COMMUNICATIONS IN
ANALYSIS AND GEOMETRY
Volume 10, Number 5, 999-1016, 2002

Superharmonic functions in $\mathbf{R}^n$ and the Penrose Inequality in General Relativity

HUBERT L. BRAY AND KEVIN IGA

1. Introduction.

In this paper we prove a nonlinear property of superharmonic function in $\mathbf{R}^n$ which is closely related to the Penrose Inequality [1] in general relativity. The Penrose inequality is the statement that the total ADM mass of an asymptotically flat, space-like slice of a spacetime is at least the mass of the black holes which it contains, assuming nonnegative energy density everywhere in the spacetime. This statement is a generalization of the positive mass theorem, which states that the total ADM mass of a space-like slice of a spacetime is nonnegative, again assuming nonnegative energy density globally.

The Riemannian cases of the Penrose inequality and the positive mass theorem occur when the space-like slice (M^n, g), $n \geq 3$, is assumed to have zero second fundamental form in the (n+1)-dimensional spacetime. Then the assumption of nonnegative energy density implies that (M^n, g) has nonnegative scalar curvature, and apparent horizons of black holes correspond to outermost minimal surfaces in (M^n, g). The total ADM mass of the asymptotically flat Riemannian manifold (M^n, g) is then related to how quickly the manifold becomes flat at infinity.

After considering spherically symmetric manifolds, the next simplest special case to consider are manifolds which are conformal to the standard flat metric $(\mathbf{R}^n, \delta_{ij})$. Hence, let's assume that (M^n, g) is isometric to $(\mathbf{R}^n, u(x)^{4/(n-2)}\delta_{ij})$, $n \geq 3$, where $u(x)$ is positive and goes to a constant at infinity. Conveniently, the assumption of nonnegative scalar curvature is then equivalent to $u(x)$ being superharmonic in $(\mathbf{R}^n, \delta_{ij})$. Hence, it is natural to try to prove the positive mass theorem and the Penrose inequality, in this conformally flat case, using only known properties of superharmonic functions.

As it turns out, the n-dimensional positive mass theorem in this case follows from the maximum principle applied to superharmonic functions in $\mathbf{R}^n$. However, the story is not nearly as simple for the Penrose inequality, as we will see later. In fact, we are only able to treat the $n = 3$ case for the Penrose inequality in this special case so far. However, in this paper we are able to show that the 3-dimensional Penrose inequality (but with suboptimal constant) follows from a new nonlinear property of superharmonic function in $\mathbf{R}^3$. We also generalize this property to superharmonic functions in $\mathbf{R}^n$, $n > 3$, although the property does not imply the Penrose inequality in the conformally flat case in these dimensions, even with suboptimal constants.

The reason that the Penrose inequality is more difficult than the positive mass theorem in the above mentioned conformally flat case can be seen by the following definition, theorem, and conjecture.

Definition 1. Suppose that (M^n, g), $n \geq 3$, is isometric to $(\mathbf{R}^n, u(x)^{\frac{4}{n-2}}\delta_{ij})$, where $u(x)$ is positive and superharmonic in $(\mathbf{R}^n, \delta_{ij})$ and converges to a constant $a > 0$ at infinity. Suppose also that $u(x)$ is harmonic outside a compact set K, so that we may expand $u(x)$ using spherical harmonics to get

$$u(x) = a + \frac{b}{|x|^{n-2}} + \mathcal{O}(\frac{1}{|x|^{n-1}}). \tag{1.1}$$

Then we define the total mass of (M^n, g) to be

$$m = 2ab, \tag{1.2}$$

also known as the ADM mass of (M^n, g).

For the rest of the paper we will restrict our attention to the manifolds (M^n, g) described in the above definition. We can then state the positive mass theorem for such manifolds:

Theorem 1.1. *Suppose (M^n, g) is as described in definition 1. Then $m \geq 0$, and $m = 0$ if and only if $u(x) \equiv 1$.*

The above theorem can be proved using either the maximum principle or the divergence theorem, and we leave this as an exercise for the reader. This next conjecture (which has been proven in the case n=3 in much more general settings by Huisken and Ilmanen [2] and the first author [3]) is equivalent to the Penrose inequality in our conformally flat setting:

Conjecture 1.2. *Suppose (M^n, g) is as described in definition* 1 *and contains an outer-minimizing horizon Σ (see below) of area A. Then*

$$m \geq \frac{1}{2}\left(\frac{A}{\omega_{n-1}}\right)^{\frac{n-2}{n-1}}, \tag{1.3}$$

with equality if and only if $u(x) \equiv a + \frac{b}{|x-x_0|^{n-2}}$ outside Σ, for some $a, b > 0$ and for some $x_0 \in \mathbf{R}^3$.

In the above conjecture, ω_n is defined to be the volume of the unit n-sphere in $\mathbf{R}^{n+1}$. Also, let $\mathcal{S}$ be the set of all hypersurfaces in (M^n, g) which are smooth, compact boundaries of open sets. Then in the above conjecture, we define horizons to be hypersurfaces $\in \mathcal{S}$ with zero mean curvature and outer-minimizing hypersurfaces to be hypersurfaces which are not enclosed by hypersurfaces $\in \mathcal{S}$ of less area.

Even in dimension three where the above conjecture is known to be true, there does not exist a proof of this conjecture using only properties of superharmonic functions in $\mathbf{R}^3$. One reason the above conjecture is so challenging is that it is difficult to figure out how to use the fact that Σ is outer-minimizing. Also, without this assumption, the conjecture is not true.

In the next section, using only a new nonlinear property of superharmonic functions, we will prove a modified version of conjecture 1.2 for $n = 3$ (theorem 2.2), but with suboptimal constant in inequality 1.3 and, consequently, without the case of equality. We point out that this nonlinear property of superharmonic functions is also used in the first author's proof of the Riemannian Penrose inequality in [3]. Hence, the main result of this paper is really theorem 3.1 of section 3, with a very nice application of this result being theorem 2.2 of the next section.

2. Proof of a Special Case of the Penrose Conjecture Using Superharmonic Functions in $\mathbf{R}^3$.

As mentioned in the previous section, one of the difficulties is dealing with conjecture 1.2 is understanding how to use the fact that the surface Σ is outer-minimizing. Going back to the very general Riemannian Penrose inequality, it happens that it is possible to reflect the manifold through its outer-minimizing horizon to get a manifold with two asymptotically flat ends. (The details of this reflection procedure are carried out in section 6 of [3].) The benefit is that the Penrose inequality for the reflected manifold can be shown to be equivalent to the Penrose inequality for the original manifold.

Then translating this modified Penrose inequality to our setting, we get the following conjecture:

Conjecture 2.1. *Suppose (M^n, g) is as described in definition 1. Suppose also that $u(x)$ has a pole at $x = 0$, and that every surface $\Sigma \in \mathcal{S}$ which encloses the origin has area in (M^n, g) greater than or equal to A. Then*

$$m \geq \frac{1}{2}\left(\frac{A}{\omega_{n-1}}\right)^{\frac{n-2}{n-1}}, \tag{2.1}$$

with equality if and only if $u(x) \equiv a + \frac{b}{|x-x_0|^{n-2}}$ outside the outermost horizon of (M^n, g), for some $a, b > 0$ and for some $x_0 \in \mathbf{R}^3$.

As mentioned previously, the above conjecture is known to be true for $n = 3$, although an "elementary" proof using only facts about superharmonic functions is not known. However, we do have an elementary proof of the following theorem, which relies on theorem 3.1 from the next section.

Theorem 2.2. *Suppose (M^3, g) is as described in definition 1. Suppose also that $u(x)$ has a pole at $x = 0$, and that every surface $\Sigma \in \mathcal{S}$ which encloses the origin has area in (M^3, g) greater than or equal to A. Then*

$$m \geq \lambda A^{1/2}, \tag{2.2}$$

for some universal constant $\lambda > 0$.

Proof. Since $u(x)$ is a superharmonic function, we will bound the mass m from below by contructing a harmonic barrier function $\phi(x)$ which bounds $u(x)$ from below.

By the hypothesis of the theorem, the area of the sphere of radius r around the origin with respect to the metric g is at least A. Thus, we have

$$\int_{S_r(0)} u(x)^4 dA(x) \geq A \tag{2.3}$$

for all $r > 0$. By theorem 3.1 of the next section, we therefore can conclude that

$$u(x) \geq CA^{1/4}|x|^{-1/2} \tag{2.4}$$

for all $x \neq 0$, where $C = C(3,4)$ from theorem 3.1.

Now consider the sphere of radius R, where $R^{1/2} = \frac{CA^{1/4}}{2a}$, where we recall that the constants a and b are defined in definition 1. Thus,

$$u(x) \geq 2a \tag{2.5}$$

for $x \in S_R(0)$. Hence, if we let $\phi(x) = a + \beta/|x|$, where $\beta = aR$, it follows from applying the maximum principle to the difference of the superharmonic function $u(x)$ and the harmonic function $\phi(x)$ that $u(x) \geq \phi(x)$ for all x outside $B_R(x)$. Then by comparing the spherical harmonic expansions of $u(x)$ and $\phi(x)$ at infinity, we conclude that $b \geq \beta$. Hence,

$$m = 2ab \geq 2a\beta = 2a^2R = \frac{C^2}{2}A^{1/2}, \tag{2.6}$$

proving the theorem. □

One reason we are not able to prove conjecture 2.1 using the above technique is that we have not used all of the hypotheses of the conjecture. In fact, we have only used the fact that the area of all of the *spheres* centered around the origin have area in (M^3, g) at least A, which is of course a much weaker restriction than requiring this area bound for all surfaces enclosing the origin.

3. Superharmonic functions in $\mathbf{R}^n$.

In this section we prove theorem 3.1, which is at the heart of the proof of theorem 2.2. Theorem 3.1 is also of independent interest, since it is a new nonlinear property of superharmonic functions in $\mathbf{R}^n$.

Theorem 3.1. *Let $r_0 > 0$ be a fixed positive constant. Suppose that $u : \mathbf{R}^n \to [0, \infty]$ for $n \geq 3$ is a continuous superharmonic function with*

$$\int_{S_r(0)} u(x)^p \geq A \tag{3.1}$$

for all $r > r_0$, where $S_r(0)$ is the sphere of radius r centered at the origin, and $1 < p < (n-1)/(n-3)$ (if $n = 3$, then take $p > 1$). Then

$$u(\vec{x}) \geq CA^{1/p}|x|^{-(n-1)/p}, \tag{3.2}$$

for all x and for some positive constant $C(n,p)$.

Proof. Fix a particular radius $|\vec{x}| = r$ at which we verify

$$u(\vec{x}) \geq CA^{1/p}r^{-(n-1)/p}.$$

We first claim that without loss of generality, u is harmonic outside a compact set, and converges to zero as $|\vec{x}|$ goes to ∞. This is true because we can define (given any positive constants ε and K) the function

$$\tilde{u} = \min\left\{u + \varepsilon, K|\vec{x}|^{2-n}\right\}.$$

Fix any $\varepsilon > 0$. Because u is continuous, $u + \varepsilon$ attains a maximum on the closed ball of radius r around the origin. By choosing K appropriately, we can ensure Kr^{2-n} is larger than $u + \varepsilon$ on this ball. On the other hand, for $|\vec{x}|$ large enough, $\tilde{u} = K|\vec{x}|^{2-n}$ which is harmonic.

The theorem, applied to this function $\tilde{u}$, would then prove that $u(\vec{x})+\varepsilon \geq C$ for all $\varepsilon > 0$, and thus, $u(\vec{x}) \geq C$.

We can thus assume without loss of generality that u is harmonic outside a compact set, and converges to zero as $|\vec{x}|$ goes to ∞. The consequence of this is that $f = \Delta u$, which is defined as a distribution, has compact support, and $u = f(\vec{x}) * \frac{1}{\omega_n(n-2)}|\vec{x}|^{2-n}$. Since u was superharmonic, f is a positive measure.

We next mollify u so that it is smooth: let $\psi_j : \mathbf{R}^n \to \mathbf{R}$ be a sequence of spherically symmetric smooth functions supported in balls of radius $1/j$ around the origin, with $\int_{\mathbf{R}^n} \psi_j = 1$. Then $u_j = u * (\psi_j * \psi_j)$ is a smooth function.

Since a continuous function defined on a compact set is uniformly continuous (see baby Rudin, theorem 4.19), u is uniformly continuous over compact subsets of $\mathbf{R}^n$. Therefore, as j goes to infinity, u_j converges to u uniformly on compact sets.

Furthermore, if we let $f_j = f * \psi_j$, and $g_j = \frac{1}{\omega_n(n-2)}r^{2-n} * \psi_j$, then $u_j = f_j * g_j$, by associativity of convolutions.

Since u_j converges to u uniformly on compact sets, we know that for j large enough,

$$\int_{S_r(0)} u_j^p \geq A/2$$

for all $r > r_0$. Thus if we can prove the theorem for u_j smooth, then we will know that for all $\vec{x}$

$$u_j(\vec{x}) \geq 2^{-1/p}CA^{1/p}|\vec{x}|^{-(n-1)/p}. \tag{3.3}$$

Furthermore, by dividing u by $A^{1/p}$, we see that the statement of the theorem scales correctly, and hence it suffices to prove the theorem for $A = 1$. Thus, we need to prove the following:

Theorem 3.2. *Suppose that* $u : \mathbf{R}^n \to [0, \infty]$ *for* $n \geq 3$ *is a smooth superharmonic function*

$$u = f * g$$

with f *positive, continuous and of compact support, and* $g(\vec{z})$ *a spherically symmetric decreasing function of* $|\vec{z}|$*, and*

$$\int_{S_r(0)} u(\vec{x})^p \geq 1 \tag{3.4}$$

for $r \in [a, b]$*, where* $S_r(0)$ *is the sphere of radius* r *centered at the origin,* $[a, b]$ *is a closed interval in* $(0, \infty)$*, and* $1 < p < (n-1)/(n-3)$ *(if* $n = 3$*, then take* $p > 1$*). Then*

$$u(\vec{x}) \geq C \tag{3.5}$$

for all unit vectors $\vec{x}$*, for some positive constant* $C(n, p, a, b)$*.*

By the spherical symmetry of the problem, it suffices to prove that $u(-\vec{e}_1) \geq C$, where $\vec{e}_1$ is the basis vector in the x_1 direction.

The equation $u = f * g$ can be written

$$u(\vec{x}) = \int_{\mathbf{R}^n} f(\vec{y}) g(\vec{x} - \vec{y})\, d^n\vec{y}.$$

Let $\bar{f}(t) = \int_{S_t(0)} f(\vec{y}) d^{n-1}\vec{y}$ for $t \geq 0$ and let

$$\bar{u}(\vec{x}) = \int_0^\infty \bar{f}(t) g(\vec{x} - t\vec{e}_1)\, dt.$$

We view this procedure of moving from f to $\bar{f}$ as that of concentrating the source of u to the positive x_1 axis. Then as the following lemma will demonstrate, this will increase $\int_{S_r(0)} u^p$ but decrease its value at $-\vec{e}_1$, thus reducing us to the axially symmetric case.

If $\vec{v}$ is a unit vector in $\mathbf{R}^n$, then we define a reflection

$$R_{\vec{v}}(\vec{x}) = \vec{x} - 2\frac{\vec{v} \cdot \vec{x}}{|\vec{v}|^2}\vec{v}$$

as the reflection through the hyperplane perpendicular to $\vec{v}$. Define the corresponding "fold" map:

$$\rho_{\vec{v}}(\vec{x}) = \begin{cases} \vec{x}, & \text{if } \vec{x} \cdot \vec{v} \geq 0 \\ R_{\vec{v}}(\vec{x}), & \text{otherwise} \end{cases}$$

If $f(\vec{x})$ is a real-valued function on $\mathbf{R}^n$, we will denote by $f^{\wedge\vec{v}}$ (or $f^{\wedge}$ if $\vec{v}$ is understood), the "folded" function as follows:

$$f^{\wedge\vec{v}}(\vec{x}) = \begin{cases} f(\vec{x}) + f(R_{\vec{v}}(\vec{x})), & \vec{x}\cdot\vec{v} \geq 0 \\ 0, & \vec{x}\cdot\vec{v} < 0 \end{cases}$$

For convenience we will define the following functions:

$$f^{+}(\vec{x}) = \begin{cases} f(\vec{x}), & \vec{x}\cdot\vec{v} \geq 0 \\ 0, & \vec{x}\cdot\vec{v} < 0 \end{cases}$$

$$f^{-}(\vec{x}) = \begin{cases} f(\vec{x}), & \vec{x}\cdot\vec{v} < 0 \\ 0, & \vec{x}\cdot\vec{v} \geq 0 \end{cases}$$

so that

$$f(\vec{x}) = f^{+}(\vec{x}) + f^{-}(\vec{x})$$

and

$$f^{\wedge}(\vec{x}) = f^{+}(\vec{x}) + f^{-}(R_v(\vec{x})).$$

Lemma 3.3. *Let $\phi(\vec{x})$ be a bounded continuous real-valued function on $\mathbf{R}^n$, spherically symmetric around the origin, and decreasing as a function of $|\vec{x}|$. For any $f(\vec{x})$ a non-negative integrable function on $\mathbf{R}^n$ with compact support, define $u_f = \phi * f$ be the convolution of ϕ with f, and let*

$$A_f = \int_{S^{n-1}} |u_f|^p \, d^{n-1}\omega$$

where $p \geq 1$. Then

$$A_f \leq A_{f^{\wedge}}.$$

for any unit vector $\vec{v} \in S^{n-1}$.

Proof. Let $\vec{x} \in \mathbf{R}^n$ be an arbitrary point with $\vec{x}\cdot\vec{v} \geq 0$, and let $\vec{x}' = R_{\vec{v}}(\vec{x})$. Now $u_{f^+}(\vec{x}') \leq u_{f^+}(\vec{x})$ since ϕ is a decreasing function of distance, and for any $\vec{y}$ with $\vec{x}\cdot\vec{y} \geq 0$, we have that $\vec{x}$ is at least as close to $\vec{y}$ as $\vec{x}'$ is.

Similarly, $u_{f^-}(\vec{x}) \leq u_{f^-}(\vec{x}')$, since for any $\vec{y}$ with $\vec{x}\cdot\vec{y} < 0$, we have that $\vec{x}$ is at least as far from $\vec{y}$ as $\vec{x}'$ is.

If we let $a = u_{f^+}(\vec{x}')$, $b = u_{f^+}(\vec{x})$, $c = u_{f^-}(\vec{x})$, $d = u_{f^-}(\vec{x}')$, then we have so far concluded that $a \leq b$ and $c \leq d$. We will next prove that when $a \leq b$ and $c \leq d$,

$$(a+d)^p + (b+c)^p \leq (a+c)^p + (b+d)^p.$$

First, without loss of generality, assume $b - a \geq d - c$. Then note that $[a + c, b + d]$ and $[a + d, b + c]$ are nonempty intervals centered on the same average point, but $[a+d, b+c] \subset [a+c, b+d]$, as can be verified by comparing the lengths.

By convexity of the function $g(x) = x^p$ for $p \geq 1$, we have that the graph of x^p on $[a + c, b + d]$ lies below the secant line on the same interval. Since $[a + d, b + c] \subset [a + c, b + d]$, the secant line for $[a + d, b + c]$ lies below the secant line on $[a+c, b+d]$. Thus, the midpoints of these secant line segments compare as follows:

$$\frac{(a+d)^p + (b+c)^p}{2} \leq \frac{(a+c)^p + (b+d)^p}{2}$$

which gives us

$$(a+d)^p + (b+c)^p \leq (a+c)^p + (b+d)^p$$

as claimed above.

Applied to our case, we have

$$\begin{aligned}&\left(u_{f^+}(\vec{x}') + u_{f^-}(\vec{x}')\right)\right)^p + \left(u_{f^+}(\vec{x}) + u_{f^-}(\vec{x})\right)\right)^p \leq \\ &\leq \left(u_{f^+}(\vec{x}') + u_{f^-}(\vec{x})\right)\right)^p + \left(u_{f^+}(\vec{x}) + u_{f^-}(\vec{x}')\right)\right)^p.\end{aligned}$$

Since $f = f^+ + f^-$ and u depends linearly on f, $u = u_{f^+} + u_{f^-}$. Similarly, from $f^\wedge = f^+ + f^- \circ R_{\vec{v}}$, we get $u = u_{f^+} + u_{f^-} \circ R_{\vec{v}}$. This gives us

$$(u_f(\vec{x}'))^p + (u_f(\vec{x}))^p \leq (u_{f^\wedge}(\vec{x}'))^p + (u_{f^\wedge}(\vec{x}))^p.$$

This holds whenever $\vec{x} \cdot \vec{v} \geq 0$, but by symmetry under interchanging $\vec{x}$ and $\vec{x}' = R_{\vec{v}}(\vec{x})$ in the formula, it also holds for all $\vec{x} \in S^{n-1}$.

Integrating this formula over all $\vec{x} \in S^{n-1}$ we get

$$\begin{aligned}&\int_{S^{n-1}} (u_f(\vec{x}'))^p \, d^{n-1}\vec{x} + \int_{S^{n-1}} (u_f(\vec{x}))^p \, d^{n-1}\vec{x} \leq \\ &\leq \int_{S^{n-1}} (u_{f^\wedge}(\vec{x}'))^p \, d^{n-1}\vec{x} + \int_{S^{n-1}} (u_{f^\wedge}(\vec{x}))^p \, d^{n-1}\vec{x}.\end{aligned}$$

By changing variables from $\vec{x}$ to $\vec{x}'$ in the first and third integrals, we obtain the second and fourth integrals, respectively, and, dividing by two, we have

$$\int_{S^{n-1}} (u_f(\vec{x}))^p \, d^{n-1}\vec{x} \leq \int_{S^{n-1}} (u_{f^\wedge}(\vec{x}))^p \, d^{n-1}\vec{x}.$$

□

Lemma 3.4. *Let $\phi(\vec{x})$, $f(\vec{x})$, and u_f be as in the previous lemma. As before, define*

$$A_f = \int_{S^{n-1}} |u_f|^p \, d^{n-1}\omega$$

where $p \geq 1$. Similarly, define $\bar{f}$ as above:

$$\bar{f}(t) = \int_{S_t(0)} f(\vec{y}) d^{n-1}\vec{y}$$

and

$$\bar{u}_f(\vec{x}) = \int_0^\infty \frac{\bar{f}(t)}{\phi(\vec{x} - t\vec{e}_1)} \, dt$$

$$A_{\bar{f}} = \int_{S^{n-1}} |\bar{u}_f|^p \, d^{n-1}\omega.$$

Then

$$A_f \leq A_{\bar{f}}.$$

Proof. We will generate a sequence f_n, where $f_0 = f$, and

$$A_{f_n} \leq A_{f_{n+1}}.$$

This is obtained by choosing a sequence of unit vectors $\vec{v}_n$, and setting $f_n = (f_{n-1})^{\wedge \vec{v}_{n-1}}$, applying the previous lemma.

The $\vec{v}_n$ will be chosen to make the domain of f concentrate on the positive x_1 axis. In order to keep track of this, we define inductively the maximal possible domain of f as follows:

Let $g_0(\vec{x})$ be the constant function 1, and recursively define $g_n = (g_{n-1})^{\wedge \vec{v}_{n-1}}$. Then for each n, the support of f_n will be a subset of the support of g_n. Let $W_n = supp(g_n) \cap S^{n-1}$. The $\vec{v}_n$ should be chosen so that $W_n \supset W_{n+1}$, which means that the reflection $\rho_{\vec{v}_n}$ should send W_n into itself. We should choose the $\vec{v}_n$ satisfying this restraint, in order to get a nested sequence of W_n that converges to the point $\vec{e}_1$.

Let α be an angle between 0 and $\pi/2$. Consider the set of vectors of the form

$$\vec{u}_{\alpha,j,\pm} = \sin\alpha\, \vec{e}_1 \pm \cos\alpha\, \vec{e}_j$$

where j ranges over all the basis vectors that are not $\vec{e}_1$. Let C_α be the "hypercube" set centered at $\vec{e}_1$:

$$C_\alpha = \left\{ \vec{x} \in S^{n-1} | \vec{x} \cdot \vec{u}_{\alpha,j,\pm} \geq 0 \text{ for all } j \text{ and choices of } \pm \right\}$$

We take $\vec{v}_0 = \vec{e}_1$, and thus W_1 is the right hemisphere with $x_1 \geq 0$. We then have that W_1 is the hypercube set $C_{\pi/2}$ centered at $\vec{e}_1$.

Inductively, whenever W_k lies in the hypercube $C_{2\alpha}$ centered at $\vec{e}_1$, we will choose $\vec{v}_{k+1}, ..., \vec{v}_{k+2n-2}$ to be the $2n-2$ vectors $\vec{u}_{\alpha,j,\pm}$. We will then prove that W_{k+2n-2} lies in the hypercube C_α centered at $\vec{e}_1$.

It is clear that each fold map will make sure $\vec{x}\cdot\vec{u}_{\alpha,j,\pm} \geq 0$ for the particular choice of α, j, and $\pm$. But we need to show that we are not at the same time making $\vec{x}\cdot\vec{u}_{\beta,j',\pm} < 0$ for some previous choice of β, j', and $\pm$. We will in fact show that the sequence of W_m's is a nested sequence:

$$W_1 \supset W_2 \supset \cdots .$$

To verify this, suppose some $\vec{v}_{k+i}$ sent a vector $\vec{x}$ in some W_{k+i} to something outside W_{k+i}. Since W_{k+i} was formed by "folding" using the $\vec{v}$'s, we know that for some $\vec{v}_m$, we have that $\vec{x}\cdot\vec{v}_m \geq 0$ but $\rho_{\vec{v}_{k+i}}(\vec{x})\cdot\vec{v}_m < 0$. Furthermore, since $\vec{x}$ was in W_{k+i} to begin with, $\vec{x}\cdot\vec{v}_{k+i} < 0$.

We have that $\vec{v}_{k+i} = \vec{u}_{\alpha,j_1,\pm}$ for some j_1 and choice of $\pm$, and that $\vec{v}_m = \vec{u}_{\beta,j_2,\pm}$ for some other similar choices. If $j_1 \neq j_2$, then $\vec{v}_{k+i}\cdot\vec{v}_m \geq 0$, and therefore

$$\begin{aligned}
\rho_{\vec{v}_{k+i}}(\vec{x})\cdot\vec{v}_m &= R_{\vec{v}_{k+i}}(\vec{x})\cdot\vec{v}_m \\
&= (\vec{x} - 2\vec{v}_{k+i}\cdot\vec{x}\vec{v}_{k+i})\cdot\vec{v}_m \\
&= \vec{x}\cdot\vec{v}_m - 2(\vec{v}_{k+i}\cdot\vec{x})(\vec{v}_{k+i}\cdot\vec{v}_m) \\
&\geq -2(\vec{v}_{k+i}\cdot\vec{x})(\vec{v}_{k+i}\cdot\vec{v}_m) \\
&\geq 0
\end{aligned}$$

which gives the contradiction. We therefore consider $j_1 = j_2$. If the choices of $\pm$ for $\vec{v}_{k+i}$ and $\vec{v}_m$ are the same, then again $\vec{v}_{k+i}\cdot\vec{v}_m \geq 0$ and by the same argument arrive at the contradiction. So we know that the choices of $\pm$ are different, and without loss of generality

$$\begin{aligned}
\vec{v}_{k+i} &= \vec{u}_{\alpha,j,+} = \sin\alpha\vec{e}_1 + \cos\alpha\vec{e}_j \\
\vec{v}_m &= \vec{u}_{\beta,j,-} = \sin\beta\vec{e}_1 - \cos\beta\vec{e}_j.
\end{aligned}$$

If we write $\vec{x} = a\vec{e}_1 + b\vec{e}_j + \vec{y}$, then by induction (that all the W_k are in W_1) we know that $a \geq 0$. Furthermore, since by assumption $\vec{x}\cdot\vec{v}_{k+i} < 0$, we have that $b < 0$. Also since $\vec{x} \in W_k \subset C_{2\alpha}$, we know that $a\sin(2\alpha) - b\cos(2\alpha) \geq 0$, which together with $a^2 + b^2 + |\vec{y}|^2 = 1$, implies that $b \geq -\sin(2\alpha)$ and

$a \geq \cos(2\alpha)$. Then since $\vec{x} \cdot \vec{v}_{k+i} < 0$,

$$\begin{aligned}
\rho_{\vec{v}_{k+i}}(\vec{x}) \cdot \vec{e}_j &= (\vec{x} - 2\vec{v}_{k+i} \cdot \vec{x}\vec{v}_{k+i}) \cdot \vec{e}_j \\
&= b - 2(a \sin\alpha + b\cos\alpha)\cos\alpha \\
&\leq -\sin(2\alpha) - 2(\cos(2\alpha)\sin\alpha - \sin(2\alpha)\cos\alpha)\cos\alpha \\
&= 0
\end{aligned}$$

and similarly

$$\begin{aligned}
\rho_{\vec{v}_{k+i}}(\vec{x}) \cdot \vec{e}_1 &= (\vec{x} - 2\vec{v}_{k+i} \cdot \vec{x}\vec{v}_{k+i}) \cdot \vec{e}_1 \\
&= a - 2(\vec{v}_{k+i} \cdot \vec{x}) \sin\alpha \\
&\geq a \geq 0.
\end{aligned}$$

Thus

$$\rho_{\vec{v}_{k+i}}(\vec{x}) \cdot \vec{v}_m = \rho_{\vec{v}_{k+i}}(\vec{x}) \cdot (\sin(\beta)\,\vec{e}_1 - \cos(\beta)\,\vec{e}_j) \geq 0,$$

which is, again, a contradiction.

In this way, we see that we have a nested sequence $W_k \supset W_{k+1}$, each of which is contained in the hypercubes C_α centered at $\vec{e}_1$, with α decreasing to zero, so that $\bigcap W_k = \{\vec{e}_1\}$, and f_n converges in measure to a delta function concentrated on the positive x_1 axis.

By the previous lemma, we have that

$$A_{f_n} \leq A_{f_{n+1}}.$$

By the uniform convergence of u_n to $\bar{u}$, we have that A_{f_n} converges to $A_{\bar{f}}$. □

Thus, if $\int_{S_r(0)} u^p$ is greater than or equal to one, then so is $\int_{S_r(0)} \bar{u}^p$.

Also, note that $u(-\vec{e}_1) \geq \bar{u}(-\vec{e}_1)$, since of all the points on $S_r(0)$, the point farthest away from $-\vec{e}_1$ is $r\vec{e}_1$.

We have thus reduced our situation to the case where the source function f is supported on the positive x axis, that is, using $\bar{f}$. So we will have to show $\bar{u}(-\vec{e}_1) \geq C$.

To accomplish this, we interpret our inequality

$$\int_{S_r(0)} \bar{u}(x)^p \, d^{n-1}x \geq 1$$

which holds for $1 \leq r \leq 2$. Using the notation σ_n for the area of an n-dimensional sphere of radius 1, we have:

$$\begin{aligned} 1 &\leq \int_{S_r(0)} \bar{u}(x)^p \, d^{n-1}x \\ &= \int_0^\pi \sigma_{n-2}(r\sin\theta)^{n-2} r \, d\theta \, (\bar{u}(r\cos\theta e_1 + r\sin\theta e_2))^p \\ &= \int_0^\pi \sigma_{n-2} r^{n-1} \sin^{n-2}\theta \, d\theta \left(\int_0^\infty \frac{\bar{f}(t)\, dt}{|(r\cos\theta - t)e_1 + r\sin\theta e_2|^{n-2}} \right)^p \\ &= \sigma_{n-2} r^{n-1} \int_0^\pi \sin^{-1+\varepsilon}\theta \, d\theta \left(\int_0^\infty \bar{f}(t) Q_{t,r}(\theta)\, dt \right)^p \end{aligned}$$

where

$$Q_{t,r}(\theta) = \frac{\sin^{\frac{n-2+1-\varepsilon}{p}}\theta}{(t^2 + r^2 - 2tr\cos\theta)^{(n-2)/2}},$$

and $\varepsilon > 0$ is any positive constant.

Let

$$S_r(t) = \sup_{0 \leq \theta \leq \pi} Q_{t,r}(\theta)$$

Then we have that

$$1 \leq \sigma_{n-2} r^{n-1} \int_0^\pi d\theta \left[\int_0^\infty \bar{f}(t) S_r(t)\, dt \right]^p$$

so that

$$\int_0^\infty \bar{f}(t) S_r(t)\, dt \geq \left(\pi \sigma_{n-2} r^{n-1}\right)^{-1/p}$$

for $a \leq r \leq b$.

Furthermore, since $k^{n-2} S_{kr}(kt) = S_r(t)$, setting $k = 1/r$ gives

$$S_r(t) = \frac{1}{r^{n-2}} S_1\left(\frac{t}{r}\right).$$

Hence,

$$\int_0^\infty \bar{f}(t) \frac{1}{r^{n-2}} S_1\left(\frac{t}{r}\right) dt \geq \left(\pi \sigma_{n-2} r^{n-1}\right)^{-1/p}$$

so that

$$\int_0^\infty \bar{f}(t) \frac{1}{t} \frac{t}{r^2} S_1\left(\frac{t}{r}\right) dt \geq (\pi\sigma_{n-2})^{-1/p} r^{(n-4)-(n-1)/p}$$

Now integrate with respect to r from a to b:

$$\int_0^\infty \bar{f}(t)\, dt \frac{1}{t} \int_a^b \frac{t}{r^2} S_1\left(\frac{t}{r}\right) dr \geq \int_a^b (\pi\sigma_{n-2})^{-1/p} r^{(n-4)-(n-1)/p}\, dr.$$

Let $y = t/r$, and substitute for r. Then

$$\begin{aligned}
&\int_0^\infty \bar{f}(t)\,dt\,\frac{1}{t}\int_{t/b}^{t/a} S_1(y)\,dy \geq \\
&\geq (\pi\sigma_{n-2})^{-1/p}\left(\frac{1}{(n-3)-(n-1)/p}\right) r^{(n-3)-(n-1)/p}\Big|_a^b \\
&= (\pi\sigma_{n-2})^{-1/p}\,\frac{1}{(n-3)-(n-1)/p}\left(b^{(n-3)-(n-1)/p} - a^{(n-3)-(n-1)/p}\right).
\end{aligned}$$

The possibility that $(n-3)-(n-1)/p = 0$ is excluded since $p < (n-1)/(n-3)$.

Define

$$k = \sup_{0<t<\infty} \frac{t+1}{t}\int_{t/b}^{t/a} S_1(y)\,dy.$$

We will next prove that $k < \infty$. Then our theorem will be proved since

$$\begin{aligned}
&u(-\vec{e}_1) \geq \\
&\geq \bar{u}(-\vec{e}_1) = \int_0^\infty \frac{\bar{f}(t)}{t+1}\,dt \\
&\geq \frac{1}{k}\int_0^\infty \frac{\bar{f}(t)}{t+1}\,dt\,\frac{t+1}{t}\int_{t/b}^{t/a} S_1(y)\,dy \\
&\geq \frac{1}{k}(\pi\sigma_{n-2})^{-1/p}\,\frac{1}{(n-3)-(n-1)/p}\left(b^{(n-3)-(n-1)/p} - a^{(n-3)-(n-1)/p}\right) \\
&= C.
\end{aligned}$$

Lemma 3.5. *The constant* $k = \sup_{0\leq t\leq\infty} \frac{t+1}{t}\int_{t/b}^{t/a} S_1(y)\,dy$ *is finite.*

Proof. We will bound $S_1(y)$ from above. Recall that

$$S_1(t) = \sup_{0\leq\theta\leq\pi} Q_{t,1}(\theta) = \sup_{0\leq\theta\leq\pi} \frac{(\sin\theta)^{(n-2+1-\varepsilon)/p}}{(t^2+1-2t\cos\theta)^{(n-2)/2}}.$$

Our first estimate involves the fact that by $\sin\theta \leq 1$ and $\cos\theta \leq 1$, we have

$$S_1(t) \leq \frac{1}{|t-1|^{n-2}}. \tag{3.6}$$

Using this estimate, we see that

$$\frac{t+1}{t}\int_{t/b}^{t/a} S_1(y)\,dy$$

is continuous in t in $\{0 < t < a\} \cup \{t > b\}$, so that we will have a bound on this function of t as long as we can control it when t is between a and b, when t goes to 0, and when t goes to ∞.

For t going to ∞, we use the above estimate (3.6) to compute the following supremum:

$$\sup_{2b \le t < \infty} \frac{t+1}{t} \int_{t/b}^{t/a} S_1(y)\, dy$$

This is done in a straighforward way when $n \ge 3$:

$$\begin{aligned}
\sup_{2b \le t < \infty} \frac{t+1}{t} \int_{t/b}^{t/a} S_1(y)\, dy &\le \sup_{2b \le t < \infty} \frac{t+1}{t} \int_{t/b}^{t/a} \frac{1}{|y-1|^{n-2}}\, dy \\
&\le \sup_{2b \le t < \infty} \frac{t+1}{t} \int_{t/b}^{t/a} \frac{1}{|y-1|}\, dy \\
&= \sup_{2b \le t < \infty} \frac{t+1}{t} \log \left| \frac{t/a - 1}{t/b - 1} \right| \\
&= \sup_{2b \le t < \infty} \frac{t+1}{t} \log \left(1 + \frac{t/a - t/b}{t/b - 1} \right) \\
&\le \sup_{2b \le t < \infty} \frac{t+1}{t} \frac{t(1/a - 1/b)}{t/b - 1} \\
&= \sup_{2b \le t < \infty} \frac{t+1}{t/b - 1} (1/a - 1/b) \\
&= (2b+1)(1/a - 1/b)
\end{aligned}$$

Now consider

$$\lim_{t \to 0} \frac{t+1}{t} \int_{t/b}^{t/a} S_1(y)\, dy.$$

The same estimate (3.6) gives us, with L'hôpital's rule,

$$\begin{aligned}
\lim_{t \to 0} \frac{t+1}{t} \int_{t/b}^{t/a} S_1(y)\, dy &\le \lim_{t \to 0} \frac{t+1}{t} \int_{t/b}^{t/a} \frac{1}{|y-1|^{n-2}}\, dy \\
&= \lim_{t \to 0} \frac{1}{1} \frac{d}{dt} \int_{t/b}^{t/a} \frac{1}{|y-1|^{n-2}}\, dy \\
&= \lim_{t \to 0} \frac{1}{a} \frac{1}{|t/a - 1|^{n-2}} - \frac{1}{b} \frac{1}{|t/b - 1|^{n-2}} \\
&= \lim_{t \to 0} 1/a - 1/b = 1/a - 1/b
\end{aligned}$$

Thus, the function of which we are taking the supremum is bounded as $t \to \infty$ and as $t \to 0$.

To estimate
$$\frac{t+1}{t}\int_{t/b}^{t/a} S_1(y)\,dy$$
in the interval $(a-\varepsilon, b+\varepsilon)$, for some ε, we will need a better estimate than (3.6). Now
$$\begin{aligned} Q_{t,1}(\theta)^2 &= \frac{(1-\cos^2\theta)^{(n-2+1-\varepsilon)/p}}{(t^2+1-2t+2t(1-\cos\theta))^{n-2}} \\ &= \frac{[(1-\cos\theta)(1+\cos\theta)]^{(n-1-\varepsilon)/p}}{((t-1)^2+2t(1-\cos\theta))^{n-2}} \end{aligned}$$
Hence
$$S_1(t)^2 \le \sup_{0\le z\le 2} \frac{(2z)^{(n-1-\varepsilon)/p}}{((t-1)^2+2tz)^{n-2}}.$$
If we define
$$h(z) = \frac{(2z)^{(n-1-\varepsilon)/p}}{((t-1)^2+2tz)^{n-2}}$$
then we wish to find the maximum value of $h(z)$.

First, $h(z)$ is defined for all $z \ge 0$, unless the denominator is zero. This can only happen if $t = 1$, in which case $h(z) = (2z)^{(n-1-\varepsilon)/p-(n-2)}$, and we see that $h(z)$ is still defined and smooth (except at $z = 0$). In this circumstance $h(z)$ grows without bound as z approaches zero. But we can ignore $S_1(1)$ since $\{1\}$ is only a single point and is thus of measure zero in the integral.

So suppose $t \ne 1$. Since $h(z)$ is a positive differentiable function that goes to 0 as z goes to 0 or to ∞, we see that any maximum occurs at a critical point. To find critical points, we solve:
$$\begin{aligned} 0 = \frac{d}{dz}\log h(z) &= \frac{(n-1-\varepsilon)}{p}\frac{1}{z} - (n-2)\frac{2t}{(t-1)^2+2tz} \\ (n-1-\varepsilon)\left((t-1)^2+(2t)z\right) &= (n-2)(2t)pz \\ (n-1-\varepsilon)(t-1)^2 &= 2t((n-2)p-(n-1-\varepsilon))z \\ z_{CRIT} &= \frac{(n-1-\varepsilon)(t-1)^2}{2t((n-2)p-(n-1-\varepsilon))} \end{aligned}$$
When $p > \frac{n-1-\varepsilon}{n-2}$, we always have a unique positive critical point. Otherwise, there is no critical point, and hence h takes on its maximum value on the boundary. We consider these two cases:

Case I: $p \le \frac{n-1-\varepsilon}{n-2}$:

Under this circumstance, $h(z)$ has no critical points, and so attains its maximum at $z = 2$.

$$\begin{aligned}\sup h(z) &= \frac{4^{(n-1-\varepsilon)/p}}{((t-1)^2+4t)^{n-2}} \\ &= 4^{(n-1-\varepsilon)/p}(t+1)^{-2(n-2)}\end{aligned}$$

Thus $S_1(t)$ is continuous in $0 < t < \infty$, and

$$\frac{t+1}{t}\int_{t/b}^{t/a} S_1(y)\,dy$$

is continuous in t. Hence it is bounded in any compact set, and since it is bounded near $t = 0$ and $t \to \infty$, we have that this function is bounded, and k is finite.

Case II: $p > \frac{n-1-\varepsilon}{n-2}$:

Then as discussed above, $h(z)$ must attain its maximum at this unique critical point:

$$\begin{aligned}\sup_{0<z<\infty} h(z) &= h(z_{CRIT}) \\ &= \frac{\left(\frac{(n-1-\varepsilon)(t-1)^2}{t((n-2)p-(n-1-\varepsilon))}\right)^{(n-1-\varepsilon)/p}}{\left((t-1)^2+\frac{(n-1-\varepsilon)(t-1)^2}{(n-2)p-(n-1-\varepsilon)}\right)^{n-2}} \\ &= t^{-(n-1-\varepsilon)/p}|t-1|^{\frac{2(n-1-\varepsilon)}{p-2(n-2)}}\frac{\left(\frac{n-1-\varepsilon}{(n-2)p-(n-1-\varepsilon)}\right)^{(n-1-\varepsilon)/p}}{\left(1+\frac{n-1-\varepsilon}{(n-2)p-(n-1-\varepsilon)}\right)^{n-2}}\end{aligned}$$

Hence we have improved the estimate (3.6) to the following:

$$S_1(t) \leq \text{const. } |t-1|^{(n-1-\varepsilon)/p-(n-2)}t^{-(n-1-\varepsilon)/2p} \tag{3.7}$$

Now recall that we are trying to show k is finite, which is to find a universal bound for

$$\frac{t+1}{t}\int_{t/b}^{t/a} S_1(y)\,dy.$$

We have already investigated the behavior of this function as t approaches zero and as t grows without bound, so we are left with the circumstance when the integral involves $y = 1$, because that is where (3.7) is unbounded.

The above bound shows that

$$S_1(t) \leq \text{bounded parts } |t-1|^{(n-1-\varepsilon)/p-(n-2)}$$

and an integral over t involving $t = 1$ will be bounded whenever the exponent $(n-1-\varepsilon)/p - (n-2)$ is greater than -1; i.e. whenever

$$p < (n-1-\varepsilon)/(n-3).$$

But we know that if $n > 3$, then $p < (n-1)/(n-3)$, so we can achieve this simply by choosing ε small enough (when $n = 3$, $(n-1-\varepsilon)/p - (n-2)$ will automatically be greater than -1, for ε small).

Therefore the integral is bounded, and the resulting function

$$\frac{t+1}{t}\int_{t/b}^{t/a} S_1(y)\,dy$$

is a continuous function of t. Thus, there is a finite supremum, and k is finite. □

References.

[1] R. Penrose, "Naked Singularities," *Ann. New York Acad. Sci.* **224** (1973) 125-134.

[2] G. Huisken, T. Ilmanen, "The Inverse Mean Curvature Flow and the Riemannian Penrose Inequality," *Jour. Diff. Geom.* **59** (2001), issue 3, 353-437.

[3] H. L. Bray, "Proof of the Riemannian Penrose Inequality Using the Positive Mass Theorem," *Jour. Diff. Geom.* **59** (2001), issue 2, 177-267.

RECEIVED JANUARY 19, 2001.