

# BLACK HOLE FORMATION AND STABILITY: A MATHEMATICAL INVESTIGATION

LYDIA BIERI

**ABSTRACT.** The dynamics of the Einstein equations feature the formation of black holes. These are related to the presence of trapped surfaces in the spacetime manifold. The mathematical study of these phenomena has gained momentum since D. Christodoulou’s breakthrough result proving that, in the regime of pure general relativity, trapped surfaces form through the focusing of gravitational waves. (The latter were observed for the first time in 2015 by Advanced LIGO.) The proof combines new ideas from geometric analysis and nonlinear partial differential equations, and it introduces new methods to solve large data problems. These methods have many applications beyond general relativity. D. Christodoulou’s result was generalized by S. Klainerman and I. Rodnianski, and more recently by these authors and J. Luk. Here, we investigate the dynamics of the Einstein equations, focusing on these works. Finally, we address the question of stability of black holes and what has been known so far, involving recent works of many contributors.

## CONTENTS

1. Introduction	1
2. Mathematical general relativity	3
3. Gravitational radiation	12
4. The formation of closed trapped surfaces	14
5. Stability of black holes	25
6. Cosmic censorship	26
Acknowledgments	27
About the author	27
References	27

## 1. INTRODUCTION

The Einstein equations exhibit singularities that are hidden behind event horizons of black holes. A black hole is a region of spacetime that cannot be observed from infinity. The first encounters with its intriguing properties go back to

---

Received by the editors April 27, 2017.

2010 *Mathematics Subject Classification.* Primary 83C05, 83C57, 35A20, 35A01, 35A02, 53C10.

The author acknowledges her NSF support and is supported by NSF CAREER Grant No. DMS-1253149.

the famous Schwarzschild solution of the Einstein vacuum (EV) equations in four spacetime dimensions

$$(1) \quad \text{Ric}(g) = 0.$$

$\text{Ric}(g)$  denotes the Ricci curvature of the Lorentzian metric  $g$  of the four-dimensional spacetime. In 1916 K. Schwarzschild [64] found the first solution to these equations, after A. Einstein in 1915 had formulated the general theory of relativity and had derived the Einstein equations [31], [32]. The Schwarzschild solution is spherically symmetric and depends on the mass  $M$  of the body that it describes. Very importantly, Birkhoff [13] in 1923 proved that the Schwarzschild solution is the only spherically symmetric solution of the EV equations. It describes the gravitational field outside a nonrotating star or black hole, generally outside of any spherically symmetric body. The evolution of the body itself does not change the gravitational field in the exterior. We note at this point that a spherically symmetric object does not generate any gravitational waves. In the coordinate system, in which the Schwarzschild solution was first discovered, it has a singularity at  $r = 2M$ , where  $r$  denotes the radius of the spheres being the orbits of the rotation group. In 1924 A. Eddington [30] used a coordinate transformation getting rid of this singularity but did not comment on it. It was G. Lemaître [51] who in 1933 observed that this is only a coordinate singularity and that the Schwarzschild solution behaves “nicely” there in other coordinates. There is a true singularity at  $r = 0$ . By works of D. Finkelstein [34], M. D. Kruskal [48], J.L. Synge [65], G. Szekeres [70], systems of coordinates for the complete analytic extension of the Schwarzschild solution had been found, and dynamical properties of the region  $r < 2M$  had been addressed. Finkelstein [34] in 1958 mentioned that the hypersurface  $r = 2M$  is an event horizon, that is the boundary of the region, which is causally connected to infinity. More interesting is the behavior of other black hole spacetimes, namely the Kerr solutions discovered by R. Kerr [41] in 1963. This is a two-parameter family of axisymmetric solutions of the EV equations (1) having an event horizon, and the spacetime outside this horizon is a regular asymptotically flat region. Besides the mass parameter  $M$  (positive) this family is characterized also by the angular momentum  $a$  about the axis of symmetry with  $|a| \leq M^2$ .

Are black holes rare phenomena, or do we expect them to occur often in the universe? How do they form, and can they form in the evolution of initial data that do not contain any black holes? The latter was investigated by Demetrios Christodoulou in 2008 in his celebrated monograph [25]. Christodoulou’s answer is yes, and he provides a detailed description of black hole formation in his main proof. In order to state the main result of [25], we now turn to the notion of a *closed trapped surface*.

A concept directly related to the formation of black holes is a closed trapped surface introduced by R. Penrose [59] in 1965. He defines a trapped surface to be a spacelike surface such that the expansion scalars with respect to every family of future-directed null geodesic normals are negative, i.e., infinitesimally virtual displacements along these normals imply pointwise decrease of the area element. Penrose proved the following.

**Theorem 1** (R. Penrose [59]). *A spacetime  $(\mathcal{M}, g)$  is future null geodesically incomplete if the following three conditions hold:*

1.  $\text{Ric}(V, V) \geq 0$  for all null vectors  $V$ .

2. *There exists a noncompact Cauchy hypersurface  $H$  in  $\mathcal{M}$ .*
3. *There is a closed trapped surface  $S$  in  $\mathcal{M}$ .*

A modern version of this incompleteness theorem can be formulated as follows:

**Theorem 2.** *Consider regular characteristic initial data on a complete null geodesic cone  $\mathcal{C}$ . Denote by  $(\mathcal{M}, g)$  the maximal future development of the data on  $\mathcal{C}$ . Assume that  $\mathcal{M}$  contains a closed trapped surface  $S$ . Then  $(\mathcal{M}, g)$  is future null geodesically incomplete.*

At this point it was not clear at all if closed trapped surfaces form in the evolution of data that does not contain any such surface. In particular, one can ask what happens in the situation where the initial conditions are very far from containing a closed trapped surface. We would like to study the long-time evolution for the Einstein equations and show that closed trapped surfaces form under physical conditions. Christodoulou did this through analyzing the dynamics of gravitational collapse.

The first and simplest version of Christodoulou's main result on the formation of closed trapped surfaces for the Einstein vacuum equations (1) can be stated as follows:

**Theorem 3** (D. Christodoulou [25]). *Closed trapped surfaces form in the Cauchy development of initial data, which are arbitrarily dispersed, if the incoming energy per unit solid angle in each direction in a suitably small time interval is large enough.*

We can rephrase this result, saying that if enough energy through gravitational waves has been concentrated in a small enough time interval, then a closed trapped surface will form.

Christodoulou's result was generalized by Sergiu Klainerman and Igor Rodnianski [44] as well as by these authors and Jonathan Luk [42] to allow for more general initial data.

Next, we can ask if black holes are stable. This question is topic of ongoing research in the field. Many contributors have studied the first step towards understanding this problem.

In the present article, we will first give an introduction to the main ideas of mathematical general relativity, then investigate the main steps and methods of the proof of black hole formation by Christodoulou and the generalization by Klainerman, Luk, and Rodnianski, then we shall address the stability problem of black holes. Along the way, we will highlight gravitational waves that were observed for the first time in 2015 by Advanced LIGO [1].

## 2. MATHEMATICAL GENERAL RELATIVITY

### 2.1. Einstein equations and spacetime manifold.

**The equations.** Albert Einstein in 1915 derived the famous field equations of gravitation and established the general relativity (GR) theory [31], [32]. Much different from Newtonian physics where space and time are separate and independent concepts, already special relativity (1905) combines space and time into a (flat) spacetime manifold known as the Minkowski spacetime. General relativity gives that manifold a curved metric whose curvature encodes the properties of the

gravitational field. Thus gravitation acts through curvature. Our spacetime manifold carries all the information. “It is all there is, and nothing lives independently from it.” The Newtonian law under which each mass makes a gravitational field is replaced by the Einstein field equations

$$(2) \quad R_{ij} - \frac{1}{2}Rg_{ij} = \frac{8\pi G}{c^4}T_{ij}.$$

Here  $c$  denotes the speed of light,  $G$  is the Newtonian gravitational constant, the indices  $i, j$  take on values  $0, 1, 2, 3$ , and the tensors are as follows:  $R_{ij}$  is the Ricci curvature tensor,  $R$  is the scalar curvature tensor,  $g_{ij}$  is the metric tensor, and  $T_{ij}$  denotes the energy-momentum tensor. The latter contains matter or energy present in the spacetime such as a fluid or electromagnetic fields. If there are matter (or energy) fields, thus  $T_{ij} \neq 0$ , then they obey their own evolution equations and together with the Einstein equations (2) they form a coupled system. One then solves the Einstein system for the metric tensor  $g_{ij}$ . If there are no other fields, then  $T_{ij} = 0$  and equations (2) reduce to the Einstein vacuum (EV) equations (1).

**The spacetime.** A crucial difference from studying a partial differential equation on a (Euclidean or curved) background is that we are constructing the manifold itself by solving the Einstein equations. The resulting spacetime may feature intriguing properties, including black hole formation or gravitational waves. The main goal of mathematical GR is to investigate classes of these manifolds, their structures and dynamics as well as their stability. This can only be achieved by solving the Cauchy problem for physical settings via geometric analysis and often combining various areas of mathematics.

**Definition 1.** A spacetime manifold is defined to be a four-dimensional, oriented, differentiable manifold  $\mathcal{M}$  with a Lorentzian metric  $g$ .

*Remark.* An  $n$ -dimensional spacetime is defined in the corresponding way.

**Definition 2.** A Lorentzian metric  $g$  is defined to be a continuous assignment of a nondegenerate quadratic form  $g_p$ , being of index 1, in  $T_p\mathcal{M}$  for every  $p$  in  $\mathcal{M}$ .

The simplest example of a Lorentzian metric is the Minkowski metric  $\eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$ . In our Lorentzian manifold  $(\mathcal{M}, g)$  the tangent space  $T_p\mathcal{M}$  at each  $p$  is flat Minkowskian. We observe that a Lorentzian metric is a special case of a pseudo-Riemannian metric.

In this article we are interested in purely gravitational questions; therefore, we are looking for spacetimes  $(\mathcal{M}, g)$  with  $g$  solving the EV equations (1). At this point, we note that these equations can be written as a system of nonlinear, hyperbolic partial differential equations (pde).

The Lorentzian structure of our metric allows us to distinguish the following three types of vectors at a point  $p$  in  $T_p\mathcal{M}$ : A vector  $X \in T_p\mathcal{M}$  is called timelike if  $g_p(X, X) < 0$ , it is called null or lightlike if  $g_p(X, X) = 0$ , and it is called spacelike if  $g_p(X, X) > 0$ . A hypersurface is called spacelike if its normal vector is timelike. As in GR nothing travels faster than the speed of light; trajectories of massless particles are null, whereas those for massive objects are timelike. A curve that is timelike or null is called *causal*.

We know that in special relativity light travels along light cones (Figure 1).

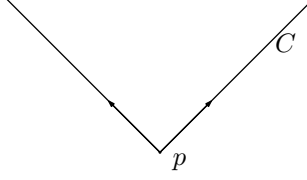


FIGURE 1

In general relativity light travels along null hypersurfaces which are generated by the congruence of null geodesics (Figure 2).

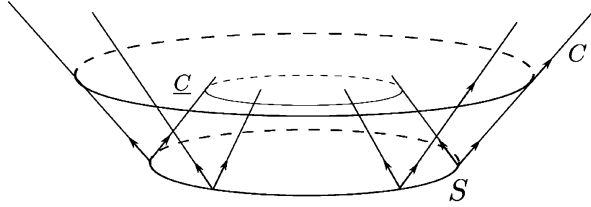


FIGURE 2

Let us start with a spacelike 2-surface  $S$  in  $\mathcal{M}$ . At each  $p \in T_p\mathcal{M}$ , we identify two orthogonal future-directed null vectors, namely  $\mathcal{L}_p$  which is outward pointing, and  $\underline{\mathcal{L}}_p$  which is inward pointing. The corresponding vectorfields defined in this way on  $S$  are  $\mathcal{L}$ , respectively  $\underline{\mathcal{L}}$ . The null hypersurfaces generated by the corresponding sets of null geodesics orthogonal to  $S$  are denoted by  $C$  and  $\underline{C}$ . Viewing  $S$  as a hypersurface in  $C$ , we denote its second fundamental form by  $\chi$ , and the second fundamental form of  $S$  in  $\underline{C}$  by  $\underline{\chi}$ . Their traceless parts are called the shears and are denoted by  $\hat{\chi}$ ,  $\hat{\underline{\chi}}$ , respectively. The traces  $\text{tr}\chi$  and  $\text{tr}\underline{\chi}$  are the expansion scalars.

From above and with this notation, we know that  $S$  is trapped if

$$\text{tr}\chi < 0 \quad \text{and} \quad \text{tr}\underline{\chi} < 0.$$

Thus the null hypersurfaces look as in Figure 3.

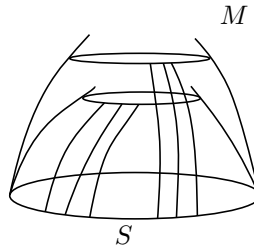


FIGURE 3

The mathematical structures behind this picture will be explained in Section 4.2. For the moment, let us note that in order to prove Theorem 3, Christodoulou

in [25] uses a “natural” foliation of the spacetime into null hypersurfaces  $C_u$  and  $\underline{C}_u$ , the double-null foliation.

We can also think of  $S$  as the intersection of  $C_u$  with a spacelike hypersurface  $H_t$ , where the  $C_u$  are the null hypersurfaces of a foliation of the spacetime  $(\mathcal{M}, g)$  with respect to an optical function  $u$ . Then we write  $S_{t,u}$  for the 2-surface. Note that the  $S_{t,u}$  are diffeomorphic to  $S^2$ . The main foliations of the spacetime in [27] are given first by the time function  $t$ , of which the level sets are maximal spacelike hypersurfaces  $H_t$  with vanishing linear momentum, and second by the optical function  $u$  (known as retarded time), for which the level sets are the  $C_u$ . The foliations are such that the density of the  $S_{t,u}$  in the  $\mathcal{H}_t$  tends to 1 as  $t \rightarrow \infty$ .

These foliations were crucial in the works by Christodoulou and Klainerman [27] on the global nonlinear stability of Minkowski space. Christodoulou in [25] combines the double-null structure and methods from [27] with new features to investigate black hole formation. Beyond that, these structures prove natural to describe radiation. In a radiative spacetime, gravitational waves (fluctuations of the curvature) travel along these null hypersurfaces  $C_u$  from their sources, such as mergers of binary black holes, of neutron stars, or as core-collapse supernovae. When we observe these waves, we can think of ourselves as located at *future null infinity*  $\mathcal{I}^+$ , which is defined to be the endpoints of all future-directed null geodesics along which for the surfaces  $S$  the area radius  $r \rightarrow \infty$ . It has the topology of  $\mathbb{R} \times \mathbb{S}^2$  with the function  $u$  taking values in  $\mathbb{R}$ . In other words, the null hypersurface  $C_u$  intersects  $\mathcal{I}^+$  at infinity in a 2-sphere  $S_{\infty,u}$ .

Let us come back to the notion of black holes. In general, we define the black hole region of an asymptotically flat spacetime  $(\mathcal{M}, g)$  to be the set of points  $B \subset \mathcal{M}$  not in the past of future null infinity  $\mathcal{I}^+$ . We write  $B = \mathcal{M} \setminus J^-(\mathcal{I}^+)$ .<sup>1</sup>

Above we encountered the Schwarzschild and Kerr black holes. In these cases, all causal geodesics  $c(s)$  entering  $B$  are incomplete towards the future. We say that  $(\mathcal{M}, g)$  is future causally geodesically incomplete. In the Schwarzschild solution the curvature grows along all incomplete  $c(s)$  when the affine parameter  $s$  tends to its supremum. The situation for Kerr is more colorful, as causality breaks down.

**2.2. Cauchy problem.** The above-mentioned Schwarzschild and Kerr spacetimes are examples of exact solutions of the Einstein equations. Whereas there exist quite a few closed-form solutions, many important physical situations do not have them. Nor do they help to investigate the space of solutions nor understand the dynamics of GR. In order to study stability problems, gravitational waves, and questions about the dynamics of the gravitational field, we have to solve the Cauchy problem (initial value problem). Exact solutions certainly inspired insights into particular cases, but only solving the Cauchy problem can answer these important questions. We shall see that for the Einstein equations this mainly means proving theorems with geometric-analytic methods. Whereas in other fields of pde analysis, geometry plays less of a role, the geometric nature of GR is crucial and features extra challenges, but it also can be used to our advantage to obtain estimates.

Within GR, other techniques have been used to approximate solutions to the Einstein equations. Among them we find methods in perturbation theory and numerical relativity. Even though the main focus in these other fields is not on proving theorems, they provide important insights into physical problems.

---

<sup>1</sup>This notion of a black hole actually has to be investigated more. We address some issues in this article. For a nice discussion of black hole spacetimes, see [29].

It is clear that all these endeavors have to be understood within the larger realm of the Cauchy problem for the Einstein equations. Moreover, only the mathematical treatment of the latter yields a full understanding of the physical picture with its intricate features.

Let us consider a solution of the EV equations (1). Denote by  $\nabla$  the covariant derivative. Then the Bianchi identities

$$(3) \quad \nabla_{[\alpha} R_{\beta\gamma]\delta\epsilon} := \nabla_{\alpha} R_{\beta\gamma\delta\epsilon} + \nabla_{\beta} R_{\gamma\alpha\delta\epsilon} + \nabla_{\gamma} R_{\alpha\beta\delta\epsilon} = 0$$

are equivalent to the contracted Bianchi identities

$$(4) \quad \nabla^{\alpha} R_{\alpha\beta\gamma\delta} = 0$$

with  $\nabla^{\alpha} := (g^{-1})^{\alpha\beta} \nabla_{\beta}$ .

Under the four constraints from the Bianchi identities, the EV system (1) provides six independent equations for the ten unknowns of the metric  $g_{ij}$ . Here, we encounter the general covariance of the Einstein equations and remark that uniqueness of solutions holds up to the equivalence under diffeomorphisms. This mathematical fact has the following physical meaning: the laws of nature do not depend on the coordinates.

**Classic approach.** The Einstein equations split into a set of constraint equations that the initial data have to obey and a set of evolution equations. The “classic” initial value problem in GR considers a three-dimensional manifold  $H$  with a complete Riemannian metric  $\bar{g}$  and a symmetric 2-tensor  $k$  solving the constraint equations

$$(5) \quad \bar{\nabla}^i k_{ij} - \bar{\nabla}_j \text{tr} k = 0,$$

$$(6) \quad \bar{R} + (\text{tr} k)^2 - |k|^2 = 0,$$

where barred quantities are with respect to  $H$ . The data evolves according to

$$(7) \quad \frac{\partial \bar{g}_{ij}}{\partial t} = -2\Phi k_{ij} + \mathcal{L}_X \bar{g}_{ij},$$

$$(8) \quad \frac{\partial k_{ij}}{\partial t} = (\bar{R}_{ij} + k_{ij} \text{tr} k - 2k_{is} k_j^s) \Phi + \mathcal{L}_X k_{ij} - \bar{\nabla}_i \bar{\nabla}_j \Phi,$$

with  $\Phi := 1/\sqrt{-g^{ij} \partial_i t \partial_j t}$  denoting the lapse function and  $X$  denoting the shift vector. The time vector field is  $T = \Phi N + X$ , and  $\mathcal{L}$  is the Lie derivative. The initial data set  $(H, \bar{g}_{ij}, k_{ij})$  embeds into the Cauchy development  $(\mathcal{M}, g)$ , namely the Lorentzian spacetime, as a spacelike hypersurface. The imbedding  $H \rightarrow \mathcal{M}$  has first, respectively second, fundamental forms  $i_*(\bar{g})$  and  $i_*(k)$ . From this point of view, it is easy to see that the constraint equations (5)–(6) are then implied by the contracted Codazzi and Gauss equations.

A general starting point to attack a pde is local and global well-posedness, followed by proving existence and uniqueness of solutions, and finally an analysis of the solutions. However, in GR we face a few subtleties. One of these was mentioned above already, namely the general covariance of the Einstein equations. Moreover, the differential structure of the spacetime is not known a priori. Then what should be the ordering on the regions where solutions are defined? One would wish for a domain of dependence theorem to hold globally. (We may think of the wave equation as a simple example to inspire our intuition.) The “magic concept” is known as *global hyperbolicity* and means that  $(\mathcal{M}, g)$  admits a Cauchy hypersurface (that is a complete, spacelike hypersurface  $H$  in  $\mathcal{M}$  with each causal curve in the

spacetime intersecting  $H$  exactly once). Then the *maximal Cauchy development* is given in a unique way as the globally hyperbolic spacetime into which all other such spacetimes imbed isometrically.

Among the active players in the very early years of GR, we find D. Hilbert and H. Weyl who contributed substantially to the theory. Let us remind ourselves that in those days the involved mathematical branches had not yet been as developed as today. This caused much confusion about mathematical properties of the Einstein equations and their physical implications. For instance, the pioneers argued about what it means for a solution to behave differently in different coordinate systems. Nowadays, the resolution of these issues is not more than an elegant lemma in geometry. In this context, we understand how important Weyl’s “causally connected” world emerged in 1923 hinting at the contents of the domain of dependence theorem which would be established much later. As it turned out, an important tool was introduced by T. de Donder and C. Lanczos and later used by G. Darmois, namely the wave coordinates, as we shall see below. Many people contributed towards a formulation and understanding of the Cauchy problem in GR, we point out also A. Lichnerowicz, K. Stellmacher, and K. Friedrichs. On the analysis side, important progress that influenced GR came with the works by H. Lewy, J. Hadamard, J. Schauder, and S. Sobolev among many others. In these years, a young woman had made her first and important steps in GR, Yvonne Choquet-Bruhat. She achieved the big breakthrough in the Cauchy problem in her celebrated works summarized below. As the purpose of the present article does not allow us to delve deeper into the history of the mathematical crescendo of the first half of the 20th century, we refer to Y. Choquet-Bruhat’s paper [20] for a more detailed discussion of the mathematical progress in GR in these years, whereas the historical facts will be described in her forthcoming autobiography. See also [12] for a discussion of the Cauchy problem in view of gravitational waves. Whereas many of the initial problems in GR have been solved, other hurdles have remained tough nuts to crack and bear challenges for future mathematical research.

These are some of the reasons why it took a long time until the Cauchy problem for the Einstein equations was even formulated properly. The breakthrough had to wait until 1952 when Y. Choquet-Bruhat [19] proved a local existence and uniqueness theorem for the Einstein equations. And only later were the aforementioned issues resolved. In 1953, J. Leray [52] discussed global hyperbolicity. The second breakthrough took place in 1969 when Y. Choquet-Bruhat and R. Geroch [21] proved the global existence of a unique maximal future development for every given initial data set.

We state the fundamental theorems by Choquet-Bruhat and Choquet-Bruhat with Geroch as follows:

**Theorem 4** (Y. Choquet-Bruhat, 1952 [19]). *Let  $(H, \bar{g}, k)$  be an initial data set satisfying the vacuum constraint equations. Then there exists a spacetime  $(\mathcal{M}, g)$  satisfying the Einstein vacuum equations with  $H \hookrightarrow \mathcal{M}$  being a spacelike surface with induced metric  $\bar{g}$  and second fundamental form  $k$ .*

**Theorem 5** (Y. Choquet-Bruhat and R. Geroch, 1969 [21]). *Let  $(H, \bar{g}, k)$  be an initial data set satisfying the vacuum constraint equations. Then there exists a unique, globally hyperbolic, maximal spacetime  $(\mathcal{M}, g)$  satisfying the Einstein vacuum equations with  $H \hookrightarrow \mathcal{M}$  being a Cauchy surface with induced metric  $\bar{g}$  and second fundamental form  $k$ .*

The main tool in the proof was the use of wave coordinates (often called harmonic coordinates even though the metric is Lorentzian). By definition, wave coordinates  $x^\alpha$  satisfy the wave equation

$$\square_g x^\alpha = 0.$$

This is equivalent to the connection coefficients of these local wave coordinates satisfying

$$g^{mn}\Gamma_{mn}^\alpha = 0.$$

From the fact that the Riemann curvature tensor can be expressed in terms of the connection coefficients, it follows that the EV equations in wave coordinates become

$$(9) \quad \square_g g_{\alpha\beta} = N_{\alpha\beta}(g, \nabla g)$$

with  $N_{\alpha\beta}(g, \nabla g)$  denoting nonlinear terms with quadratics in  $\nabla g$ . Thus, we have a system of quasilinear wave equations. (9) are the so-called reduced Einstein equations. Choquet-Bruhat in her proof studies the Cauchy problem for this reduced system. Combined with the other main idea, which relies on the domain of dependence theorem, this allowed her to prove Theorem 4.

Early analysis of equations of the type like (9) include works by Friedrichs and Lewy and by Schauder via energy methods and by Hadamard, Petrovsky, and Sobolev by constructing a parametrix.

The above Theorems 4 and 5 have been generalized to hold for many matter systems. Moreover, improvements were obtained by Dionne, Fisher and Marsden, and Hughes, Kato, and Marsden using the energy method for initial data given in specific classes of Sobolev spaces. Further improvements followed by Tataru, Smith and Tataru, Klainerman and Rodnianski, and Planchon and Rodnianski. Recently, the  $L^2$  curvature conjecture was proven by Klainerman, Rodnianski, and Szeftel. The latter show that under certain assumptions the regularity of the data can be relaxed so far that the existence of the solution depends only on the  $L^2$ -norms of the Riemannian curvature tensor and on the gradient of the second fundamental form. We only cite the references of the latter, namely [47], [46], [66], [67], [68], [69]; please see [47] for a detailed discussion and an extensive list of references.

**Characteristic approach.** The characteristic initial value problem for the EV equations (1) starts from initial data given on null hypersurfaces. The data are prescribed on either an outgoing null hypersurface or an incoming and an outgoing null hypersurface intersecting in a spacelike 2-surface. A. Rendall [61] in 1990 proved the following theorem.

**Theorem 6** (A. Rendall [61]). *Let characteristic smooth initial data for the Einstein vacuum equations be given on null hypersurfaces  $C_1$  and  $C_2$  that intersect transversely on a spacelike surface  $S = C_1 \cap C_2$ . Then there exists a (nonempty) maximal development  $(\mathcal{M}, g)$  of the initial data bounded in the past by a neighborhood of  $S$  in  $C_1 \cup C_2$ .*

Rendall's proof reduces the problem to the classic Cauchy problem.

If initial data is given on a single outgoing null hypersurface  $C$ , then one has to introduce adapted conditions at the vertex  $o$  of  $C$  for the solution.

In Christodoulou's work [25], which we shall discuss in detail below, the characteristic approach is crucial. To establish his main result, Christodoulou considers data which is trivial up to a surface  $S$ . More precisely, in the context of Rendall's Theorem 6, this corresponds to  $C$  to the future of  $S$  being  $C_1$ , whereas  $C_2$  being the incoming Minkowski cone  $\underline{C}$  rooting in  $S$ .

A basic difference between the classic and the characteristic initial value problem is that in the former the constraints for the initial data are given by elliptic pde, whereas in the latter they can be written as ode (ordinary differential equations). That is, we can specify some data freely and solve propagation equations along the generators of null hypersurfaces. Thus, this aspect is much simpler in the characteristic situation.

The characteristic treatment is more natural for questions concerning gravitational radiation: First, in the investigation of black hole formation through the focussing of gravitational waves, as is the main present topic. Second, to analyze gravitational waves coming from sources like the mergers of black holes. The reason is that these waves travel at the speed of light along null hypersurfaces of the spacetime.

**Stability of Minkowski space.** Considering asymptotically flat systems under gravitation, we would like to understand under which conditions there exist global solutions of a certain smoothness and what their structures are, respectively, when do singularities (black holes) form. The above theorems do not answer these questions but constitute the way to start. The major breakthrough [27] was achieved in 1993 by Christodoulou and Klainerman proving that for asymptotically flat initial data being small in weighted Sobolev spaces there exists a complete maximal development as a solution of the EV equations (1). This is known as the global nonlinear stability of the Minkowski space. We present a summarized version of their theorem:

**Theorem 7** (D. Christodoulou and S. Klainerman, 1993 [27]). *Let be given strongly asymptotically flat initial data for the EV equations (1) being sufficiently small. Then there exists a unique, causally geodesically complete and globally hyperbolic solution  $(\mathcal{M}, g)$ , that itself is globally asymptotically flat.*

The geometric-analytic **proof** is monumental, does not depend on coordinates, and lays open the structures of the solution spacetimes. In a first part, suitable energies are identified in the Bel–Robinson tensor. The latter is basically a quadratic of the Weyl tensor and is used heavily also in Christodoulou’s work [25]. We will give the formula below. Next, the curvature components are estimated from these energies via a comparison argument. Then, in the main part of the proof, constituting a large bootstrap argument, under assumptions on the curvature it is shown that the remaining geometric quantities are controlled. Many of the new features and ideas in the proof have had impact far beyond GR in the study of other nonlinear hyperbolic pde. See also the semiglobal result by Friedrich [37], later proofs under more assumptions and using wave coordinates by Lindblad and Rodnianski [54], [55], as well as the proof for the exterior part with a double null foliation by Klainerman and Nicolò [43]. The Christodoulou–Klainerman result [27] was generalized in 2000 by N. Zipser [79], [80] for the Einstein–Maxwell system, and in 2007 by L. Bieri [6], [7] for the EV equations with fewer assumptions on the decay at infinity and less regularity, thereby establishing the borderline case for decay of the data in the EV situation. The proofs in both these works are geometric-analytic.

A specific feature in the proof of [27] turns out to be crucial not only in order to establish Theorem 7 but also in Christodoulou’s new constructions in [25]: The above foliation into null hypersurfaces  $C_u$  is not arbitrary, but depends on a suitably chosen optical function  $u$ , whereas the spacelike hypersurfaces  $H_t$  are generated by a foliating maximal time function  $t$ . It follows from the proof in [27] that null infinity

$\mathcal{I}^+$  is complete for the data considered; however, what is known as “peeling” does not hold for all the curvature components. To understand the latter, let us focus for a moment on radiation, which we introduced as fluctuation of the curvature. Thus, in order to investigate gravitational waves coming from far-away sources, we need to determine the properties of the curvature components at null infinity. These components are obtained through contraction with vector fields of a null frame given by the above foliations. As the waves propagate at the speed of light (and light travels along null geodesics), we have to follow them along outgoing null geodesics. This type of question was addressed already during the 1960s. Trautman [74], Bondi [15], Bondi, van der Burg, and Metzner [16], Sachs [63], and Penrose [58] pioneered the use of null hypersurfaces to describe gravitational radiation. Other discussions were given by Pirani [60], Newman and Penrose [57], Geroch [39], Ashtekar and Hansen [3], Ashtekar and Schmidt [4], and Ashtekar and Streubel [5].

All these works address gravitational radiation in some way. One of the problems in studying  $\mathcal{I}^+$  arises when (as in some of the cited papers) one would like to expand the metric in power series in  $r^{-1}$  with coefficients depending on  $u$  and the angular coordinates. Or in general we can ask, How smooth should null infinity be? In the later cited papers, the authors replaced the assumptions about the power series expansion by another assumption which also requires a minimal regularity. Thus, if one conformally compactifies the boundary at null infinity, this implies a minimal regularity of the data, which in the aforementioned works would be at least  $C^2$ . However, Christodoulou showed that for physical spacetimes  $C^2$  is impossible. In fact, in the general case considered here, the conformal factor extends to  $\mathcal{I}^+$  as a function in  $C^{1,\alpha}$ . The works by Christodoulou and Klainerman [27] are within that regime. From the smoothness follows a specific hierarchy of decay for the curvature components, which is called peeling. Today, we mainly refer to the stronger assumptions as the *Newman–Penrose picture* and to the later more general situation as the *Christodoulou–Klainerman picture*.

The above foliation developed in [27] is natural in the way that one follows the waves along the  $C_u$ . In the corresponding null frame the structure equations relate curvature components with the connection coefficients. A particularly interesting representative of the latter in view of radiation is the shear  $\hat{\chi}$  which satisfies

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

where slashed quantities are on the surfaces  $S_{t,u}$ ,  $\beta$  denotes a curvature component, and  $\zeta$ , the torsion 1-form, is another connection coefficient. These equations lay open structures of the spacetime that had previously been inaccessible; in particular it could not be captured by the corresponding equations of the Newman–Penrose formalism. In [27] the authors derived thereby a new method to treat the Cauchy problem for the Einstein equations (being hyperbolic), coupling elliptic equations for  $\hat{\chi}$  and related quantities on  $S_{t,u}$  with propagation equations along the null hypersurfaces  $C_u$ . With the coupling term for the above elliptic system being  $\nabla \text{tr}\chi$ , the propagation equation along the generators of  $C_u$  reads (above and here we omit the indices for simplicity)

$$\frac{\partial \nabla \text{tr}\chi}{\partial s} + \text{tr}\chi \nabla \text{tr}\chi + 2\hat{\chi} \nabla \hat{\chi} = 0.$$

This kind of structure plays an important role as well in [25]. As explained in the sketch of the proof of Theorem 7, energy estimates for the curvature precede the

handling of  $\hat{\chi}$  and related quantities. In a bootstrap argument with control on the curvature, the connection coefficients are estimated.

In Christodoulou's monumental work on black hole formation, incoming gravitational waves follow  $C_u$ . This will be treated in Section 4. The latest works on stability of black holes is addressed in Section 5. Before concentrating on ingoing waves however, in the next chapter we say a few words about outgoing gravitational waves that were detected by Advanced LIGO [1] in 2015.

### 3. GRAVITATIONAL RADIATION

**3.1. Gravitational waves.** Gravitational waves are fluctuations of the spacetime curvature traveling at the speed of light along null hypersurfaces  $C_u$ . For the first time, Advanced LIGO [1] detected such waves in 2015. This constitutes the beginning of a new era in science where these waves through Advanced LIGO and other detectors will reveal information from regions of the universe that have been opaque to telescopes or as yet unknown. Thereby, the mathematical understanding of the dynamics of the Einstein equations, in particular the Cauchy problem, will be crucial.

A typical source for gravitational waves is the merger of two black holes. They will spiral in and finally merge, thereby radiating away energy in the form of gravitational waves (Figure 4).

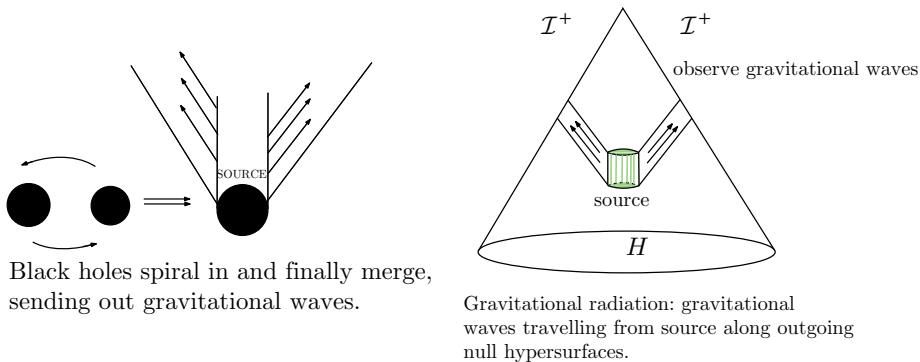


FIGURE 4

Let us consider future null infinity  $\mathcal{I}^+$  defined above and the Christodoulou–Klainerman result of Theorem 7, but now we omit the smallness assumptions and start with large initial data that might even contain black holes. More precisely, let  $(H_0, \bar{g}, k)$  be an arbitrary asymptotically flat initial data set, with only vacuum outside a compact set. By the domain of dependence theorem and the results of [27], one can show that in the new situation we can still attach a piece of asymptotic boundary  $\mathcal{I}^+$  to the Cauchy development of the initial data,  $\mathcal{I}^+$  being parametrized by  $(-\infty, u_+) \times S^2$  as opposed to  $(-\infty, +\infty) \times S^2$  from before. Thus, limits along appropriate null hypersurfaces  $C_u$  can be computed even in the new case. This follows also from [43].

Gravitational radiation is described on  $\mathcal{I}^+$ . Most important are the limits of the shears  $\underline{\hat{\chi}}$ ,  $\hat{\chi}$  and the curvature component  $\underline{\alpha}$  which is contracted twice with the incoming null vectorfield  $\underline{L}$  and decaying like  $r^{-1}$ . The radiative amplitude per unit

solid angle is given by the limit

$$(10) \quad \Xi(u, \theta) = \lim_{C_u, t \rightarrow \infty} r \hat{\chi}$$

being a symmetric, traceless 2-tensor. Similarly, one defines

$$(11) \quad \Sigma(u, \theta) = \lim_{C_u, t \rightarrow \infty} r^2 \hat{\chi}.$$

In the last chapter of [27] these limits and the asymptotic structures of the space-times are derived. The following crucial relations emerge from these studies:

$$(12) \quad \frac{\partial \Sigma}{\partial u} = -\Xi,$$

$$(13) \quad \frac{\partial \Xi}{\partial u} = -\frac{1}{4}A,$$

with  $A$  denoting the limit of the curvature component  $\underline{a}$ ,

$$(14) \quad A(u, \theta) = \lim_{C_u, t \rightarrow \infty} r \underline{a}.$$

It is interesting to see that the intricate local structures of radiative spacetimes evolve into a simpler picture at null infinity  $\mathcal{I}^+$ . This is then directly related to gravitational wave experiments. Gravitational radiation changes the spacetime while traveling through.

Doing a gravitational wave experiment, we can think of ourselves (and the detector) sitting at null infinity  $\mathcal{I}^+$  as the waves are coming from sources very far away. In fact, we identify our position at retarded time  $u_*$  as  $(u_*, \theta) \in \mathcal{I}^+$  and evolving in  $u$ . A detector like Advanced LIGO consists of three test masses  $m_0$ ,  $m_1$ ,  $m_2$  suspended by pendulums (or floating on their geodesics if the experiment is done in space). These masses are located in an  $L$ -shape (see Figure 5) and are at large distance  $r$  and angular direction  $\theta$  from the source. By laser interferometry the distances of  $m_1$  and  $m_2$  with respect to the reference test mass  $m_0$  are measured. The  $u$ -rate of change of this relative displacement is determined by  $\Xi(u, \theta)$ .

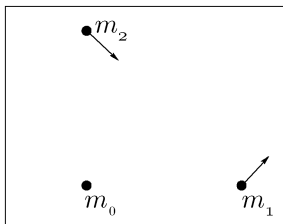


FIGURE 5

These test masses follow geodesics, and for an experiment on Earth, such as Advanced LIGO, their relative acceleration is expressed through curvature in the Jacobi equation,

$$(15) \quad \nabla_U^2 V = R(U, V)V,$$

where  $U$  denotes the tangent vector for an object in free fall separated from a second object by a vector  $V$ .

**3.2. Memory effect of gravitational waves.** So far, we have thought of gravitational waves in a region of the spacetime  $(\mathcal{M}, g)$  as changing that region “instantaneously” while traveling through. However, there is more to the story. It is predicted that gravitational radiation permanently changes the spacetime, leaving a footprint in the regions it passes. This is called the *memory effect of gravitational waves*, today known as the *Christodoulou effect*. In 1974, Ya. B. Zel’dovich and A. G. Polnarev [78] found such an effect for the linearized Einstein equations. However, it was believed to be too small for detection. In 1991, D. Christodoulou [23] investigated the full Einstein equations and derived a memory that was larger than the one of the previous work. In fact, L. Bieri and D. Garfinkle showed [11] that these are two different effects. For memory, see also the following works by Braginsky and Grishchuk [17], Braginsky and Thorne [18], Blanchet and Damour [14], Frauendiener [38], Wiseman and Will [76], and more recently Bieri, Chen, and S.-T. Yau [8], [9], Bieri and Garfinkle [10], [11], Tolish and Wald [73], Flanagan and Nichols [35], [36], and Favata [33]. We refer to these articles for further references.

To explain how this effect can be measured, we consider the Jacobi equation (15). In Figure 5, showing the three test masses, the arrows refer to a permanent displacement in the horizontal plane. For simplicity of the discussion, assume that the wave source is perpendicular to the plane of the three test masses. Equation (15) gives an acceleration on the left-hand side and curvature on the right-hand side. Here is where the two relations (12) and (13) play a crucial role. First using (13) in (15) where the leading order curvature term is  $A$ , and integrating once, second using (12) to substitute the shear terms and integrate again, and finally taking the limit as  $u \rightarrow \infty$ , we obtain an equation of the form (omitting indices)

$$(16) \quad \Delta x = -\frac{d}{r}(\Sigma^+ - \Sigma^-),$$

where  $\Sigma^\pm$  denotes the limits of  $\Sigma$  when  $u \rightarrow +\infty$ , respectively  $u \rightarrow -\infty$ , and  $\Delta x$  is the distance of the permanent relative displacement and  $d$  is the initial separation. It can be shown by geometric-analytic investigations that  $\Sigma^+ - \Sigma^-$  is related to

$$(17) \quad F = \mathcal{C} \int_{-\infty}^{+\infty} |\Xi|^2 du$$

with  $\frac{F}{4\pi}$  being the total energy radiated away in a given direction per unit solid angle. An experiment on how to detect gravitational wave memory with Advanced LIGO has recently been suggested in [49] by P. Lasky, E. Thrane, Y. Levin, J. Blackman, and Y. Chen.

#### 4. THE FORMATION OF CLOSED TRAPPED SURFACES

**4.1. Precise formulation of the result: Initial data and evolution.** Let us consider incoming gravitational waves concentrating in a small time interval. The claim of Theorem 3 by Christodoulou is that if the amount of energy is above a certain threshold, then a closed trapped surface will form. We are now going to state this result in a more precise way, explain the setting, and discuss the main ideas of the monumental proof.

We start by considering a spacetime manifold  $(\mathcal{M}, g)$  with boundary, being a smooth solution of the Einstein vacuum equations (1) such that the past boundary of  $\mathcal{M}$  is the future null geodesic cone  $C_0$  of a point  $O$ . Initial data is given on

$C_0$ , assuming that it is trivial in a neighborhood of  $O$ , and our  $(\mathcal{M}, g)$  is to be a development of this initial data. (See Figure 7.) We introduce  $T$  to be a unit future-directed timelike vector at the vertex  $O$  and denote by  $\gamma_0$  the geodesic generated by  $T$  and such that the tangent vector at  $O$  to each null geodesic generator has projection  $T$  along  $T$ . The generators of the cone  $C_0$  are parametrized by an affine parameter  $s$  measured from  $O$  and such that  $s$  is the parameter of the geodesic null vectorfield  $L'$  along  $C_0$  being the tangent field of each generator. Now, the initial data is assumed to be trivial for  $s \leq r_0$  for some  $r_0 > 1$ . This means that they coincide with the data corresponding to a truncated cone in Minkowski spacetime. Then the boundary of this region with trivial data is a round sphere of radius  $r_0$ . In particular, each generator extends up to parameter value  $r_0 + \delta$ , where  $\delta$  is a constant  $1 \geq \delta > 0$ . It is assumed that  $C_0$  does not contain any conjugate or cut points.

The domain of dependence theorem guarantees that the solution spacetime has a region that is Minkowskian and that is given by the past of a backwards light cone  $\underline{C}_e$  of a point  $e$  on  $\gamma_0$  at distance  $2r_0$  from  $O$ .

An advanced time function  $\underline{u}$  on  $C_0$  is defined by

$$(18) \quad \underline{u} = s - r_0.$$

In the proof, the spacetime will be constructed from the initial data, whereby the level sets  $C_{\underline{u}}$  of  $\underline{u}$  are required to be ingoing null hypersurfaces.

In analogy to the previous sections, we denote by  $\hat{\chi}$  the trace-free part of  $\chi$ , which is the second fundamental form of the sections  $S_s$  of  $C_0$  corresponding to constant values of the affine parameter  $s$ , and we let  $\hat{g}$  be the induced metric on  $S_s$ . Thus,  $\hat{\chi}$  is the shear of these sections.

We introduce the following crucial function:

$$(19) \quad e = \frac{1}{2} |\hat{\chi}|_{\hat{g}}^2.$$

This  $e$  is an invariant of the conformal intrinsic geometry of  $C_0$ .

With these tools, we now state the following version of the main theorem:

**Theorem 8** (Closed trapped surface formation; D. Christodoulou [25]). *Let  $k, l$  be positive constants such that  $k > 1 > l$ . Let initial data be given as described above, and assume that*

$$(20) \quad \frac{r_0^2}{8\pi} \int_0^\delta e \, d\underline{u} \geq \frac{k}{8\pi}$$

*with the integral along  $C_0$  where  $\underline{u} \in [0, \delta]$  for some  $\delta > 0$ . Then, if  $\delta$  is suitably small, the maximal development of the data contains a closed trapped surface  $S$  diffeomorphic to  $S^2$  and has area*

$$(21) \quad \text{Area}(S) \geq 4\pi l^2.$$

The notion in Theorem 3 of incoming energy per unit solid angle in each direction in a suitably small time interval is replaced in Theorem 8 by the left-hand side of (20). The reason is that the incoming energy per unit solid angle in each direction in the advanced time interval  $[0, \delta]$  is only defined at past null infinity. Thus, from Theorem 8 Christodoulou derives another result with initial data given at past null infinity, formulated for the moment in Theorem 9 below. We will analyze the details of Theorem 9 in the next sections. For these investigations one has to let  $r_0 \rightarrow \infty$  and thereby move  $C_0$  back to past null infinity  $\mathcal{I}_-$ . From our previous discussions

about radiation at future null infinity  $\mathcal{I}_+$ , it is straightforward to see that the limit of the rescaled quantity of  $e$  corresponds to the analogue of  $|\Xi|^2$  defined through (10), and because we are at past null infinity, we have to replace  $\hat{\chi}$  by  $\hat{\chi}$ .

**Theorem 9** (Closed trapped surface formation; D. Christodoulou [25]). *Let  $k, l$  be constants as in Theorem 8. Let smooth asymptotic initial data be given at past null infinity  $\mathcal{I}_-$  being trivial for  $\underline{u} \leq 0$ . Assume that the incoming energy per unit solid angle in each direction in the advanced time interval  $[0, \delta]$  is greater than or equal to  $\frac{k}{8\pi}$ . Then, if  $\delta$  is suitably small, the maximal development of the data contains a closed trapped surface  $S$  diffeomorphic to  $S^2$  and has area*

$$(22) \quad \text{Area}(S) \geq 4\pi l^2.$$

#### 4.2. The optical structure.

**Double null foliation.** This foliation relies on the two optical functions  $u$  and  $\underline{u}$ . In (18) we defined  $\underline{u}$ . We emphasize that for each value  $v$  the corresponding level set of  $\underline{u}$  is the incoming null hypersurface  $\underline{C}_v$ . Now, the function  $u$  conjugate to  $\underline{u}$  is introduced such that for each  $v$  the  $v$ -level set of  $u$  is the outgoing null hypersurface  $C_v$ . The notation hereafter for these outgoing null hypersurfaces is  $C_u$  and, for the incoming ones,  $\underline{C}_u$ . We note that the outgoing level sets  $C_u$  of  $u$  emanate from points on  $\gamma_0$ , and that  $u|_{\gamma_0}$  measures arc length from  $O$  along  $\gamma_0$  minus  $r_0$ .

The intersections

$$(23) \quad S_{\underline{u}, u} = \underline{C}_{\underline{u}} \cap C_u$$

are spacelike 2-surfaces.

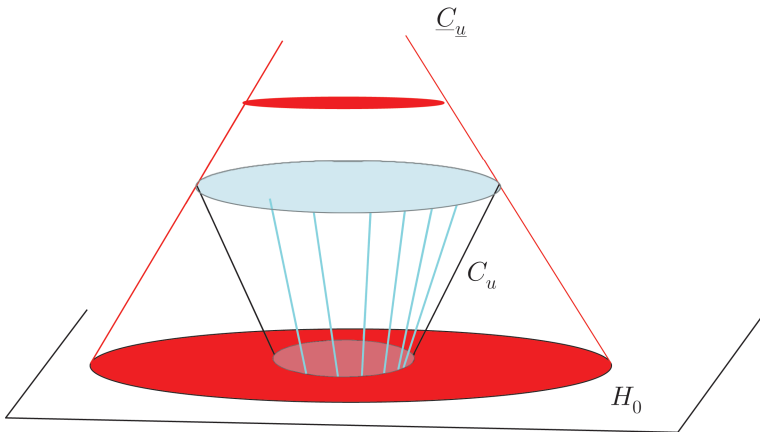


FIGURE 6

Furthermore, we define the  $H_t$  by

$$u + \underline{u} = t.$$

It is clear that we do not expect this foliation to exist for very long. Though it will exist up to a null hypersurface  $\underline{C}_\delta$  and a hypersurface  $H_c$  for small  $\delta$  and for  $u_0 < c < 0$ . We denote this region by  $M_R$ .

The proof of the main theorem (Theorems 8 and 9) will rely on a continuity argument: Regarding  $(\mathcal{M}, g)$  we suppose that the generators of  $\underline{C}_u$  and  $C_u$  have

no end points in  $\mathcal{M} \setminus \gamma_0$ . This means in particular that the  $C_u$  do not contain any conjugate or cut points in  $\mathcal{M}$  and that the  $\underline{C}_u$  do not contain any focal or cut points in  $\mathcal{M} \setminus \gamma_0$ . Then it follows that in  $M_R \setminus \gamma_0$  the  $\underline{C}_u$  and  $C_u$  are smooth,  $H_t$  is a spacelike hypersurface, and the spacelike 2-surfaces  $S_{\underline{u},u}$  are embedded in  $M_R \setminus \gamma_0$ . (We note that the  $S_{\underline{u},u}$  are diffeomorphic to  $S^2$ .)

The optical functions  $u$  and  $\underline{u}$  themselves obey the eikonal equation,

$$\begin{aligned} (g^{-1})^{\mu\nu} \partial_\mu u \partial_\nu u &= 0, \\ (g^{-1})^{\mu\nu} \partial_\mu \underline{u} \partial_\nu \underline{u} &= 0. \end{aligned}$$

Next, we are going to introduce three *null frames* related to this structure.

First, the future-directed null geodesic vectorfields  $L'$  and  $\underline{L}'$  are given by

$$(24) \quad L'^\mu = -2(g^{-1})^{\mu\nu} \partial_\nu u, \quad \underline{L}'^\mu = -2(g^{-1})^{\mu\nu} \partial_\nu \underline{u}.$$

From this is defined a positive function  $\Omega$  as follows:

$$-g(L', \underline{L}) = 2\Omega^{-2}.$$

We observe that  $\Omega$  is the inverse density of the double null foliation.

Second, the normalized vectorfields  $\hat{L}$  and  $\hat{\underline{L}}$  are defined as

$$(25) \quad \hat{L} = \Omega L', \quad \hat{\underline{L}} = \Omega \underline{L}'.$$

They satisfy

$$g(\hat{L}, \hat{\underline{L}}) = -2.$$

Third, we define the vectorfields  $L$  and  $\underline{L}$  as

$$(26) \quad L = \Omega^2 L', \quad \underline{L} = \Omega^2 \underline{L}'.$$

They satisfy

$$\begin{aligned} Lu &= 0 = \underline{L}u, \\ L\underline{u} &= 1 = \underline{L}u. \end{aligned}$$

We find that the integral curves of  $L$  are the generators of the outgoing null hypersurfaces  $C_u$  parametrized by  $\underline{u}$  and that the integral curves of  $\underline{L}$  are the generators of the incoming null hypersurfaces  $\underline{C}_u$  parametrized by  $u$ . Following the notation of [25], we define the flow  $\Phi_\tau$  generated by  $L$  on any  $C_u$  and the flow  $\underline{\Phi}_\tau$  generated by  $\underline{L}$  on any  $\underline{C}_u$ . The  $\Phi_\tau : S_{\underline{u},u} \rightarrow S_{\underline{u}+\tau,u}$  and  $\underline{\Phi}_\tau : S_{\underline{u},u} \rightarrow S_{\underline{u},u+\tau}$  are diffeomorphisms.

Thus, the above structures yield *canonical coordinate systems*. Let  $(\theta^1, \theta^2)$  be local coordinates on a domain  $U$  on  $S_{0,u_0}$ . Then we can extend these to  $\underline{\Phi}_u(\Phi_u(U)) \subset S_{\underline{u},u+u_0}$ . Now, given two domains  $U_1$  and  $U_2$  with coordinates  $(\theta^A; A = 1, 2)$ , respectively  $(\theta'^A; A = 1, 2)$ , covering  $S_{0,u_0}$ ,  $M_R \setminus \gamma_0$  is covered by two regions with coordinates  $(\underline{u}, u, \theta^A; A = 1, 2)$  and  $(\underline{u}, u, \theta'^A; A = 1, 2)$ , respectively.  $(\underline{u}, u) \in D \setminus \gamma_0$ . The domain  $D$  is depicted in Figure 7.

The vectorfields  $\underline{L}$  and  $L$  in these canonical coordinates read as

$$\underline{L} = \frac{\partial}{\partial u}, \quad L = \frac{\partial}{\partial \underline{u}} + b^A \frac{\partial}{\partial \theta^A},$$

with  $b^A$  obeying

$$\frac{\partial b^A}{\partial u} = 4\Omega^2 \zeta^{\#A}$$

for the torsion  $\zeta$ . (It is  $\zeta^{\#A} = (\not{g}^{-1})^{AB}\zeta_B$  and  $\zeta_A = \zeta(\frac{\partial}{\partial\theta^A})$ .) Finally, the metric  $g$  in these coordinates is given by

$$(27) \quad g = -2\Omega^2(du \otimes d\underline{u} + d\underline{u} \otimes du) + \not{g}_{AB}(d\theta^A - b^A d\underline{u}) \otimes (d\theta^B - b^B d\underline{u}).$$

It turns out that stereographic coordinates on the sphere are especially nice to work with in this problem.

**4.3. “The” theorems.** The overarching principle of the *proof of black hole formation* is twofold: First, the spacetime has to be constructed and it has to be shown that it exists long enough such that a closed trapped surface can form. This solution has to be “nice” enough (i.e., sufficiently smooth, no focal points for instance). Second, the formation of a trapped surface has to be proven.

These are formulated in “*the two main theorems*”, namely the *existence* theorem, Theorem 10, and the *closed trapped surface formation* theorem, Theorem 8.

The former is much more difficult and subtle to establish than the latter. Therefore, the sketch of the proof of Theorem 10 will take most of the remaining part of this article, whereas Theorem 8 is shown in a more straightforward manner. One main reason for this lies in the existence of the double null foliation requiring that (in the region of interest) the incoming and outgoing null hypersurfaces do not produce caustics. Generally, it is not unusual for null hypersurfaces to generate caustics, especially when gravity is strong. Thus, Christodoulou [25] had to investigate (and thereby mathematically describe) physical situations where gravity is sufficiently strong to form a trapped surface but not too strong to form caustics (or at least not until after the trapped surface has formed).

In order to state the existence theorem, Theorem 10, we briefly revisit the initial data from Section 4.1. There we give initial data on a complete future null geodesic cone  $C_0$ , the data being trivial for  $s \leq r_0$ . We consider the restriction of the initial data to  $s \leq r_0 + \delta$ , thus restricting to  $\underline{u} \in [0, \delta]$  where the data is trivial for  $\underline{u} \leq 0$ . At the vertex  $O$  and thereby on  $C_{u_0}$ , we set  $u$  equal to  $u_0 = -r_0$ . As  $2(u - u_0)$  along  $\gamma_0$  equals the arc length from  $O$ , our  $u$  is determined everywhere. Therefore, the spacetime that we want to construct will be bounded in the future by  $H_{-1}$  where  $\underline{u} + u = -1$ , and by  $\underline{C}_\delta$ . Following the notation of [25], we call this development  $M_{-1}$ , and we call  $M'_{-1}$  the nontrivial region of  $M_{-1}$  given by  $\underline{u} > 0$ .

The existence theorem is stated as follows.

**Theorem 10** (Existence; D. Christodoulou [25]). *Let initial data be given as described in Section 4.1 and in the previous paragraph. Let  $\delta$  be sufficiently small. Then the maximal development under the Einstein vacuum equations (1) contains a region  $M_{-1}$  where the foliation given in Section 4.2 can be constructed such that the null hypersurfaces  $C_u$  and  $\underline{C}_{\underline{u}}$  contain no cut or conjugate points.  $M_{-1}$  is bounded in the future by the spacelike hypersurface  $H_{-1}$  and the incoming null hypersurface  $\underline{C}_\delta$ .*

This existence theorem not only establishes the solution but it also yields important information on the geometric and analytic structures of these solutions, thus on spacetimes where closed trapped surfaces begin to form.

**4.4. Proof of the existence theorem.** Before we start discussing the essentials of the proof, let us ask the following question: Knowing the energy method from pde theory, does there exist something similar for the Einstein equations? Yes, there is a generalized form of this idea as we mentioned already above in the ideas

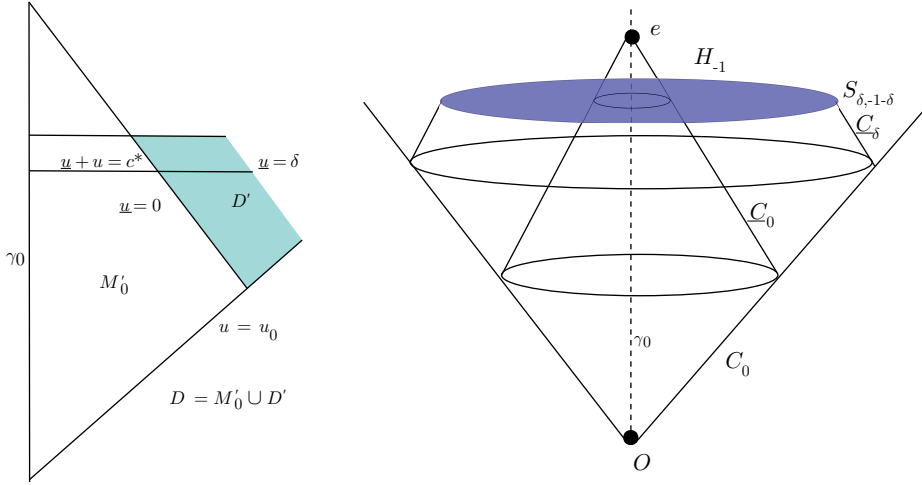


FIGURE 7

of the proof of Theorem 7. As an ultra short summary we may recall that energy controls curvature which controls the other geometric quantities in a bootstrap argument. Of course, interesting structures are hiding behind these concepts that have to be unraveled. It is important that the energies are constructed with respect to “useful” vectorfields. In the following outline of the main points of the proof of Theorem 10, we are going to investigate these structures and connect them with new features of the problem under study.

In this section, we give a sketch of the main methods and ideas of the proof of the existence theorem, Theorem 10. While two of these methods were introduced by Christodoulou and Klainerman in [27], the third method was developed by Christodoulou in [25].

One of the methods from [27] is specific to the Einstein equations, whereas the other can be applied to a broad range of nonlinear hyperbolic pde, in particular to all Euler–Lagrange systems of hyperbolic pde. The first of these methods concerns the Bianchi identities (3) with the goal to obtain estimates for the spacetime curvature  $W$ . We remark that in an Einstein vacuum spacetime  $(\mathcal{M}, g)$  the curvature  $R_{\alpha\beta\gamma\delta}$  is equal to the Weyl curvature  $W_{\alpha\beta\gamma\delta}$ . (This is clear from the fact that the Riemannian curvature  $R_{\alpha\beta\gamma\delta}$  splits into its traceless part  $W_{\alpha\beta\gamma\delta}$  and a part containing the Ricci and scalar curvature which are zero by virtue of the EV equations (1).) More generally, a *Weyl field*  $W$  on  $(\mathcal{M}, g)$  is a 4-covariant tensorfield with the algebraic properties of the Weyl or conformal curvature tensor. Left  ${}^*W$  and right  $W^*$  Hodge duals of  $W$  can be defined, and it is shown that  ${}^*W = W^*$ . There is a nice analogy with Maxwell’s theory of electromagnetism. A Weyl field satisfies equations that are similar to Maxwell’s equations for an electromagnetic field. The *Bianchi equations* for  $W$  are the linear equations

$$(28) \quad \nabla^\alpha W_{\alpha\beta\gamma\delta} = J_{\beta\gamma\delta},$$

with  $J_{\beta\gamma\delta}$  being a *Weyl current*. Moreover, the following holds

$$\nabla_{[\alpha} W_{\beta\gamma]\delta\epsilon} = \epsilon_{\mu\alpha\beta\gamma} J^{*\mu}_{\delta\epsilon},$$

where  $\epsilon$  denotes the volume form for of the spacetime manifold and  $J_{\beta\gamma\delta}^* = \frac{1}{2} J_{\beta}^{\mu\nu} \epsilon_{\mu\nu\gamma\delta}$ . If the Weyl field  $W$  is the spacetime curvature itself, then the corresponding Weyl current vanishes and we are back to the Bianchi identities (3).

Let us define the deformation tensor  $^{(Y)}\hat{\pi}$  of  $Y$  to be the trace-free part of  $\mathcal{L}_Y g$ . We may think of this quantity to measure how the conformal geometry of  $(\mathcal{M}, g)$  changes under the flow of  $Y$ .

We have to deal with more general Weyl fields. And as the Lie derivative  $\mathcal{L}_Y W$  of a Weyl field is in general not a Weyl field because it has trace, Christodoulou and Klainerman introduce the modified Lie derivative  $\hat{\mathcal{L}}_Y W$ , which is trace free and has all the other properties of a Weyl field, thus it is a Weyl field. Similarly, this holds for  $J$ . Due to certain conformal covariance properties of the Bianchi equations, it follows that the Weyl current associated to  $\hat{\mathcal{L}}_Y W$  is the sum of  $\hat{\mathcal{L}}_Y J$  and a bilinear term that is itself linear in  $^{(L)}\hat{\pi}$  and its first covariant derivative and in  $W$  and its first covariant derivative.

New Weyl fields are generated from the original curvature tensor of  $(\mathcal{M}, g)$  by consecutively applying  $\hat{\mathcal{L}}_{Y_i}$  for  $i = 1, \dots, n$  where  $Y_1, \dots, Y_n$  are *commutation vectorfields*.

In order to make use of some form of the energy method for the Einstein equations, we introduce the Bel–Robinson tensor  $Q(W)$  associated to  $W$ . This  $Q$  plays a role similar to the energy momentum–stress tensor  $T$  for the electromagnetic field.

$$(29) \quad Q_{\alpha\beta\gamma\delta} = \frac{1}{2} (W_{\alpha\rho\gamma\sigma} W_{\beta}^{\rho\sigma} + {}^*W_{\alpha\rho\gamma\sigma} {}^*W_{\beta}^{\rho\sigma}) .$$

It satisfies the positivity condition

$$(30) \quad Q(X_1, X_2, X_3, X_4) \geq 0,$$

where  $X_1, X_2, X_3$ , and  $X_4$  are future-directed causal vectors.  $Q$  is symmetric and trace free in any pair of indices. Moreover, if  $W$  satisfies the Bianchi equations, then  $Q$  is divergence free,

$$(31) \quad D^\alpha Q_{\alpha\beta\gamma\delta} = 0.$$

In general,  $\text{div} Q$  equals an expression linear in  $W$  and in  $J$ .

From  $Q$  for a given  $W$  we define the *energy-momentum density* vectorfield  $P(W; X_1, X_2, X_3)^\alpha$  associated to  $W$  and to the three *multiplier vectorfields*  $X_1, X_2, X_3$ , which are future-directed causal:

$$(32) \quad P(W; X_1, X_2, X_3)^\alpha = -Q(W)^\alpha_{\beta\gamma\delta} X_1^\beta X_2^\gamma X_3^\delta .$$

It follows that  $\text{div} P$  equals the sum of  $-(\text{div} Q(W))(X_1, X_2, X_3)$  and a bilinear expression that is linear in  $Q(W)$  and in  $^{(X_1)}\hat{\pi}$ ,  $^{(X_2)}\hat{\pi}$ ,  $^{(X_3)}\hat{\pi}$ .

The divergence theorem in spacetime together with the positivity property of  $Q(W)$  yield control of all the derivatives of the curvature up to required order  $m$ . This latter control is realized via the integrals on the future boundary.

The second method from [27] (with wide applications) used in [25] is tightly interwoven with the first one and is rooted in the specific foliations of the spacetime (discussed above). Whereas the optical function  $u$  lay open the structures of the natural flow of outgoing gravitational waves along the  $C_u$  and played a crucial role in [27] and [23], the optical function  $\underline{u}$  is most crucial in [25] because it follows incoming gravitational radiation along the  $\underline{C}_{\underline{u}}$ . The trapped surfaces that form in the evolution in [25] are sections  $S_{\underline{u}, u}$  given in (23) of “late”  $C_u$  everywhere along

which precisely these  $C_u$  have negative expansion. These structures also allow us to identify natural vectorfields that are used within the energy method. One requirement is that the set of commutation vectorfields has to span the tangent space to  $\mathcal{M}$  at each point.

Finally, the third and newest method of Christodoulou in [25] is the so-called *short pulse method*. It has a wide range of applications in other nonlinear pde. It is a specific way of treating the focusing of incoming waves. A main feature is that initial data has to be sufficiently *large* so that a closed trapped surface will form. In a broader context, one can consider Euler–Lagrange systems of (nonlinear) hyperbolic pde and establish an existence theorem for large initial data and study the evolution to understand interesting phenomena that may occur.

We recall that a heuristic version of the main theorem above includes “large” incoming energy during a “small” time interval. Thus, there is a small parameter involved that we call  $\delta$ . In what follows we will analyze the role of this parameter in the data thereby explaining the short pulse method.

After a technical setup, one starts with the simple task to analyze the equations along  $C_{u_0}$ . As  $C_{u_0}$  is a null hypersurface and one faces a characteristic initial value problem, one can prescribe free data not worrying about constraints. Then the full set of data (including all the curvature components and their transversal derivatives) is easily obtained by integrating ode along the generators of  $C_{u_0}$ .

The free data can be given as 2-covariant, symmetric, positive definite tensor density  $m$  on the sphere and depending on  $\underline{u}$ , moreover  $m$  being of weight  $-1$  and  $\det m = 1$ . In particular, we write

$$(33) \quad m = \exp \psi$$

with  $\psi \in \hat{S}$  the latter denoting the space of symmetric, trace-free, two-dimensional matrices. Thus,  $\exp : \hat{S} \rightarrow H_1^+$  is an analytic diffeomorphism. The transformation rule turns out to be especially simple for stereographic charts on  $S^2$ .

For the short pulse ansatz in [25] Christodoulou considers an arbitrary, two-dimensional, smooth, symmetric, trace-free matrix-valued map  $\psi_0$  on  $S^2$  that depends on  $s \in [0, 1]$  and that extends smoothly to  $s \leq 0$ . Introduce the following:

$$(34) \quad \psi(\underline{u}, \theta) = \frac{\delta^{\frac{1}{2}}}{|u_0|} \psi_0\left(\frac{\underline{u}}{\delta}, \theta\right), \quad (\underline{u}, \theta) \in [0, \delta] \times S^2.$$

Then the equations along  $C_0$  are analyzed and yield a specific structure of the spacetime curvature along  $C_{u_0}$ . In order to state these, we decompose the spacetime curvature  $R_{\alpha\beta\gamma\delta}$  with respect to the natural foliation discussed above. This yields the following for any vectors  $X, Y \in T_P S_{\underline{u}, u}$  at a point  $p$ :

$$\begin{aligned} \alpha(X, Y) &= R(X, \hat{L}, Y, \hat{L}), \\ \underline{\alpha}(X, Y) &= R(X, \underline{\hat{L}}, Y, \underline{\hat{L}}), \\ \beta(X) &= \frac{1}{2} R(X, \hat{L}, \underline{\hat{L}}, \hat{L}), \\ \underline{\beta}(X) &= \frac{1}{2} R(X, \underline{\hat{L}}, \underline{\hat{L}}, \hat{L}), \\ \rho &= \frac{1}{4} R(\underline{\hat{L}}, \hat{L}, \underline{\hat{L}}, \hat{L}), \\ \sigma_{\ell}(X, Y) &= \frac{1}{2} R(X, Y, \underline{\hat{L}}, \hat{L}), \end{aligned}$$

where  $\not\!\!d$  denotes the area form of  $S_{u,u}$ . We note that  $\alpha, \underline{\alpha}$  are symmetric, 2-covariant, trace-free tensorfields on  $S_{u,u}$ , whereas  $\beta, \underline{\beta}$  are 1-forms on  $S_{u,u}$  and  $\rho, \sigma$  are functions on these surfaces. The following structures are obtained for these components along  $C_{u_0}$ :

$$\begin{aligned}
 \sup_{C_{u_0}} |\alpha| &\leq O_2(\delta^{-\frac{3}{2}} |u_0|^{-1}), \\
 \sup_{C_{u_0}} |\beta| &\leq O_2(\delta^{-\frac{1}{2}} |u_0|^{-2}), \\
 \sup_{C_{u_0}} |\rho|, \sup_{C_{u_0}} |\sigma| &\leq O_3(|u_0|^{-3}), \\
 \sup_{C_{u_0}} |\underline{\beta}| &\leq O_4(\delta |u_0|^{-4}), \\
 \sup_{C_{u_0}} |\underline{\alpha}| &\leq O_5(\delta^{\frac{3}{2}} |u_0|^{-5}).
 \end{aligned}
 \tag{35}$$

Pointwise quantities are taken with respect to the induced metric  $\not\!\!g$  on the 2-surfaces.<sup>2</sup>

One of the deep insights of the work [25] is disclosed on the right-hand side of (35): This particular dependence on  $\delta$  of the curvature components is called the *short pulse hierarchy*. Indeed, we read off a *nonlinear* hierarchy. (A direct computation shows that the linearized equations would give a different hierarchy.)

With the ansatz (34) one observes that the amplitude of the pulse is proportional to the square root of the pulse length. This interesting relation only shows in nonlinear theory, it does not exist in a linear one. A closer look makes evident that this hierarchy is required for trapped surfaces to form in  $M_{-1}$ . Thus, it is the heart of the proof of the existence theorem, Theorem 10.

The main challenge of this method is to prove that the specific hierarchy is preserved in the evolution. Towards this goal, vectorfields with specific weights are chosen in connection with the energies and currents defined above. In particular, as multiplier vectorfields we take  $L$  and  $K$  with

$$K = u^2 \underline{L}.$$

Then for each Weyl field (curvature and modified Lie derivatives of curvature) one defines energy-momentum density vectorfields  $P(W; X_1, X_2, X_3)^\alpha$  (see (32)) with the vectorfields  $K, L$  in place of  $X_1, X_2, X_3$ .

The commutation vectorfields are  $L, S$  and  $O_i$  with  $i = 1, 2, 3$ , where the latter are the rotation fields and  $S$  is defined by

$$S = u \underline{L} + \underline{u} L.$$

The modified Lie derivatives  $\hat{\mathcal{L}}_Y$  of the curvature are taken with respect to the commutation vectorfields, thus  $Y$  replaced by  $L, S, O_i : i = 1, 2, 3$ . In particular there are first-order and second-order modified Lie derivatives of the spacetime curvature.

Next, one defines the total second-order energy-momentum densities  $P_2^{(n)}$  for  $n = 0, 1, 2, 3$  as the sum of

$$\delta^{2l} P^{(n)}(W)$$

---

<sup>2</sup>The  $O_p(\delta^l |u_0|^s)$  denote the product of  $\delta^l |u_0|^s$  with a nonnegative and nondecreasing continuous function of the  $C^p$ -norm of  $\psi_0$  on  $[0, 1] \times S^2$ .

over all the Weyl fields in a specific way, where  $l$  denotes an index according to the number of  $\hat{\mathcal{L}}_L$  operators applied to  $R$ . Then specific energies  $E_2^{(n)}(u)$  are defined via integrals on  $C_u$  and fluxes  $F_2^{(n)}(\underline{u})$  via integrals on the  $\underline{C}_{\underline{u}}$  of the 3-forms dual to the  $P_2^{(n)}$ . From these, one defines the following quantities:

$$(36) \quad \mathcal{E}_2^{(n)} = \sup_u (\delta^{2q_n} E_2^{(n)}(u)), \quad n = 0, 1, 2, 3,$$

$$(37) \quad \mathcal{F}_2^{(3)} = \sup_{\underline{u}} (\delta^{2q_3} F_2^{(3)}(\underline{u})),$$

with exponents  $q_n : n = 0, 1, 2, 3$  given as

$$q_0 = 1, \quad q_1 = 0, \quad q_2 = -\frac{1}{2}, \quad q_3 = -\frac{3}{2}.$$

One of the main goals then is to bound the quantities in (36) and (37) in terms of the initial data.

We recall from above that the final piece in the proof of the existence theorem is to estimate all the other geometric quantities in a bootstrap argument under corresponding assumptions on the curvature. In particular, the behavior of the shear  $\hat{\chi}$  will be crucial in view of (19) and Theorem 8, respectively 9. In general, one may have the first idea to just integrate propagation equations for the connection coefficients along the generators of  $C_u$  and  $\underline{C}_{\underline{u}}$ , but in such a procedure a derivative would be “lost” because the connection coefficients are estimated at the same level as the curvature components. One has to make use of elliptic estimates for the connection coefficients on the  $S_{u,u}$  and couple these with propagation equations. We already referred to this above, as Christodoulou and Klainerman developed this method in [27] to prove the global nonlinear stability of Minkowski spacetime.

These are the main ideas [25] of Christodoulou’s proof of existence. Beyond these, more intricate challenges had to be overcome and the proof bears many conceptual and technical novelties that we do not address here. However, the main achievements summarized in this section allow us to understand the formation of closed trapped surfaces in Theorems 8 and 9. This will be discussed in the next section.

**4.5. Formation of closed trapped surfaces.** The closed trapped surface formation theorems, Theorems 8 and 9, are simpler to prove than the existence theorem, Theorem 10. From the proof of the latter, not only is the spacetime constructed that is required in Theorems 8, respectively 9, but also the structure of the curvature and connection coefficients has been revealed. In particular the expansion  $\text{tr}\underline{\chi}$  and the shear  $\hat{\chi}$  are estimated.

From the existence result it follows that on  $M'_{-1}$  it is

$$(38) \quad |\text{tr}\underline{\chi} + \frac{2}{|u|}| \leq O(\delta|u|^{-2}).$$

Thus if  $\delta$  is sufficiently small, then  $\text{tr}\underline{\chi}$  is negative everywhere on  $M'_{-1}$ . Therefore, a surface  $S_{u,u}$  in  $M'_{-1}$  is indeed a *trapped* sphere if and only if everywhere on  $S_{u,u}$  it is  $\text{tr}\underline{\chi} < 0$ .

One also finds that on  $M'_{-1}$  it is

$$(39) \quad |\hat{\chi}|^2 \leq O(\delta^{-1}|u|^{-2}).$$

In the proof of the formation of closed trapped surfaces, it is then shown that  $\text{tr}\chi < 0$ , which is deduced from the behavior of  $\hat{\chi}$  that is from  $e$  of (19).

This concludes our discussion of the pioneering result [25] establishing the formation of closed trapped surfaces for the Einstein equations.

**4.6. Generalization.** Christodoulou's work [25] was generalized by Klainerman and Rodnianski in [44] and [45]. Further extensions were given also in [42] by Klainerman, Luk, and Rodnianski.

Klainerman and Rodnianski in [44] and [45] relax the initial assumptions and therefore the short pulse hierarchy for closed trapped surfaces to form in the evolution of that data. Their proof mainly follows the lines of [25] by Christodoulou, but it introduces new solutions to new problems due to the more general situation.

The proof by Klainerman and Rodnianski on the one hand induces more challenges whereas on the other hand it simplifies certain aspects. A major simplification comes from the fact that the proof is established at one lower order of differentiability. However, as less is assumed from the beginning, less control is gained on the solutions than in the pioneering result.

Klainerman, Luk, and Rodnianski, using the original existence theorem, have derived a result [42] which considerably extends the original trapped surface formation theorem, as it does not require a lower bound on the incoming energy in all directions. In particular, they show the following: Consider the outer boundary  $\underline{C}_\delta$  of the existence domain  $M_{-1}$ . Look at some neighborhood in  $S^2$  of some direction, and assume that the incoming energy in the directions corresponding to this neighborhood is sufficiently large depending on the angular size of this very neighborhood. Then  $\underline{C}_\delta$  contains a trapped surface. It is interesting to note that even though none of the sections  $S_{\delta,u}$  may be trapped, it is shown that there is another section of  $\underline{C}_\delta$  that is trapped. Namely, this is a surface represented as a graph  $u = G(\theta)$  over  $S^2$ . In fact, this surface attains large negative values of  $u$  in the directions corresponding to the neighborhood of large incoming energy, however, in the antipodal directions it comes near  $S_{\delta,-1+\delta}$ , the future boundary of  $\underline{C}_\delta$  in  $M_{-1}$ .

Recently, P. Le [50] greatly simplified and clarified the latter work by looking at the intersection of a hyperplane with a lightcone in Minkowski spacetime. In view of the extrinsic geometry of the intersection, Le shows that in the case of a null hyperplane intersecting the cone, we have a noncompact marginally trapped surface. For this situation, he gives a geometric interpretation of Green's function of the Laplacian on the standard sphere.

See also the proof by Reiterer and Trubowitz [62]. Further, see Yu [77], and Li and Yu [53]. Miao and Yu [56] applied the short pulse method successfully to study shock formation in quasilinear wave equations. Christodoulou's work [25] and his short pulse method have sparked a wealth of activity in GR and related fields. For a more complete discussion of these, see the latter references.

**4.7. Incompleteness theorem revisited.** As concluding remarks about black hole formation, we address again the incompleteness theorem in connection with the formation of closed trapped surfaces. Consider therefore the incompleteness theorems, Theorem 1, respectively 2.

From this it can be shown in a straightforward manner that indeed the spacetime, which is the maximal development of “appropriate” initial data, contains a black hole region.

Christodoulou constructed explicit initial data in [25, Chapter 17, page 579]. For an appropriate choice of initial data, which is also asymptotically flat (which is made precise in terms of decay of the data), there exists a solution spacetime for at least a finite value of  $u$  such that future null infinity  $\mathcal{I}^+$  can be defined. Indeed, one can then prove the following corollary.

**Corollary 1.** *Consider complete initial data on  $C_0$  as in Theorem 8. Let  $(\mathcal{M}, g)$  be the maximal development of this data. Then*

$$S_{\delta, -1-\delta} \cap J^-(I^+) = \emptyset.$$

*In particular, the spacetime  $\mathcal{M}$  contains a black hole region  $B$ .*

This corollary is obtained via the following road: Consider Christodoulou’s initial data mentioned above together with his Theorem 8 and Penrose’s incompleteness theorem, Theorem 1 (respectively 2). From this, the next statement follows directly.

**Corollary 2.** *Consider complete initial data on  $C_0$  as in Theorem 8. Let  $(\mathcal{M}, g)$  be the maximal development of this data. Then  $(\mathcal{M}, g)$  is future causally geodesically incomplete.*

We can summarize simply that the presence of a closed trapped surface in a spacetime  $(\mathcal{M}, g)$ , which solves the Einstein vacuum equations, implies the existence of a black hole. By a simple contradiction argument, we can prove that if such an  $(\mathcal{M}, g)$  contains a closed trapped surface  $S$ , then  $S \cap J^-(I^+) = \emptyset$ , which is the content of Corollary 1. In particular, it says that  $S$  cannot lie in the domain of outer communications  $J^-(I^+)$ , but  $\mathcal{M}$  must have a black hole region that must contain  $S$ .

See the following references on these topics: [40], [75], [26], as well as [27], [43].

## 5. STABILITY OF BLACK HOLES

If we think of the simplest black hole solutions of the EV equations (1), namely Schwarzschild or Kerr spacetimes, then we may ask, What happens if we perturb such a specific solution? Do we expect it to settle down to (another) solution of that type? Is the answer yes, if we start with initial data that is sufficiently close to Schwarzschild or Kerr? It turns out that this is an open problem of very active research in mathematical GR with goal of proving the nonlinear stability of the Kerr family. In fact, one would like to prove a conjecture of the following form: (Mass is denoted by  $M$  and angular momentum by  $a$ .)

**Conjecture 1.** *Let  $(H, \bar{g}, k)$  be a vacuum initial data set sufficiently close to two-ended Kerr data for some subextremal parameters  $0 \leq |a_i| \leq M_i$ . Then the resulting vacuum spacetime  $(\mathcal{M}, g)$  has a complete future null infinity  $\mathcal{I}^+$  such that the metric restricted to  $J^-(\mathcal{I}^+)$  remains close to for all time, and moreover it asymptotically settles down to a nearby Kerr solution in a uniform way with quantitative decay rates.*

An important fact for the Einstein equations is that the problem of completeness and asymptotic stability in the above sense are coupled. Therefore, any progress

towards understanding those problems has to come with a quantitative description of decay rates of the solution.

As Kerr stability is a huge field with many contributors; instead of citing the extremely long list of partial achievements, we refer to the following survey articles for a precise account of the history and references including the latest works by Dafermos [28], Tataru [71], and Andersson, Bäckdahl, and Blue [2].

One can formulate the following three statements of “stability” in this context: linear mode stability, linear stability and nonlinear stability (see Conjecture 1). Whiting proved linear mode-stability based on work by Teukolsky. (In the Schwarzschild case, see Regge and Wheeler, and also Zerilli.) A related problem, which one wants to solve before attacking the main case above, is the study of the wave equation on a fixed black hole background. For this case, versions of linear stability on Schwarzschild were proven by Wald, Kay, Friedman.

In recent years, the main focus was on studying the wave equation on Kerr background (see works by Andersson, Blue, Dafermos, Dyatlov, Finster, Ionescu, Kamran, Klainerman, Rodnianski, Smoller, Sterbenz, Tataru, Tohaneanu, Yau, and many more). We may summarize the findings as follows. Solutions to the wave equation  $\square_g \Psi = 0$  on subextremal Kerr  $|a| < M$  remain bounded in the exterior, and they decay inverse-polynomially to zero. An interesting linear instability arises (see Aretakis, and Lucietti and Reall) in the extremal case, which is not captured by “mode stability”.

It was recently established that solutions to the linearization of the EV equations around a Schwarzschild metric for regular initial data remain globally bounded on the black hole exterior, and they decay to a linearized Kerr metric. See the recent paper by Dafermos, Holzegel and Rodnianski, and the recent paper by Finster and Smoller.

The only global nonlinear stability result (Theorem 7) proven so far is the one on the global nonlinear stability of Minkowski spacetime of [27] by Christodoulou and Klainerman. See also generalizations and other related works cited above.

On the road towards proving the fully nonlinear result of the above conjecture, there are expected to be many beautiful insights into the structures of the Einstein equations.

## 6. COSMIC CENSORSHIP

One of the great open problems of GR is the so-called *weak cosmic censorship* conjecture, which is the following:

**Conjecture 2** (Weak cosmic censorship). *For generic asymptotically flat vacuum initial data, the resulting vacuum spacetime has a complete future null infinity  $\mathcal{I}^+$ .*

There is of course room to make clear what “generic” data should look like.

The Einstein equations also allow other types of singularities, namely the so-called naked singularities, which are not surrounded by an event horizon but can be seen from infinity. Cosmic censorship conjectures that the latter do not form during a gravitational collapse. Christodoulou investigated this in a series of papers in the 1980s and 1990s, where he showed for certain classes of initial data (studying a scalar-field model) that such naked singularities may occur, but they are unstable. See his work [25] for a summary and the references.

For a nice explanation of the above conjecture and also the *strong cosmic censorship* statement, see the introduction of [25].

There is still a long way to go in order to understand trapped surface formation for the general Einstein equations (2), when the right-hand side is not equal to zero. A huge step happened with Christodoulou's result [25] constituting a major breakthrough establishing the picture outlined above for the Einstein vacuum equations (1).

Beyond the topics addressed in this article, mathematical GR bears many more exciting challenges for geometric analysis and other mathematical fields to be investigated in the future.

#### ACKNOWLEDGMENTS

The author is grateful to the AMS for the invitation to give a talk in the Current Events Seminar. She thanks Demetrios Christodoulou as well as David Garfinkle for helpful comments on the first draft of this article.

#### ABOUT THE AUTHOR

Lydia Bieri is associate professor in the Department of Mathematics at the University of Michigan in Ann Arbor. She obtained her PhD in 2007 from ETH in Zurich, and she was a Benjamin Pierce Lecturer at Harvard University before moving to Ann Arbor. Bieri's main research is in mathematical general relativity and geometric analysis.

#### REFERENCES

- [1] Abbott, B.P., et al., LIGO Scientific Collaboration and Virgo Collaboration, *Observation of gravitational waves from binary black hole merger*, Phys. Rev. Lett. **116** (2016) 061102.
- [2] L. Andersson, T. Bäckdahl, and P. Blue, *Spin geometry and conservation laws in the Kerr spacetime*, Surveys in Differential Geometry, vol. 20, International Press, Boston, MA, 2015, pp. 183–226, DOI 10.4310/SDG.2015.v20.n1.a8. MR3467368
- [3] A. Ashtekar and R. O. Hansen, *A unified treatment of null and spatial infinity in general relativity. I. Universal structure, asymptotic symmetries, and conserved quantities at spatial infinity*, J. Math. Phys. **19** (1978), no. 7, 1542–1566, DOI 10.1063/1.523863. MR0503432
- [4] A. Ashtekar and B. G. Schmidt, *Null infinity and Killing fields*, J. Math. Phys. **21** (1980), no. 4, 862–867, DOI 10.1063/1.524467. MR565734
- [5] A. Ashtekar and M. Streubel, *Symplectic geometry of radiative modes and conserved quantities at null infinity*, Proc. Roy. Soc. London Ser. A **376** (1981), no. 1767, 585–607, DOI 10.1098/rspa.1981.0109. MR626172
- [6] L. Bieri, *An extension of the stability theorem of the Minkowski space in general relativity*, ETH Zurich, PhD thesis, **17178**, Zurich (2007).
- [7] L. Bieri, *Extensions of the stability theorem of the Minkowski space in general relativity. Part I: Solutions of the Einstein vacuum equations*, AMS/IP Studies in Advanced Mathematics. American Mathematical Society, Providence, RI; International Press, Cambridge, MA, 2009. MR2531716 (2010k:83009)
- [8] L. Bieri, P. Chen, and S.-T. Yau, *Null asymptotics of solutions of the Einstein-Maxwell equations in general relativity and gravitational radiation*, Adv. Theor. Math. Phys. **15** (2011), no. 4, 1085–1113. MR2929683
- [9] L. Bieri, P. Chen, and S.-T. Yau, *The electromagnetic Christodoulou memory effect and its application to neutron star binary mergers*, Classical Quantum Gravity **29** (2012), no. 21, 215003. MR2994213
- [10] L. Bieri and D. Garfinkle, *Neutrino radiation showing a Christodoulou memory effect in general relativity*, Ann Henri Poincaré. **16** (2015), no. 3, 801–839. MR3311889
- [11] L. Bieri and D. Garfinkle, *Perturbative and gauge invariant treatment of gravitational wave memory*, Phys. Rev. D. **89** (2014), 084039.
- [12] L. Bieri, D. Garfinkle, and N. Yunes, *Gravitational waves and their mathematics*, To appear.
- [13] G. D. Birkhoff, *Relativity and modern physics*, Harvard Univ. Press., Cambridge MA (1923).

- [14] L. Blanchet and T. Damour, *Hereditary effects in gravitational radiation*, Phys. Rev. D (3) **46** (1992), no. 10, 4304–4319, DOI 10.1103/PhysRevD.46.4304. MR1189306
- [15] H. Bondi, *Gravitational waves in general relativity*, Nature **186** (1960), 535.
- [16] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, *Gravitational waves in general relativity. VII. Waves from axi-symmetric isolated systems*, Proc. Roy. Soc. A. **269** (1962), 21–52.
- [17] V. B. Braginsky and L. P. Grishchuk, *Kinematic resonance and the memory effect in free mass gravitational antennas*, Zh. Eksp. Teor. Fiz. **89** (1985) 744 [Sov. Phys. JETP **62** (1986) 427].
- [18] V. B. Braginsky, K. S. Thorne, *Gravitational-wave bursts with memory and experimental prospects*, Nature (London) **327** (1987) 123–125.
- [19] Y. Choquet-Bruhat, *Théorème d'existence pour certain systèmes d'équations aux dérivées partielles nonlinéaires*, Acta Math. **88** (1952), 141–225. MR0053338
- [20] Y. Choquet-Bruhat, *Beginning of the Cauchy problem for Einstein's field equations*, Surveys in Differential Geometry. 2015. One hundred years of general relativity, 1–16, Surv. Differ. Geom. 20, International Press, Boston, MA, 2015. MR3467361
- [21] Y. Choquet-Bruhat and R. Geroch, *Global aspects of the Cauchy problem in general relativity*, Comm. Math. Phys. **14** (1969), 329–335.
- [22] D. Christodoulou, *The global initial value problem in general relativity*, The Ninth Marcel Grossmann Meeting (Rome, 2000). ed. by V. G. Gurzadyan et al., World Scientific, Singapore, 2002, pp. 44–54.
- [23] D. Christodoulou, *Nonlinear nature of gravitation and gravitational-wave experiments*, Phys. Rev. Lett. **67** (1991), no. 12, 1486–1489. MR1123900 (92f:83009)
- [24] D. Christodoulou, *Mathematical problems of general relativity theory I and II*, Zürich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008. MR2391586 (2008m:83008)
- [25] D. Christodoulou, *The formation of black holes in general relativity*, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich, 2009. MR2488976 (2009k:83010)
- [26] D. Christodoulou, *Topics on black holes and gravitational waves*, Lectures held at ETH Zürich, (2011).
- [27] D. Christodoulou and S. Klainerman, *The global nonlinear stability of the Minkowski space*, Princeton Mathematical Series, vol. 41, Princeton University Press, Princeton, NJ, 1993. MR1316662
- [28] M. Dafermos, *The mathematical analysis of black holes in general relativity*, Proceedings of the ICM Seoul, Vol. III, (2014), 747–772.
- [29] M. Dafermos and I. Rodnianski, *Lectures on black holes and linear waves*, Evolution equations, Clay Math. Proc., vol. 17, Amer. Math. Soc., Providence, RI, 2013, pp. 97–205. MR3098640
- [30] A. S. Eddington, *A comparison of Whitehead's and Einstein's formulas*, Nature **113** (1924), 192.
- [31] A. Einstein, *Zur Allgemeinen Relativitätstheorie*, Sitzungsber. K. Preuss. Akad. Wiss., Berlin, (1915), 778–786.
- [32] A. Einstein, *Die Feldgleichungen der Gravitation*, Sitzungsber. K. Preuss. Akad. Wiss., Berlin, (1915), 844–847.
- [33] M. Favata, *Nonlinear gravitational-wave memory from binary black hole mergers*, Astrophys. J. **696** (2009), L159–L162.
- [34] D. Finkelstein, *Past-future asymmetry of the gravitational field of a point particle*, Phys. Rev. **110** (1958), 965–967.
- [35] É. É. Flanagan and D. A. Nichols, *Conserved charges of the extended Bondi-Metzner-Sachs algebra*, Phys. Rev. D **95** (2017), no. 4, 044002.
- [36] É. É. Flanagan and D. A. Nichols, *Erratum: Observer dependence of angular momentum in general relativity and its relationship to the gravitational-wave memory effect*, Phys. Rev. D **93** (2016), no. 4, 049905(E), 1. MR3499432
- [37] H. Friedrich, *On the existence of  $n$ -geodesically complete or future complete solutions of Einstein's field equations with smooth asymptotic structure*, Comm. Math. Phys. **107** (1986), 587–609.

- [38] J. Frauendiener, *Note on the memory effect*, Classical Quantum Gravity **9** (1992), no. 6, 1639–1641. MR1166696
- [39] R. Geroch, *Asymptotic structure of space-time*, Asymptotic structure of space-time (Proc. Sympos., Univ. Cincinnati, Cincinnati, Ohio, 1976), Plenum, New York, 1977, pp. 1–105. MR0484240
- [40] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time*, Cambridge Monographs on Mathematical Physics, No. 1, Cambridge University Press, London–New York, 1973. MR0424186
- [41] R. P. Kerr, *Gravitational field of a spinning mass as an example of algebraically special metrics*, Phys. Rev. Lett. **11** (1963), 237–238, DOI 10.1103/PhysRevLett.11.237. MR0156674
- [42] S. Klainerman, J. Luk, and I. Rodnianski, *A fully anisotropic mechanism for formation of trapped surfaces*, Invent Math. **195** (2014).
- [43] S. Klainerman and F. Nicolò, *The evolution problem in general relativity*, Progress in Mathematical Physics, vol. 25, Birkhäuser Boston, Inc., Boston, MA, 2003. MR1946854
- [44] S. Klainerman and I. Rodnianski, *On the formation of trapped surfaces*, Acta. Math. **208** (2012), no. 2, 211–213. MR2931382
- [45] S. Klainerman and I. Rodnianski, *On emerging scarred surfaces for the Einstein vacuum equations*, Discrete Contin. Dyn. Syst. **28** (2010), no. 3, 1007–1031. MR2644776
- [46] S. Klainerman, I. Rodnianski, and J. Szeftel, *The bounded  $L^2$  curvature conjecture*, Invent. Math. **202** (2015), no. 1, 91–216. MR3402797
- [47] S. Klainerman, I. Rodnianski, and J. Szeftel, *Overview of the proof of the bounded  $L^2$  curvature conjecture* (2012), arXiv:1204.1772v1.
- [48] M. D. Kruskal, *Maximal extension of Schwarzschild metric*, Phys. Rev. (2) **119** (1960), 1743–1745. MR0115757
- [49] P. Lasky, E. Thrane, Y. Levin, J. Blackman, and Y. Chen, *Detecting gravitational-wave memory with LIGO: implications of GW150914*, Phys. Rev. D. **117** (2016), no. 6, 061102.
- [50] P. Le, *The intersection of a hyperplane with a lightcone in the Minkowski spacetime* (2016), arXiv:1601.01567v1
- [51] G. Lemaître, *L’univers en expansion*, Ann. Soc. Sci. Bruxelles I **A53** (1933), 51–85.
- [52] J. Leray, *Hyperbolic differential equations*, The Institute for Advanced Study, Princeton, NJ, 1953. MR0063548
- [53] J. Li and P. Yu, *Construction of Cauchy data of vacuum Einstein field equations evolving to black holes*, Ann. of Math. (2) **181** (2015), no. 2, 699–768, DOI 10.4007/annals.2015.181.2.6. MR3275849
- [54] H. Lindblad and I. Rodnianski, *Global existence for the Einstein vacuum equations in wave coordinates*, Comm. Math. Phys. **256** (2005), 43–110. MR2134337
- [55] H. Lindblad and I. Rodnianski, *The global stability of Minkowski space-time in harmonic gauge*, Ann. of Math. (2) **171** (2010), no. 3, 1401–1477. MR2680391
- [56] S. Miao and P. Yu, *On the formation of shocks for quasilinear wave equations*, Invent. Math. **207** (2017), no. 2, 697–831, DOI 10.1007/s00222-016-0676-2. MR3595936
- [57] E. T. Newman and R. Penrose, *An approach to gravitational radiation by a method of spin coefficients*, J. Mathematical Phys. **3** (1962), 566–578. MR0141500
- [58] R. Penrose, *Asymptotic properties of fields and space-times*, Phys. Rev. Lett. **10** (1963), 66–68, DOI 10.1103/PhysRevLett.10.66. MR0149912
- [59] R. Penrose, *Gravitational collapse and space-time singularities*, Phys. Rev. Lett. **14** (1965), 57–59.
- [60] F. Pirani, *Invariant formulation of gravitational radiation theory*, Phys. Rev. (2) **105** (1957), 1089–1099. MR0096537
- [61] A. D. Rendall, *Reduction of the characteristic initial value problem to the Cauchy problem and its applications to the Einstein equations*, Proc. Roy. Soc. London Ser. A **427** (1990), no. 1872, 221–239. MR1032984
- [62] M. Reiterer and E. Trubowitz, *Strongly focused gravitational waves*, Comm. Math. Phys. **307** (2011), no. 2, 275–313, DOI 10.1007/s00220-011-1314-x. MR2837117
- [63] R. K. Sachs, *Gravitational waves in general relativity. VIII. Waves in asymptotically flat space-time*, Proc. Roy. Soc. London Ser. A **270** (1962), 103–126. MR149908
- [64] K. Schwarzschild, *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie*, Sitz. Deut. Akad. Wiss. Berlin, Kl. Math.-Phys. Tech. (1916), 189–196.

- [65] J. L. Synge, *The gravitational field of a particle*, Proc. Roy. Irish Acad. Sect. A. **53** (1950), 83–114. MR0039426
- [66] J. Szeftel, *Parametrix for wave equations on a rough background I: Regularity of the phase at initial time* (2012), arXiv:1204.1768v1.
- [67] J. Szeftel, *Parametrix for wave equations on a rough background II: Construction of the parametrix and control at initial time* (2012), arXiv:1204.1769v1
- [68] J. Szeftel, *Parametrix for wave equations on a rough background III: Space-time regularity of the phase* (2012), arXiv:1204.1770v1
- [69] J. Szeftel, *Parametrix for wave equations on a rough background IV: Control of the error term* (2012), arXiv:1204.1771v1
- [70] Gy. Szekeres, *On the singularities of a Riemannian manifold*, Publ. Math. Debrecen **7** (1960), 285–301. MR0125541
- [71] D. Tataru, *Decay of linear waves on black hole space-times*, Surveys in differential geometry 2015. One hundred years of general relativity, Surv. Differ. Geom., vol. 20, International Press, Boston, MA, 2015, pp. 157–182, DOI 10.4310/SDG.2015.v20.n1.a7. MR3467367
- [72] K. S. Thorne, *Gravitational-wave bursts with memory: the Christodoulou effect*, Phys. Rev. D (3) **45** (1992), no. 2, 520–524, DOI 10.1103/PhysRevD.45.520. MR1144215
- [73] A. Tolish and R. M. Wald, *Retarded fields of null particles and the memory effect*, Phys. Rev. D **89**, 064008 (2014).
- [74] A. Trautman, *Radiation and boundary conditions in the theory of gravitation*, Bull. Acad. Polon. Sci. Sér. Math. Astr. Phys. **6** (1958). 407–412. MR0097266
- [75] R. M. Wald, *General relativity*, University of Chicago Press, Chicago, IL, 1984. MR757180
- [76] A. G. Wiseman and C. M. Will, *Christodoulou’s nonlinear gravitational-wave memory: Evaluation in the quadrupole approximation*, Phys. Rev. D. **44** (1991), R2945.
- [77] P. Yu, *Energy estimates and gravitational collapse*, Comm. Math. Phys. **317** (2013), no. 2, 273–316, DOI 10.1007/s00220-012-1617-6. MR3010185
- [78] Ya. B. Zel’dovich and A. G. Polnarev, *Radiation of gravitational waves by a cluster of superdense stars*, Astron. Zh. **51**, 30. (1974). [Sov.Astron. **18**, 17. (1974)].
- [79] N. Zipser, *The global nonlinear stability of the trivial solution of the Einstein-Maxwell equations*, PhD thesis, Harvard University, Cambridge, MA (2000).
- [80] N. Zipser, *Extensions of the stability theorem of the Minkowski space in general relativity. Part II: Solutions of the Einstein-Maxwell equations*, AMS/IP Studies in Advanced Mathematics, 45. American Mathematical Society, Providence, RI; International Press, Cambridge, MA, 2009. MR2531716

UNIVERSITY OF MICHIGAN, DEPARTMENT OF MATHEMATICS, ANN ARBOR, MICHIGAN  
*E-mail address:* lbieri@umich.edu