, it follows that

$$\lim_{\varepsilon\to 0}\frac{d}{dt}\mathcal{S}_-^2(g_{\varepsilon,t})|_{t=0} = 0. \tag{2.34}$$

Thus, on the limit (Ω_y, g_o), one has

$$0 = \int_{\Omega_y} < L^*\bar{\tau}, \eta z > = \int_{\Omega_y} < D^2\bar{\tau}, \eta z > - \int_{\Omega_y} \eta\bar{\tau}|z|^2. \tag{2.35}$$

We note that $\int_{\Omega_y} < D^2\bar{\tau}, \eta z > \to 0$ as $R \to \infty$. This follows by integrating by parts, i.e. applying the divergence theorem, twice, together with the Bianchi identity and the fact that $s \cdot \bar{\tau} \equiv 0$

Hence, (2.35) implies that either $z \equiv 0$, in which case (Ω_y, g_o) is flat, giving a contradiction as before, or the limit $\bar{\tau} \equiv 0$. This is treated in the last case.

Case (iii). Suppose $\bar{\tau} \equiv 0$. In this case, the limit equations are

$$\nabla\mathcal{Z}^2 + \bar{c}\cdot g = 0, \tag{2.36}$$

$$\frac{1}{6}\Delta s + \frac{1}{2}|z|^2 = 3\bar{c}. \tag{2.37}$$

We will prove below that necessarily $\bar{c} = 0$. This then implies that $\int_{\Omega_y}|z|^2 dV_o = 0$, so that (Ω_y, g_o) is flat, and one has a contradiction as before.

Now (Ω, g_o) is of finite volume and collapsing at infinity. Hence, for any divergent sequence $\{x_i\} \in \Omega_y$ and any $R < \infty$, the pointed sequence $(B_{x_i}(R), g_o, x_i)$ in (Ω_y, g_o) collapses with bounded curvature along an injective F-structure. We may then unwrap this collapse by passing to sufficiently large finite covers; choosing a sequence $R_j \to \infty$, and a suitable diagonal subsequence gives rise to a complete limit solution (N, g_∞, x_∞) of (2.36)-(2.37) with a free isometric S^1 action. The convergence to such limits is smooth, by elliptic regularity applied to the equations (2.36)-(2.37), (c.f. [4, §3]). The constant $\bar{c}$ of course remains the same in passing to this geometric limit. Thus it suffices to evaluate $\bar{c}$ on these simpler manifolds (N, g_∞).

Let V be the orbit space of the S^1 action, so that V is a complete Riemannian surface. Let $f : V \to \mathbb{R}$ denote the length of the S^1 fibers and A the curvature form of the S^1 bundle. In [4, Prop. 4.1] it is proved, (via a Gauss-Bonnet type argument), that

$$\int_V |\nabla \log f|^2 < \infty, \text{ and } \int_V |A|^2 < \infty.$$

Suppose first that there exists $v_o > 0$ and a sequence of points $p_i \in V$ such that $areaD_{p_i}(1) \geq v_o$, where $D_{p_i}(1)$ is the geodesic disc about p_i of radius 1 in V. Then $\nabla \log f \to 0$ and $A \to 0$ in $D_{p_i}(R)$ as $i \to \infty$, for any fixed R. This implies that any geometric limit $(N', g'_\infty, p_\infty)$ of (N, g_∞, p_i) is a complete product metric of the form $V' \times S^1$, of non-negative scalar curvature, and satisfying (2.36)-(2.37), again with the same $\bar{c}$. It follows that the Gauss curvature of V' is non-negative and hence there are points $q_i \in V'$ such that the metric g'_∞ on $D_{q_i}(1) \subset V'$ converges to the flat metric, (in covers if necessary). This of course implies $\bar{c} = 0$, as required.

On the other hand, suppose that $areaD_{p_i}(1) \to 0$ for any divergent sequence p_i in V, so that V itself collapses everywhere at infinity. Then repeating the construction above on a divergent sequence, unwrapping the collapse as above, gives rise to a further complete limit $(N', g'_\infty, x'_\infty)$ which is still a solution of (2.36)-(2.37), and which now has a free isometric $S^1 \times S^1$ action. The second S^1 action arises from the unwrapping of the collapse of V at infinity. However, for instance by [15, Thm.8.4], the only complete metrics of non-negative scalar curvature with such an action are flat. Hence again $\bar{c} = 0$, as claimed.

Thus in all cases the existence of cusps leads to a contradiction, which proves the result.

■

Remark 2.8. A much simpler proof of Proposition 2.7 is possible if it is assumed that some (non-empty) component Ω_o of Ω is irreducible. Namely, there exists an exhaustion of Ω_o by compact sets K_j, such that $\Omega_o \setminus K_j$ is a graph manifold and ∂K_j is a collection of tori. If any such torus is incompressible in Ω_o, then as above, [15, Thm.8.4] implies that Ω_o is flat, and one has a contradiction as before. Thus, all such tori must be compressible. If now Ω_o is irreducible, then standard arguments in 3-manifold topology, (c.f. also the proof of Theorem 2.9 below), imply that Ω_o must be a solid torus $D^2 \times S^1$. But this means that Ω_o is a graph manifold, and so the discussion in (2.14)-(2.15) holds. This means that Ω_o is either empty or flat, either of which is a contradiction.

With some further topological arguments, this argument can be extended to the situation where it is assumed that M is irreducible in place of Ω_o.

§2.2. In this subsection, we assemble the results of §2.1 to complete the proof of Theorem 0.2.

Proof of Theorem 0.2: $\sigma(M) < 0$.

Suppose M is a closed, oriented tame 3-manifold with

$$\sigma(M) < 0.$$

Again, let $\{(\Omega_\varepsilon, g_\varepsilon)\}, \varepsilon = \varepsilon_i$ be a sequence of minimizers of $I_\varepsilon^{\,-}$ satisfying (2.1), so that as discussed in §2.1, a subsequence of $\{g_\varepsilon\}$ either converges, collapses, or forms cusps. By Proposition 2.3, the condition $\sigma(M) < 0$ implies that no subsequence of $\{g_\varepsilon\}$ can collapse.

If a subsequence of $\{g_\varepsilon\}$ converges, (modulo diffeomorphisms of M), then Proposition 2.1 shows it converges to a unit volume constant curvature metric g_o on M, with scalar curvature $s_{g_o} = \sigma(M)$. In particular, (by rescaling g_o by the factor $\sigma(M)/6), M$ admits a hyperbolic structure, and if

$vol_{-1}M$ is the volume of M in the hyperbolic metric, (of constant curvature -1), then

$$|\sigma(M)| = 6(vol_{-1}M)^{2/3}, \tag{2.38}$$

giving (0.9).

If a subsequence of $\{g_\varepsilon\}$ forms cusps, then by Proposition 2.5, there is a maximal open set $\Omega \subset M$, consisting of a bounded number of components, on which the subsequence converges, (modulo diffeomorphisms of M), to a limit metric g_o. The metric g_o is a complete, constant curvature metric satisfying

$$s_{g_o} = \sigma(M), \quad vol_{g_o}\Omega = 1, \tag{2.39}$$

so that the pair (Ω, g_o) realizes the Sigma constant of M in this generalized sense. In particular, as in (2.38), this proves (0.8).

The open manifold Ω has a finite number of components Ω_i; each Ω_i has a finite number of ends, each diffeomorphic to $T^2 \times \mathbb{R}^+$. The domain Ω embeds as an open domain in M, and the complement $G = M \setminus \Omega$ has the structure of a manifold with boundary, with a finite number of components G_j. Each G_j is a graph manifold, with a finite number of boundary components, each diffeomorphic to T^2. Thus the decomposition

$$M = \Omega \cup G \tag{2.40}$$

gives a decomposition of M into hyperbolic and graph manifold regions, as in (0.7).

With the above understood, it remains to prove that each torus T^2 in a hyperbolic cusp $T^2 \times \mathbb{R}^+ \subset \Omega$ is incompressible in M. This is of course a crucial issue, since without it the geometric decomposition (2.40) may have no topological significance. In the same vein, one also needs to establish the uniqueness of the decomposition of M into Ω and $M \setminus \Omega$. Otherwise, some subsequences of $(\Omega_\varepsilon, g_\varepsilon)$ may converge everywhere, while others may form cusps. We first prove the incompressibility of the tori; the uniqueness then follows easily after this.

In the following, we assume for simplicity that the metric g_o above is scaled to give a hyperbolic metric, i.e. a metric of constant sectional curvature -1, on Ω.

Theorem 2.9. *Each torus T^2 in a hyperbolic cusp $T^2 \times \mathbb{R}^+ \subset \Omega$ is incompressible in M.*

Proof: Let $T = T^2$ and suppose D is a compressing disc in M, with $\partial D \subset T$. By Dehn's Lemma, c.f. [18], we may assume that D is embedded. Since T is incompressible in Ω, D intersects G, and we may assume that $D \subset G$. Let G' be the component of G containing D. Perturb D to obtain another disjoint, isotopic disc $D', \partial D' \subset T$, so that $\partial D \cup \partial D'$ bounds a small annulus A in T. Note that $A \cup D \cup D'$ is the boundary of a small 3-ball in M, while the surface $(T \setminus A) \cup (D \cup D')$ is an embedded 2-sphere $S^2 \subset G'$.

If G' is irreducible, or more precisely if the S^2 above bounds a 3-ball in G', then it is clear that T is the boundary of a solid torus $D^2 \times S^1$, (c.f. [19, II.2.4]). In this case, the component G' is $D^2 \times S^1$.

If the S^2 above does not bound a 3-ball in G', then we may write G' as

$$G' = G_1 \# ... \# G_q, \tag{2.41}$$

where each G_i is irreducible. By definition of connected sum, one factor in (2.41), say G_1, is then a solid torus $D^2 \times S^1$ glued onto T. Thus, one has

$$G' = (D^2 \times S^1) \# G'', \tag{2.42}$$

where G'' is a graph manifold. The factor G'' may either be closed or have non-empty boundary; in the latter case, $\partial G''$ again a union of tori, each contained in distinct ends of the hyperbolic part Ω of M.

It follows that part of the original manifold M is obtained by glueing on a solid torus, i.e. performing a Dehn surgery, to the torus T in a cusp of Ω, and possibly adding on other graph manifold components by connected sum.

We first prove Theorem 2.9 in the case $G'' = \emptyset$ in (2.42), so that G' is a solid torus. The general case is then easily obtained from this.

The idea here is to metrically glue on a solid torus explicitly and in such a way as to decrease $\mathcal{S}_{-}^{2}$ a definite amount below the value $|\sigma(M)|$, which of course gives a contradiction to (0.4). This turns out to be possible because the most natural metrics on a solid torus are of positive scalar curvature, and so contribute nothing to the value of $\mathcal{S}_{-}^{2}$.

To begin, choose an $S^1 \times S^1$ product structure for T^2 in the hyperbolic cusp $T^2 \times \mathbb{R}^+$. This choice is of course not unique; it can be changed by an automorphism of T^2, i.e. element of $SL(2, \mathbb{Z})$. Thus, we may choose a basis $S^1 \times S^1$, so that the Dehn surgery glues on a disc D^2 onto the first factor, while the second factor is left fixed.

In this basis, the flat metric on T^2 may be written as

$$d_1^2 d\theta_1^2 + d_2^2 d\theta_2^2 + 2ad_1 d_2 d\theta_1 d\theta_2, \tag{2.43}$$

where $a = \cos\alpha$, α the angle between the two S^1 factors, and d_1, d_2 are constants. The hyperbolic metric on the cusp is then given by

$$dt^2 + e^{-2t}(d_1^2 d\theta_1^2 + d_2^2 d\theta_2^2 + 2ad_1 d_2 d\theta_1 d\theta_2). \tag{2.44}$$

Since we are only concerned with the end behavior, assume (2.44) to hold for $t \geq t_o$, for t_o a free large parameter, and assume the torus T corresponds to the slice at t_o. Now metrically glue on a disc D^2 to the first S^1 factor, with metric of the form,

$$dr^2 + f_1^2 d\theta_1^2 + f_2^2 d\theta_2^2 + 2af_1 f_2 d\theta_1 d\theta_2. \tag{2.45}$$

Here, we choose functions f_1, f_2 depending only on r, where $r \in [0, \frac{\pi}{2}]$. For the moment, it is required that the metric is piecewise smooth (at least C^2), and is C^1 at the seam where the disc is glued to T. Comparing the forms (2.44) and (2.45), this amounts to the requirement that f_1, f_2 satisfy the boundary conditions

$$f_1(0) = 0,\ f_1'(0)(1-a^2)^{1/2} = 1,\ f_1(\frac{\pi}{2}) = d_1 e^{-t_o},\ f_1'(\frac{\pi}{2}) = d_1 e^{-t_o}, \tag{2.46}$$

$$f_2(0) = a > 0,\ f_2'(0) = 0,\ f_2(\frac{\pi}{2}) = d_2 e^{-t_o},\ f_2'(\frac{\pi}{2}) = d_2 e^{-t_o}, \tag{2.47}$$

It is simpler for computations to follow to orthogonalize the basis θ_1, θ_2. Thus, set

$$d\bar{\theta}_2 = d\theta_2 + a\frac{f_1}{f_2}d\theta_1.$$

A simple substitution shows that the metric (2.45) may be rewritten as

$$dr^2 + f_1^2(1-a^2)d\theta_1^2 + f_2^2 d\bar{\theta}_2^2. \tag{2.48}$$

In the following, specific warping functions f_1 and f_2 are chosen to satisfy (2.46)-(2.47), giving a specific metric glueing of D^2 onto S^1. We are not interested in any optimal choices, and so just choose specific forms for f_1 and f_2 that suffice for the needs of the argument.

As a first approximation to the glueing, set

$$f_1(r) = c_1 \tan\frac{r}{2},\quad f_2(r) = c_2 e^{-\cos r}, \tag{2.49}$$

where $c_1 = d_1 e^{-t_o}, c_2 = d_2 e^{-t_o}$. A simple calculation shows that this metric is C^1 at $t = t_o$, i.e. where $r = \frac{\pi}{2}$. Further, f_2 satisfies (2.47) and f_1 satisfies $f_1(0) = 0$, but $f_1'(0) \cdot (1-a^2)^{1/2} = \frac{1}{2}c_1(1-a^2)^{1/2} \neq 1$. Note that $c_1 << 1$ for t_o sufficiently large. In other words, the glueing of $D^2 \times S^1$ is a C^o glueing, but not C^1. The glued solid torus is metrically a cone manifold, with cone angle $\pi c_1(1-a^2)^{1/2}$

along the core geodesic $\gamma = \{r = 0\}$. In particular, there is a concentration of positive curvature along γ, for t_o large. This cone singularity along the core curve will be smoothed later.

For metrics of the form (2.48), it is easily computed that the scalar curvature is given by

$$\frac{1}{2}s = -\frac{f_1''}{f_1} - \frac{f_2''}{f_2} - \frac{f_1'f_2'}{f_1f_2}. \tag{2.50}$$

For the choices (2.49), one has

$$\frac{f_1''}{f_1} = \frac{1}{2}(\cos(r/2))^{-2}, \quad \frac{f_1'}{f_1} = (\sin r)^{-1},$$

$$\frac{f_2''}{f_2} = \cos r + \sin^2 r, \quad \frac{f_2'}{f_2} = \sin r,$$

which gives

$$\frac{1}{2}s = -1 - \cos r - \sin^2 r - \frac{1}{2}\cos^{-2}(r/2).$$

An exercise in calculus gives the bound

$$s \geq -6, \tag{2.51}$$

on the full solid torus $r^{-1}[0, \frac{\pi}{2}]$, with $s > -6$ on $r^{-1}[0, \frac{\pi}{2})$. It is worth pointing out that this metric does not have sectional curvature $K \geq -1$ everywhere.

Next, we estimate the volume of the glueing, compared with the volume of the hyperbolic cusp. First, the volume of the hyperbolic cusp, cut off at $t = t_o$ is given by

$$V_C(t_o, \infty) = d_1 d_2 \int_{t_o}^{\infty} e^{-2t}dt = \frac{1}{2}d_1 d_2 e^{-2t_o}. \tag{2.52}$$

On the other hand the volume of the solid torus glued in at $t = t_o$ is given by

$$V_D(0, \frac{\pi}{2}) = \int_0^{\pi/2} f_1 f_2 dr = d_1 d_2 e^{-2t_o} \int_0^{\pi/2} e^{-\cos r}\tan(r/2)dr,$$

and, for instance, a numerical evaluation shows that

$$\int_0^{\pi/2} e^{-\cos r}\tan(r/2)dr < .464 < \frac{1}{2}. \tag{2.53}$$

Thus, the volume of the glued solid torus is less than the volume of the hyperbolic cusp.

We now turn to the smoothing of the cone singularity along the core curve. Since there is a concentration of positive curvature at the core curve, we will smooth it to have positive scalar curvature nearby.

In detail, the function f_2 satisfies the boundary condition (2.47) at $r = 0$, and so need not be changed. Now recall that $f_1(0) = 0$, while

$$f_1'(0) = c_1/2 << (1 - a^2)^{-1/2}. \tag{2.54}$$

The latter estimate follows since we are free to choose c_1 arbitrarily small, by going sufficiently far down the hyperbolic cusp. Given then a fixed small choice of c_1, choose $r_o << c_1$. One may then bend f_1 on the interval $[r_o/2, r_o]$, to a new function $\bar{f}_1$ satisfying the boundary conditions (2.46) at $r_o/2$ in place of 0, i.e.

$$\bar{f}_1(r_o/2) = 0, \bar{f}_1'(r_o/2) = (1 - a^2)^{-1/2}, \tag{2.55}$$

while $\bar{f}_1$ agrees C^1 with f_1 at the value r_o. This bending then obviously has the property that

$$\frac{\bar{f}_1''}{\bar{f}_1} << 0. \tag{2.56}$$

One may then readjust f_2 so that it satisfies the boundary conditions (2.47) at $r_o/2$ in place of 0. This requires only a small C^2 perturbation of f_2, so that its contribution to the curvature only changes slightly. Together with (2.50) and (2.51), (2.56) implies that the scalar curvature of the resulting metric is (very) positive in the small solid torus given by the region $r^{-1}[r_o/2, r_o]$. Obviously the change to the volume by this bending can be made arbitrarily small.

It follows that the metric on $D^2 \times S^1 \cup T^2 \times \mathbb{R}^+$, given by (2.45) on $[r_o/2, \frac{\pi}{2}]$, for the above choices for f_1 and f_2, and by (2.44) on $[t_o, t_o+1]$ is a C^1 smooth metric, which is C^∞ off the seams at $t_o = \{r = \frac{\pi}{2}\}$, and r_o. This metric has smaller volume than the volume of the hyperbolic cusp and has scalar curvature $s \geq -6$ everywhere, except at the seams, where s is not defined. Note that s is a piecewise smooth function, with only jump discontinuities at the seams t_o, r_o, and that $s > -6$ for $r < \frac{\pi}{2}$. We may take a smooth approximation $\bar{g}$ to this metric by smoothing the warping functions. The scalar curvature $\bar{s}$ of $\bar{g}$ then interpolates the values of s, i.e. smooths the jump discontinuity. Clearly, one may thus choose a smoothing so that $\bar{s} \geq -6$ everywhere.

The metric $\bar{g}$ is a complete smooth metric on a domain $\bar{M} \subset M$ satisfying

$$s_{\bar{g}} \geq -6, \quad \text{and} \quad vol_{\bar{g}}\bar{M} < vol_{g_o}\Omega. \tag{2.57}$$

In case $\bar{M} \neq M$, so that there are graph manifold components of M remaining in $M \setminus \bar{M}$, one may use [2, Thm 5.7] as above in the proof of (2.24), to obtain a metric $\widetilde{g}$ on M satisfying

$$s_{\widetilde{g}} \geq -6 \text{ on } M \setminus K, \ |s_{\widetilde{g}}| \leq C, \ vol_{\widetilde{g}}M \setminus K \leq \delta, \tag{2.58}$$

where K is an arbitrarily prescribed compact domain in $\bar{M}, C$ is independent of K, and δ may be made arbitrarily small by choosing K sufficiently large.

Now from (2.39), $\sigma(M)$ is realized by the complete hyperbolic metric g_o on $\Omega \subset M$. It follows from the last estimates in (2.57) and (2.58) that, for δ sufficiently small,

$$vol_{\widetilde{g}}M < vol_{g_o}\Omega, \tag{2.59}$$

while by the remaining estimates in (2.57)-(2.58),

$$\int_M (s_{\widetilde{g}}^-)^2 dV_{\widetilde{g}} < \int_\Omega s_{g_o}^2 dV_g = \int_\Omega (s_{g_o}^-)^2 dV_g. \tag{2.60}$$

This implies that

$$\mathcal{S}_-^2(\widetilde{g}) < |\sigma(M)|, \tag{2.61}$$

which is of course impossible by (0.4).

This completes the proof of Theorem 2.9 in case G' is a solid torus. Now suppose that the component G' of G is of the form

$$G' = (D^2 \times S^1)\#G'', \tag{2.62}$$

with G'' a graph manifold.

Suppose first that G'' is closed. Then one has the decomposition

$$M = M'\#G'', \quad \text{where} \quad M' = (M \setminus G') \cup_{T^2} (D^2 \times S^1),$$

and M' is closed. As above, construct the metric $\widetilde{g}$ on the closed manifold M', where M' is obtained by performing Dehn surgery on the given T^2. As in (2.61), this gives

$$\mathcal{S}_-^2(\widetilde{g}) < |\sigma(M')|. \tag{2.63}$$

If however G'' is not closed, then one may write

$$G'' = \#_1^k(D^2 \times S^1)\#G''',$$

where G''' is a closed graph manifold and G'' has $k \geq 1$ toral boundary components. Each such torus is (isotopic to) a torus in a hyperbolic cusp in Ω. Hence,

$$M = M'\#G''',$$

where M' is the closed manifold obtained from $M \setminus G'$ by glueing on solid tori to its boundary components. Thus, in the same manner as before, construct the metric $\widetilde{g}$ on M' satisfying (2.63).

Now since G'' or G''' is a closed graph manifold, it admits metrics $h = h_\delta$ with arbitrarily small volume, and arbitrarily small L^2 norm of scalar curvature. In particular, $\mathcal{S}^2_-(h_\delta) \leq \delta$, for any prescribed $\delta > 0$. As discussed in [2,§7], [5,§6.1], or also in Case (i) of Theorem 2.14 below, one may then form a metric g^* on the sum $M'\#G''$, (or $M'\#G'''$), agreeing with $\widetilde{g}$ outside a very small ball in M', and with h_δ outside a very small ball in G'', and so that on the neck $S^2 \times I, g^*$ has uniformly bounded scalar curvature and arbitrarily small volume. In fact one may choose $g^*|_{S^2\times I}$ to be a truncation of the isometrically doubled Schwarzschild metric g_S in (1.23) with mass $m << \delta$. Thus we have constructed metrics g^* on M satisfying

$$\mathcal{S}^2_-(g^*) \leq \mathcal{S}^2_-(\widetilde{g}) + \delta, \tag{2.64}$$

for any given $\delta > 0$; of course g^* depends on δ. Choosing δ sufficiently small, from (2.63) one again obtains

$$\mathcal{S}^2_-(g^*) < |\sigma(M)|,$$

which is impossible. It follows that each torus T_i in the hyperbolic cusps of Ω is incompressible in M.

■

This result is similar in spirit, (although the proof is quite different), to Thurston's cusp closing theorem, c.f. [36].

Theorem 2.9 and the preceding work imply the existence of the decomposition (0.7) of M. In particular, the tori T^2_i in the hyperbolic cusps give a partial torus decomposition of M, in the sense of Jaco-Shalen-Johannson, [19], [21]. We discuss the full torus decomposition of M, corresponding to the further decomposition of G into Seifert fibered spaces S_k, in §2.3 below. The manifold $G = \cup S_k$ is called the *characteristic variety* of M, c.f. [19], [21]; Theorem 0.2 gives a new (geometric) proof of its existence for tame 3-manifolds.

Finally, we discuss the issue of uniqueness of this decomposition. Thus, given one decomposition

$$M = \Omega \cup G, \tag{2.65}$$

where the union is along incompressible tori $\{T^2_i\}$, let T' be any other incompressible torus. We claim that T' can be isotoped into G; my thanks to Yair Minsky for assistance with the argument below. First, since Ω is hyperbolic, any incompressible torus T' embedded in Ω is boundary parallel, and thus may be isotoped into G. If T' intersects Ω and G, then there is a torus $T \in \{T^2_i\}$ such that $T \cap T' \neq \emptyset$. Now $T \cap T'$ is a collection of embedded essential circles, bounding annuli $\{A_i\}$ in T. It follows that one may form a new torus $\bar{T}$ by matching the annuli $T' \cap \Omega$ with the annuli $\{A_i\}$. The torus $\bar{T}$ remains incompressible, and lies in the hyperbolic manifold Ω. Thus, it may be isotoped into G, which induces an isotopy of T' into G.

Hence, if $\{T'_j\}$ is another torus decomposition of M into Ω' and G', then $\{T'_j\} \subset G$. Further, one may assume that the collection $\{T'_j\}$ is disjoint from the collection $\{T_i\}$. Now if $\{T'_j\}$ is not isotopic to $\{T_i\}$, then it follows that one of the tori $T_k \in \{T_i\}$ is contained in one of the hyperbolic manifolds Ω'_k and is not isotopic to a boundary torus in a cusp of Ω'_k. Since T_k is incompressible, this is impossible. Thus, the collections $\{T_i\}$ and $\{T'_j\}$ of tori are isotopic.

This completes the proof of Theorem 0.2 when $\sigma(M) < 0$.

■

Proof of Theorem 0.2: $\sigma(M) = 0$.

Suppose M is a closed, oriented tame 3-manifold with

$$\sigma(M) = 0.$$

Consider a sequence $\{(\Omega_\varepsilon, g_\varepsilon)\}$ of minimizers of $I_\varepsilon{}^-, \varepsilon = \varepsilon_i$, satisfying (2.1). Thus, as discussed in §2.1, a subsequence of $\{(\Omega_\varepsilon, g_\varepsilon)\}$ either converges, collapses, or forms cusps.

If $\{g_\varepsilon\}$ converges, (modulo diffeomorphisms of M), then by Proposition 2.1, it converges everywhere on M to a smooth limit metric g_o of constant curvature, and scalar curvature $s_{g_o} = \sigma(M) = 0$. Thus, g_o is a flat metric on M, so that (M, g_o) is a flat 3-manifold, and in particular a graph manifold with infinite π_1. In this case, the metric g_o on M realizes $\sigma(M)$.

If $\{g_\varepsilon\}$ collapses, then as in Proposition 2.3, it collapses along a sequence of F-structures on M. In particular, M is a graph manifold. In this case, if M is not a flat 3-manifold, no smooth metric on M realizes $\sigma(M)$.

We claim that $|\pi_1(M)| = \infty$. For it is standard, c.f. [26, §5], that a closed oriented graph manifold M with $\pi_1(M)$ finite must be a spherical space form S^3/Γ, where Γ is a finite subgroup of $SO(4)$. Such manifolds have Yamabe metrics of positive scalar curvature, so that $\sigma(M) > 0$. Since $\sigma(M) = 0$, this gives the claim.

By Proposition 2.7, the sequence $\{g_\varepsilon\}$ cannot form cusps. This completes the proof of Theorem 0.2 in case $\sigma(M) = 0$.

∎

Remark 2.10. (i). Referring to Remarks 2.2 and 2.6, the uniqueness results above prove that if M is tame, then minimizing pairs $(\Omega_\varepsilon, g_\varepsilon)$ are unique, and independent of ε, for all $\varepsilon > 0$, (with the exception of flat manifolds). In particular, the decomposition (0.7) is the same as the decomposition given in Theorem 1.1.

Note also that the proof of Theorem 0.2 does not require that M is irreducible.

(ii). For completeness, we make a few remarks on the Sigma constant of graph manifolds. By Remark 2.4, an arbitrary closed oriented graph manifold G necessarily satisfies

$$\sigma(G) \geq 0.$$

Standard 3-manifold topology, c.f. [26, §6] or [38], implies that an irreducible graph manifold of infinite π_1 necessarily has a $\mathbb{Z} \oplus \mathbb{Z} \subset \pi_1$. From results of Schoen-Yau, [31], this implies that M has no metric of positive scalar curvature, and thus in particular, no Yamabe metric of positive scalar curvature. Thus, we have the implication

$$|\pi_1(N)| = \infty \Rightarrow \sigma(N) = 0, \tag{2.66}$$

for irreducible graph manifolds N. The opposite implication

$$\sigma(M) = 0 \Rightarrow |\pi_1(M)| = \infty,$$

for arbitrary closed graph manifolds follows from the proof of Theorem 0.2 above.

Remark 2.11. As noted in the Introduction, the geometrization of closed oriented irreducible 3-manifolds which are tame and of infinite π_1 is now a simple consequence of Theorem 0.2. Namely, as noted following (0.1), such a manifold M must be a $K(\pi, 1)$, and hence by (0.6), $\sigma(M) \leq 0$.

If in addition M is atoroidal, then by Remark 2.10, M cannot be an irreducible graph manifold, and so the $\sigma(M) = 0$ case of Theorem 0.2 implies that $\sigma(M) < 0$. Hence Theorem 0.2 implies that M is hyperbolic.

If M contains essential tori, then Theorem 0.2 implies that M is a union along essential tori of complete hyperbolic manifolds and graph manifolds. The hyperbolic manifolds are of course geometric, and graph manifolds are also geometric, since they are unions along tori of Seifert fibered spaces; the geometrization of the graph manifolds will be discussed next in §2.3.

§2.3. We complete this section with an analysis of how the torus decomposition and the geometrization of graph manifolds can be obtained from the near-limiting behavior of suitable minimizing sequences for $\mathcal{S}^2_-$ or $I_\varepsilon{}^-$. Thus we describe in particular how the other Seifert fibered geometries arise in this context. On the other hand, this section is not logically necessary for any later work.

The graph manifold structures in Theorem 0.2 in either case $\sigma(M) < 0$ or $\sigma(M) = 0$ are obtained from the near-limiting behavior of a minimizing sequence $\{g_i\}$ for $I_\varepsilon{}^-, \varepsilon > 0$; as noted in Remark 2.10, the parameter ε no longer plays any role and the conclusions of Theorem 0.2 may be obtained for any choice of $\varepsilon > 0$. The results are independent of ε.

In the situation $\sigma(M) < 0$, on the hyperbolic domain $H = \cup H_j$, the metrics g_i on M converge to the unique limit metric g_o on H. On the graph manifold domain $G = \cup G_k$, the sequence $\{g_i\}$ collapses with bounded curvature (in L^2). Of course, there are many possible choices for the minimizing sequence $\{g_i\}$, and so apriori there may be many possible ways to collapse G.

Thus, throughout §2.3, let G be an oriented graph manifold, either closed and satisfying $\sigma(G) = 0$, or with boundary consisting of a finite number of incompressible tori corresponding to the situations $\sigma(M) = 0$ or $\sigma(M) < 0$ in Conjectures II and I respectively. In the latter case, it follows from a well-known result of Gromov-Lawson [15, Thm.8.4], that G admits no complete metrics of non-negative scalar curvature which are not flat. Since such a G does admit complete metrics with $\mathcal{S}_-^2$ arbitrarily small, we will say that $\sigma(G) = 0$ in this case also. Note that we do not assume that G is irreducible.

We begin with a general discussion of the decomposition of graph manifolds obtained from the geometry of a collapsing sequence of metrics. Since G admits polarized F-structures, $\inf I_\varepsilon{}^- = \inf \mathcal{S}_-^2 = 0$, and there are minimizing sequences $\{g_i\}$ for $I_\varepsilon{}^-$ which have uniformly bounded curvature in L^∞ and which volume collapse G along a sequence $\mathcal{F}_i$ of polarized F-structures. (The uniqueness of this collapse will be discussed below). From the definition of polarized F-structure, it follows that G is partitioned into a collection of regions, possibly depending on i, on which the orbits of $\mathcal{F}_i$ are circles and tori. Let $S_j = S_j(i)$ be the components of the regions with S^1 fibers, and $L_k = L_k(i)$ the components of regions with T^2 fibers; thus $\{S_j, g_i\}$ collapses along S^1 orbits, while $\{(L_k, g_i)\}$ collapses along T^2 orbits. We assume w.l.o.g. that all domains S_j and L_k have only toral boundary components, invariant under the group actions. Assuming the collection $\{S_j\}$ is non-empty, the components L_k are considered as regions where the S_j components are glued together by toral diffeomorphisms.

Although this decomposition D of G into these types of regions, i.e. the placement of the tori separating the S and L factors, is not precisely determined by the geometry of $\{g_i\}$, it is a consequence of the fact that the F-structure is polarized that the transition between any two distinct components of this division necesssarily takes place over domains of larger and larger diameter as $i \to \infty$, c.f. [9]. Thus, as $i \to \infty$, the diameter of each component S_j or L_k is forced to go to ∞ and the metric g_i restricted to any such component becomes more and more complete.

The Seifert fibered components $S = S_j$ are S^1 fibrations over a surface $\Sigma = \Sigma_j$. Assuming that Σ is not closed, Σ is thus either an open hyperbolic surface, an annulus $S^1 \times I$, a Möbius band, or a disc D^2. The latter three cases are somewhat exceptional and we make some further remarks on them.

If Σ is an annulus, so that $S_j = T^2 \times I$, one has two possibilities. If the S^1 structure on S_j is compatible with the S^1 structure on some neighbor S_k, (near the boundary), we just extend the Seifert fibered structure on S_k to include S_j and drop S_j from the list. If the S^1 structure on S_j is not compatible with either of its neighbors in this sense, then the collapse theory [8] implies that S_j must be an L factor, i.e. S_j is collapsed along both S^1 factors. Thus in either case, there are no such S factors.

If Σ is a Möbius band, then S is the oriented S^1 bundle over Σ. Note that ∂S is connected, and is an incompressible torus in S. This manifold admits a complete flat geometry.

The situation where $\Sigma = D^2$, so that S is a solid torus $D^2 \times S^1$ is the most important, in that this is the only Seifert fibered space with compressible boundary. Although D^2 admits complete hyperbolic and complete flat metrics, it is the spherical metric on D^2 embedded as a hemisphere in S^2 which is geometrically the most natural, c.f. the proof of Theorem 2.9. Such a metric is of course incomplete.

In the next result we show that any decomposition D of G containing solid tori components in the Seifert fibered factors above may be altered to a new decomposition without such components.

Lemma 2.12. *Let G be either a closed oriented graph manifold with $\sigma(G) = 0$, or a compact oriented graph manifold with incompressible boundary, so that ∂G is a finite union of tori. Let*

$$G = G_1 \# ... \# G_k \tag{2.67}$$

be the sphere decomposition of G into prime factors.

Then each G_p is a prime graph manifold with $\sigma(G_p) \geq 0$, and $\sigma(G_j) = 0$, for some j. Further, there exists a decomposition D_p of each G_p such that no Seifert fibered component S in G_p above is a solid torus.

Proof: Each G_p is a graph manifold, either closed or with incompressible torus boundary. Hence by Remark 2.4 and the discussion above, $\sigma(G_p) \geq 0$ for each p. If $\sigma(G_p) > 0$, then G_p is either $S^2 \times S^1$ or a spherical space form S^3/Γ. If all G_p satisfy $\sigma(G_p) > 0$, then $\sigma(G) > 0$, since positive scalar curvature is preserved under connected sums, c.f. [14], [32]; thus $\sigma(G_j) = 0$, for some j.

Let D be any decomposition of G as above into S_j and L_k factors. Suppose some $S = S_j$ factor is a solid torus $D^2 \times S^1$. The factor S must then be glued, (via an L component), onto a boundary component T of a Seifert fibered space $S' = S_{j'}$. We may assume that T is incompressible in S', for otherwise G is a union of two solid tori, and hence a lens space or $S^2 \times S^1$, either of which have positive Sigma constant, contradicting $\sigma(G) = 0$.

There are exactly two possibilities for the resulting topology of $S' \cup_{T^2} (D^2 \times S^1)$, as discussed by Waldhausen, c.f. [39].

(i). First, if the fiber of S' is not glued to the meridian, i.e. to D^2, then the resulting manifold is again a Seifert fibered space $\bar{S}$, possibly with an exceptional fiber at the core of the solid torus, c.f. [39, §10]. Hence, in this case one may just alter the decomposition D by eliminating the L factor between S' and $D^2 \times S^1$ and enlarging S' to the Seifert fibered space $S' \cup_{T^2} (D^2 \times S^1)$. (Waldhausen's definition of graph manifold excludes this operation, but we do not do so here).

(ii). On the other hand, if the fiber is glued to the meridian, then the resulting manifold is reducible, c.f. [39, p.90]. This implies that G has a non-trivial sphere decomposition, $G = G^1 \# G^2$. The essential 2-sphere is formed as in the proof of Theorem 2.9. The factors G^1 and G^2 may be disconnected by glueing in a 3-ball on each side. The original decomposition D is then altered or split to give a decomposition D^i of each G^i, c.f. [39, p.90]. By the Kneser finiteness theorem [22], or by [39, p.92], there are only finitely many isotopy classes of essential 2-spheres in G, so that this process terminates after a finite number of iterations.

The result of this process is then the decomposition (2.67) with each G_p prime and a corresponding decomposition D_p of G_p without any solid torus components.

■

Note that the proof above shows that any initial graph manifold decomposition of G may be altered in a canonical way to a decomposition D satisfying the conclusions of Lemma 2.12. In particular, this process carries out the sphere decomposition of G. The union of the decompositions D_p of each prime factor G_p in (2.67) does not extend to a graph manifold decomposition of G. This corresponds to the fact that the necks $S^2 \times I$ about essential 2-spheres in G are no longer decomposed into S and L factors.

Next we discuss the uniquenesss of the decomposition of each factor G_p in (2.67).

Lemma 2.13. *Let G be a closed oriented graph manifold with $\sigma(G) = 0$, or a compact oriented graph manifold with incompressible toral boundary components, (and hence also $\sigma(G) = 0$). Suppose*

G is irreducible, and not a closed flat, Nil or Sol manifold, or the oriented S^1 bundle over a Möbius band.

Then the decomposition D of G as in Lemma 2.12, into Seifert fibered and toral factors without solid torus components, is unique up to isotopy of G. The tori in the L_k factors give the JSJ torus decomposition of G. Further, the decomposition is injective in the sense that $\pi_1(S_j)$ and $\pi_1(L_k)$ inject in $\pi_1(G)$.

Proof: The existence of the decomposition follows from Lemma 2.12 and its uniqueness up to isotopy is proved by Waldhausen [39, §8], c.f. also [26, p.132] and [28, Lemma 5.5]. The four exceptions arise from the fact that in these cases one may either choose D to be empty, (the manifolds are geometric), or may one choose D to be non-empty in inequivalent ways. The injectivity statement is proved in [28, Thm. 4.2]. The fact that the tori in the L_k factors give the JSJ decomposition is then immediate. ■

We recall that the prime factors G_p of the sphere decomposition (2.67) with $\sigma(G_p) > 0$ are either $S^2 \times S^1$ or space forms S^3/Γ.

Given these results on the structure of graph manifolds, we now return to the geometric behavior of suitable minimizing sequences. Let G be as in Lemma 2.12. We will say that a minimizing sequence $\{g_i\}$ for $\mathcal{S}^2_-$ on G *performs* the geometrization of G if the following conditions hold:

(i). Let $\{S^2_q\}$ be a fixed collection of essential 2-spheres defining the sphere decomposition (2.67) of G. The sequence $\{g_i\}$ crushes each $S^2 \in \{S^2_q\}$ to a point $x = x_q \in G$, in the sense that

$$diam_{g_i} S^2 \to 0 \ \text{ and } \ \rho_i(x) \to 0, \ \text{ as } \ i \to \infty, \tag{2.68}$$

where ρ is the L^2 curvature radius. Further, the blow-ups of the metrics $\{g_i\}$ by $\rho_i(x)^{-2}$ converge to the complete doubled Schwarzschild metric (1.23), with mass $m \sim 1$.

(ii). The metrics $\{g_i\}$ collapse each irreducible factor G_p with $\sigma(G_p) = 0$, (except possibly factors which are closed flat manifolds), along a polarized F-structure, (unique in the non-exceptional cases), determined by Lemma 2.13, on a scale however much larger than the crushing of the essential 2-spheres, i.e.

$$diam_{g_i} \mathcal{O}(y) >> diam_{g_i} S^2, \tag{2.69}$$

for any y, where $\mathcal{O}$ is the orbit of the F-structure through y.

(iii). After passing to suitable covers of each S component of G_p with $\sigma(G_p) = 0$, the metrics g_i converge to one of the six Seifert fibered geometries on S, i.e. $SL(2,\mathbb{R})$, Nil, $H^2 \times \mathbb{R}$, $\mathbb{R}^3$, $S^2 \times \mathbb{R}$, S^3, or suitable covers of G_p converge to the Sol geometry. On each factor G_p with $\sigma(G_p) > 0$, the metrics g_i converge to the constant curvature metric of curvature $+1$ on S^3/Γ or the product metric on $S^2 \times S^1$ of the form $S^2(1) \times S^1(1)$.

Note that in general, such a sequence $\{g_i\}$ is not tame, in that the L^2 norm of the curvature z_{g_i} diverges to infinity, due to the crushing of the essential 2-spheres. On the other hand, any graph manifold G as above does admit tame minimizing sequences which volume collapse G with bounded curvature, as discussed in the beginning of §2.3. In particular, on reducible 3-manifolds satisfying the conclusions of Conjectures I and II, a tame minimizing sequence does not perform the geometrization of M.

The following result is now however a simple consequence of this discussion.

Theorem 2.14. *Let G be either a closed oriented graph manifold with $\sigma(G) = 0$, or a compact oriented graph manifold with incompressible boundary, so that ∂G is a finite union of tori.*

Then there exists a minimizing sequence $\{g_i\}$ for $\mathcal{S}^2_-$ on G, i.e.

$$\mathcal{S}^2_-(g_i) \to 0,$$

which performs the geometrization of G.

If in addition G is irreducible, then there exist minimizing sequences for ${I_{\varepsilon}}^{-}$, for any $\varepsilon > 0$, which perform the geometrization of G.

Proof: We construct the minimizing sequence $\{g_i\}$ to satisfy each of the conditions (i)-(iii) in turn.

(i). Let $\{S_q^2\}$ be a collection of embedded 2-spheres in G giving the sphere decomposition of G. For each such S_q^2, associate the isometrically doubled Schwarzschild metric $g_i = g_S(m_i)$ in (1.23), with mass $m_i \to 0$, defined on $S^2 \times \mathbb{R} = \mathbb{R}^3 \# \mathbb{R}^3$. Observe that this sequence of metrics converges to the union of two copies of $\mathbb{R}^3$, glued together at one point; the 2-sphere connecting the $\mathbb{R}^3$ factors is crushed to a point as $i \to \infty$. Further note that $\rho_i(x) \sim m_i$, for any $x \in S_q^2$ and g_i is scalar-flat.

Choose a sequence $\delta_i \to 0$, with $m_i << \delta_i$, and let $g_i|_{B_q(\delta_i)}$ be the metric $g_S(m_i)$ restricted to the δ_i tubular neighborhood of the horizon $\Sigma = S^2(m_i)$. Thus, g_i in the annular region $A_q(\delta_i/2, \delta_i)$ differs from the flat metric on the order of $m_i/\delta_i << 1$. The blow-ups $g_i' = \rho_i(x)^{-2} \cdot g_i$ based at x converge to the complete isometrically doubled Schwarzschild metric g_S of mass $m \sim 1$.

(ii). Next consider minimizing sequences for $\mathcal{S}_{-}^2$ on the prime factors G_p of G first with $\sigma(G_p) = 0$. Any minimizing sequence $\{g_i\}$ for $\mathcal{S}_{-}^2$ with bounded curvature will volume collapse these factors, except possibly when G_p is flat. In accordance with case (i) of Lemma 2.12, any such sequence may be altered to a sequence, still called $\{g_i\}$ and volume collapsing with bounded curvature, which has no solid tori components in its Seifert fibered factors. Lemma 2.13 implies that such a decomposition is unique up to isotopy, and induces the JSJ torus decomposition.

For the exceptional cases where G_p is a closed flat, Nil or Sol manifold, or the S^1 bundle over the Möbius band, (so that G_p has an empty torus decomposition), choose $\{g_i\}$ to be geometric, volume collapsing in the last three cases.

Similarly on the factors with $\sigma(G_p) > 0$, choose $\{g_i\}$ to be geometric, and so $\{g_i\}$ is the constant sequence, c.f. (iii) above.

We then glue together the metrics g_i on the prime factors G_p to form a metric on G. First, consider the factors G_p with $\sigma(G_p) = 0$. On such factors, although the metric g_i is very collapsed for i large, on scales much smaller than the collapse scale the metric g_i on each G_p is close to a flat metric on a ball. Thus, choose δ_i in (i) above so that $\delta_i << diam_{g_i}\mathcal{O}(y)$, for all $y \in G_p$. The geodesic balls in (G_p, g_i) are then almost flat, and in particular are topological balls. It follows that we may remove such balls from each G_p, and metrically glue in the Schwarzschild metrics g_i from (i). This can be done so that the full curvature, and so in particular the scalar curvature, of the glueing remains bounded in the glueing region $A_q(\delta_i/2, \delta_i)$ as $\delta_i \to 0$. This construction is the same as the end of the proof of Theorem 2.9 and is carried out in detail in [4, §6.2] or [5, §6.1]; we refer there for further details.

Similarly, the prime factors with $\sigma(G_p) > 0$, i.e. S^3/Γ and $S^2 \times S^1$ with geometric metric, are glued to the other prime factors by Schwarzschild necks. Note that for g a geometric metric on S^3/Γ or $S^2 \times S^1$, $\mathcal{S}_{-}^2(g) = 0$. However, for any metric on $S^2 \times S^1$, ${I_{\varepsilon}}^{-} > 0$, for any fixed $\varepsilon > 0$; one can make ${I_{\varepsilon}}^{-} \to 0$ only by collapsing the S^1 factor with bounded curvature.

(iii). It remains to geometrize the Seifert fibered factors in each prime factor G_p. This has already been done in (ii) for the factors with $\sigma(G_p) > 0$, or for factors with flat, Nil or Sol geometries.

Thus assume $\sigma(G_p) = 0$, and so $\pi_1(G_p)$ is infinite. It follows that $\pi_1(S_j)$ is infinite, for all j. For if $\pi_1(S)$ is finite, for some $S = S_j$, then $S = S^3/\Gamma$. Hence, since S is closed, $S = G_p$, and $\sigma(G_p) > 0$, a contradiction.

By standard results, c.f. [26, Ch.5], the S^1 orbits inject in $\pi_1(S_j)$, for each j, (and thus inject in $\pi_1(G)$ by Lemma 2.13). Thus, the S^1-collapsed geometry of (S_j, g_i) may be unwrapped by passing to sufficiently large, (depending on i), finite covers $\bar{S}_j$, so that the geometry of $(\bar{S}_j, g_i)$ is bounded; thus the injectivity radius of $\{g_i\}$ at a fixed base point $x_j \in \bar{S}_j$ is bounded away from 0 and ∞. The metrics g_i then (sub)-converge to a limit metric $\bar{g}$ which is a complete S^1 invariant metric on $\bar{S}_j$.

Now of course for an arbitrary minimizing sequence $\{g_i\}$ constructed as in (ii) above, the limit metric $\bar{g}$ on $\bar{S}_j$ will not be geometric, i.e. will not be given by one of the Seifert fibered geometries. However, we simply choose $\{g_i\}$ so that its limit $\bar{g}$ on $\bar{S}_j$ is a complete Seifert fibered geometry.

For the same reasons as above, the metrics g_i on the L_k factors converge, when lifted to suitable $\mathbb{Z}\oplus\mathbb{Z}$ covers, to complete T^2 invariant metrics on $T^2\times\mathbb{R}$. Observe that the Seifert fibered geometries are rigid within their class, in that one cannot continuously deform one Seifert fibered geometry class to another class. Similarly, distinct S_j components with the same *complete* non-compact Seifert geometry cannot be glued together within any fixed geometry. Thus, the L_k factors, which serve to glue together the Seifert fibered factors S_j, cannot be made geometric.

This completes the description of the minimizing sequence $\{g_i\}$ for $\mathcal{S}^2_-$. Finally, note that if G is irreducible, so that Step (i) above is vacuous, one may choose $\{g_i\}$ to be volume collapsing with uniformly bounded curvature and hence $\{g_i\}$ is a minimizing sequence for $I_\varepsilon{}^-$, for any $\varepsilon > 0$.

■

3. Metric Surgery on Spheres in Asymptotically Flat Ends.

In this section, we prove Theorem 0.4. Thus, throughout this section, we assume that M is a closed, oriented 3-manifold, which is not tame, but which is spherically tame. Hence, as discussed in Theorem 1.3, there is a sequence $\varepsilon = \varepsilon_i \to 0$, and a sequence of minimizing pairs $(\Omega_\varepsilon, g_\varepsilon), \Omega_\varepsilon \subset\subset M$, together with a complete blow-up limit (N, g') which has an asymptotically flat end E, topologically of the form $S^2 \times \mathbb{R}^+$ outside a compact subset of E. Note that since $N \subset\subset \Omega_\varepsilon$, for ε sufficiently small, one has $N \subset\subset M$, so that in particular the 2-sphere S^2 in E embeds in M, i.e.

$$S^2 \subset M. \tag{3.1}$$

Theorem 3.1. *Suppose M is a spherically tame, but not tame, closed and oriented 3-manifold, with $\sigma(M) \leq 0$. Then the 2-sphere $S^2 \subset M$ in (3.1) is essential in M. In particular, M is reducible.*

Proof: The proof is by contradiction, so we assume that the 2-sphere S^2 in (3.1) bounds a 3-ball in M. The idea is similar to the proof of Theorem 2.9, in that we glue on a specific comparison 3-ball onto S^2, analogous to the specific metric Dehn surgery carried out in Theorem 2.9. The details of this construction are somewhat more involved however.

Since (E, g') is asymptotically flat, there is a compact set $K \subset E$ and a ball $D \subset \mathbb{R}^3$ such that $E \setminus K$ is diffeomorphic to $\mathbb{R}^3 \setminus D$, and in a suitable chart on $E \setminus K$, the metric g' has the expansion

$$g'_{ij} = (1 + \frac{2m}{r})\delta_{ij} + h, \tag{3.2}$$

where

$$|h| = O(r^{-2}), |D^k h| = O(r^{-2-k}), k = 1, 2 \tag{3.3}$$

with $m > 0$.

As observed following (1.26) or in the proof of Proposition 2.1, the estimate (1.26) implies that the metrics g'_ε from (1.19) converge in the weak $L^{3,p}$ and (strong) $C^{2,\alpha}$ topology to the limit metric g', uniformly on compact subsets of (N, g'). Thus for any given R large but fixed, the metric g'_ε from (1.19) also has the form (3.2)-(3.3) on $B_{y_\varepsilon}(2R)$, i.e. for large $r \leq 2R$,

$$(g'_\varepsilon)_{ij} = (1 + \frac{2m}{r})\delta_{ij} + h_\varepsilon, \tag{3.4}$$

where

$$|h_\varepsilon| = O(r^{-2}), |D^k h_\varepsilon| = O(r^{-2-k}), k = 1, 2, \tag{3.5}$$

on $B(2R) \setminus D \subset \mathbb{R}^3$. The expressions (3.4)-(3.5) are valid for all ε sufficiently small depending on the choice of R, and h_ε converges in the $C^{2,\alpha}$ topology to h as $\varepsilon \to 0$.

The smooth domain $B(2R)\setminus D$ embeds smoothly in Ω_ε for ε sufficiently small, and hence also in M. In particular, for all R large but fixed, the canonical round 2-spheres $S^2(R)\subset\mathbb{R}^3\setminus D\cong E\setminus K$, centered at $0\in\mathbb{R}^3$, embed in M.

By the assumption above, $S^2(R)$ bounds a 3-ball B^3 in M, so that in particular $S^2(R)$ disconnects M into two components, consisting of the *inside* component containing the base point y of N, and the *outside* component containing the unbounded part of the end $E\subset N$. We will say that $S^2(R)$ bounds a 3-ball B^3 on the inside, (or outside), if the inside (or outside) component is a 3-ball. Obviously this notion is independent of R, for all R large.

Case I. Suppose that $S^2(R)$ bounds B^3 on the inside.
First, for clarity, we glue in a comparison 3-ball $\bar{B}$ onto the spherically symmetric metric

$$g_S=(1+\frac{2m}{r})\delta_{ij}, \tag{3.6}$$

at $S^2(R)$. Following this, it is shown that essentially the same construction and estimates are valid when glueing onto the limit metric g' in (3.2) or the approximations g'_ε in (3.4). A simple computation shows that g_S has positive scalar curvature, $s_{g_S}>0$, so that $(s_{g_S})^-\equiv 0$.

In the metric g_S, the sphere $S^2(R)$ is isometric to a constant curvature sphere of Euclidean radius $R(1+\frac{2m}{R})^{1/2}$. Let (B_o,g_o) be the flat Euclidean ball of the same radius, so that $S^2(R)$ is filled in on the "inside" with a flat Euclidean metric g_o. The metrics g_S and g_o agree C^o at the boundary, but not C^1. Let $\bar{g}$ denote the union of these two metrics, as a metric on $\mathbb{R}^3$.

The main point now is that $S^2(R)$ is more convex in the Euclidean metric than in the g_S metric. Thus, let A denote the 2nd fundamental form of $S^2(R)$, w.r.t. the outward normal, with A_{g_o} denoting the 2nd fundamental form w.r.t. the Euclidean metric, and A_{g_S} the form w.r.t. the g_S metric; the sign conventions are such that $A>0$ for spheres in Euclidean space. Then a short computation using (3.6) gives

$$A_{g_o}=\frac{1}{R(1+\frac{2m}{R})^{1/2}}I\sim\big(\frac{1}{R}-\frac{m}{R^2}+O(R^{-3})\big)I,\ A_{g_S}=\frac{1-\frac{m}{R}(1+\frac{2m}{R})^{-1}}{R(1+\frac{2m}{R})}I\sim\big(\frac{1}{R}-\frac{2m}{R^2}+O(R^{-3})\big)I, \tag{3.7}$$

where I denotes the identity matrix. Thus

$$A_{g_o}-A_{g_S}=\frac{m}{R^2}I+O(R^{-3})>0, \tag{3.8}$$

since $m>0$ and R is assumed sufficiently small.

This means that there is a concentration of positive curvature for $\bar{g}$ at the seam $S^2(R)$. It follows that the metric $\bar{g}$ may be smoothed to a metric $\widetilde{g}$ in the annulus $A(R-1,R+1)$ so that it agrees with the flat metric at $S^2(R-1)$ and with g_S at $S^2(R+1)$, and satisfies

$$(\widetilde{s})^-=(s_{g_S})^-=0, \tag{3.9}$$

in $A(R-1,R+1)$. Of course $\widetilde{g}$ agrees with g_S outside the ball $B(R+1)$.

To verify (3.9), one may write $\bar{g}$ explicitly as a conformally flat warped product

$$\bar{g}=dt^2+f^2(t)ds^2_{S^2}; \tag{3.10}$$

here $f(t)=t$, for $t\leq\bar{R}\equiv R(1+\frac{2m}{R})^{1/2}$, while $f(t)=r(1+\frac{2m}{r})^{1/2}$, for $r=r(t)\geq\bar{R}$. The scalar curvature of a metric of the form (3.10) is given by

$$\frac{1}{2}s=-2\frac{f''}{f}+\frac{1-(f')^2}{f^2},$$

while $A=(f'/f)\cdot I$. The estimate (3.8) means that when smoothing the Lipschitz function f at the seam $\bar{R}$ to obtain a smooth function $\widetilde{f}$, one has $\widetilde{f}''\leq 0$ near the seam. Hence the smoothing $\widetilde{g}$ increases the scalar curvature, which gives (3.9).

Next, we make a similar estimate for the change in volume, as above working with g_S. Clearly $vol_{g_o}B(\bar{R}) = \omega_3\bar{R}^3 = \nu_2\bar{R}^3/3$, where ω_3 (ν_2) is the volume of the Euclidean unit 3-ball (2-sphere). Another short computation using (3.6) shows that the volume of the region $B(R)$ in the end (E, g_S) satisfies

$$vol_{g_S}B(R) \geq (\frac{1}{3}R^3 + \frac{3}{2}mR^2 - O(R))\nu_2 = (R^3 + \frac{9}{2}mR^2 - O(R))\omega_3.$$

From the definition of $\bar{R}$, it follows that

$$vol_{g_o}B_o - vol_{g_S}B(R) \leq -\frac{3}{2}mR^2\omega_3 + O(R),$$

so that, for R sufficiently large,

$$vol_{g_o}B_o < vol_{g_S}B(R). \tag{3.11}$$

Hence, one sees that $\widetilde{g}$ has non-negative scalar curvature and less volume than g_S, and so is an "admissible" comparison metric to g_S, in the sense of the minimizing properties stated in Theorem 1.3. We now claim that essentially the same construction and estimates may be carried out w.r.t. the metric g' in (3.2) in place of the metric g_S.

Thus, consider the sphere $S^2(R)$ w.r.t. the metric g', for R large. Since h in (3.2) is sufficiently small in the C^2 topology, $(S^2(R), g')$ has positive Gauss curvature. Hence, by the Weyl-Alexandrov embedding theorem, c.f. [35], $(S^2(R), g')$ may be isometrically embedded in $\mathbb{R}^3$ as a convex surface. In particular, $S^2(R)$ bounds a convex domain $\bar{B} \subset \mathbb{R}^3$, where $\mathbb{R}^3$ is given the flat metric g_o as before. One may then glue in $\bar{B}$ to the boundary $S^2(R)$, to obtain a new complete metric $\bar{g}'$ on $\bar{N}$, diffeomorphic to $\mathbb{R}^3$. This metric is C^o, and piecewise C^∞, but is not C^1 at the seam $S^2(R)$.

Using the decay estimates on h, it follows as in (3.7)-(3.8) that

$$A_{g_o} - A_{g'} = \frac{m}{R^2}I + o(R^{-2}) > 0, \tag{3.12}$$

for R sufficiently large. As before, the metric $\bar{g}'$ may be smoothed to a metric $\widetilde{g}$ near the seam $S^2(R)$, in a manner similar to the smoothing following (3.10). More precisely, at the seam $(S^2(R), g')$ consider exponential normal coordinates w.r.t. g' on the outside and w.r.t. g_o on the inside. The metric $\bar{g}$ then has the form $\bar{g} = dt^2 + g_t$, where t is the signed distance to $S^2(R)$ and g_t is a curve of metrics on S^2. The family g_t is piecewise C^∞ and Lipschitz through the seam $S^2(R)$. The second fundamental form A in these coordinates is given by $A = dg/dt$. As above with (3.10), the estimate (3.12) implies there is a smoothing $\widetilde{g}$ near $S^2(R)$ so that $\widetilde{s} \geq 0$ near $S^2(R)$. Hence

$$(\widetilde{s})^- \geq (s_{g'})^-, \tag{3.13}$$

as in (3.9). For the same reasons, the volume estimate (3.11) holds w.r.t. $\widetilde{g}$ and g'.

Finally, we claim that these same estimates hold on the approximations g'_ε in place of the limit g'. This holds since by (3.4)-(3.5), g'_ε has the same form as g', i.e. the lower order terms h_ε obey the same estimates as the lower order term h for g'.

It follows that there are metrics $\widetilde{g}'_\varepsilon$, (close to $\widetilde{g}$), such that for ε sufficiently small,

$$(s_{\widetilde{g}'_\varepsilon})^- \geq (s_{g'_\varepsilon})^-, \tag{3.14}$$

in $A(R-1, R+1)$, while $s_{\widetilde{g}'_\varepsilon} \equiv 0$ in $B(R-1)$ and $\widetilde{g}'_\varepsilon = g'_\varepsilon$ outside $B(R+1)$. In addition, one has

$$vol_{\widetilde{g}'_\varepsilon}B(R+1) < vol_{g'_\varepsilon}B(R+1),$$

and hence

$$vol_{\widetilde{g}'_\varepsilon}\widetilde{\Omega}_\varepsilon < vol_{g'_\varepsilon}\Omega_\varepsilon, \tag{3.15}$$

where $\widetilde{\Omega}_\varepsilon$ is Ω_ε with the component bounding S^2 on the inside replaced by a 3-ball.

We now compare the values of the functional $I_\varepsilon{}^-$ on $\widetilde{g}'_\varepsilon$ and g'_ε. From (3.14), and the fact $\widetilde{g}'_\varepsilon$ is scalar-flat in $B(R-1)$,

$$\int_{\widetilde{\Omega}_\varepsilon}(s^-_{\widetilde{g}'_\varepsilon})^2 dV_{\widetilde{g}'_\varepsilon} \leq \int_{\Omega_\varepsilon}(s^-_{g'_\varepsilon})^2 dV_{g'_\varepsilon}. \tag{3.16}$$

Thus, the estimates (3.15) and (3.16) imply that $\mathcal{S}^2_-(\widetilde{g}'_\varepsilon) < \mathcal{S}^2_-(g'_\varepsilon)$.

Next, to compare the L^2 norm of z on both metrics, the flat metric and the metric g'_ε differ on the order of R^{-1} near $S^2(R)$, while their curvatures differ on the order of R^{-3}. It follows that the smoothing $\widetilde{g}'_\varepsilon$ in $A(R-1,R+1)$ may be done so that the curvatures of $\widetilde{g}'_\varepsilon$ near the seam $S^2(R)$ are on the order of at most R^{-2}. In terms of the discussion above on (3.10), this arises from the fact that since $f = O(R)$ and the jump in f' at the seam is on the order of $O(R^{-1})$, one may choose the smoothing $\widetilde{f}$ with $|\widetilde{f}''| = O(R^{-1})$, so that $|\widetilde{f}''/f| = O(R^{-2})$. Thus,

$$\int_{A(R-1,R+1)} |z_{\widetilde{g}'_\varepsilon}|^2 dV_{\widetilde{g}'_\varepsilon} \leq cR^{-2}, \tag{3.17}$$

since $vol A(R-1,R+1) = O(R^2)$; here c is a constant independent of R.

In $B(R-1), \widetilde{g}'_\varepsilon$ is flat, so that of course $z = 0$ in this region. On the other hand, g'_ε has a definite amount of curvature in this region, (inside $S^2(R-1)$). For instance, since at the base point $y_\varepsilon, \rho'(y_\varepsilon) = 1$, one has

$$\int_{B_{y_\varepsilon}(1)} |r_{g'_\varepsilon}|^2 dV_{g'_\varepsilon} \geq c_o vol B_{y_\varepsilon}(1),$$

and the same for limit metric g' based at y. Since the limit (N, g') is not of constant curvature on any open set, (by analyticity), we then also have

$$\int_{B_{y_\varepsilon}(1)} |z_{g'_\varepsilon}|^2 dV_{g'_\varepsilon} \geq d_o vol B_{y_\varepsilon}(1), \tag{3.18}$$

for ε sufficiently small, where $d_o = d_o(c_o)$. Comparing (3.17) and (3.18), it follows that by choosing R sufficiently large, and ε sufficiently small, one obtains

$$\int_{\widetilde{\Omega}_\varepsilon} |z_{\widetilde{g}'_\varepsilon}|^2 dV_{\widetilde{g}'_\varepsilon} < \int_{\Omega_\varepsilon} |z_{g'_\varepsilon}|^2 dV_{g'_\varepsilon}. \tag{3.19}$$

Combining the estimates (3.15), (3.16), (3.19), giving

$$I_\varepsilon{}^-(\widetilde{g}'_\varepsilon) < I_\varepsilon{}^-(g'_\varepsilon). \tag{3.20}$$

For $\varepsilon > 0$ fixed, the metrics $\widetilde{g}'_\varepsilon$ and g'_ε are not (necessarily) defined on M. However, both are limits of sequences $\{\widetilde{g}_i\}$, $\{g'_i\}$ of metrics on M for which $I_\varepsilon{}^-$ converges to its value on $\widetilde{g}'_\varepsilon$ and g'_ε. Hence, (3.20) contradicts the minimizing property of g'_ε or g_ε in (1.8); we recall that $I_\varepsilon{}^-$ is scale invariant. This proves Theorem 3.1 in case $S^2(R)$ bounds a 3-ball on the inside.

Case II. Suppose $S^2(R)$ bounds B^3 on the outside.
We proceed as above to construct a suitable comparison metric, although this case requires a little more delicate consideration. In particular we will need to make stronger use of the decay estimate (3.3) on h; this estimate plays only a minor role in Case I, (c.f. also Theorem 4.2 below).

Observe that if $S^2(R)$ bounds on the inside, in place of glueing in a flat 3-ball $\subset \mathbb{R}^3$ as in Case I above, one may instead glue in the 3-ball contained in a very large 3-sphere $S^3(\delta^{-1})$ of radius δ^{-1}; for δ sufficiently small, all the estimates above remain valid. The idea now is that if instead $S^2(R)$ bounds on the outside, we glue in the complementary (very) large 3-ball in $S^3(\delta^{-1})$.

Note that in this case, since the end E has infinite volume,

$$vol_{g'_\varepsilon} B^3 \to \infty, \quad \text{as} \quad \varepsilon \to 0, \tag{3.21}$$

where B^3 is the 3-ball bounding $S^2(R)$ on the outside.

We first construct the comparison metric $\widetilde{g}$ on the limit, and then approximate it to obtain a comparison metric $\widetilde{g}'_\varepsilon$ for g'_ε, as in Case I. The end E of (N, g') is asymptotically flat, so satisfies (3.2). It is convenient, although not necessary, to rescale the metric g' so that the mass is normalized to $m = \frac{1}{2}$; observe that the mass scales as distance. This normalization eliminates the dependence of the estimates to follow on the mass m. In particular, assuming $m = \frac{1}{2}$ from now on, one has

$$\int_{A(R,\infty)} |z|^2 dV_{g'} \sim R^{-3}, \tag{3.22}$$

as $R \to \infty$. Now as before, write the metric g' in (3.2) as

$$g' = g_S + h, \tag{3.23}$$

where

$$g_S = (1 + \frac{1}{r})g_{Eucl} \ \text{ and } \ h = O(r^{-2}). \tag{3.24}$$

Given R large, let δ and D be the solutions to the equations

$$\sin \delta D = \delta R(1 + \frac{1}{R})^{1/2}, \tag{3.25}$$

$$\cos \delta D = 1 - \frac{1}{2R}(1 - \frac{1}{R})^{-1}.$$

These equations mean that the geodesic sphere $S^2(D)$ of radius D in $S^3(\delta^{-1})$, about some base point x_o, is isometric to the sphere $S^2(R) \subset (E, g_S)$, and the 2^{nd} fundamental form of $S^2(D) \subset S^3(\delta^{-1})$ satisfies

$$A = \frac{\delta \cos \delta D}{\sin \delta D} I = R^{-1}(1 + \frac{1}{R})^{-1/2}(1 - \frac{1}{2R}(1 + \frac{1}{R})^{-1})I, \tag{3.26}$$

agreeing with the 2^{nd} fundamental form of $S^2(R) \subset (E, g_S)$. Thus in the metric g_S, the boundary $S^2(R)$ is isometric, and has identical 2^{nd} fundamental form, to $S^2(D) \subset S^3(\delta^{-1})$. Let g_δ denote the (round) metric on $S^3(\delta^{-1})$. Note that $\delta^2 = O(R^{-3})$, so that the curvatures of g_S and (S^3, g_δ) are on the order of R^{-3}.

As in Case I, for the moment, we work with the metric g_S. It follows that if one attaches the complementary geodesic ball $B = B_{y_o}(2\pi\delta^{-1} - D) \subset S^3(\delta^{-1})$, to $S^2(D)$, where y_o is the antipodal point to x_o, then the resulting metric $\bar{g}$ consisting of g_S on the inside and (B, g_δ) on the outside is piecewise C^∞, and is C^1 smooth at the seam $S^2(R)$. We then smooth the seam $S^2(R)$, in a conformally flat way as in (3.10), in a band $A = A(R-1, R+1)$ about $S^2(R)$. Because the curvatures of g_S and g_δ are on the order of $O(R^{-3})$ and the metrics agree C^1 at the seam, one obtains in this way a smooth metric $\widetilde{g}$ satisfying

$$\int_A |z|^2 dV_{\widetilde{g}} \leq c \cdot R^{-4}. \tag{3.27}$$

Note that $z = 0$ past $S^2(R+1)$, i.e. outside $S^2(R+1)$. Further, since the scalar curvature of $S^3(\delta^{-1})$ is positive and g_S is scalar-flat, the smoothing may be done so that the metric $\widetilde{g}$ has $\widetilde{s} \geq 0$ pointwise. It follows that $\widetilde{g}$ is an admissible comparison metric to g_S, in that it has less volume than g_S and has non-negative scalar curvature. By comparing (3.22) and (3.27), one sees that

$$\int |z_{\widetilde{g}}|^2 < \int |z_{g_S}|^2. \tag{3.28}$$

Next, we deal with the lower order term h in (3.2). First, the curvatures of g' and g_S in the glueing region above differ on the order of at most $O(R^{-4})$, since the metrics differ by $O(R^{-2})$ and

$|D^k h| = O(R^{-2-k}), k = 1,2$. Now redefine δ and D so as to solve the system

$$\sin \delta D = \delta R(1 + \frac{1}{R})^{1/2}, \tag{3.29}$$

$$\cos \delta D = 1 - \frac{1}{2R}(1 - \frac{1}{R})^{-1} - \frac{1}{R^{\lambda}},$$

where λ is a fixed number in $(\frac{3}{2}, 2)$. For this choice of δ and D, the geodesic sphere $S^2(D) \subset S^3(\delta^{-1})$ is isometric to $S^2(R)$ in g_S, while its 2^{nd} fundamental form A_{g_δ} is smaller than the 2^{nd} fundamental form A_{g_S}; in fact,

$$A_{g_S} - A_{g_\delta} \sim \frac{1}{R^{1+\lambda}} I > 0. \tag{3.30}$$

Thus, as in (3.8), there is a concentration of positive curvature at the seam.

As above, use the Weyl embedding theorem to isometrically embed $(S^2(R), g')$ in $S^3(\delta^{-1})$. Since h satisfies the decay estimates (3.3), $(S^2(R), g')$ is close to $(S^2(D), g_\delta)$, and has 2^{nd} fundamental form A_{g_δ} still satisfying (3.30), i.e.

$$A_{g'} - A_{g_\delta} \sim \frac{1}{R^{1+\lambda}} I. \tag{3.31}$$

Thus, to the metric g' at the boundary $S^2(R)$, attach on the outside the large domain $\bar{B}$ in $S^3(\delta^{-1})$ with $\partial \bar{B} = S^2(R)$, so that $\bar{B}$ is a small perturbation of $B_{y_o}(2\pi\delta^{-1} - D)$.

The resulting metric $\bar{g}'$ is C^o, and satisfies the estimate (3.31) at the seam $S^2(R)$. Exactly as discussed earlier following (3.12), this metric may be smoothed within the annulus $A(R-1, R+1)$ to a metric $\widetilde{g}'$ so that

$$(\widetilde{s}')^- \geq (s')^- = 0, \tag{3.32}$$

everywhere. From the estimate (3.21), it is obvious that $\widetilde{g}'$ has less volume than g'.

Further, using (3.31), the curvature of $\widetilde{g}'$ in $A(R-1, R+1)$ is on the order of

$$|\widetilde{z}'| = O(R^{-1-\lambda}). \tag{3.33}$$

This follows from the fact that the curvatures of g' and g_δ are on the order of $O(R^{-3})$ together with (3.31) and the Gauss-Codazzi equations. Thus,

$$\int_{A(R-1,R+1)} |\widetilde{z}'|^2 dV_{\widetilde{g}'} = O(R^{-2\lambda}) = o(R^{-3}). \tag{3.34}$$

Comparing (3.34) with (3.22), $\widetilde{g}'$ hence has non-negative scalar curvature, less volume and less L^2 norm of curvature than g'.

We are now in position to construct a comparison metric $\widetilde{g}'_\varepsilon$ to g'_ε and compare the values $I_\varepsilon{}^-(\widetilde{g}'_\varepsilon)$ and $I_\varepsilon{}^-(g'_\varepsilon)$. As $\varepsilon = \varepsilon_i \to 0$, the metrics g'_ε converge smoothly to the limit g', uniformly on compact sets. One may then fix a choice of R sufficiently large, and then choose ε sufficiently small so that in the region $A(R-10, R+10)$, the metric g'_ε is of the form (3.2). It follows that the estimates and constructions above on g' are equally valid for g'_ε. Now from (3.21), for ε sufficiently small,

$$vol_{\widetilde{g}'_\varepsilon}(\widetilde{\Omega}_\varepsilon) < vol_{g'_\varepsilon}(\Omega_\varepsilon),$$

where as in Case I, $\widetilde{\Omega}_\varepsilon$ is Ω_ε with the component bounding S^2 on the outside replaced by a 3-ball. Further, from the construction above, one has $(\widetilde{s}'_\varepsilon)^- \geq (s'_\varepsilon)^-$. It follows that

$$\int_{\widetilde{\Omega}_\varepsilon} (s^-_{\widetilde{g}'_\varepsilon})^2 dV_{\widetilde{g}'_\varepsilon} < \int_{\Omega_\varepsilon} (s^-_{g'_\varepsilon})^2 dV_{g'_\varepsilon}.$$

and similarly,

$$\int_{\widetilde{\Omega}_\varepsilon} |z_{\widetilde{g}'_\varepsilon}|^2 dV_{\widetilde{g}'_\varepsilon} < \int_{\Omega_\varepsilon} |z_{g'_\varepsilon}|^2 dV_{g'_\varepsilon}.$$

Thus, again one has

$$I_{\varepsilon}{}^{-}(\widetilde{g}'_{\varepsilon}) < I_{\varepsilon}{}^{-}(g'_{\varepsilon}), \tag{3.35}$$

contradicting, as in (3.20), the minimizing property of g'_{ε}.

It follows that $S^2(R)$ cannot bound a 3-ball either on the inside or the outside, which completes the proof.

■

Remark 3.2. We point out that the comparison argument above strongly makes use of the flexibility in the functional $\mathcal{S}^2_-$ or $I_{\varepsilon}{}^{-}$, in that, because of the cutoff $s^- = \min(s, 0)$, one may ignore regions of the manifold where the scalar curvature of the metric is positive. A similar comparison argument for the more rigid functional $\mathcal{S}^2$ in (0.2), or the associated I_{ε} as in (0.10), (with $\mathcal{S}^2$ in place of $\mathcal{S}^2_-$), would be much more difficult to carry out. In this situation, the blow-up limit (N, g') is necessarily scalar-flat, and the allowable comparison metrics on the limit must also be scalar-flat. However, it is not clear that g' admits any compact scalar-flat perturbations.

Similarly, the comparison argument in Case II does not hold if one uses $|r|^2$ in place of $|z|^2$ in the definition (0.10) of $I_{\varepsilon}{}^{-}$.

4. Asymptotically Flat Ends and Annuli.

Theorems 0.2 and 0.4 are the main results of this paper, and as discussed in §0, reduce Conjectures I and II to the Sphere conjecture. This final section of the paper presents some remarks and results related to the Sphere conjecture, and so serves as a bridge to the sequel paper.

Regarding the Sphere conjecture, Theorem 1.4 of course gives a natural condition implying that all ends of complete $\mathcal{Z}^2_c$ solutions (N, g') are asymptotically flat. Proposition 4.1 below gives a relatively simple characterization of such limits which admit at least one asymptotically flat end, thus generalizing in a sense Theorem 1.4. (The proof however is just a minor variation of that of Theorem 1.4). The main result of this section, Theorem 4.2, shows that one may carry out metric sphere surgeries, as in the proof of Theorem 3.1, under much weaker conditions than the assumption of an asymptotically flat end, at least from the inside. Hence, this result extends the domain of validity of Theorem 0.4. We include Theorem 4.2 in this paper since the main ideas of the proof are similar to those of Theorem 3.1, although the technical details are somewhat different. These more technical issues in fact serve as an introduction to methods used more extensively in the sequel.

We consider complete, non-flat $\mathcal{Z}^2_c$ solutions (N, g', ω), i.e. complete metrics satisfying the conclusions of Theorem 1.3, with $\omega = \tau + \alpha s/12$. In particular, the metric satisfies the equations (1.16)-(1.17). For convenience, set

$$u = -\omega, \tag{4.1}$$

so that $u > 0$ in the interior of the region where $s = 0$. We begin with the following:

Proposition 4.1. *Let (N, g, u) be a complete, non-flat $\mathcal{Z}^2_c$ solution.*

(i). *Suppose $u_o = sup_N u < \infty$. Then there is a constant $\delta_o > 0$, depending only on α/u_o, such that if a compact subset C_o of the level set $L_o = \{u = u_o(1 - \delta_o)\}$ bounds a component of the superlevel set $U^o = \{u \geq u_o(1 - \delta_o)\}$ in N, then (N, g) has an asymptotically flat end.*

(ii). *There exists $K < \infty$, depending only on α, such that if a compact subset C_1 of the level set $L_1 = \{u = K\}$, (assumed non-empty), bounds a component of $U^1 = \{u \geq K\}$ in N, then (N, g) has an asymptotically flat end on which u is bounded.*

Proof: (i). Let D^o be a component of U^o such that $C_o = \partial D^o \subset L_o$ is compact. By the maximum principle applied to the trace equation (1.17), D^o must be non-compact, and hence defines an end E of N.

We claim that E, or possibly a sub-end of E, is asymptotically flat. This follows by assembling results from [5,§7]. Namely one has

$$u \geq u_o(1-\delta_o) > 0, \tag{4.2}$$

everywhere on E. Further, the oscillation of u on E satisfies $osc_E u \leq \delta_o \cdot u_o$. By [5, Lemma 7.2], this implies that if δ_o is sufficiently small, then the curvature of (E, g') is everywhere small, in the sense that

$$|r|(x) \leq \delta_1; \tag{4.3}$$

the constant δ_1 depends only on δ_o.

It follows from the estimates (4.2) and (4.3), together with [5, Prop. 7.17] that

$$\limsup_{t\to\infty} |r| \to 0 \tag{4.4}$$

in E, where $t(x) = \mathrm{dist}(x, \partial E)$. Now as noted at the end of [5,Prop.7.17], the estimates (4.4) and $u \geq u_1$, for some $u_1 > 0$, are all that are required to carry out the proof of Theorem C in [5], i.e. Theorem 1.4 here, the point being that the proof takes place on each end individually. Hence a subend of E is asymptotically flat.

(ii). As above, the component D^1 of $U^1 = \{u \geq K\}$ bounding C_1 defines an end E of N. Now divide the Euler-Lagrange equations (1.16)-(1.17) by K, to obtain the equations

$$\frac{\alpha}{K}\nabla\mathcal{Z}^2 - L^*\bar{u} = 0,$$

$$\Delta\bar{u} = \frac{\alpha}{4K}|z|^2,$$

where $\bar{u} = \frac{u}{K}$. If, for a given α, K is sufficiently large, these equations are close to the static vacuum equations. Since there are no complete non-flat static vacuum solutions by [4, Thm.3.2], it follows that the curvature is small sufficiently far out in D^1, i.e.

$$|r|(x) \leq \delta_1,$$

for all $x \in D^1$ of distance at least T to C_1, where δ_1 depends only on a sufficiently large choice of K and T; c.f. the proof of [5, Prop.7.17]. It then follows as before that (4.4) holds, and the remainder of the proof follows as in (i) above from [5, Thm.C].

■

For the work in the sequel paper, we will need a generalization of the metric sphere surgery given in Theorem 3.1. It is clear that one does not need a "complete" asymptotically flat end to carry out the proof of this result. It suffices to have a suitable spherical annulus $A = S^2 \times I$ where the glueing takes place on which the metric g' is sufficiently close to the flat metric, i.e. has the form (3.2)-(3.3).

While the glueing on the outside, i.e. Case II of the proof of Theorem 3.1, requires a strong estimate on the deviation of g' from the flat metric, i.e. that g' have the form (3.2) with $h = O(r^{-2})$, a much weaker estimate suffices for glueings of 3-balls on the inside; this is often the more important case anyway, c.f. Remark 4.3.

To describe this quantitatively, let (N, g', y) be a complete non-flat $\mathcal{Z}_c^2$ solution, arising as a blow-up limit of $(\Omega_\varepsilon, g_\varepsilon, y_\varepsilon), \varepsilon = \varepsilon_i \to 0$, as in Theorem 1.3. Let v_o, κ be (arbitrary) small positive constants less than $\frac{1}{2}$, let u_o, d be any positive constants, and let D, R be large positive constants, with R sufficiently large, (depending only on $\frac{\alpha}{u_o}$). We suppose there exists a component A of the geodesic annulus $A_y((1-d)R, (1+d)R)$ about $y \in N$, which is topologically of the form $S^2 \times I$, and which satisfies the global size bounds

$$volA \geq v_o \cdot R^3, \quad diamA \leq D \cdot R. \tag{4.5}$$

Further, suppose the potential function u satisfies the oscillation bounds

$$osc_A u = \delta_o \cdot u_o, \quad \sup_A u \equiv u_o > 0. \tag{4.6}$$

and that there is some level set $L_o = \{u = u_o(1-\delta')\}$ of u in A such that

$$L_o \subset A_{\kappa d} = \{x \in A : dist(x, \partial A) = \kappa d R\}, \tag{4.7}$$

for some $\kappa > 0$, and that the sub-level set $U_o = \{u \leq u_o(1-\delta')\}$ contains the inner boundary $\bar{A} \cap S_y((1-d)R)$ of A.

Any asymptotically flat end on which u is bounded away from 0 at infinity satisfies these conditions; in fact these conditions are much weaker than such an assumption. We then have:

Theorem 4.2. *Let (N, g', y) be a complete non-flat $\mathcal{Z}_c^2$ solution, as above, arising as a blow-up limit of $(\Omega_\varepsilon, g_\varepsilon, y_\varepsilon), \varepsilon = \varepsilon_i \to 0$, and satisfying the size and potential hypotheses (4.5)-(4.7).*

Then if δ_o is sufficiently small, depending only on v_o, d, κ and D, the essential 2-sphere $S^2 \subset A \subset N$ cannot bound a 3-ball in M on the inside of S^2. In particular, N itself is a reducible 3-manifold, while M is reducible on the inside of S^2.

Proof: For convenience, throughout the proof, we work in the scale where $R = 1$, i.e. rescale g' by R^{-2}. For simplicity, we do not change the notation for g', and often ignore the prime in the equations to follow. Thus, in this scale, (4.5) becomes

$$volA \geq v_o, \quad diamA \leq D. \tag{4.8}$$

As in the proof of Proposition 4.1, there exists $\delta_1 = \delta_1(\delta_o, \kappa)$, which may be made arbitrarily small by requiring that δ_o is sufficiently small, such that

$$|r|(x) \leq \delta_1, \tag{4.9}$$

for all $x \in A_{\kappa d}$. We assume that δ_o is sufficiently small so that (4.9) implies that $A_{\kappa d}$ is diffeomorphic to a flat manifold. Hence $A_{\kappa d}$ carries a flat metric g_o, which is $\delta_2 = \delta_2(\delta_1)$ close to g' in the $C^{1,\alpha}$ topology.

The distance function t from the base point y in N, or equivalently (by subtracting a constant) from the inner boundary $\partial_i A = S_y(1-d)$ of A, is close to a distance function t_o w.r.t. the flat metric g_o on $A_{\kappa d}$. Since A is simply connected, the global size bounds (4.8) then imply that $(A_{\kappa d}, g_o)$ may be isometrically immersed into a bounded domain C in $\mathbb{R}^3$; the volume and diameter of C are bounded, independent of δ. This immersion is a covering map when restricted to $S^2 \subset A$ and hence $A_{\kappa d}$ embeds isometrically onto the domain $C \subset \mathbb{R}^3$. One has $dist_{\mathbb{R}^3}(\partial_i C, \partial_o C)$ approximately equal to $2d$; here ∂_i and ∂_o correspond to the inner and outer boundaries of A, where $\partial_o A = S_y(1+d)$. In particular, there is a 2-sphere S^2 embedded in $C \subset \mathbb{R}^3$, and essential in C in the sense that S^2 separates $\partial_i C$ from $\partial_o C$.

We now analyse in detail the structure of the metric g' in $A_{\kappa d}$. To do this, we essentially linearize the $\mathcal{Z}_c^2$ equations at the "limit" metric g_o and "limit" potential $u \equiv u_o$, i.e. we analyse the first order deviation of (g, u) from the flat pair (g_o, u_o) in $A_{\kappa d}$. For simplicity, assume that

$$u_1 \equiv \sup_{A_{\kappa d}} u = 1; \tag{4.10}$$

this may be achieved, w.l.o.g, by renormalizing the defining equations (1.16)-(1.17), i.e. changing α to α/u_1. Of course $u_1 \sim u_o$ in (4.6).

Consider the $\mathcal{Z}_c^2$ equations (1.16)-(1.17) in this scale, i.e. (using (4.1)),

$$\alpha \nabla \mathcal{Z}^2 = L^* u, \tag{4.11}$$

$$\Delta u = \frac{\alpha}{4}|z|^2. \tag{4.12}$$

The coefficient α scales as the square of the distance, (so that the equations (4.11)-(4.12) are scale-invariant, c.f. [5, §4.1]). Thus, in the scale above, $\alpha \sim R^{-2} << 1$ while $|\nabla \mathcal{Z}^2| < O(\delta_1^2) << 1$. This latter estimate follows from uniform elliptic regularity estimates for the $\mathcal{Z}_c^2$ equations in regions where u satisfies (4.6), c.f. [5, Thm. 4.2] together with (4.9). Let

$$\mu = \Big(\int_{A_{d/2}} |r|^2 dV\Big)^{1/2}, \tag{4.13}$$

so that μ is small, depending on δ_1. We formally linearize the equations (4.11)-(4.12) at the flat metric g_o, by dividing by μ, to obtain

$$\frac{\alpha}{\mu}\nabla\mathcal{Z}^2 = L^*(\frac{u}{\mu}),$$

$$\Delta(\frac{1-u}{\mu}) = \frac{\alpha}{4\mu}|z|^2.$$

By the remarks on regularity above, $\delta_1 \leq c\cdot\mu$, for a constant c indpendent of δ_o. Hence $|\frac{\alpha}{\mu}\nabla\mathcal{Z}^2| <<$ 1 and $\frac{\alpha}{\mu}|z|^2 << 1$ for δ_o sufficiently small. This gives

$$|L^*(\frac{u}{\mu})| << 1 \quad \text{and} \quad |\Delta(\frac{1-u}{\mu})| << 1, \tag{4.14}$$

away from ∂A. Thus, the function

$$\psi \equiv \frac{1-u}{\mu} \tag{4.15}$$

is almost harmonic w.r.t. the g' metric, and hence almost harmonic w.r.t. the g_o metric. Further, expanding L^* in (4.14) gives

$$|D^2\psi - u(\frac{r}{\mu})| << 1,$$

and since u is close to 1,

$$\frac{r}{\mu} = D^2\psi + o(1), \tag{4.16}$$

away from ∂A, where $o(1)$ is (arbitrarily) small if δ_o is sufficiently small. Note that by construction, i.e. by (4.13) and the elliptic regularity, $\frac{r}{\mu}$ is uniformly bounded, independent of the smallness of δ_o, and hence so is $D^2\psi$, i.e. there exists K, independent of δ_o, such that in $A_{\kappa d/2}$,

$$|D^2\psi| \leq K. \tag{4.17}$$

We claim that $|\psi|$ is also uniformly bounded above, i.e. for $x \in A_{\kappa d}$,

$$|\psi|(x) \leq L, \tag{4.18}$$

for some constant L, independent of δ_o small. To see this, let

$$\delta = osc_{A_{\kappa d}} u \leq \delta_o. \tag{4.19}$$

Then for $\lambda = \frac{\delta}{\mu}$, the function $\bar{\psi} = \psi/\lambda = (1-u)/\delta$ has oscillation equal to 1 on $A_{\kappa d}$. Now if $\lambda >>$ 1, it follows from (4.17) that

$$|D^2\bar{\psi}| << 1,$$

so that $\bar{\psi} = (1-u)/\delta$ is almost an affine function on $A_{\kappa d}$. For λ sufficiently large, this however contradicts the fact that the level set L_o in (4.7) is compact and contained in $A_{\kappa d}$. Hence, λ is bounded above for all δ_o sufficiently small and so, via (4.10), (4.18) holds.

By (4.13) and elliptic regularity again, the L^∞ norm of $\frac{r}{\mu}$ is on the order of 1, and hence by (4.16), ψ cannot be too close to 0, or to any affine function on $A_{\kappa d}$. In particular since (4.19) gives

$osc_{A_{\kappa d}}\psi = \frac{\delta}{\mu}$, λ is also bounded below away from 0. Thus the ratio μ/δ is bounded away from 0 and ∞, independent of δ_o small, i.e.

$$\mu \sim \delta. \tag{4.20}$$

Hence, from (4.15) one may write

$$u = 1 - \delta\nu, \tag{4.21}$$

where ν differs from ψ by a bounded scale factor; in particular ν is almost a positive harmonic function on $(A_{\kappa d}, g')$, uniformly bounded away from 0 and ∞, independent of δ_o.

Similarly, since $\frac{r}{\delta} \sim 1$, i.e. formally the linearization of the curvature r is bounded, we may write

$$g = g_o + \delta h + o(\delta), \tag{4.22}$$

where $|h| \sim 1$ in the C^o topology, and hence $|h| \sim 1$ in the C^k topology, by elliptic regularity for the $\mathcal{Z}_c^2$ equations as above.

Now as in [5, §7] for instance, consider the conformally equivalent metric

$$\widetilde{g} = u^2 \cdot g. \tag{4.23}$$

A standard computation of the Ricci curvature under conformal changes, c.f. [6, Ch.1J], gives

$$\widetilde{r} = r - u^{-1}D^2u + 2(d\log u)^2 - u^{-1}\Delta u \cdot g =$$
$$= 2(d\log u)^2 - \frac{\alpha}{2u}|z|^2 \cdot g - \frac{\alpha}{u}\nabla\mathcal{Z}^2.$$

where the second equality follows from (4.11)-(4.12). Since $|\nabla\mathcal{Z}^2| = O(\delta^2), \alpha = O(R^{-2})$ and, from (4.19), $|d\log u| = O(\delta)$, it follows that

$$\widetilde{r} \sim 2(du)^2 = O(\delta^2), \tag{4.24}$$

and hence $\widetilde{g}$ is flat to order δ^2, i.e.

$$\widetilde{g} = g_o + O(\delta^2).$$

Since $u = 1 - \delta\nu$, we obtain the expansion

$$g' = (1 + 2\nu\delta)g_o + o(\delta) \tag{4.25}$$

in $A_{\kappa d}$. This improves the estimate (4.22), i.e. shows that

$$h = 2\nu g_o,$$

so that to first order in δ, the metric g_o differs only conformally from the flat metric g_o. Note in particular that this form of the metric agrees with the form (0.17) of g in an asymptotically flat end, with ν then corresponding to m/r - a multiple of the Green's function on $\mathbb{R}^3$.

Having identified the form of g' and u to first order in δ, we now are in position to carry out the metric sphere surgery of Theorem 3.1 under these circumstances. Thus, as before, we argue by contradiction and assume that the 2-sphere S^2, essential in A, bounds a 3-ball in M.

Note first that we need only consider the situation where $S^2 \subset A$ does not bound an end $E \subset N$ on which osc $u \leq \delta_o$. For in this situation, the assumptions of Proposition 4.1 are satisfied, and so (N, g') has an asymptotically flat end; the proof of Theorem 4.2 then proceeds as in Theorem 3.1. Hence for the remainder of the proof, assume that S^2 does not bound such an end in N, but does bound in M.

Thus, suppose the S^2 in A bounds on the inside in M, i.e. the component containing the base point y and bounding S^2 is a 3-ball in M. This implies that (a smooth approximation to) the outer boundary $\partial_o A_{\kappa d}$ of $A_{\kappa d}$ bounds a compact domain W in $\mathbb{R}^3$, with $W \subset M$. Hence the level set L_o of ν from (4.7) bounds a compact subdomain, (not necessarily a ball), in W.

The metric g' is real-analytic in A, as is the potential function u, c.f. Theorem 1.3. Hence the level sets of u, and so ν, are real-analytic. Assuming the level set L_o in (4.7) is regular, i.e. there are no critical points of ν on L_o, view the level set L_o as isometrically embedded in $(\mathbb{R}^3, g_o)$. Since ν is constant on L_o, the Riemannian surface (L_o, g') embeds to first order in δ isometrically in $\mathbb{R}^3$; in fact (L_o, g') to first order is just the dilation of (L_o, g_o) by the factor $(1+2\nu\delta)$ by (4.25). It then follows from a result of [20], that (L_o, g') itself embeds isometrically in $\mathbb{R}^3$, for δ sufficiently small. Similarly, all the other regular level sets L of ν which are compactly contained in $A_{\kappa d}$ embed isometrically in (W, g_o). Such level sets L then bound a smooth compact domain $V = V(L, \delta) \subset (W, g_o)$, which as above vary with δ.

For a suitable choice of the level L, (near L_o), we use the metric

$$\bar{g} = g_o \cup g', g_o = g_o|_V, \tag{4.26}$$

as a comparison metric to g', as in Case I of the proof of Theorem 3.1. This metric is piecewise smooth, but is only C^o at the seam L.

The two main ingredients in the comparison argument in Case I of Theorem 3.1 are the relations between the 2^{nd} fundamental forms (3.8) and the volumes (3.11) of the flat metric g_o and the metric g'.

First, we prove an analogue of the estimate (3.8). From the form of the metric g' in (4.25), an easy computation for conformal metrics, (c.f. [6, Ch.1J]), gives

$$A_{g_o} - A_{g'} = \ - < \nabla\nu, X > \delta \cdot I + o(\delta), \tag{4.27}$$

where X is the unit outward normal at L and $I = g'|_L$. By construction, i.e. from the hypothesis following (4.7), we have $- < \nabla\nu, X >= |\nabla\nu| > 0$ on any regular level L. Further, since $\nu \sim 1$ and osc $\nu \sim 1$ in $A_{\kappa d}$, independent of δ_o, and so ν is uniformly controlled independent of δ_o, it follows that there exist regular levels L near L_o such that

$$|\nabla\nu|_L \geq \mu_o > 0, \tag{4.28}$$

for some positive constant μ_o, independent of δ_o.

Thus, as in (3.8), L is more convex in the flat metric g_o than in the g' metric, and so there is a concentration of positive curvature on L, provided δ_o is sufficiently small. As in (3.13), one may then smooth the metric $\bar{g}$ near L to a metric $\widetilde{g}$ satisfying

$$(s_{\widetilde{g}})^- \geq (s_{g'})^- = 0, \tag{4.29}$$

everywhere.

Observe that the blow-up metrics g'_ε limiting on g' have the same form (4.25) on $A_{\kappa d}$ provided ε is chosen sufficiently small, depending only on δ_o. Hence, one may carry out the construction above, (with the same g_o), to obtain comparison metrics $\bar{g}_\varepsilon$ to g'_ε, for which there are smoothings $\widetilde{g}_\varepsilon$ satisfying

$$(s_{\widetilde{g}_\varepsilon})^- \geq (s_{g'_\varepsilon})^-, \tag{4.30}$$

for ε sufficiently small.

To estimate the difference in the volumes, suppose first that S^2 does not bound in N; of course it does bound a 3-ball B^3 in M. In this situation,

$$vol_{g'_\varepsilon} B^3 \to \infty,$$

and so obviously, since the volume of the limit comparison flat 3-ball is finite,

$$vol_{\widetilde{g}_\varepsilon} B^3 < vol_{g'_\varepsilon} B^3, \tag{4.31}$$

for ε sufficiently small.

Hence we may assume that S^2 bounds a compact 3-ball (on the inside) in N. It follows from the discussion above that the domain in N bounded by the level set L is diffeomorphic to the flat domain V above. Thus we have two metrics, the metric g', (rescaled by R^{-2}), and the flat metric g_o, on V.

We first claim that the expansion (4.25) is valid outside a subset of (arbitrarily) small g_o volume in V, depending only on δ_o. To see this, the expression (4.25) is valid in the subdomain V_k of V on which ν satisfies

$$\nu \leq k,$$

where k may be made arbitrarily large if δ_o is chosen sufficiently small. Now the function ν is approximately a harmonic function on (V_k, g'), with boundary value $\nu \equiv$ const. on $\partial V = L$; in fact from (4.14), $|\Delta_{g'}\nu| \leq cR^{-2} \cdot \delta$. From the expression (4.25), a short computation then shows that

$$|\Delta_{g_o}(e^{\nu} - 1)| = R^{-2}O(\delta).$$

Let $v = e^{\nu} - 1$ and $v_k = \min(v, e^k)$, so that (essentially) $v_k \equiv e^k$ on $V \setminus V_k$. It follows that v_k is close to a positive superharmonic function ϕ_k on (V, g_o). It is well known, c.f. [17, Thm.5.8] that the measure of the set where such a function is very large is small, and hence

$$vol_{g_o}\{\nu \geq k\} \leq \delta_2, \tag{4.32}$$

where $\delta_2 = \delta_2(\delta, k)$. This proves the claim above.

Given these preliminaries, we now do the volume comparison. We have $vol_{g'}V \geq vol_{g'}V_k$. From the expression (4.25) and from the construction of V, c.f. the discussion preceding (4.26),

$$vol_{g'}V_k = vol_{g_o}V_k^o + 3(\int_{V_k^o} \nu dV_{g_o})\delta + o(\delta). \tag{4.33}$$

Observe here that $V_k^o \neq V_k$, but by construction, V_k^o is the domain in $(\mathbb{R}^3, g_o)$ with $\partial V_k^o = L$, where L, (and *not* (L, g_o)) is isometrically embedded in $(\mathbb{R}^3, g_o)$; this relation is just the same as the relation between R and $\bar{R}$ in Case I of Theorem 3.1. On the other hand, since to 1^{st} order in δ, (L, g') is just the dilation of (L, g_o) by the factor $(1 + 2\nu|_L\delta)$, one has

$$vol_{g_o}V_k = (1 + 3\nu|_L\delta) \cdot vol_{g_o}V_k^o + o(\delta). \tag{4.34}$$

Thus, modulo lower order terms in δ,

$$vol_{g_o}V_k^o \sim (1 - 3\nu|_L\delta) \cdot vol_{g_o}V_k. \tag{4.35}$$

Since, again by the hypothesis following (4.7) and the maximum principle,

$$\frac{1}{vol_{g_o}V_k}\int_{V_k} \nu dV_{g_o} > \nu|_L, \tag{4.36}$$

independent of the size of δ, it follows from (4.33)-(4.36) that

$$vol_{g'}V > vol_{g_o}V_k, \tag{4.37}$$

for δ sufficiently small. This estimate together with (4.32), gives

$$vol_{g'}V > vol_{g_o}V, \tag{4.38}$$

for δ small. The estimate (4.38) will of course also hold for g'_ε in place of g', and $\widetilde{g}_\varepsilon$ in place of g_o, for ε sufficiently small, and so we have the analogue of (3.11), i.e.

$$vol_{g'_\varepsilon}V > vol_{\widetilde{g}_\varepsilon}V. \tag{4.39}$$

The estimate for the comparison for $\mathcal{Z}^2$ is essentially the same as in Case I of Theorem 3.1. Thus the metric $\widetilde{g}$ has $z \equiv 0$ inside the seam L, while in a small band T about L where $\bar{g}$ is smoothed,

$$\int_T |\widetilde{z}|^2 dV_{\widetilde{g}} \leq \delta_3, \tag{4.40}$$

where δ_3 may be made arbitrarily small by choosing δ_o sufficiently small. On the other hand, the limit (N, g', y) has a definite amount of curvature inside S^2, exactly as in (3.18), since R is large. Thus, for the same reasons as in (3.19)-(3.20), one obtains in this situation

$$I_{\varepsilon}^{-}(\widetilde{g}_{\varepsilon}) < I_{\varepsilon}^{-}(g'_{\varepsilon}), \tag{4.41}$$

which is impossible by the minimizing property of g'_{ε}. This completes the proof.

■

Remark 4.3. In the context of Theorem 4.2, suppose the 2-sphere S^2 in A bounds a 3-ball on the outside in M. As remarked in the proof above, if S^2 bounds in $U^o \subset N$, then again Proposition 4.1 proves that there is an asymptotically flat sub-end, and one may apply Theorem 3.1.

Thus suppose S^2 bounds a 3-ball B on the outside, but does not bound in U^o. In this case, if there is another larger annulus A', i.e. a component of a larger geodesic annulus $A((1-d)R', (1+d)R')$ with $R' > 2R$, satisfying the assumptions of Theorem 4.2, then $A' \subset M$ is topologically contained in the 3-ball B. Hence, the 2-sphere $(S^2)'$ in A' bounds a 3-ball on the inside, and one obtains a contradiction again from the proof of Theorem 4.2. Hence, under such circumstances, the (original) S^2 cannot bound a 3-ball in M on either side.

As an example of such a situation, one might (possibly) have blow-up limit $\mathcal{Z}_c^2$ solutions (N, g') which for instance are topologically the double of $\mathbb{R}^3 \setminus \cup B_i$, where B_i is a countable collection of disjoint 3-balls in $\mathbb{R}^3$, i.e. N is an infinite connected sum of $\mathbb{R}^3$'s. Such a manifold cannot be asymptotically flat, but might satisfy the preceding condition.

References

[1] *M. Anderson*, Scalar curvature and geometrization conjectures for 3-manifolds, Comparison Geometry, MSRI Publ., **30**, (1997), 49-82.

[2] *M. Anderson*, Extrema of curvature functionals on the space of metrics on 3-manifolds, Calc. Var. & P.D.E., **5**, (1997), 199-269.

[3] *M. Anderson*, Extrema of curvature functionals on the space of metrics on 3-manifolds, II, Calc.Var. & P.D.E., **12**, (2001), 1-58.

[4] *M. Anderson*, Scalar curvature, metric degenerations and the static vacuum Einstein equations on 3-manifolds, I, Geom. & Funct. Anal., **9**, (1999), 855-967.

[5] *M. Anderson*, Scalar curvature, metric degenerations and the static vacuum Einstein equations on 3-manifolds, II, Geom. & Funct. Anal., **11**, (2001), 273-381.

The papers above are also available at http://www.math.sunysb.edu/ ∼ anderson/

[6] *A. Besse*, Einstein Manifolds, Ergebnisse der Mathematik, 3. Folge, Band 10, Springer Verlag, New York (1987).

[7] *J. Cheeger and M. Gromov*, Collapsing Riemannian manifolds while keeping their curvature bounded I, Jour. Diff. Geom., **23**, (1986), 309-346.

[8] *J. Cheeger and M. Gromov*, Collapsing Riemannian manifolds while keeping their curvature bounded II, Jour. Diff. Geom., **32**, (1990), 269-298.

[9] *K. Fukaya*, Collapsing Riemannian manifolds to ones of lower dimension, Jour. Diff. Geom., **25**, (1987), 139-156.

[10] *S. Gallot*, Isoperimetric inequalities based on integral norms of Ricci curvature, Asterisque **157-158**, (1988), 191-217.

[11] *D. Gilbarg and N. Trudinger*, Elliptic Partial Differential Equations of the Second Order, 2nd Edition, Springer Verlag, New York, 1983.

[12] *M. Gromov*, Metric Structures for Riemannian and Non-Riemannian Spaces, Progress in Math., **152**, Birkhauser Verlag, Boston, (1999).

[13] *M. Gromov*, Volume and bounded cohomology, Publ. Math. I.H.E.S., **56**, (1982), 5-100.

[14] *M. Gromov and H. Lawson, Jr.*, Spin and scalar curvature in the presence of a fundamental group, Annals of Math., **111**, (1980), 209-230.

[15] *M. Gromov and H. Lawson, Jr.*, Positive scalar curvature and the Dirac operator on complete Riemannian manifolds, Publ. Math. I.H.E.S., **58**, (1983), 83-196.

[16] *R. Hamilton*, Non-singular solutions of the Ricci flow on three-manifolds, Comm. Geom. Anal, **7**, (1999), 695-729.

[17] *W. Hayman*, Subharmonic Functions, Academic Press, London, (1976).
[18] *W. Jaco*, Lectures on Three-Manifold Topology, Conf. Bd. Math. Sci., A.M.S., **43**, (1980).
[19] *W. Jaco and P. Shalen*, Seifert fibered spaces in 3-manifolds, Memoirs A.M.S., **220**, (1979).
[20] *H. Jacobowitz*, Local isometric embeddings of surfaces into Euclidean four space, Indiana Univ. Math. Jour., **21**, (1971), 249-254.
[21] *K. Johannson*, Homotopy equivalence of 3-manifolds with boundaries, Lect. Notes in Math., **761**, Springer Verlag, (1979).
[22] *M. Kneser*, Geschlossene Flächen in dreidimensionale Mannifaltigkeiten, Jahresber. Deutsch. Math. Verein, **38**, (1929), 248-260.
[23] *O. Kobayashi*, Scalar curvature of a metric of unit volume, Math. Annalen, **279**, (1987), 253-265.
[24] *J. Milnor*, A unique factorization theorem for 3-manifolds, Amer. J. Math., **84**, (1962), 1-7.
[25] *G. Mostow*, Quasiconformal mappings in n-space and the strong rigidity of hyperbolic space forms, Publ. Math. I.H.E.S., **34**, (1968), 53-104.
[26] *P. Orlik*, Seifert Manifolds, Lect. Notes in Math., vol. 291, Springer Verlag, New York, (1972).
[27] *G. Prasad*, Strong rigidity of Q-rank 1 lattices, Inventiones Math., **21**, (1973), 255-286.
[28] *X. Rong*, The limiting eta invariant of collapsed 3-manifolds, Jour. Diff. Geom., **37**, (1993), 535-568.
[29] *R. Schoen*, Conformal deformation of a metric to constant scalar curvature, Jour. Diff. Geom., **20**, (1985), 479-495.
[30] *R. Schoen*, Variational theory for the total scalar curvature functional for Riemannian metrics and related topics, Lect. Notes Math. **1365**, Springer Verlag, Berlin, (1987), 120-154.
[31] *R. Schoen and S.-T. Yau*, Existence of incompressible minimal surfaces and the topology of three dimensional manifolds with non negative scalar curvature, Annals of Math., **110**, (1979), 127-142.
[32] *R. Schoen and S.-T. Yau*, On the structure of manifolds with positive scalar curvature, Manuscripta Math., **29**, (1979), 159-183.
[33] *P. Scott*, The geometries of 3-manifolds, Bull. London Math. Soc., **15**, (1983), 401-487.
[34] *T. Soma*, The Gromov invariant of links, Inventiones Math., **64**, (1981), 445-454.
[35] *M. Spivak*, A Comprehensive Introduction to Differential Geometry, V, Publish or Perish, Houston, 1979.
[36] *W. Thurston*, The Geometry and Topology of 3-Manifolds, (preprint), Princeton, 1979.
[37] *W. Thurston*, Three dimensional manifolds, Kleinian groups and hyperbolic geometry, Bulletin A.M.S., **6**, (1982), 357-381.
[38] *F. Waldhausen*, Eine Klasse von 3-dimensionalen Mannigfaltigkeiten I, Inventiones Math., **3**, (1967), 308-333.
[39] *F. Waldhausen*, Eine Klasse von 3-dimensionalen Mannigfaltigkeiten II, Inventiones Math., **4**, (1967), 87-117.

September, 2000

Department of Mathematics
S.U.N.Y. at Stony Brook
Stony Brook, N.Y. 11794-3651
anderson@math.sunysb.edu