



ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus

Mécanique


Yakov Shlapentokh-Rothman

Weak cosmic censorship, trapped surfaces, and naked singularities for the Einstein vacuum equations

Volume 353 (2025), p. 379-410

Online since: 6 February 2025

<https://doi.org/10.5802/crmeca.284>

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



*The Comptes Rendus. Mécanique are a member of the
Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1873-7234



Review article / Article de synthèse

Weak cosmic censorship, trapped surfaces, and naked singularities for the Einstein vacuum equations

Censure cosmique faible, surfaces piégées et singularités nues pour les équations d'Einstein dans le vide

Yakov Shlapentokh-Rothman ^{a, b}

^a Department of Mathematics, University of Toronto, 40 St. George Street, Toronto, ON, Canada

^b Department of Mathematical and Computational Sciences, University of Toronto Mississauga, 3359 Mississauga Road, Mississauga, ON, Canada

E-mail: yakovsr@math.toronto.edu

Abstract. The weak cosmic censorship conjecture posits that, generically, all singularities in General Relativity arising from regular asymptotically flat initial data should have a complete future null infinity. While this conjecture remains wide open, it has inspired many mathematical works concerning topics such as trapped surface formation and the construction of naked singularities. In this article we will review some of these works and attempt to emphasize their interconnectedness.

Résumé. La conjecture de censure cosmique faible postule que, de manière générale, toutes les singularités en relativité générale provenant de données initiales régulières asymptotiquement plates devraient avoir un infini nul futur complet. Bien que cette conjecture reste largement ouverte, elle a inspiré de nombreux travaux mathématiques portant sur des sujets tels que la formation de surfaces piégées et la construction de singularités nues. Dans cet article, nous passerons en revue certains de ces travaux et tenterons de mettre en évidence leurs interconnexions. interconnectedness.

Keywords. Weak cosmic censorship, Trapped surfaces, Naked singularities.

Mots-clés. Censure cosmique faible, Surfaces piégées, Singularités nues.

Funding. NSERC discovery grants (RGPIN-2021-02562 and DGEGR-2021-00093).

Manuscript received 11 October 2024, revised and accepted 13 January 2025.

1. Introduction

We are interested here in General Relativity as a part of classical physics, and thus the Cauchy problem will be paramount for us. We now quickly recall the relevant notions. We say that the pair of a 3-dimensional Riemannian manifold (Σ, h) and a symmetric $(0, 2)$ -tensor k on Σ satisfy the *constraint equations* if

$$R(h) + (\operatorname{tr}_h k)^2 - |k|_h^2 = 0, \quad \operatorname{div}_h k - \nabla \operatorname{tr}_h k = 0.$$

In this case we say that (Σ, h, k) forms an *initial data set*. The fundamental theorem concerning the Cauchy problem is the following celebrated result of Choquet-Bruhat and Geroch:

Theorem 1. [1, 2] *For every sufficiently regular initial data set (Σ, h, k) , there exists a unique maximal globally hyperbolic spacetime (\mathcal{M}, g) solving the Einstein vacuum equations*

$$\text{Ric}(g) = 0,$$

so that Σ may be embedded in (\mathcal{M}, g) with first and second fundamental form h and k respectively.

The proof of Theorem 1 also yields that continuous dependence on the data holds in a suitable sense. We will assume in this article that the reader is familiar with the basics of Lorentzian geometry (such as the term “globally hyperbolic development” above in Theorem 1). See [3] for an introduction to Lorentzian geometry and causality theory.

As with any evolutionary PDE from mathematical physics, once a locally well-posed initial value problem has been established, the next natural question is, supposing we have sufficiently “nice” initial data, do we obtain a corresponding “global and singularity free” solution? Note, however, that unlike the case for the more traditional PDEs of mathematical physics, in view of the geometric nature of the Einstein equations, it is not immediately clear what “global and singularity free” should mean! Nevertheless, some necessary conditions are immediate; for example, on physical grounds it is clear that we should require that global existence at least include causal geodesic completeness. As we shall discuss in the next section, suitably understood, the famous Schwarzschild solution already illustrates the ubiquitousness of the failure of global existence for the Einstein vacuum equation.

1.1. The Schwarzschild spacetime and geodesic incompleteness

The first discovered non-trivial exact solution for the Einstein equations was the Schwarzschild spacetime [4]. For every $M > 0$, the *exterior region* of the Schwarzschild spacetime of mass M may be covered by so-called “Schwarzschild” coordinates $(t, r, \theta, \phi) \in \mathbb{R} \times (2M, \infty) \times \mathbb{S}^2$:

$$g_{\mathcal{S}} = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

It took over 30 years for the global geometry of the Schwarzschild spacetime to be properly understood. We do not have the space here to review this in full detail (see the detailed exposition in Section 2 of [5]), but it will be useful to recall some of the most salient points. The maximal extension of the Schwarzschild spacetime is a globally hyperbolic spacetime with a two-ended Cauchy hypersurface $\Sigma \approx \mathbb{R} \times \mathbb{S}^2$ and may be globally covered by a double-null coordinate system where

$$g_{\mathcal{S}} = -\Omega^2(U, V) dU dV + r^2(U, V)(d\theta^2 + \sin^2\theta d\phi^2),$$

for suitable functions $\Omega(U, V)$ and $r(U, V)$. Furthermore, we may rescale the U and V coordinates so that their ranges lie in a bounded set. In Figure 1 we then depict the range of the U and V coordinates with the corresponding Penrose diagram. We emphasize that each point in the diagram corresponds to a sphere in the actual spacetime.

The original Schwarzschild coordinates correspond to the rightmost diamond in the Penrose diagram. (The leftmost diamond corresponds to an isometric copy of the rightmost copy. We will mostly ignore this leftmost diamond in what follows.) Introducing a new coordinate $v = t + r + 2M \log(r - 2M)$, we obtain in (v, r, θ, ϕ) coordinates

$$g_{\mathcal{S}} = -\left(1 - \frac{2M}{r}\right) dv^2 + 2dv dr + r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

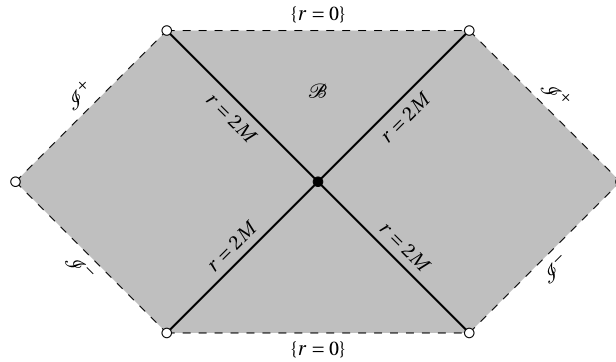


Figure 1. Penrose diagram of maximally extended Schwarzschild.

We may then let these coordinates run over the range $(v, r, \theta, \phi) \in \mathbb{R} \times (0, \infty) \times \mathbb{S}^2$, and these coordinates cover the union of the original exterior region with the region \mathcal{B} from the Penrose diagram. Time-orienting by ∂_v , one now easily checks that for each $(c_0, \theta_0, \phi_0) \in \mathbb{R} \times \mathbb{S}^2$, after reparameterization, the curve

$$s \mapsto (v, r, \theta, \phi) = (c_0, -s, \theta_0, \phi_0), \quad s < 0,$$

represents a future oriented incomplete null geodesic. Furthermore, as $r \rightarrow 0$ along the geodesic, one may compute that the Kretschmann scalar $R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta}$ blows-up. It immediately follows that there is no way to acquire future null geodesic completeness by extending the Schwarzschild spacetime in a C^2 fashion¹, and thus that the Schwarzschild spacetime represents an example of singularity formation for the Cauchy problem!

Having established the existence of a singular spacetime, one immediately asks if the presence of this incompleteness is a generic feature of solutions? Could it be the case that this incompleteness is a special feature of highly symmetric spacetimes? Penrose’s notion of a trapped surface, which we will discuss in the next section, quickly yields insight into these questions.

1.2. Trapped surfaces

Penrose’s concept of a trapped surface will play a central role in this article.

Definition 1.1. [8] Let \mathcal{S} be a closed oriented spacelike surface in a 3 + 1 dimensional time-oriented Lorentzian manifold. At every point $p \in \mathcal{S}$ one may pick two linearly independent future oriented null and normal vectors L and \underline{L} . We say that \mathcal{S} is trapped if both of the corresponding mean curvatures are negative. If one mean curvature is negative and the other vanishes, we say that \mathcal{S} is marginally trapped. We say that \mathcal{S} is anti-trapped if both of the corresponding mean curvatures are positive, and if one mean curvatures is positive and the other vanishes, we say that \mathcal{S} is marginally anti-trapped.

Note that by the first variation formula for the area form, a surface \mathcal{S} is trapped if the area form locally contracts under any flow of null vector congruences. This is physically interpreted as a sign that one is in a regime of strong gravity. A straightforward calculation shows that a sphere of radius r in the Schwarzschild spacetime is trapped if and only if $t > 0$ and $r < 2M$. This region of trapped spheres exactly corresponds to the region \mathcal{B} in Figure 1.

The usefulness of the trapped surface notion follows from two fundamental facts. The first is that being a trapped surface is a manifestly stable notion. Trapped surfaces cannot disappear

¹In fact, it is not even possible to extend maximal Schwarzschild in a C^0 fashion! See [6, 7].

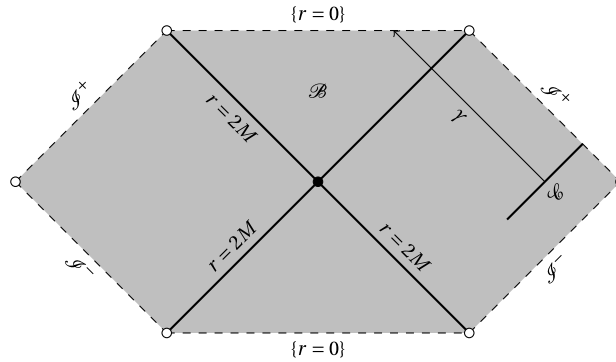


Figure 2. The hypersurface \mathcal{C} and null geodesic γ in Schwarzschild from Proposition 3.

from small perturbations of a metric, and, more importantly, in view of Cauchy stability, any sufficiently small perturbation of the Cauchy data of a spacetime with trapped surfaces will still lead to a spacetime which possesses trapped surfaces. In particular, any sufficiently small perturbation of Schwarzschild’s Cauchy data will still lead to a spacetime with a trapped surface. The second fact underlying the usefulness of the trapped surface notion is the following famous theorem due to Penrose.

Theorem 2. [8] *Suppose that (\mathcal{M}, g) is a globally hyperbolic spacetime with a non-compact Cauchy surface and satisfies*

$$\text{Ric}(V, V) \geq 0, \quad \forall V \text{ with } g(V, V) = 0.$$

Then, if (\mathcal{M}, g) possesses a trapped surface, it must be future causally geodesically incomplete.

We immediately conclude that the geodesic incompleteness of Schwarzschild is a stable property to perturbations of the Cauchy data! We are thus forced to take seriously the failure of global existence in General Relativity.

1.3. Completeness of future null infinity and the weak cosmic censorship conjecture

While the Schwarzschild spacetime poses serious conceptual questions concerning its singularity as $r \rightarrow 0$, the spacetime geometry also suggests a “way out”. On the Penrose diagram in Figure 1, we have denoted part of the boundary by \mathcal{S}^+ . This denotes “future null infinity” and corresponds to future end points of the null hypersurfaces $t = r + 2M \log(r - 2M) + c$ for any choice of $c \in \mathbb{R}$. Physically, one imagines that a far away observer propagates along a causal curve which is close to \mathcal{S}^+ . The following proposition is straightforward to establish.

Proposition 3. *For $r_0 \gg 1$, let \mathcal{C} denote the outgoing null hypersurface $\{t = r + 2M \log(r - 2M)\} \cap \{r \geq r_0\}$. For each $\bar{r} \geq r_0$ and $(\bar{\theta}, \bar{\phi}) \in \mathbb{S}^2$, we may consider the ingoing null geodesic $\gamma_{\bar{r}, \bar{\theta}, \bar{\phi}}$ uniquely defined by the requirement that $\gamma_{\bar{r}, \bar{\theta}, \bar{\phi}}(0) = \mathcal{C}|_{(r, \theta, \phi) = (\bar{r}, \bar{\theta}, \bar{\phi})}$ and $\dot{\gamma}_{\bar{r}, \bar{\theta}, \bar{\phi}}(0) = \partial_t - (1 - (2M/r))\partial_r$. We then define $A(\bar{r})$ to be the affine length of γ (it is immediate that A does not depend on $(\bar{\theta}, \bar{\phi})$). We then have that*

$$\lim_{\bar{r} \rightarrow \infty} A(\bar{r}) = \infty.$$

See Figure 2 for the Penrose diagram corresponding to Proposition 3. Proposition 3 has the physical interpretation that far away observers “at” future null infinity \mathcal{S}^+ live forever and may make observations without ever needing to interact with the Schwarzschild singularity. This good property of future null infinity is referred to as *completeness*. Our next goal is formulate a general definition of this property, but we need a few preliminaries.

Definition 4. Let (Σ, h, k) be an initial data set. We say that (Σ, h, k) is strongly asymptotically flat with n -ends if there exists a compact set K so $\Sigma \setminus K$ is diffeomorphic to n copies of \mathbb{R}^3 minus a ball, and there exists coordinates $\{x^a\}$ on each of these so that

$$h_{ab} = \left(1 + \frac{2M}{r}\right) \delta_{ab} + o_4(r^{-3/2}), \quad k_{ab} = o_3(r^{-5/2}),$$

where the $o_i(\cdot)$ means that up to i derivatives may be applied and that each derivative yields an extra power of decay in r .

It is a consequence of the stability of Minkowski space [9, 10], that associated to each end of a strongly asymptotically flat initial data set, the corresponding maximal development will possess an outgoing null hypersurface \mathcal{C} with future complete null generators and whose geometry asymptotes to a null cone from Minkowski space. (The details of the asymptotic geometry will not be important for us here.)

We now are ready to give a definition of a complete future null infinity. We note that there are various definitions existing in the literature. Our definition is in the spirit of the definition given in [11]. A key point in both is that we do not formally define future null infinity as a boundary of a conformal compactification of spacetime. This avoids various technical regularity issues concerning such compactifications which are largely irrelevant for the physical content of the definition.

Definition 1.2. Let (Σ, h, k) be a strongly asymptotically flat initial data set and (\mathcal{M}, g) be the corresponding maximal development. Associated to each asymptotically flat end, choose an outgoing null hypersurface \mathcal{C} in \mathcal{M} with future complete null generators whose geometry approaches that of a standard Minkowski cone. Let L denote a geodesic null normal along \mathcal{C} . Then let \underline{L} be a choice of a transversal null vector field along \mathcal{C} which is parallel transported by L . For every point $p \in \mathcal{C}$, we let $A(p)$ denote the maximal affine length of the null geodesic starting at p with initial tangent \underline{L} , and we let $p(s)$ denote the outgoing null geodesic which starts at p and has initial tangent vector equal to L . We then say that (\mathcal{M}, g) has a *complete future null infinity* if, for every $p \in \mathcal{C}$, $\lim_{s \rightarrow \infty} A(p(s)) = \infty$.

We note that we have chosen to work with strongly asymptotically initial data sets so that we may invoke [9, 10]. However, newer proofs of the stability of Minkowski space which allow weaker versions of asymptotic flatness (cf. [12]) in principle allow for weaker asymptotic flatness assumptions in Definition 1.2. The corresponding reformulations of Definition 1.2 are straightforward to state.

Finally, we are ready for the weak cosmic censorship conjecture.

Conjecture 5 (Weak cosmic censorship). *Generic strongly asymptotically flat initial data sets possess a complete future null infinity.*

Spacetimes with an incomplete future null infinity are called *naked singularities*². The earliest version of Conjecture 5 originates in [13], which, in particular, introduced the notion of a “cosmic censor” who prevents the appearance of naked singularities. However, the addition of the word “generic” into the conjecture only followed the discovery of naked singularity solutions for the spherically symmetric Einstein-scalar field system [14]. See the discussion in Section 2.3 below.

We close this section by noting some important ambiguities underlying Conjecture 5. The most apparent of these is the presence of the word “generic;” a successful resolution of Conjecture 5 will have to make this precise. A more subtle ambiguity is the fact that the maximal development of an initial data set depends on the functional framework used to solve the Einstein

²In this review we will restrict attention to naked singularities which arise in evolution from complete and sufficiently regular initial data. This requirement rules out solutions such as negative mass Schwarzschild or the super extremal Kerr black hole.

equations. In particular, as is the case famously with the strong cosmic censorship conjecture (see [15]), different regularity requirements on initial data or on what constitutes an admissible development of data could lead to different outcomes for Conjecture 5. See the discussions below in Sections 2.3–2.5.

1.4. *Outline of rest of the article*

Though the focus on this article is on the Einstein vacuum equations, much of our intuition about the vacuum equations derives from an influential sequence of works carried out by Christodoulou in the 1990s concerning the spherically symmetric Einstein-scalar field system. In Section 2, we will review Christodoulou’s work; in particular we will discuss Christodoulou’s examples of naked singularities, the study of the instabilities of naked singularities, and the central role of trapped surface formation. Finally we will briefly discuss recent work of Jaydeep Singh which yields new insights on the naked singularities of Christodoulou.

In Section 3 we will start with a discussion of Christodoulou’s short-pulse method and corresponding breakthrough work on trapped surface formation for the Einstein vacuum equations. We will then survey various works which lead to significant extensions of Christodoulou’s original work.

Finally, in Section 4 we will discuss the recent construction of naked singularity solutions for the Einstein vacuum equations. We will explain various notions of self-similarity and give a high level overview of the proof of existence of the naked singularities.

2. The spherically symmetric Einstein-scalar field system

Before trying to tackle the weak cosmic censorship conjecture for the Einstein vacuum equations, it is natural to work first with easier “model problems”. For many PDEs of mathematical physics, such a model problem, which is consistent with asymptotic flatness, may be obtained by restricting considerations to spherically symmetric solutions. Unfortunately, for the Einstein equations, Birkhoff’s theorem implies that any solution to the Einstein vacuum equations which is spherically symmetric must be static and isometric to a Schwarzschild spacetime. Arguably, the simplest way to remove this rigidity while retaining many fundamental features of the Einstein vacuum equations is to consider the spherically symmetric Einstein-scalar field system.

The unknowns for the spherically symmetric Einstein (real) scalar-field system are a 1 + 1 dimensional Lorentzian manifold with boundary (\mathcal{Q}, g) which arises from quotienting out the original spacetime by the spherical symmetry, an area radius function $r : \mathcal{Q} \rightarrow [0, \infty)$, and a scalar field $\phi : \mathcal{Q} \rightarrow \mathbb{R}$. The equations are

$$\begin{aligned} r\nabla_\alpha\nabla_\beta r &= \frac{1}{2}g_{\alpha\beta}(1 - |\nabla r|_g^2) - r^2(\partial_\alpha\phi\partial_\beta\phi - \frac{1}{2}g_{\alpha\beta}|\nabla\phi|_g^2), \\ \nabla^\alpha(r^2\nabla_\alpha\phi) &= 0 \\ K(g) &= r^{-2}(1 - |\nabla r|_g^2) + |\nabla\phi|_g^2, \end{aligned} \tag{2.1}$$

where $K(g)$ denotes the Gauss curvature of g . It is straightforward to give analogues of Definition 1.2 and Conjecture 5 for this system.

In the Sections 2.1–2.4 we review the results from an important sequence of papers due to Christodoulou [14, 16–18]. In our discussion of these papers we will follow closely the review article [11].

Finally, in Section 2.5 we discuss the recent work [19] concerning the behavior of Christodoulou’s naked singularities under relatively smooth perturbations.

2.1. *Scale-invariance and BV solutions*

We set $\Gamma \doteq \partial\mathcal{Q}$ to be the set of fixed points of the $SO(3)$ action. Using that every contractible $1 + 1$ dimensional Lorentzian manifold is conformally equivalent to a suitable subset of \mathbb{R}^{1+1} , it is straightforward to show that the metric g may be written as

$$g = -\Omega^2 du dv, \tag{2.2}$$

for some function $\Omega(u, v)$ and where the null coordinates u and v satisfy $v = u$ along Γ . We note that the double-null form (2.2) is preserved by any coordinate change $u \mapsto f(u)$ and $v \mapsto h(v)$ for $f', h' > 0$. The system (2.1) then becomes the following system of $1 + 1$ dimensional PDEs for the functions $\Omega(u, v)$, $r(u, v)$, and $\phi(u, v)$:

$$\begin{aligned} r \frac{\partial^2 r}{\partial u \partial v} + \frac{\partial r}{\partial u} \frac{\partial r}{\partial v} &= -\frac{1}{4} \Omega^2, \\ \frac{\partial^2 r}{\partial u^2} - 2\Omega^{-1} \frac{\partial r}{\partial u} \frac{\partial \Omega}{\partial u} &= -r \left(\frac{\partial \phi}{\partial u} \right)^2, \\ \frac{\partial^2 r}{\partial v^2} - 2\Omega^{-1} \frac{\partial r}{\partial v} \frac{\partial \Omega}{\partial v} &= -r \left(\frac{\partial \phi}{\partial v} \right)^2, \\ r^2 \frac{\partial^2 \log \Omega}{\partial u \partial v} - \frac{1}{4} \Omega^2 &= -r^2 \frac{\partial \phi}{\partial u} \frac{\partial \phi}{\partial v} + \frac{\partial r}{\partial u} \frac{\partial r}{\partial v}, \\ r \frac{\partial^2 \phi}{\partial u \partial v} + \frac{\partial r}{\partial u} \frac{\partial \phi}{\partial v} + \frac{\partial r}{\partial v} \frac{\partial \phi}{\partial u} &= 0. \end{aligned} \tag{2.3}$$

Closely related to the rescaling gauge freedom in the double-null coordinate system (2.2) is the fact that there is a natural group action of \mathbb{R}_+ on solutions to the system (2.3) defined by

$$\{r(u, v), \Omega(u, v), \phi(u, v)\} \mapsto \{a^{-1}r(au, av), \Omega(au, av), \phi(au, av)\}, \quad a \in (0, \infty). \tag{2.4}$$

Following Christoudoulou [17], we refer to solutions which are fixed points of the map (2.4) as being *scale-invariant*. In [17] Christoudoulou classifies scale-invariant solutions possessing a past complete incoming null curve emanating from $(u, v) = (0, 0)$. The derivative $\partial_v \phi$ will have a jump discontinuity along $\{v = 0\}$ and the spacetimes fall into three classes depending on the value of this jump

$$A \doteq \frac{1}{2} \left(\lim_{v \rightarrow +0} r \partial_v \phi - \lim_{v \rightarrow -0} r \partial_v \phi \right).$$

- (1) When $A^2 < 1$, the spacetimes is global. The scalar field ϕ is constant to the past of $\{v = 0\}$ and future of $\{u = 0\}$, the transversal null derivatives of ϕ jump along $\{u = 0\}$ and $\{v = 0\}$, and in the region $\{v \geq 0\} \cap \{u \leq 0\}$ we have

$$\phi = \frac{1}{2} \log \left| \frac{u - (1 + A)v}{u - (1 - A)v} \right|.$$

- (2) When $A^2 = 1$, the spacetime exists up to $\{u = 0\}$ where r vanishes and the scalar field ϕ blows-up.
- (3) When $A^2 > 1$, the spacetime exists up to the curve $\{u = -(A - 1)v\}$ and is preceded by an *apparent horizon* along $\{u = -(A^2 - 1)v\}$. The spheres in between the apparent horizon and the final singular curve are all trapped.

Though these solutions are low-regularity in that the derivative of ϕ may jump, the scale-invariant spacetimes may be considered to be physically reasonable in the following sense. They admit an interpretation as an incoming spherical impulsive gravitational wave. When the amplitude of the wave is small, the wave bounces off the center and flies back out to infinity. Once the amplitude of the wave is large enough, upon collapse to the center we have singularity formation. However, generically this singularity is preceded by a trapped region which is consistent with the singularity begin hidden from infinity (cf. the region of trapped surfaces in the Schwarzschild spacetime and also the discussion in Section 2.2 below).

Christodoulou was thus inspired in [17] to build a functional framework which includes the scale-invariant solutions, has norms which are invariant under the map (2.4), and allows one to establish a local well-posedness result for the spherically symmetric Einstein-scalar field system. The class of spacetimes considered are referred to as *spacetimes of bounded variation*. In view of our space constraints in this article we will not describe in detail the norms in this space, but we note that the scalar field ϕ is required to be an absolutely continuous function, $r(\partial\phi/\partial v)$ is required to be a function of bounded variation on each hypersurface $\{u = \text{const}\}$, and $r(\partial\phi/\partial u)$ is required to be a function of bounded variation on each hypersurface $\{v = \text{const}\}$. In addition to having scale-invariant norms defining the space, solutions of bounded variation are intimately tied to the scale-invariant solutions discussed in the paragraph above in that for any solution of bounded variation, one may consider the rescaling map (2.4) as $a \rightarrow 0$ and obtain an exactly scale invariant solution in the limit.

Finally, focusing again only on the scalar field, we note that the work [17] shows that a sufficient condition for global existence as a solution of bounded variation is, in addition to natural normalization assumptions on the geometry, for the scalar field ϕ to satisfy the following along an initial outgoing null hypersurface $\{u = \text{const}\}$:

$$\int_0^\infty |\partial_v^2(r\phi)|^2 dv \ll 1. \tag{2.5}$$

Finally, for later use, we use the term *absolutely continuous spacetimes* to refer to the situation when the scalar field satisfies that $\partial_v(r\phi)$ and $\partial_u(r\phi)$ are absolutely continuous functions along constant u and v hypersurfaces respectively.

2.2. Trapped surface formation

We have seen in our earlier discussions that the Schwarzschild spacetime possesses a complete future null infinity and that all of the spheres immediately preceding the $\{r = 0\}$ singularity are trapped. This turns out not to be a coincidence and, in fact, it was shown in [16] that for the spherically symmetric Einstein-scalar field system, the existence of a single trapped surface guarantees the completeness of future null infinity. (Similar statements can in fact be made for a large class of spherically symmetric Einstein matter systems [20].) It is thus in principle quite useful to have various criteria which imply the eventual formation of a trapped surface.

Before proceeding, we define an important quantity, the *Hawking mass* m via the formula

$$1 - \frac{2m}{r} = |\nabla r|^2.$$

The Hawking mass of a sphere is always non-negative and can be considered a quasi-local measure of the energy/mass contained inside the sphere. It is useful to keep in mind that that the ratio m/r is scale-invariant in that it is left invariant under the map (2.4), and that we always have $\partial_v m \geq 0$.

Christodoulou’s key result concerning trapped surface formation for the spherically symmetric Einstein-scalar field system is the following.

Theorem 6. [16] *Let \mathcal{S}_{u_0, v_0} denote the sphere lying at the intersection $\{u = u_0\} \cap \{v = v_0\}$. Consider two spheres \mathcal{S}_{0, v_0} and \mathcal{S}_{0, v_1} lying along $\{u = 0\}$ with $v_1 > v_0 > 0$. We define the following two quantities:*

$$\delta \doteq \frac{r(\mathcal{S}_{0, v_1})}{r(\mathcal{S}_{0, v_0})} - 1, \quad \eta \doteq \frac{2(m(\mathcal{S}_{0, v_1}) - m(\mathcal{S}_{0, v_0}))}{r(\mathcal{S}_{0, v_1})}. \tag{2.6}$$

Assume that the axis Γ is regular up to the point where it intersects $\{v = v_0\}$. Then there exists a constant $c_1, c_2 > 0$ so that $0 < \delta < c_1$ and $\eta > c_2 \delta \log(\delta^{-1})$ implies that there exists $\tilde{u} > 0$ with $v_0 > \tilde{u}$ and so that the sphere $\mathcal{S}_{\tilde{u}, v_1}$ is a trapped sphere.

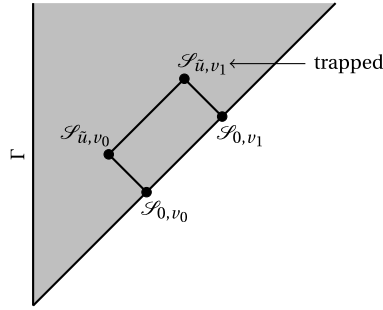


Figure 3. Penrose diagram for Theorem 6.

In Figure 3 we provide the corresponding Penrose diagram.

The quantities δ and η defined in (2.6) are both invariant under the map (2.4), and the condition that $\eta > c\delta \log(\delta^{-1})$ is sharp up to the log factor. More specifically, let $\delta > 0$ be a small constant. Then it is straightforward to construct smooth initial data for the Einstein-scalar field system along $\{u = 0\}$ so that $r|_{u=0} = v$ and so that scalar field is smooth, satisfies (2.5), is compactly supported in $v \in (1, 1 + \delta)$, and also satisfies

$$\int_1^{1+\delta} |\partial_v(r\phi)|^2 dv \gtrsim \delta, \tag{2.7}$$

where the implied constant in (2.7) depends on the implied constant in (2.5). In view of Christodoulou’s theory of BV solutions (see Section 2.1 above) such data will give rise to a solution which is free of trapped surfaces. Furthermore, we will have that

$$r(\mathcal{S}_{0,1+\delta}) = 1 + \delta, \quad r(\mathcal{S}_{0,1}) = 1,$$

and as consequence of the ∂_v -propagation equation for the Hawking mass³ that

$$m(\mathcal{S}_{0,1}) = 0, \quad m(\mathcal{S}_{0,1+\delta}) \gtrsim \delta.$$

Thus, for such solutions we will have both the absence of trapped surfaces in evolution and that $\eta \gtrsim \delta$ (for η and δ as in (2.6)).

Finally, we note that it is a remarkable fact that the quantities considered in (2.6) only involve the behavior of the solution tangentially along $\{u = 0\}$ and that, other than the regularity of the axis Γ , no a priori assumptions are made about the solution on any transversal hypersurface $\{v = \text{const}\}$!

2.3. Naked singularities

In the work [14] Christodoulou constructed solutions to the spherically symmetric Einstein-scalar field system which correspond to naked singularities. The first step in the construction of these solutions is to introduce a generalization of the scale-invariant solutions that we discussed in Section 2.1. More specifically, we observe that the map (2.4) in fact sits inside a 1-parameter family of \mathbb{R}_+ actions parameterized by $k \in \mathbb{R}$ and defined by

$$\{r(u, v), \Omega(u, v), \phi(u, v)\} \mapsto \{a^{-1}r(au, av), \Omega(au, av), \phi(au, av) + k \log(a)\}, \quad a \in (0, \infty). \tag{2.8}$$

³This equation is

$$\partial_v m = \frac{1}{2} \left(\frac{\partial r}{\partial v} \right)^{-1} \left(1 - \frac{2m}{r} \right) r^2 (\partial_v \phi)^2.$$

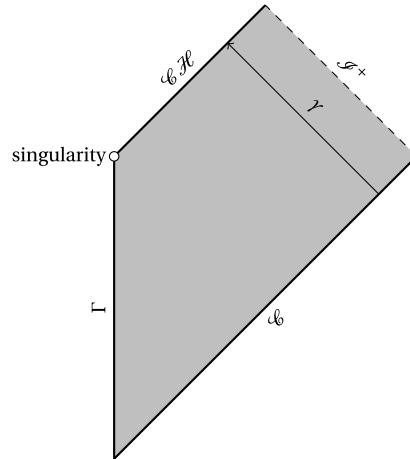


Figure 4. Penrose diagram for the naked singularities of [14] with a depicted incoming null geodesic γ of finite affine length.

A spacetime which is invariant under the map (2.8) is called *k-self similar*. Since it can be useful to not work in the double-null coordinates underlying (2.8), we observe that we can phrase *k*-self similarity in a coordinate invariant way: a spacetime is *k*-self similar if there exists a scaling vector field \mathcal{K} so that

$$\mathcal{L}_{\mathcal{K}} g = 2g, \quad \mathcal{L}_{\mathcal{K}} r = r, \quad \mathcal{L}_{\mathcal{K}} \phi = -k.$$

When $k \neq 0$ it is no longer possible to write down the self-similar solutions explicitly; however, Christodoulou showed that it is still possible to reduce the *k*-self similar spherically symmetric Einstein-scalar system to a two dimensional autonomous system. In [14] Christodoulou carried out a systematic phase plane analysis of this system, and he found that when $0 < k^2 < 1/3$ and after a suitable truncation of the scalar field in the large r region, the *k*-self similar spacetimes have a Penrose diagram as depicted in Figure 4, are asymptotically flat, and moreover have an incomplete future null infinity. This showed, for the first time, that one needs the word “generic” in Conjecture 5. It will be useful for the discussion later in Section 2.4 to observe that along the past light cone of the singularity, the ratio m/r is constant and non-zero.

We make a brief note on global gauges for the *k*-self similar spacetimes: if one has a *k*-self similar spacetime in the double-null coordinate system of the map (2.4), then one may easily check that $\mathcal{K} = u\partial_u + v\partial_v$. However, it turns out that for *k*-self similar spacetimes, it is not possible to have such a coordinate system which covers the past light cone of the singularity, a hypersurface which is of fundamental importance to the analysis. Christodoulou in fact works in *Bondi coordinates* (u, r) where u is an eikonal function. Another option is to work in a rescaled double-null coordinate system where the map (2.8) and hence \mathcal{K} do not take their usual form (see Appendix A of [21]).

These naked singularities do not have smooth data along the initial outgoing hypersurface \mathcal{C} ; however, the lack of regularity there only occurs at the intersection of the past cone of the singularity and \mathcal{C} . At this point the tangential null derivative of the scalar field ϕ is Hölder continuous. This is more regular than initial data for a solution of bounded variation and we may thus consider the initial data to be “regular”. Moreover, Christodoulou shows that the singular point is such it is impossible to extend the solution to the singular point as a solution of bounded variation. The Cauchy horizon $\mathcal{C}\mathcal{H}$ emanating from the singularity is regular in that the derivative of the scalar field extends to $\mathcal{C}\mathcal{H}$ in a Hölder continuous function; in fact, it is

possible to extend the spacetime as a k -self similar solution across \mathcal{CH} . (Thus these spacetimes are relevant for the strong cosmic censorship conjecture.)

We will discuss the instability of these spacetimes in Section 2.4 below. Before addressing the issue of instability, a natural question that arises is whether it is possible to construct naked singularity spacetimes which do not possess any exact self-similarity. Exactly such spacetimes were constructed using a perturbative scattering approach in [21].

Though we do not have the space here to discuss these in detail, it is also interesting to construct naked singularity solutions for other spherically symmetric Einstein matter systems, for example, see [22–24]. Finally, we note that it is an open problem to rigorously construct smooth (or even C^2 !) initial data for the spherically symmetric Einstein-scalar field system which leads to a naked singularity. With regards to this, it is worth observing that the physics literature around *critical collapse* (see [25, 26]) does suggest that naked singularities arising from smooth initial data should exist.

2.4. Instability of naked singularities

In the work [18] Christodoulou studied the instability of a general class of naked singularities for the spherically symmetric Einstein-scalar field system. Before discussing this result in its full generality, it is conceptually useful to focus on the instability of the specific (truncated) k -self similar naked singularities discussed in Section 2.3 above.

Let $\underline{\mathcal{C}}$ denote the past light cone of the singularity. The scaling vector field \mathcal{K} is tangent to $\underline{\mathcal{C}}$, and we may introduce a coordinate $u < 0$ along $\underline{\mathcal{C}}$ so that $\mathcal{K}|_{\underline{\mathcal{C}}} = u\partial_u$ and so that $\underline{\mathcal{C}} \cap \mathcal{C}$ corresponds to $u = -1$. The singularity then corresponds to the limit $u \rightarrow 0$ along $\underline{\mathcal{C}}$. Let $\mathfrak{p} = (\partial_v r)^{-1} \partial_v \phi$ be a suitably normalized transversal derivative of the scalar field ϕ along $\underline{\mathcal{C}}$, so that in view of the underlying self-similarity, we have $|\mathfrak{p}| \sim u^{-1}$. Along $\underline{\mathcal{C}}$ the wave equation may be considered a transport equation for \mathfrak{p} :

$$\partial_u \mathfrak{p} + \frac{1+k^2}{u} \mathfrak{p} = F(u), \quad (2.9)$$

where the $1+k^2$ and $F(u) \sim u^{-2}$ depend only on the behavior of the geometry and scalar field tangential to $\underline{\mathcal{C}}$. Even though $|\mathfrak{p}| \sim u^{-1}$, it is straightforward to see that a solution $\check{\mathfrak{p}}$ of \mathfrak{p} 's equation with generic initial data at $\{u = -1\}$ will have $|\check{\mathfrak{p}}| \sim |u|^{-1-k^2}$, which is a faster blow-up rate than the self-similar rate. Heuristically, we expect that if there exists a perturbation of the initial data so that the corresponding \mathfrak{p} is more singular as $u \rightarrow 0$ than the self-similar rate, then we will have an instability for the naked singularity background. Optimistically, we could then hope to apply Theorem 6 to obtain the formation of a trapped surface and then deduce the existence of a complete future null infinity. We note that one may connect this instability in the evolution equation for transversal null derivatives of ϕ along $\underline{\mathcal{C}}$ to an analogous growth of the energy of null geodesics traveling along $\underline{\mathcal{C}}$; this leads us to refer to the instability as a *blue-shift instability*.

This instability can indeed be triggered by considering a perturbation of the scalar field where the solution ϕ is kept constant in the causal past of the singularity, so as to preserve the coefficients and right hand side of the Equation (2.9), and then introducing a perturbation of \mathfrak{p} exactly along $\underline{\mathcal{C}} \cap \mathcal{C}$. Since the transversal derivative of ϕ thus must have a jump discontinuity, this leads to the perturbation being only in the class of solutions of bounded variation. However, it turns out that by exploiting the gap between $1+k^2$ and 1, one can also excite a similar instability with a perturbation which again keeps ϕ constant in the causal past of singularity and arranging for the perturbed \mathfrak{p} to be Hölder continuous with a small k -dependent exponent $\gamma > 0$. See [19] for more details.

We now turn to the much more general instability results of [18].

Theorem 7. [18] *We parameterize initial data for the spherically symmetric Einstein-scalar field system by a choice of $\alpha = \partial_\nu(r\phi)$ along an initial outgoing cone \mathcal{C} , where $\partial_\nu r = 1/2$. Suppose that for some $\alpha = \alpha_0$, a function of bounded variation, the corresponding spacetime has an incomplete future null infinity and that, along the past light cone of the singularity, $(m/r) \not\rightarrow 0$ as $r \rightarrow 0$. Then there exists f_1 which is a function of bounded variation and f_2 which is an absolutely continuous function so that whenever $(c_1, c_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ the spacetime arising from the initial data $\alpha_0 + c_1 f_1 + c_2 f_2$ has a complete future null infinity.*

We note that the proof also applies to the case of a *locally naked singularities*. As in the instability of the specific naked singularities discussed above, a key role is played by a blue-shift instability along the past light cone of the singularity. However, in this more general scenario, we do not have precise information about the geometry available. Remarkably, using only the information on m/r as a starting point, Christodoulou is able to show that a blue-shift instability may be triggered by a generic (in the sense of Theorem 7) perturbation, and eventually leads to the conditions of Theorem 6 being satisfied. Finally, we note that [18] justifies the assumption on m/r along the past light cone of the singularity by a connection to continuation criteria for solutions of bounded variation from [17] (see also the corresponding discussion in [11]).

These instability proofs have been revisited with techniques developed originally for use with the Einstein vacuum equations (outside of symmetry): such techniques also allow for the study of certain instabilities generated by non-spherically symmetric and anisotropic gravitational perturbations (see [27–29]). We note in particular the recent work [29] where An revisits the instability for the specific k -self similar naked singularities of Christodoulou. By considering anisotropic perturbations An is able to construct infinitely many more unstable directions. For these perturbations he is furthermore able to construct a corresponding anisotropic apparent horizon and to consider perturbations which vanish at the past light cone of the singularity when measured using suitable scale-invariant norms. This final fact is related to the discussion earlier in the section that the blue-shift instability can be triggered by scalar field perturbations which have a Hölder continuous first derivative as opposed to the derivative merely being absolutely continuous.

2.5. Epilogue: the importance of the functional framework

The naked singularities considered in Theorem 7 arose from data at least as regular as a solution of bounded variation and the perturbations considered could be as regular as a solution with $\partial_\nu(r\phi)$ being an absolutely continuous function. It is natural to ask what happens if we consider naked singularities arising from more regular initial data but then only allow perturbations which are also more regular. In fact, this question is already interesting if we restrict to the k -self similar naked singularities constructed by Christodoulou; for these solutions $\partial_\nu(r\phi)$ is Hölder continuous.

In a very interesting recent work [19], Singh has initiated the study of these questions by considering the dynamics of the wave equation $\square_g \phi = 0$ (with no symmetry restriction) on the k -self similar naked singularity backgrounds. If the initial data for $(\partial_\nu r)^{-1} \partial_\nu \phi$ lies in a more regular Hölder space than the regularity of the background scalar field, then Singh showed that the blue-shift instability is not present, and the corresponding solution to $\square_g \phi = 0$ obeys bounds which are less singular than the scalar field of the background. For initial data which lies in the same Hölder space as the background scalar field, the solution obeys bounds which are analogous to the background scalar field. In so far as solutions to the wave equation $\square_g \phi = 0$ are a proxy for perturbations to the Einstein-scalar field system and non-spherically symmetric gravitational perturbations, these results are consistent with the k -self similar singularities being nonlinearly

stable to sufficiently regular perturbations! We conclude that the validity of the weak cosmic censorship conjecture may depend sensitively on the precise functional framework which is used both to define admissible spacetimes and “generic” perturbations. A similar connection between the validity of the strong cosmic censorship conjecture and the specific functional framework used has been an active area of research, for example, see [15, 30–32].

3. Trapped surface formation in vacuum

We have seen in Section 2 that trapped formation plays an important role in issues related to weak cosmic censorship. In moving from the spherically symmetric Einstein-scalar field system to the vacuum Einstein equations with no symmetry restrictions, the following simple sounding questions already turn out to be highly non-trivial:

Question 8. *Is it possible for a trapped surface to form in evolution from trapped, marginally trapped, anti-trapped, and marginally anti-trapped surface free initial data with one asymptotically flat end? What about the case of scattering initial data posed on past null infinity \mathcal{I}^- ?*

Note that every Cauchy hypersurface in a Schwarzschild or a Kerr black hole spacetime is both two-ended and possesses either a trapped, anti-trapped, or marginally trapped surface. In contrast, the natural setting for the study of gravitational collapse is in the evolution of Cauchy data with one asymptotically flat end which is far from possessing any trapped or anti-trapped surface.

Question 8 was answered positively by Christodoulou in the monumental work [33]. Moreover, what is even more significant than the knowledge of a positive answer to Question 8, is that the work [33] introduced an entirely new paradigm for understanding solutions to the Einstein vacuum equations for (relatively) long times in certain large data regimes and with no symmetry. In the following sections we will start with a brief description of [33] and then we will proceed to describe various follow-up works which have introduced powerful new ideas and lead to many extensions of the original result of Christodoulou. The end result is that we currently have a quite sophisticated understanding of trapped surface formation even if we still do not quite have a theorem for the Einstein vacuum equations which is as strong as Theorem 6. (It should be noted, however, that it is far from clear that such an analogue can exist without symmetry assumptions.)

Christodoulou’s work [33] and essentially all of the follow-up works on trapped surface formation have worked with the Einstein vacuum equations expressed in a double-null gauge (note, however, the interesting recent exception [34]). Many of these works have furthermore fundamentally exploited the characteristic initial value problem. We assume familiarity with these in the following sections and have included brief reviews in Sections A–C of the Appendix.

3.1. The short pulse method of Christodoulou

In this section we will review the short pulse method introduced in the work [33]. For a more thorough review of the original work [33] and a broader discussion of the relevant context, in addition to the original text [33], we recommend the article [35].

A key idea in [33] was to consider a certain class of characteristic initial data for the Einstein vacuum equations which allows for some “largeness” to be pumped into the system in a controlled way. We now give a brief overview of this strategy. We consider a characteristic initial value problem with data posed along two transversely intersecting null hypersurfaces $\{u = u_0\}$ and $\{v = 0\}$ for $u_0 < 0$ and $|u_0| \gg 1$. In order to eventually obtain scattering data on \mathcal{I}^- , one should keep in mind that we will eventually take the limit $u_0 \rightarrow -\infty$. Along the incoming hypersurface $\{v = 0\}$ Christodoulou placed trivial Minkowskian data. Along $\{u = u_0\}$, we take $\Omega = 1$.

The key starting point for the short pulse method is to choose the remaining characteristic data along $\{u = u_0\}$ so that, in orthonormal frames on \mathbb{S}^2 ,

$$\hat{\chi}_{AB}(v, \theta^C) = |u_0|^{-1} \delta^{-1/2} f_{AB} \left(\frac{v}{\delta}, \theta^C \right), \tag{3.1}$$

where $0 < \delta \ll 1$ may be taken arbitrarily small, and $f_{AB}(x, \theta^A)$ is a smooth symmetric $(0, 2)$ -tensor compactly supported in $(x, \theta^A) \in (0, 1) \times \mathbb{S}^2$ and satisfying, for each $\theta^C \in \mathbb{S}^2$,

$$\int_0^1 (f_{AB}(x, \theta^C))^2 dx = c_0, \quad \|f_{AB}\|_{C^k((1,2) \times \mathbb{S}^2)} \lesssim 1, \tag{3.2}$$

where $c_0 \sim 1$ is a suitable δ -independent constant and k is a suitably large δ -independent integer. The gravitational radiation field along \mathcal{S}^- is given by $\lim_{u \rightarrow -\infty} \hat{\chi}_{AB}(v, \theta^C)|u|$, and, in the limit as $u_0 \rightarrow -\infty$, Equations (3.1) and (3.2) end up corresponding to a lower bound for the incoming energy per unit solid angle in each direction in the advanced time interval $[0, \delta]$.

There are now two main interrelated issues to understand. First, what is the explicit mechanism we expect to lead to the formation of the trapped surface? Second, in view of the fact that (3.1) manifestly puts us in a large data regime, why can we expect to construct enough of the spacetime to show that this mechanism takes place?

The first step in the analysis is, still at the level of the initial null hypersurface $\{u = u_0\}$, to obtain suitable estimates for all of the Ricci coefficients and null curvature components. Given the complexity of the equations of the double-null gauge (see Section A) this at first appears like a task of daunting complexity. However, Christodoulou exploits a remarkable inductive structure in the Einstein equations, which enables one to sequentially estimate all of the required terms in terms by solving transport and elliptic equations only in terms of quantities already controlled. A key point in this analysis is that while (3.1) is large in L_v^∞ , it is small in any L_v^p with $p \in [1, 2)$, and thus solutions to transport equations in this process often gain smallness in δ . We list here the results of this analysis for the null curvature components:

$$|\alpha| \lesssim \delta^{-3/2} |u_0|^{-1}, \quad |\beta| \lesssim \delta^{-1/2} |u_0|^{-2}, \quad |\rho| + |\sigma| \lesssim |u_0|^{-3}, \tag{3.3}$$

$$|\underline{\beta}| \lesssim \delta |u_0|^{-4}, \quad |\underline{\alpha}| \lesssim \delta^{3/2} |u_0|^{-5}. \tag{3.4}$$

We note that the u_0 weights here reflect a hierarchy consistent with *peeling* (see the discussion in [35]). It is also worth remarking that the powers of δ one sees in (3.3) and (3.4) deviate from what one would expect purely based on linear considerations.

The bulk of the technical work in [33] is concerned with showing that these hierarchies of estimates understood for the characteristic initial data can, with a few modifications, be propagated into the spacetime in a region $\{(u, v, \theta^A) \in [u_0, \check{u}] \times [0, \delta] \times \mathbb{S}^2\}$. Here $\check{u} < 0$ may be taken arbitrarily small as long δ is then taken sufficiently small depending on \check{u} . In order to prove this “semi-global” existence result, in addition to transport and elliptic equation estimates, Christodoulou introduces an intricate scheme of energy estimates using weighted vector fields in conjunction with the Bel–Robinson tensor. While we will not discuss these estimates in more detail here, it is worth emphasizing that it is a remarkable fact that the structure behind the initial data hierarchy can be exploited in evolution all the way from $\{u = u_0\}$ to $\{u = \check{u}\}$!

Now that we have discussed the spacetime construction, we can explain at a schematic level how the formation of the trapped surface occurs. First we must guarantee that there is no trapped surface in the initial data. In view of (3.1) and (3.2), we have that $\|\hat{\chi}\|_{L_{\mathbb{S}^2}^\infty L_{u=u_0}^2}^2 \lesssim c_0 |u_0|^{-2}$. It then follows by integrating Raychaudhuri’s Equation (A.3), the relation $\mathcal{L}_{e_4} \mathbf{g} = \chi$, the fact that $\Omega = 1$ along $\{u = u_0\}$, and the triviality of the data along $\{v = 0\}$, that we have, along $\{u = u_0\}$

$$\left| \text{tr} \chi - \frac{2}{|u_0|} \right| \lesssim \frac{c_0}{|u_0|^2}.$$

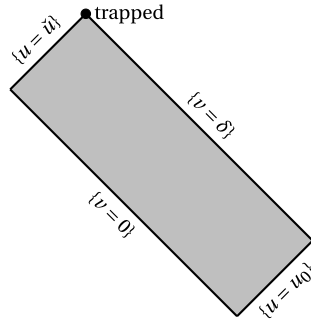


Figure 5. Penrose diagram for short pulse spacetime.

In particular, it is clear that $\text{tr}\chi > 0$ and we are far from having a trapped surface along $\{u = u_0\}$. Next we would like to show that a trapped surface emerges at $(u, v) = (\check{u}, \delta)$. Our plan will be again to eventually integrate the Raychaudhuri Equation (A.3); however, first we will show $\hat{\chi}$ undergoes an amplification while evolving from $u = u_0$ to $u = \check{u}$. In order to see we use the ∇_3 propagation equation for $\hat{\chi}$ (A.5) as well as the following upper bounds which are a consequence of the estimates in the aforementioned semi-global existence result:

$$\left| \text{tr}\underline{\chi} - \frac{2}{u} \right| \lesssim \delta |u|^{-2}, \quad |\Omega\underline{\omega}| \lesssim \delta |u|^{-3}, \quad |\eta| + |u|\delta^{1/2}|\nabla\eta| \lesssim \delta^{1/2}|u|^{-2},$$

$$|\hat{\chi}| \lesssim \delta^{1/2}|u|^{-2}, \quad |\log\Omega| \lesssim \delta |u|^{-3}.$$

In particular we may write (A.5) as

$$\nabla_3 \hat{\chi} - (|u|^{-1} + O(\delta|u|^{-2}))\hat{\chi} = O(|u|^{-3}).$$

Keeping in mind that δ may be assumed small depending on \check{u} , integrating from $\{u = u_0\}$ to $\{u = \check{u}\}$ then yields, in an orthonormal frame

$$\hat{\chi}_{AB}(v, \theta^C)|_{u=\check{u}} = |\check{u}|^{-1} \delta^{-1/2} f_{AB} \left(\frac{v}{\delta}, \theta^C \right) + O(|\check{u}|^{-2}).$$

Now we turn again to the Raychaudhuri Equation (A.3) which implies

$$\partial_v(\Omega^{-1} \text{tr}\chi) \leq -|\hat{\chi}|^2. \tag{3.5}$$

Since we have Minkowskian data along $\{v = 0\}$, we integrate (3.5) from $v = 0$ along $\{u = \check{u}\}$ and obtain

$$\Omega^{-1} \text{tr}\chi|_{(u,v)=(\check{u},\delta)} \leq \frac{2}{|\check{u}|} - \frac{c_0}{2|\check{u}|^2} \Rightarrow \text{tr}\chi < 0.$$

In view of the smallness of \check{u} we have the formation of a trapped surface.

The initial data for Christodoulou’s short pulse spacetimes are given either on null hypersurfaces, or, once the limit $u_0 \rightarrow -\infty$ is taken, on past null infinity. It is natural to also ask for a version of this result with asymptotically flat initial data given along a spacelike hypersurface free of trapped surfaces. Exactly this construction was achieved in [36]. The proof used both the geometric information obtained from Christodoulou’s estimates as well as gluing techniques for the constraint equations developed by Corvino–Schoen [37]. We close the section by noting the recent work [38] which provides a new proof of the formation of trapped surfaces from trapped surface free Cauchy data in the context of C^2 spacetimes. The proof is quite short compared to other such results and employs a short pulse ansatz and null gluing techniques. An interesting aspect of the proof is that the only estimates for the Einstein equations in evolution needed are those following from Cauchy stability (in particular it bypasses any need to understand a semi-global existence problem).

In Figure 5 we have drawn the Penrose diagram for Christodoulou’s short pulse spacetime.

3.2. A relaxed hierarchy and anisotropic trapped surface formation

Soon after Christodoulou’s original work [33], Klainerman–Rodnianski, focusing on the case of fixed u_0 , introduced a new strategy for the semi-global existence result which leads to significant simplifications and opens the door to considering trapped surface formation by anisotropic initial data [39, 40]. The basic idea can be explain as follows (we take $u_0 = -1$ in what follows). Christodoulou’s choice of data for $\hat{\chi}$ (3.1) leads to the following upper bounds for any pair of integers n and m with $n + m \leq k$ (here k is as in (3.2))

$$\delta^n \|\nabla_4^n \nabla^m \hat{\chi}\|_{L^2_{\{u=-1\}}} \lesssim 1. \tag{3.6}$$

Klainerman–Rodnianski instead consider a wider class of initial data, where one only requires the following weaker bound along $u = -1$:

$$\delta^{n+m/2} \|\nabla_4^n \nabla^m \hat{\chi}\|_{L^2_{\{u=-1\}}} \lesssim 1. \tag{3.7}$$

There are two main motivations for considering data of the form (3.7). First, the upper bound (3.7) is consistent with considering an initial $\hat{\chi}$ along $\{u = -1\}$ of the form

$$\hat{\chi}_{AB}(v, \theta^C) = \delta^{-1/2} f_{AB} \left(\frac{v}{\delta}, \frac{\theta^C}{\delta^{1/2}} \right).$$

This allows one to consider initial data which is highly localized in a particular angular region of the sphere. This is of interest in order to study various anisotropic phenomenon.

Before explaining the second motivation, we need to review the notion of *signature* for the Einstein vacuum equations. For any Ricci coefficient or curvature component ψ , we define

$$s(\psi) = N_4(\psi) + \frac{1}{2} N_A(\psi) - 1,$$

where N_4 is the number of times L appears in the definition of the component ψ and N_A is the number of angular frames:

$$\begin{aligned} s(\alpha) &= 2, & s(\beta) &= 3/2, & s(\rho, \sigma) &= 1, & s(\underline{\beta}) &= 1/2, & s(\underline{\alpha}) &= 0, \\ s(\underline{\chi}, \underline{\omega}) &= 1, & s(\underline{\zeta}, \underline{\eta}, \underline{\eta}) &= 1/2, & s(\underline{\chi}, \underline{\omega}) &= 0, \\ s(\nabla_4 \phi) &= 1 + s(\phi), & s(\nabla \phi) &= 1/2 + s(\phi), & s(\nabla_3 \phi) &= s(\phi). \end{aligned}$$

We also set

$$s(\phi_1 \cdot \phi_2) = s(\phi_1) + s(\phi_2),$$

where \cdot stands for any type of contraction. A key point behind the usefulness of the signature concept is that the equations of the double-null gauge preserve signature in that all terms in any given equation have the same signature. Signature is furthermore preserved upon commutation with ∇_4 , ∇_3 , and ∇ . The weaker bound (3.7) turns out to mesh nicely with a hierarchy of δ -bounds which is based on signature, and this ultimately allows for many simplifications in the corresponding semi-global existence result. We close by noting that this relaxed hierarchy turns to not quite be strong enough to see the amplification of $\hat{\chi}$ discussed at the end of Section 3.1, however, by interpolating the relaxed bounds with some other relatively straightforward bounds based on (3.6), the work [39] is also able to recover the trapper surface formation result of Christodoulou. The work [41] extended these methods to also allow for data on past null infinity as in Christodoulou’s original work.

While the works [39, 40] allowed for highly anisotropic data (3.7) they were not quite able to prove that trapped surfaces arise from such data (though they did show the formation of so-called pre-scarred surfaces where $\text{tr}\chi < 0$ on an arbitrarily large portion of a sphere). In the work [42] Klainerman, Luk, and Rodnianski were able to establish a fully anisotropic trapped surface formation result. The assumptions of the main theorem can be taken to be the same as

Christodoulou's work [33] except with one significant change, namely that instead of the left hand side of (3.2) holding for every point in \mathbb{S}^2 , we only now require that

$$\sup_{\theta^A \in \mathbb{S}^2} \int_0^1 (f_{AB}(x, \theta^A))^2 dx \sim 1. \quad (3.8)$$

The key idea in the work [42] was to look for a trapped surface lying along $\{v = \delta\}$ but with respect to a foliation by a new function \tilde{u} as opposed to u . At a technical level, the work [42] shows that if one perturbs the null coordinate u along $\{v = \delta\}$ with the addition of a function $R(\theta^A)$ defined along \mathbb{S}^2 , then the outgoing second fundamental forms for the new spheres have a transformation formula which involves an elliptic operator applied to the function R . It is then shown that the assumption (3.8) is sufficient to find an R satisfying a corresponding elliptic differential inequality leading to a trapped surface in the new foliation.

3.3. Renormalization: impulsive waves and applications to trapped surface formation

A connection can be drawn between the semi-global existence result underlying Christodoulou's trapped surface formation result and low-regularity well-posedness for the Einstein vacuum equations. Namely, since the ansatz (3.1) corresponds to the data for the metric being large in H^s , when $s > 1$, bounded when $s = 1$, and small when $s < 1$, one is lead to ask, in general, what type of initial Sobolev regularity is required for well-posedness of the initial value problem. As it turns out, a large literature exists around low-regularity well-posedness results for the Einstein equations which, in view of space constraints, we will not attempt to survey here. We will just point out that the culmination of this line of research was the remarkable collection of works [43–47] which, roughly speaking, establish that the Einstein vacuum equations are locally well-posed when the initial spacetime curvature lies in L^2 along a Cauchy hypersurface. Moreover, in view of forthcoming work of Luk–Moschidis (see Theorem 4.6 of [48]), this well-posedness result is sharp when regularity is measured in standard Sobolev spaces.

The works [49, 50] by Luk and Rodnianski initiated a new paradigm for considering certain classes of low-regularity solutions to the Einstein vacuum equations. These works were motivated by the desire to establish a local well-posedness result for the propagation (and possible interaction) of *impulsive gravitational waves* (see [51, 52]). Certain components of the curvature tensors of such spacetimes possess δ -function singularities which propagate along null hypersurfaces; for one or two propagating impulsive waves in a suitable double-null foliation, the singular curvature components correspond to α and $\underline{\alpha}$. One key insight from [49, 50], which has played a crucial role in many later developments, is that α and $\underline{\alpha}$ can be removed from the Bianchi equations for curvature in such a way that the remaining system still allows for energy estimates (see the equations of Section B). The upshot of this theory is that one has local well-posedness for characteristic initial value problems when, schematically put, the shears $\hat{\chi}$ and $\underline{\hat{\chi}}$ satisfy that

$$\sum_{i=0}^3 \|\nabla^i \hat{\chi}\|_{L^2_{u=u_0}} + \sum_{i=0}^3 \|\nabla^i \underline{\hat{\chi}}\|_{L^2_{v=v_0}} \doteq \mathcal{A}(\hat{\chi}, \underline{\hat{\chi}}) < \infty.$$

Remarkably, Christodoulou's initial data (3.1) exactly satisfies that $\mathcal{A}(\hat{\chi}, \underline{\hat{\chi}}) \lesssim 1$, and so from the “renormalized” point of view, the semi-global existence result now becomes a question of local existence. Using this, the work [50] is able to attain another proof both of Christodoulou's semi-global existence result and the corresponding trapped surface formation. Furthermore, this new proof allowed, for the first time, for non-trivial data along the incoming null hypersurface $\{v = 0\}$.

One natural question left open by all of these various results on trapped surface formation, is whether anything can be said about the short pulse spacetimes as the parameter $\delta \rightarrow 0$. This question has been addressed in [53] where the authors employ their renormalization techniques

to understand the connection between certain types of high-frequency limits of the Einstein equations and null dust solutions⁴. In particular, they show that one may take the $\delta \rightarrow 0$ (weak) limit of the short pulse spacetimes and obtain a solution to the Einstein-null dust system which also exhibits the formation of a trapped surface.

3.4. Scale-invariant trapped surface formation

Christodoulou’s trapped surface criteria in Theorem 6 is scale invariant in that it is invariant under the rescaling map (2.4). It is natural to ask if we can also have a scale invariant criteria for trapped surface formation in vacuum, though first we must formulate the question more precisely. In any given set of local coordinates $\{x^\alpha\}$, the analogue of the rescaling map (2.4) is simply the map $x^\alpha \mapsto \lambda x^\alpha$ for $\lambda > 0$. The Sobolev space $\dot{H}^{3/2}$ is the unique homogeneous Sobolev space which is left invariant under this rescaling. Thus, we ask if one can find a trapped surface formation criteria which only requires largeness of the initial data for the metric in H^s for some $s \geq 3/2$.

The ansatz (3.1) leads to initial data which large in $H^{1+\epsilon}$, for any $\epsilon > 0$. Thus a new type of data is required. This new class of data was provided in the work [56] by An and Luk which succeeded in establishing the desired scale-invariant trapped surface formation result. While the work [56] in fact includes a large family of data which interpolates between Christodoulou like short pulse data and scale-invariant data, here we will just describe an example of the scale-invariant like data considered. Let $A \gg 1$ and $B \gg 1$ and then choose $0 < \delta \ll 1$, depending on A and B . We then take the initial data along $\{v = 0\}$ to be trivial, and suppose that $\hat{\chi}$ along $\{u = -1\}$ satisfies the following upper and lower bounds (see the discussion in the introduction to [57]):

$$\sum_{i=0}^5 \|\nabla^i \hat{\chi}\|_{L^\infty L^2(\mathbb{S}^2_{-1,v})} \lesssim A, \quad \inf_{\omega \in \mathbb{S}^2} \delta^{-1} \int_0^\delta |\hat{\chi}(v, \omega)|^2 dv \geq 4A^2. \tag{3.9}$$

Then the solution exists in a characteristic rectangle $(u, v, \theta^A) \in [-1, -\delta BA] \times [0, \delta] \times \mathbb{S}^2$, and the sphere at $(u, v) = (-\delta BA, \delta)$ is trapped.

To see the connection to $\hat{\chi}$ being small in $H^{1/2-s}$ for $s > 0$ (and hence the characteristic data being small in $H^{3/2-s}$), it is useful to first imagine that $\hat{\chi}$ is a scalar function $\tilde{\chi}(v, \theta^A)$, in which case, we could set $\tilde{\chi}(v, \theta^A) = CAf(v/\delta)$ for any smooth, non-zero, compactly supported function $f(x) : [0, 1] \rightarrow \mathbb{R}$ and suitable constant C (independent of δ). This would satisfy (3.9), and we then have, for any $s \in [-1/2, 1/2]$

$$\|\tilde{\chi}\|_{H^{1/2-s}} \lesssim \|\tilde{\chi}\|_{L^2}^{1/2+s} \|\tilde{\chi}\|_{H^1}^{1/2-s} \lesssim (A\delta^{1/2})^{1/2+s} (\delta^{-1/2}A)^{1/2-s} \lesssim A\delta^s.$$

To actually construct true scale-invariant characteristic initial data for $\hat{\chi}$, one has to work a bit harder since topological obstructions prevent the existence of a nowhere vanishing trace-free $(0, 2)$ -tensor on \mathbb{S}^2 . Nevertheless, it remains possible, see Appendix C of [57] for an explicit such construction.

We close the section by noting the sequence of works [57–60] where An and collaborators have studied the emergence of apparent horizons and further refined the various trapped surface formation results discussed above. We note in particular the work [58] which by use of signature considerations and a rescaling argument is able to provide new simpler proofs of the scale-invariant trapped surface formation results from [56] and also to connect the analysis to the peeling properties of the Einstein equations.

⁴These questions are motivated by Burnett’s conjecture [54]; see the survey article [55].

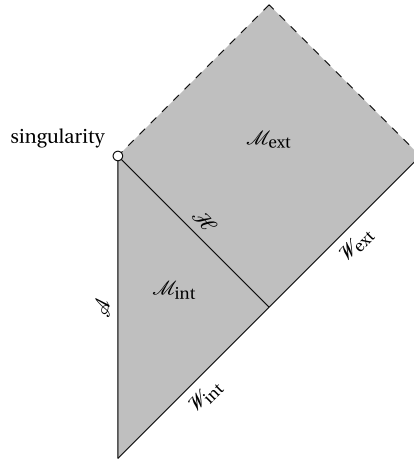


Figure 6. Penrose diagram for vacuum naked singularity.

4. Self-similarity and naked singularities in vacuum

In view of Christodoulou’s construction [14] of a naked singularity for the spherically symmetric Einstein-scalar field system, it is natural to try to construct naked singularity solutions for the Einstein vacuum equations. Unfortunately, there is no direct way to import Christodoulou’s construction to the Einstein vacuum equations in view of the following:

- (1) Birkhoff’s theorem states that all spherically symmetric solutions to the Einstein vacuum equations are isometric to Schwarzschild. Thus, we cannot expect to simply turn off the scalar field in Christodoulou’s example.
- (2) Christodoulou’s analysis relied heavily on the fact that the spherically symmetric Einstein-scalar field system reduced, upon the imposition of self-similarity, to a two dimension autonomous system. There is no such reduction for the Einstein vacuum equations which is also consistent with asymptotic flatness.

Nevertheless, we have the following theorem:

Theorem 9. [61, 62] *There exists naked singularity solutions for the Einstein vacuum equations.*

In Figure 6 we have depicted a Penrose diagram of the naked singularity solution. (Note that other than the use of different labels for the hypersurfaces, the diagram is the same as Figure 4.) In the following sections we will describe in more details various properties of these solutions and discuss some of the main ideas which go into their construction.

4.1. Qualitative properties and comparison with the solutions of Christodoulou

In this section we will compare the naked singularity solutions of Christodoulou with the solutions produced by Theorem 9. When drawing analogies between solutions to the Einstein vacuum equations and solutions to the spherically symmetric Einstein-scalar field system, it is useful to identify the following:

$$\partial_u \phi \leftrightarrow \hat{\chi}, \quad \partial_v \phi \leftrightarrow \hat{\chi}.$$

We will see that both classes of naked singularities share many qualitative properties.

As with Christodoulou’s naked singularities, the vacuum naked singularities arise from asymptotically flat data along a null hypersurface. We recall that for Christodoulou’s solutions, the initial

characteristic data for the scalar field satisfies that $(\partial_\nu r)^{-1} \partial_\nu \phi$ is smooth except when it crosses the past light cone of the singular point where it is Hölder continuous. The initial characteristic data for the solutions of Theorem 9 can be considered analogous in that, with respect to a suitable ν -coordinate and for any $N \in \mathbb{Z}$ with $N \gg 1$, Theorem 9 produces solutions so that the initial $\hat{\chi}$ is C^N except where it crosses the past light cone of the singularity. Across the past light cone of the singularity, we have that $\hat{\chi}$ is Hölder continuous. The loss of regularity for $\hat{\chi}$ is only with respect to ∂_ν ; it in fact is $C_{\mathbb{S}^2}^N C_v^{1,\gamma}$.

We cannot appeal to Christodoulou’s well-posedness result in BV in order to justify that this class of initial data is “sufficiently regular”. Instead we start by observing that $\hat{\chi}$ being Hölder continuous is more than what it is needed for the local theory developed by [49, 50]. The analysis from [49, 50] did not attempt to cover the solution near the “axis” but, for this type of data with a Hölder continuous $\hat{\chi}$, one expects such a well-posedness result to hold.

We now turn to a discussion of the main singularity of the spacetime. We have already mentioned that Christodoulou’s naked singularities cannot be extended to the singular point in such a way as to remain a solution of bounded variation. One way to see this is to observe that along the past light cone of singularity \mathcal{C} we have

$$\int_{\mathcal{C}} |\partial_u \phi| du = \infty.$$

For the naked singularities of Theorem 9 we have the following analogous fact. There exists (at least one) null geodesic γ converging to the singularity so that

$$\int_{\gamma} |\hat{\chi}| ds = \infty.$$

In fact, one can also check that there exists Jacobi fields which blow-up along γ and that the singularity is a *strong singularity* in the sense of Tipler [63].

4.2. Self-similarity for the Einstein vacuum equations

A fundamental role in the naked singularity construction is played by various types of *self-similarity*. In Section 4.2.1 we discuss an important precursor, the self-similar spacetimes of Fefferman–Graham. Then in Section 4.2.2 we will discuss a new type of self-similar spacetimes and discuss why one could expect them to be useful for the construction of naked singularities.

4.2.1. The self-similarity of Fefferman–Graham

Fefferman–Graham [64, 65] were the first to systematically study a class of self-similar solutions to the Einstein vacuum equations without any symmetry assumptions. Their class of spacetimes was the following:

Definition 10. We say that a spacetime (\mathcal{M}, g) is Fefferman–Graham self-similar (FGSS) if (\mathcal{M}, g) may be equipped with a double-null foliation (A.1) which covers $u \in (-\infty, 0)$ and at least one of $(v/u) \in [0, c)$ or $(v/u) \in (-c, 0]$ for some small constant c , and so that $K = u\partial_u + v\partial_v$ satisfies

$$\mathcal{L}_K g = 2g. \tag{4.1}$$

Using (4.1), a straightforward computation shows that the metric components Ω , b , and g of a FGSS spacetime must satisfy in coordinate frames that

$$\begin{aligned} \Omega(u, v, \theta^A) &= \tilde{\Omega}\left(\frac{v}{u}, \theta^A\right), & b^A(u, v, \theta^B) &= u^{-1} \tilde{b}^A\left(\frac{v}{u}, \theta^B\right), \\ g_{AB}(u, v, \theta^C) &= u^2 \tilde{g}_{AB}\left(\frac{v}{u}, \theta^C\right), \end{aligned} \tag{4.2}$$

for some $\tilde{\Omega}$, \tilde{b} , and \tilde{g} .

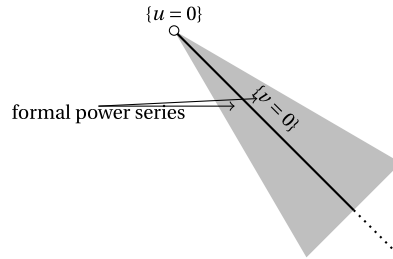


Figure 7. Region covered the Fefferman–Graham formal expansions.

Fefferman–Graham did not actually work in the double-null gauge, but it is possible to translate their study of self-similar solutions to the double-null gauge. As a preliminary step in their program to understand local conformal invariants, Fefferman–Graham studied formal power series expansions for FGSS spacetimes. If we specialize their results concerning these formal expansions to spacetime dimension 4 and recast them in the language of the double null-gauge, we can informally⁵ phrase their result as follows:

Theorem 11. *For every smooth metric g_0 on S^2 and smooth g_0 -trace-free symmetric $(0,2)$ -tensor k_0 on S^2 , there exists a smooth FGSS metric g defined in the double-null gauge (A.1) for $\{0 \leq (v-u) < 1/2\} \cap \{u \in (-\infty, 0)\}$ so that g has a unique formal Taylor series expansion in v/u , $g \approx \sum_{j=1}^{\infty} g_j$, and*

- (1) *The induced metric g on $S_{u,v}^2$ satisfies $g|_{v=0} = u^2 g_0$ in the coordinate frame, the quantities b and $\log \Omega$ associated to g vanish along $\{v = 0\}$, and the restriction of $\hat{\chi}_{AB}$ to $\{v = 0\}$ satisfies $\hat{\chi}_{AB} = u(k_0)_{AB}$ in the coordinate frame.*
- (2) *For every sufficiently large positive integer n we have $\text{Ric}(\sum_{j=1}^n g^{(j)})(X, Y) = O(|v/u|^n)$, for $X, Y \in \{\partial_v, \partial_u, u^{-1}\partial_{\theta^A}\}$.*

In Figure 7 we depict the region of spacetime formally covered by these expansions:

We note that if g_0 and k_0 are analytic, then the formal expansions produced by the above theorem will actually converge (see [65]). If we leave the analytic category, then we can appeal to the work [66] to show the existence of a true solution in the region to the future of $\{v = 0\}$ provided g and k_0 lie in suitable Sobolev spaces.

In particular, in view of (4.2) it is plausible that a FGSS spacetime cannot be extended in a regular fashion to $u = 0$ along curves of constant v/u . Thus it may appear a reasonable strategy to use a FGSS spacetime to try to build a naked singularity⁶. However, the formal theory of Fefferman–Graham and its extensions in [66] only concern the description of the spacetime locally near $\{v = 0\}$, while the construction of a naked singularity involves a global analysis, that is, we need our spacetime to contain a hypersurface corresponding to complete and asymptotically flat initial data. The global analysis turns out to be problematic because there is no known way to extend a FGSS spacetime to the past of $\{v = 0\}$ in such a way which is compatible with the existence of a past complete Cauchy hypersurface other than by gluing a *flat* spacetime to the past of $\{v = 0\}$. Unfortunately, such a flat extension to the past of $\{v = 0\}$ prevents the resulting spacetime from being considered (part of) a naked singularity for (at least) the following two reasons:

⁵See [66, 67] for precise statements.

⁶We note that even if one did not know about the work of Fefferman–Graham, in view of the long tradition of the use of self-similarity to produce singular solutions to PDEs in mathematical physics, it would be reasonable to try to search for self-similar solutions of the Einstein equations as part of the pursuit of naked singularities.

- (1) If the spacetime is flat to the past of $\{v = 0\}$, then it is a consequence of the formal theory of Fefferman–Graham that in order for the spacetime to not be globally flat, we must allow $\hat{\chi}$ to have a jump discontinuity along $\{v = 0\}$. This lowers the regularity of whatever we pick to be our Cauchy hypersurface, the resulting spacetime represents an impulsive gravitational wave with an ingoing spherical front, and it is difficult to argue that the “singularity” at $\{u = 0\}$ represents a loss of regularity.
- (2) If the spacetime is flat to the past of $\{v = 0\}$ then any causal curve entering the “singularity” at $\{u = 0\}$ will only pass through a flat spacetime and thus not actually encounter any truly singular behavior.

4.2.2. Twisted self-similarity

One of the key ideas in our work [61] was the introduction of a generalization of FGSS spacetimes which is consistent with the possibility of non-flat, past complete extensions to the past of $\{v = 0\}$. We will give a brief introduction in this section to this new type of self-similarity. We direct the reader to the work [67] for a systematic treatment.

The rigidity of FGSS spacetimes that we have mentioned in the previous section has its source in the following three closely related facts about the geometry of the null hypersurface $\{v = 0\}$ in the FGSS spacetimes:

- (1) The induced (indefinite) metric along $\{v = 0\}$ is $u^2 \tilde{g}_{AB}(\theta^C) d\theta^A d\theta^B$. In view of the uniformization theorem, we can write $\tilde{g}_{AB} = e^{2\varphi} \mathring{g}_{AB}$ where \mathring{g}_{AB} denotes the round metric on the sphere and φ is a function on \mathbb{S}^2 . Then, defining a new coordinate $\tilde{u} \doteq ue^\varphi$, in the (\tilde{u}, θ^A) coordinates, the induced metric along $\{v = 0\}$ becomes $\tilde{u}^2 \mathring{g}_{AB} d\theta^A d\theta^B$. This is identical to the induced metric along $\{v = 0\}$ in Minkowski space!
- (2) A computation shows that all components of the curvature tensor which are tangential to $\{v = 0\}$ must vanish.
- (3) The orbits of the restriction of the conformal Killing field $K = u\partial_u + v\partial_v$ to $\{v = 0\}$ coincide with the orbits of the null generators of $\{v = 0\}$. Equivalently, the shift vector vanishes on $\{v = 0\}$.

In [61] we modified Definition 10 by only requiring that the double-null coordinate system was valid for $(v/-u) \in (0, c)$ instead of $[0, c)$. We then required that the spacetime extends to $\hat{v} = 0$ in a new coordinate system (\hat{v}, u, θ^A) where

$$\hat{v} \doteq v^{1-2\kappa}$$

for some $0 < \kappa \ll 1$. We refer to such a self-similar spacetime as a κ -self-similar (κ SS) spacetime. Of fundamental importance for the construction is the ingoing Raychaudhuri Equation (A.4) along the null hypersurface $\{\hat{v} = 0\}$. If $\lim_{v \rightarrow 0} \Omega^2 (v/-u)^{2\kappa} = 1$, then this becomes the following equation along $\{\hat{v} = 0\}$:

$$\operatorname{div} b - \mathcal{L}_b \operatorname{div} b - \frac{1}{2} (\operatorname{div} b)^2 = \frac{1}{4} |\nabla \hat{\otimes} b|^2 - 4\kappa. \quad (4.3)$$

The presence of the free constant -4κ turns out to allow us to find many solutions to (4.3) with a non-trivial $\nabla \hat{\otimes} b$. If we can construct solutions with this new type of self-similarity and non-trivial $\nabla \hat{\otimes} b$, then the vector field K will still extend to $\{\hat{v} = 0\}$, but now the orbits of K will *not* coincide with null generators of the null hypersurface $\{\hat{v} = 0\}$, and the intrinsic geometry of the $\{\hat{v} = 0\}$ hypersurface will not agree with the intrinsic geometry of a standard null hypersurface in Minkowski space.

Under a suitable smallness assumption on $b|_{v=0}$, one may establish an analogue of the formal expansions of Fefferman–Graham for κ SS spacetimes. While a systematic discussion of these formal expansions is outside the scope of the survey article we will observe here some of the consequences of this analysis:

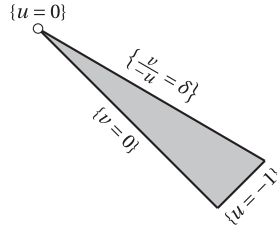


Figure 8. Depiction of first constructed region of exterior of naked singularity.

- (1) The best regularity in the \hat{v} -variable we can expect for a κ SS spacetime is $C^{1,\gamma}$ for $0 < \gamma \ll 1$. This is compensated, however, by additional regularity in the u and angular variables as long as we make certain specific (non-trivial) choices of the shift vector b ; for any $N \geq 1$, if the shift vector takes a certain specific form along $\{v = 0\}$ (to leading order in ϵ) we expect that our spacetime is at least in the class $C_{\hat{v}}^{1,\gamma} C_{u,\theta^A}^N$.
- (2) A key role is played by developing an existence theory and a priori estimates for equations which are of the form $\psi + \mathcal{L}_X \psi + W\psi = H$, where ψ is an unknown tensor on \mathbb{S}^2 and (X, W, H) are given tensors satisfying a suitable smallness condition (cf. (4.3)). It is also important to study the evolutionary analogue of these equations: $(-v)\mathcal{L}_{\partial_v} \psi + \psi + \mathcal{L}_X \psi + W\psi = H$, where now all of the tensors additionally depend on $v \in (0, c)$ for some $c > 0$.
- (3) As in the choice of (g_0, k_0) in the Fefferman–Graham expansions, there is a certain set of “free data” which parametrizes the expansions for κ SS solutions.

Finally, we note that one may in fact consider versions of the above modification of FGSS where we allow for the form of the regular coordinate \hat{v} to be much more general than $\hat{v} \doteq v^{1-2\kappa}$. We refer to this class of more general spacetimes as “twisted self-similar” (TSS) spacetimes, and we systematically studied them in the work [67] (some specific instances of TSS spacetimes are also important in [62], see Section 4.4 below).

4.3. The exterior

In this section we will discuss construction of the *exterior* region of the naked singularity which was carried out in [61]. This is the region to the future of the past light cone of the singularity and is labeled \mathcal{M}_{ext} in Figure 6.

Let $0 < \epsilon \ll \delta \ll 1$. We start by constructing our desired spacetime in the region $\{(v/-u) \in [0, \delta]\} \cap \{u \in (0, -1]\}$ as depicted in Figure 8. While our spacetime will not be an exact κ SS spacetime, it will be approximately κ SS for $\kappa \sim \epsilon^2$ as $v \rightarrow 0$. As with true κ SS spacetimes, we will have a regular coordinate system (\hat{v}, u, θ^A) for $\hat{v} \doteq v^{1-2\kappa}$ where the spacetime extends to $\{\hat{v} = 0\}$. Tangentially along $\{\hat{v} = 0\}$ the spacetime will be exactly κ SS in that

$$(v^\kappa \Omega)|_{\hat{v}=0}(u, \theta^A) = (-u)^\kappa, \quad b^A|_{\hat{v}=0}(u, \theta^B) = u^{-1} \tilde{b}^A(\theta^B), \quad \mathfrak{g}_{AB}|_{\hat{v}=0}(u, \theta^C) = u^2 \tilde{\mathfrak{g}}_{AB}(\theta^C),$$

for suitable \tilde{b} and $\tilde{\mathfrak{g}}$ so that, in particular, the fundamental constraint Equation (4.3) holds. We also require, further, that along $\{\hat{v} = 0\}$, $|b| \sim \epsilon$, and that b takes a certain precise form to leading order as $\epsilon \rightarrow 0$.

Our method for constructing this spacetime is to solve a characteristic initial value problem with data along $\{\hat{v} = 0\}$ as described in the paragraph above, and data along $\{u = -1\}$ as described below. Then, the construction of the spacetime essentially boils down to suitable a priori estimates. We do not have space in this survey article to discuss all of these, in the following we will only discuss some of the ideas involved in the energy estimates for curvature.

In order to keep the presentation here simpler, in our discussion of the outgoing characteristic initial data along $\{u = -1\}$, we will only discuss the choice of $\Omega^{-1}\hat{\chi}$. Our a priori estimates will rely on a renormalization scheme which in turn relies on our spacetime being self-similar to leading order as $(v/-u) \rightarrow 0$ (see the discussion below). In particular, it turns out to be necessary that $\Omega^{-1}\hat{\chi}|_{u=-1} = \vartheta|_{u=-1} + O(v^a)$ for suitable $a \gtrsim 1$ and where ϑ must satisfy

$$v\mathcal{L}_{\partial_v}\vartheta_{AB} + \mathcal{L}_{b|_{v=0}}\vartheta_{AB} - (\frac{1}{2}d\mathbf{v}b|_{v=0} + 2\kappa)\vartheta_{AB} - ((\nabla\hat{\otimes}b)|_{v=0})^C_{(A}\vartheta_{B)C} = H_{AB}, \tag{4.4}$$

where H_{AB} is some tensor independent of v , which is determined in terms of $b|_{v=0}$ and satisfies $|H| \sim \epsilon$. We also have that $\kappa \sim \epsilon^2$ and that the operator

$$\vartheta \mapsto \mathcal{L}_{b|_{v=0}}\vartheta_{AB} - \frac{1}{2}d\mathbf{v}b|_{v=0}\vartheta_{AB} - ((\nabla\hat{\otimes}b)|_{v=0})^C_{(A}\vartheta_{B)C}, \tag{4.5}$$

is anti-symmetric with respect to $L^2(g)$.

Some of the key phenomenon ultimately associated with the solutions ϑ to (4.4) can already be understood if we examine the much simpler scalar equation:

$$v\mathcal{L}_{\partial_v}f - \epsilon^2 f = \epsilon. \tag{4.6}$$

One solution to (4.6) is

$$f = -\frac{1 - v\epsilon^2}{\epsilon}.$$

Even though we will have that $f \sim \epsilon^{-1}$ when $v = 0$, we also have the bound $|f| \lesssim \epsilon|\log(v)|$. For ϑ we also obtain similar estimates; however, a significant additional difficulty, which is not present in the scalar Equation (4.6), is that we also need to control angular derivatives of ϑ , and naive commutation of the Equation (4.4) destroys the underlying anti-symmetric structure. Thus a careful commutation scheme must be designed.

We now briefly discuss the commutation scheme. The first key fact is that we consider characteristic data which is almost axisymmetric. Thus, the commutation errors with the axisymmetric vector field ∂_ϕ will come with a small constant. The next realization is that one can commute with the operator (4.5). The vector field b turns out to be linearly independent of ∂_ϕ whenever it does not vanish and it thus turns out these commutators suffice to control all derivatives of the solution except at the points where ∂_ϕ or b vanish. (We note, however, that in view of the fact that $|b| \lesssim \epsilon$ the estimates we obtain for higher derivatives have a very bad dependence on ϵ^{-1} .) At the points where ∂_ϕ or b vanish we simply commute with ∇ and the potentially problematic terms that show up turn out to have a good sign.

A key idea to obtain the necessary a priori estimates is to subtract off the expected leading order self-similar behavior of certain curvature and Ricci components before carrying energy, transport, or elliptic estimates. For example, the leading order self-similar behavior of α is generated by a self-similar extension of the ϑ we discussed above. We obtain a new renormalized unknowns $\tilde{\alpha}$ by subtracting off a self-similar extension of ϑ . In particular, with the renormalized unknowns we may use weights are (mildly) singular as $(v/-u) \rightarrow 0$. We note that one may anticipate that the error terms generated by this renormalization procedure are small because we know that we can produce formal power series expansions for κ SS solutions which are valid as $(v/-u) \rightarrow 0$. It is furthermore crucial for this scheme, that even though ϑ is size ϵ^{-1} when $v = 0$, it quickly becomes size ϵ once $\log(v) \sim 1$. One particular consequence of these estimates will be that along the hypersurface $(v/-u) = \delta$ all curvature and Ricci components will be small in that (letting Ψ denote a curvature component and ψ a Ricci coefficient)

$$\left| \Psi|_{\frac{v}{-u}=\delta} \right| \lesssim \epsilon|u|^{-2}, \quad \left| \psi|_{\frac{v}{-u}=\delta} \right| \lesssim \epsilon|u|^{-1}.$$

The next region we study is the region where $v/-u$ is bounded away from 0 and ∞ . The relevant Penrose diagram is depicted in Figure 9. The region covered by the dashed lines corresponds to the previous section, and the light-gray filled in-region is what we will discuss in

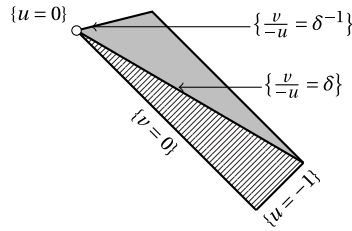


Figure 9. Penrose diagram for second constructed region.

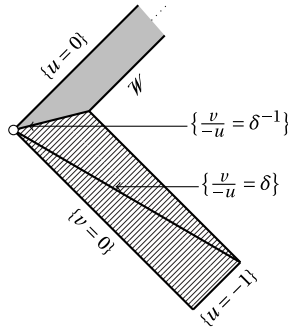


Figure 10. Penrose diagram for the third constructed region.

this section. In this region the estimates are considerably simpler than in the previous region; for example, the key curvature bounds essentially follow from a Grönwall estimate with self-similar weights.

Finally, we come to the construction of the solution up to the hypersurface $\{u = 0\}$. This corresponds to the light gray region in Figure 10 (and the dashed region corresponds to the part we have already constructed).

As a consequences of the of the analysis of the previous region we will already have suitable estimates for all of the double-null quantities along $\{v/-u = \delta^{-1}\}$. We then extend data to an asymptotically flat null hypersurface \mathcal{W} . As usual, the key to then constructing the spacetime to the future of $\{v/-u = \delta^{-1}\}$ and \mathcal{W} is to establish suitable a priori estimates. (After constructing the spacetime to the future of \mathcal{W} we can rescale the u -coordinate so that \mathcal{W} corresponds to $\{u = -1\}$.)

This region is, in many ways, dual to the region near $\{v = 0\}$ that we discussed above. We expect the geometry near $\{u = 0\}$ to be modeled on a twisted self-similar solution (see [67]); in particular, we expect to $\Omega \hat{\chi}$ to be, at best, Hölder continuous at $\{u = 0\}$ and for $\underline{\alpha}$ to not lie in L^2 . In the region near $\{v = 0\}$ there were similar difficulties for $\hat{\chi}$ and α and we solved these via a renormalization scheme which explicitly subtracted off from $\hat{\chi}$ and α a self-similar term generated by the free data ϑ which captured the singular and large parts of $\hat{\chi}$ and α . However, in this region, because we are not posing initial data along $\{u = 0\}$, it is not possible to carry out such an explicit renormalization procedure. Instead, we simply eliminate $\underline{\alpha}$ from our Bianchi system using the renormalized Bianchi equations derived in [49, 50].

4.4. The interior

In this section we will discuss construction of the *interior* region of the naked singularity which was carried out in [62]. This is the region to the future of the past light cone of the singularity and

is labeled \mathcal{M}_{int} in Figure 6. Our plan will first be to construct the solution in a double-null gauge which holds in a region

$$(u, v, \theta^A) \in \left\{ u \in (-\infty, 0) \text{ and } \frac{v}{-u} \in (-1, 0) \right\}, \tag{4.7}$$

and where the self-similar vector field takes the form $u\partial_u + v\partial_v$. We do not expect this coordinate system to be regular at $\{v = 0\}$ (which will correspond to the past light cone of the singularity) or along the “axis” $\{u = v\}$. In order to carry out this construction we start by identifying a subset of the null-structure equations and use these to derive various equations for the metric components Ω , b , and g . (It is outside the scope of this survey article to describe these equations in specific detail.) Using self-similarity we reduce these to a set of equations along the null hypersurface $\{u = -1\}$, and then show the existence of solutions along this hypersurface with suitable boundary conditions at $\{v = -1\}$ and $\{v = 0\}$. After using self-similarity to expand our metric off of $\{u = -1\}$ to the entire region (4.7), the result will be a spacetime where a certain subset of the Einstein vacuum equations hold. We then run a “propagation of constraints” argument to show that the spacetime is actually Ricci flat. Finally, we then construct regular coordinate systems both at $\{v = 0\}$ and at the axis $\{v = u\}$ and show that the interior spacetime may then be glued to our previously constructed exterior spacetime.

Many of the equations we solve will be roughly of the following form⁷:

$$(-v)\partial_v^2\psi + b^A\nabla_A\partial_v\psi + \frac{(-v)A_1}{v+1}\partial_v\psi + A_2\partial_v\psi + \Omega^2(\Delta + A_3(v+1)^{-2})\psi = H, \quad \psi|_{v=0} = h. \tag{4.8}$$

Here $\psi(v, \theta^A) : (-1, 0) \times \mathbb{S}^2 \rightarrow \mathbb{R}$ is the unknown, the A_i are suitable constants, H and h can be thought to be given tensors or represent nonlinear terms, (Ω, b, g) will be assumed to satisfy various bootstrap estimates, and Δ is the Laplacian associated to g .

A key difficulty is that because the shift vector b does not vanish identically when $v = 0$, our equation is of a mixed elliptic–hyperbolic type. A relatively simple model equation for the elliptic–hyperbolic aspects of the problem (4.8) is

$$(-v)\partial_v^2\psi + \epsilon\partial_\phi\partial_v\psi + \partial_\phi^2\psi = 0, \quad \psi|_{v=-1} = 0, \quad \psi|_{v=0} = h,$$

where our unknown is $\psi(v, \phi) : (-1, 0) \times (-\pi, \pi) \rightarrow \mathbb{R}$, we have $0 < \epsilon \ll 1$, and $h : (-\pi, \pi) \rightarrow \mathbb{R}$ is given. We consider a suitable family of smoothing operators $\{\Pi_\delta\}_{\delta>0}$ on \mathbb{S}^2 which converge to the identity as $\delta \rightarrow 0$, and then, for each $0 < \delta \ll 1$, introduce the regularized equation

$$(-v)\partial_v^2\psi_\delta + \epsilon\Pi_\delta(\partial_\phi\partial_v\psi_\delta) + \partial_\phi^2\psi_\delta = 0, \quad \psi_\delta|_{v=-1} = 0, \quad \psi_\delta|_{v=-\delta} = h,$$

for an unknown $\psi_\delta : (-1, -\delta) \times (-\pi, \pi) \rightarrow \mathbb{R}$. This equation is elliptic⁸ using $-(\epsilon\partial_\phi\psi_\delta + \psi_\delta)$ as a multiplier, one may establish suitable a priori estimates which hold uniformly as $\delta \rightarrow 0$. Then it is relatively straightforward to establish the existence of a solution ψ_δ and to obtain the solution ψ as an appropriate limit $\psi_\delta \rightarrow_{\delta \rightarrow 0} \psi$. Though we will not discuss it in more detail, we note in passing that another important difficulty in the analysis of equations of the form (4.8) is that they are singular near $v = -1$, which corresponds to the breakdown of the double-null foliation at the axis.

In order to construct a good coordinate system near the axis $\{u = v\}$ we introduce self-similar wave coordinates $(\xi_t, \xi_x, \xi_y, \xi_z)$, in the region $\{(v/-u) \in (-1, -\delta)\}$ for suitable $\delta > 0$, which satisfy

$$\square_g \xi = 0, \quad K\xi = \xi, \tag{4.9}$$

⁷These are relevant both for the original existence argument and for the “propagation of constraints” argument. Also, though we suppress this feature for our survey article, many of the equations considered additionally involve various projections on suitable spaces of spherical harmonics.

⁸Note that the principle symbol of the regularized equation is $(-v)\partial_v^2 + \partial_\phi^2$ and that $-v$ is always positive in $v \in (-1, -\delta)$.

as well as suitable boundary conditions. If we work sufficiently close to the axis, then K will be timelike in the region we are defining $(\xi_t, \xi_x, \xi_y, \xi_z)$ and the system (4.9) reduces to an elliptic equation. Since, as is well-known, the Einstein vacuum equations become quasilinear wave equations in harmonic coordinates, we are eventually able to exploit various elliptic estimates to show that g is smooth in these harmonic coordinates.

Near $\{v = 0\}$ we cannot expect our self-similar wave coordinates to provide a good coordinate system. The solution we construct in the interior region will not directly correspond to the κ SS solutions discussed in Section 4.2.2; instead it will be of the more general TSS solutions discussed in [67]. The key good coordinate system turns out to be “lapse renormalized coordinates” (\hat{v}, u, θ^A) where we define \hat{v} by

$$\hat{v}(v, u, \theta^A) \doteq \int_0^v \Omega^2(\tilde{v}, u, \theta^A) d\tilde{v}.$$

One shows that these coordinates allows for the exterior metric and interior metric to be glued together along $\{\hat{v} = 0\}$ to produce a $C^{1,\gamma}$ metric (with additional regularity in the θ^A and u direction). The resulting spacetime is the desired naked singularity solution.

4.5. A few open questions

We close the section by quickly listing some open questions and further directions about naked singularities for the Einstein vacuum questions:

- (1) It will be interesting to study both the stability and instability problems of the solutions produced by Theorem 9. Do there exists similar phenomenon to those discovered in Singh's work [19]?
- (2) Can one construct naked singularities arising from smooth, or even C^2 initial data?
- (3) In higher dimensions, naked singularities are expected to be much more ubiquitous [68, 69]. It would be very interesting to establish existence of asymptotically flat higher dimensional naked singularities and to understand to what extent their presence is generic.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Dedication

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

The author acknowledges support from an Alfred P. Sloan Fellowship in Mathematics and from NSERC discovery grants RGPIN-2021-02562 and DGEGR-2021-00093.

Appendix A. Equations of the double-null gauge

In this appendix we collect the equations of the double-null gauge for ease of reference and to fix the notation. For a proper introduction to the double-null gauge we refer the reader to the works [10, 33]. We will say that (\mathcal{M}, g) is in the double-null gauge with the shift in the u -direction if

$$g = -2\Omega^2(du \otimes dv + dv \otimes du) + \mathfrak{g}_{AB}(d\theta^A - b^A du) \otimes (d\theta^B - b^B du), \quad (\text{A.1})$$

holds, and that (\mathcal{M}, g) is in the double-null gauge with the shift in the v -direction if (A.1) holds with the roles of u and v interchanged. In the following, unless indicated otherwise, we assume that the shift in the u -direction.

We define the null vector fields e_4 and e_3 by

$$e_4 \doteq \Omega^{-1} \partial_v, \quad e_3 \doteq \Omega^{-1} (\partial_u + b).$$

The $\mathbb{S}_{u,v}^2$ -projected derivatives in the e_4 and e_3 direction are denoted by ∇_4 and ∇_3 . The induced covariant derivative along each $\mathbb{S}_{u,v}^2$ is denoted by ∇ , $*$ refers to the Hodge star operator on $(\mathbb{S}^2, \mathfrak{g})$, slashed operators are defined with \mathfrak{g} , and we use K to denote the Gaussian curvature of \mathfrak{g} . We will use \mathcal{L} to denote the Lie derivative.

The Ricci coefficients $(\omega, \underline{\omega}, \zeta_A, \eta_A, \underline{\eta}_A, \chi_{AB}, \underline{\chi}_{AB})$ satisfy

$$\begin{aligned} \omega &= -\frac{1}{2} \nabla_4 \log \Omega, & \underline{\omega} &= -\frac{1}{2} \nabla_3 \log \Omega, & \mathcal{L}_{\partial_v} b^A &= -4\Omega^2 \zeta^A, \\ \eta_A &= \zeta_A + \nabla_A \log \Omega, & \underline{\eta}_A &= -\zeta_A + \nabla_A \log \Omega, \\ \mathcal{L}_{e_4} \mathfrak{g}_{AB} &= \chi_{AB}, & \mathcal{L}_{e_3} \mathfrak{g}_{AB} &= \underline{\chi}_{AB}. \end{aligned} \quad (\text{A.2})$$

We split χ_{AB} and $\underline{\chi}_{AB}$ into their trace and trace-free parts via

$$\chi_{AB} = \hat{\chi}_{AB} + \frac{1}{2} \text{tr} \chi \mathfrak{g}_{AB}, \quad \underline{\chi}_{AB} = \hat{\underline{\chi}}_{AB} + \frac{1}{2} \text{tr} \underline{\chi} \mathfrak{g}_{AB}.$$

The null curvature components $(\alpha_{AB}, \beta_A, \rho, \sigma, \underline{\beta}_A, \underline{\alpha}_{AB})$ are related to the Ricci coefficients via the various null-structure equations:

$$\nabla_4 \text{tr} \chi + \frac{1}{2} (\text{tr} \chi)^2 = -|\hat{\chi}|^2 - 2\omega \text{tr} \chi, \quad (\text{A.3})$$

$$\nabla_4 \hat{\chi} + \text{tr} \chi \hat{\chi} = -\alpha - 2\omega \hat{\chi},$$

$$\nabla_3 \text{tr} \underline{\chi} + \frac{1}{2} (\text{tr} \underline{\chi})^2 = -|\hat{\underline{\chi}}|^2 - 2\underline{\omega} \text{tr} \underline{\chi}, \quad (\text{A.4})$$

$$\nabla_3 \hat{\underline{\chi}} + \text{tr} \underline{\chi} \hat{\underline{\chi}} = -\underline{\alpha} - 2\underline{\omega} \hat{\underline{\chi}},$$

$$\nabla_3 \hat{\chi} + \frac{1}{2} \text{tr} \underline{\chi} \hat{\chi} = 2\underline{\omega} \hat{\chi} + \nabla \hat{\otimes} \eta + \eta \hat{\otimes} \eta - \frac{1}{2} \text{tr} \chi \hat{\chi}, \quad (\text{A.5})$$

$$\nabla_3 \text{tr} \chi + \frac{1}{2} \text{tr} \chi \text{tr} \underline{\chi} = 2\rho + 2\underline{\omega} \text{tr} \chi + 2 \text{div} \eta + 2|\eta|^2 - \hat{\chi} \cdot \hat{\underline{\chi}}, \quad (\text{A.6})$$

$$\nabla_4 \hat{\underline{\chi}} + \frac{1}{2} \text{tr} \chi \hat{\underline{\chi}} = 2\omega \hat{\underline{\chi}} + \nabla \hat{\otimes} \eta + \eta \hat{\otimes} \eta - \frac{1}{2} \text{tr} \chi \hat{\underline{\chi}},$$

$$\nabla_4 \text{tr} \underline{\chi} + \frac{1}{2} \text{tr} \chi \text{tr} \underline{\chi} = 2\rho + 2\omega \text{tr} \underline{\chi} + 2 \text{div} \underline{\eta} + 2|\underline{\eta}|^2 - \hat{\underline{\chi}} \cdot \hat{\chi},$$

$$\nabla_4 \eta = -\chi \cdot (\eta - \underline{\eta}) - \beta,$$

$$\nabla_3 \underline{\eta} = -\underline{\chi} \cdot (\eta - \underline{\eta}) + \underline{\beta},$$

$$\text{curl} \eta = \sigma + \frac{1}{2} \hat{\underline{\chi}} \wedge \hat{\chi},$$

$$\text{curl} \underline{\eta} = -\sigma - \frac{1}{2} \hat{\chi} \wedge \hat{\underline{\chi}},$$

$$\nabla_4 \underline{\omega} = \frac{1}{2} \rho + 2\underline{\omega} \omega + \frac{1}{2} |\eta|^2 - \eta \cdot \underline{\eta},$$

$$\nabla_3 \omega = \frac{1}{2} \rho + 2\underline{\omega} \omega + \frac{1}{2} |\underline{\eta}|^2 - \underline{\eta} \cdot \eta,$$

$$K = -\rho + \frac{1}{2} \hat{\chi} \cdot \hat{\underline{\chi}} - \frac{1}{4} \text{tr} \chi \text{tr} \underline{\chi},$$

$$\text{div} \hat{\chi} - \frac{1}{2} \nabla \text{tr} \chi = -\beta + \frac{1}{2} \text{tr} \chi \zeta - \zeta \cdot \hat{\chi},$$

$$\text{div} \hat{\underline{\chi}} - \frac{1}{2} \nabla \text{tr} \underline{\chi} = \underline{\beta} - \frac{1}{2} \text{tr} \underline{\chi} \zeta + \zeta \cdot \hat{\underline{\chi}},$$

Finally we have the Bianchi equations for the curvature components:

$$\begin{aligned}
\nabla_3 \underline{\alpha} + \frac{1}{2} \text{tr} \underline{\chi} \underline{\alpha} &= \nabla \hat{\otimes} \underline{\beta} + 4 \underline{\omega} \underline{\alpha} - 3(\hat{\chi} \underline{\rho} + {}^* \hat{\chi} \underline{\sigma}) + (\zeta + 4 \underline{\eta}) \hat{\otimes} \underline{\beta}, \\
\nabla_4 \underline{\beta} + 2 \text{tr} \underline{\chi} \underline{\beta} &= \text{div} \underline{\alpha} - 2 \underline{\omega} \underline{\beta} + (2 \zeta + \underline{\eta}) \cdot \underline{\alpha}, \\
\nabla_3 \underline{\beta} + \text{tr} \underline{\chi} \underline{\beta} &= \nabla \underline{\rho} + 2 \underline{\omega} \underline{\beta} + {}^* \nabla \underline{\sigma} + 2 \hat{\chi} \cdot \underline{\beta} + 3(\underline{\eta} \underline{\rho} + {}^* \underline{\eta} \underline{\sigma}), \\
\nabla_4 \underline{\sigma} + \frac{3}{2} \text{tr} \underline{\chi} \underline{\sigma} &= -\text{div} {}^* \underline{\beta} + \frac{1}{2} \hat{\chi} \wedge \underline{\alpha} - \zeta \wedge \underline{\beta} - 2 \underline{\eta} \wedge \underline{\beta}, \\
\nabla_4 \underline{\rho} + \frac{3}{2} \text{tr} \underline{\chi} \underline{\rho} &= \text{div} \underline{\beta} - \frac{1}{2} \hat{\chi} \cdot \underline{\alpha} + \zeta \cdot \underline{\beta} - 2 \underline{\eta} \cdot \underline{\beta} \\
\nabla_3 \underline{\sigma} + \frac{3}{2} \text{tr} \underline{\chi} \underline{\sigma} &= -\text{div} {}^* \underline{\beta} - \frac{1}{2} \hat{\chi} \wedge \underline{\alpha} + \zeta \wedge \underline{\beta} - 2 \underline{\eta} \wedge \underline{\beta}, \\
\nabla_3 \underline{\rho} + \frac{3}{2} \text{tr} \underline{\chi} \underline{\rho} &= -\text{div} \underline{\beta} - \frac{1}{2} \hat{\chi} \cdot \underline{\alpha} + \zeta \cdot \underline{\beta} + 2 \underline{\eta} \cdot \underline{\beta}, \\
\nabla_4 \underline{\beta} + \text{tr} \underline{\chi} \underline{\beta} &= -\nabla \underline{\rho} + 2 \underline{\omega} \underline{\beta} + {}^* \nabla \underline{\sigma} + 2 \hat{\chi} \cdot \underline{\beta} - 3(\underline{\eta} \underline{\rho} - {}^* \underline{\eta} \underline{\sigma}), \\
\nabla_3 \underline{\beta} + 2 \text{tr} \underline{\chi} \underline{\beta} &= -\text{div} \underline{\alpha} - 2 \underline{\omega} \underline{\beta} - (-2 \zeta + \underline{\eta}) \cdot \underline{\alpha}, \\
\nabla_4 \underline{\alpha} + \frac{1}{2} \text{tr} \underline{\chi} \underline{\alpha} &= -\nabla \hat{\otimes} \underline{\beta} + 4 \underline{\omega} \underline{\alpha} - 3(\hat{\chi} \underline{\rho} - {}^* \hat{\chi} \underline{\sigma}) - (-\zeta + 4 \underline{\eta}) \hat{\otimes} \underline{\beta}.
\end{aligned}$$

Appendix B. The renormalized Bianchi equations from [49, 50]

We list here the renormalized Bianchi Equations from [49, 50]. We first define

$$\check{\rho} \doteq \underline{\rho} - \frac{1}{2} \hat{\chi} \cdot \hat{\chi}, \quad \check{\sigma} \doteq \underline{\sigma} - \frac{1}{2} \hat{\chi} \wedge \hat{\chi}.$$

Then we have

$$\begin{aligned}
\nabla_3 \underline{\beta} + \text{tr} \underline{\chi} \underline{\beta} &= \nabla \check{\rho} + {}^* \nabla \check{\sigma} + 2 \underline{\omega} \underline{\beta} + 2 \hat{\chi} \cdot \underline{\beta} + 3(\underline{\eta} \check{\rho} + {}^* \underline{\eta} \check{\sigma}) \\
&\quad + \frac{1}{2} (\nabla(\hat{\chi} \cdot \hat{\chi}) + {}^* \nabla(\hat{\chi} \wedge \hat{\chi})) + \frac{3}{2} (\underline{\eta} \hat{\chi} \cdot \hat{\chi} + {}^* \underline{\eta} \hat{\chi} \wedge \hat{\chi}), \\
\nabla_4 \check{\sigma} + \frac{3}{2} \text{tr} \underline{\chi} \check{\sigma} &= -\text{div} {}^* \underline{\beta} - \zeta \wedge \underline{\beta} - 2 \underline{\eta} \wedge \underline{\beta} + \frac{1}{2} \hat{\chi} \wedge (\nabla \hat{\otimes} \underline{\eta}) - \frac{1}{2} \hat{\chi} \wedge (\underline{\eta} \hat{\otimes} \underline{\eta}), \\
\nabla_4 \check{\rho} + \frac{3}{2} \text{tr} \underline{\chi} \check{\rho} &= \text{div} \underline{\beta} + \zeta \cdot \underline{\beta} + 2 \underline{\eta} \cdot \underline{\beta} - \frac{1}{2} \hat{\chi} \cdot (\nabla \hat{\otimes} \underline{\eta}) - \frac{1}{2} \hat{\chi} \cdot (\underline{\eta} \hat{\otimes} \underline{\eta}) + \frac{1}{4} \text{tr} \underline{\chi} |\hat{\chi}|^2, \\
\nabla_3 \check{\sigma} + \frac{3}{2} \text{tr} \underline{\chi} \check{\sigma} &= \text{div} {}^* \underline{\beta} + \zeta \wedge \underline{\beta} - 2 \underline{\eta} \wedge \underline{\beta} + \frac{1}{2} \hat{\chi} \wedge (\nabla \hat{\otimes} \underline{\eta}) + \frac{1}{2} \hat{\chi} (\underline{\eta} \hat{\otimes} \underline{\eta}), \\
\nabla_3 \check{\rho} + \frac{3}{2} \text{tr} \underline{\chi} \check{\rho} &= -\text{div} \underline{\beta} + \zeta \cdot \underline{\beta} - 2 \underline{\eta} \cdot \underline{\beta} - \frac{1}{2} \hat{\chi} \cdot (\nabla \hat{\otimes} \underline{\eta}) - \frac{1}{2} \hat{\chi} \cdot (\underline{\eta} \hat{\otimes} \underline{\eta}) + \frac{1}{2} \text{tr} \underline{\chi} |\hat{\chi}|^2, \\
\nabla_4 \underline{\beta} + \text{tr} \underline{\chi} \underline{\beta} &= -\nabla \check{\rho} + {}^* \nabla \check{\sigma} + 2 \underline{\omega} \underline{\beta} + 2 \hat{\chi} \cdot \underline{\beta} - 3(\underline{\eta} \check{\rho} - {}^* \underline{\eta} \check{\sigma}) \\
&\quad - \frac{1}{2} (\nabla(\hat{\chi} \cdot \hat{\chi}) + {}^* \nabla(\hat{\chi} \wedge \hat{\chi})) - \frac{3}{2} (\underline{\eta} \hat{\chi} \wedge \hat{\chi} + {}^* \underline{\eta} \hat{\chi} \wedge \hat{\chi}).
\end{aligned}$$

Appendix C. Characteristic initial value problem

In this section we will quickly review the well-posedness of the characteristic initial value problem for the Einstein equations. Choose real numbers $u_0, u_1, v_0,$ and v_1 with $u_0 < u_1$ and $v_0 < v_1$. We say a triple of $(\Omega^{(\text{in})}, (b^A)^{(\text{in})}, \mathfrak{g}_{AB}^{(\text{in})})$ defined along $[u_0, u_1] \times \mathbb{S}^2$, a pair $(\Omega^{(\text{out})}, \mathfrak{g}_{AB}^{(\text{out})})$ defined along $[v_0, v_1] \times \mathbb{S}^2$, and a 1-form ζ_A on \mathbb{S}^2 form a characteristic initial data set if

- (1) We have $(\Omega^{(\text{in})}, \mathfrak{g}^{(\text{in})})|_{u=u_0} = (\Omega^{(\text{out})}, \mathfrak{g}^{(\text{out})})|_{v=v_0}$.
- (2) We consider $(\Omega^{(\text{in})}, (b^A)^{(\text{in})}, \mathfrak{g}_{AB}^{(\text{in})})$ as being defined along a hypersurface $\{v = v_0\}$ and $(\Omega^{(\text{out})}, \mathfrak{g}_{AB}^{(\text{out})})$ as being defined along a hypersurface $\{u = u_0\}$. After defining $\underline{\omega}, \omega, \underline{\chi}$, and $\underline{\chi}$ using (A.2), we require that the Raychaudhuri Equations (A.3) and (A.6) hold along $\{u = u_0\}$ and $\{v = v_0\}$ respectively.

We can now state a local existence theorem.

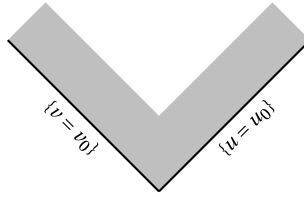


Figure 11. Penrose diagram for spacetime arising from characteristic initial value problem.

Theorem 12. [70, 71] Let $(\Omega^{(\text{in})}, (b^A)^{(\text{in})}, \mathfrak{g}_{AB}^{(\text{in})})$, $(\Omega^{(\text{out})}, \mathfrak{g}_{AB}^{(\text{out})})$, and $\tilde{\zeta}_A$ form an initial data set of sufficiently high regularity. Then there exists a constant $c > 0$ and a unique solution $(\mathcal{M}, \mathfrak{g})$ to the Einstein vacuum equations expressed in a double-null gauge and defined in a region $(u, v, \theta^A) \in [u_0, u_1] \times [v_0, v_0 + c] \times \mathbb{S}^2 \cup [u_0, u_0 + c] \times [v_0, v_1] \times \mathbb{S}^2$ so that

$$\begin{aligned} (\Omega^{(\text{in})}, (b^A)^{(\text{in})}, \mathfrak{g}_{AB}^{(\text{in})}) &= (\Omega, b^A, \mathfrak{g}_{AB})|_{v=v_0}, & (\Omega^{(\text{out})}, \mathfrak{g}_{AB}^{(\text{out})}) &= (\Omega, \mathfrak{g}_{AB})|_{u=u_0} \\ \tilde{\zeta}_A &= \zeta_A|_{(u,v)=(u_0,v_0)}. \end{aligned}$$

In Figure 11 we have drawn a Penrose diagram depicting the range of the solution.

References

- [1] Y. Fourès-Bruhat, “Théorème d’existence pour certains systèmes d’équations aux dérivées partielles non linéaires”, *Acta Math.* **88** (1952), pp. 141–225.
- [2] Y. Choquet-Bruhat and R. P. Geroch, “Global aspects of the Cauchy problem in general relativity”, *Commun. Math. Phys.* **14** (1969), pp. 329–335.
- [3] R. M. Wald, *General Relativity*, The University of Chicago Press: Chicago, 1984.
- [4] K. Schwarzschild, “Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie”, *Sitzungsber. K. Preuss. Akad. Wiss.* **1** (1916), pp. 189–196.
- [5] M. Dafermos and I. Rodnianski, “Lectures on black holes and linear waves”, in *Evolution Equations*, Clay Mathematics Proceedings, American Mathematical Society: Providence, RI, 2013, pp. 97–205.
- [6] J. Sbierski, “The C_0 -inextendibility of the Schwarzschild spacetime and the spacelike diameter in Lorentzian geometry”, *J. Differ. Geom.* **108** (2018), no. 2, pp. 319–378.
- [7] J. Sbierski, “On the proof of the C^0 -inextendibility of the Schwarzschild spacetime”, *J. Phys. Conf. Ser.* **968** (2018), article no. 012012.
- [8] R. Penrose, “Gravitational collapse and space-time singularities”, *Phys. Rev. Lett.* **14** (1965), pp. 57–59.
- [9] D. Christodoulou and S. Klainerman, *The Global Nonlinear Stability of the Minkowski Space*, Princeton Mathematical Series, Princeton University Press: Princeton, NJ, 1993, pp. x+514.
- [10] S. Klainerman and F. Nicolò, *The Evolution Problem in General Relativity*, Progress in Mathematical Physics, Birkhäuser Boston Inc.: Boston, MA, 2003, pp. xiv+385.
- [11] D. Christodoulou, “On the global initial value problem and the issue of singularities”, *Class. Quantum Gravity* **16** (1999), A23–A35.
- [12] D. Shen, *Global stability of Minkowski spacetime with minimal decay*, preprint, 2023, 2310.07483.
- [13] R. Penrose, “Gravitational collapse: the role of general relativity”, *Riv. Nuovo Cim.* **1** (1969), pp. 252–276.
- [14] D. Christodoulou, “Examples of naked singularity formation in the gravitational collapse of a scalar field”, *Ann. Math. (2)* **140** (1994), no. 3, pp. 607–653.
- [15] M. Dafermos and J. Luk, “The interior of dynamical vacuum black holes I: the C^0 -stability of the Kerr Cauchy horizon”, *Ann. Math.* (2017). (to appear).
- [16] D. Christodoulou, “The formation of black holes and singularities in spherically symmetric gravitational collapse”, *Commun. Pure Appl. Math.* **44** (1991), no. 3, pp. 339–373.
- [17] D. Christodoulou, “Bounded variation solutions of the spherically symmetric Einstein-scalar field equations”, *Commun. Pure Appl. Math.* **46** (1993), no. 8, pp. 1131–1220.
- [18] D. Christodoulou, “The instability of naked singularities in the gravitational collapse of a scalar field”, *Ann. Math.* **149** (1999), pp. 183–217.
- [19] J. Singh, *High regularity waves on self-similar naked singularity interiors: decay and the role of blue-shift*, preprint, 2024, 2402.00062.

- [20] M. Dafermos, “Spherically symmetric spacetimes with a trapped surface”, *Class. Quantum Gravity* **22** (2005), no. 11, pp. 2221–2232.
- [21] J. Singh, “A construction of approximately self-similar naked singularities for the spherically symmetric Einstein-scalar field system”, *Ann. Henri Poincaré* (2024).
- [22] D. Christodoulou, “Violation of cosmic censorship in the gravitational collapse of a dust cloud”, *Commun. Math. Phys.* **93** (1984), no. 2, pp. 171–195.
- [23] Y. Guo, M. Hadzic and J. Jang, “Naked singularities in the Einstein–Euler system”, *Ann. PDE* **9** (2023), no. 1, article no. 4.
- [24] P. Bizoń and A. Wasserman, “On the existence of self-similar spherically symmetric wave maps coupled to gravity”, *Class. Quantum Gravity* **19** (2002), no. 12, pp. 3309–3321.
- [25] M. Choptuik, “Universality and scaling in gravitational collapse of a massive scalar field”, *Phys. Rev. Lett.* **70** (1993), pp. 9–12.
- [26] C. Gundlach and J. Martin-Garcia, “Critical phenomena in gravitational collapse”, *Living Rev. Relativ.* **10** (2007), article no. 5.
- [27] J. Liu and J. Li, “A robust proof of the instability of naked singularities of a scalar field in spherical symmetry”, *Commun. Math. Phys.* **363** (2018), no. 2, pp. 561–578.
- [28] J. Li and J. Liu, “Instability of spherical naked singularities of a scalar field under gravitational perturbations”, *J. Differ. Geom.* **120** (2022), no. 1, pp. 97–197.
- [29] X. An, “Naked singularity censoring with anisotropic apparent horizon”, *Ann. Math.* (2024). (to appear).
- [30] V. Cardoso, J. A. L. Costa, K. Destounis, P. Hintz and A. Jansen, “Quasinormal modes and strong cosmic censorship”, *Phys. Rev. Lett.* **120** (2018), article no. 031103.
- [31] M. Dafermos and Y. Shlapentokh-Rothman, “Rough initial data and the strength of the blue-shift instability on cosmological black holes with Λ ensuremath> 0”, *Class. Quantum Gravity* **35** (2018), no. 19, article no. 195010.
- [32] O. J. C. Dias, H. S. Reall and J. E. Santos, “Strong cosmic censorship: taking the rough with the smooth”, *J. High Energy Phys.* (2018), no. 10, article no. 1.
- [33] D. Christodoulou, *The Formation of Black Holes in General Relativity*, EMS Monographs in Mathematics, European Mathematical Society (EMS): Zürich, 2009, pp. x+589.
- [34] X. Chen and S. Klainerman, *Formation of trapped surfaces in geodesic foliation*, preprint, 2024, 2409.14582.
- [35] M. Dafermos, “The formation of black holes in general relativity [after D. Christodoulou]”, *Sémin. Bourbaki* **2011/2012** (2013), no. 352, pp. 243–313. Exposé 1043–1058, p. Exp. No. 1051, viii.
- [36] J. Li and P. Yu, “Construction of Cauchy data of vacuum Einstein field equations evolving to black holes”, *Ann. Math. (2)* **181** (2015), no. 2, pp. 699–768.
- [37] J. Corvino and R. M. Schoen, “On the asymptotics for the vacuum Einstein constraint equations”, *J. Differ. Geom.* **73** (2006), no. 2, pp. 185–217.
- [38] C. Kehle and R. Unger, “Event horizon gluing and black hole formation in vacuum: the very slowly rotating case”, *Adv. Math.* **452** (2024), article no. 109816.
- [39] S. Klainerman and I. Rodnianski, “On the formation of trapped surfaces”, *Acta Math.* **208** (2012), no. 2, pp. 211–333.
- [40] S. Klainerman and I. Rodnianski, “On emerging scarred surfaces for the Einstein vacuum equations”, *Discrete Contin. Dyn. Syst.* **28** (2010), no. 3, pp. 1007–1031.
- [41] X. An, *Formation of trapped surfaces in general relativity*, PhD thesis, Princeton University: ProQuest LLC, Ann Arbor, MI, 2014.
- [42] S. Klainerman, J. Luk and I. Rodnianski, “A fully anisotropic mechanism for formation of trapped surfaces in vacuum”, *Invent. Math.* **198** (2014), no. 1, pp. 1–26.
- [43] S. Klainerman, I. Rodnianski and J. Szeftel, “The bounded L^2 curvature conjecture”, *Invent. Math.* **202** (2015), no. 1, pp. 91–216.
- [44] J. Szeftel, “Sharp Strichartz estimates for the wave equation on a rough background”, *Ann. Sci. Éc. Norm. Supér. (4)* **49** (2016), no. 6, pp. 1279–1309.
- [45] J. Szeftel, “Parametrix for wave equations on a rough background III: space–time regularity of the phase”, *Astérisque* (2018), no. 401, pp. viii+321.
- [46] J. Szeftel, “Parametrix for wave equations on a rough background. I: regularity of the phase at initial time. II: construction and control at initial time”, *Astérisque* (2023), no. 443, pp. ix+275.
- [47] J. Szeftel, “Parametrix for wave equations on a rough background: IV. Control of the error term”, *Astérisque* (2023), no. 444, pp. viii+314.
- [48] J. Luk, “Singularities in general relativity”, in *ICM—International Congress of Mathematicians. Vol. 5. Sections 9–11*, EMS Press: Berlin, 2023, pp. 4120–4141. ©2023.
- [49] J. Luk and I. Rodnianski, “Local propagation of impulsive gravitational waves”, *Commun. Pure Appl. Math.* **68** (2015), no. 4, pp. 511–624.

- [50] J. Luk and I. Rodnianski, “Nonlinear interaction of impulsive gravitational waves for the vacuum Einstein equations”, *Camb. J. Math.* **5** (2017), no. 4, pp. 435–570.
- [51] R. Penrose, “The geometry of impulsive gravitational waves”, in *General Relativity (Papers in Honour of J. L. Synge)*, Clarendon Press: Oxford, 1972, pp. 101–115.
- [52] K. A. Khan and R. Penrose, “Scattering of two impulsive gravitational plane waves”, *Nature* **229** (1971), no. 5281, pp. 185–186.
- [53] J. Luk and I. Rodnianski, *High-frequency limits and null dust solutions in general relativity*, preprint, 2020, 2009.08968.
- [54] G. A. Burnett, “The high-frequency limit in general relativity”, *J. Math. Phys.* **30** (1989), no. 1, pp. 90–96.
- [55] C. Huneau and J. Luk, “High-frequency solutions to the Einstein equations”, *Class. Quantum Gravity* **41** (2024), article no. 143002.
- [56] X. An and J. Luk, “Trapped surfaces in vacuum arising dynamically from mild incoming radiation”, *Adv. Theor. Math. Phys.* **21** (2017), no. 1, pp. 1–120.
- [57] X. An, “Emergence of apparent horizon in gravitational collapse”, *Ann. PDE* **6** (2020), no. 2, article no. 10.
- [58] X. An, “A scale-critical trapped surface formation criterion: a new proof via signature for decay rates”, *Ann. PDE* **8** (2022), no. 1, article no. 3.
- [59] X. An and T. He, “Dynamics of apparent horizon and a null comparison principle”, *Ann. PDE* **10** (2024), article no. 15.
- [60] X. An and Q. Han, *Anisotropic dynamical horizons arising in gravitational collapse*, preprint, 2020, 2010.12524.
- [61] I. Rodnianski and Y. Shlapentokh-Rothman, “Naked singularities for the Einstein vacuum equations: the exterior solution”, *Ann. Math. (2)* **198** (2023), no. 1, pp. 231–391.
- [62] Y. Shlapentokh-Rothman, *Naked singularities for the Einstein vacuum equations: the interior solution*, preprint, 2022, 2204.09891.
- [63] F. J. Tipler, “Singularities in conformally flat spacetimes”, *Phys. Lett. A* **64** (1977), no. 1, pp. 8–10.
- [64] C. Fefferman and C. R. Graham, “Conformal invariants”, *Astérisque* (1985), pp. 95–116. Numéro Hors Série The Mathematical Heritage of Élie Cartan (Lyon, 1984).
- [65] C. Fefferman and C. R. Graham, *The Ambient Metric*, Annals of Mathematics Studies, Princeton University Press: Princeton, NJ, 2012, pp. x+113.
- [66] I. Rodnianski and Y. Shlapentokh-Rothman, “The asymptotically self-similar regime for the Einstein vacuum equations”, *Geom. Funct. Anal.* **28** (2018), no. 3, pp. 755–878.
- [67] Y. Shlapentokh-Rothman, “Twisted self-similarity and the Einstein vacuum equations”, *Commun. Math. Phys.* **401** (2023), no. 2, pp. 2269–2325.
- [68] L. Lehner and F. Pretorius, “Black strings, low viscosity fluids, and violation of cosmic censorship”, *Phys. Rev. Lett.* **105** (2010), no. 10, article no. 101102.
- [69] X. An and X. Zhang, “Examples of naked singularity formation in higher-dimensional Einstein-vacuum spacetimes”, *Ann. Henri Poinc.* **19** (2018), no. 2, pp. 619–651.
- [70] A. D. Rendall, “Reduction of the characteristic initial value problem to the Cauchy problem and its applications to the Einstein equations”, *Proc. R. Soc. Lond. A* **427** (1990), no. 1872, pp. 221–239.
- [71] J. Luk, “On the local existence for the characteristic initial value problem in general relativity”, *Int. Math. Res. Not. IMRN* (2012), no. 20, pp. 4625–4678.