An overview on the Penrose inequality

Marc Mars
Facultad de Ciencias, Universidad de Salamanca, Plaza de la Merced s/n, 37008 Salamanca, Spain
E-mail: marc@usal.es

Abstract. A summary of the main results and ideas on the Penrose inequality is given. The original heuristic argument by Penrose leading to the inequality is reviewed and the main ideas behind the known proofs of particular cases are discussed. Some of the recent approaches to treat the general Penrose inequality, including the use of the Jang equation, the Bergqvist mass and the uniformly expanding flows are also described.

1. Introduction
The Penrose inequality is a conjecture which states roughly speaking that a black hole of given area $A$ contributes with a minimum amount to the total mass of the spacetime. In units where $c = G = 1$ this minimum mass is given by $\sqrt{A/16\pi}$. It can therefore be regarded as an strengthening of the positive mass theorem which states that the total mass of an asymptotically flat spacetime is non-negative and zero if and only if the spacetime is Minkowski space (see Sect. 2 for definitions). The Penrose inequality conjecture also has a rigidity part stating that if the total mass equals $\sqrt{A/16\pi}$ then the spacetime is Schwarzschild.

The Penrose inequality appears in two different versions. One which may be called “global” where strong global assumptions on the spacetime are made and another “local” one where the inequality is formulated directly in terms of an initial data set, therefore local in time.

In the global version, first proposed by Penrose [1], the spacetime is asymptotically flat in the sense that it admits a complete future null infinity $\mathcal{I}^+$, the spacetime is strongly asymptotically predictable and $J^{-}(\mathcal{I}^+) \neq \mathcal{M}$ (see e.g. [2] for definitions). The event horizon $\mathcal{H}$ is the boundary of $J^{-}(\mathcal{I}^+)$. This is a null Lipschitz hypersurface. Assume moreover that $\mathcal{H}$ is differentiable. A cut $S$ of $\mathcal{H}$ is a smooth, closed (i.e. compact and without boundary) surface embedded in $\mathcal{H}$ which intersects each null generator of $\mathcal{H}$ at most on one point. Let $M_{ADM}$ be the total mass of the spacetime $(\mathcal{M}, g)$. The “global Penrose inequality” states

$$M_{ADM} \geq \sqrt{\frac{|S|}{16\pi}},$$

where $|S|$ is the area of $S$. For general event horizons, non-necessarily smooth, inequality (1) still makes sense if $S$ is the intersection of $\mathcal{H}$ with an achronal, spacelike, embedded, smooth hypersurface provided the area is interpreted as the 2-dimensional Hausdorff measure. The Hausdorff measurability of $S$ is demonstrated in [3].

In this global setting, inequality (1) is strongly supported by the following argument, due to Penrose [1]. Let $S_1$ and $S_2$ be two cross sections of $\mathcal{H}$ with $S_2$ lying to the future of $S_1$. 
Then the area theorem for black holes states $|S_1| \leq |S_2|$. This was proven by Hawking [4] when $\mathcal{H}$ is smooth and generalized to arbitrary event horizons $\mathcal{H}$ in [3] assuming also much milder asymptotic conditions. Along its dynamical evolution, the black hole will emit gravitational waves and accrete matter. From physical principles, the spacetime is expected to settle down to some equilibrium configuration. Thus, at late times the spacetime is expected to approach an asymptotically stationary black hole configuration. Assuming also that all the matter fields are swallowed by the black hole in the process (an external electromagnetic field would not alter the conclusions), the black hole uniqueness theorem (see e.g. [5]) imply that the spacetime must approach the Kerr metric (modulo the technical assumptions that still remain open for these theorems, see e.g. [6]). In Kerr, the area of the horizon $A_{\text{Kerr}}$ does not depend on the cut (as for any Killing horizon) and the inequality

$$A_{\text{Kerr}} \leq 16\pi M_{\text{Kerr}}^2,$$

where $M_{\text{Kerr}}$ is the ADM mass of the spacetime, holds by direct computation. $M_{\text{Kerr}}$ corresponds to the asymptotic value of the Bondi mass of the evolving black hole. Gravitational waves can carry only positive energy and therefore the Bondi mass is monotonically non-increasing in any asymptotically flat spacetime [7]. The Bondi mass approaches the ADM mass at early retarded times ([8] and references therein). Putting together all this chain of inequalities, we conclude (1). This argument is obviously not a theorem since it uses results which have been shown with various levels of rigour and moreover relies on the completely heuristic argument that the spacetime should settle down to a stationary state. Nevertheless, the argument gives strong support for (1).

Inequality (1) is global in the sense that the area corresponds to a cut of the event horizon and locating the event horizon requires knowledge of the whole future history of the spacetime. Penrose’s original idea was to translate this global inequality into an inequality which could be stated (and verified) in terms of quantities defined independently of the evolution of the spacetime, i.e. directly in terms of the initial data of the spacetime. The total mass is a functional on the initial data, but the area of the event horizon must be substituted by some alternative local property.

The singularity theorems of Penrose [9], Hawking [10] and others (see [11] for a review) state that initial data sets fulfilling a suitable “strong gravitational field” condition must develop, provided the matter satisfied the strong energy condition, into a maximal globally hyperbolic spacetime that contains a singularity (i.e. an incomplete inextendible causal geodesic). The condition of strong gravitational field may be of different types. Two cases specially relevant in the asymptotically flat context include (i) the topology of the initial data set is non-simply connected [12] or (ii) the initial data set admits a trapped surface (see below for definitions). Besides the existence of singularity in the Cauchy development, virtually nothing is known in general on the type of singularities that may occur. More specifically it is not known whether the future development admits a complete $\mathcal{I}$ and therefore defines a black hole spacetime. Physical arguments on the lack of predictability that would occur if the singularities are not hidden from infinity via an event horizon make it reasonable that initial data sets generating singular spacetimes in fact generate black hole spacetimes. This is known under the name of weak cosmic censorship hypothesis and was first proposed by Penrose [13]. The validity of this hypothesis remains an important open problem, see [14] for a relatively recent review.

If weak cosmic censorship holds, then an initial data set containing a future trapped surface $S$ must develop into a black hole spacetime and hence have an event horizon. A general result on black hole spacetimes [4, 15] is that any future trapped surface does not enter into the domain of outer communications of the black hole, defined as the causal past of $\mathcal{I}^+$. Therefore, the intersection of the event horizon and the initial data set defines a spacelike surface $\mathcal{H}_\Sigma$ that separates $S$ from the asymptotic region. The location of $\mathcal{H}_\Sigma$ cannot be determined from the
initial data alone. Moreover, it need not be true in general that the area of $S$ is smaller that the area of $\mathcal{H}_S$. However, it is true that $|\mathcal{H}_S| \geq A_{\text{min}}(S)$ where $A_{\text{min}}(S)$ is the infimum of the areas of all spacelike surfaces enclosing the trapped surface $S$. Thus, if weak cosmic censorship holds, the global Penrose inequality implies the following inequality which involves only objects defined in terms of the local geometry of the initial data set

$$M_{\text{ADM}} \geq \sqrt{\frac{A_{\text{min}}(S)}{16\pi}}. \quad (2)$$

The plausibility argument behind this inequality rests strongly on the validity of the weak cosmic censorship hypothesis. In particular, finding initial data sets which violate (2) would give initial data which very likely evolve into a spacetime containing naked singularities, thus contradicting the cosmic censorship. In fact, Penrose’s original intention [1] for studying (2) was to find suitable counterexamples to the “standard view of gravitational collapse”, namely counterexamples of weak cosmic censorship. However, no counterexample has been found so far and the point of view has shifted towards conjecturing the validity of (2). Proving this conjecture would obviously not say anything about the validity of weak cosmic censorship, but it would certainly give strong indirect support for it (because, why should otherwise such an inequality be true?). Inequality (2) (or modifications thereof, see below) is known as the Penrose inequality, sometimes also “Isoperimetric inequality for black holes” [16] and its proof remains one of the important challenges in general relativity and in differential geometry to date. In this review we will summarize some of the existing results on the Penrose inequality. Since the topic is vast, I will only aim at giving a general overview on the subject, with no claim of exhaustivity. For an alternative excellent review, the reader is referred to [17].

2. Definitions

A spacetime $(\mathcal{M}, g)$ is a 4-dimensional connected Hausdorff manifold endowed with a smooth metric of signature $+2$. We will further assume that $\mathcal{M}$ is orientable and $(\mathcal{M}, g)$ is time-orientable, i.e. it admits a smooth timelike vector field. We assume $C^\infty$ just for simplicity, most of the results hold under much weaker differentiability assumptions.

The Penrose inequality involves the area of certain compact codimension-two spacetime surfaces. Let us therefore start with some basic definitions on the geometry of surfaces.

Let $S$ be a compact, embedded, codimension-two surface in $(\mathcal{M}, g)$ defined via an embedding $\Phi : S \rightarrow \mathcal{M}$. Assume that the induced metric $h$ on $S$ is positive definite, i.e. $S$ is spacelike. We will often identify $S$ and its image $\Phi(S)$ in the spacetime. At $p \in S$ denote by $T_p S$ and $N_p S$ the tangent and normal spaces of $S$. Being $S$ non-null it follows $T_p \mathcal{M} = T_p S \oplus N_p S$. According to this decomposition, the tangential and normal components of a vector $\vec{V}$ are denoted respectively by $\vec{V}^\parallel$ and $\vec{V}^\perp$. The second fundamental form vector of $S$, denoted by $\vec{K}$ is defined as usual as

$$\vec{K}(\vec{X}, \vec{Y}) = -\left(\nabla_\vec{X} \vec{Y}\right)^\perp$$

where $\vec{X}$, $\vec{Y}$ are tangent vectors to $S$. The second fundamental form is symmetric and its trace $\vec{H} = \text{tr}_h \vec{K}$ is the mean curvature vector. Our sign conventions are such that the mean curvature vector of a sphere in $\mathbb{R}^3$ points outwards.

Assume next that $S$ is orientable and that there exist globally on $S$ two linearly independent, nowhere zero, null vector fields $\vec{l}^+$ and $\vec{l}^-$ orthogonal to $S$. They can be chosen without loss of generality future directed and satisfying $(\vec{l}^+ \cdot \vec{l}^-) = -2$. They are uniquely defined up to rescalings $\vec{l}^+ \rightarrow F \vec{l}^+$, $\vec{l}^- \rightarrow F^{-1} \vec{l}^-$, $F > 0$ defined on $S$, plus interchange $\vec{l}^+ \leftrightarrow \vec{l}^-$. The null expansions of $S$ are defined as $\theta_{\pm} = (\vec{H} \cdot \vec{l}_{\pm})$. An alternative definition of $\theta_{\pm}$ can be given in terms of the divergence of $\vec{l}_{\pm}$ as

$$\theta_{\pm} = h^{\mu \nu} \nabla_\mu \vec{l}_{\pm}^\nu |_S.$$ \quad (3)
where $h^{\mu\nu}$ is the projector tangent to $S$, explicitly $h_{\mu\nu} = g_{\mu\nu} + 1/2(l^\mu l^\nu + l^\nu l^\mu)$. Formally, expression (3) requires an extension of $\tilde{l}^\pm$ off $S$ but the result is clearly independent of the extension. Whenever the extension is chosen so that $\tilde{l}^\pm$ is geodesic and affinely parametrized the expansion can be written as $\theta_{\pm} = \nabla_{\alpha} l^\alpha |_{S}$. The null expansions contain the same information as the mean curvature vector as the following decomposition holds: $\vec{H} = -\frac{1}{2} \left( \theta_{-} \tilde{l}^{-} + \theta_{+} \tilde{l}^{+} \right)$. The main geometrical content of the mean curvature vector comes from the first variation of area. Take any vector field $\vec{\xi}$ defined on $S$ and extend it off $S$ arbitrarily. For $s \in \mathbb{R}$ small enough, let $\Phi_s : S \rightarrow \mathcal{M}$ be the embedding defined so that any $p \in S$ is moved a parametric amount $s$ along the integral curve of $\vec{\xi}$ starting at $p$. We will call $\Phi_s$ a variation of $S$ along the vector $\vec{\xi}$.

For any embedded compact spacelike surface $S$ we denote by $|S|$ its area. Then, the first variation of area reads (see e.g. [18])

$$\frac{d|S_s|}{ds} \bigg|_{s=0} = \int_S \left( \vec{H} \cdot \vec{\xi} \right) \eta_S,$$

where $\eta_S$ is the metric volume form of $S$. In particular, the change of area of $S$ for variations along the null direction $\tilde{l}^\pm$ is determined by the integral of $\theta_{\pm}$. A future trapped surface is defined by the property that the area strictly decreases for any future null variation. Equivalently $\vec{H}$ is future timelike everywhere or $\theta_{+} < 0$, $\theta_{-} < 0$. Past trapped surfaces are defined interchanging future and past and satisfy $\theta_{+} > 0$ and $\theta_{-} > 0$.

A future marginally trapped surface is defined by the conditions $\theta_{+} = 0$ and $\theta_{-} < 0$ for some choice of $\tilde{l}^\pm$. Past is defined by reversing the signs. It is often useful to consider surfaces which have one of the null expansions identically vanishing and the other arbitrary. These surfaces are called outer marginally trapped and the outer direction is defined to be the null direction pointing outside the compact region).

It will be often necessary to consider surfaces embedded in a spacelike hypersurface $\Sigma$ which itself is embedded in the spacetime $\mathcal{M}$. Assume as before that $\Sigma$ is orientable. Then it admits a unique (up to orientation) globally defined unit normal vector $\vec{\nu}$ which is at the same time tangent to $\Sigma$. The future unit normal to $\Sigma$ denoted by $\vec{\nu}$ is obviously also normal to $S$. Null normal vectors can then be defined in terms of $\vec{\nu}$ as $\tilde{l}^+ = \vec{\nu} + \vec{m}$ and $\tilde{l}^- = \vec{\nu} - \vec{m}$. Changing the orientation of $\vec{m}$ amounts to interchanging $\tilde{l}^+$ and $\tilde{l}^-$. The spacelike hypersurface $\Sigma$ has a positive definite induced metric, which will be called by $\gamma_{ij}$, and a second fundamental form w.r.t. $\vec{\nu}$, denoted by $A_{ij}$ (Latin indices refer to tensors defined on $\Sigma$). As a submanifold of $\Sigma$, $S$ has a mean curvature w.r.t. $\vec{m}$ which is denoted by $p$. Then (see e.g. [18]) $\vec{H} = p\vec{m} - q\vec{\nu}$, where $q = \text{tr}_{h} A$, which implies $\theta_{+} = \pm p + q$.

A second basic ingredient of the Penrose inequality is the concept of total mass of a spacetime or of an initial data set. An initial data set is a quintuple $(\Sigma, \gamma, A_{ij}, \rho, J_i)$ where $(\Sigma, \gamma)$ is a Riemannian manifold, $A_{ij}$ a symmetric tensor, $\rho$ a scalar field called energy density and $J_i$ a vector called energy flux, satisfying the constraint equations

$$R(\gamma) - A_{ij}A^{ij} + A^2 = 16\pi \rho,$$

$$\nabla_j A^j_i - \nabla_i \text{tr}_\gamma A = -8\pi J_i,$$

where $R(\gamma)$ is the Ricci scalar of $\gamma$, $\text{tr}_\gamma A = A^j_j$ and $\nabla_i$ denotes the metric covariant derivative in $(\Sigma, \gamma)$. If the matter model is such that the Einstein-matter equations define a well-posed system, the initial data set will generate a spacetime in the usual sense that there exists a
unique globally hyperbolic spacetime \((\mathcal{M}, g)\) satisfying the Einstein-matter equations admitting a Cauchy surface isometric to \((\Sigma, \gamma)\) and with second fundamental form \(A_{ij}\).

A spacetime \((\mathcal{M}, g)\) is said to satisfy the dominant energy condition (DEC), if the Einstein tensor \(\mathcal{G}_{\alpha\beta}\) of \(g\) maps future causal vectors into past causal vectors. DEC implies for the initial data the inequality

\[
\rho \geq \sqrt{J_i J_i} \quad \text{(DEC)}.
\]

With a slight, but generally used, abuse of notation we will say that the initial data set satisfies DEC provided (7) holds.

An initial data set is asymptotically flat provided \(\Sigma\) is the disjoint union of a compact set \(K\) and a set \(\Sigma_{\infty}\) diffeomorphic to \(\mathbb{R}^3 \setminus B\) where \(B\) is a closed ball. Moreover, in Cartesian coordinates \(x^i\) in \(\Sigma_{\infty}\) induced by this diffeomorphism, the metric and second fundamental forms have the following asymptotic behaviour

\[
\gamma_{ij} = \delta_{ij} + O\left(\frac{1}{r}\right), \quad \partial_i \gamma_{ij} = O\left(\frac{1}{r^2}\right), \quad \partial \partial_i \gamma_{ij} = O\left(\frac{1}{r^3}\right),
\]

\[
A_{ij} = O\left(\frac{1}{r^2}\right), \quad \partial A_{ij} = O\left(\frac{1}{r^3}\right),
\]

where \(r = \sqrt{\delta_{ij} x^i x^j}\). The Penrose inequality involves the total mass of a spacetime. In the asymptotically flat setting, the total ADM energy-momentum vector [19] is defined through the coordinate expressions

\[
E = \lim_{r \to \infty} \frac{1}{16\pi} \int_{S_r} (\partial \gamma_{ij} - \partial_i \gamma_{jj}) \, dS^i,
\]

\[
P_i = \lim_{r \to \infty} \frac{1}{8\pi} \int_{S_r} (A_{ij} - \gamma_{ij} \text{tr} \, \gamma \, A) \, dS^j,
\]

where \(S_r\) is the surface at constant \(r\) and \(dS^i = \vec{n}^i dS\) with \(\vec{n}\) being the outward unit normal to \(S_r\) and \(dS\) the surface element. This definition depends a priori on the choice of coordinates \(x^i\). However, they can be shown to define geometric quantities on \((\Sigma, \gamma_{ij}, A_{ij})\) provided the constraints satisfy the additional decay properties

\[
R(\gamma) = O\left(\frac{1}{r^4}\right), \quad \nabla_i A^i_j - \nabla_j A = O\left(\frac{1}{r^4}\right).
\]

In fact, the ADM energy-momentum is shown in [20], generalizing results in [21], to be well-defined under much weaker conditions, where the metric and second fundamental forms belong to appropriate weighted Sobolev spaces involving just two derivatives for \(h\) and one for \(A_{ij}\), and that the left hand sides of (8) are integrable on \(\Sigma\). The total ADM mass of an asymptotically flat initial data set is defined as \(M_{\text{ADM}} = \sqrt{E^2 - \delta^{ij} P_i P_j}\). The statement that \(M_{\text{ADM}}\) is real and in fact strictly positive except whenever \(\rho = 0, J_i = 0\) and \((\Sigma, \gamma, A)\) corresponds to a slice of Minkowski spacetime, is the content of the positive mass theorem of Schoen and Yau [22, 23].

### 3. First order variations of the null expansions

An important technical tool that will be used often below is how the null expansions of a spacelike surface vary to first order when the surface is perturbed along arbitrary directions. In this section we write down the variations of \(\theta^\pm\) along the null normals \(\vec{l}^\pm\). The first thing that needs to be noticed is that \(\theta^\pm\) are not functions on the spacetime, but functionals on surfaces. Thus, when writing e.g. \(\vec{l}_+(\theta_-)\) what is really meant is the following:
A variation of $S$ is defined through a differential map $\Phi : S \times I \to (\mathcal{M}, g)$, where $I \ni 0$ is an open interval such that $\Phi_\lambda \equiv \Phi(\cdot, \lambda)$ is an embedding for all $\lambda \in I$. The variation vector $\tilde{\xi}$ is defined as $\tilde{\xi} = d\Phi(\partial_\lambda)|_{\lambda=0}$ and is therefore defined only on $S$ (we identify $S$ with its image in $\Sigma$ under $\Phi_0$). Let $S_\lambda = \Phi_\lambda(S)$. This defines a codimension two surface which we assume to be spacelike. We can therefore calculate the various geometric objects of $S_\lambda$. Let $\tilde{\xi}$ be a geodesic satisfying $\frac{d}{d\lambda}(\tilde{\xi}(\lambda)) = Q\tilde{\xi}_\lambda$. We allow for arbitrary $Q : I \times S \to \mathbb{R}$ i.e. we do not assume the parameter $\lambda$ be affine. A direct calculation shows that

$$\tilde{\xi}(F)|_S = \frac{dF(\lambda, \cdot)}{d\lambda} \bigg|_{\lambda=0},$$

where the entry $\cdot$ stands for the point on $S$ where the calculation is being done.

Let us next consider the derivatives along the null normal $\tilde{\xi}$. Since the derivative at $S$ depends on the variation vector $\tilde{\xi}$ at $S$ only (and is independent of its extension off $S$), let us without loss of generality extend $\tilde{\xi}_\lambda$ to be a geodesic satisfying $\tilde{\xi}_\lambda^\alpha \nabla_\alpha \tilde{\xi}_\lambda^\beta = Q\tilde{\xi}_\lambda^\beta$. We allow for arbitrary $Q : I \times S \to \mathbb{R}$ i.e. we do not assume the parameter $\lambda$ be affine. A direct calculation shows that

$$\tilde{\xi}_\lambda^A = 2\tilde{\xi}_\lambda^\beta \cdot \tilde{R}_AB^\lambda.$$

Using this expressions a straightforward calculation gives the Raychaudhuri equation

$$\tilde{\xi}_\lambda^A(\theta_+) = Q\theta_+ - \left( \tilde{\xi}_\lambda^\beta \cdot \tilde{R}_AB^\lambda \right) \left( \tilde{\xi}_\lambda^\beta \cdot \tilde{R}^{AB}_\lambda \right) - \text{Ric} \left( \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta \right),$$

where $\text{Ric}$ is the Ricci tensor of the spacetime $(\mathcal{M}, g)$. A longer calculation gives the derivative of $\theta_-$ along $\tilde{\xi}_\lambda$

$$\tilde{\xi}_\lambda^A(\theta_-) = -Q\theta_- - \left( \tilde{\xi}_\lambda^\beta \cdot \tilde{R}_AB^\lambda \right) \left( \tilde{\xi}_\lambda^\beta \cdot \tilde{R}^{AB}_\lambda \right) - \text{Ric} \left( \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta \right) - \frac{1}{2} \text{Riem} \left( \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta \right) + 2 \left( D_A S^A + S_A S^A \right),$$

where $\text{Riem}$ is the Riemann tensor of $(\mathcal{M}, g)$, $D$ denotes the Levi-Civita covariant derivative of $(S_\lambda, h)$ and the one-form $S_A$ tangent to $S_\lambda$ is defined by

$$S_A = -\frac{1}{2} \left( \tilde{\xi}^\beta \cdot \nabla_\beta \tilde{\xi}^\lambda \right)$$

and $e^\alpha_A = \frac{\partial \Phi}{\partial y^\alpha}$ with $y^\alpha$ being local coordinates on $S$. The Gauss equation applied to $S_\lambda$ implies

$$R(g) = R(h) - \tilde{H}^2 + \tilde{R}^{AB} \cdot \tilde{R}^{AB}_\lambda - 2 \text{Ric} \left( \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta \right) + \frac{1}{2} \text{Riem} \left( \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta \right),$$

where $R(g)$ is the scalar curvature of $(\mathcal{M}, g)$. This allows for an alternative form of (10), namely

$$\tilde{\xi}_\lambda^A(\theta_-) = -Q\theta_- + \text{Ein} \left( \tilde{\xi}_\lambda^\beta, \tilde{\xi}_\lambda^\beta \right) - \left( R(h) - \tilde{H}^2 \right) + 2 \left( D_A S^A + S_A S^A \right),$$

where Ein is the Einstein tensor.
4. Formulations of the Penrose inequality

The heuristic derivation of the Penrose inequality gives \( M_{ADM} \geq \sqrt{A_{\text{min}}(S)/16\pi} \) where \( A_{\text{min}}(S) \) is the minimum area required to enclose any given future trapped surface \( S \). Since this inequality should hold for any future trapped surface \( S \) it should also be true

\[
M_{ADM} \geq \sup_S \sqrt{\frac{A_{\text{min}}(S)}{16\pi}},
\]

where the supremum is taken with respect to all future trapped surfaces \( S \). This, however, is a rather complicated object. It is natural to ask whether it can be substituted by something simpler. In order to do that, it is useful to define the future trapped region \( T \subset \Sigma \) by deciding that \( x \in T \) if and only if there exists a future trapped surface \( S \) contained in \( \Sigma \) and containing \( x \). Very little is known about the set \( T \) in general, despite its physical relevance as it characterizes locally (in time) the region where the gravitational field is “intense”. Two known facts about \( T \) are that it is open and it has compact closure, i.e. it does not extend to infinity in any direction (recall that \( \Sigma \) is complete as part of the asymptotic flatness definition). Moreover, if the topological boundary \( \partial T \) is smooth, then it must be a future marginally trapped surface [24]. However, it is unclear whether \( \partial T \) should be expected to be smooth. Another obvious property is that if a surface encloses \( \partial T \), then it encloses all trapped surfaces in \( \Sigma \). If we take the minimum area \( S_{\text{min}} \) enclosure of \( \partial T \), it is clear from the definitions that \( |S_{\text{min}}| \geq \sup_S A_{\text{min}}(S) \) and it is plausible that they are equal (although I am not aware of a proof). If they were equal then the Penrose inequality could be formulated as

\[
M_{ADM} \geq \sqrt{\frac{|S_{\text{min}}|}{16\pi}}.
\]

In any case, (13) implies (12), and it has the advantage that it does not involve taking any supremum. It becomes specially useful with \( \partial T \) is a marginally trapped surface. Thus one can take (13) to be the Penrose inequality on initial data sets. As far as I know this formulation is due to Horowitz [25].

Although the Penrose inequality, seen as a consequence of cosmic censorship, should be expected to hold only for the minimum area enclosure of \( \partial T \), it is often the case that the inequality is written down directly for an outermost marginally trapped surface \( S_H \) (provided such surface exists), i.e.

\[
M_{ADM} \geq \sqrt{\frac{|S_H|}{16\pi}}.
\]

So far no counterexample of this stronger inequality has been found (though not many examples either, see [26] for a numerical analysis of a special class of initial data sets). In the literature, inequality (14) is often called Penrose inequality as well.

5. The Penrose inequality in spherical symmetry

The only case where the Penrose inequality has been proven without the assumption of time symmetry is in spherical symmetry, i.e. assuming that \( \gamma_{ij} \) and \( A_{ij} \) are invariant under an \( SO(3) \) action with two-dimensional orbits (which may degenerate to points). This inequality was first established by Malec and ´O Murchadha [27] assuming that the initial data is maximal, i.e. \( \text{tr}_\gamma A = 0 \) and in full generality by Hayward [28].

Since this a simple and instructive case, we outline a proof, which differs slightly but is close in spirit from the argument in [28]. Let \( (\Sigma, \gamma, A) \) be an asymptotically flat initial data set with spherical symmetry. The trapped region \( T \) introduced in the previous section is geometrically
defined and therefore spherically symmetric (if a point $x \in T$, the orbit containing $x$ must also belong to $T$ as the action of any element of $SO(3)$ on a trapped surface gives another trapped surface). It follows that the topological boundary $\partial T$ is also spherically symmetric and therefore smooth and in fact topologically $S^2$ if non-empty, which we assume from now on. Two possibilities arise: (i) either in the region outside $T \cup \partial T$ there are minimal surfaces (possibly intersecting $\partial T$) or alternatively (ii) $\partial T$ has less or equal area than any surface that encloses it, i.e. $\partial T$ is area outer minimizing. That these two conditions are exclusive follows from the fact that if $\partial T$ is not area outer minimizing, then finding the global minimum of the area functional among all surfaces enclosing $\partial T$ we would find a minimal surface which lies in the exterior of $T \cup \partial T$, hence we would fall into case (i). Conversely, assume there is a minimal surface outside $T \cup \partial T$. Consider the outermost minimal surface $S_{\text{min}}$ of $(\Sigma, \gamma)$, which exists in every asymptotically flat Riemannian manifold containing minimal surfaces. From spherical symmetry, $S_{\text{min}}$ must be spherically symmetric and therefore must enclose $\partial T$. The area of the concentric spheres lying between $\partial T$ and $S_{\text{min}}$ cannot monotonically increase from $\partial T$ to $S_{\text{min}}$ (otherwise $S_{\text{min}}$ would not be a strict minimum of the area functional). Taking the minimum of these areas we find a surface that encloses $\partial T$ and has smaller area, which shows that $\partial T$ is not area outer minimizing.

In case (i) we restrict our attention to the exterior region of $S_{\text{min}}$ and in case (ii) to the exterior of $\partial T$. In both cases, we have an asymptotically flat spherically symmetric initial data $(\Sigma_{\text{ext}}, \gamma_{\text{ext}}, A_{\text{ext}})$ with spherical boundary which is free of any minimal surfaces in the interior. The metric $h_{\text{ext}}$ can be written in the form

$$h_{\text{ext}} = \frac{dr^2}{1 - 2m(r)} + r^2 \left( d\theta^2 + \sin^2 \theta d\phi^2 \right),$$

where $m(r)$ is a function that satisfies $2m(r) > r$ and $\lim_{r \to \infty} m(r) = E_{\text{ADM}}$. The second fundamental form can be written in these coordinates as $A_{\text{ext}} = C(r)dr^2 + B(r)\left( d\theta^2 + \sin^2 \theta d\phi^2 \right)$. Asymptotic flatness implies $C(r) = O(\frac{1}{r})$, $B(r) = O(\frac{1}{r^2})$. The dominant energy condition (7) reads $\rho \leq \left| J_r \sqrt{1 - 2m(r)} \right|$ where $J_r = (J, \partial_r)$. The constraint equations (5,6) become simply

$$\frac{dm(r)}{dr} = 4\pi r^2 - \frac{B^2}{2r^2} - BC \left( 1 - \frac{2m(r)}{r} \right),$$

$$\frac{dB(r)}{dr} = 4\pi r^2 J_r + \frac{B}{r} + C (r - 2m(r)).$$

If we define $M_H(r) \equiv m(r) + \frac{B^2}{2r}$, these two equations imply

$$\frac{dM_H}{dr} = 4\pi r^2 \left( \rho + \frac{B}{r} J_r \right).$$

Moreover, the mean curvature $p$ and trace $q$ of the second fundamental forms on the surfaces \{r = const\} are $p = \frac{2}{r} \sqrt{1 - \frac{2m}{r}} > 0$, $q = 2B$. In addition, the decay at infinity implies $\theta_+ > 0$ and $\theta_- < 0$ for sufficiently large spheres.

Let us first deal with case (i). Since $\partial \Sigma_{\text{ext}}$ is a minimal surface, we have $p = 0$ and moreover $\theta_+ = \theta_- = q \geq 0$ because it lies outside $T$ and hence cannot be a future trapped surface (it is therefore either a past trapped surface if $q > 0$ or a marginally trapped surface if $q = 0$). Thus, there must exist an outermost spherically symmetric past marginally trapped surface $S^-$ (which may coincide with $\partial \Sigma_{\text{ext}}$). Outside of it, we have $\theta_+ \leq 0$ and $\theta_- < 0$, which implies

$$| Br^{-1} \left( 1 - \frac{2m}{r} \right)^{-1/2} | \leq 1.$$
and hence $M_H$ non-decreasing, provided DEC (7) holds. Being $S^-$ past marginally trapped, we have $M_H(S^-) = \frac{r^2}{2} = \sqrt{|S^-|}/16\pi$. Applying the monotonicity of $M_H$ from $S^-$ to infinity and using that $\partial \Sigma_{\text{ext}}$ is minimal surface, the following inequalities follow

$$\sqrt{\frac{|\partial \Sigma_{\text{ext}}|}{16\pi}} \leq \sqrt{\frac{|S^-|}{16\pi}} \leq E_{\text{ADM}},$$

which establishes the Penrose inequality for case (i). In case (ii) a similar argument applies because $\partial T$ is area outer minimizing. If there is a past marginally trapped surface $S^-$ outside $\partial T$, apply the monotonicity of $M_H$ from $S^-$ to infinity. If there is none, apply monotonicity of $M_H$ from $\partial T$ to infinity. In either case one concludes

$$E_{\text{ADM}} \geq \sqrt{\frac{|\partial T|}{16\pi}}.$$  \hspace{1cm} (16)

Notice that the Penrose inequality does not state (16) in case (i). In that case the minimum area needed to enclose $\partial T$ must be used. This in agreement with the heuristic argument in Sect. 1. In fact, Ben-Dov [29] has found an explicit example in spherical symmetry where the inequality (16) is violated.

The Penrose inequality proven in spherical symmetry involves the total energy of the slice. This is weaker inequality than the expected Penrose inequality in terms of the total ADM mass. The stronger inequality remains open. In my opinion proving the stronger version would probably give some ideas on how the general Penrose inequality might be approached.

6. Riemannian Penrose inequality

The Penrose conjecture has seen notable advances in the last ten years. Although a general proof is still lacking, fundamental work by Huisken and Ilmanen [30] and Bray [31] has shown the inequality in the special but relevant case of time-symmetry.

By definition, an initial data set is called time-symmetric whenever $A_{ij} = 0$. This has two immediate consequences, namely that the ADM three-momentum vanishes identically, so that $M_{\text{ADM}} = E_{\text{ADM}}$, and only one constraint equation remains,

$$R(\gamma) = 16\pi \rho$$

which gives $R(\gamma) \geq 0$ provided the dominant energy condition holds. In fact, the weak energy condition (Ein$(\vec{u}, \vec{u}) \geq 0$ for all causal vectors) suffices to reach the same conclusion in this case.

Another immediate consequence is that there exist no trapped surfaces in the initial data because $\theta_+ = -\theta_-$, and we have $T = \emptyset$. Moreover, the only future or past marginally trapped surfaces in $(\Sigma, \gamma)$ are the minimal surfaces. The Penrose inequality therefore becomes an inequality relating the total mass and the area of the outermost minimal surface in $(\Sigma, \gamma)$. For asymptotically flat Riemannian metrics with $R(\gamma) \geq 0$, the results of Meeks III and Yau [32], see also [30], imply the existence on an outermost minimal surface $S_m$ which separates $\Sigma$ into an asymptotically flat region $\Sigma$ and a compact manifold $K$ such that $\Sigma$ does not contain any immersed minimal surface. Moreover, each connected component of $S_m$ has spherical topology. The Penrose inequality on this setting would read $M_{\text{ADM}} \geq \sqrt{|S_m|/(16\pi)}$. This is an inequality than can be stated directly in terms of Riemannian manifolds, and it is often named “Riemannian Penrose inequality”.

6.1. Huisken and Ilmanen’s proof

The heuristic idea behind the proof of Huisken and Ilmanen was first proposed by Geroch [33] and is based on the observation that there exists a scalar functional on surfaces which is monotonic
under a suitable flow of surfaces and which interpolates between $\sqrt{|S_m|/16\pi}$ on the outermost minimal surface and $M_{ADM}$ at infinity. This functional generalizes to arbitrary surfaces the scalar function $M_H(r)$ which we used in Sect. 5. This functional is now often called Geroch mass and is defined, on any compact surface $S$ embedded in $(\Sigma, \gamma)$

$$M_G(S) = \sqrt{|S|/16\pi} \left( \frac{\chi(S)}{2} - \int_S p^2 \eta_S \right)$$

where $\eta_S$ is the volume form of the induced metric on $S$ and $\chi(S)$ is the Euler characteristic of $S$. Hence $\chi(S) = 2$ for a connected surface with spherical topology. For an arbitrary compact surface $\chi(S) = \sum 2(1 - g)$ where the sum is over the connected components and $g$ is the genus.

The aim is to calculate how $M_G$ changes under a general variation. Since any tangential component of the variation vector $\xi$ would correspond to a diffeomorphism within $S$ and would not affect the geometric variation of $S$ within $\Sigma$ we will assume, without loss of generality that $\xi$ is orthogonal to $S$ so that we can write $\xi = e^\psi \vec{m}$ where $\vec{m}$ is the unit outer normal vector to $S$.

An obvious property of $M_G$ is that its value on any connected, topologically $S^2$ minimal surface is $M_G = \sqrt{|S|/(16\pi)}$. For surfaces with asymptotically flat end $\Sigma_\infty$, the asymptotic decay of $\gamma$ implies $\lim_{r \to \infty} M_G(S_r) = M_{ADM}$. Thus, $M_G$ indeed interpolates between the left and right hand sides of the Penrose inequality. In order to write down a general variation formula for the Geroch mass under normal variation vectors, we need to introduce the second fundamental form $k_{AB}$ of $S$ as a submanifold of $\Sigma$ and its trace free part $\Pi_{AB} = k_{AB} - \frac{1}{2} \pi \eta_{AB}$. This general variation formula was first obtained by Geroch [33] and can be written in the following form

$$\frac{dM_G(S_\lambda)}{d\lambda} = \frac{1}{8\pi} \sqrt{|S_\lambda|/16\pi} \int_{S_\lambda} \left[ p e^\psi \left( \frac{R(\gamma)}{2} + \frac{1}{2} \Pi_{AB} \Pi^{AB} + D_A \psi D^A \psi \right) + \left( p e^\psi - a(\lambda) \right) \left( \Delta_{h_\lambda} \psi - \frac{1}{2} R(h_\lambda) + \frac{1}{4} p^2 \right) \right] \eta_{S_\lambda},$$

where $R(h)$ denotes the curvature scalar of $h$ and the constant $a(\lambda)$ is defined as the average of $pe^\psi$, i.e.

$$a(\lambda) = \frac{\int_{S_\lambda} pe^\psi \eta_{S_\lambda}}{|S_\lambda|}.$$ 

The calculation leading to (18) requires using the Gauss-Bonnet formula

$$\int_{S_\lambda} R(h_\lambda) \eta_{S_\lambda} = 4\pi \chi(S_\lambda),$$

which is valid in two dimensions. Expression (18) is therefore valid for any two-dimensional surface embedded into a three-dimensional Riemannian manifold.

Two of the terms in the first line of (18) are clearly non-negative. The term $R(\gamma)$ is non-negative provided the weak energy condition holds and $\text{tr} \gamma A = 0$ (in particular in the time symmetric case $A_{ij} = 0$).

One way of making $M_G$ monotonic under the flow is to demand that $pe^\psi = 1$ so that the first parenthesis in the second line vanishes identically. Such a flow has the property that the velocity is inversely proportional to the mean curvature of the surface at each point, and is therefore called Inverse Mean Curvature Flow, IMCF.
If the flow could be started on the outermost minimal surface (which has $p = 0$) and existed for all $\lambda$ while approaching large round spheres at infinity, the Penrose inequality would follow. In fact, Geroch original idea was to use this flow starting from a point (so that $M_G = 0$ initially) and therefore prove the positive mass theorem. Later Jang and Wald [34] realized that the same argument starting from an outermost minimal surface would prove the Penrose inequality. However, it was immediately realized that the flow will in general have singularities. A simple example of this consists of an IMCF starting on any surface with non-trivial genus (say a torus) in Euclidean space. Since the flow is by construction outer directed, and the velocity can never be zero for a smooth surface (so that the evolution cannot freeze) it is clear that after some time the flow, if it can be extended that far, will eventually evolve the surface so that two points on the hole touch each other, thus giving rise to a non-regular surface. Another instance of the same singular behaviour is observed by taking a surface composed of two round spheres. By symmetry, an the fact that the flow is locally defined, it is clear that each sphere will evolve into a larger and larger sphere until they will eventually touch each other (that they touch instead of approaching some asymptotic state follows from the fact that larger spheres have smaller mean curvature and hence flow at a larger velocity under IMCF).

The presence of singularities made this idea dormant for decades. Huisken and Ilmanen’s fundamental contribution was to define the flow in a suitably weak sense so that the singularities could be treated (and in fact basically avoided). The argument rests on two complementary ideas. First of all, write down the flow, which is a geometric parabolic flow, in terms of level sets i.e. the leaves of the flow become the level sets of a real function $u$ on $\Sigma$. This function satisfies a degenerate elliptic equation

$$\nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) = |\nabla u|$$

where $\nabla \cdot V$ is the divergence of $V$. Equation (20) translates directly the IMCF condition into the level set formulation. In principle it is possible that $u$ remains constant on open sets, which has the immediate consequence that the flow may jump across regions with positive measure. The fundamental idea is to use these jumps precisely to avoid the singularities that the smooth IMCF would otherwise have. In order to achieve this Huisken and Ilmanen find a variational formulation for (20). This equation is not the Euler-Lagrange equation of any functional. However, by freezing the right-hand side to $|\nabla u|$, the authors write down the functional (which now depends on $u$)

$$E_u(v) = \int_\Sigma (|\nabla v| + v|\nabla u|) \eta_\gamma.$$  

The critical points of this functional with respect to compactly supported variations of $v$ gives (20) with $u$ replaced by $v$ on the left hand side. One then looks for functions $u$ which minimize their own functional, thus giving (20).

This variational formulation has a geometric counterpart which implies that each of the level sets of $u$ (i.e. the surface $S_\lambda$) is area outer minimizing. This means that it has less of equal area that any other surface $S'$ that encloses it. Thus, if we start from a smooth surface $S_0$ that is area outer minimizing, the IMCF which, being defined by a geometric parabolic equation enjoys short time existence, will evolve the surface smoothly for some “time”. For small enough values of $\lambda$, each level set will be outer area minimizing. However, there may well exist a value for which $S_\lambda$ ceases to be area outer minimizing. Take the smallest of such values, $\lambda_1$, so that there exists a surface $S'_{\lambda_1}$ which encloses it and has less or equal area. In fact, it must have equal area because if it had strictly less area, it would also have less area than some close enough previous leaf (the flow is smooth before the jump and the area changes smoothly). The surface $S'_{\lambda_1}$ may have common points with $S_{\lambda_1}$. However, the surface $S'_{\lambda_1} \setminus S_{\lambda_1}$ must be a minimal surface. If it
were not, it would admit a compactly supported variation that would have less area and would remain outside $S_{\lambda_1}$, still enclosing it. This again means that the jump should have happened before because this varied surface enclose all surfaces in the flow from $S_0$ up to $S_{\lambda_1}$, has larger area than $S_0$ (because the latter is area outer minimising) and has less area that $S_{\lambda_1}$. So, for some value of the flow parameter smaller than $\lambda_1$ it would have the same area, and it would encloses it, against definition of $\lambda_1$ as being the smallest with such property. Notice that the smooth flow could have become singular before such a non-outter minimizing leaf is achieved. That this cannot happen is one of the several statements that Huisken and Ilmanen had to show, and which makes their proof technically difficult.

In the torus example before, the smooth flow would thicken the torus (in an axially symmetric way) up to some surface which has exactly the same area than the surface obtained by closing the hole by two horizontal planes. It is clear in this example that the new pieces have vanishing mean curvature.

Since the level set function $u$ is such that each level set is outer area minimising, this means that the IMCF evolves smoothly for as long as the leaves remains area outer minimizing until a surface is reached (if at all) which is not area outer minimizing. At this point, the surface jumps (meaning that $u$ remains constant between the two surfaces). The flow should be continued from the new surface outwards. Since the new surface has pieces with $p = 0$, the IMCF cannot be defined just there in a classical sense. However, the weak formulation in terms of level sets also takes care of this.

A crucial condition for the monotonicity argument to go through in this formulation is that the Geroch mass of the new surface after the jump is still larger or equal than before the jump. We know the total area does not change. Moreover, the new pieces of the surface have $p = 0$, while the deleted pieces have $p > 0$ because the flow was smooth up to and including $S_{\lambda_1}$. Thus, across the jump the Geroch mass is non-decreasing provided the Euler characteristic of the surface does not decrease. The example with the torus in Euclidean space shows that in general the topology of the surface may change across the jump, in that case changing from a surface with genus one to a topological sphere. In this particular case, the Euler characteristic increases and the Geroch mass would be monotonically increasing. The question is therefore whether this is a general property or not. The example with two spheres in Euclidean space show that this cannot be true in general. There the Euler characteristic starts being four and after the jump we have a topological sphere, so $\chi = 2$. A general result by Huisken and Ilmanen [30] is that a topologically $S^2$ surface in an asymptotically flat euclidean three-manifold $(\Sigma, \gamma)$ outside the outermost minimal surface cannot jump, via the weak formulation of the IMCF, to another surface with smaller Euler characteristic, i.e. no holes can appear on the surface after the jump. Nevertheless, the outermost minimal surface $S$ in $(\Sigma, \gamma)$ need not be connected and the IMCF must start on $S$ in order to interpolate between the area of $S$ and the total mass of $(\Sigma, \gamma)$. Here is where the restriction on the topology of the horizon comes into play. The Geroch mass is therefore monotonic under the weak IMCF starting on any connected component of the outermost minimal surface. Hence the inequality

$$M_{\text{ADM}} \geq \max_i \sqrt{\frac{|S_i|}{16\pi}}$$

holds, where $S_i$ is each one of the connected components of the outermost minimal surface in $(\Sigma, \gamma)$.

The Penrose inequality has a rigidity part, namely that equality is achieved only for the Schwarzschild metric. More precisely, the time symmetric initial data set for the Kruskal extension of the Schwarzschild metric is $\Sigma_{\text{Sch}} = \mathbb{R}^3 \setminus \{0\}$ with metric

$$g_{\text{Sch}} = \left(1 + \frac{m}{2r}\right)^{\frac{4}{2}} \left(dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2\right).$$
The surface \( r = m/2 \) is minimal and separates the manifold into two isometric pieces (corresponding to the two asymptotic end of the Kruskal metric). The rigidity part of the Riemannian Penrose inequality states that if equality is achieved in (22) then the region in \((\Sigma, \gamma)\) outside its outermost minimal surface \( S_m \) is isometric the domain \( r > m/2 \) of \((\Sigma_{Sch}, g_{Sch})\) with \( m = M_{ADM} \). Heuristically, it is clear that some rigidity is to be expected already from the monotonicity property of the Geroch mass. If equality is achieved, the derivative of the Geroch mass is everywhere zero. In particular, in the smooth part of the flow \( R(\gamma) = 0, \psi = \text{const} \) and each leaf is totally umbilical at each point \( (\Pi_{AB} = 0) \). The IMCF condition gives \( p = \text{const} \). Showing spherical symmetry for each one of the leaves follows easily from the condition \( R(\gamma) = 0 \). Huisken and Ilmanen are able to show the same results using only the weak flow.

Huisken and Ilmanen’s proof has interesting side consequences. The most notable one is that the Bartnik capacity is non-zero for any domain which is not locally isometric to Euclidean space. The Bartnik capacity [35] is defined for 3-dimensional Riemannian manifolds \((\Omega, \gamma)\) with no boundary and compact metric closure. Such a domain is called admissible if it can be isometrically embedded into a complete and connected asymptotically flat Riemannian three-dimensional manifold \((\Sigma, \gamma)\), with non-negative Ricci scalar, a boundary which is either empty or a minimal surface, and free of any other minimal surfaces. \((\Sigma, \gamma)\) is called an admissible extension. The Bartnik capacity is the infimum of the ADM masses of all possible extensions of \((\Omega, \gamma)\). By the positive mass theorem, it is immediately non-negative and it was conjectured in [35] to be positive if \((\Omega, \gamma)\) is not a subset of Euclidean space. Huiskens and Ilmanen used the weak IMCF to reach a slightly weaker conclusion, namely that if the Bartnik capacity is zero, then \((\Omega, \gamma)\) is locally isometric to Euclidean space. The idea of the proof is to take a point \( p \) in \( \Omega \) where the metric is non-flat and any admissible extension \((\Sigma, \gamma)\). Consider the weak IMCF starting at \( p \). For some small enough value of the flow parameter the corresponding leaf has a positive Geroch mass and is shown to belong to \( \Omega \) irrespectively of the extension. The Penrose inequality therefore gives a positive lower bound for the ADM mass independent of the extension.

6.2. Bray’s proof
H.L. Bray was able [31] to prove the Riemannian Penrose inequality for arbitrary outermost minimal surfaces, with no restriction of connectedness. Bray’s proof also uses geometric flows in an essential way, but instead of flowing surfaces in a fixed Riemannian manifold, as in Huiskens and Ilmanen’s proof, a flow of metrics with non-negative Ricci scalar is used. The idea is to modify the metric so that the horizon with respect to the new metrics has the same area as the original one while the total ADM mass does not increase. If moreover, the flow has the property that outside the outermost horizon the Riemannian manifolds approach, in a suitable sense, the \( r > m/2 \) portion of the Schwarzschild initial data set (23) then the Penrose inequality would follow as Schwarzschild saturates the inequality.

The first step in Bray’s proof is to reduce the class of metrics to consider. Schoen and Yau [23] proved that given an asymptotically flat Riemannian manifold \((\Sigma, \gamma)\) with \( R \geq 0 \), and any positive number \( \epsilon \), there always exists an asymptotically flat metric \( g_\epsilon \) on \( \Sigma \) which is conformally flat outside a compact sect \( K \), i.e. \( g_\epsilon = u_\epsilon^4 \delta \) on \( \Sigma^\infty \equiv \Sigma \setminus K \), where \( u_\epsilon : \Sigma^\infty \rightarrow \mathbb{R} \) is a positive function approaching a positive constant at infinity, and which fulfills the following properties

\[
\begin{align*}
M_{ADM}(\gamma) - M_{ADM}(g_\epsilon) &\leq \epsilon, \\
\frac{g_\epsilon(X, X)}{\gamma(X, X)} - 1 &\leq \epsilon, \quad \forall X \in T_x \Sigma \quad \text{and} \quad \forall x \in \Sigma, \\
R(g_\epsilon) &\geq 0.
\end{align*}
\]

i.e. the total mass changes within \( \epsilon \) and the metric itself changes pointwise at most by \( \epsilon \) (in
the sense that unit vectors change their norm at most by $\epsilon$, all this while still maintaining a non-negative Ricci scalar. A metric which is conformally flat on the asymptotic end $\Sigma^{\infty}$ is called \textit{harmonically flat}.

This result implies that in order to prove the Riemannian Penrose inequality it is sufficient to consider harmonically flat metrics. Indeed, if there were a counterexample to the Penrose inequality, i.e. a metric with total mass being strictly smaller than the appropriate function of the area of its outermost minimal surface, we could find a harmonically flat metric which still has an outermost minimal surface with total area that differs at most in $\epsilon$ with the previous one. The total ADM mass is changed by at most a positive constant times $\epsilon$. By choosing $\epsilon$ small enough, it would follow that the Penrose inequality is violated also by a harmonically flat space.

Bray’s version of the Penrose inequality deals with three-dimensional asymptotically flat, Riemannian manifolds $(\Sigma, \gamma)$ with one or several asymptotic ends (in the latter case, one is selected to define the concept of “outer” and to compute the mass) and containing an outer minimizing horizon $S_0$ (in the case of several ends the existence of such a horizon is guarantied). A horizon is a smooth compact surface with vanishing mean curvature which is a boundary of an open set $\Omega$ containing all possible asymptotic ends on $(\Sigma, \gamma)$ except the one that has been selected. It is called outer minimizing if it has less or equal area than any other surface $S$ enclosing it, where “enclosing” means that $S$ is the boundary of some open set containing $\Omega$.

Bray’s theorem states $M_{\text{ADM}} \geq \sqrt{\frac{2\pi}{3\alpha}}$ irrespectively of any connectedness condition on $S_0$.

Bray assumes without loss of generality that $\gamma$ is harmonically flat and restricts attention to the exterior of $\Omega$, which we call $\Sigma_0$, so that $\partial \Sigma_0 = S_0$. The flow of metrics is defined via a conformal rescaling $g_t = u_t^{\gamma}\tilde{g}$ depending on a parameter $t$. The function $u_t$ is defined via an elliptic equation for $v_t = \frac{du_t}{dt}$ together with the initial value $u_0 = 1$. $v_t$ is defined by solving the Dirichlet problem

$$
\begin{align*}
\Delta_{\gamma}v_t &= 0 \quad \text{on } \Sigma_t, \\
v_t &= 0 \quad \text{on } \Sigma_0 \setminus \Sigma_t, \\
v_t|_{\Sigma_t} &= 0, \\
v_t &\to -e^{-t} \quad \text{at } \infty
\end{align*}
$$

where $S_t$ is the outermost minimal area surface enclosing $S_0$ with respect to the metric $g_t$ and $\Sigma_t$ is the exterior of $S_t$ in $\Sigma_0$. Since $S_t$ is defined using the metric one is trying to construct, it is not obvious a priori that such problem admits a solution. Bray uses a discretization argument whereby the function $u_k(x)$, for fixed value of $x$, is discretized as a continuous piecewise linear function in $t$, where the jumps in the derivatives occur at fixed intervals of length $\epsilon$. The slope of $u_t$ on the $(k+1)$-th interval, i.e. on $[k\epsilon, (k+1)\epsilon)$, $k \in \mathbb{N}$ is defined by solving the Laplace equation with respect to the metric $\gamma$ with vanishing boundary conditions on surfaces $S_k^t$ and approaching a suitable constant at infinity. At $t = (k+1)\epsilon$, the boundary surface $S_{k+1}^t$ is defined simply as the outermost minimal area enclosure of $S_k^t$ with respect to the metric $u_k^{\gamma}$. Thus, while the metric is changed continuously of $t$, the boundaries $S_k^t$ change only at discrete values. Existence follows directly from a simple inductive argument. This construction gives existence of a flow of metrics depending on $\epsilon$. It is an important ingredient of Bray’s proof to show that a suitable limit $\epsilon \to 0$ exists, thus defining the flow of metrics in (24).

The limit is performed in two ways, first the conformal factors are seen to converge when $\epsilon \to 0$ to a locally Lipschitz function $u_t$. This already defines the metric $g_t$ and the surfaces $S_t$ as before. On the other hand the collection of surfaces $S_k^t$ are seen to converge, in Hausdorff distance sense, to a collection of surfaces $\{S_\ast(t)\}$. For each $t$ this collection may in general have several elements. The surfaces $S_t$ and $\{S_\ast(t)\}$ are however, closely related. Except for $t$ in a countable set $J$, they must coincide and $S_t$ is continuous in $t$. For $t \in J$, the left and right
limits of $S_t$ may be different, but the right limit always encloses the left limit. Moreover, for $t_2 > t_1$, $S_{t_2}$ encloses $S_{t_1}$. Thus, the flow of surfaces $S_t$ is everywhere outward and may jump, also outwards, at a countable number of places.

In addition, and this is crucial for the proof of the Penrose inequality, the area of $S_t$ is constant for all $t$, even at the jumps. Except for the jumps, this constancy can be understood heuristically because $S_t$ is a closed minimal surface and hence its area does not change to first order with respect to any variation, in particular the variation with transforms $S_t$ into $S_{t+\delta}$ in the metric $g_t$. There is however a second source of change of area because the metric itself is changed. However, since $v_t = 0$ on $S_t$ by construction, the metric $g_{t+\delta}$ coincides with $g_t$, to first order, on $S_t$ and therefore the area of $S_t$ with the metric $g_t$ coincides to first order with its area with respect to $g_{t+\delta}$ from which constancy of the area follows.

The second crucial ingredient in Bray’s proof is the fact that the total mass $m(t)$ of the conformally rescaled metric $g_t$ is a non-increasing function of $t$. The proof of this fact relies on two facts. First of all, the definition of $v_t$ in (24) seems to depend on $\gamma$ and on $t$. However, as a simple consequence of how the Laplacian changes under conformal rescalings, it is not difficult to see that $v_t$ depends only on $g_t$ (i.e. a Dirichlet problem with respect to the metric $g_t$ can be written down which has $v_t$ as its only solution). Therefore, proving that $m'(t) \leq 0$, prime being derivative w.r.t $t$, amounts to proving that $m'(0) \leq 0$ because there is nothing that distinguishes $\gamma = g_0$ from any other metric in the flow. For $t = 0$, the function $v_0$ is just minus the Green function of $g_0$ defined as

$$
\begin{align*}
\Delta_{g_0} G(S_0) &= 0 \quad \text{on } \Sigma_0, \\
G(S_0)\big|_{S_0} &= 0, \\
G(S_0) &\to 1 \quad \text{at } \infty.
\end{align*}
$$

The total mass for a harmonically metric $u^4\delta$, can be directly computed from the first two leading orders of $u$ at infinity. Namely, if $u = a + b/(2r) + O(1/r^2)$ then the ADM mass is $m = ab$. Expanding the Green function $G$ in the asymptotic region as $G = 1 - c/(2r) + O(1/r^2)$, it follows directly from its definition that $m(t) = M + t(c - 2M) + o(t)$ where $M$ is the ADM mass of $g_0$. Thus, the mass will not be increasing provided $c \leq 2M$. This turns out to be a general property of the Green function $G(S_0)$ of any asymptotically flat metric when the boundary $S_0$ is a minimal surface. This result, also proven by Bray [31], has independent interest and has been already used in another context by P. Miao [36] to characterize the Schwarzschild initial data as the only static initial data with a non-empty minimal boundary.

The proof of $c \leq 2M$ relies on the positive mass theorem and the idea is related to the method used by Bunting and Masood-ul-Alam [37] to prove uniqueness of the Schwarzschild black hole. The idea is to double $(\Sigma_0, g_0)$ across its boundary $S_0$ (i.e. take two copies and identify the boundaries) to define a new manifold $\hat{\Sigma}$. Define also a function $\Phi(x) = \frac{1 + G(x)}{2}$ on one of the copies and $\Phi(x) = \frac{1 - G(x)}{2}$ on the other copy. This defines a function which is harmonic away from $S_0$ (which remains a minimal surface) and which approaches one on one asymptotic end and zero on the other asymptotic end. This function is also $C^1$ everywhere. Consider the conformally rescaled metric $\hat{g} = \Phi^4 g_0$ on $\hat{\Sigma}$. The asymptotic behaviour of $G$ near the infinity where it vanishes shows that $(\hat{\Sigma}, \hat{g})$ admits a one-point compactification there, thus defining a complete asymptotically flat Riemannian manifold. Furthermore, from the fact that $g_0$ has non-negative Ricci scalar, the same happens to $\hat{g}$. If this manifold was smooth, the positive mass theorem could be invoked to conclude that the mass of $\hat{g}$ is non-negative and zero if and only if $(\hat{\Sigma}, \hat{g})$ is flat Euclidean space. From the way the mass changes under conformal rescalings, it would follow $c \leq 2M$ with equality if and only if $(\Sigma_0, g_0)$ is flat. However, the metric $\hat{g}$ is not smooth across $S_0$ and some approximation argument is needed whereby the surface $S_0$ is replaced by a small cylindrical neck that allows for a smoothing of the metric.
In fact, with related techniques, Miao [38] has extended these results and has proven that the positive mass theorem holds for complete, asymptotically flat Riemannian manifolds \((\Omega, \gamma)\) of dimension \(n \leq 7\) with non-negative Ricci scalar so that the metric is non-smooth across a compact codimension one hypersurface \(S\) provided (i) this hypersurface separates \(\Omega\) into two pieces \(\Omega^\pm\) with \(\Omega^+\) unbounded, (ii) the induced metric on \(S\) from both sides coincide and (iii) the mean curvature with respect to the outer direction satisfies the inequality

\[
p(S, \gamma_-) \geq p(S, \gamma_+),
\]

(25)

where \(\gamma_\pm\) is the metric \(\gamma\) restricted to \(\Omega^\pm\). In the case considered by Bray, the metrics are \(\hat{g}_\pm = [(1 \pm G)/2]^4 g_0\). Under a conformal rescaling \(\hat{g} = u^4 g\), the mean curvature \(p\) of a hypersurface changes according to \(\hat{p} = (p + 4\tilde{n}(u))/u^2\), where \(\tilde{n}\) is the unit normal with respect to which \(p\) is calculated. Applying this to the metrics \(\hat{g}_\pm\) and using that \(S_0\) is minimal, it follows \(p(S_0, \hat{g}_-) = p(S_0, \hat{g}_+)\) and therefore the total mass must be positive and zero only for Euclidean space, as claimed. In particular \(c \leq 2M\).

It is worth mentioning that another related positive mass theorem for metrics which are smooth on \(\Omega^+\) and on \(\Omega^-\), only Lipschitz across \(S\) but such that equality in (25) holds has been proven by Shi and Tam [39] for spin manifolds of arbitrary dimension (this is no restriction in three dimensions) using spinor techniques.

Once constancy of area and non-increasing of the mass is shown, Bray’s proof of the Penrose inequality requires one final ingredient, namely that the metric outside \(S_t\) approaches the Schwarzschild metric in a suitable sense. The surfaces \(S_t\) flow outwards in \(\Sigma_0\). In fact, Bray proves that \(S_t\) encloses, for large enough \(t\) any bounded set of \(\Sigma_0\). Thus, the manifold \(\Sigma_t\) shrinks and it may seem that it disappears in the limit. However, due to the asymptotic condition \(u_t \to e^{-t}\) at infinity which also tends to zero when \(t \to \infty\), the domain \((\Sigma_t, g_t)\) stays unbounded for all \(t\). The idea is that the factor \(e^{-t}\) expands the domain for an ever shrinking region back to a “fixed” one. More precisely, Bray shows that for each \(\epsilon > 0\) there exists a \(T\) so that for \(t > T\) there exists a diffeomorphism \(\Phi_t\) between \((\Sigma_t, g_t)\) and a fixed Schwarzschild manifold of mass \(m_f\) outside its horizon such that both metrics are \(\epsilon\)-close to each other (in the sense that the length of any unit vector w.r.t \(g_0\) has length in \(1 - \epsilon, 1 + \epsilon\) w.r.t \(\Phi^*(g_{sch})\)). Moreover, \(m(t)\) and \(m_f\) also differ at most by \(\epsilon\). This concludes the proof of the inequality

\[
M \geq \sqrt{\frac{|S_0|}{16\pi}}
\]

(26)

for the original metric. As Bray points out, no property from \((\Sigma, g_0)\) inside \(S_0\) is in fact used in the proof, therefore establishing (26) also for manifolds with boundary \(S_0\), provided this is an outermost minimal surface.

7. On the general Penrose inequality
The validity of the Penrose inequality for arbitrary initial data sets (with non-zero second fundamental form) is completely open. The proofs by Huisken and Ilmanen and Bray of the Riemannian Penrose inequality use positive definite metrics with non-negative Ricci scalar as well as minimal surfaces. The proofs therefore apply not only in the time-symmetric case but also for maximal hypersurfaces, \(\text{tr}_\gamma A = 0\) provided the marginally trapped surface \(S\) in \((\Sigma, \gamma, A)\) is also a minimal surface, i.e. \(A^{ij} m^i m^j = 0\) on \(S\). However, these proofs do not settle the general Penrose inequality even in this case, because they involve the total energy instead of the total mass. In the time symmetric case they of course coincide.

7.1. Null shells of dust
Probably the simplest situation with no symmetries where the non-time-symmetric Penrose inequality can be studied is Penrose’s original example involving a collapsing shell of null dust
in Minkowski space. From causality theory, and assuming that the imploding shell has no self-intersections in the past, the metric inside the shell is just Minkowski space. A convenient way of ensuring regularity of the null hypersurface to the past is that, on certain instant of Minkowski time \( t = t_0 \) the shell defines a convex surface. The surfaces \( S_t \equiv N \cap \{ t = \text{const} \} \), where \( N \) is the null hypersurface of the shell, can be regarded, after an obvious identification of different constant time hyperplanes, as the surface at distance \( (t_0 - t) \) from \( S_{t_0} \) towards its exterior. The distance level sets from a convex surface stay convex to its exterior, from which regularity of \( N \) to the past of \( t_0 \) follows. Outside the null shell the metric is no longer flat, as gravitational waves are emitted by the collapsing dust. Towards the future, \( N \) will obviously become singular at the first focal point of the inner normal rays of \( S_{t_0} \). There a spacetime singularity will form.

Penrose devised this physical process as a potential counterexample of the now called Penrose inequality. The fundamental simplification of this problem is that the inequality can be translated into an inequality directly in Minkowski space, as follows [1, 40, 41, 42]. Choose arbitrarily a Minkowskian time \( t \) in the interior of the shell and define \( \xi = -dt \). Let \( \tilde{I}^- \) be the future null tangent to \( N \) normalized to \( \xi(\tilde{I}^-) = -1 \). Take any closed, spacelike surface embedded in \( N \) and let \( \tilde{I}^+ \) be its future null normal satisfying \( (\tilde{I}^+, \tilde{I}^-) = -2 \), as in Sect. 2. The energy momentum of the spacetime is a distribution supported on \( N \) and reads \( T_{\alpha\beta} = -8\pi\mu l^-_{\alpha}l^-_{\beta} \delta \), where \( \mu \) is the energy density of the shell and the Dirac \( \delta \) is defined with respect to the volume form \( da \) induced by the normal \( l^-_{\alpha} \) to \( N \), i.e. \( l^-_{\mu} da = \eta_{\mu\alpha\beta\gamma} e_1^\alpha e_2^\beta e_3^\gamma \) with \( e_\alpha^\mu = \frac{\partial x^\alpha}{\partial y^\mu} \) and \( y^\mu \rightarrow x^\alpha(y^\mu) \) a coordinate expression of the embedding defining \( N \) (see e.g. [43] for details).

The null expansion \( \theta_+ \) jumps across \( N \). This can be determined using the Raychaudhuri equation (9). One way of doing it is extending the null expansion \( \theta_+ \) to a (discontinuous) function on the spacetime by extending the null vector \( \tilde{I}_+ \) to a geodesic null congruence and taking its divergence. This defines a distribution \( \theta_+ = \theta_+E \theta_+ + \theta_+(1 - \theta) \), where \( \theta \) acts on tests functions as integration on the exterior domain outside the shell and \( I(E) \) stand for interior (exterior) of the shell. Since \( \partial_\mu \theta = -l^-_{\mu} \delta \), the derivative of \( \theta_+ \) along \( \tilde{I}_+ \) gives

\[
l^+_{\mu} \partial_\mu \theta_+ = l^+_{\mu} \partial_\mu \theta_+^E (1 - \theta) + \theta_+^l \partial_\mu \theta^l_+ + 2 [\theta_+] \delta,
\]

where the jump \( [\theta_+] = (\theta_+^E - \theta_+^l)|_{N} \). The Raychaudhuri equation, which in this case is a distributional equation, has a singular part supported on \( N \) only through the term \( -\text{Ric}(\tilde{I}_+, \tilde{I}_+) \delta \equiv 32\pi \mu \delta \). The jump must therefore be \( \theta_+^E = \theta_+^l \), \( |_{N} = 16\pi \mu \). The incoming null expansion \( \theta_- \) of the shell is continuous across the shell and therefore negative everywhere. The surface \( S \) will therefore be marginally trapped iff \( \theta_+^E = 0 \) (and will in fact be future marginally trapped). Assume such an \( S \) exists along the shell. Although it is not obvious that there cannot be another tube of future marginally trapped tube shielding \( S \) from infinity, it is reasonable to expect that the Penrose inequality should hold for \( S \), i.e \( M_B \geq \sqrt{\frac{16\pi S}{3\pi}} \) where \( M_B \) is the Bondi mass of past null infinity at the cut defined by \( N \). Using the conservation equation \( \nabla_\alpha T^{\alpha\beta} = 0 \), it follows that the integral \( \int_S \mu dS \) does not depend on the cut \( S \) of \( N \). Evaluating this integral at past null infinity gives precisely the Bondi mass (this can be easily shown for instance using the Hawking mass introduced in Sect. 7.4). Thus, the Penrose inequality in this setting becomes

\[
\int_S \theta_+ |dS \geq \sqrt{16\pi S}.
\]

(27)

The remarkable property of this expression is that it involves only closed surfaces \( S \) in Minkowski space with the only restriction that the null hypersurface generated by outer past null normal geodesics is regular everywhere. For surfaces \( S \) lying on the past null cone of a point this inequality becomes \( 4\pi \int_{S^2} r^2 |\eta_{S^2} \leq \int_{S^2} (r^2 + |D r|^2 / r) \eta_{S^2} \) where \( r \) is any positive function on the sphere \( S^2 \) endowed with the unit round metric. This inequality was first written down by
Penrose [1] and proven by Tod [40] using a suitable Sobolev inequality (see [44] for an alternative proof).

If $S$ lies in a spacelike hyperplane in Minkowski, then $\theta_\flat$ equals the mean curvature $p$ of $S$ as a subset of Euclidean space and the inequality becomes $\int_S pdS \geq \sqrt{16\pi|S|}$. As first noticed by Gibbons in his Ph.D. thesis, this inequality is exactly the Minkowski inequality for convex bodies, see e.g. [45]. As proven by Trudinger [46], the Minkowski inequality $\int_S pdS \geq \sqrt{16\pi|S|}$ holds for arbitrary mean convex bodies in Euclidean space. Using Trudinger’s inequality and a normal projection of $S$ into a constant time hyperplane in Minkowski space has led Gibbons [42] to claim the general validity of the inequality (27) and therefore of the Penrose inequality for collapsing shells of null dust.

7.2. Jang equation

Schoen and Yau’s proof of the positive mass theorem proceeded in two steps. First the positive mass theorem was proven for Riemannian manifolds with non-negative Ricci scalar. The general case $(\Sigma, \gamma, A)$ with arbitrary second fundamental form was treated by modifying the metric $\gamma$ through a transformation invented by Jang [47], namely $\hat{\gamma}_{ij} = \gamma_{ij} + \partial_i f \partial_j f$, where $f$ solves a quasilinear elliptic equation now called Jang’s equation. This transformation has the property that $\hat{\gamma}$ satisfies a suitable positivity property that allows to solve the Lichnerowicz equation and conformally transform $\hat{\gamma}$ into a metric of vanishing scalar curvature, to which the previous positive mass theorem can be applied. The proof is however involved because the Jang equation does not admit regular solution when the initial data set $(\Sigma, \gamma, A)$ contains marginally trapped surfaces, but the idea can nevertheless be made to work [23].

A natural question, already asked in [31], is whether a similar idea can be applied to prove the general Penrose inequality. Notice that the Penrose inequality has been proven in the Riemannian setting so at present the status is similar to the situation of the proof of the positive mass theorem after its Riemannian proof [22]. This idea has been analyzed recently by Malec and O’Murchadha [48] with negative results. The argument used in [48] is based on the observation that if the Jang equation could be used to prove the general Penrose inequality, it should be able to do it in particular in the spherically symmetric case. The Jang equation becomes a simple ODE in that case. For the method to work, the marginally trapped surface in the initial data must be transformed into a minimal surface after the conformal rescaling. Although a complete proof of impossibility was not given, the results in [48] strongly point towards the fact that the Jang equation method cannot handle the spherically symmetric case and therefore neither the general one.

7.3. Spinor techniques: Ludvigsen-Vickers and Bergqvist approaches

Shortly after Shoen and Yau’s proof of the positive mass theorem, Witten [49] proposed a completely different method using spinors (see [50] for a rigorous version of Witten’s ideas). It is natural to ask whether spinorial techniques can be also successfully applied to the Penrose inequality. After all, the spinor techniques had been successfully extended to prove the positive mass theorem in the presence of black holes (i.e marginally trapped surfaces) in the initial data [51] (see [52] for a rigorous proof). The main difficulty lies in finding suitable boundary conditions on Witten’s equation on the boundary of the black holes so that the boundary term arising by integrating the Schrödinger-Lichnerowicz [53] identity can be related to the area of the black hole. A serious attempt to achieve this in the Riemannian case (i.e. with vanishing fundamental form and minimal surfaces) is due to M. Herzlich [54], who however obtained a Penrose-like inequality involving not just the total mass and area of the minimal surface but also a Sobolev type constant. Depending on the space, Herzlich’s inequality may be stronger or weaker than the Penrose inequality.
Spinor techniques have also been used to prove (or attempt to prove) positivity of the Bondi mass at past null infinity \[55\]. All these arguments however have been criticised in \[56\], where the first correct proof of positivity of the Bondi energy is claimed. Regarding the Penrose inequality, Ludvigsen and Vickers \[57\] gave a spinor-based argument leading to the inequality 
\[
M_B \geq \sqrt{|S|/16\pi}
\]
for a future marginally trapped surface \(S\) in spacetimes where the outer past null cone of \(S\) could be extended to past null infinity. This is, of course, a strong global assumption on the spacetime. However, Bergqvist \[58\] found a a gap in the proof which at present remains open. Bergqvist also reformulated Ludvigsen and Vickers’ argument so that all spinors could be completely dispensed of. The idea is, in some sense, similar to the IMCF flow for the Geroch mass and is based on using a functional on spacelike closed surfaces \(S_\mu\) constructed as follows. Start with a closed spacelike surface \(S\) and let \(\vec{l}\) and \(\vec{k}\) be past directed null normals to \(S\) partially fixed by the normalization \(\vec{l} \cdot \vec{k} = -2\). Assume also that \(\vec{l}\) points outwards of \(S\) in the sense that the null geodesics starting on \(S\) with tangent vector \(\vec{l}\) extend to infinite values of the affine parameter and intersect \(\mathcal{I}^-\) in a cut (this, of course, entails a strongly global assumption on the spacetime). Let \(\mu\) be the affine parameter of this geodesic with \(\mu = \mu_0\) on \(S\) where \(\mu_0\) is a constant to be chosen later. The surfaces \(\{\mu = \text{const}\}\) define a collection of spacelike surfaces all lying in \(\mathcal{N}\), the outer past null cone of \(S\). The Bergqvist mass is defined on each leave \(\{S_\mu\}\) of this foliation as

\[
M_b(S_\mu) = \frac{1}{16\pi} \int_{S_\mu} \theta_k \eta S_\mu + 4\pi \chi(S_\mu) \mu.
\]

Using the expressions (4) and (10), it is immediate to obtain

\[
\vec{l} \left( \int_{S_\mu} \theta_k \eta S_\mu \right) = \int_{S_\mu} \left( \text{Ein}(\vec{l}, \vec{k}) + 2S_A S^A \right) \eta S_\mu - 4\pi \chi(S_\mu),
\]

where the Gauss-Bonnet theorem \(\int_S R(h) \eta S = 4\pi \chi(S)\) has been used. We therefore get

\[
\vec{l}(M_b) = \frac{1}{16\pi} \int_{S_\mu} \left( \text{Ein}(\vec{l}, \vec{k}) + 2S_A S^A \right) \eta S_\mu \geq 0,
\]

provided the dominant energy condition is satisfied in the spacetime. Assuming that \(\mu_0\) can be chosen so that \(\theta_k = -2/\mu + O(1/\mu^3)\), i.e. without term in \(\mu^{-2}\), then the Bergqvist mass can be seen to approach the Bondi mass. If the value of \(M_B\) could be somehow related to the area of the initial marginally trapped surface, a Penrose inequality would follow. Bergqvist \[58\] shows that this can indeed be done but only after assuming that the metric of the surfaces \(S_\mu\) approach the round metric, after a suitable constant rescaling, at infinity. This a strong restriction and it is not clear whether the argument can be accommodated to deal with the general case. The idea however remains interesting and deserves further investigation.

7.4. Uniformly expanding flows
Huisken and Ilmanen’s proof is based on the observation that the Geroch mass is monotonic under inverse mean curvature flows. These flows involve surfaces in three-dimensional Riemannian manifolds. The Geroch mass is a particular case for time-symmetric initial data sets of a similar object defined for surfaces in spacetime, the Hawking mass \[59\]. This object also approaches the ADM mass for suitable large surfaces and gives the right hand side of (14) on any marginally trapped surface. Thus, in principle it is possible that the Hawking mass may also play an important role in proving the general Penrose inequality. The first step is of course to check whether the Hawking mass is monotonic under suitable flows. Monotonicity of
the Hawking mass along null directions was analyzed by Hayward [60]. More recently, the rate of change of the Hawking mass along an IMCF in an arbitrary initial data set \((\Sigma, \gamma_{ij}, A_{ij})\) was studied in [61]. A general monotonicity theorem for spacetime flows (i.e. flows where the two-dimensional surface is varied in spacetime) has been recently been obtained [62]. This spacetime approach, in particular, allows to unify and extend the null and spacelike approaches mentioned above. The following discussion is based on this paper.

The Hawking mass is a functional on closed spacelike surfaces \(S\) embedded in spacetime defined by

\[
M_H(S) = \sqrt{\frac{|S|}{16\pi}} \left( \frac{\chi(S)}{2} - \frac{1}{16\pi} \int_S H^2 \eta_S \right),
\]

(28)

where \(H^2 = (\vec{H} \cdot \vec{H})\) and \(\vec{H}\) is the mean curvature vector of \(S\) as a submanifold of the spacetime.

In order to find how \(M_H\) is changed under an arbitrary variation vector \(\vec{\xi}\), supposed orthogonal to \(S\), we need to introduce the tangential derivative \(\nabla^\perp\) on normal vectors \(\vec{V}\) along \(S\) as

\[
\nabla^\perp \vec{V} = (\nabla \vec{V})^\perp.
\]

(29)

After a somewhat involved calculation, the variation of the Hawking mass along \(\vec{\xi}\) is obtained in [62] to be

\[
\frac{dM_H(S_\lambda)}{d\lambda} = \frac{1}{16\pi} \sqrt{\frac{|S_\lambda|}{16\pi}} \left[ \frac{4\pi \chi(S_\lambda)a(\lambda)}{16\pi} + \int_{S_\lambda} \left( 2(\tilde{K}^{AB} \cdot \vec{H})(\tilde{K}^{AB} \cdot \vec{\xi}) - \left( \tilde{\vec{\xi}} \cdot \vec{H} \right) \right) \right] H^2 + 2h^{AB} \left( \partial_t(\tilde{\vec{\xi}}, e_A) e_B \cdot \tilde{H} \right) + 2tr_{S_\lambda} \left( \tilde{H} \cdot \nabla^\perp \nabla^\perp \tilde{\vec{\xi}} \right) \eta_{S_\lambda},
\]

(30)

where

\[
a(\lambda) \equiv \frac{\int_{S_\lambda} \left( \tilde{\vec{\xi}} \cdot \vec{H} \right) \eta_{S_\lambda}}{|S_\lambda|}.
\]

This expression can be elaborated further by introducing a Hodge dual operation on the normal space to the surface \(S\), which is a Lorentzian two dimensional vector space at each point. For any normal vector field \(\vec{W}\), we define the dual by

\[
W^* = \frac{1}{2} \eta^a_\beta \rho_\delta W^\beta e^\gamma_A e^\delta_B \eta^{AB},
\]

where \(\eta\) is the spacetime volume form and \(\eta_{S_\lambda}\) is the volume form of \(S\). It follows \(W^{**} = \vec{W}\) and \((\vec{W} \cdot \vec{U}^*) = -(\vec{W}^* \cdot \vec{U})\). Therefore any normal vector is orthogonal to its dual and both have opposite norms. When \(\vec{\xi}\) is non-null we decompose it in the null basis \(\{\vec{l}_+, \vec{l}_-\}\)

\[
\vec{\xi} = \vec{\xi}_+ + \vec{\xi}_- \quad \text{where} \quad \vec{\xi}_+ = A\vec{l}_+, \quad \vec{\xi}_- = B\vec{l}_- \quad \text{for some functions} \quad A, B.
\]

(31)

Define \(\psi\) by \(\xi^a \xi_\alpha = e^{2\psi} = -4AB \neq 0 \ (\epsilon = \pm 1)\) and introduce the one-form on \(S_\lambda\)

\[
U_C = \frac{1}{4} \left( \frac{\vec{l}_+ \cdot \nabla_C \vec{\xi}_-}{B} - \frac{\vec{l}_- \cdot \nabla_C \vec{\xi}_+}{A} \right).
\]
Then, the variation of the Hawking mass can be rewritten as

\[
\frac{dM_H(S_\lambda)}{d\lambda} = \frac{1}{8\pi} \sqrt{\frac{S_\lambda}{16\pi}} \int_{S_\lambda} \left[ \text{Ein}(\vec{H}^*, \vec{\xi}^*) + 8\pi \Theta^T(\vec{H}^*, \vec{\xi}^*) + (|U|^2 + |D\psi|^2)(\vec{\xi} \cdot \vec{H}) - (2U \cdot D\psi)(\vec{\xi} \cdot \vec{H}^*) - D \cdot U(\vec{\xi} \cdot \vec{H}^*) + \left[ (\vec{\xi} \cdot \vec{H}) - a(\lambda) \right] \left( \Delta \psi - \frac{1}{2} R(h_\lambda) + \frac{1}{4} H^2 \right) \right] \eta_{S_\lambda},
\]

(32)

where \( D \cdot U \) denotes the divergence of \( U \) and

\[
8\pi \Theta^T(\vec{H}^*, \vec{\xi}^*) = \left( \vec{H}_{AB} \cdot \vec{H}^* \right) \left( \vec{\xi}^* \cdot \vec{H}^{AB} \right) - \frac{1}{2} \left( \vec{H}_{AB} \cdot \vec{H}^{AB} \right) \left( \vec{\xi}^* \cdot \vec{H}^* \right).
\]

(33)

Using (30) and (32) it can be seen that there are at least two cases where the Hawking mass can be made monotonic. In both cases the dominant energy condition is required.

The first one involves spacelike flow vectors, spacelike or null mean curvature satisfying

\[
(\vec{\xi} \cdot \vec{H}) \geq 0
\]

and imposes two restrictions on the flow vector. Firsty, \( (\vec{\xi} \cdot \vec{H}) = c(\lambda) \), which we can call dual inverse mean curvature condition and secondly the elliptic equation

\[
\Delta \psi - \frac{1}{2} R(h_\lambda) + \frac{1}{4} H^2 = \alpha(\lambda).
\]

(34)

The constant \( \alpha(\lambda) \) is determined by the condition that (34) admits solutions, i.e. \( \alpha(\lambda)|_{S_\lambda} = \int_{S_\lambda} \left( -\frac{1}{2} R(h_\lambda) + \frac{1}{4} H^2 \right) \eta_{S_\lambda} \).

The second situation where monotonicity of the Hawking mass is achieved assumes that \( \vec{H} \) is spacelike and restricts \( \vec{\xi} \) to be of the form

\[
\vec{\xi} = \frac{a(\lambda)}{H^2} \left( \vec{H} - c(\lambda) \vec{H}^* \right),
\]

(35)

where \( a(\lambda) > 0 \) and \( |c(\lambda)| \leq a(\lambda) \). The case with \( c(\lambda) = \pm a(\lambda) \) corresponds to a null variation vector and corresponds exactly to the flow studied by Hayward [60]. The case \( |c(\lambda)| < a(\lambda) \) corresponds to a spacelike flow vector and therefore can be rephrased in terms of initial data sets. Monotonicity of the Hawking mass in this setting was first considered in [61]. The case \( c(\lambda) = 0 \) can be naturally called inverse mean curvature flow vector and monotonicity of the Hawking mass in this case was first mentioned in [30] and studied in detail by Frauendiener [63]. In [62] the flows (35) have been collectively called uniformly expanding flows since the mean curvature along \( \vec{\xi} \) and its dual \( \vec{\xi}^* \) are both constant.

With this monotonicity property, the situation regarding the general Penrose inequality can be compared to the status of the Riemannian Penrose inequality after Geroch’s heuristic argument. It is conceivable that the uniformly expanding flows might be useful for the proof of the general Penrose inequality. However, many issues remain open. For instance, in a spacetime formulation, the jumps that occurred in the Riemannian setting will remain, but it is completely unclear how and when the jumps should take place, even from a purely heuristic point of view. Moreover, the inverse mean curvature flow is a parabolic equation in the Riemannian setting, so that local existence was granted. This is not so for the uniformly expanding flows because they form a so-called forward-backward parabolic system, for which no local existence theory is known (see [62] for a proof and discussion of this point). Nevertheless a weak formulation of the uniformly expanding flows does exist [62]. Whether this can give useful hints on how to define the jumps and solve local and global existence problems is at present unclear.
8. Concluding remarks
The Penrose inequality can be strengthened when suitable matter fields are present. The most relevant case being asymptotically flat spacetimes with electromagnetic fields having all the charge sources hidden from infinity inside the marginally trapped surface. Penrose’s heuristic argument leading to the Penrose inequality can be repeated taking into account that the final state is the Kerr-Newman black hole. The explicit inequality between area and mass for this metric leads to the following Penrose inequality

\[ \sqrt{\frac{|S_{\text{min}}|}{16\pi}} \leq \frac{1}{2} \left( M_{\text{ADM}} + \sqrt{M_{\text{ADM}}^2 - q^2} \right), \]  

where \( q \) is the total charge of the spacetime. A stronger version of this inequality which is nicer looking in the sense that it bounds the total mass from below is

\[ M_{\text{ADM}} \geq \frac{1}{2} \left( \sqrt{\frac{|S_{\text{min}}|}{4\pi}} + q^2 \sqrt{\frac{4\pi}{|S_{\text{min}}|}} \right). \]  

If this second inequality were true and equality was achieved only by the Reissner-Nordström black hole, this would provide a variational characterization of the Reissner-Nordström metric in a similar way as the Riemannian Penrose inequality (and the full inequality if true) gives a minimum characterization of the Schwarzschild metric for the functional \( M_{\text{ADM}} / \sqrt{|S|} \). Remarkably, Weinstein and Yamada [64] have found a counterexample to (37) (not to (36)) by considering a suitable modification of a Papapetrou-Majumdar solution with two equal masses.

All the discussion above is restricted to four spacetime dimensions. However, the Penrose inequality makes sense in any dimension. Although no conclusive results have been obtained so far, some partial results and ideas have been put forward. The interested reader is addressed to the (most likely incomplete) list of references [65].

A second interesting generalization of the Penrose inequality deals with spacetimes which are not asymptotically flat, but e.g. asymptotically anti-de Sitter. In this case, the Geroch mass can be suitably modified and similar ideas as those used by Huisken and Ilmanen can be applied. This has led Chruściel and Simon [66], adapting previous ideas of Gibbons [67], to prove a Penrose inequality for time symmetric asymptotically anti-de Sitter initial data sets, provided a suitable smooth global solution of the level set formulation of the inverse mean curvature flow exists.

Acknowledgements
I would like to thank the organizers of the Spanish Relativity meeting in Palma for the opportunity to talk about this topic in their interesting and nicely organized congress. I thank José M.M. Senovilla and Raúl Vera for useful comments on the manuscript. Financial support from the Spanish Ministerio de Educación y Tecnología under project BFME2003-02121 and the Junta de Castilla y León, project number SA010CO is acknowledged.

References
[1] Penrose R 1973 “Naked singularities” Ann. N. Y. Acad. Sci. 224 125-34
[2] Wald R “General Relativity” Chicago University Press 1984
[3] Chruściel P T, Delay E, Galloway G J and Howard R 2001 “Regularity of horizons and the area theorem” Annales Henri Poincare 2 109-78
[4] Hawking S W and Ellis G F R 1973 The large scale structure of space-time (Cambridge Univ. Press, Cambridge)
[5] Heusler M 1996 Black hole uniqueness theorems (Cambridge Lecture Notes in Physics vol 6) (Cambridge University Press, Cambridge)
[6] Chrusciel P T 1994 “No Hair” Theorems - folklore, conjectures, results” (Differential Geometry and Mathematical Physics, Contemporary Math. vol 170) Ed. J. Beem and K.L. Duggal (American Mathematical Soc. Providence) pp 23-49
Chrusciel P T 2002 “Black holes” (Proceedings of the Tübingen conference on conformal structure of space-time, Lecture Notes in Physics vol 604) Ed J. Frauendiener and H. Friedrich (Springer) pp 61-102

[7] Bondi H, van der Burg M G J and Metzner A W K 1962 “Gravitational waves in general relativity VII. Waves from axi-symmetric isolated systems” Proc. Roy. Soc. London A269 21-52

[8] Zhang X 2006 “On the relation between ADM and Bondi energy-momenta” Adv. Theor. Math. Phys. 10 261-82

[9] Penrose R 1965 “Gravitational collapse and space-time singularities” Phys. Rev. Lett. 14 57-9

[10] Hawking S W 1967 “The occurrence of singularities in cosmology. III Causality and singularities” Proc. Roy. Soc. London A300 187-201
Hawking S W and Penrose R 1970 “The singularities of gravitational collapse and cosmology” Proc. Roy. Soc. London A314 529-48

[11] Senovilla J M M 1998 “Singularity theorem and their consequences” Gen. Rel. Grav. 30 701-848

[12] Gannon D 1975 “Singularities in non-simply connected spacetimes” J. Math. Phys. 16 2364-67

[13] Penrose R 1969 “Gravitational collapse: the role of generalrelativity” NC 1 252-76

[14] Wald R W “Gravitation and geometric analysis” Comm. Analysis and Geometry

[15] Bartnik R 2005 “Phase space for the Einstein equations” Commun. Pure and Applied Math.

[16] Bartnik R 1986 “The mass of an asymptotically flat manifold” Bartnik R 1986 “The mass of an asymptotically flat manifold” Commun. Math. Phys. 87-102

[17] Bray H L, Chrusciel P T 2004 “The Penrose Inequality” (The Einstein Equations and the Large Scale Behavior of Gravitational Fields, 50 years of the Cauchy problem in general relativity), Ed H. Friedrich and P.T. Chrusciel (Birkhäuser)

[18] Jost J 2001 Riemannian Geometry and geometric analysis, 3rd ed. (Springer Berlin)

[19] Arnowitt R, Deser S and Misner C W 1962 “The dynamics of general relativity” (Gravitation: An introduction to current research) Ed L. Witten (Wiley, New York)

[20] Bartnik R 2005 “Phase space for the Einstein equations” Comm. Analysis and Geometry 13 845-85

[21] Bartnik R 1986 “The mass of an asymptotically flat manifold” Commun. Pure and Applied Math. 39 661-93

[22] Schoen R and Yau S-T 1979 “On the proof of the positive mass conjecture in General Relativity” Commun. Math. Phys. 65 45-76
Schoen R and Yau S-T 1981 “The energy and the linear momentum of spacetimes in General Relativity” Commun. Math. Phys. 79 47-51
Schoen R and Yau S-T 1981 “Proof of the positive mass theorem II” Commun. Math. Phys. 79 231-60

[23] Kriele M and Hayward S A 1997 “Outer trapped surfaces and their apparent horizon” J. Math. Phys. 38 1593-604

[24] Horowitz G 1984 “The positive energy theorem and its extensions” (Asymptotic behavior of mass and spacetime geometry, Springer Lecture Notes in Physics vol 202) Ed. F. Flaherty, (Springer, New York)

[25] Karkowski K and Malec E 2005 “The general Penrose inequality: lessons from numerical evidence” Acta Physica Polonica B36 59-74

[26] Malec E and O Murchadha N 1994 “Trapped surfaces and the Penrose inequality in spherically symmetric geometries” Phys. Rev. D 49 6931-34

[27] Malec E and O Murchadha N 1994 “Trapped surfaces and the Penrose inequality in spherically symmetric geometries” Phys. Rev. D 49 6931-34

[28] Hayward S A 1996 “Gravitational energy in spherical symmetry” Phys. Rev. D 53 1938-49

[29] Ben-Dov I 2004 “The Penrose inequality and apparent horizons” Phys. Rev. D 70 124031

[30] Huisken G, Ilmanen T 2001 “The inverse mean curvature flow and the Riemannian Penrose inequality” J. Diff. Geom. 59 353-437

[31] Bray H L 2001 “Proof of the Riemanian Penrose inequality using the positive mass theorem” J. Diff. Geom. 59 177-207

[32] Meeks III W H and Yau S-T 1980 “Topology of three-dimensional manifolds and the embedding problems in minimal surface theory” Ann. Math. 112 441-84

[33] Geroch R 1973 “Energy Extraction” Ann. N. Y. Acad. Sci. 224 108-17

[34] Jang P S and Wald R 1977 “The positive energy conjecture and the cosmic censorship” J. Math. Phys. 18 41-4
[35] Bartnik R 1989 “New definition of quasi-local mass” PRL 62 2346-48
Bartnik R 1989 “Results and conjectures in mathematical relativity”, (Miniconference on Geometry and Physics, Proc. Centre Math. Anal.) (Austr. Nat. Univ. Canberra) pp 110-40
Bartnik R 1997 “Energy in General Relativity” (Tsing Hua Lectures on Geometry and Analysis, Hsinchu) (Internat. Press, Cambridge) pp 5-27

[36] Miao P 2005 “A remark on boundary effects in static vacuum initial data sets” Class. Quantum Grav. 22 L53-L59

[37] Bunting G L and Masood-ul-Alam A K M 1987 “Nonexistence of multiple back holes in Asymptotically Euclidean Static Vacuum Space-Time” Gen. Rel. Grav. 19 147-54

[38] Miao P 2002 “Positive Mass Theorem on Manifolds admitting Corners along a Hypersurface” Adv. Theor. Math. Phys. 6 1163-82

[39] Shi Y and Tam L 2002 “Positive mass theorem and the boundary behaviors of compact manifolds with nonnegative scalar curvature” J. Diff. Geom. 62 79-125

[40] Tod K P 1985 “Penrose quasi-local mass and the isoperimetric inequality for static black-holes” Class. Quantum Grav. 2 L65-L68

[41] Tod K P 1992 “The hoop conjecture and the Gibbons-Penrose construction of trapped surfaces” Class. Quantum Grav. 9 1581-91

[42] Pelath M A, Tod K P and Wald R 1998 “Trapped surfaces in prolate collapse in the Gibbons-Penrose construction” Class. Quantum Grav. 15 3917-34

[43] Miao P 2002 “Positive Mass Theorem on Manifolds admitting Corners along a Hypersurface” Adv. Theor. Math. Phys. 6 1163-82

[44] Bunting G L and Masood-ul-Alam A K M 1987 “Nonexistence of multiple back holes in Asymptotically Euclidean Static Vacuum Space-Time” Gen. Rel. Grav. 19 147-54

[45] Tod K P 1992 “The hoop conjecture and the Gibbons-Penrose construction of trapped surfaces” Class. Quantum Grav. 9 1581-91

[46] Trudinger N S 1994 “Isoperimetric inequalities for quermassintegrals” Ann. Inst. H. Poincaré, Sect. C, 11 411-25

[47] Jang P S 1978 “On the positivity of energy in general relativity” J. Math. Phys. 19 1152-55

[48] Malec E and Ó. Murchadha N 2004 “The Jang equation, apparent horizons, and the Penrose inequality” Class. Quantum Grav. 21 5777-88

[49] Witten E 1981 “A new proof of the positive energy theorem” Commun. Math. Phys. 80 381-402

[50] Parker T and Taubes C H 1982 “On Witten’s Proof of the Positive Energy Theorem” Commun. Math. Phys. 84 223-38

[51] Reula O 1982 “Existence theorem for solutions of Witten’s equation and nonnegativity of total mass” J. Math. Phys. 23 810-14

[52] Choquet-Bruhat Y 1984 “Positive energy theorems” (Relativity, Groups and Topology II, Proceedings of the Les Houches Summer School, June-August 1983) Ed B. S. DeWitt and R. Stora (North Holland, Amsterdam)

[53] Ludvigsen M and Vickers J A G 1983 “An inequality relating the total mass and the area of a trapped surface in general relativity” J. Phys. A: Math. Gen. 16 3349-53

[54] Ludvigsen M and Vickers J A G 1981 “The positivity of the Bondi mass” J. Phys. A: Math. Gen. 14 L389-L391

[55] Ludvigsen M and Vickers J A G 1982 “A simple proof of the positivity of the Bondi mass” J. Phys. A: Math. Gen. 15 L67-L70

[56] Reula O and Tod K P 1984 “Positivity of the Bondi energy” J. Math. Phys. 25 1004-08.

[57] Chrusciel P T, Jezierski J and Leski S 2004 “The Trautman-Bondi mass of initial data sets” Adv. Theor. Math. Phys. 8 83-139

[58] Ludvigsen M and Vickers J A G 1983 “An inequality relating the total mass and the area of a trapped surface in general relativity” J. Phys. A: Math. Gen. 16 3349-53

[59] Bergqvist G 1997 “On the Penrose inequality and the role of auxiliary spinor fields” Class. Quantum Grav. 14 2577-83

[60] Hawking S W 1968 “Gravitational radiation in an expanding universe” J. Math. Phys. 9 598-604.

[61] Hayward S A 1994 “Quasi-localization of Bondi-Sachs energy-loss” Class. Quantum Grav. 11 3037-48.
[61] Malec E, Mars M and Simon W 2002 “On the Penrose inequality for general horizons” Phys. Rev. Lett. 88 121102.

[62] Bray H L, Hayward S A, Mars M and Simon W “Generalized inverse mean curvature flows in spacetime” To appear in Communications in Mathematical Physics, Preprint gr-qc/0603014.

[63] Frauendiener J 2001 “On the Penrose inequality” Phys. Rev. Lett. 87 101101-1

[64] Weinstein G and Yamada S 2005 “On a Penrose inequality with charge” Commun. Math. Phys. 257 703-23

[65] Ida D and Nakao K 2002 “Isoperimetric inequality for higher-dimensional black holes” Phys. Rev. D 66 064026

Barrabes C, Frolov V P and Lesigne E 2004 “Geometric inequalities and trapped surfaces in higher dimensional spacetimes” Phys. Rev. D 69 101501

Gibbons G W and Holzegel G 2006 “The Positive mass and isoperimetric inequalities for axisymmetric black holes in four and five dimensions” Class. Quantum Grav. 23 6459-78

Dafermos M and Holzegel G “On the nonlinear stability of higher-dimensional triaxial Bianchi IX black holes” Preprint gr-qc/0510051

[66] Chruściel P T and Simon W 2001 “Towards the classification of static vacuum spacetimes with negative cosmological constant” J. Math. Phys. 42 1779-817

[67] Gibbons G W 1999 “Some comments on gravitational entropy and the inverse mean curvature flow” Class. Quantum Grav. 16 1677-87