The Existence of a Black Hole
Due to Condensation of Matter

Richard Schoen[1] and S.-T. Yau[2]

1 Department of Mathematics, University of California, Berkeley, CA 94720, USA
2 School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA

Abstract. When enough matter is condensed in a small region, gravitational effects will be strong enough to cause collapse and a black hole will be formed. We formulate and prove here such a statement in the language of general relativity. (This is Theorem 2 of this paper.)

The main result of this paper is that for an asymptotically flat initial data set, with the mass density large on a large region, there is an apparent horizon (and a closed trapped surface) in the initial data. It was shown by Penrose [2] and Hawking [1] that under physically reasonable assumptions, the existence of a closed trapped surface implies that the spacetime which evolves from the initial data contains a black hole. Simple examples show that the mass density can be large on a set of arbitrarily large *diameter* without the existence of an apparent horizon in the initial data. Therefore, the notion of a "large region" must be suitably defined. We formulate a notion which measures more than one direction in Ω.

Definition. Let Γ be a simple closed curve in Ω which bounds a disk in Ω. We let $N_r(\Gamma)$ denote the set of points within a distance r of Γ. Define the *H-radius of Ω with respect to Γ* by

$$\mathrm{Rad}(\Omega, \Gamma) = \sup\{r : \mathrm{dist}(\Gamma, \partial\Omega) > r, \ \Gamma \text{ does not bound a disk in } N_r(\Gamma)\}.$$

We define the *H-radius of Ω*, denoted $\mathrm{Rad}(\Omega)$, by

$$\mathrm{Rad}(\Omega) = \sup\{\mathrm{Rad}(\Omega, \Gamma) : \Gamma \text{ as above}\}.$$

Remark 1. Note that if Ω were a ball of radius R in $\mathbb{R}^3$, then $\mathrm{Rad}(\Omega) = R/2$; on the other hand if Ω is the cross product of $S^2(R)$ with an interval $(-L, L)$, then $\mathrm{Rad}(\Omega) = \min\left\{\dfrac{\pi R}{2}, L\right\}$.

Remark 2. One can alternatively replace homotopy by homology in the above definition and define a homology radius of Ω with respect to a curve (or family of curves).

Now suppose (N, g, P) is an initial data set. We define μ, J^i by

$$\mu = \tfrac{1}{2}\Big[R - \sum_{i,j} P^{ij}P_{ij} + \Big(\sum_i P^i{}_i\Big)^2\Big],$$
$$J^i = \sum_j D_j\Big[P^{ij} - \Big(\sum_k P^k{}_k\Big)g^{ij}\Big],$$

where R is the scalar curvature of the Riemannian metric g and P_{ij} is the second fundamental form of N in the space-time. Our first result deals with the case of a maximal initial data set.

Theorem 1. *Suppose* $\mathrm{Tr}(P) = \sum_{i,j} g^{ij}P_{ij} \equiv 0$. *If Ω is any open set in N such that for some number $\Lambda > 0$ we have $\mu \geq \Lambda$ on Ω, then we have the inequality*

$$\mathrm{Rad}(\Omega) \leq \frac{2\pi}{\sqrt{3}\sqrt{\Lambda}}.$$

Remark 3. There is no requirement that N be asymptotically flat in Theorem 1. The result is purely local.

Theorem 1 will be an immediate consequence of the following result.

Proposition 1. *Suppose N is a three dimensional Riemannian manifold and $\Omega \subset N$ is a bounded region such that the first Dirichlet eigenvalue on Ω of the operator $-\Delta + \tfrac{1}{2}R$ is at least Λ. Then* $\mathrm{Rad}(\Omega) \leq \sqrt{\dfrac{3}{2}} \dfrac{\pi}{\sqrt{\Lambda}}$.

Proof. Let $f > 0$ be the first Dirichlet eigenfunction of $-\Delta + \tfrac{1}{2}R$ on Ω. Thus we have

$$\Delta f + (\lambda - \tfrac{1}{2}R)f = 0, \qquad \lambda \geq \Lambda. \tag{1}$$

Let ϱ be any positive number less than $\mathrm{Rad}(\Omega)$, and let Γ be a curve such that $\mathrm{Rad}(\Omega, \Gamma) > \varrho$. Without loss of generality we may assume N is a complete manifold and $f > 0$ is defined on N; no requirement is made on $N \sim \Omega$. For any disk Σ spanning Γ, define $A_f(\Sigma)$ by

$$A_f(\Sigma) = \int_\Sigma f d\sigma, \qquad d\sigma = \text{area element}.$$

Let Σ be an immersed disk in N with $\partial\Sigma = \Gamma$ such that Σ minimizes $A_f(\Sigma)$. Since Σ minimizes A_f, it has nonnegative second variation of A_f. This implies, by a calculation, that the first Dirichlet eigenvalue of the operator L given below is nonnegative on Σ

$$L(\varphi) = -[\Delta\varphi + f^{-1}\nabla\varphi \cdot \nabla f + (\tfrac{1}{2}R - f^{-1}\Delta_N f - \tfrac{1}{2}K + f^{-1}\Delta f)\varphi],$$

where Δ, ∇ are taken on Σ and we have used Δ_N to denote the three-dimensional Laplace operator; K is twice the usual Gauss curvature of Σ. Substituting (1) into this expression, we get on $\Sigma \cap \Omega$,

$$L(\varphi) = -[\Delta\varphi + f^{-1}\nabla\varphi \cdot \nabla f + (\lambda - \tfrac{1}{2}K + f^{-1}\Delta f)\varphi].$$

Let $g > 0$ on Σ be the first eigenfunction of L. Thus g satisfies the inequality

$$\Delta g + f^{-1} Vf \cdot Vg + gf^{-1}\Delta f + (\lambda - \tfrac{1}{2}K)g \leqq 0. \tag{2}$$

Since $\mathrm{Rad}(\Omega, \Gamma) > \varrho$, there is a point $x \in \Sigma \cap \partial N_\varrho(\Gamma)$. Consider curves γ lying in Σ which connect x to $\Gamma = \partial\Sigma$, and define $I[\gamma]$ by

$$I[\gamma] = \int_\gamma fg\,ds,$$

where ds is arclength along γ. Let γ be a curve which minimizes $I(\gamma)$ over all curves from x to Γ. By replacing Γ by a terminal segment of γ if necessary, we construct a curve γ connecting a point $\Sigma \cap \partial N_\varrho(\Gamma)$ to Γ which minimizes I and is contained in $N_\varrho(\Gamma)$. In particular, we have

$$\mathrm{Length}(\gamma) = \int_\gamma ds \geqq \varrho. \tag{3}$$

The fact that γ has nonnegative second variation for I implies, by a calculation, that the operator L_0 on γ has nonnegative first eigenvalue

$$\begin{aligned}
L_0 \psi = -\Bigg[& \frac{d^2\psi}{ds^2} + f^{-1}\frac{d\psi}{ds}\frac{df}{ds} + g^{-1}\frac{d\psi}{ds}\frac{dg}{ds} \\
& + \left(\lambda + g^{-1}Lg + f^{-1}\frac{d^2f}{ds^2} + g^{-1}\frac{d^2g}{ds^2} + f^{-1}g^{-1}\frac{df}{ds}\frac{dg}{ds}\right)\psi \Bigg].
\end{aligned}$$

Combining this with (2) and letting $h(s)$ denote the first eigenfunction of L_0 on $[0, l]$, $l = \mathrm{Length}(\gamma)$, we then have $h > 0$ and

$$h^{-1}h'' + f^{-1}f'' + g^{-1}g'' + h^{-1}f^{-1}h'f' + h^{-1}g^{-1}h'g' + f^{-1}g^{-1}f'g' + \lambda \leqq 0$$

on $[0, l]$, where we use primes to denote $\dfrac{d}{ds}$. Let φ be any function vanishing at $s = 0$ and $s = l$, and multiply both sides of the above inequality by φ^2, and integrate by parts to get

$$\int_0^l \left\{ \frac{1}{2}(h^{-2}(h')^2 + f^{-2}(f')^2 + g^{-2}(g')^2)\varphi^2 + \frac{1}{2}\left(\frac{d}{ds}\log fgh\right)^2 \varphi^2 + \lambda\varphi^2 \right\} sS$$

$$\leqq 2\int_0^l \varphi\varphi'\frac{d}{dS}\log fgk\,ds, \tag{4}$$

where we have done a rearrangement on the left. The following inequality is easily checked

$$2\left|\varphi\varphi'\frac{d}{ds}\log fgh\right| \leqq \frac{3}{2}(\varphi')^2 + [\text{First two terms of (4)}]. \tag{5}$$

Thus we finally get

$$\lambda\int_0^l \varphi^2\,ds \leqq \frac{3}{2}\int_0^l (\varphi')^2\,dS,$$

which implies the operator $-\dfrac{d^2}{ds^2} - \dfrac{2}{3}\lambda$ has nonnegative first eigenvalue on $[0, l]$. This clearly implies

$$l \leq \sqrt{\frac{3}{2}} \frac{\pi}{\sqrt{\lambda}} \leq \sqrt{\frac{3}{2}} \frac{\pi}{\sqrt{\Lambda}}.$$

By (3) and the fact that ϱ is any number less than $\mathrm{Rad}(\Omega)$, we get the conclusion of Proposition 1.

Proof of Theorem 1. Theorem 1 is essentially the case of Proposition 1 with the function f constant. (Note that $\Lambda \leq \mu \leq \frac{1}{2}R$.) We then observe that the coefficient of $(\varphi')^2$ in (5) can be improved to $\frac{4}{3}$, which then gives the conclusion of Theorem 1.

Our main theorem will rely on the following existence result which was proved in [4].

Proposition 2. *Suppose $\Omega_1 \subset N$ is a bounded region. Assume that H_1 denotes the mean curvature of $\partial\Omega_1$ taken relative to the inward normal. Assume that on $\partial\Omega_1$ the inequality $H_1 > |\mathrm{Tr}_{\partial\Omega_1}(P)|$ holds at each point. Here $\mathrm{Tr}_{\partial\Omega_1}(P)$ refers to the trace relative to the induced metric on $\partial\Omega_1$ of the restriction of P to $\partial\Omega_1$. If Ω_1 contains no apparent horizons, then there is a unique solution f on Ω_1 to the equation*

$$\sum_{i,j=1}^{3} \left(g^{ij} - \frac{f^i f^j}{1 + |\nabla f|^2} \right)(f_{ij} - P_{ij}) = 0. \tag{6}$$

Remark 4. In our convention the unit ball in $\mathbb{R}^3$ has boundary of positive mean curvature relative to the inward normal. If N is asymptotically flat and Ω_1 is the interior of a large ball, then the boundary condition assumed in Proposition 2 is satisfied.

We now state our main theorem.

Theorem 2. *Let $\Omega \subset N$ be a bounded region on which the inequality $\mu - |J| \geq \Lambda$ holds. Assume either that N is asymptotically flat or, more generally, that $\Omega \subset \Omega_1$, where Ω_1 is as in Proposition 1. If $\mathrm{Rad}(\Omega) \geq \sqrt{\dfrac{3}{2}} \dfrac{\pi}{\sqrt{\Lambda}}$, then N (respectively Ω_1) contains an apparent horizon Σ, that is a closed two sphere satisfying $H_\Sigma = \pm \mathrm{Tr}_\Sigma(P)$. [In particular, there is a closed (future or past) trapped surface.] Moreover, any such apparent horizon Σ lying entirely within Ω has diameter at most $\dfrac{2\pi}{\sqrt{3}\sqrt{\Lambda}}$. Any such Σ intersecting $\partial\Omega$ has the property that $\Omega \cap \Sigma$ lies everywhere within a distance $\dfrac{2\pi}{\sqrt{3}\sqrt{\Lambda}}$ of $\partial\Omega$.*

Proof. By Proposition 2, if such an apparent horizon does not exist, then (6) is solvable on Ω_1. If such a solution f of (6) exists, then by inequality (2.29) of [4] we have on Ω

$$\Lambda \leq \frac{1}{2}\bar{R} - \sum_i (h_{i4} - P_{i4})^2 + \sum_i D_i(h_{i4} - P_{i4}),$$

where $\bar{R}$ denotes the scalar curvature of N with metric $\bar{g}$ given by $\bar{g}_{ij} = g_{ij} + f_i f_j$, and the other terms are defined in [3]. This inequality easily implies that for φ vanishing on $\partial\Omega$ we have

$$\Lambda \int_\Omega \varphi^2 \overline{dv} \leqq \int_\Omega (\tfrac{1}{2}\bar{R}\varphi^2 + |D\varphi|^2)\overline{dv}.$$

In other words, the first Dirichlet eigenvalue of the operator $-\bar{\Delta} + \tfrac{1}{2}\bar{R}$ is at least Λ. Since the metric $\bar{g}$ is larger than g, it follows that $\text{Rad}(\Omega)$ is greater than $\sqrt{\dfrac{3}{2}}\dfrac{\pi}{\sqrt{\Lambda}}$ relative to both g and $\bar{g}$. This is a contradiction to Proposition 1. Therefore no solution f to (6) can exist on Ω_1, and an apparent horizon Σ exists in Ω_1. The statements about the diameter of Σ follow also from Proposition 1 because such Σ have the property (see [5]) that the operator $-\Delta + \tfrac{1}{2}K_\Sigma$ has first eigenvalue at least Λ, and hence any curve γ which minimizes the functional $I[\gamma]$ considered in Proposition 1 cannot have a segment longer than $\dfrac{2\pi}{\sqrt{3}\sqrt{\Lambda}}$ lying within Ω. This completes the proof of Theorem 2.

Note. In [4], we mentioned the classification of asymptotically flat space which satisfies the local energy condition. We should point out that in the classification mentioned there, we assume the initial data set does not contain any apparent horizon. In this regard, we have proved that a compact three-dimensional manifold which admits a metric with positive scalar curvature is diffeomorphic to the connected sum of copies of $S^2 \times S^1$ and copies of three dimensional manifolds with finite fundamental group.

References

1 Hawking, S, Penrose, R : The singularities of gravitational collapse and cosmology Proc R Soc. London Ser A **314**, 529–548 (1970)
2 Penrose, R : Gravitational collapse and space-time singularities Phys Rev Lett **14**, 57–59 (1965)
3 Schoen, R, Yau, S -T : Positivity of the total mass of a general space-time Phys Rev Lett **43**, 1457–1459 (1979)
4 Schoen, R., Yau, S -T : Proof of the positive mass theorem. II Commun Math Phys **79**, 231–260 (1981)
5 Fischer-Colbrie, D, Schoen, R : The structure of complete stable minimal surfaces in 3-manifolds of non-negative scalar curvature Commun Pure Appl Math **33**, 199–211 (1980)