On a recent proposal of the Penrose inequality

Alberto Carrasco\textdegree{} and Marc Mars\textdegree{2}
Dept. Física Fundamental, Univ. de Salamanca, Pl. de la Merced s/n, 37008 Salamanca, Spain
E-mail: 1acf@usal.es, 2marc@usal.es

Abstract. Recently Bray and Khuri \cite{BrayK} have proposed a version of the Penrose inequality in terms of a new type of horizons, the so-called generalized apparent horizons. In \cite{Carrasco} we find counter-examples to this conjecture in the Kruskal spacetime. The main technical tool for this result is the implicit function theorem. In this short note we explain in detail how this theorem can be applied in our situation.

The standard version of the Penrose inequality bounds the total ADM mass of an asymptotically flat initial data set in terms of the minimal area needed to enclose the boundary of the outer trapped region. The heuristic argument supporting its validity \cite{Penrose} is based on a number of expected properties of gravitational collapse, most notably the weak cosmic censorship conjecture. The Penrose inequality, however, is logically independent of these hypotheses and proving (or disproving) its validity would therefore give rather convincing support for the standard viewpoint of gravitational collapse.

At present, the Penrose inequality has been proven only in spherical symmetry and when the second fundamental form of the initial data set vanishes (see \cite{York} for a recent review). In a very interesting proposal to address the general case of the Penrose inequality, Bray and Khuri \cite{BrayK} have proposed a new version of the inequality which involves a novel type of surfaces called generalized trapped horizons, defined as follows. On an asymptotically flat initial data set \((\Sigma, \gamma_{ij}, A_{ij})\), consider bounding surfaces, i.e. compact and embedded surfaces \(S\) without boundary such that \(\Sigma \setminus S\) consists of two disjoint open regions \(\Omega^+\) and \(\Omega^-\). Let \(\Omega^+\) denote the region containing the asymptotically flat end. Let \(p\) be the mean curvature of \(S \subset \Sigma\) with respect to the unit normal pointing towards \(\Omega^+\), and let \(q\) be the trace of the pull-back of the second fundamental form \(A_{ij}\) of \(\Sigma\) onto \(S\). A bounding surface \(S\) is a generalized apparent horizon if \(p = |q|\). Eichmair \cite{Eichmair} has proven that, if \(\Sigma\) contains at least one bounding surface satisfying \(p \leq |q|\), then the initial data set contains a unique outermost generalized apparent horizon \(S_{\text{out}}\). This surface is \(C^{2,\alpha}\) and has the property of having less area than any other surface enclosing it. The Bray and Khuri proposal of the Penrose inequality is

\[
M \geq \sqrt{\frac{|S_{\text{out}}|}{16\pi}},
\]

where \(M\) is the ADM mass of the initial data and \(|S_{\text{out}}|\) the area of the outermost generalized apparent horizon.

As discussed in \cite{York}, there exist slices in the Kruskal spacetime such that the outermost generalized apparent horizon penetrates, at least partially, inside the domain of outer communications. It becomes natural to consider whether the conjecture (1) holds true for
such slices. In a recent paper [2] we have argued that, in fact, there exist slices of the Kruskal spacetime where the inequality (1) is violated. The main idea exploited in [2] is to consider slices $\Sigma_{\epsilon}$ which are small (but finite) perturbations of the time symmetric slice of the Kruskal spacetime and to use the implicit function theorem to locate a generalized apparent horizon $S_{\text{out}}$ near the bifurcation surface on each one of the perturbed slices $\Sigma_{\epsilon}$. An explicit calculation of the area of $S_{\epsilon}$ shows $|S_{\epsilon}| \geq 16\pi M^2$ for $\epsilon \neq 0$ small enough. In a second step, we prove that the outermost generalized apparent horizon $S_{\text{out}}$ either coincides with $S_{\epsilon}$ or else satisfies $|S_{\text{out}}| \geq |S_{\epsilon}|$. In either case, the inequality (1) is violated in the slice $\Sigma_{\epsilon}$. An important issue that has been recently put forward by M. Khuri [6] is whether the implicit function theorem can be applied to the functional $p - |q|$, since the appearance of an absolute value in $|q|$ makes the derivative potentially discontinuous. The aim of this contribution is to describe in detail how the implicit function theorem can be applied in our situation.

Consider the Kruskal spacetime of mass $M > 0$ with metric $ds^2 = \frac{2M^3}{r}e^{-r/2M}dudv + r^2(d\theta^2 + \sin^2 \theta d\phi^2)$, where $r(wv)$ solves $uv = e^{r/2M}(r - 2M)/2M$. The null vectors $\partial_u$ and $-\partial_v$ are future directed. Consider the one-parameter family of axially symmetric spacelike hypersurfaces $\Sigma_{\epsilon} = \mathbb{R} \times S^2$ defined by $\Sigma_{\epsilon} = \{ u = \hat{y} - \epsilon x, v = \hat{y} + \epsilon x, \cos \theta = x, \phi = \phi \}$, where $\hat{y} \in \mathbb{R}$, $x \in [-1,1]$, $\phi \in [0,2\pi]$. The discrete isometry of the Kruskal spacetime defined by $\{ u, v \} \rightarrow \{ v, u \}$ implies that, under reflection with respect to the equatorial plane, i.e. $(\hat{y}, x, \phi) \rightarrow (\hat{y}, -x, \phi)$, the induced metric of $\Sigma_{\epsilon}$ remains invariant, while the second fundamental form of $\Sigma_{\epsilon}$ changes sign. Uniqueness of the outermost generalized apparent horizon implies that $S_{\text{out}}^\epsilon$ must be axially symmetric and have equatorial symmetry. We therefore restrict ourselves to the subclass of surfaces $S_{\epsilon} \subset \Sigma_{\epsilon}$ defined by $S_{\epsilon} = \{ \hat{y} = y(x, \epsilon), x = \phi = \phi \}$, with $y(-x, \epsilon) = y(x, \epsilon)$. We are seeking solutions of the equation $p - |q| = 0$. Thus, it becomes natural to restrict the space of functions to $U^{m,\alpha} = \{ y \in C^{m,\alpha}(S^2) : \partial_\theta y = 0, y(-x) = y(x) \}$, i.e. $m$-times differentiable functions on the unit sphere, with Hölder continuous $m$-th derivatives with exponent $\alpha$ and invariant under the axial Killing vector on $S^2$ and under reflection about the equatorial plane. $U^{m,\alpha}$ is a closed subset of the Banach space $C^{m,\alpha}(S^2)$ hence a Banach space itself. Let $I \subset \mathbb{R}$ be the closed interval where $\epsilon$ takes values. For each function $y \in U^{2,\alpha}$, the expression $p - |q|$ defines a non-linear map $\mathcal{F} : U^{2,\alpha} \times I \rightarrow U^{0,\alpha}$. We are therefore looking for solutions of $f(y, \epsilon) = 0$. At $\epsilon = 0$, the hypersurface $\Sigma_0$ is totally geodesic, which implies that $q = 0$ for any surface on it. Consequently, all generalized apparent horizons satisfy $p = 0$ and are, in fact, minimal surfaces. The only closed minimal surface in $\Sigma_0$ is the bifurcation surface $\{ u = 0, v = 0 \}$. Thus, the equation $f(y, \epsilon) = 0$ has $y = 0$ as the unique solution when $\epsilon = 0$. We want to use the implicit function theorem to show that there exists a unique solution $y(x, \epsilon)$ of $f(y, \epsilon) = 0$ in a suitable neighbourhood of $y = 0$ and $\epsilon = 0$. This requires $f$ to be $C^1(U^{2,\alpha} \times I)$ and the partial derivative $D_y f|_{(y=0, \epsilon=0)}$ to be an isomorphism between $U^{2,\alpha} \rightarrow U^{0,\alpha}$. As emphasized by Khuri, the appearance of an absolute value in $|q|$ makes it unlikely that the functional $f$ is $C^1(U^{2,\alpha} \times I)$. To address this issue, we restrict ourselves to surfaces of the form $y = eY$, where $Y \in U^{2,\alpha}$. An explicit computation of the mean curvature $p$ on such surfaces gives $p = \epsilon \mathcal{P}(Y(x), \dot{Y}(x), \dot{Y}(x), x, \epsilon)$, where $\mathcal{P}$ denotes derivative with respect to $x$ and where $\mathcal{P} : \mathbb{R}^3 \times [-1,1] \times I \rightarrow \mathbb{R}$ is a smooth (in fact, analytic) function. Similarly $q = \epsilon \mathcal{Q}(Y(x), \dot{Y}(x), x, \epsilon)$, where $\mathcal{Q} : \mathbb{R}^2 \times [-1,1] \times I \rightarrow \mathbb{R}$ is an analytic function. Moreover, the function $\mathcal{Q}$ has the symmetry $\mathcal{Q}(x_1, x_2, x_3, x_4) = -\mathcal{Q}(x_1, -x_2, -x_3, x_4)$, which reflects the fact that the extrinsic curvature of $\Sigma_{\epsilon}$ changes sign under a transformation $x \rightarrow -x$. Let us write $P(Y, \epsilon)(x) = \mathcal{P}(Y(x), \dot{Y}(x), \dot{Y}(x), x, \epsilon)$ and similarly $\mathcal{Q}(Y, \epsilon)(x) = \mathcal{Q}(Y(x), \dot{Y}(x), x, \epsilon)$.

Instead of analysing the functional $f$, we set out to consider the alternative functional $F : U^{2,\alpha} \times I \rightarrow U^{0,\alpha}$ defined by $F(Y, \epsilon) = P(Y, \epsilon) - |Q(Y, \epsilon)|$. This functional has the property that, for $\epsilon > 0$, the solutions of $F(Y, \epsilon) = 0$ correspond exactly to the solutions of $f(y, \epsilon) = 0$ via the relation $y = eY$. However, the functional $F$ is well-defined even at $\epsilon = 0$ (in fact, for all $\epsilon \in I$). By proving that $F = 0$ admits solutions in a neighbourhood of $\epsilon = 0$, we will conclude
that \( f = 0 \) admits solutions for \( \epsilon > 0 \) and the solutions will in fact belong to a neighbourhood of \( y = 0 \) since \( y = \epsilon Y \).

In order to show that \( F \) admits solutions we will use the implicit function theorem. A direct calculation yields \( F(Y, 0)(x) = G(L(Y)(x) - 3|x|) \), where \( G \) is a positive function of \((Y(x), \dot{Y}(x), x) \) and \( L(Y) \equiv -(1 - x^2)\dot{Y} + 2x\dot{Y} + Y \). This operator is an isomorphism between \( U^{2,\alpha} \) and \( U^{0,\alpha} \). Let \( Y_{sol} \in U^{2,\alpha} \) be the unique solution of the equation \( L(Y) = 3|x| \). For later use, we note that \( Q(Y_{sol}, 0) = -3x \). This vanishes only at \( x = 0 \). This is the key property that allows us to prove that \( F \) is \( C^1(U^{2,\alpha} \times I) \).

The \( C^1(U^{2,\alpha} \times I) \) property of the functional \( P(Y, \epsilon) \) is standard. More subtle is to show that \(|Q| \) is \( C^1(U^{2,\alpha} \times I) \) in a suitable neighbourhood of \((Y_{sol}, \epsilon) = 0 \). Let \( r_0 > 0 \) and define \( V_0 = \{(Y, \epsilon) \in U^{2,\alpha} \times I : \|Y - Y_{sol}, \epsilon\|_{U^{2,\alpha} \times I} \leq r_0\} \). First of all we need to show that \(|Q|\) is (Fréchet-)differentiable on \( V_0 \), i.e. that for all \((Y, \epsilon) \in V_0 \) there exists a continuous linear mapping \( D_Y Q|_\epsilon : U^{2,\alpha} \times I \rightarrow U^{0,\alpha} \) such that, for all \((h, \delta) \in U^{2,\alpha} \times I\), \(|Q(Y + h, \epsilon + \delta) - Q(Y, \epsilon)| = D_Y Q|_\epsilon(h, \delta) + R_Y(h, \delta)\) where \( \|R_Y(h, \delta)\|_{U^{0,\alpha}} = o(\|(h, \delta)\|_{U^{2,\alpha} \times I}) \).

The key observation is that

\[
|Q(Y, \epsilon)(x)| = -\sigma(x)Q(Y, \epsilon)(x) \tag{2}
\]

where \( \sigma(x) \) is the sign function, (i.e. \( \sigma(x) = +1 \) for \( x \geq 0 \) and \( \sigma(x) = -1 \) for \( x < 0 \)). For \( x \) away from a neighbourhood of \( 0 \), this follows from the fact that \( Q(Y_{sol}, \epsilon) = -3x \), which is negative for \( x > 0 \) and positive for \( x < 0 \). Taking \( r_0 \) small enough, and using that \( Q \) is a smooth function of their arguments, the same holds for any \((Y, \epsilon) \in V_0 \). Moreover, the function \( Q(Y, \epsilon)(x) \) is odd in \( x \), so it passes through zero at \( x = 0 \). Hence, in a small enough neighbourhood of \( x = 0 \), the relation \( (2) \) holds provided we can prove that \( Q(Y, \epsilon) \) is strictly decreasing at \( x = 0 \). But this follows easily from the fact that \( \frac{dQ(Y_{sol}, \epsilon)}{dx} |_{x = 0} = -3 \) and \( Q \) is a smooth function.

From its definition, it is immediate that \( Q(Y, \epsilon)(x) \) is \( C^{1,\alpha} \) and has derivative \( D_Y Q|_\epsilon(h, \delta) = A_{Y_{\epsilon}}(x)\dot{h}(x) + B_{Y_{\epsilon}}(x)\dot{h}(x) + C_{Y_{\epsilon}}(x)\delta \), where \( A_{Y_{\epsilon}}(x) \equiv \partial_1 Q|_{(Y_{\epsilon}(x), \dot{Y}(x), x, \epsilon)} \), \( B_{Y_{\epsilon}}(x) \equiv \partial_2 Q|_{(Y_{\epsilon}(x), \dot{Y}(x), x, \epsilon)} \) and \( C_{Y_{\epsilon}}(x) \equiv \partial_3 Q|_{(Y_{\epsilon}(x), \dot{Y}(x), x, \epsilon)} \). We note that these three functions are \( C^{1,\alpha} \) and that \( A_{Y_{\epsilon}}, C_{Y_{\epsilon}} \) are odd, while \( B_{Y_{\epsilon}} \) is even (as a consequence of the symmetries of \( Q \)). Defining the linear map \( D_Y Q|_\epsilon(h, \delta) = -\sigma(A_{Y_{\epsilon}}h + B_{Y_{\epsilon}}\dot{h} + C_{Y_{\epsilon}}\delta) \), it follows from \( (2) \) that

\[
|Q(Y + h, \epsilon + \delta) - Q(Y, \epsilon)| = D_Y Q|_\epsilon(h, \delta) + R_Y(h, \delta) \quad \text{with} \quad \|R_Y(h, \delta)\|_{U^{0,\alpha}} = o(\|(h, \delta)\|_{U^{2,\alpha} \times I})
\]

In order to conclude that \( D_Y Q|_\epsilon \) is the derivative of \( Q(Y, \epsilon) \), we only need to check that, it is (i) well-defined, i.e. its image belongs to \( U^{0,\alpha} \), and (ii) that it is continuous, i.e. \( \|D_Y Q|_\epsilon(h, \delta)\|_{U^{0,\alpha}} \leq C\|\epsilon\|_{U^{2,\alpha} \times I} \) for some constant \( C \). To show (i), the most difficult term is \( -\sigma B_{Y_{\epsilon}}\dot{h} \), because \( B_{Y_{\epsilon}}(x) \) is even and need not vanish at \( x = 0 \). However \( \dot{h} \) is an odd function, and hence \( -\sigma B_{Y_{\epsilon}}\dot{h} \) is continuous. To show it is also Hölder continuous, we only need to consider points \( x_1 = a \) and \( x_2 = b \) with \( 0 < a < b \) (if \( x_1 \cdot x_2 \geq 0 \), the sign function remains constant, so \( -\sigma B_{Y_{\epsilon}}\dot{h} \) is in fact \( C^{1,\alpha} \)). Calling \( w(x) \equiv -\sigma(x)B_{Y_{\epsilon}}(x)\dot{h}(x) \) and using that \( w(x) \) is even, we find

\[
|w(x_2) - w(x_1)| = |w(b) - w(a)| = |(w(b) - w(a))| = \left| \frac{d(B_{Y_{\epsilon}}\dot{h})}{dx} \right|_{x = \xi} |x_2 - x_1|^\alpha
\]

where \( \xi \in (a, b) \) and we have used that \( |b - a|^\alpha \leq |b + a|^\alpha = |x_2 - x_1|^\alpha \) and \( |b - a| < 1 \). This proves that \( -\sigma B_{Y_{\epsilon}}\dot{h} \) is Hölder continuous with exponent \( \alpha \).
To check (ii), we first notice that \( w(x) \) obviously satisfies \( \sup_x |w| < C \| (h, \delta) \|_{U^2,0} \) because \( B_{Y,\epsilon}(x) \) is \( C^{1,\alpha} \). It remains to bound the H"older constant \( |w|_\alpha = \sup_{x_1 \neq x_2 \in [a,b]} \frac{|w(x_2) - w(x_1)|}{|x_2 - x_1|^{\alpha}} \).

Combining (3) with the fact that \( B_{Y,\epsilon}(x) \) is \( C^{1,\alpha} \), the bound \( |w|_\alpha \leq C \| (h, \delta) \|_{U^2,0} \) follows at once. This proves (ii) for the term \( -\sigma B_{Y,\epsilon} h \). A similar argument applies to \( -\sigma A_{Y,\epsilon} h \) and \( -\sigma C_{Y,\delta} \), and we conclude that \( D_{Y,\epsilon}[Q] \) is indeed a continuous operator.

In order to apply the implicit function theorem, it is furthermore necessary that \( |Q| \in C^1(U^{2,\alpha} \times I) \) (i.e. that \( D_{Y,\epsilon}[Q] \) depends continuously on \((Y,\epsilon)\)). This means that given any convergent succession \((Y_n, \epsilon_n) \in V_\epsilon\), the corresponding operators \( D_{Y_n,\epsilon_n}[Q] \) also converge.

Denoting by \((Y,\epsilon) \in V_\epsilon\), the limit of the succession, we need to prove that \( \| D_{Y_n,\epsilon_n}[Q] - D_{Y,\epsilon}[Q] \|_{U^{2,\alpha} \times I} \to 0 \). It suffices to find a constant \( K \) (which may depend on \((Y,\epsilon)\)), such that

\[
\| (D_{Y_n,\epsilon_n}[Q] - D_{Y,\epsilon}[Q])(h, \delta) \|_{U^0} < K \| (h, \delta) \|_{U^{2,\alpha} \times I} \| (Y_n - Y, \epsilon_n - \epsilon) \|_{U^2,0} \times I \tag{4}
\]

for all \((h, \delta) \in U^{2,\alpha} \times I\). Again, the most difficult case involves \( \sigma(B_{Y,\epsilon} - B_{Y_n,\epsilon_n}) \hat{h} \), so we concentrate on this term. Using the mean value theorem on the function \( B \equiv \partial_2 Q \) (recall that \( B_{Y,\epsilon} = B(Y(x),Y(x),x,\epsilon) \)),

\[
\sup_x |\sigma(x)(B_{Y,\epsilon} - B_{Y_n,\epsilon_n}) \hat{h}| \leq 2 \sup_x \| \nabla B \| \sup_x |\hat{h}| \| (Y_n - Y, \epsilon_n - \epsilon) \|_{U^{2,\alpha} \times I}, \tag{5}
\]

where \( \nabla B \) is the gradient of \( B \) and \( K \subset \mathbb{R}^4 \) is a compact domain depending only on \( r_0 \) and \( Y_{sol} \) defined so that, for all \((Y,\epsilon) \in V_\epsilon\), the quadruple \((Y(x),Y(x),x,\epsilon) \in K\), for all \( x \in [-1,1] \).

Inequality (5) is already of the form (4) (recall that \( B \) is smooth). It only remains to bound the H"older constant of \( z \equiv \sigma(B_{Y,\epsilon} - B_{Y_n,\epsilon_n}) \hat{h} \) in a similar way. As before, this is done by distinguishing two cases, namely when \( x_1 \cdot x_2 \geq 0 \) and when \( x_1 \cdot x_2 < 0 \). Obtaining an inequality of the form \( \sup_{x_1 \neq x_2, x_1, x_2 \geq 0} |z(x_2) - z(x_1)| \leq K \| (h, \delta) \|_{U^{2,\alpha} \times I} \| (Y_n - Y, \epsilon_n - \epsilon) \|_{U^2,0} \times I \) is standard, because \( \sigma(x) \) is a constant function. When \( x_1 \cdot x_2 < 0 \), we exploit the parity of the functions as in (3) to get \( |z(x_2) - z(x_1)| \leq \left( \frac{d(B_{Y_n,\epsilon_n} - B_{Y,\epsilon})}{dx} \right)_{x = \xi} |x_2 - x_1| \), where \( \xi \in (a,b) \) and we are assuming \( x_1 = -a, x_2 = b, 0 < a < b \) without loss of generality. Bounding the right hand side in terms of \( K_z \| (h, \delta) \|_{U^{2,\alpha} \times I} \| (Y_n - Y, \epsilon_n - \epsilon) \|_{U^2,0} \times I \) is again standard, since the sign function \( \sigma(x) \) has already disappeared. This, combined with (5) gives (4) and hence continuity of the derivative of \( D_{Y,\epsilon}[Q] \) with respect to \((Y,\epsilon) \in V_\epsilon\).

The final requirement to apply the implicit function theorem to \( F = P - |Q| \) is to check that \( D_Y F |_{(Y_{sol},\epsilon=0)} \) is invertible. A simple computation gives \( D_Y F |_{(Y_{sol},\epsilon=0)}(h) = cL(h) \), where \( c \) is a positive function of \( x \) and \( L \) is the elliptic operator defined above, which is an isomorphism between \( U^{2,\alpha} \) and \( U^0 \). Thus, the implicit function theorem can be used to conclude that there exists an open neighbourhood \( I \subset \epsilon \) of \( \epsilon = 0 \) and a \( C^1 \) map \( \hat{Y} : I \to U^{2,\alpha} \) such that \( \hat{Y}(\epsilon = 0) = Y_{sol} \) and \( y = c \hat{Y}(\epsilon) \) defines a \( C^{2,\alpha} \) generalized apparent horizon embedded in \( \Sigma_\epsilon \).

Acknowledgments

We acknowledge financial support from the projects FIS2009-07238 (MEC), Gr-234 (JCyL) and P09-FQM-4496 (JA and FEDER funds). AC acknowledges the PhD grant AP2005-1195 (MEC).

References

[1] Bray H L and Khuri M Preprint arXiv:0905.2622
[2] Carrasco A and Mars M Preprint arXiv:0911-0883v1
[3] Penrose R 1973 Ann. N. Y. Acad. Sci. 224 125-134
[4] Mars M 2009 Class. Quantum Grav. 26 193001
[5] Eichmair M Preprint arXiv:0805.4554
[6] Khuri M, private communication