Black Holes, Geometric Flows, and the Penrose Inequality in General Relativity

Hubert L. Bray

In 1973 R. Penrose [13] made a physical argument that the total mass of a spacetime containing 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*. The Riemannian Penrose inequality was first proved by G. Huisken and T. Ilmanen in 1997 for a single black hole [8] and then by the author in 1999 for any number of black holes [1]. The two approaches use two different geometric flow techniques. The most general version of the Penrose inequality is still open.

A natural interpretation of the Penrose inequality is that the mass contributed by a collection of black holes is (at least) $\sqrt{A/16\pi}$. More generally, the question, How much matter is in a given region of a spacetime? is still very much an open problem [6]. In this paper we will discuss some of the qualitative aspects of mass in general relativity, look at some informative examples, and describe the two very geometric proofs of the Riemannian Penrose inequality.

Hubert L. Bray is assistant professor of mathematics at the Massachusetts Institute of Technology. His email address is `bray@math.mit.edu`.

Research supported in part by NSF grants #DMS-9706006 and #DMS-9971960.

Total Mass in General Relativity

Two notions of mass which are well understood in general relativity are local energy density at a point and the total mass of an asymptotically flat spacetime. However, defining the mass of a region larger than a point but smaller than the entire universe is not at all well understood.

Suppose (M^3, g) is a Riemannian 3-manifold isometrically embedded in a (3+1) dimensional Lorentzian spacetime N^4. Suppose that M^3 has zero second fundamental form in the spacetime. This is a simplifying assumption which allows us to think of (M^3, g) as a "$t = 0$" slice of the spacetime. (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 allows for M^3 to have general second fundamental form) is known as the Riemannian Penrose inequality when the second fundamental form is set to zero.

We also want to consider only (M^3, g) that are asymptotically flat at infinity, which means that for some compact set K, the "end" $M^3 \setminus K$ is diffeomorphic to $\mathbf{R}^3 \setminus B_1(0)$, where the metric g is asymptotically approaching (with certain decay conditions) the standard flat metric δ_{ij} on $\mathbf{R}^3$ at infinity. The simplest example of an asymptotically flat manifold is $(\mathbf{R}^3, \delta_{ij})$ itself. Other good examples are the conformal metrics $(\mathbf{R}^3, u(x)^4\delta_{ij})$, 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 $M^3 \setminus K$ must have the property described above.) A qualitative picture of an asymptotically flat 3-manifold is shown in Figure 1.

The point of these assumptions on the asymptotic behavior of (M^3, g) at infinity is that they imply the existence of the limit

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

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) of (M^3, g) and does not depend on the choice of an asymptotically flat coordinate chart.

The above equation is where many people would stop reading an article like this. But before you do, we will promise not to use this definition of the total mass in this paper. In fact, it turns out that total mass can be quite well understood with an example. Going back to the example $(\mathbf{R}^3, u(x)^4\delta_{ij})$, if we suppose that $u(x) > 0$ has the asymptotics at infinity

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

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

(2) $$m = 2ab.$$

Furthermore, suppose that (M^3, g) is any metric whose "end" is isometric to $(\mathbf{R}^3 \setminus K, u(x)^4\delta_{ij})$, where $u(x)$ is harmonic in the coordinate chart of the end $(\mathbf{R}^3 \setminus K, \delta_{ij})$ and goes to a constant at infinity. Then expanding $u(x)$ in terms of spherical harmonics demonstrates that $u(x)$ satisfies condition (1). We will call such Riemannian manifolds (M^3, g) *harmonically flat at infinity*, and we note that the total mass of these manifolds is also given by equation (2).

A very nice lemma by Schoen and Yau states that, given any $\epsilon > 0$, it is always possible to perturb an asymptotically flat manifold to become harmonically flat at infinity in such a way 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 need to consider only harmonically flat manifolds! Thus, we can use equation (2) as our definition of total mass. As an example, note that $(\mathbf{R}^3, \delta_{ij})$ 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.

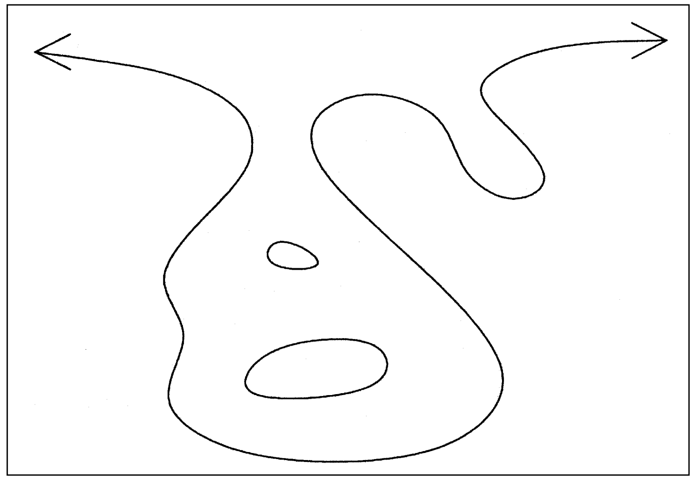

Figure 1.

The Phenomenon of Gravitational Attraction

What do the above definitions of total mass have to do with anything physical? That is, if the total mass is the $1/r$ rate at which the metric becomes flat at infinity, what does this have to do with our real-world intuitive idea of mass?

The answer to this question is very nice. Given a Schwarzschild spacetime metric

$$\left(\mathbf{R}^4, \left(1 + \frac{m}{2|x|}\right)^4 (dx_1^2 + dx_2^2 + dx_3^2) - \left(\frac{1 - m/2|x|}{1 + m/2|x|}\right)^2 dt^2\right), |x| > m/2,$$

for example, the $t = 0$ slice (which has zero second fundamental form) is the spacelike Schwarzschild metric $(\mathbf{R}^3 \setminus B_{m/2}(0), (1 + \frac{m}{2|x|})^4\delta_{ij})$ (discussed more later). According to equation (2), the parameter m is in fact the total mass of this 3-manifold.

On the other hand, suppose we were to release a small test particle, initially at rest, a large distance r from the center of the Schwarzschild spacetime. If this particle is not acted upon by external forces, then it should follow a geodesic in the spacetime. It turns out that with respect to the asymptotically flat coordinate chart, these geodesics "accelerate" towards the middle of the Schwarzschild metric at a rate proportional to m/r^2 (in the limit as r goes to infinity). Thus, our Newtonian notion of mass also suggests that the total mass of the spacetime is m.

Local Energy Density

Another well-understood quantification of mass is local energy density. In fact, in our setting the local energy density at each point is

$$\mu = \frac{1}{16\pi} R,$$

where R is the scalar curvature of the 3-manifold (which has zero second fundamental form in the spacetime) at each point. Note that $(\mathbf{R}^3, \delta_{ij})$ has zero energy density at each point as well as zero total mass. This is appropriate since $(\mathbf{R}^3, \delta_{ij})$ is in fact a "$t = 0$" slice of Minkowski spacetime, which represents a vacuum. Classically, physicists consider $\mu \geq 0$ to be a physical assumption. Hence, from this point on, we will assume not only that (M^3, g) is asymptotically flat but also that it has nonnegative scalar curvature,

$$R \geq 0.$$

This notion of energy density also helps us understand total mass better. After all, we can take any asymptotically flat manifold and then change the metric to be perfectly flat outside a large compact set, thereby giving the new metric zero total mass. However, if we introduce the physical condition that both metrics have nonnegative scalar curvature, then it is a beautiful theorem that such a modification is not possible unless the original metric was already $(\mathbf{R}^3, \delta_{ij})$! (This theorem is actually a corollary to the positive mass theorem discussed in a moment.) Thus, the curvature obstruction of having nonnegative scalar curvature at each point is a very interesting condition.

Also, notice the indirect connection between the total mass and the local energy density. At this point, there does not seem to be much of a connection at all. Total mass is the $1/r$ rate at which the metric becomes flat at infinity, and local energy density is the scalar curvature at each point. Furthermore, if a metric is changed in a compact set, local energy density is changed, but the total mass is unaffected.

Indeed, the total mass is *not* the integral of the local energy density over the manifold. In fact, this integral fails to take into account either potential energy (which would be expected to contribute a negative energy) or gravitational energy (discussed in a moment). Hence, it is not initially clear what we should expect the relationship between total mass and local energy density to be, so let us begin with an example.

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

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

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

Hence, since the physical assumption of nonnegative energy density implies nonnegative scalar curvature, we see that the positive function $u(x)$ must be superharmonic (that is, $\Delta u \leq 0$). For simplicity, assume also 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). By the maximum principle, it follows that the minimum value for $u(x)$ must be a, referring to equation (1). Hence, $b \geq 0$, which implies that $m \geq 0$! Thus we see that the assumption of nonnegative energy density at each point of $(\mathbf{R}^3, u(x)^4\delta_{ij})$ implies that the total mass is also nonnegative, which is what one would hope.

The Positive Mass Theorem

Why would one hope this? What would be the difference if the total mass were negative? This would mean that a gravitational system of positive energy density could collectively act as a net negative total mass. This phenomenon has not been observed experimentally, and so it is not a property that we would expect to find in general relativity.

More generally, suppose that 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, proved first by Schoen and Yau [14] in 1979 using minimal surface techniques and then by Witten [17] in 1981 using spinors. In the case of zero second fundamental form, the positive mass theorem is known as the Riemannian positive mass theorem and is stated below.

Theorem 1. [16] *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 $(\mathbf{R}^3, \delta)$.*

Gravitational Energy

The preceding example fails to illustrate all of the subtleties of the positive mass theorem. For example, it is easy to construct asymptotically flat manifolds (M^3, g) (not conformal to $\mathbf{R}^3$) that have zero scalar curvature everywhere and yet have *nonzero* total mass. By the positive mass theorem, the mass of these manifolds is positive. Physically, this corresponds to a spacetime that has zero energy density everywhere and yet still has positive total mass. From where did this mass come? How can a vacuum have positive total mass?

Physicists refer to this extra energy as gravitational energy. There is no known local definition of the energy density of a gravitational field, and presumably such a definition does not exist. The curious phenomenon then is that for some reason gravitational energy always makes a nonnegative contribution to the total mass of the system.

Black Holes

Another very interesting and natural phenomenon in general relativity is the existence of black holes. Instead of thinking of black holes as singularities in a spacetime, we will think of black holes in terms of their horizons. For example, suppose we are exploring the universe in a spacecraft capable of traveling any speed less than the speed of light. If we are investigating a black hole, we want to make sure that we do not get too close and get trapped by the "gravitational forces" of the black hole. We can imagine a "sphere of no return" beyond which escape from the black hole is impossible. It is called the event horizon of a black hole.

However, one limitation of the notion of an event horizon is the difficulty of determining its location. One way is to let daredevil spacecraft see how close they can get to the black hole and still escape from it eventually. The problem with this approach (besides the cost in spacecraft) is that it is hard to know when to stop waiting for a daredevil spacecraft to return. Even if it has been fifty years, it could be that this particular daredevil was not trapped by the black hole but got so close that it will take one thousand or more years to return. Thus, to define the location of an event horizon even mathematically, we need to know the entire evolution of the spacetime. Hence, event horizons cannot be computed based only on the local geometry of the spacetime.

This problem is solved (at least for the mathematician) by the notion of apparent horizons of black holes. Given a surface in a spacetime, 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. The outermost boundary of these trapped surfaces is called the apparent horizon of the black hole. Apparent horizons can be computed based on their local geometry, and an apparent horizon always implies the existence of an event horizon outside of it [7].

Now let us return to the case we are considering in this paper where (M^3, g) is a "$t = 0$" slice of a spacetime with zero second fundamental form. Then it is a very nice geometric fact that 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 outermost is well defined with respect to a chosen end of the manifold [1]. 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 in Figure 2.

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

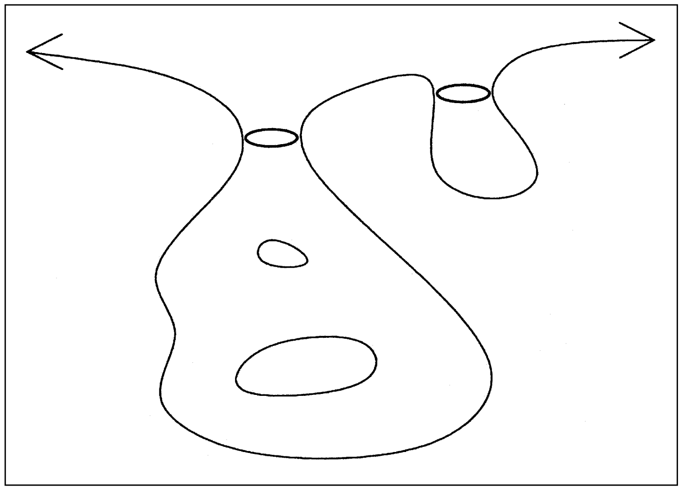

Figure 2.

It also follows from a stability argument (which interestingly uses the Gauss-Bonnet theorem) that each component of an outermost minimal surface (in a 3-manifold with nonnegative scalar curvature) must have the topology of a sphere. Furthermore, there is a physical argument, based on [13], which suggests that the mass contributed by the black holes (thought of as the connected components of Σ_0) should be defined to be $\sqrt{A_0/16\pi}$, where A_0 is the area of Σ_0. Hence, the physical argument that the total mass should be greater than or equal to the mass contributed by the black holes yields the following geometric statement.

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

$$m \geq \sqrt{\frac{A_0}{16\pi}}, \tag{3}$$

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

The above statement has been proved by the author [1], and Huisken and Ilmanen [8] proved it when A_0 is defined instead to be the area of the largest connected component of Σ_0. In this article we will discuss both approaches. They 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. Nothing very important is gained by considering manifolds with more than one end, since extra ends can always be compactified by connect summing them (around a neighborhood of infinity) with large spheres while still preserving nonnegative scalar curvature. 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.

The Schwarzschild Metric

The Schwarzschild metric $(\mathbf{R}^3 \setminus \{0\}, (1 + \frac{m}{2|x|})^4 \delta_{ij})$ referred to in the above statement of the Riemannian Penrose Inequality is a particularly important example to consider. It corresponds to a zero second fundamental form, spacelike slice of the usual (3+1)-dimensional Schwarzschild metric (which represents a spherically symmetric static black hole in a vacuum). The 3-dimensional Schwarzschild metrics have total mass $m > 0$ and are characterized by being the only spherically symmetric, geodesically complete, zero scalar curvature 3-metrics other than $(\mathbf{R}^3, \delta_{ij})$. They 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 fixing 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.

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 (3) using a very clever physical argument [13], which we will not have room to repeat here. His observation was that a counterexample to inequality (3) 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 spacetime unless they are inside a black hole).

In 1977 Jang and Wald [11], extending ideas of Geroch, gave a heuristic proof of inequality (3) 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 it is nondecreasing under this "inverse mean curvature flow". Indeed, under inverse mean curvature flow it follows from the Gauss equation and the second variation formula that

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

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, since $R \geq 0$, and

$$\int_\Sigma K \leq 4\pi \tag{4}$$

(true for any connected surface by the Gauss-Bonnet Theorem), it follows that

$$\frac{d}{dt} m_{Hawking}(\Sigma) \geq 0. \tag{5}$$

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 (3) follows from inequality (5).

As noted by Jang and Wald, this argument works only 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. The problem 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 [8] in 1997. With a very clever new approach, Huisken and Ilmanen discovered how to reformulate inverse mean curvature flow using an energy minimization principle in such a way that the new generalized 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 (3) for a single black hole.

Coincidentally, the author found another proof of inequality (3), submitted in 1999, which works for any number of black holes. 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 (3), 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 outwards 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 on the Penrose Conjecture have been made by Herzlich using the Dirac operator, which Witten used to prove the positive mass theorem; by Gibbons in the special case of collapsing shells; by Tod; by Bartnik for quasi-spherical metrics; and by the author using isoperimetric surfaces. There is also some interesting work of Ludvigsen and Vickers using spinors and of Bergqvist, both concerning the Penrose inequality for null slices of a spacetime.

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. Furthermore, the Hawking mass is still monotone under this new modified flow. Notice that when $\Sigma(t)$ jumps outwards 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,

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

because (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 cannot have $|\bar{\Sigma}(t)| < |\Sigma(t)|$ (since $\Sigma(t)$ would have jumped outwards at some earlier time). This is only a heuristic argument, but we can then see by the preceding two equations that the Hawking mass is nondecreasing during a jump.

This new flow can be rigorously defined, always exists, and the Hawking mass is monotone. In [8] 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

$$\text{div}\left(\frac{\nabla u}{|\nabla u|}\right) = |\nabla u| \tag{6}$$

in an appropriate weak sense. Since the left-hand side of this 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 equation characterizes inverse mean curvature flow for the level sets of $u(x)$ when $|\nabla u(x)| \neq 0$.

Huisken and Ilmanen use an energy minimization principle to define weak solutions to equation (6). Equation (6) is said to be weakly satisfied in Ω by the locally Lipschitz function u if for every locally Lipschitz function v with $\{v \neq u\} \subset\subset \Omega$,

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

where

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

The Euler-Lagrange equation of the above energy functional yields equation (6).

In order to prove that a solution u exists to the preceding two equations, Huisken and Ilmanen regularize the degenerate elliptic equation (6) to the elliptic equation

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

Solutions to this equation are shown to exist using the existence of a subsolution; then taking the limit as ϵ goes to zero yields a weak solution to equation (6). We are skipping many details here, but these are the main ideas.

As it turns out, weak solutions $u(x)$ to equation (6) often have flat regions where $u(x)$ equals a constant. The level sets $\Sigma(t)$ of $u(x)$ are discontinuous in t in this case, which corresponds to the "jumping out" phenomenon referred to earlier.

We also note that since the Hawking mass of the level 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 asymptotically hyperbolic 3-manifolds having $R \geq -6$. This situation occurs if (M^3, g) is chosen to be a constant mean curvature slice of the spacetime or if the spacetime is defined to solve the Einstein equation with nonzero cosmological constant. In these cases there exists a modified Hawking mass that is monotone under inverse mean curvature flow—it is the usual Hawking mass plus $4(|\Sigma|/16\pi)^{3/2}$. However, 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 (4), inverse mean curvature flow proves the Riemannian Penrose Inequality only 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 generalized to higher dimensions.

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) for $0 \leq t < \infty$. 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. 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. 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

$$A'(t) = 0,$$
$$m'(t) \leq 0$$

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

$$m(0) \geq m(\infty) = \sqrt{\frac{A(\infty)}{16\pi}} = \sqrt{\frac{A(0)}{16\pi}}, \tag{7}$$

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.

The Definition of the Flow

In fact, 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 ordinary differential equation in t defined by equations (8), (9), (10), and (11). Let

$$g_t = u_t(x)^4 g_0 \tag{8}$$

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

$$\Sigma(t) = \text{the outermost minimal area enclosure of } \Sigma_0 \text{ in } (M^3, g_t), \tag{9}$$

where Σ_0 is the original outer minimizing 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 minimizing horizon of (M^3, g_t). Given the horizon $\Sigma(t)$, define $v_t(x)$ such that

$$\begin{cases} \Delta_{g_0} v_t(x) & \equiv 0 \text{ outside } \Sigma(t) \\ v_t(x) & = 0 \text{ on } \Sigma(t) \\ \lim_{x\to\infty} v_t(x) & = -e^{-t} \end{cases} \tag{10}$$

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

$$u_t(x) = 1 + \int_0^t v_s(x)\,ds, \tag{11}$$

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

Now equation (11) 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, of g_0, and of $v_t(x)$, 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.

Theorem 2. *Taken together, equations (8), (9), (10), and (11) define a first-order ordinary differential equation in t for $u_t(x)$ having a solution which is Lipschitz in the t variable, class C^1 in the x variable everywhere, and smooth in the x variable outside $\Sigma(t)$. Furthermore, $\Sigma(t)$ is a smooth, strictly outer minimizing 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 (11) 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

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

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

Even so, it still may not seem that g_t is naturally defined, since the rate of change of g_t appears to depend on both t and the original metric g_0 in equation (10). 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 [1] we prove this very important fact and define a new equivalence class of metrics called the harmonic conformal class. 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, we are basically forced to choose the particular conformal flow of metrics defined above.

Theorem 3. *The function $A(t)$ is constant in t, and $m(t)$ is nonincreasing in t, for all $t \geq 0$.*

That $A'(t) = 0$ follows because to first order the metric is not changing on $\Sigma(t)$ (since $v_t(x) = 0$ there) and 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 is proving that $m'(t) \leq 0$. Curiously, this follows from a nice trick using the Riemannian positive mass theorem, which we describe later.

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 $(\mathbf{R}^3 \setminus \{0\}, s)$ in the limit as t goes to ∞. More precisely:

Theorem 4. *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 $(\mathbf{R}^3 \setminus \{0\}, s)$ 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 4 is not really surprising, although a careful proof is rather long. However, if one is willing to believe that the flow of metrics converges to a spherically symmetric metric outside the horizon, then Theorem 4 follows from two observations. The first 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 (12). The second observation 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 (3) then follows for harmonically flat manifolds [1] from equation (7) using Theorems 2, 3, and 4. 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.

Qualitative Discussion

The two diagrams below are meant to help illustrate some of the properties of the conformal flow of the metric. Figure 3 is the original metric, which has a strictly outer minimizing horizon Σ_0. As t increases, $\Sigma(t)$ moves outwards but never inwards. In Figure 4 we can observe one of the consequences of the fact that $A(t) = A_0$ is constant in t. Since the

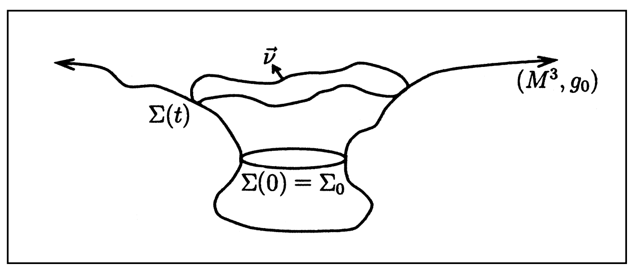

Figure 3.

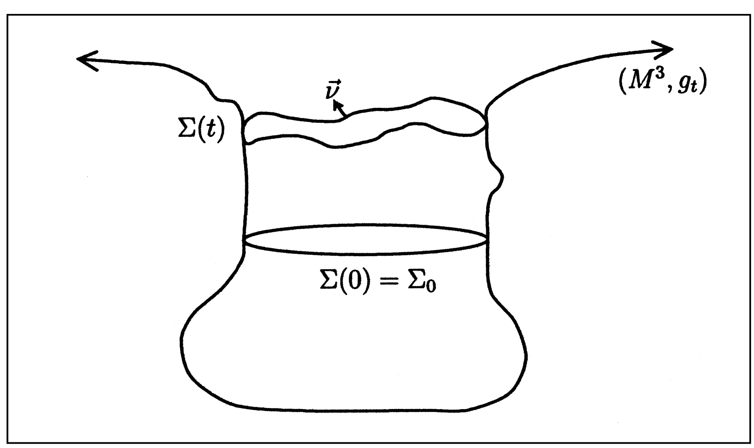

Figure 4.

metric is not changing inside $\Sigma(t)$, all of the horizons $\Sigma(s)$, $0 \le s \le t$, have area A_0 in (M^3, g_t). Hence, inside $\Sigma(t)$, the manifold (M^3, g_t) becomes cylinder-like in the sense that it is laminated (meaning foliated but with some gaps allowed) by all of the previous horizons, which all have the same area A_0 with respect to the metric g_t.

Now suppose that the original horizon Σ_0 of (M^3, g) had two components, for example. Then each of the components of the horizon will move outwards as t increases, and at some point before they touch they will suddenly jump outwards to form a horizon with a single component enclosing the previous horizon with two components. Even horizons with only one component will sometimes jump outwards, but no more than a countable number of times. It is interesting that this phenomenon of surfaces jumping is also found in the Huisken-Ilmanen approach to the Penrose conjecture using the generalized $1/H$ flow.

Proof that $m'(t) \le 0$. The most surprising aspect of the conformal flow of metrics is that $m'(t) \le 0$. As mentioned above, this important fact follows from a nice trick using the Riemannian positive mass theorem.

The first step is to realize that while the rate of change of g_t appears to depend on both t and g_0, this is in fact an illusion. As described in detail in [1], the rate of change of g_t can be described purely in terms of g_t (and Σ_0). It is also true that the rate of change of g_t depends only on g_t and $\Sigma(t)$. Hence, there is no distinguished value of t, so proving $m'(t) \le 0$ is equivalent to proving $m'(0) \le 0$. Thus, without loss of generality, we take $t = 0$ for convenience.

Now expand the harmonic function $v_0(x)$, defined in equation (10), using spherical harmonics at infinity to get

$$v_0(x) = -1 + \frac{c}{|x|} + \mathcal{O}\left(\frac{1}{|x|^2}\right) \tag{13}$$

for some constant c. Since the rate of change of the metric g_t at $t = 0$ is given by $v_0(x)$ and since the total mass $m(t)$ depends on the $1/r$ rate at which the metric g_t becomes flat at infinity (see equation (2)), it is not surprising that direct calculation gives us

$$m'(0) = 2(c - m(0)).$$

Hence, to show that $m'(0) \le 0$, we need to show that

$$c \le m(0). \tag{14}$$

In fact, counterexamples to equation (14) can be found if we remove either of the requirements that $\Sigma(0)$ (which is used in the definition of $v_0(x)$) be a minimal surface or that (M^3, g_0) have nonnegative scalar curvature. Hence, we quickly see that equation (14) is a fairly deep conjecture which says something quite interesting about manifolds with nonnegative scalar curvature. Now the Riemannian positive mass theorem is also a deep conjecture which says something quite interesting about manifolds with nonnegative scalar curvature. Hence, it is natural to try to use the Riemannian positive mass theorem to prove equation (14).

Thus, we want to create a manifold whose total mass depends on c from equation (13). The idea is to use a reflection trick similar to one used by Bunting and Masood-ul-Alam for another purpose in [5]. First, remove the region of M^3 inside $\Sigma(0)$ and then reflect the remainder of (M^3, g_0) through $\Sigma(0)$. Define the resulting Riemannian manifold to be $(\bar{M}^3, \bar{g}_0)$, which has two asymptotically flat ends, since (M^3, g_0) has exactly one asymptotically flat end not contained by $\Sigma(0)$. Note that $(\bar{M}^3, \bar{g}_0)$ has nonnegative scalar curvature everywhere except on $\Sigma(0)$, where the metric has corners. Moreover, the fact that $\Sigma(0)$ has zero mean curvature (since it is a minimal surface) implies that $(\bar{M}^3, \bar{g}_0)$ has *distributional* nonnegative scalar curvature everywhere, even on $\Sigma(0)$. This notion is made rigorous in [1]. Thus we have used the fact that $\Sigma(0)$ is minimal in a crucial way.

Recall from equation (10) that $v_0(x)$ was defined to be the harmonic function equal to zero on $\Sigma(0)$ which goes to -1 at infinity. We want to reflect $v_0(x)$ to be defined on all of $(\bar{M}^3, \bar{g}_0)$. The trick here is to define $v_0(x)$ on $(\bar{M}^3, \bar{g}_0)$ to be the harmonic function which goes to -1 at infinity in the original end and goes to 1 at infinity in the reflected end. By symmetry, $v_0(x)$ equals 0 on $\Sigma(0)$ and so agrees with its original definition on (M^3, g_0).

The next step is to compactify one end of $(\bar{M}^3, \bar{g}_0)$. By the maximum principle, we know that $v_0(x) > -1$ and $c > 0$, so the new Riemannian manifold $(\bar{M}^3, (v_0(x) + 1)^4 \bar{g}_0)$ does the job quite nicely and compactifies the original end to a point. In fact, the compactified point at infinity and the metric there can be filled in smoothly (using the fact that (M^3, g_0) is harmonically flat). It then follows from equation (12) that this new compactified manifold has nonnegative scalar curvature since $v_0(x) + 1$ is harmonic.

The last step is simply to apply the Riemannian positive mass theorem to $(\bar{M}^3, (v_0(x) + 1)^4 \bar{g}_0)$. It is not surprising that the total mass $\tilde{m}(0)$ of this manifold involves c, but it is quite lucky that direct calculation yields

$$\tilde{m}(0) = -4(c - m(0)),$$

which must be positive by the Riemannian positive mass theorem. Thus, we have that

$$m'(0) = 2(c - m(0)) = -\frac{1}{2}\tilde{m}(0) \le 0.$$

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 inequality, in which M^3 is allowed to have any second fundamental form in the spacetime. 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. For physicists this problem is definitely a top priority since most spacetimes do not even admit zero second fundamental form spacelike slices.

Another interesting problem is to ask these same questions in higher dimensions. The author is currently working on a paper to prove the Riemannian Penrose inequality in dimensions less than 8. Dimensions 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 conjecture is still open in dimensions 8 and higher for manifolds which are not spin manifolds.

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 R_g is the scalar curvature at each point, dV_g is the volume form, and g is scaled so that the total volume of (M^3, g) is equal to 1. 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 (topological) Yamabe invariant of M^3, denoted here as $Y(M^3)$, is defined to be the supremum of $C(g)$ over all metrics g. It is known that $Y(S^3) = 6 \cdot (2\pi^2)^{2/3} \equiv Y_1$ is the largest possible value for the Yamabe invariant of a 3-manifold. 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 nonorientable S^2-bundle over S^1.

The author, working with Andre Neves on a problem suggested by Richard Schoen, recently was able to compute the Yamabe invariant of RP^3 using inverse mean curvature flow techniques [3] and found that $Y(RP^3) = Y_1 / 2^{2/3} \equiv Y_2$. A corollary is that $Y(RP^2 \times S^1) = Y_2$ 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 the invariant 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 $RP^2 \times S^1$.

References

[1] H. L. Bray, Proof of the Riemannian Penrose inequality using the positive mass theorem, *J. Differential Geom.* (to appear).

[2] ____, The Penrose inequality in general relativity and volume comparison theorems involving scalar curvature, thesis, Stanford University, 1997.

[3] H. L. Bray and A. Neves, Classification of prime 3-manifolds with Yamabe invariant greater than RP^3 (in preparation).

[4] H. L. Bray and R. M. Schoen, Recent proofs of the Riemannian Penrose conjecture, *Current Developments in Mathematics 1999* (S.-T. Yau, ed.), International Press.

[5] Gary L. Bunting and A. K. M. Masood-ul Alam, Nonexistence of multiple black holes in asymptotically Euclidean static vacuum space-time, *Gen. Relativity Gravitation* **19** (1987).

[6] D. Christodoulou and S.-T. Yau, Some remarks on the quasi-local mass, *Mathematics and General Relativity,* Contemp. Math., vol. 71, Amer. Math. Soc., Providence, RI, 1988, pp. 9–14.

[7] S. W. Hawking and G. F. R. Ellis, *The Large-Scale Structure of Space-Time*, Cambridge University Press, Cambridge, 1973.

[8] G. Huisken and T. Ilmanen, The inverse mean curvature flow and the Riemannian Penrose inequality, *J. Differential Geom.* (to appear).

[9] ____, The Riemannian Penrose inequality, *Internat. Math. Res. Notices* **20** (1997), 1045–58.

[10] ____, A note on inverse mean curvature flow, *Proceedings of the Workshop on Nonlinear Partial Differential Equations (Saitama University, Sept. 1997)*, available from Saitama University.

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

[12] T. H. Parker and C. H. Taubes, On Witten's proof of the positive energy theorem, *Commun. Math. Phys.* **84** (1982), 223–38 .

[13] R. Penrose, Naked singularities, *Ann. New York Acad. Sci.* **224** (1973), 125–34.

[14] R. Schoen and S.-T. Yau, On the proof of the positive mass conjecture in general relativity, *Comm. Math. Phys.* **65** (1979), 45–76.

[15] ____, Positivity of the total mass of a general spacetime, *Phys. Rev. Lett.* **43** (1979), 1457–9.

[16] ____, Proof of the positive mass theorem II, *Comm. Math. Phys.* **79** (1981), 231–60.

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