The Penrose Inequality

Hubert L. Bray*
Piotr T. Chruściel†

August 31, 2004

Abstract

In 1973, R. Penrose presented an argument that the total mass of a space-time which contains black holes with event horizons of total area A should be at least $\sqrt{A/16\pi}$. An important special case of this physical statement translates into a very beautiful mathematical inequality in Riemannian geometry known as the Riemannian Penrose inequality. This inequality was first established by G. Huisken and T. Ilmanen in 1997 for a single black hole and then by one of the authors (HB) in 1999 for any number of black holes. The two approaches use two different geometric flow techniques and are described here. We further present some background material concerning the problem at hand, discuss some applications of Penrose-type inequalities, as well as the open questions remaining.

AMS Classification Scheme: 83C05, 53C80

Contents

*Mathematics Department, 2-179, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA, bray@math.mit.edu. Research supported in part by NSF grants #DMS-0206483 and #DMS-9971960 and by the Erwin Schrödinger Institute, Vienna.

†Département de Mathématiques, Faculté des Sciences, Parc de Grandmont, F37200 Tours, France, Piotr.Chrusciel@lmpt.univ-tours.fr. Partially supported by a Polish Research Committee grant 2 P03B 073 24, by the Erwin Schrödinger Institute, Vienna, and by a travel grant from the Vienna City Council; URL `www.phys.univ-tours.fr/~piotr`

1 Introduction

1.1 What is the Penrose Conjecture?

We will restrict our attention to statements about space-like slices (M^3, g, h) of a space-time, where g is the positive definite induced metric on M^3 and h is the second fundamental form of M^3 in the space-time. From the Einstein equation $G = 8\pi T$, where G is the Einstein curvature tensor and T is the stress-energy tensor, it follows from the Gauss and Codazzi equations that

$$\mu = \frac{1}{8\pi} G^{00} = \frac{1}{16\pi}[R - \sum_{i,j} h^{ij} h_{ij} + (\sum_i h_i{}^i)^2], \tag{1.1}$$

$$J^i = \frac{1}{8\pi} G^{0i} = \frac{1}{8\pi} \sum_j \nabla_j [h^{ij} - (\sum_k h_k{}^k) g^{ij}], \tag{1.2}$$

where μ and J are respectively the energy density and the current vector density at each point of M^3. Then the physical assumption of nonnegative energy density everywhere in the space-time as measured by observers moving in all future-pointing, time-like directions (known as the dominant energy condition) implies that

$$\mu \geq |J| \tag{1.3}$$

everywhere on M^3. Hence, we will only consider Cauchy data (M^3, g, h) which satisfy inequality (1.3).

The final assumption we will make is that (M^3, g, h) is *asymptotically flat*, which will be discussed in more detail below. Typically, one assumes that M^3 consists of a compact set together with one or more asymptotically flat "ends", each diffeomorphic to the complement of a ball in $\mathbb{R}^3$. For example, $\mathbb{R}^3$ has one end, whereas $\mathbb{R}^3 \# \mathbb{R}^3$ has two ends.

Penrose's motivation for the Penrose Conjecture [70] goes as follows: Suppose we begin with Cauchy data (M^3, g, h) which is asymptotically flat (so that total mass of a chosen end is defined) and satisfies $\mu \geq |J|$ everywhere. Using this as initial data, solve the Einstein equation forward in time, and suppose that the resulting space-time is asymptotically flat in null directions so that the Trautman-Bondi mass is defined for all retarded times. Suppose further that the space-time eventually settles down to a Kerr solution, so that the Trautman-Bondi mass asymptotes to the ADM mass of the relevant Kerr solution. By the Hawking Area Theorem [48] (compare [29]), the total area of the event horizons of any black holes does not decrease, while the total Trautman-Bondi mass of the system — which is expected to approach the ADM mass at very early advanced times — does not increase. Since Kerr solutions all have

$$m \geq \sqrt{A_e/16\pi}, \tag{1.4}$$

where m is total ADM mass [5, 38] and A_e is the total area of the event horizons, we must have this same inequality for the original Cauchy data (M^3, g, h).

The reader will have noticed that the above argument makes a lot of global assumptions about the resulting space-times, and our current understanding of the associated mathematical problems is much too poor to be able to settle those one way or another. The conjecture that (all, or at least a few key ones of) the above global properties are satisfied is known under the name of *Penrose's cosmic censorship hypothesis*. We refer the reader to the article by Lars Andersson in this volume and references therein for more information about that problem.

A natural interpretation of the Penrose inequality is that the mass contributed by a collection of black holes is not less than $\sqrt{A/16\pi}$. More generally, the question "How much matter is in a given region of a space-time?" is still very much an open problem [23]. In this paper, we will discuss some of the qualitative aspects of mass in general relativity, look at examples which are informative, and describe the two very geometric proofs of the Riemannian Penrose inequality. The most general version of the Penrose inequality is still open and is discussed in section 4.2. The notes here are partly based on one of the author's (HB) lectures at the "Fifty Years of the Cauchy Problem in General Relativity" Summer School held in August 2002 in Cargèse (videos of lectures available at URL `http://fanfreluche.math.univ-tours.fr`, or on the DVD enclosed with this volume), and some sections draw substantially on his review paper [17], following a suggestion of the editors of this volume. The mathematically oriented reader with limited knowledge of the associated physics might find it useful to become acquainted with [17] before reading the current presentation.

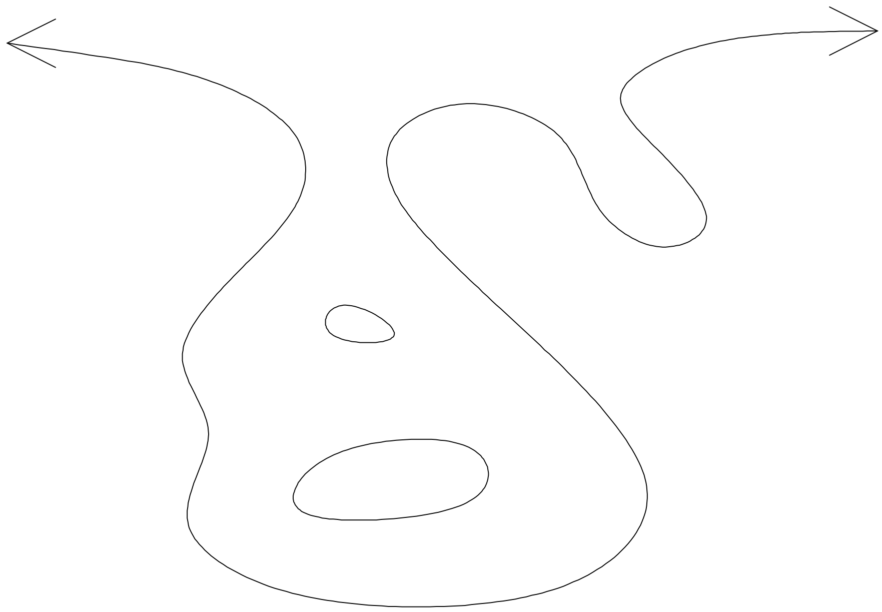
1.2 Total Mass in General Relativity

Amongst the notions of mass which are well understood in general relativity are local energy density at a point, the total mass of an asymptotically flat space-time (whether at spacelike or at null infinity; the former is usually called the ADM mass while the latter the Trautman-Bondi mass), and the total mass of an asymptotically anti-de Sitter space-time (often called the Abbott-Deser mass). On the other hand, defining the mass of a region larger than a point but smaller than the entire universe is not very well understood at all. While we will return to this last question in Section 4.3 below, we start here with a discussion of the ADM mass.

Suppose (M^3, g) is a Riemannian 3-manifold isometrically embedded in a (3+1) dimensional Lorentzian space-time N^4. Suppose that M^3 has zero second fundamental form in the space-time. (Recall that the second fundamental form is a measure of how much M^3 curves inside N^4. M^3 is also sometimes called "totally geodesic" since geodesics of N^4 which are tangent to M^3 at a point stay inside M^3 forever.) The Penrose inequality (which in its full generality allows for M^3 to have non-vanishing second fundamental form) is known as the *Riemannian Penrose inequality* when the second fundamental form is set to zero.[1]

In this work we will mainly consider (M^3, g) that are asymptotically flat at infinity, which means that for some compact set K, the "end" $M^3\backslash K$ is diffeomorphic to $\mathbb{R}^3\backslash B_1(0)$, where the metric g is asymptotically approaching (with the decay conditions (1.7) below) the standard flat metric δ_{ij} on $\mathbb{R}^3$ at infinity. The simplest example of an asymptotically flat manifold is $(\mathbb{R}^3, \delta_{ij})$ itself. Other good examples are the conformal metrics $(\mathbb{R}^3, u(x)^4\delta_{ij})$, where $u(x)$ approaches a constant sufficiently rapidly at infinity. (Also, sometimes it is convenient to allow (M^3, g) to have multiple asymptotically flat ends, in which case each connected component of $M^3\backslash K$ must have the property described above.) A qualitative picture of an asymptotically flat 3-manifold is shown below.

[1] This terminology is somewhat misleading, in the following sense: the results discussed below hold as soon as the scalar curvature is non-negative. This will certainly be the case if $h_{ij} = 0$ and $\mu \geq 0$ in (1.1), but *e.g.* $\sum_i {h_i}^i = 0$, or various other conditions in this spirit, suffice.



The assumptions on the asymptotic behavior of (M^3, g) at infinity will be tailored to imply the existence of the limit

$$m = \frac{1}{16\pi} \lim_{\sigma\to\infty} \int_{S_\sigma} \sum_{i,j} (g_{ij,i}\nu_j - g_{ii,j}\nu_j)\, d\mu \tag{1.5}$$

where S_σ is the coordinate sphere of radius σ, ν is the unit normal to S_σ, and $d\mu$ is the area element of S_σ in the coordinate chart. The quantity m is called the *total mass* (or ADM mass [5]) of (M^3, g). Equation (1.5) begs the question of the geometric character of the number m: the integrand contains partial derivatives of a tensor, which makes it coordinate dependent. For example, if $g = \delta$ is the flat metric in the standard orthogonal coordinates x^i, one clearly obtains zero. On the other hand, we can introduce a new coordinate system (ρ, θ, ϕ) by changing the radial variable r to

$$r = \rho + c\rho^{1-\alpha} \;, \tag{1.6}$$

with some constants $\alpha > 0$, $c \in \mathbb{R}$. In the associated asymptotically Euclidean coordinate system $y^i = \rho x^i / r$ the metric tensor approaches δ as $O(|y|^{-\alpha})$:

$$\delta_{ij} dx^i dx^j = g_{ij} dy^i dy^j \;,$$

with

$$g_{ij} - \delta_{ij} = \mathcal{O}(|y|^{-\alpha}) \;,\; \partial_k g_{ij} = \mathcal{O}(|y|^{-\alpha-1}) \;. \tag{1.7}$$

A short calculation gives

$$m \quad = \quad \begin{cases} \infty \;, & \alpha < 1/2 \;, \\ c^2/8 \;, & \alpha = 1/2 \;, \\ 0 \;, & \alpha > 1/2 \;. \end{cases}$$

Thus, the mass m of the flat metric in the coordinate system y^i is infinite if $\alpha < 1/2$, can have an arbitrary positive value depending upon c if $\alpha = 1/2$, and vanishes for $\alpha > 1/2$. (Negative values of m can also be obtained by deforming the slice $\{t = 0\}$ within Minkowski space-time [25] when the decay rate $\alpha = 1/2$ is allowed.) The lesson of this is that the mass appears to depend upon the coordinate system chosen, even within the class of coordinate systems in which the metric tends to a constant coefficients matrix as r tends to infinity. It can be shown that the decay rate $\alpha = 1/2$ is precisely the borderline for a well defined mass: the mass is an invariant in the class of coordinate systems satisfying (1.7) with $\alpha > 1/2$ and with $R \in L^1(M)$ [6, 26][2]. We note that the above example is essentially due to Denisov and Solov'ev [37], and that the geometric character of m in a space-time setting is established in [27].

Going back to the example $(\mathbb{R}^3, u(x)^4\delta_{ij})$, if we suppose that $u(x) > 0$ has the asymptotics at infinity

$$u(x) = a + b/|x| + \mathcal{O}(1/|x|^2) \ , \tag{1.8}$$

with the derivatives of the $\mathcal{O}(1/|x|^2)$ term being $\mathcal{O}(1/|x|^3)$, then the total mass of (M^3, g) is

$$m = 2ab. \tag{1.9}$$

Furthermore, suppose (M^3, g) is any metric whose "end" is isometric to $(\mathbb{R}^3 \backslash K, u(x)^4\delta_{ij})$, where $u(x)$ is harmonic in the coordinate chart of the end $(\mathbb{R}^3 \backslash K, \delta_{ij})$ and goes to a constant at infinity. Then expanding $u(x)$ in terms of spherical harmonics demonstrates that $u(x)$ satisfies condition (1.8). We will call these Riemannian manifolds (M^3, g) *harmonically flat at infinity*, and we note that the total mass of these manifolds is also given by equation (1.9).

A very nice lemma by Schoen and Yau [73] is that, given any $\epsilon > 0$, it is always possible to perturb an asymptotically flat manifold to become harmonically flat at infinity such that the total mass changes less than ϵ and the metric changes less than ϵ pointwise, all while maintaining nonnegative scalar curvature (discussed in a moment). Hence, it happens that to prove the theorems in this paper, we only need to consider harmonically flat manifolds. Thus, we can use equation (1.9) as our definition of total mass. As an example (already pointed out), note that $(\mathbb{R}^3, \delta_{ij})$ has zero total mass. Also, note that, qualitatively, the total mass of an asymptotically flat or harmonically flat manifold is the $1/r$ rate at which the metric becomes flat at infinity.

A deep (and considerably more difficult to prove) result of Corvino [34] (compare [28, 35]) shows that if m is non zero, then one can always perturb an asymptotically flat manifold as above while maintaining zero scalar curvature and achieving (1.8) without any error term.

We finish this section by noting the following "isotropic coordinates" representation of the *exterior Schwarzschild space-time metric*

$$\left(\mathbb{R} \times \left(\mathbb{R}^3 \setminus B_{m/2}(0)\right), (1 + \frac{m}{2|x|})^4(dx_1^2 + dx_2^2 + dx_3^2) - \left(\frac{1 - m/2|x|}{1 + m/2|x|}\right)^2 dt^2\right) . \tag{1.10}$$

The $t = 0$ slice (which has zero second fundamental form) is the *exterior spacelike Schwarzschild metric*

$$\left(\mathbb{R}^3 \backslash B_{m/2}(0), (1 + \frac{m}{2|x|})^4 \delta_{ij}\right) . \tag{1.11}$$

According to equation (1.9), the parameter m is of course the total mass of this 3-manifold.

The above example also allows us to make a connection between what we have arbitrarily defined to be total mass and our more intuitive Newtonian notions of

[2]Actually the results in [6] use weighted Sobolev conditions on two derivatives of the metric, suggesting that the right decay conditions in (1.7) are $o(|y|^{-1/2})$ for the metric and $o(|y|^{-3/2})$ for its derivatives. It can be checked that the argument in [26] generalises, and gives the result under those conditions.

mass. Using the natural Lorentzian coordinate chart as a reference, one can compute that geodesics in the Schwarzschild space-time metric are curved when $m \neq 0$. Furthermore, if one interprets this curvature as acceleration due to a force coming from the central region of the manifold, one finds that this fictitious force yields an acceleration asymptotic to m/r^2 for large r. Hence, a test particle left to drift along geodesics far out in the asymptotically flat end of the Schwarzschild spacetime "accelerates" according to Newtonian physics as if the total mass of the system were m.

1.3 Example Using Superharmonic Functions in $\mathbb{R}^3$

Once again, let us return to the $(\mathbb{R}^3, u(x)^4\delta_{ij})$ example. The formula for the scalar curvature is

$$R(x) = -8u(x)^{-5}\Delta u(x).$$

Hence, since the physical assumption of nonnegative energy density implies nonnegative scalar curvature, we see that $u(x) > 0$ must be superharmonic ($\Delta u \leq 0$). For simplicity, let us also assume that $u(x)$ is harmonic outside a bounded set so that we can expand $u(x)$ at infinity using spherical harmonics. Hence, $u(x)$ has the asymptotics of equation (1.8). By the maximum principle, it follows that the minimum value for $u(x)$ must be a, referring to equation (1.8). Hence, $b \geq 0$, which implies that $m \geq 0$. Thus we see that the assumption of nonnegative energy density at each point of $(\mathbb{R}^3, u(x)^4\delta_{ij})$ implies that the total mass is also nonnegative, which is what one would hope.

1.4 The Positive Mass Theorem

Suppose we have any asymptotically flat manifold with nonnegative scalar curvature, is it true that the total mass is also nonnegative? The answer is *yes*, and this fact is known as the positive mass theorem, first proved by Schoen and Yau [72] in 1979 using minimal surface techniques and then by Witten [80] in 1981 using spinors. (The mathematical details needed for Witten's argument have been worked out in [6, 21, 51, 69].) In the zero second fundamental form case, also known as the *time-symmetric* case, the positive mass theorem is known as the Riemannian positive mass theorem and is stated below.

Theorem 1.1 *(Schoen, Yau [72]) Let (M^3, g) be any asymptotically flat, complete Riemannian manifold with nonnegative scalar curvature. Then the total mass $m \geq 0$, with equality if and only if (M^3, g) is isometric to $(\mathbb{R}^3, \delta)$.*

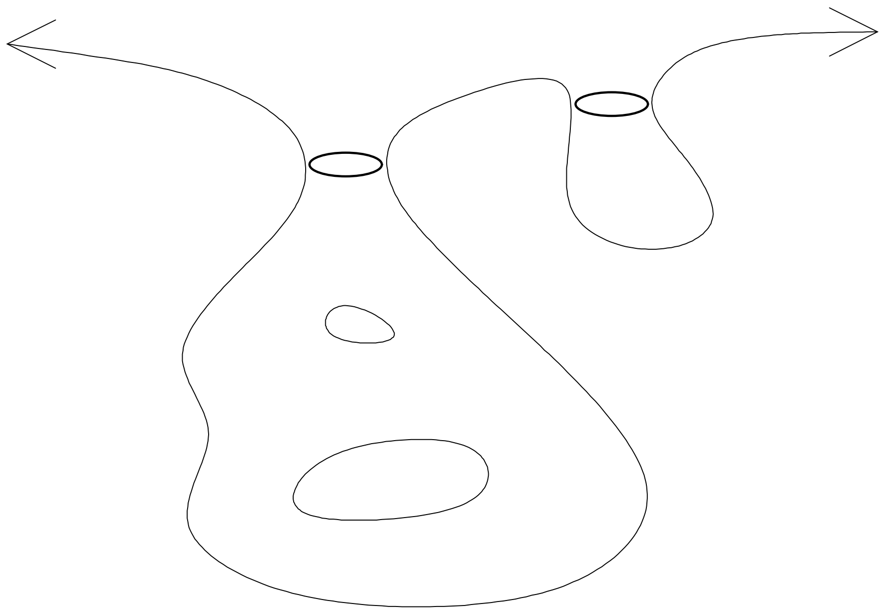
1.5 Apparent horizons

Given a surface in a space-time, suppose that it emits an outward shell of light. If the surface area of this shell of light is decreasing everywhere on the surface, then this is called a trapped surface.[3] The outermost boundary of these trapped surfaces is called the apparent horizon. Apparent horizons can be computed in terms of Cauchy data, and under appropriate global hypotheses an apparent horizon implies the existence of an event horizon outside of it [48, 78] in the time-symmetric case. The reader is referred to [24] for a review of what is known about apparent horizons; further recent results include [36, 67].

[3] The reader is warned that several authors require the trapping of both outwards and inwards shells of light in the definition of trapped surface. The inwards null directions are irrelevant for our purposes, and they are therefore ignored in the definition here.

Now let us return to the case where (M^3, g) is a "$t = 0$" slice of a space-time with zero second fundamental form. Then apparent horizons of black holes intersected with M^3 correspond to the connected components of the outermost minimal surface Σ_0 of (M^3, g).

All of the surfaces we are considering in this paper will be required to be smooth boundaries of open bounded regions, so that outermost is well-defined with respect to a chosen end of the manifold [15]. A minimal surface in (M^3, g) is a surface which is a critical point of the area function with respect to any smooth variation of the surface. The first variational calculation implies that minimal surfaces have zero mean curvature. The surface Σ_0 of (M^3, g) is defined as the boundary of the union of the open regions bounded by all of the minimal surfaces in (M^3, g). It turns out that Σ_0 also has to be a minimal surface, so we call Σ_0 the *outermost minimal surface*. A qualitative sketch of an outermost minimal surface of a 3-manifold is shown below.



We will also define a surface to be *(strictly) outer minimising* if every surface which encloses it has (strictly) greater area. Note that outermost minimal surfaces are strictly outer minimising. Also, we define a *horizon* in our context to be any minimal surface which is the boundary of a bounded open region.

It also follows from a stability argument (using the Gauss-Bonnet theorem interestingly) that each component of an outermost minimal surface (in a 3-manifold with nonnegative scalar curvature) must have the topology of a sphere [68].

Penrose's argument [70], presented in Section 1.1, suggests that the mass contributed by the black holes (thought of as the connected components of Σ_0) should be at least $\sqrt{A_0/16\pi}$, where A_0 is the area of Σ_0. This leads to the following geometric statement:

The Riemannian Penrose Inequality *Let (M^3, g) be a complete, smooth, 3-manifold with nonnegative scalar curvature which is harmonically flat at infinity with total mass m and which has an outermost minimal surface Σ_0 of area A_0. Then*

$$m \geq \sqrt{\frac{A_0}{16\pi}}, \tag{1.12}$$

with equality if and only if (M^3, g) *is isometric to the Schwarzschild metric* $(\mathbb{R}^3\backslash\{0\}, (1 + \frac{m}{2|x|})^4\delta_{ij})$ *outside their respective outermost minimal surfaces.*

The above statement has been proved by one of us (HB) [15], and Huisken and Ilmanen [57] proved it when A_0 is defined instead to be the area of the largest connected component of Σ_0. In this paper we will discuss both approaches, which are very different, although they both involve flowing surfaces and/or metrics.

We also clarify that the above statement is with respect to a chosen end of (M^3, g), since both the total mass and the definition of outermost refer to a particular end. In fact, nothing very important is gained by considering manifolds with more than one end, since extra ends can always be compactified as follows: Given an extra asymptotically flat end, we can use a lemma of Schoen and Yau [73] to make the end harmonically flat outside a bounded region. By an extension of this result in the thesis of one of the authors (HB) [14], or using the Corvino-Schoen construction [34], we can make the end exactly Schwarzschild outside a bounded set while still keeping nonnegative scalar curvature. We then replace the interior Schwarzschild region by an object often referred to as "a bag of gold", one way of doing it proceeds as follows: Since we are now in the class of spherically symmetric manifolds, we can then "round the metric up" to be an extremely large spherical cylinder outside a bounded set. This can be done while keeping nonnegative scalar curvature since the Hawking mass increases during this procedure and since the rate of change of the Hawking mass has the same sign as the scalar curvature in the spherically symmetric case (as long as the areas of the spheres are increasing). Finally, the large cylinder can be capped off with a very large sphere to compactify the end.

Hence, we will typically consider manifolds with just one end. In the case that the manifold has multiple ends, we will require every surface (which could have multiple connected components) in this paper to enclose all of the ends of the manifold except the chosen end.

1.6 The Schwarzschild Metric

The (spacelike) Schwarzschild metric $(\mathbb{R}^3\backslash\{0\}, (1 + \frac{m}{2|x|})^4\delta_{ij})$ (compare (1.11)), referred to in the above statement of the Riemannian Penrose Inequality, is a particularly important example to consider, and corresponds to a zero-second fundamental form, space-like slice of the usual (3+1)-dimensional Schwarzschild metric.

The 3-dimensional Schwarzschild metrics with total mass $m > 0$ are characterised by being the only spherically symmetric, geodesically complete, zero scalar curvature 3-metrics, other than $(\mathbb{R}^3, \delta_{ij})$. Note that this flat metric on $\mathbb{R}^3$ may be interpreted as the $m = 0$ case of the Schwarzschild metric. Negative values of m also give Schwarzschild metrics, but these metrics are not geodesically complete since they have a curvature singularity at the coordinate sphere $r = -m/2$. If this singularity is smoothed out in a spherically symmetric way, the resulting metric has very concentrated negative energy density (and scalar curvature) in the smoothed out region, which violates the assumption of positive energy density used throughout this paper.

The 3-dimensional Schwarzschild metrics with total mass $m > 0$ can also be embedded in 4-dimensional Euclidean space (x, y, z, w) as the set of points satisfying $|(x, y, z)| = \frac{w^2}{8m} + 2m$, which is a parabola rotated around an S^2. This last picture allows us to see that the Schwarzschild metric, which has two ends, has a Z_2 symmetry which fixes the sphere with $w = 0$ and $|(x, y, z)| = 2m$, which is clearly minimal. Furthermore, the area of this sphere is $4\pi(2m)^2$, giving equality in the Riemannian Penrose Inequality.

1.7 A Brief History of the Problem

The Riemannian Penrose Inequality has a rich history spanning nearly three decades and has motivated much interesting mathematics and physics. In 1973, R. Penrose in effect conjectured an even more general version of inequality (1.12) using a very clever physical argument [70], described in Section 1.1. His observation was that a counterexample to inequality (1.12) would yield Cauchy data for solving the Einstein equations, the solution to which would likely violate the Cosmic Censor Conjecture (which says that singularities generically do not form in a space-time unless they are inside a black hole).

In 1977, Jang and Wald [58], extending ideas of Geroch [43], gave a heuristic proof of inequality (1.12) by defining a flow of 2-surfaces in (M^3, g) in which the surfaces flow in the outward normal direction at a rate equal to the inverse of their mean curvatures at each point. The Hawking mass of a surface (which is supposed to estimate the total amount of energy inside the surface) is defined to be

$$m_{Hawking}(\Sigma) = \sqrt{\frac{|\Sigma|}{16\pi}} \left(1 - \frac{1}{16\pi}\int_\Sigma H^2\right),$$

(where $|\Sigma|$ is the area of Σ and H is the mean curvature of Σ in (M^3, g)) and, amazingly, is nondecreasing under this "inverse mean curvature flow." This is seen by the fact that under inverse mean curvature flow, it follows from the Gauss equation and the second variation formula that

$$\frac{d}{dt} m_{Hawking}(\Sigma) = \sqrt{\frac{|\Sigma|}{16\pi}} \left[\frac{1}{2} + \frac{1}{16\pi}\int_\Sigma 2\frac{|\nabla_\Sigma H|^2}{H^2} + R - 2K + \frac{1}{2}(\lambda_1 - \lambda_2)^2\right] \quad (1.13)$$

when the flow is smooth, where R is the scalar curvature of (M^3, g), K is the Gauss curvature of the surface Σ, and λ_1 and λ_2 are the eigenvalues of the second fundamental form of Σ, or principal curvatures. Hence,

$$R \geq 0,$$

and

$$\int_\Sigma K \leq 4\pi \quad (1.14)$$

(which is true for any connected surface by the Gauss-Bonnet Theorem) imply

$$\frac{d}{dt} m_{Hawking}(\Sigma) \geq 0. \quad (1.15)$$

Furthermore,

$$m_{Hawking}(\Sigma_0) = \sqrt{\frac{|\Sigma_0|}{16\pi}}$$

since Σ_0 is a minimal surface and has zero mean curvature. In addition, the Hawking mass of sufficiently round spheres at infinity in the asymptotically flat end of (M^3, g) approaches the total mass m. Hence, if inverse mean curvature flow beginning with Σ_0 eventually flows to sufficiently round spheres at infinity, inequality (1.12) follows from inequality (1.15).

As noted by Jang and Wald, this argument only works when inverse mean curvature flow exists and is smooth, which is generally not expected to be the case. In fact, it is not hard to construct manifolds which do not admit a smooth inverse mean curvature flow. One of the main problems is that if the mean curvature of the evolving surface becomes zero or is negative, it is not clear how to define the flow.

For twenty years, this heuristic argument lay dormant until the work of Huisken and Ilmanen [57] in 1997. With a very clever new approach, Huisken and Ilmanen discovered how to reformulate inverse mean curvature flow using an energy minimisation principle in such a way that the new generalised inverse mean curvature flow always exists. The added twist is that the surface sometimes jumps outward. However, when the flow is smooth, it equals the original inverse mean curvature flow, and the Hawking mass is still monotone. Hence, as will be described in the next section, their new flow produced the first complete proof of inequality (1.12) for a single black hole.

Coincidentally, one of the authors (HB) found another proof of inequality (1.12), submitted in 1999, which provides the correct inequality for any number of black holes. (When the outermost horizon is not-connected, the Huisken-Ilmanen proof bounds the mass in terms of the area of its largest component, while the new argument gives the full inequality, with the sum of areas of all components.) The approach involves flowing the original metric to a Schwarzschild metric (outside the horizon) in such a way that the area of the outermost minimal surface does not change and the total mass is nonincreasing. Then since the Schwarzschild metric gives equality in inequality (1.12), the inequality follows for the original metric.

Fortunately, the flow of metrics which is defined is relatively simple, and in fact stays inside the conformal class of the original metric. The outermost minimal surface flows outward in this conformal flow of metrics, and encloses any compact set (and hence all of the topology of the original metric) in a finite amount of time. Furthermore, this conformal flow of metrics preserves nonnegative scalar curvature. We will describe this approach later in the paper.

Other contributions to the Penrose Conjecture have been made by O'Murchadha and Malec in spherical symmetry [65], by Herzlich [50,52] using the Dirac operator with spectral boundary conditions (compare [9,66]), by Gibbons in the special case of collapsing shells [44], by Tod [75] as it relates to the hoop conjecture, by Bartnik [7] for quasi-spherical metrics, by Jezierski [59,60] using adapted foliations, and by one of the authors (HB) using isoperimetric surfaces [14]. A proof of the Penrose inequality for conformally flat manifolds (but with suboptimal constant) has been given in [18]. We also mention work of Ludvigsen and Vickers [63] using spinors and Bergqvist [11], both concerning the Penrose inequality for null slices of a space-time.

Various space-time flows which could be used to prove the full Penrose inequality (see Section 4.2 below) have been proposed by Hayward [49], by Mars, Malec and Simon [64], and by Frauendiener [39]. It was independently observed by several researchers (HB, Hayward, Mars, Simon) that those are special cases of the same flow, namely flowing in the direction $\vec{I} + c(t)\vec{I}'$, where $\vec{I}$ is the inverse mean curvature vector $-\vec{H}/ < \vec{H}, \vec{H} >$ (which is required to be spacelike outward pointing), $\vec{I}'$ is the future pointing vector with the same length as $\vec{I}$ and orthogonal to $\vec{I}$ in the normal bundle to the surface, and $\vec{H}$ is the mean curvature vector of the surface in the spacetime. The function $c(t)$ is required to satisfy $-1 \leq c(t) \leq 1$ but is otherwise free, with its endpoint values corresponding to Hayward's null flows, $c(t) = 0$ corresponding to Frauendiener's flow, and $-1 \leq c(t) \leq 1$ yielding hypersurfaces satisfying the Mars, Malec, Simon condition which implies the monotonicity of the spacetime Hawking mass functional. The catch, however, is that this flow is not parabolic and therefore only exists for a positive amount of time under special circumstances. However, as observed by HB at the Penrose Ineqalities Workshop in Vienna, July 2003, there does exist a way of defining what a weak solution to the above flow is using a max-min method analogous to the notion of weak solution to inverse mean curvature flow (which minimizes an energy functional) defined by Huisken and Ilmanen [57]. Finding ways of constructing solutions which exist for an infinite amount of time (analogous to the time-symmetric inverse mean curvature flow due to Huisken and Ilmanen) is a very interesting problem to consider.

2 Inverse Mean Curvature Flow

Geometrically, Huisken and Ilmanen's idea can be described as follows. Let $\Sigma(t)$ be the surface resulting from inverse mean curvature flow for time t beginning with the minimal surface Σ_0. Define $\bar{\Sigma}(t)$ to be the outermost minimal area enclosure of $\Sigma(t)$. Typically, $\Sigma(t) = \bar{\Sigma}(t)$ in the flow, but in the case that the two surfaces are not equal, immediately replace $\Sigma(t)$ with $\bar{\Sigma}(t)$ and then continue flowing by inverse mean curvature.

An immediate consequence of this modified flow is that the mean curvature of $\bar{\Sigma}(t)$ is always nonnegative by the first variation formula, since otherwise $\bar{\Sigma}(t)$ would be enclosed by a surface with less area. This is because if we flow a surface Σ in the outward direction with speed η, the first variation of the area is $\int_\Sigma H\eta$, where H is the mean curvature of Σ.

Furthermore, by stability, it follows that in the regions where $\bar{\Sigma}(t)$ has zero mean curvature, it is always possible to flow the surface out slightly to have positive mean curvature, allowing inverse mean curvature flow to be defined, at least heuristically at this point.

It turns out that the Hawking mass is still monotone under this new modified flow. Notice that when $\Sigma(t)$ jumps outward to $\bar{\Sigma}(t)$,

$$\int_{\bar{\Sigma}(t)} H^2 \le \int_{\Sigma(t)} H^2$$

since $\bar{\Sigma}(t)$ has zero mean curvature where the two surfaces do not touch. Furthermore,

$$|\bar{\Sigma}(t)| = |\Sigma(t)|$$

since (this is a neat argument) $|\bar{\Sigma}(t)| \le |\Sigma(t)|$ (since $\bar{\Sigma}(t)$ is a minimal area enclosure of $\Sigma(t)$) and we can not have $|\bar{\Sigma}(t)| < |\Sigma(t)|$ since $\Sigma(t)$ would have jumped outward at some earlier time. This is only a heuristic argument, but we can then see that the Hawking mass is nondecreasing during a jump by the above two equations.

This new flow can be rigorously defined, always exists, and the Hawking mass is monotone, if the scalar curvature is positive. In [57], Huisken and Ilmanen define $\Sigma(t)$ to be the level sets of a scalar valued function $u(x)$ defined on (M^3, g) such that $u(x) = 0$ on the original surface Σ_0 and satisfies

$$\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) = |\nabla u| \tag{2.1}$$

in an appropriate weak sense. Since the left hand side of the above equation is the mean curvature of the level sets of $u(x)$ and the right hand side is the reciprocal of the flow rate, the above equation implies inverse mean curvature flow for the level sets of $u(x)$ when $|\nabla u(x)| \neq 0$.

Huisken and Ilmanen use an energy minimisation principle to define weak solutions to equation (2.1). Equation (2.1) is said to be weakly satisfied in Ω by the locally Lipschitz function u if for all locally Lipschitz v with $\{v \neq u\} \subset\subset \Omega$,

$$J_u(u) \le J_u(v)$$

where

$$J_u(v) := \int_\Omega |\nabla v| + v|\nabla u|.$$

It can then be seen that the Euler-Lagrange equation of the above energy functional yields equation (2.1).

In order to prove that a solution u exists to the above two equations, Huisken and Ilmanen regularise the degenerate elliptic equation (2.1) to the elliptic equation

$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{|\nabla u|^2+\epsilon^2}}\right) = \sqrt{|\nabla u|^2+\epsilon^2}.$$

Solutions to the above equation are then shown to exist using the existence of a subsolution, and then taking the limit as ϵ goes to zero yields a weak solution to equation (2.1). There are many details which we are skipping here, but these are the main ideas.

As it turns out, weak solutions $u(x)$ to equation (2.1) often have flat regions where $u(x)$ equals a constant. Hence, the levels sets $\Sigma(t)$ of $u(x)$ will be discontinuous in t in this case, which corresponds to the "jumping out" phenomenon referred to at the beginning of this section.

We also note that since the Hawking mass of the levels sets of $u(x)$ is monotone, this inverse mean curvature flow technique not only proves the Riemannian Penrose Inequality, but also gives a new proof of the Positive Mass Theorem in dimension three. This is seen by letting the initial surface be a very small, round sphere (which will have approximately zero Hawking mass) and then flowing by inverse mean curvature, thereby proving $m \geq 0$.

The Huisken and Ilmanen inverse mean curvature flow also seems ideally suited for proving Penrose inequalities for 3-manifolds which have $R \geq -6$ and which are asymptotically hyperbolic; this is discussed in more detail in Section 4.1.

Because the monotonicity of the Hawking mass relies on the Gauss-Bonnet theorem, these arguments do not work in higher dimensions, at least so far. Also, because of the need for equation (1.14), inverse mean curvature flow only proves the Riemannian Penrose Inequality for a single black hole. In the next section, we present a technique which proves the Riemannian Penrose Inequality for any number of black holes, and which can likely be generalised to higher dimensions.

3 The Conformal Flow of Metrics

Given any initial Riemannian manifold (M^3, g_0) which has nonnegative scalar curvature and which is harmonically flat at infinity, we will define a continuous, one parameter family of metrics (M^3, g_t), $0 \leq t < \infty$. This family of metrics will converge to a 3-dimensional Schwarzschild metric and will have other special properties which will allow us to prove the Riemannian Penrose Inequality for the original metric (M^3, g_0).

In particular, let Σ_0 be the outermost minimal surface of (M^3, g_0) with area A_0. Then we will also define a family of surfaces $\Sigma(t)$ with $\Sigma(0) = \Sigma_0$ such that $\Sigma(t)$ is minimal in (M^3, g_t). This is natural since as the metric g_t changes, we expect that the location of the horizon $\Sigma(t)$ will also change. Then the interesting quantities to keep track of in this flow are $A(t)$, the total area of the horizon $\Sigma(t)$ in (M^3, g_t), and $m(t)$, the total mass of (M^3, g_t) in the chosen end.

In addition to all of the metrics g_t having nonnegative scalar curvature, we will also have the very nice properties that

$$\begin{aligned} A'(t) &= 0, \\ m'(t) &\leq 0 \end{aligned}$$

for all $t \geq 0$. Then since (M^3, g_t) converges to a Schwarzschild metric (in an appropriate sense) which gives equality in the Riemannian Penrose Inequality as described in the introduction,

$$m(0) \geq m(\infty) = \sqrt{\frac{A(\infty)}{16\pi}} = \sqrt{\frac{A(0)}{16\pi}} \tag{3.1}$$

which proves the Riemannian Penrose Inequality for the original metric (M^3, g_0). The hard part, then, is to find a flow of metrics which preserves nonnegative scalar curvature and the area of the horizon, decreases total mass, and converges to a Schwarzschild metric as t goes to infinity. This proceeds as follows:

The metrics g_t will all be conformal to g_0. This conformal flow of metrics can be thought of as the solution to a first order o.d.e. in t defined by equations (3.2)-(3.5). Let

$$g_t = u_t(x)^4 g_0 \tag{3.2}$$

and $u_0(x) \equiv 1$. Given the metric g_t, define

$$\Sigma(t) = \text{the outermost minimal area enclosure of } \Sigma_0 \text{ in } (M^3, g_t) \tag{3.3}$$

where Σ_0 is the original outer minimising horizon in (M^3, g_0). In the cases in which we are interested, $\Sigma(t)$ will not touch Σ_0, from which it follows that $\Sigma(t)$ is actually a strictly outer minimising horizon of (M^3, g_t). Then given the horizon $\Sigma(t)$, define $v_t(x)$ such that

$$\begin{cases} \Delta_{g_0} v_t(x) & \equiv & 0 & \text{outside } \Sigma(t) \\ v_t(x) & = & 0 & \text{on } \Sigma(t) \\ \lim_{x\to\infty} v_t(x) & = & -e^{-t} & \end{cases} \tag{3.4}$$

and $v_t(x) \equiv 0$ inside $\Sigma(t)$. Finally, given $v_t(x)$, define

$$u_t(x) = 1 + \int_0^t v_s(x)ds \tag{3.5}$$

so that $u_t(x)$ is continuous in t and has $u_0(x) \equiv 1$.

Note that equation (3.5) implies that the first order rate of change of $u_t(x)$ is given by $v_t(x)$. Hence, the first order rate of change of g_t is a function of itself, g_0, and $v_t(x)$ which is a function of g_0, t, and $\Sigma(t)$ which is in turn a function of g_t and Σ_0. Thus, the first order rate of change of g_t is a function of t, g_t, g_0, and Σ_0. (All the results in this section are from [15].)

Theorem 3.1 *Taken together, equations (3.2)-(3.5) define a first order o.d.e. in t for $u_t(x)$ which has a solution which is Lipschitz in the t variable, C^1 in the x variable everywhere, and smooth in the x variable outside $\Sigma(t)$. Furthermore, $\Sigma(t)$ is a smooth, strictly outer minimising horizon in (M^3, g_t) for all $t \geq 0$, and $\Sigma(t_2)$ encloses but does not touch $\Sigma(t_1)$ for all $t_2 > t_1 \geq 0$.*

Since $v_t(x)$ is a superharmonic function in (M^3, g_0) (harmonic everywhere except on $\Sigma(t)$, where it is weakly superharmonic), it follows that $u_t(x)$ is superharmonic as well. Thus, from equation (3.5) we see that $\lim_{x\to\infty} u_t(x) = e^{-t}$ and consequently that $u_t(x) > 0$ for all t by the maximum principle. Then since

$$R(g_t) = u_t(x)^{-5}(-8\Delta_{g_0} + R(g_0))u_t(x) \tag{3.6}$$

it follows that (M^3, g_t) is an asymptotically flat manifold with nonnegative scalar curvature.

Even so, it still may not seem like g_t is particularly naturally defined since the rate of change of g_t appears to depend on t and the original metric g_0 in equation (3.4). We would prefer a flow where the rate of change of g_t can be defined purely as a function of g_t (and Σ_0 perhaps), and interestingly enough this actually does turn out to be the case. In [15] we prove this very important fact and define a new equivalence class of metrics called the harmonic conformal class. Then once we decide to find a flow of metrics which stays inside the harmonic conformal class of the original metric (outside the horizon) and keeps the area of the horizon $\Sigma(t)$ constant, then we are basically forced to choose the particular conformal flow of metrics defined above.

Theorem 3.2 *The function $A(t)$ is constant in t and $m(t)$ is non-increasing in t, for all $t \geq 0$.*

The fact that $A'(t) = 0$ follows from the fact that to first order the metric is not changing on $\Sigma(t)$ (since $v_t(x) = 0$ there) and from the fact that to first order the area of $\Sigma(t)$ does not change as it moves outward since $\Sigma(t)$ is a critical point for area in (M^3, g_t). Hence, the interesting part of theorem (3.2) is proving that $m'(t) \leq 0$. Curiously, this follows from a nice trick using the Riemannian positive mass theorem.

Another important aspect of this conformal flow of the metric is that outside the horizon $\Sigma(t)$, the manifold (M^3, g_t) becomes more and more spherically symmetric and "approaches" a Schwarzschild manifold $(\mathbb{R}^3\backslash\{0\}, s)$ in the limit as t goes to ∞. More precisely,

Theorem 3.3 *For sufficiently large t, there exists a diffeomorphism ϕ_t between (M^3, g_t) outside the horizon $\Sigma(t)$ and a fixed Schwarzschild manifold $(\mathbb{R}^3\backslash\{0\}, s)$ outside its horizon. Furthermore, for all $\epsilon > 0$, there exists a T such that for all $t > T$, the metrics g_t and $\phi_t^*(s)$ (when determining the lengths of unit vectors of (M^3, g_t)) are within ϵ of each other and the total masses of the two manifolds are within ϵ of each other. Hence,*

$$\lim_{t\to\infty} \frac{m(t)}{\sqrt{A(t)}} = \sqrt{\frac{1}{16\pi}}.$$

Theorem 3.3 is not that surprising really although a careful proof is reasonably long. However, if one is willing to believe that the flow of metrics converges to a spherically symmetric metric outside the horizon, then theorem 3.3 follows from two facts. The first fact is that the scalar curvature of (M^3, g_t) eventually becomes identically zero outside the horizon $\Sigma(t)$ (assuming (M^3, g_0) is harmonically flat). This follows from the facts that $\Sigma(t)$ encloses any compact set in a finite amount of time, that harmonically flat manifolds have zero scalar curvature outside a compact set, that $u_t(x)$ is harmonic outside $\Sigma(t)$, and equation (3.6). The second fact is that the Schwarzschild metrics are the only complete, spherically symmetric 3-manifolds with zero scalar curvature (except for the flat metric on R^3).

The Riemannian Penrose inequality, inequality (1.12), then follows from equation (3.1) using theorems 3.1, 3.2 and 3.3, for harmonically flat manifolds [15]. Since asymptotically flat manifolds can be approximated arbitrarily well by harmonically flat manifolds while changing the relevant quantities arbitrarily little, the asymptotically flat case also follows. Finally, the case of equality of the Penrose inequality follows from a more careful analysis of these same arguments.

We refer the reader to [16,17,20] for further review-type discussions of the results described above.

4 Open Questions and Applications

Now that the Riemannian Penrose conjecture has been proved, what are the next interesting directions? What applications can be found? Is this subject only of physical interest, or are there possibly broader applications to other problems in mathematics?

Clearly the most natural open problem is to find a way to prove the general Penrose conjecture (discussed in the next subsection) in which M^3 is allowed to have any second fundamental form in the space-time. There is good reason to think that this may follow from the Riemannian Penrose inequality, although this is a bit delicate. On the other hand, the general positive mass theorem followed from the Riemannian positive mass theorem as was originally shown by Schoen and Yau using an idea due to Jang [71, 73]. For physicists this problem is definitely a top priority since most space-times do not even admit zero second fundamental form space-like slices. We note that the Riemannian Penrose inequality does give a result which applies to situations more general than time symmetric, as the condition $R \geq 0$ holds, *e.g.*, for maximal initial data sets $\mathrm{tr}_g h = 0$, as well as in several other situations ("polar gauge", and so on). However, the general situation remains open.

Another interesting question is to ask these same questions in higher dimensions. One of us (HB) is currently working on a paper to prove the Riemannian Penrose inequality in dimensions less than 8. Dimension 8 and higher are harder because of the surprising fact that minimal hypersurfaces (and hence apparent horizons of black holes) can have codimension 7 singularities (points where the hypersurface is not smooth). This curious technicality is also the reason that the positive mass theorem in dimensions 8 and higher for manifolds which are not spin has only been announced very recently by Christ and Lohkamp [22], using a formidable singularity excision argument, and it is conceivable that this technique will allow one to extend the Riemannian Penrose Inequality proof to all dimensions.

Naturally it is harder to tell what the applications of these techniques might be to other problems, but already there have been some. One application is to the famous Yamabe problem: Given a compact 3-manifold M^3, define $E(g) = \int_{M^3} R_g dV_g$ where g is scaled so that the total volume of (M^3, g) is one, R_g is the scalar curvature at each point, and dV_g is the volume form. An idea due to Yamabe was to try to construct canonical metrics on M^3 by finding critical points of this energy functional on the space of metrics. Define $C(g)$ to be the infimum of $E(\bar{g})$ over all metrics $\bar{g}$ conformal to g. Then the (smooth) Yamabe invariant of M^3, denoted here as $Y(M^3)$, is defined to be the supremum of $C(g)$ over all metrics g. $Y(S^3) = 6 \cdot (2\pi^2)^{2/3} \equiv Y_1$ is known to be the largest possible value for Yamabe invariants of 3-manifolds. It is also known that $Y(T^3) = 0$ and $Y(S^2 \times S^1) = Y_1 = Y(S^2 \tilde{\times} S^1)$, where $S^2 \tilde{\times} S^1$ is the non-orientable S^2 bundle over S^1.

One of the authors (HB) and Andre Neves, working on a problem suggested by Richard Schoen, were able to compute the Yamabe invariant of RP^3 using inverse mean curvature flow techniques [19] (see also [13, Lecture 2]) and found that $Y(RP^3) = Y_1/2^{2/3} \equiv Y_2$. A corollary is $Y(RP^2 \times S^1) = Y_2$ as well. These techniques also yield the surprisingly strong result that the only prime 3-manifolds with Yamabe invariant larger than RP^3 are S^3, $S^2 \times S^1$, and $S^2 \tilde{\times} S^1$. The Poincaré conjecture for 3-manifolds with Yamabe invariant greater than RP^3 is therefore a corollary. Furthermore, the problem of classifying 3-manifolds is known to reduce to the problem of classifying prime 3-manifolds. The Yamabe approach then would be to make a list of prime 3-manifolds ordered by Y. The first five prime 3-manifolds on this list are therefore S^3, $S^2 \times S^1$, $S^2 \tilde{\times} S^1$, RP^3, and $RP^2 \times S^1$.

4.1 The Riemannian Penrose conjecture on asymptotically hyperbolic manifolds

Another natural class of metrics that are of interest in general relativity consists of metrics which asymptote to the hyperbolic metric. Such metrics arise when considering solutions with a negative cosmological constant, or when considering "hyperboloidal hypersurfaces" in space-times which are asymptotically flat in isotropic directions (technically speaking, these are spacelike hypersurfaces which intersect $\mathscr{I}$ transversally in the conformally completed space-time). For instance, recall that in the presence of a cosmological constant Λ the scalar constraint equation reads

$$R = 16\pi\mu + |h|_g^2 - (\mathrm{tr}\,_g h)^2 + 2\Lambda \ .$$

Suppose that $h = \lambda g$, where λ is a constant; such an h solves the vector constraint equation. We then have

$$R = 16\pi\mu - 6\lambda^2 + 2\Lambda =: 16\pi\mu + 2\Theta \ . \tag{4.1}$$

The constant Θ equals thus Λ when $\lambda = 0$, or $-3\lambda^2$ when $\Lambda = 0$. The positive energy condition $\mu \geq 0$ is now equivalent to

$$R \geq 2\Theta \ .$$

For $\lambda = 0$ the associated model space-time metrics take the form

$$ds^2 = -(k - \frac{2m}{r} - \frac{\Lambda}{3}r^2)dt^2 + (k - \frac{2m}{r} - \frac{\Lambda}{3}r^2)^{-1}dr^2 + r^2 d\Omega_k^2 \ , \quad k = 0, \pm 1 \ , \tag{4.2}$$

where $d\Omega_k^2$ denotes a metric of constant Gauss curvature k on a two dimensional compact manifold M^2. These are well known static solutions of the vacuum Einstein equation with a cosmological constant Λ; some subclasses of (4.2) have been discovered by de Sitter [74] ((4.2) with $m = 0$ and $k = 1$), by Kottler [62] (Equation (4.2) with an arbitrary m and $k = 1$). The parameter $m \in \mathbb{R}$ can be seen to be proportional to the total Hawking mass (*cf.* (4.5) below) of the foliation $t = \mathrm{const}$, $r = \mathrm{const}$. We will refer to those solutions as the generalized Kottler solutions. The constant Λ in (4.2) is an arbitrary real number, but in this section we will only consider $\Lambda < 0$.

From now on the overall approach resembles closely that for asymptotically flat space-times, as described earlier in this work. For instance, one considers manifolds which contain asymptotic ends diffeomorphic to $\mathbb{R}^+ \times M^2$. It is convenient to think of each of the sets "$\{r = \infty\} \times M^2$" as a connected component at infinity of a boundary at infinity, call it $\partial_\infty M^3$, of the initial data surface M^3. There is a well defined notion of mass for metrics which asymptote to the above model metrics in the asymptotic ends, somewhat similar to that in (1.5). In the hyperbolic case the boundary conditions are considerably more delicate to formulate as compared to the asymptotically flat one, and we refer the reader to [30–32, 79] for details. In the case when M^3 arises from a space-times with negative cosmological constant Λ, the resulting mass is usually called the Abbott-Deser mass [1]; when $\Lambda = 0$ and M^3 is a hyperboloidal hypersurface the associated mass is called the Trautman-Bondi mass. (The latter notion of mass has often been referred to as "Bondi mass" in the literature, but the name "Trautman-Bondi mass" seems more appropriate, in view of the work in [77], which precedes [12] by four years; see also [76].) A large class of initial data sets with the desired asymptotic behavior has been constructed in [2, 3, 61], and the existence of the associated space-times has been established in [40, 41].

The monotonicity argument of Geroch [43], described in Section 1.7, has been extended by Gibbons [45] to accommodate for the negative cosmological constant; we follow the presentation in [33]: We assume that we are given a three dimensional manifold (M^3, g) with connected minimal boundary ∂M^3 such that

$$R \geq 2\Theta \ ,$$

for some strictly negative constant Θ (compare (4.1)). We further assume that there exists a smooth, global solution of the inverse mean curvature flow without critical points, with u ranging from zero to infinity, vanishing on ∂M^3, with the level sets of u

$$\Sigma(s) = \{u(x) = s\}$$

being compact. Let A_s denote the area of $\Sigma(s)$, and define

$$\sigma(s) = \sqrt{A_s}\int_{\Sigma(s)} ({}^2\mathcal{R}_s - \frac{1}{2}H_s^2 - \frac{2}{3}\Theta)d^2\mu_s \;, \tag{4.3}$$

where ${}^2\mathcal{R}_s$ is the scalar curvature (equal twice the Gauss curvature) of the metric induced on $\Sigma(s)$, $d^2\mu_s$ is the Riemannian volume element associated to that same metric, and H_s is the mean curvature of $\Sigma(s)$. The hypothesis that du is nowhere vanishing implies that all the objects involved are smooth in s. At $s = 0$ we have $H_0 = 0$ and $A_0 = A_{\partial M^3}$ so that

$$\begin{aligned} \sigma(0) &= \sqrt{A_{\partial M^3}}\int_{\partial M^3} ({}^2\mathcal{R}_0 - \frac{2}{3}\Theta)d^2\mu_0 \\ &= \sqrt{A_{\partial M^3}}\left(8\pi(1-g_{\partial M^3}) - \frac{2}{3}\Theta A_{\partial M^3}\right) \;. \end{aligned} \tag{4.4}$$

Generalising a formula of Hawking [47], Gibbons [45, Equation (17)] assigns to the $\Sigma(s)$ foliation a *total mass* M_{Haw} via the formula

$$M_{Haw} \equiv \lim\nolimits_{\epsilon\to 0}\frac{\sqrt{A_{1/\epsilon}}}{32\pi^{3/2}}\int_{\{u=1/\epsilon\}} ({}^2\mathcal{R}_s - \frac{1}{2}H_s^2 - \frac{2}{3}\Theta)d^2\mu_s \;, \tag{4.5}$$

where A_α is the area of the connected component under consideration of the level set $\{u = \alpha\}$. It follows that

$$\lim_{s\to\infty}\sigma(s) = 32\pi^{3/2}M_{Haw} \;,$$

assuming the limit exists. The generalisation in [45] of (1.13) establishes the inequality

$$\frac{\partial\sigma}{\partial s} \geq 0 \;. \tag{4.6}$$

This implies $\lim_{s\to\infty}\sigma(s) \geq \sigma(0)$, which gives

$$2M_{Haw} \geq (1-g_{\partial M^3})\left(\frac{A_{\partial M^3}}{4\pi}\right)^{1/2} - \frac{\Theta}{3}\left(\frac{A_{\partial M^3}}{4\pi}\right)^{3/2} \;. \tag{4.7}$$

Here $A_{\partial M^3}$ is the area of ∂M^3 and $g_{\partial M^3}$ is the genus thereof. Equation (4.7) is sharp — the inequality there becomes an equality for the generalized Kottler metrics (4.2).

The hypothesis above that du has no critical points together with our hypothesis on the geometry of the asymptotic ends forces ∂M^3 to be connected. It is not entirely clear what is the right generalisation of this inequality to the case where several black holes occur, with one possibility being

$$2M_{Haw} \geq \sum_{i=1}^{k}\left((1-g_{\partial_i M^3})\left(\frac{A_{\partial_i M^3}}{4\pi}\right)^{1/2} - \frac{\Theta}{3}\left(\frac{A_{\partial_i M^3}}{4\pi}\right)^{3/2}\right) \;. \tag{4.8}$$

Here the $\partial_i M^3$'s, $i = 1, \ldots, k$, are the connected components of ∂M^3, $A_{\partial_i M^3}$ is the area of $\partial_i M^3$, and $g_{\partial_i M^3}$ is the genus thereof. This would be the inequality one would obtain from the Geroch–Gibbons argument if it could be carried through for u's which are allowed to have critical points, on manifolds with $\partial_\infty M^3$ connected but ∂M^3 — not connected.

As in the asymptotically flat case, the naive monotonicity calculation of [43] breaks down at critical level sets of u, as those do not have to be smooth submanifolds. Nevertheless the existence of the appropriate function u (perhaps with critical points) should probably follow from the results in [55,56]. The open questions here are 1) a proof of monotonicity at jumps of the flow, where topology change might occur, and 2) the proof that the Hawking mass (4.5) exists, and equals the mass of the end under consideration. We also note that in the hyperbolic context it is natural to consider not only boundaries ∂M^3 which are minimal, but also boundaries satisfying

$$H = \pm 2 \ .$$

This is related to the discussion at the beginning of this section: if $\lambda = 0$, then an apparent horizon corresponds to $H = 0$; if $\Lambda = 0$ and $\lambda = -1$, then a future apparent horizon corresponds to $H = 2$, while a past apparent horizon corresponds to $H = -2$.

Let us discuss some of the consequences of the (hypothetical) inequality (4.8). In the current setting there are some genus-related ambiguities in the definition of mass (see [33] for a detailed discussion of various notions of mass for static asymptotically hyperbolic metrics), and it is convenient to introduce a mass parameter m defined as follows

$$m = \begin{cases} M_{Haw} \ , & \partial_\infty M^3 = S^2 \ , \\ M_{Haw} \ , & \partial_\infty M^3 = T^2, \text{ with the normalization } A'_\infty = -12\pi/\Theta \ , \\ \dfrac{M_{Haw}}{|g_{\partial_\infty M^3} - 1|^{3/2}} \ , & g_{\partial_\infty M^3} > 1 \ . \end{cases} \tag{4.9}$$

Here A'_∞ is the area of $\partial_\infty M^3$ in the metric $d\Omega^2_k$ appearing in (4.2). For generalized Kottler metrics the mass m so defined coincides with the mass parameter appearing in (4.2) when u is the "radial" solution $u = u(r)$ of the inverse mean curvature flow.

Note, first, that if all connected components of the horizon have spherical or toroidal topology, then the lower bound (4.8) is strictly positive. For example, if $\partial M^3 = T^2$, and $\partial_\infty M^3 = T^2$ as well we obtain

$$2m \geq -\frac{\Lambda}{3}\left(\frac{A_{\partial M^3}}{4\pi}\right)^{3/2} \ .$$

On the other hand if $\partial M^3 = T^2$ but $g_{\partial_\infty M^3} > 1$ from Equation (4.8) one obtains

$$2m \geq -\frac{\Lambda}{3|g_\infty - 1|}\left(\frac{A_{\partial M^3}}{4\pi}\right)^{3/2} \ .$$

Recall that in a large class of space-times[4] the Galloway–Schleich–Witt–Woolgar inequality [42] holds:

$$\sum_{i=1}^{k} g_{\partial_i M^3} \leq g_\infty \ . \tag{4.10}$$

[4]The discussion that follows applies to all (M^3, g, h)'s that can be isometrically embedded into a globally hyperbolic space-time $\mathscr{M}$ (with timelike conformal boundary at infinity) in which the null convergence condition holds; further the closure of the image of M^3 should be a partial Cauchy surface in $\mathscr{M}$. Finally the intersection of the closure of M^3 with $\mathscr{I}$ should be compact. The global hyperbolicity here, and the notion of Cauchy surfaces, is understood in the sense of manifolds with boundary, see [42] for details.

It implies that if $\partial_\infty M^3$ has spherical topology, then all connected components of the horizon must be spheres. Similarly, if $\partial_\infty M^3$ is a torus, then all components of the horizon are spheres, except perhaps for at most one which could be a torus. It follows that to have a component of the horizon which has genus higher than one we need $g_\infty > 1$ as well.

When some — or all — connected components of the horizon have genus higher than one, the right hand side of Equation (4.8) might become negative. Minimising the generalised Penrose inequality (4.8) with respect to the areas of the horizons gives the following interesting inequality

$$M_{Haw} \geq -\frac{1}{3\sqrt{-\Lambda}} \sum_i |g_{\partial_i M^3} - 1|^{3/2} , \tag{4.11}$$

where the sum is over those connected components $\partial_i M^3$ of ∂M^3 for which $g_{\partial_i M^3} \geq 1$. Equation (4.11), together with the elementary inequality $\sum_{i=1}^N |\lambda_i|^{3/2} \leq \left(\sum_{i=1}^N |\lambda_i|\right)^{3/2}$, lead to

$$m \geq -\frac{1}{3\sqrt{-\Lambda}} . \tag{4.12}$$

Similarly to the asymptotically flat case, the Geroch–Gibbons argument establishing the inequality (4.4) when a suitable u exists can also be carried through when $\partial M^3 = \emptyset$. In this case one still considers solutions u of the differential equation (2.1) associated with the inverse mean curvature flow, however the Dirichlet condition on u at ∂M^3 is replaced by a condition on the behavior of u near some chosen point $p_0 \in M^3$. If the level set of u around p_0 approach distance spheres centred at p_0 at a suitable rate, then $\sigma(s)$ tends to zero when the $\Sigma(s)$'s shrink to p_0, which together with the monotonicity of σ leads to the positive energy inequality:

$$M_{Haw} \geq 0 . \tag{4.13}$$

It should be emphasised that the Horowitz-Myers solutions [54] with negative mass show that this argument breaks down when $g_\infty = 1$.

When $\partial_\infty M^3 = S^2$ the inequality (4.13), with M_{Haw} replaced by the Hamiltonian mass (which might perhaps coincide with M_{Haw}, but this remains to be established), can be proved by Witten type techniques [30,31] (compare [4,46,79,81]). On the other hand it follows from [10] that when $\partial_\infty M^3 \neq S^2$ there exist no asymptotically covariantly constant spinors which can be used in the Witten argument. The Geroch–Gibbons argument has a lot of "ifs" attached in this context, in particular if $\partial_\infty M^3 \neq S^2$ then some level sets of u are necessarily critical and it is not clear what happens with σ at jumps of topology. We note that the area of the horizons does not occur in (4.12) which, when $g_{\partial_\infty M^3} > 1$, suggests that the correct inequality is actually (4.12) rather than (4.13), whether or not black holes are present.

We close this section by mentioning an application of the hyperbolic Penrose inequality to the uniqueness of static regular black holes with a negative cosmological constant, pointed out in [33]. It is proved in that last reference that for such connected black holes an inequality *inverse* to (4.7) holds, with equality if and only if the metric is the one in (4.2). Hence a proof of the Penrose inequality would imply equality in (4.7), and subsequently a uniqueness theorem for such black holes.

4.2 Precise Formulations of the (full) Penrose Conjecture

In the next two subsections we discuss formulations of the Penrose Conjecture and possible applications of these statements to defining quasi-local mass functionals with good properties and to defining total mass in surprisingly large generality.

This discussion is based on the third lecture [13] given by one of us (HB) in Cargèse in the summer of 2002. Besides discussing various formulations of the conjecture in this subsection, we point out the value of its possible applications in the next subsection, which greatly motivates trying to prove the conjecture.

We begin with the question, "Given Cauchy data, where is the event horizon, and what lower bounds on its area can we make?" Inequality (1.4) is the most general version of the Penrose conjecture, but there are more "local" versions of it which have the advantage of possibly being easier to prove. Recall, for instance, that the exact location of event horizons can not be determined from the Cauchy data (M^3, g, h) without solving the Einstein equations infinitely forward in time. On the other hand, apparent horizons Σ can be computed directly from the Cauchy data and are characterised by the equation

$$H_\Sigma = \mathrm{tr}_\Sigma(h), \tag{4.14}$$

that is, the mean curvature H of Σ equals the trace of h along Σ. Note that in the $h = 0$ case, this is the assumption that $H = 0$, which is the Euler-Lagrange equation of a surface which locally minimises area. This leads to the first formulation of the Penrose Conjecture, which seems to be due to Gary Horowitz [53]:

Conjecture 4.1 *Let (M^3, g, h) be complete, asymptotically flat Cauchy data with $\mu \geq |J|$ and an apparent horizon satisfying equation (4.14). Then*

$$m \geq \sqrt{A/16\pi}, \tag{4.15}$$

where m is the total mass and A is the minimum area required for a surface to enclose Σ.

The logic is that since apparent horizons imply the existence of an event horizon outside of it, and all surfaces enclosing Σ have at least area A, then inequality (1.4) implies the above conjecture.

An alternative possibility would be to replace A_e in (1.4) by the area of the apparent horizon. We do not know the answer to this, but a counterexample would not be terribly surprising (although it would be very interesting). The point is that the physical reasoning used by Penrose does not directly imply that such a conjecture should be true for apparent horizons. Hence, a counterexample to the area of the apparent horizon conjecture would be less interesting than a counterexample to Conjecture 4.1 or 4.2 since one of the latter counterexamples would imply that there was actually something wrong with Penrose's physical argument, which would be very important to understand.

There are also good reasons to consider a second formulation of the Penrose Conjecture, due to one of the authors (HB), for (M^3, g, h) which have more than one end. We will choose one end to be special, and then note that large spheres S in the other asymptotically flat ends are actually "trapped," meaning that $H_S < \mathrm{tr}_S(h)$ (note that the mean curvatures of these spheres is actually negative when the outward direction is taken to be toward the special end and away from the other ends). We can conclude that these large spheres are trapped, if, for example, the mean curvatures of these large spheres is $-2/r$ to highest order and $|h|$ is decreasing like $1/r^2$ (or at least faster than $1/r$). Hence, this condition also allows us to conclude that there must be an event horizon enclosing all of the other ends. Thus, we conjecture

Conjecture 4.2 *Let (M^3, g, h) be complete, asymptotically flat Cauchy data with $\mu \geq |J|$ and more than one end. Choose one end to be special, and then define A to be the minimum area required to enclose all of the other ends. Then*

$$m \geq \sqrt{A/16\pi}, \tag{4.16}$$

where m is the total mass of the chosen end.

We note that, more precisely, A in the above conjecture is the infimum of the boundary area of all smooth, open regions which contain all of the other ends (but not the special end). Equivalently (taking the complement), A is the infimum of the boundary area of all smooth, open regions which contain the special end (but none of the other ends). A smooth, compact, area-minimising surface (possibly with multiple connected components) always exists and has zero mean curvature.

The advantage of this second formulation is that it removes equation (4.14) and the need to define apparent horizons. Also, preliminary thoughts by one of the authors (HB) lead him to believe that the above two formulations are equivalent via a reflection argument (although this still requires more consideration). In addition, this second formulation turns out to be most useful in the quasi-local mass and total mass definitions in the next subsection.

4.3 Applications to Quasi-local Mass and Total Mass

The ideas of this subsection are due to HB, and were greatly influenced by and in some cases are simply natural extensions of ideas due to Bartnik in [7,8]. All of the surfaces we are considering in this subsection are required to be boundaries of regions which contain all of the other ends besides the chosen one. Given such a surface Σ in a (M^3, g, h) containing at least one asymptotically flat end, let I be the inside region (containing all of the other ends) and O be the outside region (containing a chosen end). Then we may consider "extensions" of (M^3, g, h) to be manifolds which result from replacing the *outside* region O in M with any other manifold and Cauchy data such that the resulting Cauchy data $(\tilde{M}^3, g, h)$ is smooth, asymptotically flat, and has $\mu \geq |J|$ everywhere (including along the surgery naturally). We define a "fill-in" of (M^3, g, h) to be manifolds which result from replacing the *inside* region I in M with any other manifold and Cauchy data such that the resulting Cauchy data $(\tilde{M}^3, g, h)$ is smooth, asymptotically flat, and has $\mu \geq |J|$ everywhere. Also, we say that a surface is "outer-minimising" if any other surface which encloses it has at least as much area. Note that for "enclose" to make sense, we need to restrict our attention to surfaces which are the boundaries of regions as stated at the beginning of this paragraph. The notion of "outer-minimising" surfaces turns out to be central to the following definitions.

Suppose Σ is outer-minimising in (M^3, g, h). Define the *Bartnik outer mass* $m_{outer}(\Sigma)$ to be the infimum of the total mass over all extensions of (M^3, g, h) in which Σ remains outer-minimising. Hence, what we are doing is fixing (M^3, g, h) inside Σ and then seeing how small we can make the total mass outside of Σ without violating $\mu \geq |J|$. Intuitively, whatever the total mass of this minimal mass extension outside Σ is can be interpreted as an upper bound for the mass contributed by the energy and momentum inside Σ.

The definition begs the question, why do we only consider extensions which keep Σ outer-minimising? After all, we are attempting to find an extension with minimal mass, and one might naively think that the minimal mass extensions would naturally have this property anyway, and locally the minimal mass extensions we defined above probably usually do (if they exist). However, given any Σ, it is always possible to choose an extension which shrinks to a small neck outside Σ and then flattens out to an arbitrarily small mass Schwarzschild metric outside the small neck. Hence, without some restriction to rule out extensions with small necks, the infimum would always be zero. Bartnik's original solution to this problem was to not allow apparent horizons outside of Σ, and this works quite nicely. For technical reasons, however, we have chosen to preserve the "outer-minimising" condition on Σ, which allows us to prove that $m_{outer}(\Sigma) \geq m_{inner}(\Sigma)$, defined in a moment.

(Thus, the definition given here is not identical to that in [7], and we do not know whether or not it gives the same number as Bartnik's original definition, although this is a reasonable conjecture under many circumstances. We also note that the work of Huisken and Ilmanen [57] described above shows that, in the $h = 0$ case, the Hawking mass of Σ is a lower bound for the total mass if Σ is outer-minimising. They also show that $m_{outer}(\Sigma) = m$ in the case that Σ is entirely outside the black hole of a time-symmetric slice of the Schwarzschild metric of total mass m. These results support considering the outer-minimising condition in the current context.)

Suppose again that Σ is outer-minimising in (M^3, g, h). Define the *inner mass* $m_{inner}(\Sigma)$ to be the supremum of $\sqrt{A/16\pi}$ over all fill-ins of (M^3, g, h), where A is the minimum area needed to enclose all of the other ends of the fill-in besides the chosen end. Hence, what we are doing is fixing (M^3, g, h) outside Σ (so that Σ automatically remains outer-minimising) and then seeing how large we can make the area of the global area-minimising surface (which encloses all of the other ends other than the chosen one). Intuitively, we are trying to fill-in Σ with the largest possible black hole, since the event horizon of the black hole will have to be at least A. If we think of $\sqrt{A/16\pi}$ as the mass of the black hole, then the inner mass gives a reasonable lower bound for the mass of Σ (since there is a fill-in in which it contains a black hole of that mass).

Theorem 4.3 *Suppose (M^3, g, h) is complete, asymptotically flat, and has $\mu \geq |J|$. Then Conjecture 4.2 implies that*

$$m_{outer}(\Sigma) \geq m_{inner}(\Sigma) \tag{4.17}$$

for all Σ which are outer-minimising.

Sketch of proof: Consider any extension on the outside of Σ (which keeps Σ outer-minimising) and any fill-in on the inside of Σ simultaneously and call the resulting manifold $\bar{M}$. Since Σ is outer-minimising, there exists a globally area-minimising surface of $\bar{M}$ which is enclosed by Σ (since going outside of Σ never decreases area). Thus, by Conjecture 4.2,

$$m \geq \sqrt{A/16\pi}, \tag{4.18}$$

for $\bar{M}$. Taking the infimum on the left side and the supremum on the right side of this inequality then proves the theorem since the total mass m is determined entirely by the extension and the global minimum area A is determined entirely by the fill-in. □

Theorem 4.4 *Suppose (M^3, g, h) is complete, asymptotically flat, and has $\mu \geq |J|$. If Σ_2 encloses Σ_1 and both surfaces are outer-minimising, then*

$$m_{inner}(\Sigma_2) \geq m_{inner}(\Sigma_1) \tag{4.19}$$

and

$$m_{outer}(\Sigma_2) \geq m_{outer}(\Sigma_1) \tag{4.20}$$

Sketch of proof: The first inequality is straightforward since every fill-in inside Σ_1 is also a fill-in inside Σ_2. The second inequality is almost as straightforward. It is true that any extension of Σ_2 (in which Σ_2 is still outer-minimising) is also an extension of Σ_1, but it remains to be shown that such an extension preserves the outer-minimising property of Σ_1. However, this fact follows from the fact that any surface enclosing Σ_1 which goes outside of Σ_2 can be made to have less or equal area by being entirely inside Σ_2 (by the outer-minimising property of Σ_2). But since Σ_1 was outer-minimising in the original manifold, any surface between Σ_1 and Σ_2 must have at least as much area as Σ_1. □

The last three theorems inspire the definition of the *quasi-local mass* of a surface Σ in (M^3, g, h) to be the interval

$$m(\Sigma) \equiv [m_{inner}(\Sigma), m_{outer}(\Sigma)] \subset R. \tag{4.21}$$

That is, we are not defining the quasi-local mass of a surface to be a number, but instead to be an interval in the real number line. Both endpoints of this interval are increasing when we move outward to surfaces which enclose the original surface. If $\Sigma \subset (M^3, g, h)$ and (M^3, g, h) is Schwarzschild data, then this interval collapses to a point and equals the mass of the Schwarzschild data (assuming Conjecture 4.2). Conversely, if the quasilocal mass interval of Σ is a point, then we expect that Σ can be imbedded into a Schwarzschild spacetime in such a way that its Bartnik data (the metric, mean curvature vector in the normal bundle, and the connection on the normal bundle of Σ) is preserved, which is a nongeneric condition. Hence, we typically expect the quasi-local mass of a surface to be an interval of positive length. We also expect the quasi-local mass interval to be very close to a point in a "quasi-Newtonian" situation, where Σ is in the part of the space-time which is a perturbation of Minkowski space, for example. We point out that so far there are not any surfaces for which we can prove that the quasi-local mass is not a point. This is because there are very few instances in which the inner and outer masses of a surface can be computed at all. These questions will have to wait until a better understanding of the Penrose conjecture is found. This definition of quasi-local mass leads naturally to definitions of *total inner mass*, m_{inner}^{total}, and *total outer mass*, m_{outer}^{total}, where in both cases we simply take the supremum of inner mass and outer mass respectively over all Σ which are outer-minimising.

Conjecture 4.5 *If (M^3, g, h) is asymptotically flat with total mass m_{ADM}, then*

$$m_{inner}^{total} = m_{outer}^{total} = m_{ADM}. \tag{4.22}$$

Consider (M^3, g, h) which is not assumed to have any asymptotics but still satisfies $\mu \geq |J|$. Then we will say that Σ (again, always assumed to be the boundary of a region in M^3) is "legal" if Σ is outer-minimising in (M^3, g, h) and there exists an asymptotically flat extension with $\mu \geq |J|$ outside of Σ in which Σ remains outer-minimising. Note that (M^3, g, h) is not assumed to have any asymptotics. We are simply defining the surfaces for which extensions with good asymptotics exist, and giving these surfaces the name "legal." Note also that both $m_{inner}(\Sigma)$ and $m_{outer}(\Sigma)$ are well-defined for legal Σ. Thus, total inner mass and total outer mass are well-defined as long as (M^3, g, h) has at least one legal Σ. Finally, theorems 4.3 and 4.4 are still true for legal surfaces even when (M^3, g, h) is not assumed to be asymptotically flat.

During the "50 Years" conference in Cargèse, Mark Aarons asked the question, "When are the total inner mass and the total outer mass different?" This is a very hard question, but it is such a good one that it deserves some speculation. If we define

$$m_{total} \equiv [m_{inner}^{total}, m_{outer}^{total}], \tag{4.23}$$

then total mass is very often well-defined (as long as there is at least one legal Σ), but is not necessarily a single value. Mark's question is then equivalent to, "When is the total mass single-valued and therefore well-defined as a real number?"

According to conjecture 4.5, we are not going to find an example of total inner mass $m_{inner}^{total} \neq m_{outer}^{total}$, the total outer mass, in the class of asymptotically flat manifolds. In fact, we are not aware of any examples of $m_{inner}^{total} \neq m_{outer}^{total}$, though one can give arguments to the effect that such situations could occur. On the other hand we believe that in reasonable situations this will not happen:

Conjecture 4.6 *Suppose (M^3, g, h) has $\mu \geq |J|$ and that there exists a nested sequence of connected, legal surfaces $\Sigma_i = \partial D_i \subset M$, $1 \leq i < \infty$, with $\bigcup_i D_i = M$ and $\lim |\Sigma_i| = \infty$. Then*

$$m_{inner}^{total} = m_{outer}^{total} \in \mathbb{R} \cup \{\infty\}. \tag{4.24}$$

At first this conjecture seems wildly optimistic considering it is suggesting that total mass is well-defined in the extended real numbers practically all of the time, where the only assumptions we are making are along the lines of saying that the noncompact end must be "large" in some sense. Note, for example, we are ruling out cylindrical ends and certain types of cusp ends. However, the idea here is that most kinds of "crazy asymptotics" cause both m_{inner}^{total} and m_{outer}^{total} to diverge to infinity. Hence, the reason this conjecture (or one similar to it) has a decent chance of being true is the possibility that either m_{inner}^{total} or m_{outer}^{total} being finite is actually a very restrictive situation. For example, if either the total inner mass or the total outer mass is finite, then it might be true that this implies that (M^3, g, h) is asymptotic to data coming from a space-like slice of a Schwarzschild space-time (in some sense). In this case, one would expect that both the total inner and outer masses actually equal the mass of the Schwarzschild space-time and therefore are equal to each other. Certainly in the case that (M^3, g, h) is precisely a slice (even a very weird slice) of a Schwarzschild space-time, it is only natural to point out that total mass should be well-defined. These definitions seem to be an approach to defining total mass in these more general settings. However, a complete understanding of these definitions clearly depends on making further progress studying the Penrose Conjecture.

References

[1] L.F. Abbott and S. Deser, *Stability of gravity with a cosmological constant*, Nucl. Phys. **B195** (1982), 76–96.

[2] L. Andersson and P.T. Chruściel, *On asymptotic behavior of solutions of the constraint equations in general relativity with "hyperboloidal boundary conditions"*, Dissert. Math. **355** (1996), 1–100.

[3] L. Andersson, P.T. Chruściel, and H. Friedrich, *On the regularity of solutions to the Yamabe equation and the existence of smooth hyperboloidal initial data for Einsteins field equations*, Commun. Math. Phys. **149** (1992), 587–612.

[4] L. Andersson and M. Dahl, *Scalar curvature rigidity for asymptotically locally hyperbolic manifolds*, Annals of Global Anal. and Geom. **16** (1998), 1–27, dg-ga/9707017.

[5] R. Arnowitt, S. Deser, and C.W. Misner, *The dynamics of general relativity*, Gravitation (L. Witten, ed.), Wiley, N.Y., 1962, pp. 227–265.

[6] R. Bartnik, *The mass of an asymptotically flat manifold*, Comm. Pure Appl. Math. **39** (1986), 661–693.

[7] ———, *New definition of quasilocal mass*, Phys. Rev. Lett. **62** (1989), 2346–2348.

[8] R. Bartnik, *Energy in general relativity*, Tsing Hua Lectures on Geometry and Analysis (S.T. Yau, ed.), International Press, 1997, `http://www.ise.canberra.edu.au/mathstat/StaffPages/Robert2.htm`.

[9] R. Bartnik and P.T. Chruściel, *Boundary value problems for Dirac-type equations*, (2003), math.DG/0307278.

[10] H. Baum, *Complete Riemannian manifolds with imaginary Killing spinors*, Ann. Global Anal. Geom. **7** (1989), 205–226.

[11] G. Bergqvist, *On the Penrose inequality and the role of auxiliary spinor fields*, Class. Quantum Grav. **14** (1997), 2577–2583.

[12] H. Bondi, M.G.J. van der Burg, and A.W.K. Metzner, *Gravitational waves in general relativity VII: Waves from axi–symmetric isolated systems*, Proc. Roy. Soc. London A **269** (1962), 21–52.

[13] H.L. Bray, *Global inequalities*, lectures given at the Cargèse Summer School on 50 years of the Cauchy problem in general relativity, August 2002, online at `fanfreluche.math.univ-tours.fr`.

[14] ______, *The Penrose inequality in general relativity and volume comparison theorem involving scalar curvature*, Ph.D. thesis, Stanford University, 1997.

[15] ______, *Proof of the Riemannian Penrose conjecture using the positive mass theorem*, Jour. Diff. Geom. **59** (2001), 177–267, math.DG/9911173.

[16] ______, *Black holes and the Penrose inequality in general relativity*, Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002) (Beijing), Higher Ed. Press, 2002, pp. 257–271.

[17] ______, *Black Holes, Geometric Flows, and the Penrose Inequality in General Relativity*, Notices of the AMS **49** (2002), 1372–1381.

[18] H.L. Bray and K. Iga, *Superharmonic functions in* $\mathbf{R}^n$ *and the Penrose inequality in general relativity*, Comm. Anal. Geom. **10** (2002), 999–1016.

[19] H.L. Bray and A. Neves, *Classification of prime 3-manifolds with Yamabe invariant greater than* RP^3, Annals of Math. (2003), in press.

[20] H.L. Bray and R.M. Schoen, *Recent proofs of the Riemannian Penrose conjecture*, Current developments in mathematics, 1999 (Cambridge, MA), Int. Press, Somerville, MA, 1999, pp. 1–36.

[21] Y. Choquet-Bruhat, *Positive-energy theorems*, Relativity, groups and topology, II (Les Houches, 1983) (B.S. deWitt and R. Stora, eds.), North-Holland, Amsterdam, 1984, pp. 739–785.

[22] U. Christ and J. Lohkamp, in preparation (2003).

[23] D. Christodoulou and S.-T. Yau, *Some remarks on the quasi-local mass*, Cont. Math. **71** (1988), 9–14.

[24] P.T. Chruściel, *Black holes*, Proceedings of the Tübingen Workshop on the Conformal Structure of Space-times, H. Friedrich and J. Frauendiener, Eds., Springer Lecture Notes in Physics **604**, 61–102 (2002), gr-qc/0201053.

[25] ______, *A remark on the positive energy theorem*, Class. Quantum Grav. **33** (1986), L115–L121.

[26] ______, *Boundary conditions at spatial infinity from a Hamiltonian point of view*, Topological Properties and Global Structure of Space–Time (P. Bergmann and V. de Sabbata, eds.), Plenum Press, New York, 1986, pp. 49–59, URL `http://www.phys.univ-tours.fr/~piotr/scans`.

[27] ______, *On the invariant mass conjecture in general relativity*, Commun. Math. Phys. **120** (1988), 233–248.

[28] P.T. Chruściel and E. Delay, *On mapping properties of the general relativistic constraints operator in weighted function spaces, with applications*, Mém. Soc. Math. de France. **94** (2003), 1–103, gr-qc/0301073v2.

[29] P.T. Chruściel, E. Delay, G. Galloway, and R. Howard, *Regularity of horizons and the area theorem*, Annales Henri Poincaré **2** (2001), 109–178, gr-qc/0001003.

[30] P.T. Chruściel and M. Herzlich, *The mass of asymptotically hyperbolic Riemannian manifolds*, Pacific Jour. Math. **212** (2003), 231–264, dg-ga/0110035.

[31] P.T. Chruściel, J. Jezierski, and S. Łęski, *The Trautman-Bondi mass of hyperboloidal initial data sets*, (2003), gr-qc/0307109.

[32] P.T. Chruściel and G. Nagy, *The mass of spacelike hypersurfaces in asymptotically anti-de Sitter space-times*, Adv. Theor. Math. Phys. **5** (2002), 697–754, gr-qc/0110014.

[33] P.T. Chruściel and W. Simon, *Towards the classification of static vacuum spacetimes with negative cosmological constant*, Jour. Math. Phys. **42** (2001), 1779–1817, gr-qc/0004032.

[34] J. Corvino, *Scalar curvature deformation and a gluing construction for the Einstein constraint equations*, Commun. Math. Phys. **214** (2000), 137–189.

[35] J. Corvino and R. Schoen, *On the asymptotics for the vacuum Einstein constraint equations*, gr-qc/0301071, 2003.

[36] S. Dain, *Trapped surfaces as boundaries for the constraint equations*, (2003), gr-qc/0308009.

[37] V. I. Denisov and V. O. Solov'ev, *The energy determined in general relativity on the basis of the traditional Hamiltonian approach does not have physical meaning*, Theor. and Math. Phys. **56** (1983), 832–838, English translation, original pagination 301–314.

[38] P.A.M. Dirac, *The theory of gravitation in Hamiltonian form*, Proc. Roy. Soc. London **A246** (1958), 333–343.

[39] J. Frauendiener, *On the Penrose inequality*, Phys. Rev. Lett. **87** (2001), 101101, gr-qc/0105093.

[40] H. Friedrich, *Cauchy problem for the conformal vacuum field equations in general relativity*, Commun. Math. Phys. **91** (1983), 445–472.

[41] ______, *Einstein equations and conformal structure: Existence of anti-de-Sitter-type space-times*, Jour. Geom. and Phys. **17** (1995), 125–184.

[42] G.J. Galloway, K. Schleich, D.M. Witt, and E. Woolgar, *Topological censorship and higher genus black holes*, Phys. Rev. **D60** (1999), 104039, gr-qc/9902061.

[43] R. Geroch, *Energy extraction*, Ann. New York Acad. Sci. **224** (1973), 108–117.

[44] G.W. Gibbons, *Collapsing shells and the isoperimetric inequality for black holes*, Class. Quantum Grav. **14** (1997), 2905–2915, hep-th/9701049.

[45] ______, *Gravitational entropy and the inverse mean curvature flow*, Class. Quantum Grav. **16** (1999), 1677–1687.

[46] G.W. Gibbons, S.W. Hawking, G.T. Horowitz, and M.J. Perry, *Positive mass theorem for black holes*, Commun. Math. Phys. **88** (1983), 295–308.

[47] S.W. Hawking, *Gravitational radiation in an expanding universe*, Jour. Math. Phys. **9** (1968), 598–604.

[48] S.W. Hawking and G.F.R. Ellis, *The large scale structure of space-time*, Cambridge University Press, Cambridge, 1973.

[49] S.A. Hayward, *Quasilocalization of Bondi-Sachs energy loss*, Class. Quantum Grav. **11** (1994), 3037–3048, gr-qc/9405071.

[50] M. Herzlich, *A Penrose-like inequality for the mass of Riemannian asymptotically flat manifolds*, Commun. Math. Phys. **188** (1997), 121–133.

[51] ———, *The positive mass theorem for black holes revisited*, Jour. Geom. Phys. **26** (1998), 97–111.

[52] ———, *Minimal spheres, the Dirac operator and the Penrose inequality*, Séminaire Théorie Spectrale et Géométrie (Institut Fourier, Grenoble) **20** (2002), 9–16.

[53] G.T. Horowitz, *The positive energy theorem and its extensions*, Asymptotic behavior of mass and spacetime geometry (F. Flaherty, ed.), Springer Lecture Notes in Physics, vol. 202, Springer Verlag, New York, 1984.

[54] G.T. Horowitz and R.C. Myers, *The AdS/CFT correspondence and a new positive energy conjecture for general relativity*, Phys. Rev. **D59** (1999), 026005 (12 pp.).

[55] G. Huisken and T. Ilmanen, *A note on inverse mean curvature flow*, (1997), Proceedings of the Workshop on Nonlinear Partial Differential Equations (Saitama University, Sept. 1997), available from Saitama University.

[56] ———, *The Riemannian Penrose inequality*, Int. Math. Res. Not. **20** (1997), 1045–1058.

[57] ———, *The inverse mean curvature flow and the Riemannian Penrose inequality*, Jour. Diff. Geom. **59** (2001), 353–437, URL `http://www.math.nwu.edu/~ilmanen`.

[58] P. S. Jang and R. M. Wald, *The positive energy conjecture and the cosmic censor hypothesis*, J. Math. Phys. **18** (1977), 41–44.

[59] J. Jezierski, *Positivity of mass for certain space-times with horizons*, Class. Quantum Grav. **6** (1989), 1535–1539.

[60] ———, *Perturbation of initial data for spherically symmetric charged black hole and Penrose conjecture*, Acta Phys. Pol. B **25** (1994), 1413–1417.

[61] J. Kánnár, *Hyperboloidal initial data for the vacuum Einstein equations with cosmological constant*, Class. Quantum Grav. **13** (1996), 3075–3084.

[62] F. Kottler, *Über die physikalischen Grundlagen der Einsteinschen Gravitationstheorie*, Annalen der Physik **56** (1918), 401–462.

[63] M. Ludvigsen and J.A.G. Vickers, *An inequality relating the total mass and the area of a trapped surface in general relativity*, Jour. Phys. A: Math. Gen. **16** (1983), 3349–3353.

[64] E. Malec, M. Mars, and W. Simon, *On the Penrose inequality for general horizons*, Phys. Rev. Lett. **88** (2002), 121102, gr-qc/0201024.

[65] E. Malec and N. Ó Murchadha, *Trapped surfaces and the Penrose inequality in spherically symmetric geometries*, Phys. Rev. **D 49** (1994), 6931–6934.

[66] E. Malec and K. Roszkowski, *Comment on the Herzlich's proof of the Penrose inequality*, Acta Phys. Pol. **B29** (1998), 1975–1978, gr-qc/9806035.

[67] D. Maxwell, *Solutions of the Einstein constraint equations with apparent horizon boundary*, (2003), gr-qc/0307117.

[68] W.H. Meeks, III and S.T. Yau, *Topology of three-dimensional manifolds and the embedding problems in minimal surface theory*, Ann. of Math. (2) **112** (1980), 441–484.

[69] T. Parker and C. Taubes, *On Witten's proof of the positive energy theorem*, Commun. Math. Phys. **84** (1982), 223–238.

[70] R. Penrose, *Gravitational collapse — the role of general relativity*, Riv. del Nuovo Cim. (numero speziale) **1** (1969), 252–276.

[71] R. Schoen and S.-T. Yau, *Positivity of the total mass of a general space–time*, Phys. Rev. Lett. **43** (1979), 1457–1459.

[72] ———, *Proof of the positive mass theorem*, Comm. Math. Phys. **65** (1979), 45–76.

[73] ———, *Proof of the positive mass theorem II*, Comm. Math. Phys. **79** (1981), 231–260.

[74] W. De Sitter, *On the curvature of space*, Proc. Kon. Ned. Akad. Wet. **20** (1917), 229–243.

[75] K. P. Tod, *The hoop conjecture and the Gibbons-Penrose construction of trapped surfaces*, Class. Quantum Grav. **9** (1992), 1581–1591.

[76] A. Trautman, *King's College lecture notes on general relativity, May-June 1958*, mimeographed notes; reprinted in *Gen. Rel. Grav.* **34**, 721–762 (2002).

[77] ———, *Radiation and boundary conditions in the theory of gravitation*, Bull. Acad. Pol. Sci., Série sci. math., astr. et phys. **VI** (1958), 407–412.

[78] R.M. Wald, *General relativity*, University of Chicago Press, Chicago, 1984.

[79] X. Wang, *Mass for asymptotically hyperbolic manifolds*, Jour. Diff. Geom. **57** (2001), 273–299.

[80] E. Witten, *A simple proof of the positive energy theorem*, Commun. Math. Phys. **80** (1981), 381–402.

[81] X. Zhang, *A definition of total energy-momenta and the positive mass theorem on asymptotically hyperbolic 3 manifolds I*, (2001), preprint.

Index